

## IMAGING THE EPOCH OF REIONIZATION: LIMITATIONS FROM FOREGROUND CONFUSION AND IMAGING ALGORITHMS

HARISH VEDANTHAM, N. UDAYA SHANKAR, AND RAVI SUBRAHMANYAN

Raman Research Institute, Bangalore, India; [harish@rri.res.in](mailto:harish@rri.res.in)

Received 2011 April 29; accepted 2011 October 3; published 2012 January 17

### ABSTRACT

Tomography of redshifted 21 cm transition from neutral hydrogen using Fourier synthesis telescopes is a promising tool to study the Epoch of Reionization (EoR). Limiting the confusion from Galactic and extragalactic foregrounds is critical to the success of these telescopes. The instrumental response or the point-spread function (PSF) of such telescopes is inherently three dimensional with frequency mapping to the line-of-sight (LOS) distance. EoR signals will necessarily have to be detected in data where continuum confusion persists; therefore, it is important that the PSF has acceptable frequency structure so that the residual foreground does not confuse the EoR signature. This paper aims to understand the three-dimensional PSF and foreground contamination in the same framework. We develop a formalism to estimate the foreground contamination along frequency, or equivalently LOS dimension, and establish a relationship between foreground contamination in the image plane and visibility weights on the Fourier plane. We identify two dominant sources of LOS foreground contamination—“PSF contamination” and “gridding contamination.” We show that PSF contamination is localized in LOS wavenumber space, beyond which there potentially exists an “EoR window” with negligible foreground contamination where we may focus our efforts to detect EoR. PSF contamination in this window may be substantially reduced by judicious choice of a frequency window function. Gridding and imaging algorithms create additional gridding contamination and we propose a new imaging algorithm using the Chirp Z Transform that significantly reduces this contamination. Finally, we demonstrate the analytical relationships and the merit of the new imaging algorithm for the case of imaging with the Murchison Widefield Array.

*Key words:* cosmology: observations – methods: analytical – methods: data analysis – techniques: interferometric

### 1. INTRODUCTION

Several synthesis telescopes such as the Murchison Widefield Array (MWA),<sup>1</sup> Low Frequency Array (LOFAR),<sup>2</sup> and Precision Array to Probe of Reionization (PAPER)<sup>3</sup> are currently being built to study the Epoch of Reionization (EoR) by observing the redshifted 21 cm line transition from neutral hydrogen. While future instruments may map the 21 cm brightness in three dimensions (frequency corresponds to line of sight (LOS)), current experiments lack the required sensitivity and instead hope to estimate the power spectrum of 21 cm brightness fluctuations. The brightness of these fluctuations is about five orders of magnitude smaller than the Galactic and extragalactic foregrounds. Moreover, removing the effects of the telescope point-spread function (PSF) from the synthesis images may not be possible to the level of sensitivity demanded by EoR power spectrum measurements. Consequently, foreground contamination and instrumental effects present a major hurdle to 21 cm tomography.

A host of simulations have demonstrated the potential in current instruments to recover the EoR power spectrum despite thermal uncertainties, foreground contaminants, and instrumental effects. Among instrumental effects, simulations have shown the frequency dependence of PSF to be of particular concern in EoR power spectrum estimation (Bowman et al. 2009). In 21 cm tomography, frequency maps to the LOS. Most of the proposed foreground subtraction algorithms rely on the ability to distinguish smooth spectral features of Galactic

and extragalactic contaminants from the relatively rapid spectral fluctuations (expected from theory) of the 21 cm EoR signal. The sidelobe response of confusing foreground contaminants may vary with frequency and generate an instrumental spectral structure at each image pixel—a phenomenon that has been dubbed “mode-mixing” in the literature (Bowman et al. 2009). Apart from mode-mixing, calibration errors and choice of gridding algorithms can also lead to undesirable instrumental response in the frequency domain.

While simulations have evaluated such mode-mixing effects for realistic arrays like the MWA (Liu et al. 2009; Bowman et al. 2009), this paper presents a more general, analytical approach. In particular, the relationship between the PSF in image coordinates ( $lmf$ ) and the telescope sampling function in Fourier coordinates ( $uvf$ ) is established, and the mode-mixing problem is cast in Fourier space where the telescope measurements are really made. In this formalism, implications for choice of experimental parameters and foreground subtraction algorithms are discussed. Apart from the analytical treatment, various practical data processing operations that influence the PSF are discussed using an example of snapshot imaging with the MWA.

The remainder of the paper is organized as follows. Section 2 defines the problem analytically and derives an expression relating the LOS foreground contamination to the weighting function in Fourier space. To better understand the relationships established in Section 2, Section 3 applies them to the simple case of imaging a point source with a uniform visibility distribution. The analytical treatment in Section 2 does not take into account the effects of discretization of visibilities in the re-gridding and imaging algorithms. Section 4 discusses these effects and complements the results of Section 2. In particular, Section 4 describes a new imaging algorithm based on the Chirp

<sup>1</sup> See <http://www.mwatelescope.org>.

<sup>2</sup> See <http://www.lofar.org>.

<sup>3</sup> See <http://astro.berkeley.edu/%7Edbacker/eor>.

Z Transform (CZT) that alleviates the problem of foreground contamination. Section 5 demonstrates the concepts developed in Sections 2, 3, and 4 through simulating a snapshot observation with the MWA. Section 6 draws up conclusions and proposals for future work.

## 2. ANALYTICAL TREATMENT

We first analytically evaluate the effect of a frequency-variant PSF at a given image pixel due to an unsubtracted point source by evaluating its contribution to foreground contamination in LOS wavenumber space. We call this contamination PSF contamination. A single point source is considered here to attain an intuitive understanding of PSF contamination. Generalization for a realistic sky will be made in subsequent sections. Array element locations are assumed to be coplanar. This reduces the PSF from a generic four-dimensional function (three spatial dimensions and one frequency dimension) to a three-dimensional function (two spatial dimensions and one frequency dimension) for simplicity.

### 2.1. Notation

The baselines are defined on a local plane whose origin is at the latitude and longitude of the array center. A two-dimensional vector  $\mathbf{x} = [x \ y]$  (in meter units) represents the projected baseline on the local plane. The corresponding baseline  $\mathbf{u} = [u \ v]$  (in wavelengths) at a frequency  $f$  is then given by  $\mathbf{u} = \mathbf{x} \cdot (f/c)$ , where  $c$  is the speed of light. A two-dimensional direction vector  $\mathbf{l}$  toward any direction is represented by its direction cosines:  $\mathbf{l} = [l \ m]$ . An interferometer observation gives a visibility cube with two spatial axes representing the baseline and a frequency axis. An image cube on the other hand has two axes representing direction cosines on the sky and a third frequency axis. Assuming coplanarity of array elements, within the framework of Fourier synthesis imaging,  $\mathbf{u}$  and  $\mathbf{l}$  are Fourier conjugates.

### 2.2. Weighting Function

The telescope sampling function is represented as  $\bar{W}(\mathbf{x}, f)$ , which is the weight assigned to the baseline represented by the vector  $\mathbf{x}$ , and includes any visibility taper applied prior to imaging. In general,  $\bar{W}(\mathbf{x})$  depends on instrumental, observational, and processing parameters. Among instrumental parameters is the antenna layout that determines the set of available baselines. Observational parameters, which include pointing direction, duration of rotation synthesis, and mode of observation—drift scan or track mode, define the  $uv$  coverage. Observational parameters along with the antenna layout determines the natural weight at each  $\mathbf{x}$ , which is proportional to the number of measured visibilities falling into the pixel centered at  $\mathbf{x}$ . Among processing parameters are the visibility grid size and extent, convolution kernel used to grid the visibilities, and any visibility mask and/or taper used during imaging. The weighting function  $\bar{W}(\mathbf{x}, f)$  may be factored into a frequency-independent geometric component  $W(\mathbf{x})$ , a frequency-dependent taper component  $T(\mathbf{x}, f)$ , and a frequency window function  $B(f)$ :

$$\bar{W}(\mathbf{x}, f) = W(\mathbf{x}) T(\mathbf{x}, f) B(f). \quad (1)$$

Instrumental and observational parameters are geometric in nature and hence frequency independent. They may be sufficiently represented by  $W(\mathbf{x})$ . Processing parameters like visibility taper

function and visibility masks are frequency dependent and are represented by  $T(\mathbf{x}, f)$ .  $B(f)$  is the window function in frequency assigned to each visibility.

### 2.3. Derivation

An image made with a weighted baseline distribution  $\bar{W}(\mathbf{x}, f)$  has a frequency-dependent PSF  $\tilde{A}(\mathbf{l}, f)$ , where  $\mathbf{l}$  is the two-dimensional vector representing direction on the sky. Tilde notation is used for image domain functions as opposed to their Fourier domain counterparts. Consider the sidelobe contribution at zenith ( $\mathbf{l} = [0 \ 0]$ ) from a source with a flux density of unity and spectral index of zero located at any  $\mathbf{l}$ . This contribution  $\tilde{S}(\mathbf{l}, f)$  varies with frequency due to a changing PSF. We wish to compute the energy in these spectral variations on different frequency scales in terms of the telescope weighting function  $\bar{W}(\mathbf{x}, f)$ .

PSF  $\tilde{A}(\mathbf{l}, f)$  is the Fourier Transform of the  $uv$ -weighting function  $A(\mathbf{u}, f)$  and is given by

$$\tilde{A}(\mathbf{l}, f) = \frac{\int_{-\infty}^{\infty} du \int_{-\infty}^{\infty} dv A(\mathbf{u}, f) \exp(-i2\pi \mathbf{u} \cdot \mathbf{l})}{\int_{-\infty}^{\infty} du \int_{-\infty}^{\infty} dv A(\mathbf{u}, f)}, \quad (2)$$

where the denominator is included to obtain  $\tilde{A}(\mathbf{l}, f)$  as the normalized PSF for all frequencies. Any additional frequency weighting may be absorbed into the frequency window function  $B(f)$ . We have assumed that  $A(\mathbf{u}, f)$  includes the conjugate baselines, failing which  $\tilde{A}(\mathbf{l}, f)$  may be defined as the real part of the integral in Equation (2). The function  $A$  is just a scaled version of the function  $\bar{W}$  through the scaling relationship given by  $\mathbf{u} = \mathbf{x} \cdot (f/c)$ . We may represent this scaling relationship as

$$A(\mathbf{u}, f) = \bar{W}\left(\frac{\mathbf{u}c}{f}, f\right) = \bar{W}(\mathbf{x}, f). \quad (3)$$

The quantity  $uc/f$  has units of distance and is frequency invariant: it is simply the baseline distance in meters. We use this to make the substitution  $x = uc/f$ ,  $y = vc/f$  and get

$$\tilde{A}(\mathbf{l}, f) = \frac{1}{\mathcal{A}_W} \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dy \bar{W}(\mathbf{x}, f) \exp\left(-i2\pi \mathbf{x} \cdot \mathbf{l} \frac{f}{c}\right), \quad (4)$$

where  $\mathcal{A}_W = \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dy \bar{W}(\mathbf{x}, f)$  is the area under  $\bar{W}$ . The importance of this substitution is that it expresses all baselines in units of meter as opposed to wavelength and makes them independent of frequency.

Consider a point source with a flux density of unity and a spectral index of zero located at  $\mathbf{l}$ . The response at zenith due to a sidelobe on this source is given by

$$\tilde{S}(\mathbf{l}, f) = \tilde{A}(\mathbf{l}, f). \quad (5)$$

We are interested in the energy in  $\tilde{S}(\mathbf{l}, f)$  at different frequency scales. In other words, to evaluate PSF contamination, we have to compute the power spectrum  $\tilde{S}(\mathbf{l}, \eta)$  by Fourier transforming  $\tilde{S}(\mathbf{l}, f)$  with frequency offset  $\Delta f$  and delay  $\eta$  as Fourier conjugates. We use  $\Delta f$  instead of  $f$  because the spatial EoR fluctuations are evaluated about an origin  $r_0$  corresponding to a frequency  $f_0$  and redshift  $z_0$ .

$\tilde{S}(\eta, \mathbf{l})$  may be expressed as

$$\tilde{S}(\mathbf{l}, \eta) = \int_{-\infty}^{\infty} d(\Delta f) \tilde{A}(\mathbf{l}, f) \exp(+i2\pi \eta \Delta f), \quad (6)$$

where we have used a positive sign in the exponent as we are transforming a function of an image space variable ( $f$ ) to a function of a Fourier space variable ( $\eta$ ). Using  $\Delta f = f - f_0$ ,  $f_0 = (f_L + f_H)/2$ , where  $f_L$  and  $f_H$  are the lowest and highest frequencies in the instrument bandpass, and writing the conjugate variable as  $f$  instead of  $\Delta f$  we get

$$\tilde{S}(\mathbf{l}, \eta) = \exp(-i2\pi\eta f_0) \int_{-\infty}^{\infty} df \tilde{A}(\mathbf{l}, f) \exp(+i2\pi\eta f). \quad (7)$$

Substituting from Equation (4) for  $\tilde{A}(\mathbf{l}, f)$ , from Equation (1) for  $\tilde{W}(\mathbf{x}, f)$ , and interchanging the order of integrations we get

$$\begin{aligned} \tilde{S}(\mathbf{l}, \eta) &= \exp(-i2\pi\eta f_0) \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dy W(\mathbf{x}) \\ &\times \int_{-\infty}^{\infty} df \frac{T(\mathbf{x}, f)}{\mathcal{A}_W} B(f) \exp\left[-i2\pi f \left(\frac{\mathbf{x}\cdot\mathbf{l}}{c} - \eta\right)\right]. \end{aligned} \quad (8)$$

The inner integral, in general, may have to be evaluated numerically. However, to have an insight into the nature of  $\tilde{S}(\mathbf{l}, \eta)$ , we evaluate Equation (8) for the case where there is no taper or mask applied in the  $uv$  domain. In this case,  $T(\mathbf{x}, f) \equiv 1$ , and  $\mathcal{A}_W$  is frequency independent. The inner integral is now just a scaled and shifted version of the Fourier transform of the frequency window function. Consider a rectangular frequency window defined by

$$B(f) = \begin{cases} \frac{1}{B} & \text{if } f \in [f_L, f_H] \\ 0 & \text{otherwise,} \end{cases} \quad (9)$$

where  $B = f_H - f_L$  is the frequency bandwidth. The inner integral now evaluates to

$$\tilde{B}\left(\frac{\mathbf{x}\cdot\mathbf{l}}{c} - \eta\right) = \frac{\sin\left[\pi B \left(\frac{\mathbf{x}\cdot\mathbf{l}}{c} - \eta\right)\right]}{\pi B \left(\frac{\mathbf{x}\cdot\mathbf{l}}{c} - \eta\right)} \exp\left[-i2\pi f_0 \left(\frac{\mathbf{x}\cdot\mathbf{l}}{c} - \eta\right)\right]. \quad (10)$$

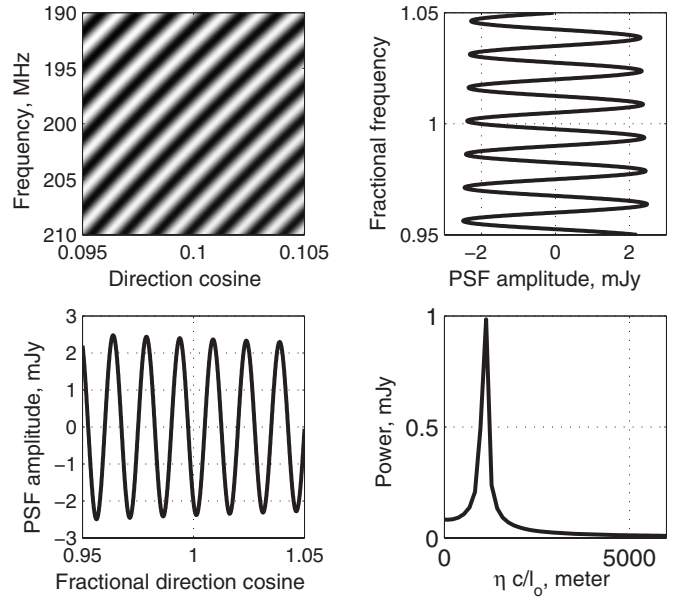
This gives us the relationship we are looking for

$$\tilde{S}(\mathbf{l}, \eta) = \frac{\exp(-i2\pi\eta f_0)}{\mathcal{A}_W} \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dy W(\mathbf{x}) \tilde{B}\left(\frac{\mathbf{x}\cdot\mathbf{l}}{c} - \eta\right). \quad (11)$$

The power spectrum  $\tilde{S}(\mathbf{l}, \eta)$  is a convolution of the weighting function  $W(\mathbf{x})$  and the convolving kernel  $\tilde{B}$ , which is a shifted version of the Fourier transform of the frequency window function  $B(f)$ . Readers conversant with the Fourier synthesis imaging literature may recognize Equation (9) as the ‘‘delay beam’’ or ‘‘fringe washing function’’ for a rectangular bandpass window function (see Thompson et al. 2001 and Bridle & Schwab 1989 for details) with  $\mathbf{x}\cdot\mathbf{l}/c$  as the geometric delay term and  $\eta$  as the instrumental delay term. The delay beam has traditionally been used to evaluate bandwidth smearing in synthesis telescopes. Equation (11) essentially shows that the PSF contamination in delay space is given by the convolution between the telescope sampling function (in meter units) and the delay beam.

### 3. IDEAL CASE OF UNIFORM VISIBILITY COVERAGE

This section provides an insight into the terms and equations of the analytical derivation, by understanding them in the case

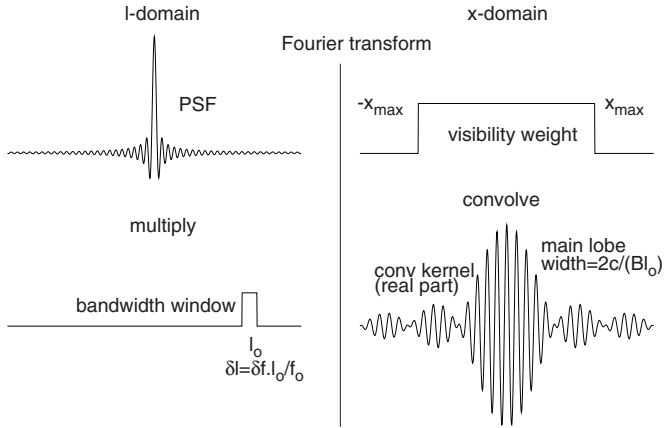


**Figure 1.** Plot showing an example of the PSF sidelobe structure in frequency for a uniformly weighted visibility distribution with uniform coverage ( $W(x) = 1$  for  $|x| < 1000$ ,  $W(x) = 0$  otherwise). Top left panel shows a small section of the PSF in the vicinity of  $(l_o, f_o) = (0.1, 200$  MHz). Top right panel shows the PSF variation at  $l_o$  as a function of fractional frequency,  $\delta f/f_o$ . Bottom left panel shows the PSF variation at  $f_o$  as a function of fractional directional cosine,  $\delta l/l_o$ . Bottom right panel shows the PSF at  $l_o$  Fourier transformed with respect to frequency to reveal the PSF contamination in  $\eta$  space due to a point source at  $l_o$ .

of a visibility distribution with uniform coverage. Uniform coverage refers to the case where there exists at least one visibility measurement in every pixel in the  $uv$  grid at all frequencies within the instrument bandwidth, and over the spatial frequency range spanned by the array configuration. It must be noted that uniform visibility coverage does not imply uniform weighting of measured visibilities. A key idea in the analytical treatment of Section 2 is the invariance of the PSF in  $lf$ . This concept is manifested in the invariance of Equation (4) in  $lf$  for a frequency-invariant weighting. Frequency-invariant weighting implies a taper function  $T(x)$  that is dependent only on  $x$ . Figure 1 shows a section of the PSF in  $lf$  domain for a uniformly weighted visibility distribution with uniform coverage. The plot extends over a 10% spread in both frequency and direction cosine ( $\Delta l/l_o = \Delta f/f_o = 0.1$ ). Due to the  $lf$  invariance, we expect the PSF to have the same value at all points defined by  $lf = l_o f_o$ . Consequently, the variation in an interval  $\delta l$  of the PSF evaluated at  $f_o$  is the same as the variation in an interval  $\delta f$  of the PSF evaluated at  $l_o$ , so long as  $\delta l/l = -\delta f/f$  or approximately  $\delta l/l_o = -\delta f/f_o$ . This can be obtained by differentiating  $lf = l_o f_o$  and then approximating  $f$  to  $f_o$  and  $l$  to  $l_o$ , which hold for  $\delta l/l \ll 1$  and  $\delta f/f \ll 1$ . Figure 1 also evaluates PSF contamination in the zenith pixel in  $\eta$  space due to PSF sidelobes on a 1 Jy point source at  $l_o$ . The localization of this contamination in light of Equations (10) and (11) is pivotal to our understanding of PSF contamination and deserves a detailed explanation.

#### 3.1. Localization of PSF Contamination

Consider a single point source at direction cosine  $l_o$  entering the zenith pixel through the PSF sidelobe at  $l_o$ . This is depicted in Figure 2. Due to the  $lf$  invariance of the PSF, the frequency structure of the PSF in a bandwidth  $\delta f$  is identical to the sidelobe structure of the PSF (evaluated at  $f_o$ ) in the interval



**Figure 2.** Plots depicting the operation in Equations (10) and (11). See Section 3.1 for a detailed explanation of the operations involved.

$[l_o - \delta f l_o / (2 f_o), l_o + \delta f l_o / (2 f_o)]$ . This truncated version of the PSF can be extracted by multiplying the PSF with a window centered at  $l_o$  and having width  $\delta f l_o / f_o$ . A Fourier transform of the product of these two functions is the convolution of their individual Fourier transforms which are the baseline weighting function  $W(x)$  and the delay beam (or convolution kernel) from Equation (11). The sinc-function envelope of the delay beam has a main lobe width equal to  $2c/(l_o \Delta f)$ , which in general is far narrower than the visibility range  $2x_{\max}$ . The rapid variation in the delay beam is a complex sinusoid in  $x$  with frequency  $l_o f_o / c$ . At any  $\eta$ , the convolution in Equation (11) is sensitive to a small range in  $W(x)$  that is centered at  $x = c\eta / l_o$  with a width  $c/(l_o \Delta f)$ . This explains the sharply peaked PSF contamination in  $\eta$  centered at  $\eta = x_{\max} l_o / c$  (see Figure 1), where  $x_{\max} l_o / c$  is indeed the geometric delay for a source at  $l_o$ . Sources from zenith to horizon ( $l_o = 0$  to  $|l_o| = 1$ ) suffer geometric delays from  $\eta = 0$  to  $\eta = x_{\max} / c$  respectively, and hence contribute to foreground contamination at different fairly localized values of  $\eta$ . This is shown in Figure 3 for sources at five different locations.

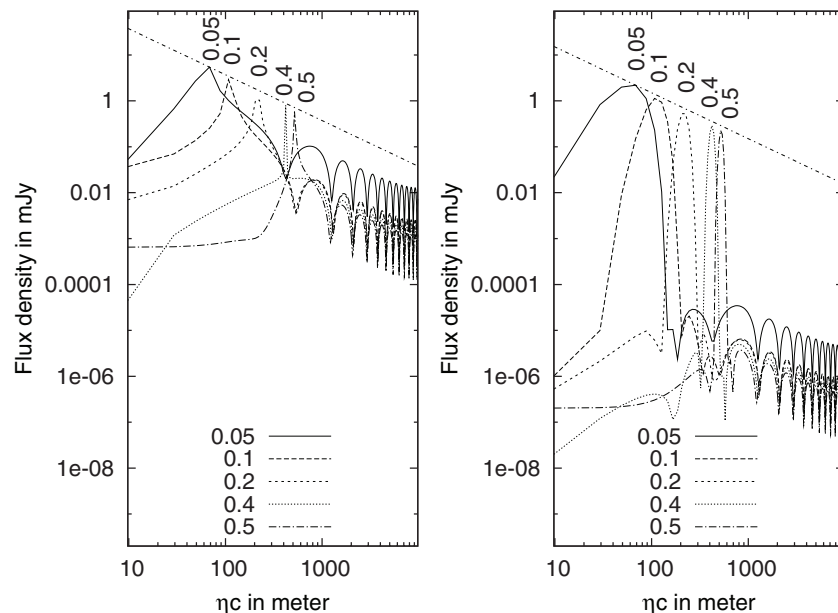
The relative magnitude of these impulses is determined by the relative sidelobe levels at the respective directions. For the case depicted in Figure 3, the rectangular shape of  $W(x)$  gives a  $1/l$  roll-off in the PSF sidelobe levels which corresponds to a  $1/\eta$  roll-off in the contamination at different values of  $\eta$ .

### 3.2. Sidelobes of the Convolution Kernel

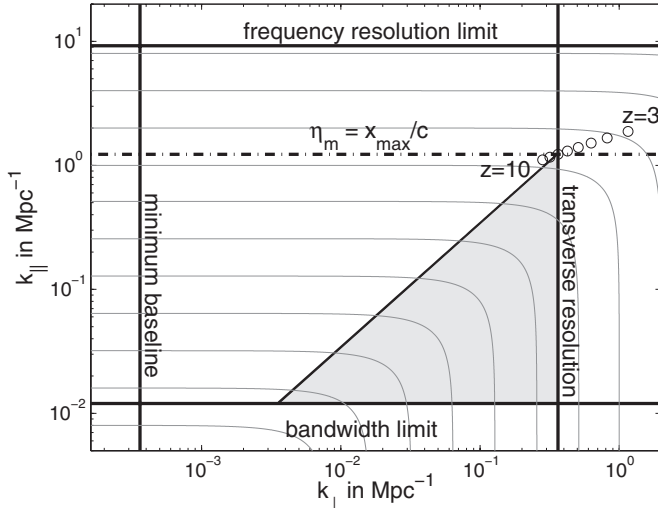
We have used the phrase sharply peaked to describe the PSF contamination for a source entering the zenith pixel from the direction  $l_o$ . The exact profile is described by the sinc envelope of the delay beam (see Figure 3). Since the maximum possible geometric delay (source at horizon) is  $x_{\max}/c$ , the delay space may be divided into two regions. Region  $\mathcal{R}_1$ :  $\eta \in [0, x_{\max}/c]$  defines a region of relatively high foreground contamination. Region  $\mathcal{R}_2$ :  $\eta > x_{\max}/c$ , however, has significantly lower but finite PSF contamination due to the sidelobes of the delay beam. The contamination in  $\mathcal{R}_2$  falls off as  $1/\eta$  since the sidelobes of a sinc-function envelope of the delay beam have a  $1/x$ -form roll-off. The level of contamination for  $\eta > x_{\max}/c$  may be reduced further by a judicious choice of a bandpass window  $B(f)$  that gives a delay beam with low sidelobe levels.  $\mathcal{R}_2$  defines an EoR window where we may focus our efforts to detect the EoR. Figure 3 demonstrates a reduction in delay beam sidelobe contamination in  $\mathcal{R}_2$  by over three orders of magnitude with a Blackman–Nuttall frequency window function (Nuttall 1981). This reduction in contamination comes at the cost of reduced resolution in  $\eta$  space. The performance of such a frequency window function will rapidly deteriorate as interference ridden frequencies are flagged. The improvement in  $\mathcal{R}_2$  contamination falls below an order of magnitude even if 1% of randomly selected channels are flagged. In any case,  $\mathcal{R}_2$  defines a region of significantly lower PSF contamination—an EoR window where we may focus our effort to detect the EoR.

### 3.3. Instrumental $k$ Space

The instrumental  $k$  space is an effective space to represent regions and extent of foreground contamination. The instrumental



**Figure 3.** PSF contamination in  $\eta$  space due to 1 Jy point sources at five different locations ( $l_o = 0.05, 0.1, 0.2, 0.4,$  and  $0.5$ ). The left panel shows the contamination for  $W(x)$  with uniform weight and a rectangular frequency bandpass, and the right panel shows the contamination when a Blackman–Nuttall window function is used for the frequency bandpass. The oblique broken line is an approximate  $1/\eta$  fit to the peaks of the impulses.



**Figure 4.** Instrumental response in  $k$  space at redshift of  $z = 8$ . The MWA instrument parameters have been used for illustration. A bandwidth limit of 30 MHz defines a minimum  $k_{\parallel\min}$  of about  $0.012 \text{ Mpc}^{-1}$ , a frequency resolution of 40 kHz defines a  $k_{\parallel\max}$  of about  $9.1 \text{ Mpc}^{-1}$ , an array extent of 1 km defines a  $k_{\perp\max}$  of about  $0.3634 \text{ Mpc}^{-1}$ , and a minimum baseline of 7.5 m gives a  $k_{\perp\min}$  of about  $3.5 \times 10^{-4}$ . The region of high foreground contamination  $\mathcal{R}_1$  extends up to  $\eta_m = x_{\max}/c = 3.33 \mu\text{s}$  that corresponds to  $k_{\parallel m}$  of about  $1.2 \text{ Mpc}^{-1}$ . The shaded triangle shows the region of PSF contamination outside of which lies the EoR window. The unfilled black circles define the point  $(k_{\perp\max}, k_{\parallel m})$  for redshifts of 3, 4, 5, 6, 7, 8, 9, and 10. Each concentric curve is the locus of constant  $k = \sqrt{k_{\perp}^2 + k_{\parallel}^2}$ .

$k$  space is two dimensional with the comoving LOS wavenumber  $k_{\parallel}$  and comoving transverse wavenumber  $k_{\perp}$  as axes. If  $k_x$ ,  $k_y$ , and  $k_z$  are comoving wavenumbers along three spatial axes with origin at some point about which the EoR fluctuations will be estimated, then Morales & Hewitt (2004) have shown that

$$\begin{aligned} k_{\perp} &= \sqrt{k_x^2 + k_y^2}, \quad k_x = \frac{2\pi u}{D_M(z)}, \quad k_y = \frac{2\pi v}{D_M(z)}, \\ k_z &= k_{\parallel} \approx \eta \frac{2\pi H_0 f_{\text{HI}} E(z)}{c(1+z)^2}, \end{aligned} \quad (12)$$

where the approximation holds for a small bandwidth (small redshift range). Here,  $f_{\text{HI}} = 1420 \text{ MHz}$  is the rest frequency of 21 cm HI line emission,  $H_0 = 70 \text{ km sec}^{-1} \text{ Mpc}^{-1}$  is the present value of the Hubble constant,  $E(z) = (\Omega_M(1+z_o)^3 + \Omega_k(1+z_o)^2 + \Omega_{\Lambda})^{\frac{1}{2}}$ , and  $D_M(z)$  is the transverse comoving distance at redshift  $z$ . Figure 4 shows a plot of the measurement space (in  $k_{\perp} - k_{\parallel}$  domain) for a Fourier synthesis array with uniform visibility coverage. A Fourier synthesis array is sensitive to measurements in a region in the  $k$  space bounded by some maximum and minimum values of  $k_{\perp}$  and  $k_{\parallel}$ . The highest  $k_{\perp}$  mode sampled by the instrument is related to the  $uv$  extent of the array and is given by

$$k_{\perp\max} = \frac{2\pi x_{\max} f_{\text{HI}}}{c(1+z)D_M}, \quad (13)$$

while the smallest  $k_{\perp}$  probed by the instrument is related to the physical extent of the primary antenna element  $d$  and is given by

$$k_{\perp\min} = \frac{2\pi d f_{\text{HI}}}{c(1+z)D_M}. \quad (14)$$

The highest value of  $k_{\parallel}$  sampled by the instrument is related to the frequency resolution and is given by

$$k_{\parallel\max} \approx \eta_{\max} \frac{2\pi H_0 f_{\text{HI}} E(z)}{c(1+z)^2}, \quad \eta_{\max} = \frac{1}{\Delta f}. \quad (15)$$

The lowest value of  $k_{\parallel}$  sampled by the instrument in a snapshot observation is related to the frequency bandwidth and is given by

$$k_{\parallel\min} \approx \eta_{\min} \frac{2\pi H_0 f_{\text{HI}} E(z)}{c(1+z)^2}, \quad \eta_{\min} = \frac{1}{B}. \quad (16)$$

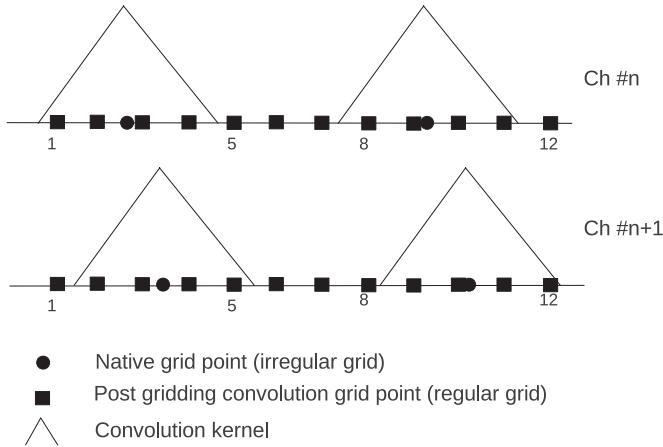
The final comment in this section is related to the implications of  $lf$  invariance of the PSF to our understanding of the location and level of contamination in the instrumental  $k$  space. Consider the foreground contamination in the neighborhood ( $\delta l \ll 1$ ) of some image pixel  $\mathcal{P}$  due to a point source at some  $l_o$  removed from  $\mathcal{P}$ . Due to the  $lf$  invariance, the transverse structure of contamination around  $\mathcal{P}$  is identical to the LOS contamination at  $\mathcal{P}$ . Since  $\delta l/l_o = \delta f/f_o$ , the energy in transverse fluctuations around  $\mathcal{P}$  at any wavenumber  $k_{\perp}$  will be the same as the energy in LOS fluctuations at  $\mathcal{P}$  at wavenumber  $\eta = (D_M l_o)/(2\pi f_o) k_{\perp}$ . Using Equation (12), the  $(k_{\perp}, k_{\parallel})$  pairs with identical contamination in them are defined by  $k_{\parallel} = k_{\perp} \frac{H_0 E(z)}{c(1+z)} l_o$ .

For a uniformly weighted visibility distribution with uniform coverage, there is significant energy only in the transverse mode  $k_{\perp o} = 2\pi x_{\max} f_o/(c D_M)$ . This gives us a corresponding  $\eta_o = l_o/f_o k_{\perp o} = x_{\max} l_o/c$ : a result we established in Section 3.1. If  $\mathcal{P}$  is the zenith pixel,  $\eta_o$  reaches a maximum value of  $\eta_m = x_{\max}/c$  for a source at the horizon ( $|l_o| = 1$ ). The corresponding values of  $k_{\parallel o}$  and  $k_{\parallel m}$  may then be obtained from Equation (12). The PSF contamination in this case will lie on the line joining  $(k_{\perp\max}, k_{\parallel m})$  and  $(k_{\perp\max}, k_{\parallel\min})$  in the instrumental  $k$  space. If the visibility plane is not uniformly filled there will be contamination in various values of  $k_{\perp} < k_{\perp o}$ . The PSF contamination however will always lie within the shaded triangle in Figure 4. The instrumental  $k$  space sampled by the telescope outside of this shaded triangle then defines the EoR window in  $k_{\perp} - k_{\parallel}$  space where we may focus our efforts to detect the EoR.

#### 4. EFFECT OF GRIDDING AND IMAGING ALGORITHMS

While Section 3 described the nature of foreground contamination along the LOS and transverse wavenumber axes, its results are a good description of the contamination in the special case of an ideal array, and hence lay the foundations needed to appreciate issues related to EoR detection with practical array geometries and data processing algorithms. Real arrays seldom have uniform visibility coverage, and the Fourier relationships between visibility and image space, and between frequency and LOS wavenumber space are, in practice, evaluated as discrete transforms. Computationally efficient transforms like the Fast Fourier Transform (FFT) are often used in Fourier synthesis arrays and almost always involve re-gridding the visibility data on a regular  $uv$  grid. This section discusses the effects of such gridding and Fourier transform algorithms used for imaging on the nature of the three-dimensional PSF and the resulting foreground contamination in  $k$  space.

We have shown in Section 3 that contamination in  $\eta$  space is substantially different in the two regions (see Figures 3 and 4)  $\mathcal{R}_1$  and  $\mathcal{R}_2$ , where  $\mathcal{R}_1$  extends from  $\eta = 0$  to  $\eta = x_{\max}/c$ , and  $\mathcal{R}_2$  (or EoR window) extends from  $\eta = x_{\max}/c$  to  $\eta = 1/\Delta f$ . We



**Figure 5.** Depiction of inter-channel baseline migration in one dimension ( $u$ ). The filled squares are grid points and the filled circles are baseline vector locations in wavelength units. The continuous line represents the post-gridding convolution visibility distribution for a triangular convolution kernel. For simplicity, we have assumed that both depicted visibility measurements have equal weights. The upper panel shows the visibility axis for some frequency channel  $n$ , and the lower panel shows the visibility axis for the adjacent channel  $n + 1$ . The depicted visibility segment is shown for a region of fairly sparse  $u$  coverage for clarity.

have also shown that contamination in  $\mathcal{R}_2$  is due to the sidelobes of the convolution kernel  $\hat{B}$  (see Equation (11)) and the sharply peaked impulses from sources at various directions are confined to  $\mathcal{R}_1$ . Consequently, if region  $\mathcal{R}_2$  exists ( $1/\Delta f > x_{\max}/c$ ), then the PSF variation in frequency at any direction cosine  $\mathbf{l}$  is oversampled by the instrument spectral resolution  $\Delta f$ . This implies that, in the ideal case, there will not be significant channel to channel variation in flux versus frequency at any sky pixel. However, processes like gridding may induce channel to channel jitter in the PSF and potentially contaminate  $\mathcal{R}_2$ . This stochastic component of foreground contamination arising from the gridding process may be termed gridding contamination. Such stochastic effects may be understood in terms of channel-to-channel migration of baseline vectors from one  $uv$  pixel to another.

#### 4.1. Baseline Migration

Fourier synthesis arrays have traditionally formed image cubes in  $lmf$  space by (1) computing the baselines (in wavelength units:  $\mathbf{u} = \mathbf{x}.f/c$ ) generated by a given antenna distribution at every frequency channel, (2) weighting the visibilities to get desirable PSF characteristics and to modify the spatial filtering, (3) performing gridding convolution on the visibilities to re-grid them on a regular  $uv$  grid at every frequency channel, and (4) using an efficient FFT algorithm to compute a dirty sky map at every frequency channel. Among these signal processing steps, gridding convolution has not been considered in the analytical treatment of Section 2. We now describe the influence of gridding convolution on foreground contamination in  $\eta$  space.

A baseline vector  $\mathbf{x}$  (in meter units) will suffer a displacement of  $\delta\mathbf{u} = \mathbf{x}\delta f/c$  (in wavelength units) between adjacent frequency channels in the  $uv$  plane. Here,  $\delta f$  is the inter-channel frequency spacing. We call this phenomenon “baseline migration” and depict it in Figure 5. The figures show a one-dimensional visibility axis at two adjacent channels. While grid pixel number 1 and 8 go from being filled to being unfilled between channel  $n$  and  $n + 1$ , visibility pixels 5 and 12 go from

being unfilled to being filled. Pixels such as 2 and 10 continue to be filled but experience a step change in their weighted visibility value. Such inter-channel variations result in step changes in gridded visibility values of pixels within the support of the convolving kernel. Many such step changes in visibilities due to the displacement of many baseline vectors results in a stochastic jitter in the frequency structure of the PSF. PSF jitter due to baseline migration may induce channel to channel fluctuations in flux versus frequency on the image plane, thereby significantly contaminating the EoR window  $\mathcal{R}_2$ . This PSF jitter due to baseline migration is expected to have high contribution (1) from longer baseline vectors due to the relatively higher quantum of baseline displacement  $\delta\mathbf{u}$  and (2) from regions of low  $uv$  density. In regions of higher  $uv$  density a given  $uv$  pixel, during gridding convolution, receives contributions from several neighboring baselines. The visibility contribution to a given  $uv$  pixel then depends on the relative displacement of the neighboring baselines, rather than their actual displacement. On the other hand, consider the extreme case of very low  $uv$  density as depicted in Figure 5. Here the actual inter-channel displacement of the baseline vectors induces relatively high steps in the post-gridding convolution visibilities. Natural weighting results in significantly lower weight assigned to regions with (1) longer baseline vectors and (2) low  $uv$  density and is expected to result in low PSF jitter in frequency. Uniform weighting, on the other hand, increases the weighting for longer baselines and is expected to result in substantially high PSF jitter.

#### 4.2. The Central $uv$ Void

The angular power spectrum of Ectic sources has been observed to have a power-law-like form (Blake & Wall 2002) with substantially greater power on small angular scales; the universe is homogeneous and isotropic on large scales. Consequently, short baselines are not expected to have substantially greater response to extragalactic continuum sources. However, the mean sky brightness will have a response at zero spacing ( $u = 0, v = 0$ ). Energy in various spatial Fourier modes is seldom perfectly isolated and the zero spacing component will leak into short baselines. Consequently, channel to channel changes in the visibility distribution close to  $u = 0, v = 0$  may result in step changes in flux versus frequency at image pixels thereby contaminating higher values of  $\eta$  in  $\mathcal{R}_2$ . While extragalactic sources are expected to have a smooth distribution on large angular scales, flux from the Galactic synchrotron emission is expected to follow a power law with significant spatial distribution power at large angular scales. Consequently, short baselines close to  $u = 0, v = 0$  will have significant response to our Galaxy and baseline migrations in this part of the  $uv$  plane will create significant PSF jitter that may further contaminate  $\mathcal{R}_2$ .

Fourier synthesis arrays usually have a central  $uv$  void as the antenna elements have a minimum spacing due to constraints imposed by their own physical size. Antenna cross-talk considerations often result in array configurations with a central  $uv$  void that is several times the antenna size. In such cases, the central  $uv$  void spans several  $uv$  pixels. We mentioned in Section 4.1 that baseline migrations in regions of high  $uv$  density do not contribute significantly to inter-channel PSF jitter. Though short spacing baselines lie in a densely populated part of the  $uv$  plane, with increasing frequency, most baseline vectors are displaced radially away from any  $uv$  grid point close to  $u = 0, v = 0$ , as opposed to other  $uv$  grid points where baseline vectors also get displaced toward them. This

asymmetric neighborhood near short spacing visibilities results in relatively higher inter-channel step changes in the visibilities on short baseline vectors. To alleviate this problem, a frequency-independent  $uv$  mask that is larger than the central  $uv$  void may be employed to reduce such step due to asymmetric migrations in short baselines.

### 4.3. The Chirp Z Transform

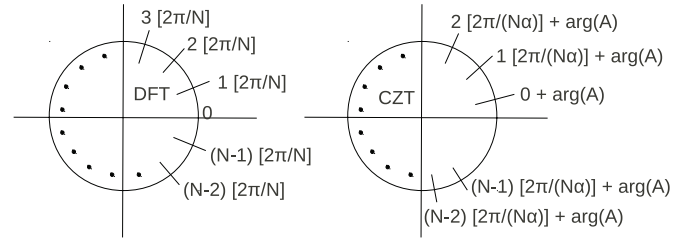
We have seen how channel to channel changes in the visibility weighting  $W(\mathbf{u}, f)$  contaminates the LOS wavenumber dimension. Here,  $W(\mathbf{u}, f)$  includes the effect of all visibility weighting and tapers used, and  $W(\mathbf{u}, f)$  determines the PSF via the imaging transform. If the visibility distribution has uniform coverage, maintaining invariance of  $W(\mathbf{u}, f)$  with frequency is possible through application of a frequency-independent visibility taper function. In such a case,  $W(\mathbf{u}, f)$  is same at all frequencies and there is no frequency dependence in the PSF. In practice, there are many unfilled  $uv$  pixels and Jelić et al. (2008) have suggested that a visibility mask may be used that rejects the  $uv$  pixels that are not filled at all (or most) frequencies prior to image formation. At one end, we may exploit the complete instrument data and sensitivity at the cost of increased contamination in  $\mathcal{R}_2$ , and at the other end we may have a frequency-independent PSF with some loss in information and sensitivity. However, there exists a middle path that assures that no data are rejected while restricting the foreground contamination in  $\mathcal{R}_2$ —image formation using the CZT (Rabiner et al. 1969). We first present a brief description of CZT.

The Z transform of any finite discrete sequence  $x_n$ ,  $n = 0, 1, 2, \dots, N-1$  is defined as

$$X(z) = \sum_{n=0}^{n=N-1} x_n z^{-n}. \quad (17)$$

The Z transform is a function of the complex variable  $z$ . When evaluated at discrete points along the unit circle defined by  $z_k = \exp(i2\pi k/N)$ , where  $k = 0, 1, 2, \dots, N-1$ , Z transform of the finite discrete sequence  $x_n$  reduces to its Discrete Fourier Transform (DFT). The CZT is another special case of the Z transform that is evaluated at discrete points on the locus given by  $z_k = AW^{-k}$  where  $A$  and  $W$  are complex constants. Such a locus describes a spiral in the complex  $z$  plane. Note how for  $A = 1$  and  $W = \exp(-i2\pi/N)$  the CZT again reduces to a DFT.

Consider a case where  $|A| = 1$  and  $W = \exp(-i2\pi/(N\alpha))$ , where  $\alpha$  is a real constant. The locus of points on the complex  $z$  plane where the transform is evaluated now lies on a unit circle like in the case of a DFT. However, angular spacing between the points on the locus is now  $2\pi/(n\alpha)$  as opposed to  $2\pi/N$  in the case of a DFT. The first point on the locus is also shifted by the argument of the complex constant  $A$ . The loci for the two cases are shown in Figure 6. Though  $|A| = 1$ ,  $W = \exp(-i2\pi/(N\alpha))$  is a special case of the generalized CZT, we refer to it hereinafter as CZT with a chirp factor  $\alpha$ . The only difference between DFT and CZT is the presence of the chirp factor  $\alpha$  and an offset due to the presence of  $A$ .  $\alpha$  simply stretches the locus on which the transform is evaluated and the CZT of a sequence is essentially a DFT of a stretched version of the sequence. We now describe how visibilities can be gridded in meter units and how a chirp factor can be used to compensate for the stretching of baseline vectors with frequency.



**Figure 6.** Sketches showing points on the  $z$  plane on which the generalized Chirp Z transform is computed for two cases: Case 1 (left panel) is for  $A = 1$  and  $W = \exp(-i2\pi/N)$  and Case 2 (right panel) is for  $|A| = 1$  and  $W = \exp(-i2\pi/(N\alpha))$ . In Case 1, the generalized CZT reduces to DFT, and for Case 2 the generalized CZT reduces to a special case of CZT that will be used in the proposed imaging algorithm of Section 4.4.

### 4.4. A CZT-based Imaging Algorithm

As described in Section 4.1, the principal cause of channel to channel PSF variations in conventional gridding and imaging processes is baseline migration due to independent visibility gridding at every channel. However, the baseline vector generated by a given antenna pair moves across the visibility grid with change in frequency in a very predictable manner. The baselines in wavelength units at any frequency  $f$  are merely the baselines in meter units stretched by a factor  $\alpha$ , where  $\alpha = f/c$ . In principle, we may grid the visibilities only once in meter units and use the CZT with chirp factor  $\alpha$  to form the image at every frequency channel. Given an image extent, the grid size in meter units is chosen to avoid aliasing even at the highest frequency in the visibility cube. Gridding the visibilities only once in meter units eliminates the problem of baseline migrations and consequently ameliorates the problem of contamination in  $\mathcal{R}_2$ . In summary, a given antenna pair will fall on exactly the same visibility grid point at all frequencies. We then compute the image on an  $lm$  grid that is constant over frequency and use the chirp factor  $\alpha$  in a CZT to compensate for baseline stretching. Such a CZT-based imaging transform is given by

$$S(\mathbf{l}_{mn}) = \sum_{i=0}^{i=N-1} \sum_{j=0}^{j=N-1} V_1(\mathbf{x}_{ij}, f) \exp\left(-\frac{i2\pi}{N\alpha(f)} \mathbf{x}_{ij} \mathbf{l}_{mn}\right), \quad (18)$$

where  $i, j$  are  $uv$  (meter units) grid pixel indices,  $m, n$  are image grid pixel indices, function  $V_1(\mathbf{x}_{ij}, f)$  represents the visibilities at frequency  $f$  gridded on to the common grid, and  $\alpha(f)$  is the chirp factor given by  $\alpha(f) = c/f$ . The transform in Equation (18) is similar to a DFT and may be evaluated by the same efficient FFT algorithms with a minor modification to the exponential terms. The CZT-based image formation results in reduced contamination in  $\mathcal{R}_2$ . In addition, it reduces the computational cost of gridding convolution because the gridding convolution at all frequency channels is done to a single common grid and the same baseline-timeslot to pixel-weight map is used for all frequencies. On the other hand, due to the introduction of  $\alpha(f)$  into the exponential (see Equation (18)), the multiplicative exponential factors in the Radix-2 butterflies need to be recomputed at every frequency channel, thereby increasing the computational cost. Hence, there are competing effects which decide the efficacy of image formation using the CZT.

## 5. THE MWA EXAMPLE

We have so far discussed the nature of foreground contamination in  $\eta$  space for an ideal array with uniform visibility

coverage. We have also outlined some of the practical considerations applicable to realistic arrays and in the process proposed a new imaging technique using the CZT. In this section we demonstrate the relationships and results of Sections 2, 3, and 4 using the case of snapshot imaging with the MWA. The MWA consists of 512 antenna elements called tiles. Each tile consists of a phased array of 16 bow-tie antennas. The tiles are distributed to form baselines up to about 1 km. The instantaneous bandwidth is about 30 MHz, and we assume a frequency resolution of about 40 kHz. MWA will produce snapshot images every 8 s. The instrumental response for this instrument configuration in the  $k_{\perp} - k_{\parallel}$  plane is shown in Figure 4. Details of the MWA telescope design may be found in Lonsdale et al. (2009).

### 5.1. Confusing Sources and Sidelobes

The threshold flux density  $S_C$  above which we expect to find one source per beam of full width at half-maximum  $\theta_{\text{arcmin}}$  at frequency  $f_{\text{GHz}}$  is given by Subrahmanyan & Ekers (2002)

$$S_C = 40 \theta_{\text{arcmin}}^{1.67} f_{\text{GHz}}^{-0.67} \mu\text{Jy}. \quad (19)$$

The stochastic response to sources weaker than  $S_C$  contributes to sky pixel rms that is close to  $S_C$ .  $S_C$  for MWA evaluates to 2.3 mJy and is about three orders of magnitude higher than the expected thermal noise in 8 s snapshot images: MWA snapshot images will be confusion limited. We expect to find one source above  $5S_C$  about in 10 beamwidths. It is common practice to use a cutoff flux corresponding to one source in 25 beamwidths (Hazard & Walsh 1959) if we need to distinguish point sources from a combination of weaker unresolved faint sources for purposes of making a source catalog. For purposes of peeling, this requirement may be relaxed, and we assume successful peeling of all sources above  $5S_C = 10$  mJy. The expected residual rms confusion noise after successful subtraction of sources above 10 mJy is about 3.6 mJy. Due to the dominance of confusion over thermal noise, we will neglect the effects of thermal noise hereinafter. However, it must be noted that post foreground modeling and successful subtraction, the residual image that will be used for EoR power spectrum estimation is expected to be dominated by thermal noise.

The threshold to which continuum sources may be identified and subtracted also depends on our ability to distinguish spillover flux due to sidelobes on the sky. We now roughly estimate the expected rms fluctuation at any image pixel due to the sidelobes on all the confusing sources in the sky—a quantity we call “sidelobe confusion.” If we assume (1) the sidelobes of the PSF to be a zero mean random variable, and (2) the confusion flux in each pixel to be a random variable that is uncorrelated with the sidelobe level, then the approximate rms sidelobe confusion can be estimated fairly easily:  $C_S = \sum_i P_i S_i$  where  $P_i$  is the sidelobe level at pixel  $i$  and  $S_i$  is the primary beam weighted true sky flux due to confusing sources at pixel  $i$ . Since  $\langle P_i \rangle = 0$ , the variance in sidelobe confusion is  $\langle C_S^2 \rangle = \sum_i \langle P_i^2 S_i^2 \rangle = \sum_i \langle P_i^2 \rangle \langle S_i^2 \rangle$ . The far sidelobes of MWA are expected to have an rms variation of about 0.05% giving  $\langle P_i^2 \rangle = 0.25 \times 10^{-6}$ , and the rms confusion is about 3.6 mJy giving  $\langle S_i^2 \rangle = 13$  mJy<sup>2</sup>. The sidelobe level assumed here is true for sidelobes far from the phase center at which sidelobe confusion is being computed. In reality sidelobe levels vary from sky pixel to pixel and a constant value has been assumed here for simplicity. Far sidelobe level has been chosen since far sidelobes result in foreground contamination at higher

values of  $\eta$  where we expect to detect the EoR with minimal contamination, and this calculation will present the worst case. While  $P_i$  is assumed to be independent of  $i$ ,  $S_i$  is assumed to be the confusion noise weighted by the primary beam gain, and the summation is performed using the expected primary beam of the MWA primary antenna element (full width to first null of  $60^\circ$ ) to evaluate sidelobe confusion. The resulting sidelobe confusion from the entire sky is  $\langle C_S \rangle \approx 1.9$  mJy. The similarity in the values of blending confusion and sidelobe confusion is an interesting consequence of the similarity between the number of resolution elements in the image (weighted by the primary beam) and inverse of the sidelobe level of the synthesized beam. This implies that after subtracting bright point sources from MWA snapshot images, the resulting residual image has statistical fluctuations with roughly equal contributions from the stochastic response to the extragalactic sources and the sidelobe levels.

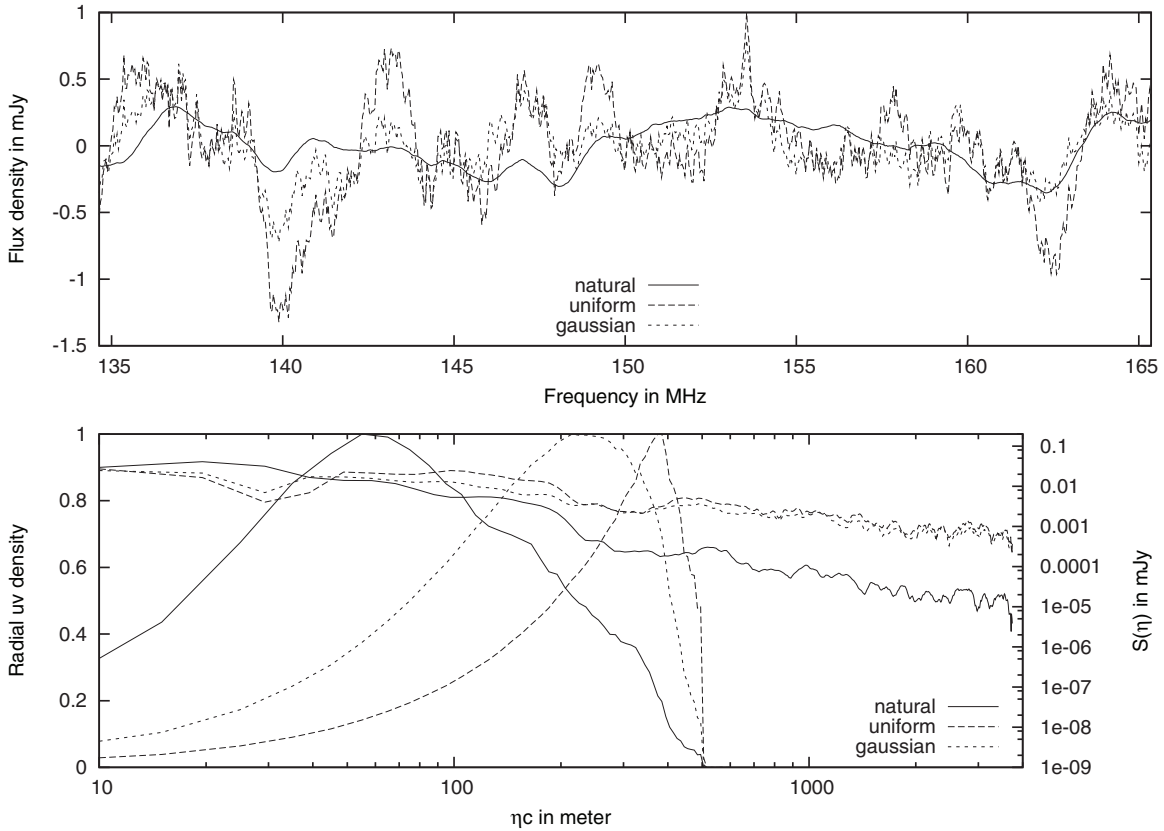
The EoR brightness fluctuations are expected to have an rms of about 25 mK on scales comparable to the MWA 5 arcmin resolution. This corresponds to an rms flux of about  $36 \mu\text{Jy}$ . Hence, apart from confusing foreground sources, sidelobe confusion can confuse any attempt to detect the EoR signal through a simple measurement of image plane variance. That said, flux from confusing sources and PSF sidelobes have structure in the frequency domain which can be exploited to detect the EoR fluctuations despite a high rms fluctuation in the image plane. While spectral structure of confusion extragalactic sources follows relatively simple power laws, spectral structure of instrumental sidelobe confusion is more complex as described in Sections 2, 3, and 4. We now quantitatively estimate the foreground contamination in frequency space (and eventually in  $\eta$  space) for the case of snapshot imaging with the MWA.

### 5.2. A Snapshot Observation

This subsection describes a simulation to estimate the foreground contamination in  $\eta$  space for the case of snapshot imaging of extragalactic point sources using the MWA. The simulation assumes a frequency-invariant sky model and assumes that all sources brighter than 10 mJy have been successfully subtracted. The sky model is then interpolated into an  $lm$  grid with grid size at least twice as fine as the MWA instrument resolution. The grid extent covers the interval  $[-0.5, 0.5]$  in direction cosines that corresponds to the MWA primary beam full width at first null, which is  $60^\circ$  at 150 MHz. Gridding convolution of visibilities is performed using a triangular kernel like the one shown in Figure 5. A simple kernel was chosen purely to illustrate the nature of contamination in  $\eta$  space. At every frequency channel, three separate PSFs are generated using natural weighting, uniform weighting, and uniform weighting with a Gaussian taper (hereinafter referred to simply as Gaussian weighting). The simulated sky in  $lm$  is then convolved with the PSF at zenith and the result is summed. This accumulated value at every frequency channel gives the dirty image flux density at the zenith pixel. This zenith pixel flux versus frequency is used to evaluate the spillover of foreground contamination by Fourier transforming to  $\eta$  domain. The simulation was performed for three cases each involving a different gridding and imaging algorithm; we discuss these below.

Case 1 employed independent gridding convolution at every frequency channel, and the dirty image was formed using an efficient FFT algorithm. Figure 7 shows the flux density versus frequency and contamination power in  $\eta$  for this case. We have shown earlier that for  $\eta < x_{\text{max}}/c$  in  $\mathcal{R}_1$ , the exact structure





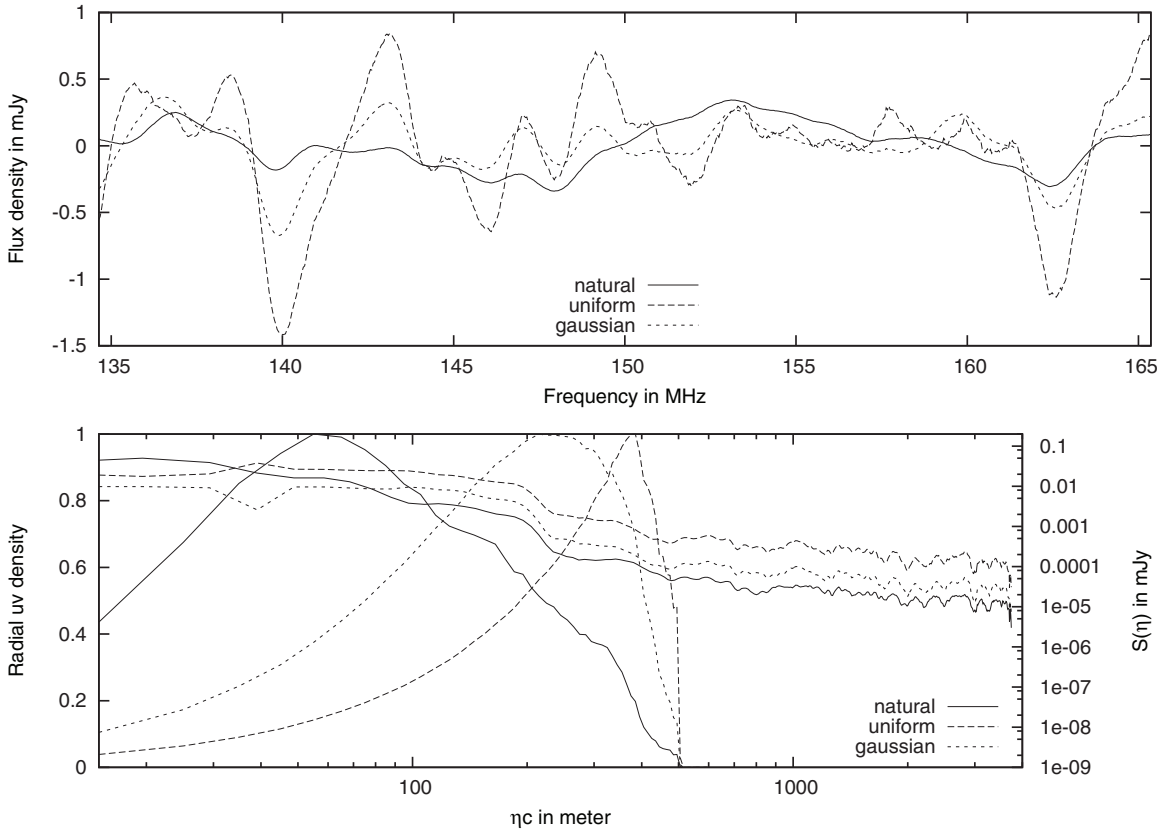
**Figure 7.** Foreground contamination along the line-of-sight dimension in the zenith pixel for the case of snapshot imaging with MWA. In generating these representations of foreground contamination, we populate sources in the sky out to  $|\mathbf{l}|_{\max} = 0.5$ . The zenith pixel flux is plotted against frequency (upper panel) and  $\eta$  (lower panel). Note the jitter in the upper panel for the case of uniform weighting and Gaussian weighting as opposed to the case of natural weighting. This jitter results in significant gridding contamination at higher values of  $\eta$  as seen in the lower panel. Also shown in the lower panel is radial  $uv$  density as a function of  $|\mathbf{x}|/|\mathbf{l}|_{\max}$  for the case of natural, uniform, and Gaussian weighting. The boundary between  $\mathcal{R}_1$  and  $\mathcal{R}_2$  lies at  $\eta c = |\mathbf{x}|_{\max}/|\mathbf{l}|_{\max} = 500$  m.

of spillover depends on the nature of relative weightings given to visibilities. Weighting here includes the effect of  $uv$  plane distribution of baselines, user defined weighting, and any tapers applied. For the case of uniformly weighted visibilities with complete  $uv$  coverage, the  $\eta$  spillover from a point source is expected to be localized since the convolution in Equation (11) has a significant contribution only due to the discontinuity in the weighting function at  $|\mathbf{x}| = x_{\max}$ . In the case of MWA snapshot visibilities, the weighting function is not smooth for  $|\mathbf{x}| < x_{\max}$  and the convolution in Equation (11) has a significant contribution from all  $|\mathbf{x}| < x_{\max}$ . Though this results in a distributed spillover, a significant part of the spillover energy is still restricted to  $\mathcal{R}_1$ , and  $\mathcal{R}_2$  will, in the absence of baseline migrations, be contaminated only due to the sidelobes of the convolving kernel from Equation (11) as discussed earlier. This is the PSF contamination that is expected to persist in  $\mathcal{R}_2$ . Any additional channel to channel variations in the PSF will contaminate  $\mathcal{R}_2$  beyond the PSF contamination. Such channel to channel variations are evident in the upper panel of Figure 7 in the form of jitter, especially for the cases of uniform weighting and Gaussian weighting. This jitter was found to be due to inter-channel baseline migrations as discussed in Section 4. To quantitatively estimate the amount of gridding contamination due to baseline migration, the Blackman–Nuttall window with very low sidelobes was used as a frequency window function  $B(f)$ . As described in Section 3.2, such a windowing function will significantly suppress PSF contamination in  $\mathcal{R}_2$  ( $\eta c > 500$  m),

and the dominant residual contamination in  $\mathcal{R}_2$  will be due to the inter-channel jitter arising from the gridding process.

It is interesting to note that baseline migrations have induced relatively higher jitter for the cases of uniform weighting and Gaussian weighting as opposed to natural weighting. As discussed in Section 4.1 this is due to two reasons. First, for an inter-channel frequency spacing of  $\delta f$ , a baseline vector  $\mathbf{x}$  (in meter units) suffers an inter-channel migration on the  $uv$  plane of length  $\mathbf{x}\delta f/c$  (in wavelength units). Second, and more importantly, a relatively higher amount of inter-channel jitter due to baseline migrations arises from regions of lower  $uv$  density (longer baseline vectors) compared to regions of higher  $uv$  density (smaller baseline vectors). Since uniform weighting, and in this case Gaussian weighting, places higher emphasis on longer baseline vectors as compared to natural weighting, the amount of jitter in Figure 7 goes in decreasing order of magnitude from uniform weighting to Gaussian weighting to natural weighting. Earth rotation synthesis and drift-scan synthesis will increase the  $uv$  density and is expected to ameliorate the jitter.

Case 2 involves image formation with the CZT algorithm with a similar image resolution at every frequency. As described in Section 4.4, image formation with the CZT algorithm is expected to lower gridding contamination in  $\mathcal{R}_2$  arising from inter-channel jitter due to baseline migrations. Figure 8 shows the LOS contamination for the case of image formation using the CZT algorithm. The jitter in flux density versus frequency



**Figure 8.** Foreground contamination along the line-of-sight dimension in the zenith pixel for the case of snapshot imaging with MWA using the Chirp Z Transform (CZT). Note how the jitter evident in Figure 7 has significantly reduced in the zenith pixel flux density vs. frequency plot (upper panel). Consequently, foreground contamination at higher values of  $\eta$  (lower plot) is significantly lower as compared to that in Figure 7. A finite amount of undesirable jitter still persists since a different set of baselines vectors were used in the gridding and imaging process at every frequency channel so as to keep the final image resolution constant at all frequencies.

(upper plot) has significantly reduced due to the absence of baseline migrations inherent to independent gridding at every frequency channel—an expected outcome of gridding the visibilities in meter units. Similar image resolution at every frequency channel was achieved by scaling the Gaussian taper to have the same form in wavelength units, and restricting the longest baseline vector that was used in the gridding and imaging process to the same value (in wavelength units) at every frequency channel. This resulted in inter-channel steps in the weighting function at the far end of the  $uv$  plane that resulted in a small but significant amount of jitter and hence a contamination in  $\mathcal{R}_2$ . Nevertheless, the contamination in  $\mathcal{R}_2$  for this case is reduced by a factor of 1.7, 6.4, and 21.7 for natural weighting, uniform weighting, and Gaussian weighting, respectively. Relative contamination for natural weighting has not improved significantly as regions in the  $uv$  plane that contribute most to jitter are down-weighted by natural weighting.

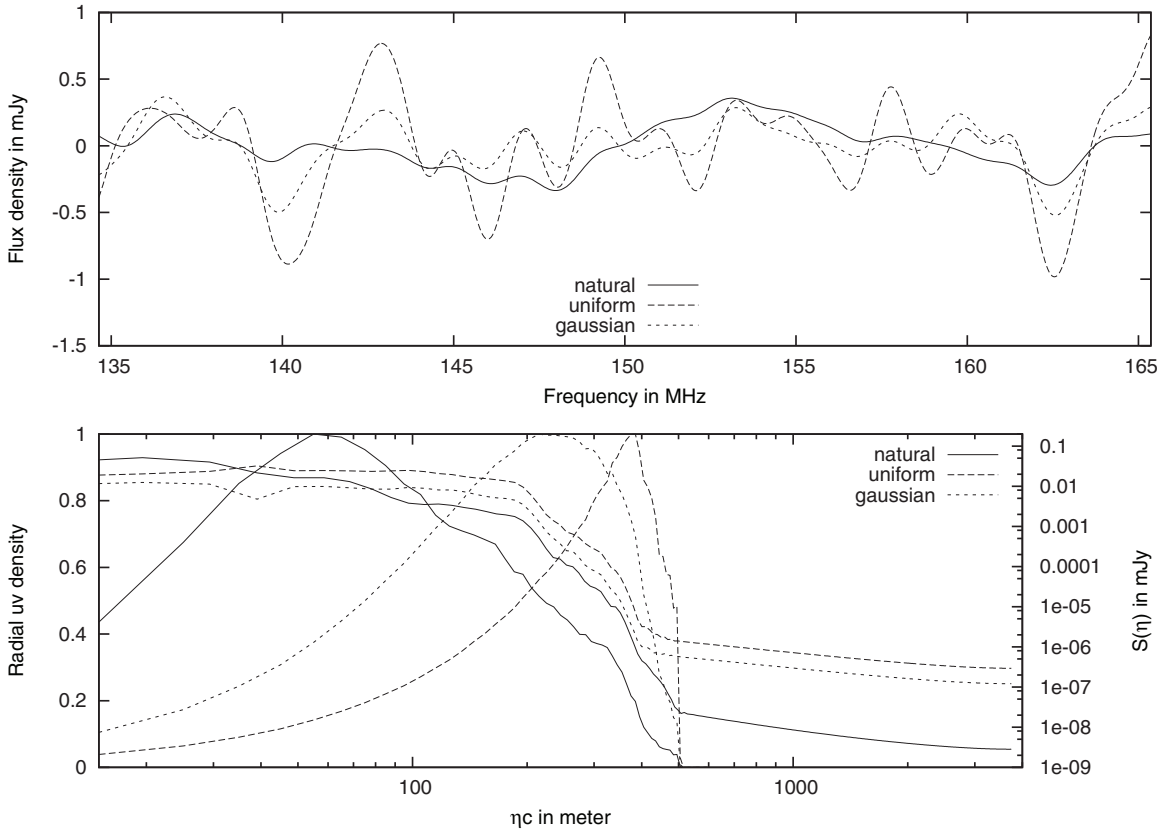
Case 3 involves image formation with the CZT algorithm with a frequency-dependent image resolution. In this case, the set of baselines that are used in the gridding and imaging routine are identical at all frequencies. It may be noted that despite a frequency-dependent image resolution, the CZT algorithm computes the sky flux density on the same  $lm$  grid at all frequencies. The results of this simulation are shown in Figure 9. The contamination in  $\mathcal{R}_2$  has dropped considerably owing to the elimination of channel to channel jitter in the PSF by this method. The extremely low sidelobes of the Blackman–Nuttall window and the absence of baseline migration in the CZT-based imaging algorithm we have proposed herein result

in extremely low levels of contamination in  $\mathcal{R}_2$ . Quantitatively, the contamination in  $\mathcal{R}_2$  for this simulation as compared to Case 1 has reduced by more than three orders of magnitude for all types of weighting schemes.

It is important to note that the amount of inter-channel jitter is largely dependent on the visibility distribution and in particular on the filling fraction in different parts of the  $uv$  plane. Therefore, the relative merits of the three different weighting schemes discussed will depend on the array configuration and observing strategy. It is however expected that the contamination in  $\mathcal{R}_2$  will be orders of magnitude lower if gridding and imaging are done using the proposed CZT algorithm along with a good frequency window. The simulations and results in this section are not exhaustive, and are included as a means to appreciate the various factors that influence the frequency dependence of the PSF and illustrate the potential reduction in foreground contamination.

## 6. CONCLUSIONS AND FUTURE WORK

Redshifted 21 cm tomography using Fourier synthesis arrays has emerged as a promising tool for EoR studies. Mitigating the confusion effects of Galactic and extragalactic foregrounds is a major challenge to such EoR experiments. Through this work we have furthered understanding of contamination arising from extragalactic continuum foreground and instrumental effects using a common framework. In particular, we have derived analytical expressions (Equation (11)) for contamination along frequency, or equivalently LOS dimension, which relates the



**Figure 9.** Foreground contamination along the line-of-sight dimension in the zenith pixel for the case of snapshot imaging using the Chirp Z Transform (CZT) with frequency-dependent image resolution. Note how there is negligible jitter in the flux density vs. frequency (upper panel). The gridding contamination at higher values of  $\eta$  (lower panel) in  $\mathcal{R}_2$  is now wholly owing to PSF contamination.

frequency-axis contamination in the image cube to the visibility weighting in Fourier space. In doing so, we have cast the problem in Fourier space where the telescope measurements indeed lie, thus enabling us to draw conclusions that are directly applicable to instrument design and data processing. Since the results in this paper address the general problem arising from a frequency-variant PSF, they are useful to all interferometric EoR experiments.

We identified two major sources of foreground contamination, “PSF contamination” and “gridding contamination,” and described their structure in LOS wavenumber space. The PSF contamination due to any point source is localized in LOS wavenumber space around  $\eta = x_{\max}l/c$ , where  $l$  is the direction cosine of the source with respect to the imaged pixel, and  $x_{\max}$  is the maximum baseline length in meters. Consequently, when imaging with continuum confusion that is offset in direction cosine up to a maximum of  $l_{\max}$  from an imaged pixel, most of the PSF contamination will be confined to a regime  $\mathcal{R}_1$  defined by  $\eta \in [0, x_{\max}l_{\max}/c]$ . It is useful to note that there does exist a regime  $\mathcal{R}_2$  defined by  $\eta \in [x_{\max}l_{\max}/c, 1/\Delta f]$ , where  $\Delta f$  is the frequency resolution, where small but finite PSF contamination persists. This contamination in  $\mathcal{R}_2$  may be further suppressed by judicious choice of a window function in frequency. For this reason,  $\mathcal{R}_2$  is an “EoR window” where we may focus our efforts to detect the EoR.

We have shown that gridding and imaging routines result in additional gridding contamination in LOS wavenumber space. In particular, we have shown that independent gridding at every frequency channel may lead to a stochastic inter-channel jitter in the PSF that potentially contaminates the EoR window  $\mathcal{R}_2$ .

To ameliorate problems associated with inter-channel jitter due to gridding, we proposed an alternative gridding and imaging algorithm based on the CZT. In this algorithm the visibilities are gridded in meter units rather than wavelength units, thereby eliminating gridding contamination. The CZT compensates for the stretching of baselines (in wavelength units) with frequency by introducing a “chirp” term—a scaling factor by which the frequency of the Fourier sinusoids are multiplied. We finally demonstrated the localization of PSF contamination and the elimination of gridding contamination in the CZT-based imaging algorithm using simulations of imaging with MWA.

Our ongoing and future work is directed toward using the concepts and ideas in this paper in an all sky simulation of MWA images that incorporates different gridding convolution kernels, Earth rotation and drift-scan synthesis, effects of array non-coplanarity, and the Galactic synchrotron emission. We are currently developing specific simulation pipelines that will accomplish the all sky simulations with the CZT-based imaging algorithm. Such simulations will quantitatively estimate not only the merit of the CZT-based imaging algorithm for EoR detection, but also the expected contamination from foreground continuum in future MWA data products.

## REFERENCES

- Blake, C., & Wall, J. 2002, *MNRAS*, **329**, L37  
 Bowman, J. D., Morales, M. F., & Hewitt, J. N. 2009, *ApJ*, **695**, 183  
 Bridle, A. H., & Schwab, F. R. 1989, in ASP Conf. Ser. 6, Wide Field Imaging I: Bandwidth and Time-average Smearing, ed. R. A. Perley, F. R. Schwab, & A. H. Bridle (San Francisco, CA: ASP), 247  
 Hazard, C., & Walsh, D. 1959, *MNRAS*, **119**, 648

- Jelić, V., Zaroubi, S., Labropoulos, P., et al. 2008, *MNRAS*, **389**, 1319
- Liu, A., Tegmark, M., Bowman, J. D., Hewitt, J. N., & Zaldarriaga, M. 2009, *MNRAS*, **398**, 401
- Lonsdale, C. J., Cappallo, R. J., Morales, M. F., et al. 2009, *Proc. IEEE*, **97**, 1497
- Morales, M. F., & Hewitt, J. N. 2004, *ApJ*, **615**, 7
- Nuttall, A. H. 1981, *IEEE Trans. Acoust. Speech Signal Process.*, **ASSP-29**, 84
- Rabiner, L. R., Schafer, R. W., & Rader, C. M. 1969, *IEEE Trans. Audio Electroacoust.*, **17**, 86
- Subrahmanyam, R., & Ekers, R. D. 2002, arXiv:[astro-ph/0209569v1](https://arxiv.org/abs/astro-ph/0209569v1)
- Thompson, A. R., Moran, J. M., & Swenson, G. W. 2001, *Interferometry and Synthesis in Radio Astronomy* (2nd ed.; New York: Wiley-Interscience)