FOREGROUND SIGNATURES AND MITIGATION IN NEUTRAL HYDROGEN POWER SPECTRA FROM THE REIONIZATION EPOCH

Nithyanandan Thyagarajan[1], Daniel C. Jacobs[1], Judd D. Bowman[1], G. Bernardi[2], F. Briggs[3], R. J. Cappallo[4], B. E. Corey[4], A. A. Deshpande[5], D. Emrich[6], B. M. Gaensler[7,8], R. Goeke[9], L. J. Greenhill[10], B. J. Hazelton[11], M. Johnston-Hollitt[12], D. L. Kaplan[13], J. C. Kasper[14,10], E. Kratzenberg[4], C. J. Lonsdale[4], M. J. Lynch[6], S. R. McWhirter[4], D. A. Mitchell[15,8], M. F. Morales[11], E. Morgan[9], D. Oberoi[16], S. M. Ord[6,8], T. Prabu[5], A. E. E. Rogers[4], A. Roshi[17], N. Udaya Shankar[5], K. S. Srivani[5], R. Subrahmanyan[5,8], S. J. Tingay[6,8], M. Waterson[6,3], R. B. Wayth[6,8], R. L. Webster[18,8], A. R. Whitney[4], A. Williams[6], C. L. Williams[9],
and MWA EoR members

Draft version October 10, 2014

ABSTRACT

Detection of 21 cm emission of neutral hydrogen from the epoch of reionization, at redshifts $z > 6$, is limited primarily by foreground emission. We investigate the chromatic signatures of an all–sky foreground model imprinted by an instrumental transfer function. Using the delay spectrum technique, we demonstrate that the foreground signatures that have the largest impact on the Hi signal arise from power received far away from the primary field of view. Comparing data from recent Murchison Widefield Array observations with simulations separated into components based on type of emission, we identify diffuse emission near the horizon as a significant contributing factor, even on wide antenna spacings. For signals entering through the primary field of view, compact objects dominate the foreground contamination. These two mechanisms imprint a characteristic *pitchfork* signature on the "foreground wedge" in Fourier space. Based on these results, we propose that selective down–weighting of baselines based on length, direction, and time will mitigate a large fraction of foreground contamination in redshifted 21 cm power spectrum analysis.

Keywords: large-scale structure of Universe — methods: statistical — radio continuum: galaxies — radio lines: general — reionization — techniques: interferometric

1. INTRODUCTION

At the end of the recombination epoch, the Universe was completely neutral. This period, referred to as the

[1] School of Earth and Space Exploration, Arizona State University, Tempe, AZ 85287, USA; e-mail: t_nithyanandan@asu.edu
[11] Department of Physics, University of Washington, Seattle, WA 98195, USA
[5] Raman Research Institute, Bangalore 560080, India
[2] Square Kilometre Array South Africa (SKA SA), Cape Town 7405, South Africa
[3] Research School of Astronomy and Astrophysics, Australian National University, Canberra, ACT 2611, Australia
[4] MIT Haystack Observatory, Westford, MA 01886, USA
[6] International Centre for Radio Astronomy Research, Curtin University, Bentley, WA 6102, Australia
[7] Sydney Institute for Astronomy, School of Physics, The University of Sydney, NSW 2006, Australia
[8] ARC Centre of Excellence for All-sky Astrophysics (CAASTRO)
[9] Kavli Institute for Astrophysics and Space Research, Massachusetts Institute of Technology, Cambridge, MA 02139, USA
[10] Harvard-Smithsonian Center for Astrophysics, Cambridge, MA 02138, USA
[12] School of Chemical & Physical Sciences, Victoria University of Wellington, Wellington 6140, New Zealand
[13] Department of Physics, University of Wisconsin–Milwaukee, Milwaukee, WI 53201, USA
[14] Department of Atmospheric, Oceanic and Space Sciences, University of Michigan, Ann Arbor, MI 48109, USA
[15] CSIRO Astronomy and Space Science (CASS), PO Box 76, Epping, NSW 1710, Australia
[16] National Centre for Radio Astrophysics, Tata Institute for Fundamental Research, Pune 411007, India
[17] National Radio Astronomy Observatory, Charlottesville and Greenbank, USA
[18] School of Physics, The University of Melbourne, Parkville, VIC 3010, Australia

Dark Ages in the Universe's history, is characterized by the localized accumulation of matter under the influence of gravity. And it ended with the formation of the first stars and galaxies which started emitting ultra–violet and X–ray radiation, thereby reionizing the neutral medium in their surroundings. This commenced the epoch of reionization (EoR) — a period of non–linear growth of matter density perturbations and astrophysical evolution. Studying the EoR holds the key to understanding this evolution.

Observations of redshifted 21 cm radiation generated by the spin flip transition of Hi has been identified as a direct probe of the EoR (Sunyaev & Zeldovich 1972; Scott & Rees 1990; Madau et al. 1997; Tozzi et al. 2000; Iliev et al. 2002). Detecting this signal has recently emerged as a very promising experiment to fill the gaps in our understanding of the Universe's history.

Sensitive instruments such as the Square Kilometre Array (SKA) are required for direct observation and tomography of redshifted Hi. Numerous precursors to the SKA such as the Murchison Widefield Array (MWA; Tingay et al. 2013; Bowman et al. 2013; Lonsdale et al. 2009), the Low Frequency Array (LOFAR; van Haarlem et al. 2013), and the Precision Array for Probing the Epoch of Reionization (PAPER; Parsons et al. 2010) have become operational with enough sensitivity for a statistical detection of the EoR Hi power spectrum (Thyagarajan et al. 2013; Beardsley et al. 2013; Bowman et al. 2006; Pober et al. 2014).

A key challenge in the statistical detection of the redshifted Hi 21 cm signal arises from the contamination by Galactic and extragalactic foregrounds (see, e.g., Di Mat-

teo et al. 2002; Zaldarriaga et al. 2004; Furlanetto et al. 2006). Morales & Hewitt (2004) show that the inherent isotropy and symmetry of the EoR signal in frequency and spatial wavenumber (k) space make it distinguishable from sources of contamination which are isolated to certain k modes by virtue of their inherent spectral smoothness (Morales et al. 2006; Bowman et al. 2009; Liu & Tegmark 2011; Parsons et al. 2012; Dillon et al. 2013).

Considerable effort has been made towards understanding the k–space behavior of foreground signatures in the observed power spectrum (Thyagarajan et al. 2013; Pober et al. 2013; Morales et al. 2012; Trott et al. 2012; Liu & Tegmark 2011; Datta et al. 2010; Bowman et al. 2009; Liu et al. 2009). A general model of scattering of smooth spectrum foregrounds to higher k–modes by the wide-field response of the instrument, into the so-called 'wedge', has emerged that provides a high level explanation for the observed foreground power spectrum. The conservative foreground strategy that has developed alongside this work is to discard k–modes which could be contaminated (*avoidance*). The more aggressive alternative is to subtract a sky model and regain access to modes lost by *avoidance*. In both cases, which parts of the sky are most critical to either *avoid* or *subtract* has remained largely uncertain. Here, we will focus primarily on extending the *avoidance* strategy by identifying foreground components most likely to "leak" from foreground modes to EoR modes and demonstrating a scheme for down–weighting these components.

Foregrounds contaminate the EoR HI power spectrum by either having intrinsic spectral variability, high instrumental chromaticity, or through approximations in power spectrum analyses. Here we use existing catalogs and a high fidelity instrumental model to capture both foreground and instrumental chromaticity. To decouple these effects from possible analysis effects, such as those pointed out by Hazelton et al. (2013), we compute power spectra using the individual baseline–based approach of Pober et al. (2013) and Parsons et al. (2014). This approximates the power spectrum as the inverse Fourier transform of the spectra generated by the instrument's correlator.

Matching real observations from the Murchison Wide-field Array with instrumental simulations of the entire known sky, we examine the signature of the power spectrum of each baseline. We report two important findings: the foregrounds most likely to obfuscate the redshifted 21 cm power spectrum are not caused by emission in the central field of view, but rather by bright objects from near the horizon; and, diffuse Galactic emission plays a larger role than previously thought. We quantify this contamination by separating the simulations into components based on type of foreground emission. We then arrive at a new method for minimizing the contribution of bright foregrounds which uses prior knowledge of the sky to down–weight adversely contaminated baselines.

In §2 we introduce the delay spectrum technique. In §3, we describe the MWA setup, summarize the observing parameters, and present the resulting data. Simulations using these observing parameters are described in §4 and analyzed by sky location and type of emission in §5. In §6, we offer an initial description of a more precise foreground avoidance technique. We present a summary of our work and findings in §7.

2. DELAY SPECTRUM

Interferometer array data known as *visibilities*, $V_f(\overline{\mathbf{u}}, f)$, represent correlations between time–series of electric fields measured by different antenna pairs with separation vectors $\overline{\mathbf{x}}$ and then Fourier transformed along the time axis to obtain a spectrum along the frequency (f) axis. Here, $\overline{\mathbf{u}} \equiv \overline{\mathbf{x}} f/c \equiv \overline{\mathbf{x}}/\lambda$, c is the speed of light, and λ is the wavelength of radio waves. $\overline{\mathbf{u}}$ is related to the transverse spatial frequency modes ($\overline{\mathbf{k}}_\perp$) of the brightness distribution of the sky as $\overline{\mathbf{k}}_\perp \equiv 2\pi\overline{\mathbf{u}}/D(z)$, where $D(z)$ is the transverse comoving distance at redshift z. If $I(\overline{\mathbf{l}}, f)$ is the emission at different frequencies on the sky as a function of direction unit vector ($\overline{\mathbf{l}}$) specified as direction cosines, $V_f(\overline{\mathbf{u}}, f)$ represents the Fourier decomposition of $I(\overline{\mathbf{l}}, f)$ attenuated by the antenna's angular power–pattern, $A(\overline{\mathbf{l}}, f)$, into transverse spatial frequency modes. Taking into account the instrumental bandpass weights $W_f(f)$, this relation can be written as:

$$V_f(\overline{\mathbf{u}}, f) = \iint\limits_{\text{sky}} A(\overline{\mathbf{l}}, f)\, I(\overline{\mathbf{l}}, f)\, W_f(f)\, e^{-i2\pi\overline{\mathbf{u}}\cdot\overline{\mathbf{l}}}\, \mathrm{d}^2\overline{\mathbf{l}}. \quad (1)$$

We define the *delay spectrum*, $V_\eta(\overline{\mathbf{u}}, \eta)$, to be the inverse Fourier transform of $V_f(\overline{\mathbf{u}}, f)$ along the frequency coordinate:

$$V_\eta(\overline{\mathbf{u}}, \eta) = \int V_f(\overline{\mathbf{u}}, f)\, W_f'(f)\, e^{i2\pi\eta f}\, \mathrm{d}f, \quad (2)$$

where, $W_f'(f)$ is a frequency window weighting function which can be chosen to control the quality of the delay spectrum (Thyagarajan et al. 2013; Vedantham et al. 2012) and η represents the signal delay between antenna pairs:

$$\eta = \frac{\overline{\mathbf{x}} \cdot \overline{\mathbf{l}}}{c} = \frac{\overline{\mathbf{u}} \cdot \overline{\mathbf{l}}}{f}. \quad (3)$$

Note that as the visibility is a complex quantity, its Fourier transform is Nyquist sampled over positive and negative delays. Negative and positive delay[19] correspond to the two hemispheres of the sky as transected by the antenna spacing vector. In this paper, $\overline{\mathbf{u}}$ and $\overline{\mathbf{x}}$ are assumed to be on a coordinate system aligned with the local east, north (along local meridian) and upward directions at the telescope site. Hence, a perfectly eastward oriented antenna spacing will observe an object in the eastern sky at positive delay, and conversely for an object in the western sky. Similarly, an object in the northern sky will appear at a positive delay at an antenna spacing oriented northward.

Since we are studying a redshifted HI spectral line, f is a measure of cosmological distance along the line of sight. Consequently, η is a measure of line–of–sight spatial frequency mode given by $k_\parallel \approx 2\pi\eta\, f_{21} H_0\, E(z)/(c(1+z)^2)$, where f_{21} is the rest frame frequency of the 21 cm spin flip transition of HI, and H_0 and $E(z) \equiv [\Omega_\mathrm{M}(1+z)^3 + \Omega_\mathrm{k}(1+z)^2 + \Omega_\Lambda]^{1/2}$ are standard terms in cosmology. We

---

[19] The terms delay and lag are used interchangeably in this paper.

use $H_0 = 100\,h$ km s^{-1} Mpc^{-1} and assume a flat ΛCDM universe with $\Omega_{\mathrm{M}} = 0.27$ and $\Omega_{\Lambda} = 0.73$.

Thus, $|V_\eta(\overline{\mathbf{u}}, \eta)|^2$ is directly related to the spatial power spectrum of EoR HI distribution, $|V(\overline{\mathbf{k}}_\perp, k_\parallel)|^2$ (Morales & Hewitt 2004). However, foreground objects also emit at radio frequencies. Due to their spatial and spectral (chromatic) structure, the measured $V_\eta(\overline{\mathbf{u}}, \eta)$ is contaminated by foreground radio objects (Thyagarajan et al. 2013; Trott et al. 2012; Morales et al. 2012; Bowman et al. 2009). These contaminations have to be characterized precisely in order to reduce their impact on EoR HI power spectrum detection sensitivity.

In summary, $V_\eta(\overline{\mathbf{u}}, \eta)$ is closely related to visibilities, $V_f(\overline{\mathbf{u}}, f)$, the basic data blocks measured by each antenna pair, as seen from equation 2. $V_\eta(\overline{\mathbf{u}}, \eta)$ captures all the effects of EoR HI signal corruption caused by foregrounds and the instrument. At the same time, it is directly related to the sought power $|V(\overline{\mathbf{k}}_\perp, k_\parallel)|^2$ in k–modes containing critical information about spatial scales. For our analysis, we use an equivalent quantity, $V_\eta(\overline{\mathbf{x}}, \eta) \equiv V_\eta(\lambda \overline{\mathbf{u}}, \eta)$. For purposes of illustration, we often use a simplified quantity, $|V_\eta(|\overline{\mathbf{x}}|, \eta)|$.

Figure 1 illustrates the Fourier space in which the delay (and power) spectra of HI from the EoR are presented in this paper. $|\overline{\mathbf{x}}|$ and $k_\perp$, denoting spatial scales in the transverse direction, form the x–axis. η and $k_\parallel$, denoting spatial scales along line–of–sight form the y–axis. Foreground emission maps to a wedge–shaped region in Fourier space, hereafter referred to as the *foreground wedge*, whose boundaries are determined by antenna spacing and the light travel time across it. These boundaries are called horizon delay limits and are shown by solid lines. The spectral transfer function of the instrument convolves the *foreground wedge* and stretches it further (unshaded narrow strips bounded by solid and dashed lines) along η–axis. This effect has been quantified in Thyagarajan et al. (2013). The width of this narrow strip is inversely proportional to the operating bandwidth. The region of Fourier space excluding the *foreground wedge* and the narrow strips is the so–called *EoR window*, shown in light and medium shades. However, the instrument used in our study, namely the MWA, has a passband constructed using coarse channels with a characteristic shape giving rise to periodic gratings thereby resulting in repetitions of the *foreground wedge* at multiples of $0.78\ \mu$s. The underlying cause of this effect will be discussed later in this paper.

In the context of EoR studies in Fourier space, the HI power spectrum from the EoR is expected to decrease rapidly with $|\overline{\mathbf{k}}|$. Hence, the brightest EoR signal will be observed on the shortest baselines and smallest delays. Thus the region of interest for EoR studies is slightly beyond the horizon delay limits (dashed lines) and inside the first grating on either side ($|\eta| \lesssim 0.78\,\mu$s, dot–dashed lines) caused by the instrument's spectral transfer function. We call this region with medium shade as the *EoR window* in this paper. The MWA EoR collaboration actively uses this as a test region in Fourier space for diagnostics on data quality.

For geometrical intuition, we restrict the orientation (θ_b, measured anti-clockwise from East) of all baselines to lie in the range $-67^\circ\!.5 \leq \theta_b < 112^\circ\!.5$. Baselines oriented in the other half–plane measure conjugate visibili-
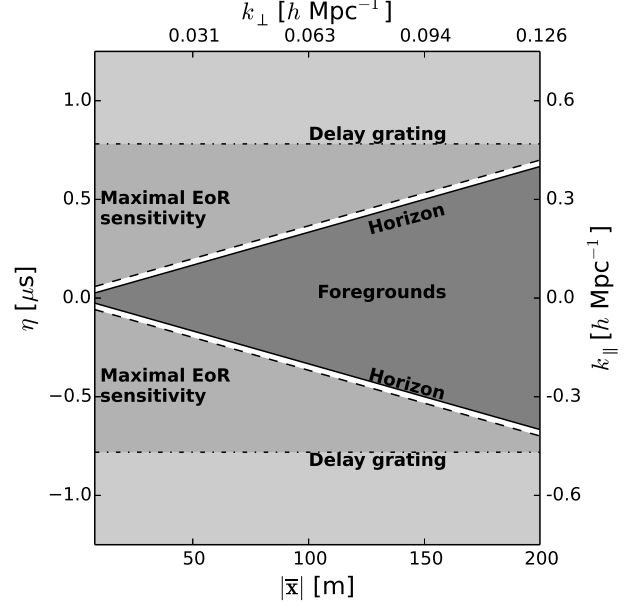

Figure 1. Illustration of Fourier space in which delay (and power) spectra of HI signals from the EoR are presented. The x–axis is denoted by $|\overline{\mathbf{x}}|$ (antenna spacing) or $k_\perp$ (transverse wavenumber). The y–axis denoted by η (delay) or $k_\parallel$ (line–of–sight wavenumber). The dark shaded region is referred to as the *foreground wedge* where smooth spectrum foregrounds reside. Its boundaries (solid lines), given by light travel time for corresponding antenna spacings, are referred to as horizon delay limits. Regions outside the *wedge* and the narrow white unshaded strips (caused by convolution with the instrumental transfer function) are expected to be relatively free of foreground contamination. Due to an undesirable grating pattern (dot–dashed lines) of the instrument used in this study, namely the MWA, we identify a restricted region of high EoR sensitivity (medium shade) and refer to it as the *EoR window* in this paper.

ties with delays of equal magnitude but of opposite sign and hence are ignored in our analysis.

3. EXAMPLE DATA

3.1. *The Murchison Widefield Array (MWA)*

The MWA was completed in 2012 and, after commissioning, began its EoR observing program in 2013. In its final configuration the MWA is a 128–tile interferometer array capable of observing a 30.72 MHz instantaneous band anywhere in the range 80–300 MHz. Each tile is a phased array of 16 dipoles, each in the shape of a bow–tie. This yields a primary field of view $\gtrsim 20^\circ$ wide and multiple secondary lobes. The array is arranged as a centrally condensed core of ~ 300 m — there are many spacings in the range 5–50 m — and a radial density that falls off as the inverse of the radius, with the longest baselines at 3 km (Beardsley et al. 2012). Here we focus on antenna spacings $|\overline{\mathbf{x}}| \leq 200$ m (spatial scales relevant to reionization) and a frequency range 170–200 MHz. The center frequency of 185 MHz corresponds to a redshift, $z \simeq 6.68$.

The MWA passband of width 30.72 MHz is divided coarsely into 24×1.28 MHz sub–bands with each sub–band weighted by a digital filter. The coarse channel shape is obtained using a 8–tap polyphase filter bank (PFB) and a Kaiser window with parameter $\beta = 5$. After correcting for the shape of these coarse channels, the fine channels (40 kHz width) at the edges of these sub–

bands are flagged because they are known to be possibly contaminated by aliasing at a low level. Any spectral flags, or other known weights constitutes $W_f(f)$ in equation 1.

Thermal noise in simulated visibilities is estimated assuming a system temperature of $T_{\mathrm{sys}} = 95$ K per polarization for all frequency channels for all antenna pairs throughout the course of the observation.

3.2. *Observation Parameters*

The MWA targets two primary low-foreground fields for reionization observations. Here, we focus on the field at RA = 0h, Dec = $-30°$. The MWA tracks a patch of sky through antenna beams formed and steered electronically by controlling delay settings of the dipoles in an MWA tile. The pointing system is capable of steering to points on a regular $\sim 7°$ grid. During observations we allow the sky to drift across the nearest available pointing, shifting between grid points as necessary (once in ~ 30 minutes). This process is repeated throughout the course of the observation ≈ 4.86 hours. Figure 2 illustrates how the apparent pointing oscillates around the desired pointing direction.

The observations shown here were undertaken on Aug 23, 2013. This day has been chosen as a reference data set for comparison within the MWA EoR collaboration and was selected on the basis of a low number of known system errors from among a long stretch of stable observing.

Here we have chosen two, 112 second long sections from this night for detailed study. These were chosen to provide a selection of possible foreground and instrumental conditions. As an example of a nominal observing setup we choose a zenith pointing; as an example of poor foreground conditions, we choose a pointing when the field is ~ 2 hours from zenith. This pointing has a significantly higher secondary lobe structure and is observed when the bright galactic center is well above the horizon.

The *local sidereal time* (LST) at these two snapshots are shown in Figure 2 as vertical lines at -1.92 hours (i.e., 22.08 hours) and 0.09 hours which are hereafter denoted as *off–zenith* and *zenith* pointings, respectively.

3.3. *Initial Data Processing*

The data are flagged[20] for interference (Offringa et al. 2012, 2010), removing 3% of the data and averaged in time and frequency from the raw 0.5 s, 40 kHz to 2s, 80 kHz. These data are then calibrated to a simulation of the sky containing 2420 point–like objects selected from the MWA Commissioning Survey (MWACS; Hurley-Walker et al. 2014, submitted). MWACS was compiled using observations taken with subsets of the array during the final buildout and commissioning; it has a flux density limit of 25 mJy and a declination range of -12°to -40°evenly covering the field of view of the observations reported here. The objects used in calibration are selected to lie inside the 5% contour of the primary lobe of the tile power–pattern. The calibration algorithm – based on forward modeling software by Sullivan et al. (2012) – computes complex gain solutions per channel per antenna averaged to two minute intervals. The solutions are fairly low signal–to–noise so we
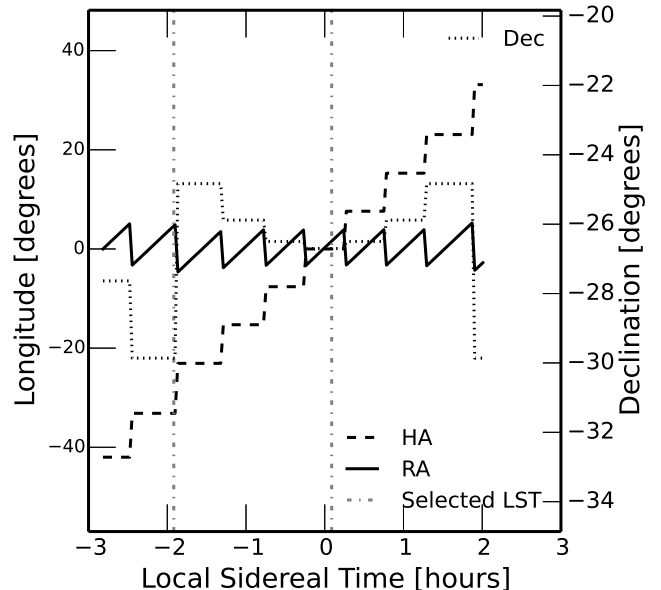
[20] http://sourceforge.net/p/aoflagger

Figure 2. MWA tile beam pointing directions during the course of the observation. The *x*–axis refers to the *local sidereal time* (LST). The axis on the left refers to longitudes, namely, Right Ascension (RA) and Hour Angle (HA). Negative values of RA, HA, and LST are to be interpreted as having been wrapped around by 360°or 24 hours. The axis on the right refers to the declination of the pointing direction. The RA, HA, and declination are plotted with solid, dashed and dotted lines respectively. The sky is allowed to drift for ~ 30 mins before the beamformer delay settings change in a discrete step to center the beam around RA = 0h, Dec = $-30°$. Dot–dashed vertical lines show the LST of *off-zenith* and *zenith* pointings used in our study.

iteratively project along the antenna and frequency dimensions to capture the relatively independent passband and antenna–to–antenna variation. First, we average the channel gains over all antennas to obtain a high signal–to–noise measurement of the bandpass. After applying this single passband, we do a second round of calibration and fit second and first order polynomials for amplitude and phase respectively for each antenna. This flattens any residual variation in bandpass and removes small phase slopes due to variations in cable delay. Finally, we fit for an additional phase known to be caused by reflections in 150 m length cables.

3.4. *Deconvolved Delay Spectrum*

We obtain the delay spectrum of these calibrated visibilities taking the delay transform of each baseline's spectrum (equation 2) choosing $W_f'(f)$ to be a *Blackman–Harris* window function. The spectrum is multiplied by the additional weight of the flagged channels which occur rarely for interference and regularly every 1.28 MHz where the edges of coarse channels are known to have a slight aliasing effect. In delay–space, these weights translate into a convolution by a grated point spread function (PSF) with gratings at multiples of $\eta = 0.78\,\mu$s. We deconvolve this PSF using a one dimensional CLEAN algorithm (Taylor et al. 1999) as described for the delay axis by Parsons & Backer (2009); Parsons et al. (2012). The CLEAN procedure iteratively finds and subtracts peak values convolved by the Fourier transform of the

weights. We limit the selection of peaks to modes inside the horizon delay limit, corresponding to smooth spectrum objects in the visible sky hemisphere.

In subsequent sections, we plot the amplitude of the delay spectrum, $|V(|\bar{\mathbf{x}}|, \eta)|$, sorted by baseline length, $|\bar{\mathbf{x}}|$. Note that visibilities from different baselines have not been averaged together in this analysis.

We show results of delay spectrum from MWA data after deconvolution along the delay axis in Figure 3. The principle feature is the *foreground wedge*, whose boundaries, marked by solid white lines, increase linearly with light travel time corresponding to baseline length. Note that the wedge is not filled out uniformly. Each baseline has a different orientation and resulting response to different objects in the sky. It is this variation that we will explore below.

The *off–zenith* pointing has emission notably higher than in the *zenith* pointing inside the *foreground wedge*. This is shown later to be due to response of the array to the bright Galactic center and Galactic plane in the westward sky. Consequently, the contamination into the *EoR window* is also higher. Emission in the *zenith* pointing is more centrally concentrated due to phasing of power–pattern to the zenith, whereas the *off–zenith* pointing has a power–pattern phased eastward of zenith but is also responsive to emission from angles far away from zenith.
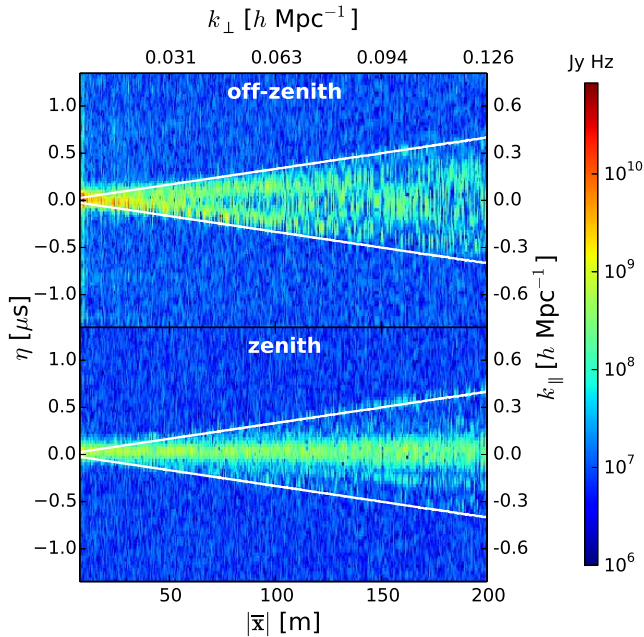

Figure 3. Delay spectra of two observations of the EoR field (RA = 0h, Dec = −30°) at an *off–zenith* pointing (top) when tiles have significant secondary lobes in the direction of the Galactic center and a *zenith* pointing (bottom) where secondary lobes are minimal and the Galaxy has set. To get the delay spectrum, each baseline is Fourier transformed into delay space (*y*–axis) and the amplitude plotted in order of baseline length (*x*–axis). For reference, the *k*–space equivalent of η and $|\bar{\mathbf{x}}|$, namely, $k_\parallel$ and $k_\perp$ are shown at the axes on the right and top respectively. The logarithmic color scale (shown at the right) is common to both panels. White lines mark the boundaries of *foreground wedge* determined by light travel time across each baseline. Smooth spectrum objects predominantly reside in this *foreground wedge*. The various features in this region depend on antenna spacing orientation, secondary lobe structure of the power–pattern, and the sky at that instant.

Despite delay deconvolution, we still see leakage beyond the cleaned regions indicating the inability of the deconvolution algorithm to fit the measurements perfectly, dominant at short antenna spacings where the number of delay modes within the horizon is lesser. This is because the maximum delay envelope (boundary of the wedge) consists of mixed emission from large portions of the sky into a narrow range of delays and the deconvolution algorithm does not have sufficient support in delay space to fit the delay spectrum. However, we wish to emphasize that the purpose of deconvolution is to rid the delay spectra of instrumental artifacts to the best extent possible in order to see the underlying signatures of foreground emission. It is not our intention in this paper to use the deconvolution as a foreground removal tool.

4. SIMULATIONS

We describe the instrumental model and foreground models we use in our simulations.

4.1. *Instrument Model*

The instrument model consists primarily of the tile beam power–pattern, $A(\bar{\mathbf{l}}, f)$ (see equation 1). It is modeled as a mutually–coupled 4–by–4 dipole array with the overall power–pattern of each individual dipole calculated via finite element electromagnetic simulation (Sutinjo et al. 2014). To speed up simulation we find that a phased array of isotropic radiators at a height of 0.3 m above an infinite ground plane provides a very good approximation to the full simulation, hence we use the idealized dipoles. We also assume that each individual dipole signal has random delay fluctuations of rms 0.05 ns, a number in line with the known repeatability and stability level of the analog signal chain (Bowman et al. 2007). The power–patterns are shown in Figure 4 for the pointings chosen in our study.

4.2. *Foreground Model*

The MWA has a very wide field of view ($\gtrsim$ 20° at 185 MHz) and has non–zero response all the way to the horizon. As already noted, foreground objects fill the *foreground wedge* with an amplitude which depends on the tile power gain in that direction, with objects at the horizon being closest to the boundaries of the *wedge*.

The degree to which these objects leak power into spectral line of sight modes in the *EoR window* will be shown to depend primarily on the brightness of foregrounds near the boundaries of the *foreground wedge*. Thus, it is important to consider an all–sky model for foreground objects in evaluating the features seen in the power spectrum instead of restricting only to the primary field of view.

We use an all–sky foreground emission model that includes both diffuse and bright compact components.

4.2.1. *Diffuse Foreground Model*

For the diffuse component, we use an all–sky radio foreground model (de Oliveira-Costa et al. 2008) to estimate the emission at 185 MHz. Since this map is predominantly based on the 408 MHz all–sky map of Haslam et al. (1982) which has an angular resolution of 0°.85, we smoothed the 185 MHz map to the same resolution. However, to avoid any artifacts from sampling this map,
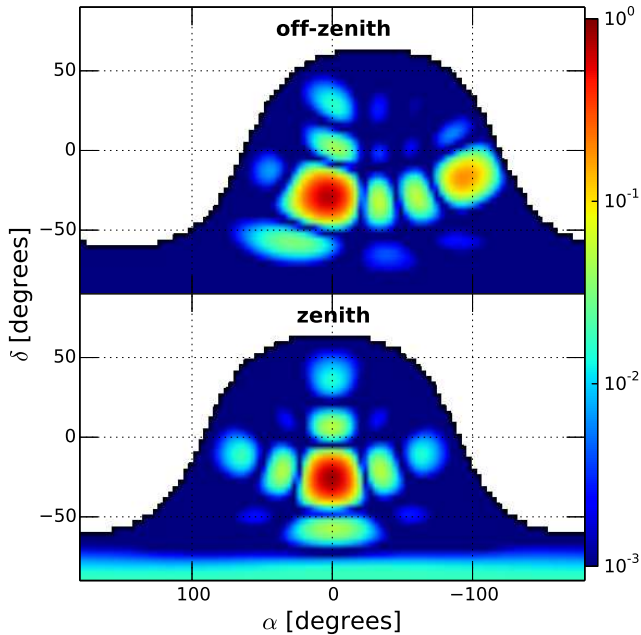
Figure 4. Models of MWA tile power–pattern for *off–zenith* (top) and *zenith* pointings (bottom) at 185 MHz. An MWA tile is modeled as a 4×4 array of isotropic radiators placed 0.3 m above an infinite ground plane. Random fluctuations of rms 0.05 ns have been added to delay corrections during the phased addition of voltages from these isotropic radiators. The logarithmic color scale (shown on the right) is common to both panels.

we sample it at $\approx 27'$ intervals. We model the diffuse foreground spectra with a single spectral index at each pixel in the map, found by comparing model maps at 170 MHz and 200 MHz.

A low resolution version of the diffuse foreground model used is shown in Figure 5. Contours of the MWA tile power–pattern shown in Figure 4 are also overlaid. Of notable significance in the *off–zenith* pointing is the presence of a portion of the Galactic plane and the bright Galactic center in the westward sky, where the MWA tile power gain is significant ($\gtrsim 12\%$). In the *zenith* pointing, the Galactic plane has set and the contour level of the power–pattern in that direction is at least 16 times smaller.

4.2.2. *Compact Foreground Model*

We use a combination of the NRAO VLA Sky Survey (NVSS; Condon et al. 1998) at 1.4 GHz and the Sydney University Molonglo Sky Survey (SUMSS; Bock et al. 1999; Mauch et al. 2003) at 843 MHz to provide our compact object catalog due to their complementary survey footprints covering the entire sky, and matched flux density sensitivity and angular resolution. The SUMSS catalog covers the sky with declination $\delta < -30°$ with a limiting peak brightness of 6–10 mJy/beam and an angular resolution of $\sim 45''$. The NVSS covers the sky with $\delta > -40°$ with a similar angular resolution and a limiting flux density of ≈ 2.5 mJy for compact objects.

From the SUMSS catalog, we select compact objects whose deconvolved major axes are equal to $0''$. From the NVSS catalog, we excluded objects that overlap with those in the SUMSS survey footprint. Compact objects from NVSS were selected if the convolved major axes were not greater than $\approx 47''$, which is almost equal to the
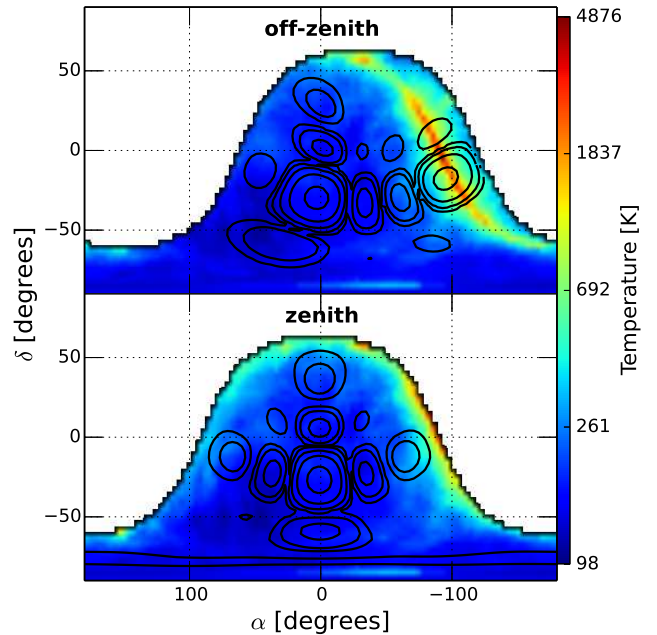

Figure 5. Sky brightness temperature of the diffuse foreground model at 185 MHz visible during *off–zenith* (top) and *zenith* (bottom) pointings. The color scale on the right is logarithmic and is common to both panels. Contours of power–pattern shown in Figure 4 are overlaid. The contour levels shown are 0.00195, 0.00781, 0.0312, 0.125, and 0.5. The Galactic center and a portion of the Galactic plane are prominently visible during the *off–zenith* pointing and the MWA tile power gain is significant ($\gtrsim 12\%$) in that direction. In contrast, emission from the Galactic plane in *zenith* pointing is significantly lesser.

angular resolution of the survey. Using a mean spectral index of $\langle \alpha_{\rm sp} \rangle = -0.83$ (flux density, $S(f) \propto f^{\alpha_{\rm sp}}$) obtained by Mauch et al. (2003) for both NVSS and SUMSS catalog objects, we calculate the corresponding flux densities at 185 MHz, S_{185}. From this subset, we choose compact objects with $S_{185} \geq 10$ Jy. The selection of such bright compact objects is not affected by minor differences in flux density sensitivity of the two surveys. We verified that our selection criteria ensure a similar areal density of objects in the two surveys.

Based on these criteria, we selected 133 objects from the SUMSS catalog and 336 objects from the NVSS catalog. Together with the diffuse foreground model, we obtain an all–sky foreground model consisting of both compact and diffuse emission.

4.3. *Comparison with Data*

With the aforementioned all–sky foreground model, and instrumental and observational parameters, we simulate visibilities using equation 1. Figure 7 shows the amplitude of the delay spectrum from *off–zenith* and *zenith* pointings. Notice the qualitative agreement of amplitude and structure with those obtained from data shown in Figure 3. The Galactic center and the Galactic plane visible in the *off–zenith* pointing make it appear brighter in the *foreground wedge*.

In order to make a quantitative comparison of delay spectra obtained with MWA data and our simulations, we consider the uncertainty in the assumed spectral index of our foreground model. Our foreground models are derived from other higher frequency catalogs and sky
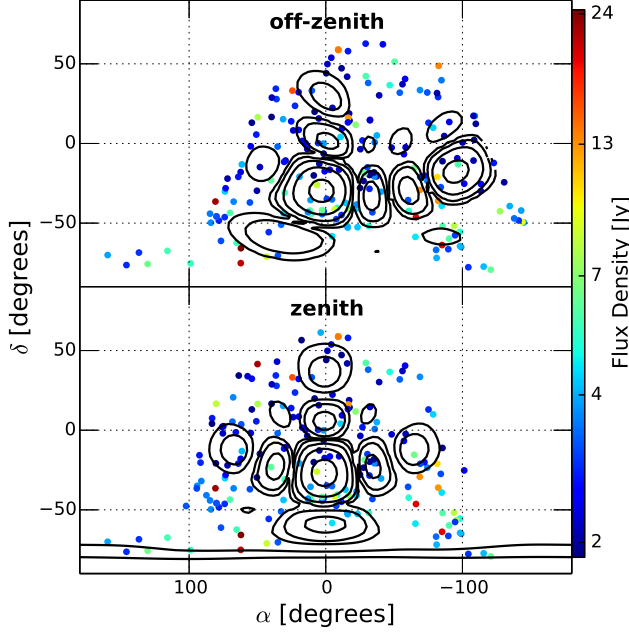
Figure 6. Flux densities of bright compact objects at 185 MHz visible during *off–zenith* (top) and *zenith* (bottom) pointings. The color scale on the right is logarithmic and is common to both panels and corresponds to the color–coded filled circles. Contours of power–pattern shown in Figure 4 are overlaid. The contour levels shown are 0.001953125, 0.0078125, 0.03125, 0.125, and 0.5. The locations of bright compact objects appear to be random.
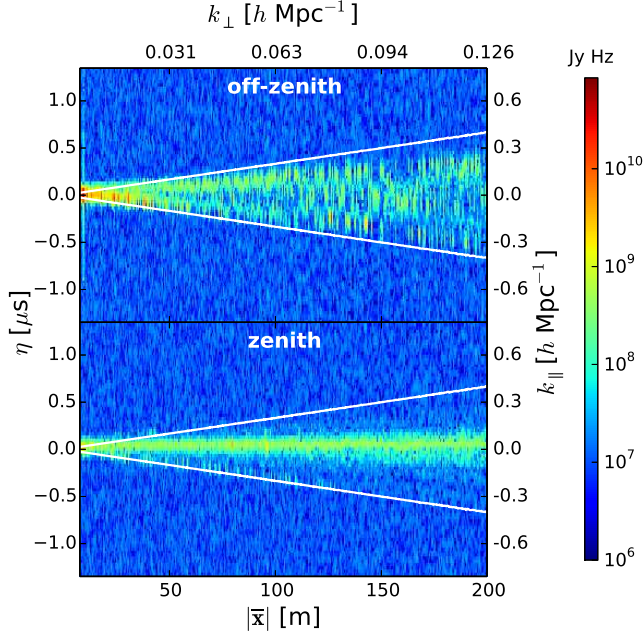
$f = 185$ MHz is the MWA observing frequency, and $\Delta\alpha_{sp}$ is the spread (HWHM) in spectral index. From Mauch et al. (2003), we assume $\Delta\alpha_{sp} \approx 0.35$ for compact objects from NVSS and SUMSS catalogs. Although the model of de Oliveira-Costa et al. (2008) yields a spectral index per direction on the sky, we assume similar uncertainties exist in spectral indices of our diffuse sky model as well, which is predominantly derived from the 408 MHz map of Haslam et al. (1982). Thus, fractional errors in delay spectrum from compact foreground objects and diffuse emission are $\sim 70\%$ and $\sim 30\%$ respectively.

In addition to intrinsic model uncertainty, delay spectra from simulations and data each have fluctuations due to thermal noise in the delay spectrum with rms $\sim 1.4\times10^7$ Jy Hz. We estimate the ratio of delay spectra from data and simulations as $\rho = |V_\eta^D(|\overline{\mathbf{x}}|, \eta)| / |V_\eta^S(|\overline{\mathbf{x}}|, \eta)|$, where superscripts D and S denote data and simulation, respectively.

Figure 8 shows normalized histograms of the distribution of $\log_{10}\rho$ for the *off–zenith* (top panel) and *zenith* (bottom panel) pointings respectively inside the *foreground wedge*. The median absolute deviation from these distributions corresponds to $\sim 90\%$ fractional difference between data and modeling on average in either case. This is in line with the aforementioned uncertainty in foreground models and thermal noise fluctuations in measurements. Currently we are significantly limited by the unavailability of all–sky foreground models at the observing frequency of 185 MHz.
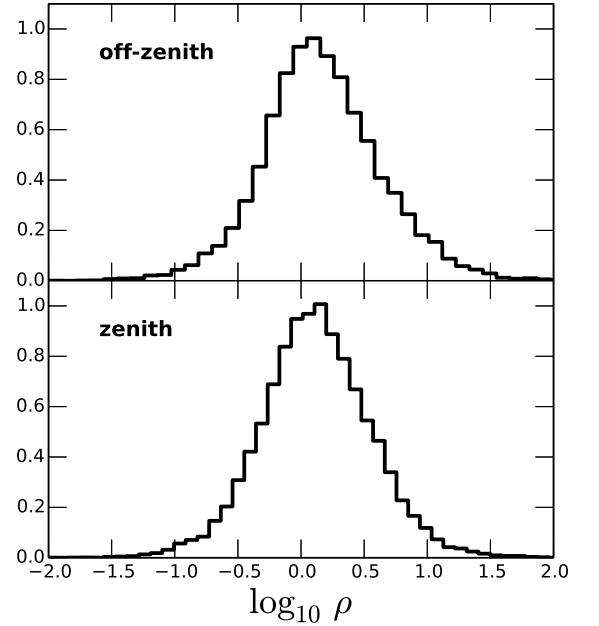

Figure 7. Same as Figure 3, but generated from simulated visibilities using the model described in §4. Note the similarity between the broad features to those in Figure 3.

Figure 8. Normalized histograms of logarithm of ratios of observed to simulated visibilities restricted to the *foreground wedge* – i.e., within the white lines of Figure 7 – for the *off–zenith* (top) and *zenith* (bottom) pointings. Median absolute deviation of $\log_{10}\rho$ is ~ 0.28 in both cases. A value of zero indicates $\rho = 1$.

maps. The inherent spread in spectral index increases the uncertainty while predicting fluxes at the observing frequency. Using simple error propagation, the fractional error in the delay spectrum caused by the spread in spectral index is $\sim \ln(f_{orig}/f)\,\Delta\alpha_{sp}$, where, f_{orig} is the original frequency at which the catalog or map was created,

These uncertainties are presented to confirm the qualitative agreement already seen between Figures 3 and 7. These uncertainty estimates are conservative. A full treatment of all uncertainties and deviations from ideal behavior such as uncertainties in tile power–pattern, cal-

ibration, data corruption due to interference, and wide–field imaging effects will bring the simulations much closer in agreement with observations. Our primary objective is to explore in detail the foreground signatures embedded in the *foreground wedge*.

5. RESULTS

Having established that the simulation matches the data to the level of expected uncertainties, we proceed to examine in further detail the key signatures seen in simulated delay spectra. A number of factors are responsible for the characteristics noted in the delay spectra obtained from data and through simulations. We address these factors below:

- *Sky Model*: Our model of the sky consists of bright compact objects and diffuse emission on diverse spatial scales as shown in Figures 5 and 6. It also consists of localized regions of strong emission such as the Galactic plane and Galactic center. The resulting sky emission is anisotropic. In fact, the patch of sky for MWA observations is chosen from regions of low foreground emission. We will investigate the signatures of each component and their potential impact on *EoR window* contamination.

- *Baseline Orientation*: Since the spatial structure of our foreground model is not expected to be isotropic, we divide our antenna spacings by their orientation, θ_b. We use the following bins: $-67°.5 \leq \theta_b < -22°.5$, $-22°.5 \leq \theta_b < 22°.5$, $22°.5 \leq \theta_b < 67°.5$, and $67°.5 \leq \theta_b < 112°.5$. The bin centers are oriented towards South–East, East, North–East, and North respectively. Since delays depend on θ_b, binning by θ_b allows us to match delay spectrum features to different sky directions.

- *Tile Pointing and Power–Pattern*: Since the MWA tiles are steered electronically, the power–pattern of the tile changes in any observing mode that tracks the object. In our study, we take into account the effect of the changing tile power–pattern on the delay spectrum (Figure 4). The LST of our data ranges from ~ 21 hours through ~ 2 hours, and for our analysis we chose the *off–zenith* and *zenith* pointings shown in Figure 2. We will show the secondary lobe structures in the tile power–pattern are tied to the foreground signatures in Fourier space.

It may be noted that numerous features overlap at varying levels of significance depending on combinations of parameters. We assign the features to their predominant causes. Secondly, we have used noiseless cases to clearly illustrate the observed foreground signatures. With the addition of noise in the visibilities which is subject to observing time, some of the weaker features may not be as prominently visible. Since the foreground signatures are far too numerous and they reside in a vast volume of parameter space such as antenna spacing length and orientation, power–pattern, patch of sky under observation, and instrumental configuration, we highlight only some examples of the most notable features observed in the delay spectra of foreground emission.

5.1. *Diffuse Foreground Signatures*

Figure 9 shows the amplitude of delay spectrum, $|V_\eta(|\bar{\mathbf{x}}|, \eta)|$, simulated with only the diffuse foreground model, omitting compact sources. The top and bottom panels correspond to *off–zenith* and *zenith* pointings, respectively.
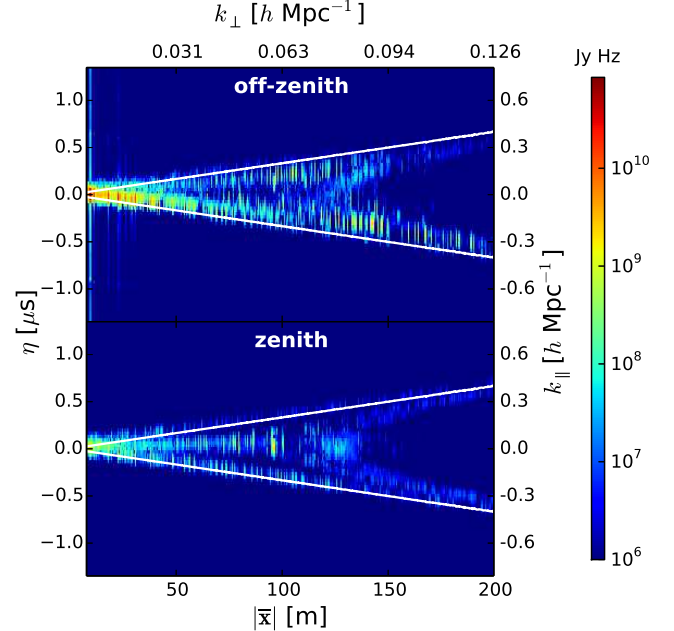

Figure 9. Simulated delay spectrum amplitudes $|V_\eta(|\bar{\mathbf{x}}|, \eta)|$ (in units of Jy Hz) for *off–zenith* (top) and *zenith* (bottom) pointings for the diffuse foreground model without any thermal noise. White lines mark the boundaries of the *foreground wedge* determined by the horizon delay limit and antenna spacing. The axes and color scale are identical to those in Figures 3 and 7. In the *zenith* pointing, delay spectrum from diffuse emission has a *two–pronged fork*–shaped structure and is present even at wide antenna spacings.

Here we examine in detail some examples of interesting features observed with the diffuse foreground model.

5.1.1. *Galactic Center on Eastward Antenna Spacings*

The most prominent signature seen in the *off–zenith* pointing (top panel of Figure 9) is due to the bright Galactic center situated on the western horizon co–located with one of the bright secondary lobes of the power–pattern. It appears as a bright branch in the delay spectrum near the negative delay horizon delay limit. This feature is strongest at short antenna spacings and fades with increasing antenna spacing. The signature is absent in the *zenith* pointing (bottom panel, Figure 9) because the Galactic center is below the horizon.

5.1.2. *Diffuse Emission Delay Spectra*

Figure 9 also demonstrates diffuse emission from within the primary field of view manifests in the *off–zenith* pointing as a branch at the top of the wedge, and at zero delay in the *zenith* pointing. The former is seen at positive delays because the primary lobe of the power–pattern is centered eastward of zenith, whereas in the latter it is centered at zenith. As we see from Equation 1, each baseline measures a single spatial mode on the

sky with an angular size scale inversely proportional to the length of the baseline projected in the direction of the emission. Thus, in the *zenith* pointing, the signature of the smooth sky model, the horizontal line at zero delay, fades away on antenna spacings $|\bar{\mathbf{x}}| \gtrsim 125$ m because the sky model is devoid of spatial structures on scales $\lesssim 0°\!.75$. In the off-zenith pointing, we see power out to much longer baselines which is described below.

5.1.3. *Diffuse Emission on Wide Antenna Spacings*

A very interesting signature of diffuse emission is revealed at regions near the positive horizon delay limit in the *off–zenith* pointing and at both positive and negative horizon delay limits in the *zenith* pointing even on wide antenna spacings (125 m $\lesssim |\bar{\mathbf{x}}| \leq 200$ m). This is contrary to expectations because longer baselines are not expected to be sensitive to diffuse emission. In the wide–field observations of the MWA, long baselines appear shortened due to projection effects for off–axis emission, especially towards the horizon. This effect is evident at all LSTs in our simulations and it naturally holds for short baselines, as well. Thus, diffuse emission from far off–axis directions manifests as an edge–heavy *two–pronged fork* across all baselines. It decreases in strength with increasing baseline length as expected but is present in all baseline orientations.

5.2. *Compact Foreground Signatures*

Figure 10 shows the delay spectrum of the compact foreground sky model (see Figure 6) for *off–zenith* and *zenith* pointings in the top and bottom panels, respectively.

In contrast to the delay spectrum of diffuse emission, compact emission manifests as a center–heavy structure in the delay spectrum with either pointing. This contributes to the *pitchfork* feature which will be discussed in detail in §5.3. Below we summarize some specific notable features obtained with our compact objects foreground model.

The amplitude response of an interferometer to a compact object is, to first order, flat with respect to baseline length. Since the main lobe of the power–pattern in the *off–zenith* pointing is centered eastward of zenith, the bulk of the compact foreground emission is seen in a branch with positive delay corresponding to the position of the primary lobe of the power–pattern. In the *zenith* pointing, compact emission from the same patch of sky is seen as a bright horizontal arm at zero delay since the primary lobe of the power–pattern is centered at zenith.

Foreground emission in zero delay and negative delay regions in the *off–zenith* pointing is caused by compact objects co–located with secondary lobes of the power–pattern. On the other hand, compact objects co–located with secondary lobes of power–pattern in the *zenith* pointing are revealed as faint but distinct branches at positive and negative delays depending on the orientation of antenna spacing and direction of emission on the sky.

5.3. *The "Pitchfork"*

Delay spectra from the all–sky foreground model in our study display a composite feature set drawn from the features of compact and diffuse foreground models. Here we
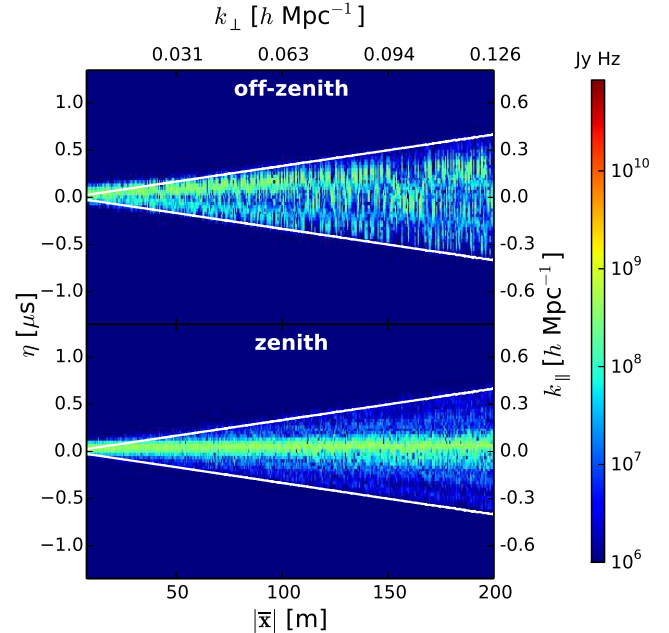

Figure 10. Simulated delay spectrum amplitudes $|V_\eta(|\bar{\mathbf{x}}|, \eta)|$ (in units of Jy Hz) for *off–zenith* (top) and *zenith* (bottom) pointings for the compact objects foreground model without any thermal noise. White lines mark the boundaries of the *foreground wedge* determined by the horizon delay limit and antenna spacing. The axes and logarithmic color scale are identical to those in Figures 3, 7 and 9. Emission from compact objects gives rise to centrally concentrated features and stays flat across antenna spacings in the *foreground wedge*.

compare the relative strengths of emission from different spatial scales in our all–sky composite foreground model.

When not dominated by the bright emission from the Galactic center, which happens only for certain pointings at certain times, the delay spectrum of the combined foreground model is dominated, in somewhat equal measure, by diffuse and compact emission. This is illustrated by a more detailed examination of the *zenith* pointing in our study.

Figure 11 shows delay spectra of three antenna pairs of different antenna spacings oriented northward during the *zenith* pointing; each is a different vertical slice of the two dimensional delay spectrum plots shown in Figures 3, 7, 9 and 10. The diffuse, compact, and composite components are are shown as dotted, dashed, and solid lines, respectively. The horizon delay limits are shown as a pair of vertical dot–dashed lines. The gray shaded area denotes the envelope of expected uncertainty in the delay spectrum. Uncertainty in emission in the *foreground* wedge (between the horizon delay limits) is dominated by the $\gtrsim 70\%$ error owing to uncertainties in predicting the spectral index of compact foreground objects, while thermal noise fluctuations dominate outside.

The zero–delay peak (corresponding to the main lobe in the power–pattern) with a value of $\sim 10^8$–10^9 Jy Hz, independent of antenna spacing, is predominantly determined by compact foreground objects. The corresponding peak at zero–delay from diffuse foreground model is at least 10 times fainter and decreases drastically with an increase in antenna spacing. This is the response expected from different antenna spacings towards compact and diffuse foreground objects.
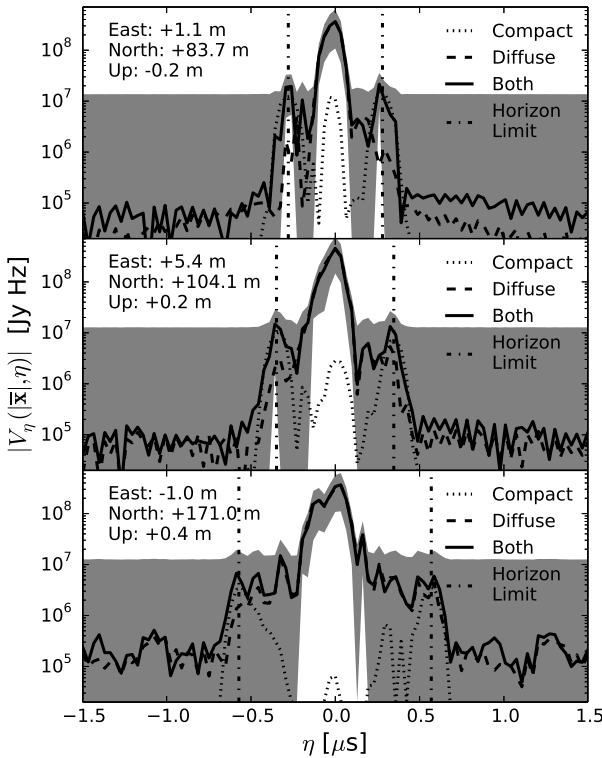
Figure 11. Simulated delay spectrum amplitudes $|V_\eta(\overline{\mathbf{x}}, \eta)|$ for three chosen antenna spacings of length: ~ 84 m (*top*), ~ 105 m (*middle*), and ~ 170 m (*bottom*). The antenna spacing is specified in each panel along East, North and Upward directions. The dotted, dashed and solid lines denote contributions from diffuse emission, compact objects and composite foreground model respectively. Vertical dot–dashed lines mark the horizon delay limits. The shaded region denotes the envelope of uncertainty around the composite model delay spectrum. This uncertainty is dictated by uncertainties in spectral indices of foreground models inside the delay horizon delay limits. Outside the horizon delay limits, it is predominantly due to thermal noise fluctuations. Compact objects dominate the central regions of delay spectrum while both components, especially diffuse emission on short antenna spacings, dominate near the horizon delay limits, giving rise to a *pitchfork*–shaped structure.

Near the horizon delay limits at shorter antenna spacing, diffuse emission is brighter than that from compact objects. Here, diffuse emission does not decrease as rapidly with increasing antenna spacing as was seen at zero–delay. In fact, even on widely spaced antennas, diffuse emission in the delay spectrum near the horizon delay limits exceeds that in the main lobe by more than two orders of magnitude. We attribute this to apparent shortening of antenna spacings projected in the direction of diffuse emission entering far from the primary field of view. This demonstrates the edge–heavy *two–pronged fork* feature discussed in §5.1.3.

Full simulations show the combination of center–heavy features on all antenna spacings from compact objects and edge–heavy features from both components especially diffuse emission even on wide antenna spacings. This results in a characteristic *pitchfork* structure imprinted in the *foreground* wedge and should be evident in observations.

All these observations are consistent with those from Figures 9 and 10. The observability of the *pitchfork* signature predicted in this paper depends on the relative levels of uncertainty in the foreground model and fluctuations from thermal noise. In our simulations, since thermal noise in these very short duration snapshots is $\gtrsim 10^7$ Jy Hz, and features near the horizon delay limits are $\sim 10^7$ Jy Hz, the *pitchfork* feature is not expected to be detected in a noisy scenario although this feature is marginally visible in the *zenith* pointing of observed data (see Figure 3). We attribute this to differences between our foreground model and the actual sky. Deeper observations should reveal the feature clearly.

We also note that increasing the antenna spacing results in progressively improving the resolution along the delay axis by increasing the number of delay bins inside the *foreground wedge*. This improves the localization of foreground objects whose signatures are imprinted in the delay spectrum. For instance, there is an increase in the number of secondary peaks in the delay spectrum between zero–delay and horizon delay limits as the antenna spacing increases from ~ 84 m to ~ 170 m. In this case, these correspond to secondary lobes of the power–pattern that lie between the primary lobe and the horizon along the local meridian. At short antenna spacings, due to relatively lower resolution along the delay axis inside the *foreground wedge* and a consequent loss of localization of foreground emission, these secondary peaks blend in with other major peaks and are not distinctly visible. These findings are consistent with the discussion in §5.2.

In summary, the brightest signature ($\gtrsim 10^{10}$ Jy Hz) is that of the Galactic center in the *off–zenith* pointing co–located at a westward secondary lobe with a significantly high gain. The next brightest signature ($\sim 10^8$–10^9 Jy Hz) is caused by compact emission appearing to be concentrated in the inner regions of the *foreground wedge* (near zero delays) rather than towards the boundaries. Diffuse emission co–located at the primary lobe of the power–pattern is fainter by approximately an order of magnitude relative to compact emission from the same region for a ~ 84 m antenna spacing. But unlike the latter, diffuse emission decreases rapidly by over two orders of magnitude as the antenna spacing is widened to ~ 170 m. However, diffuse emission is significantly higher near horizon delay limits compared to that in the primary lobe of the power–pattern as the antenna spacing is increased. This leads to edge–heavy features in the delay spectrum. Complemented by the center-heavy compact foreground features, we predict the presence of a *three–pronged pitchfork*–shaped signature in the delay spectrum of the foreground sky. We identify evidence for presence of this signature in actual MWA data, although deeper observations are needed to more clearly reveal the effect.

6. APPLICATIONS

Here, we will investigate the susceptibility of particular antenna spacings to foreground contamination arising out of bright foreground objects located near the horizon and present a technique to substantially mitigate such contamination.

The Galactic center in the *off–zenith* pointing is one such example already available in our study. Figure 12a shows the sky model (*top:* compact objects, *bottom:* dif-

fuse sky model) in this pointing. The Galactic center is the most dominant source of foreground contamination from the diffuse sky model and is co–located with a bright secondary lobe of the power pattern near the western horizon. Figure 12b shows the sky mapped to delays registered by the baseline vectors, of length 100 m for instance, oriented northward (top panel) and eastward (bottom panel). Since the Galactic center is located in the western sky, this figure demonstrates that it is observed at negative delay on a baseline vector oriented eastward and at zero delay on a baseline vector oriented northward. Figure 12c shows the delay spectra on baselines oriented northward ($67°\!.5 \leq \theta_b < 112°\!.5$) at the top and eastward ($-22°\!.5 \leq \theta_b < 22°\!.5$) at the bottom. The Galactic center manifests itself most distinctly near the negative horizon delay limit on short eastward baselines in the delay spectrum (bottom panel of Figure 12c). Consequently, the spillover caused by the instrument's transfer function from the *foreground wedge* into the *EoR window* is most severe in eastward baselines relative to northward baselines evident by the bright vertical stripes of foreground contamination.

The delay spectrum, $V_\eta(\overline{\mathbf{x}}, \eta)$, not only carries information on spatial scales of foreground emission but also offers the unique advantage of viewing the sky through a combination of antenna spacing vectors and the delay axis. With a foreground model known *a priori* in which structures and locations of very bright foreground objects such as the Galactic center or AGN are available, we can predict their response across antenna spacings as a function of observing parameters such as LST, power–pattern, etc. This allows us to programmatically screen data for antenna spacings that are severely contaminated by foregrounds near the horizon delay limits in the delay spectrum. These can be weighted appropriately during data analysis. We demonstrate such a screening technique, whereby we use the bright object's location and structure to discard antenna spacings of certain lengths and orientations to mitigate foreground contamination in the *EoR window*.

In our example, we discard eastward antenna spacings ($-22°\!.5 \leq \theta_b < 22°\!.5$) of lengths $|\overline{\mathbf{x}}| < 40$ m. Figure 13 shows the delay spectra obtained with all antenna spacings (top panel) and after applying our screening technique (bottom panel) on the *off–zenith* observation. Notice the remarkable reduction — by about an order of magnitude — in foreground contamination in this region after application of our screening technique.

This screening technique can be customized when it is necessary to optimize between foreground mitigation and loss of sensitivity from discarding data. Figure 14 shows how the typical foreground contamination in the *EoR window* depends on the orientations and lengths of discarded antenna spacings. We choose antenna spacings oriented eastward to varying degrees of directedness, i.e., $-7°\!.5 \leq \theta_b < 7°\!.5$ (solid circles), $-15° \leq \theta_b < 15°$ (solid squares), and $-22°\!.5 \leq \theta_b < 22°\!.5$ (solid stars). Among antenna spacings that satisfy these criteria, we discard data from those whose lengths are shorter than $|\overline{\mathbf{x}}|_{max}$ (x–axis) and show foreground contamination estimated in the *EoR window* from all remaining antenna spacings. Foreground contamination estimated from only northward (dashed line) and eastward (dot–dashed line) antenna spacings represent the lower and upper limits, re-

spectively, and are shown for reference. It is seen that foreground contamination can be mitigated by a factor between ~ 1.5 ($|\theta_b| \leq 7°\!.5$) and ~ 10 ($|\theta_b| \leq 22°\!.5$). It is notable that short antenna spacings play maximum role in determining foreground contamination. For instance, among antenna spacings with orientations $|\theta_b| \leq 22°\!.5$, discarding those with lengths $|\overline{\mathbf{x}}| > 40$ m has no effect on further foreground contamination and would only lead to loss of sensitivity. On the other hand, a bright compact object at the same location will give rise to foreground contamination even on longer antenna spacings. Such cases will necessitate discarding more or all of the eastward oriented antenna spacings.

In principle, instead of discarding selected antenna spacings altogether, we could down–weight them based on an optimal scheme. This technique provides a very simple and yet effective tool in adding a layer of control to mitigate effects of foreground contamination in EoR data analysis.

7. SUMMARY

Our primary motivation in this work is to understand how the various bright foregrounds will manifest in three dimensional power spectrum of HI from 21 cm reionization observations. In the units presented here, the dynamic range between bright foregrounds and the 21 cm signal is expected to be $\sim 10^8$; a detailed understanding of how foregrounds can corrupt the 21 cm power spectrum is therefore essential. This analysis extends previous work by simulating the entire sky rather than just the central field of view and by providing a comparison with early observations from the MWA. By making use of the delay spectrum technique to estimate the power spectrum, we are able to observe the effects of foregrounds while avoiding entanglements with more complex power spectrum estimators.

Simulating in all important respects the response of the MWA to an all–sky foreground model that consists of diffuse Galactic emission from de Oliveira-Costa et al. (2008) and bright compact objects from the NVSS and SUMSS catalogs, we confirm that the modeled delay spectra are in agreement with data obtained with the MWA, to the extent allowed by levels of uncertainty known in the foreground models and thermal noise fluctuations in measurements.

Looking deeper at our simulations, below the level of the instrumental noise, we identify numerous signatures of different components of foreground emission seen in the delay spectra. We establish the relationship between these signatures and observing parameters such as antenna pointing and LST, instrument parameters such as antenna power–pattern and bandpass shape, and foreground parameters such as the nature of emission, spectral index, etc.

The bright Galactic center at the edge of the western horizon co–located with one of the far secondary lobes of MWA tile power–pattern is the brightest source of foreground contamination in the *off–zenith* pointing. It manifests itself near the negative horizon delay limit in the delay spectrum on antenna spacings oriented eastward.

As expected, diffuse emission in the primary beam of the antenna power–pattern is prominent on shorter antenna spacings. However, the most interesting result is its footprint on wide antenna spacings near the horizon
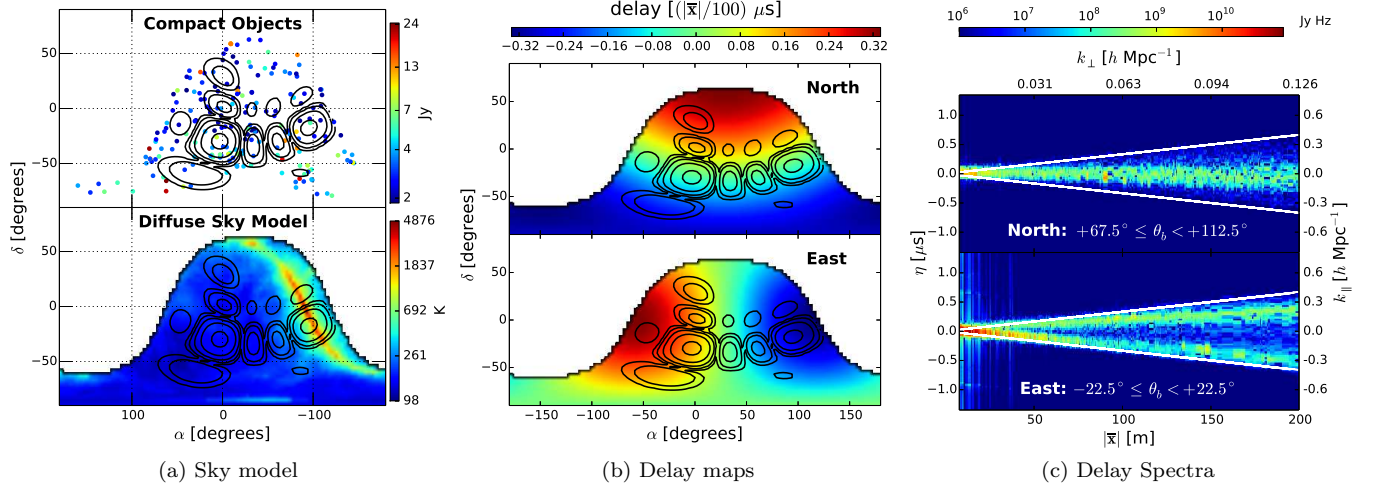
(a) Sky model

(b) Delay maps

(c) Delay Spectra

Figure 12. (a): Sky model showing compact objects (*top*) and diffuse emission (*bottom*). Contours of power pattern are overlaid whose levels are identical to those in Figures 5 and 6. The Galactic center is very prominent in diffuse emission in the western sky and is co–located with a bright secondary lobe of the power pattern. (b): Sky hemisphere mapped to delays observed on antenna spacings with northward (*top*) and eastward (*bottom*) orientations. Color scale shown is for a 100 m antenna spacing. Delays vary linearly with antenna spacing length. The bright Galactic center in the western sky is recorded at zero delays in northward antenna spacings and close to negative horizon delay limit on eastward antenna spacings. (c): Simulated delay spectra $|V_\eta(\overline{\mathbf{x}}, \eta)|$ on antenna spacings oriented northward (*top*) and eastward (*bottom*). The bright Galactic center is prominently visible close to negative horizon delay limit on eastward antenna spacings. These are also the most severely contaminated by foreground spillover. The northward antenna spacings, on the other hand, are the least contaminated.
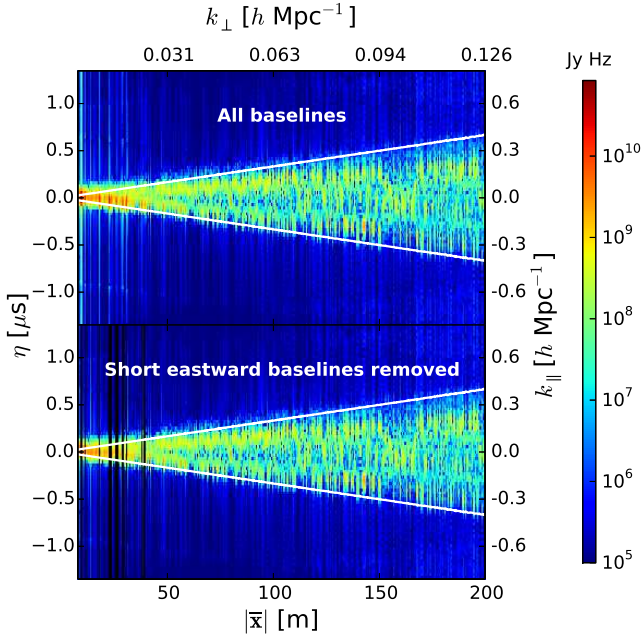
Figure 13. Simulated delay spectra amplitudes $|V_\eta(\overline{\mathbf{x}}, \eta)|$ for the *off–zenith* pointing with all antenna spacings included (*top*) and with short eastward antenna spacings discarded (*bottom*). Discarded antenna spacings (black vertical stripes) have lengths $|\overline{\mathbf{x}}| < 40$ m and orientations $-22°5 \leq \theta_\mathrm{b} < 22°5$. The spillover from the bright Galactic center near the negative horizon delay limit from the *foreground wedge* prominent when all antenna spacings are used is lowered by an order of magnitude when short eastward antenna spacings are discarded. Color scale has been adjusted to emphasize mitigation achieved in foreground contamination.

Figure 14. Foreground contamination in *EoR window* for the *off–zenith* pointing as a function of orientation and length of discarded antenna spacings. Antenna spacings oriented eastward with varying degrees of directedness are considered: $-7°5 \leq \theta_\mathrm{b} < 7°5$ (solid circles), $-15° \leq \theta_\mathrm{b} < 15°$ (solid squares), and $-22°5 \leq \theta_\mathrm{b} < 22°5$ (solid stars). Among antenna spacings that satisfy the above criteria, ones with lengths $\leq |\overline{\mathbf{x}}|_\mathrm{max}$ (*x*–axis) are discarded. Lower (dashed line) and upper (dot–dashed) limits on foreground contamination are derived from northward ($67°5 \leq \theta_\mathrm{b} < 112°5$) and eastward ($|\theta_\mathrm{b}| \leq 22°5$) antenna spacings respectively. Foreground contamination in the *EoR window* reduces by factor ~ 1.5 ($|\theta_\mathrm{b}| \leq 7°5$) to ~ 10 ($|\theta_\mathrm{b}| \leq 22°5$). Short baselines have maximum effect on foreground contamination. For instance, discarding antenna spacings oriented $|\theta_\mathrm{b}| \leq 22°5$ with lengths $|\overline{\mathbf{x}}| > 40$ m has no effect in further reducing foreground contamination.

delay limits, an edge–heavy *two–pronged fork*–shaped signature. This is due to apparent shortening of wide antenna spacings in the direction of foreground emission far off–axis, thereby retaining their response to extended emission. On the other hand, compact foreground ob-
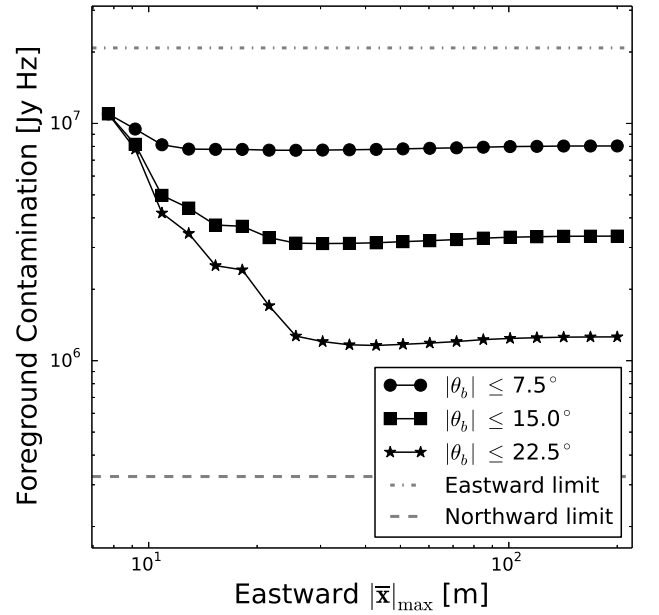
jects predominantly map onto central regions of the *fore-*

ground wedge. Features arising from compact objects co–located with primary and secondary lobes of the antenna power–pattern have been identified. In general, delay spectrum signatures of compact objects are center–heavy, in clear contrast to those from diffuse emission. A composite all–sky foreground model consisting of diffuse and compact foregrounds combines the characteristic individual shapes into a *pitchfork*–shaped structure in the *foreground wedge*. This will be distinctly visible when the thermal noise floor is sufficiently lowered, as longer observations are processed through the MWA analysis pipelines.

In conclusion, we find that inclusion of objects, both diffuse and compact, all the way to the horizon is essential to explaining the observed power spectrum. We also provide a simple and effective tool based on the delay spectrum technique to mitigate foreground contamination by an order of magnitude in EoR data analysis. It is possible to discard or down–weight antenna pairs most affected by foreground contamination as a function of antenna spacing vector and time of observation.

REFERENCES

Beardsley, A. P., Hazelton, B. J., Morales, M. F., et al. 2012, MNRAS, 425, 1781
—. 2013, MNRAS, 429, L5
Bock, D. C.-J., Large, M. I., & Sadler, E. M. 1999, AJ, 117, 1578
Bowman, J. D., Morales, M. F., & Hewitt, J. N. 2006, ApJ, 638, 20
—. 2009, ApJ, 695, 183
Bowman, J. D., Barnes, D. G., Briggs, F. H., et al. 2007, AJ, 133, 1505
Bowman, J. D., Cairns, I., Kaplan, D. L., et al. 2013, PASA, 30, 31
Condon, J. J., Cotton, W. D., Greisen, E. W., et al. 1998, AJ, 115, 1693
Datta, A., Bowman, J. D., & Carilli, C. L. 2010, ApJ, 724, 526
de Oliveira-Costa, A., Tegmark, M., Gaensler, B. M., et al. 2008, MNRAS, 388, 247
Di Matteo, T., Perna, R., Abel, T., & Rees, M. J. 2002, ApJ, 564, 576
Dillon, J. S., Liu, A., & Tegmark, M. 2013, Phys. Rev. D, 87, 043005
Furlanetto, S. R., Oh, S. P., & Briggs, F. H. 2006, Phys. Rep., 433, 181
Haslam, C. G. T., Salter, C. J., Stoffel, H., & Wilson, W. E. 1982, A&AS, 47, 1
Hazelton, B. J., Morales, M. F., & Sullivan, I. S. 2013, ApJ, 770, 156
Iliev, I. T., Shapiro, P. R., Ferrara, A., & Martel, H. 2002, ApJ, 572, L123
Liu, A., & Tegmark, M. 2011, Phys. Rev. D, 83, 103006
Liu, A., Tegmark, M., Bowman, J., Hewitt, J., & Zaldarriaga, M. 2009, MNRAS, 398, 401
Lonsdale, C. J., Cappallo, R. J., Morales, M. F., et al. 2009, IEEE Proceedings, 97, 1497
Madau, P., Meiksin, A., & Rees, M. J. 1997, ApJ, 475, 429
Mauch, T., Murphy, T., Buttery, H. J., et al. 2003, MNRAS, 342, 1117
Morales, M. F., Bowman, J. D., & Hewitt, J. N. 2006, ApJ, 648, 767
Morales, M. F., Hazelton, B., Sullivan, I., & Beardsley, A. 2012, ApJ, 752, 137
Morales, M. F., & Hewitt, J. 2004, ApJ, 615, 7
Offringa, A. R., de Bruyn, A. G., Biehl, M., et al. 2010, MNRAS, 405, 155
Offringa, A. R., van de Gronde, J. J., & Roerdink, J. B. T. M. 2012, A&A, 539, A95
Parsons, A. R., & Backer, D. C. 2009, AJ, 138, 219
Parsons, A. R., Pober, J. C., Aguirre, J. E., et al. 2012, ApJ, 756, 165
Parsons, A. R., Backer, D. C., Foster, G. S., et al. 2010, AJ, 139, 1468
Parsons, A. R., Liu, A., Aguirre, J. E., et al. 2014, ApJ, 788, 106
Pober, J. C., Parsons, A. R., Aguirre, J. E., et al. 2013, ApJ, 768, L36
Pober, J. C., Liu, A., Dillon, J. S., et al. 2014, ApJ, 782, 66
Scott, D., & Rees, M. J. 1990, MNRAS, 247, 510
Sullivan, I. S., Morales, M. F., Hazelton, B. J., et al. 2012, ApJ, 759, 17
Sunyaev, R. A., & Zeldovich, Y. B. 1972, A&A, 20, 189
Sutinjo, A., Padhi, S., Wayth, R., et al. 2014, in Antenna Technology and Applied Electromagnetics (ANTEM), 2014 16th International Symposium on, 1–2
Taylor, G. B., Carilli, C. L., & Perley, R. A., eds. 1999, Astronomical Society of the Pacific Conference Series, Vol. 180, Synthesis Imaging in Radio Astronomy II
Thyagarajan, N., Udaya Shankar, N., Subrahmanyan, R., et al. 2013, ApJ, 776, 6
Tingay, S. J., Goeke, R., Bowman, J. D., et al. 2013, PASA, 30, 7
Tozzi, P., Madau, P., Meiksin, A., & Rees, M. J. 2000, ApJ, 528, 597
Trott, C. M., Wayth, R. B., & Tingay, S. J. 2012, ApJ, 757, 101
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