

## THE RADIO SKY AT METER WAVELENGTHS: *M*-MODE ANALYSIS IMAGING WITH THE OWENS VALLEY LONG WAVELENGTH ARRAY

MICHAEL W. EASTWOOD,<sup>1</sup> MARIN M. ANDERSON,<sup>1</sup> RYAN M. MONROE,<sup>1</sup> GREGG HALLINAN,<sup>1</sup> AND OTHER PEOPLE<sup>2</sup>

<sup>1</sup> Department of Astronomy, California Institute of Technology, 1200 E California Blvd, Pasadena, CA 91125

<sup>2</sup> At some institution

### ABSTRACT

A host of new low-frequency radio telescopes are attempting to detect the 21 cm transition at cosmological distances. These measurements hold promise to directly probe star and galaxy formation at redshifts  $20 \gtrsim z \gtrsim 7$ , but are limited by the dynamic range they can achieve against foreground sources of low-frequency radio emission. Consequently, there is a growing demand for modern, high-fidelity maps of the sky at frequencies below 200 MHz for use in foreground modeling and removal. We use the Owens Valley Long Wavelength Array (OVRO-LWA) to generate 8 new maps of the sky north of  $\delta = -30^\circ$  with 15 arcmin angular resolution and at frequencies evenly spaced between 36.528 MHz and 73.152 MHz with 24 kHz bandwidth. Existing full-sky maps at these frequencies have angular resolutions  $\geq 2^\circ$  and so these maps are a 10-fold improvement in angular resolution. The thermal noise in each map is  $\sim 800$  mJy/beam. Future iterations will incorporate total power radiometry, and additional bandwidth and integration time. This will improve the thermal noise while simultaneously pushing to higher angular resolutions with the planned expansion of the OVRO-LWA to 2.6 km baselines. These maps are generated from the application of a new widefield imaging technique for drift-scanning interferometers, Tikhonov-regularized *m*-mode analysis imaging. This algorithm constructs images of the entire sky in a single synthesis imaging step with exact treatment of widefield effects. We additionally introduce this technique and describe how CLEAN can be adapted to deconvolve maps generated by *m*-mode analysis imaging. These maps serve as a first step on the path to the use of more sophisticated foreground filters in 21 cm cosmology incorporating the measured angular and frequency structure of all foreground contaminants.

*Keywords:* cosmology: observations – dark ages, reionization, first stars – radio continuum: galaxies – radio continuum: ISM

## 1. INTRODUCTION

At redshifts  $20 \gtrsim z \gtrsim 7$  the 21 cm hyperfine structure line of neutral hydrogen is expected to produce a 10 to 100 mK perturbation in the Cosmic Microwave Background (CMB) spectrum (Furlanetto et al. 2006; Pritchard & Loeb 2012). The amplitude of this perturbation on a given line-of-sight is a function of the neutral fraction of hydrogen, the baryon overdensity, the spin temperature relative to the CMB temperature at the given redshift, and the line-of-sight peculiar velocity of the gas. The spatial power spectrum of this perturbation is thought to be dominated by inhomogeneous heating of the IGM at  $z \sim 20$  (Fialkov et al. 2014), and by growing ionized bubbles during the EoR at  $z \sim 7$  where a detection can constrain the ionizing efficiency of early galaxies, the UV photon mean-free-path, and the minimum halo mass that can support star formation (Greig & Mesinger 2015).

Current 21 cm cosmology experiments can be broadly separated into two classes: global experiments that are aiming to detect the spectral signature of the cosmologically redshifted 21 cm transition after averaging over the entire sky (otherwise known as the monopole), and power spectrum experiments that incorporate angular information to attempt to measure the 3D spatial power spectrum of cosmological 21 cm perturbations. Ongoing global experiments include EDGES (Bowman & Rogers 2010; Monsalve et al. 2017), LEDA (Price et al. in prep.), BIGHORNS (Sokolowski et al. 2015), SCI-HI (Voytek et al. 2014), and SARAS 2 (Singh et al. 2017). Ongoing power spectrum experiments include PAPER/HERA (Ali et al. 2015; DeBoer et al. 2016), LOFAR (Patil et al. 2017), and the MWA (Beardsley et al. 2016; Ewall-Wice et al. 2016).

Just as for CMB experiments, foreground removal or suppression is an essential component of both classes of 21 cm cosmology experiments. The brightness temperature of the galactic synchrotron emission at high galactic latitudes is measured by Rogers & Bowman (2008) as

$$T \sim 300 \text{ K} \times \left( \frac{\nu}{150 \text{ MHz}} \right)^{-2.5}. \quad (1)$$

Therefore experiments conservatively need to achieve 5 orders of dynamic range against this foreground emission before the cosmological signal can be measured. Current foreground removal methods (for example, Parsons et al. 2012) rely on the assumption that the foreground emission is spectrally smooth. However the low-frequency radio sky is composed of several components: galactic synchrotron emission, supernova remnants, radio galaxies, free-free emission and absorption from H II regions, and a confusing background of radio

sources. Ideally a foreground removal strategy should be informed by the measured spatial structure and frequency spectrum of all foreground components. This possibility is limited by the availability of suitable high-fidelity low-frequency sky maps on angular scales ranging from tens of degrees to arcminutes.

The Global Sky Model (GSM) (de Oliveira-Costa et al. 2008; Zheng et al. 2017a) is currently the most commonly used foreground model. The GSM is a data-driven interpolation of various maps between 10 MHz and 100 GHz. However the majority of information contained in the GSM is derived at frequencies  $> 1.4$  GHz where the majority of the modern, high-fidelity input maps are located. Below 408 MHz, the GSM is a strong function of the Haslam 408 MHz map (Haslam et al. 1981, 1982; Remazeilles et al. 2015). As a consequence, the GSM is not a faithful representation of the sky below 200 MHz. For instance, Dowell et al. (2017) report errors of  $\pm 50\%$  between the GSM and their own maps at 74 MHz. This is largely due to the increasing contribution of free-free absorption and modifications to the synchrotron spectral index at low frequencies.

Recently a host of new low-frequency sky surveys have been conducted including MSSS (Heald et al. 2015), GLEAM (Wayth et al. 2015), and TGSS (Intema et al. 2017). However, the primary data product generated by these surveys is a catalog of radio point sources. At 45 MHz, Guzmán et al. (2011) created a map of the sky that captures the diffuse emission with  $5^\circ$  resolution. The LWA1 Low Frequency Sky Survey (Dowell et al. 2017) similarly maps the sky at a range of frequencies between 35 MHz and 80 MHz with resolution between  $4.5^\circ$  and  $2^\circ$ .

This paper presents a series of new low-frequency maps of the sky between 36.528 MHz and 73.152 MHz, capturing the full sky visible from the Owens Valley Radio Observatory (OVRO) with angular resolution of  $\sim 15$  arcmin. We aim for these maps to inform foreground removal strategies in 21 cm cosmology as they represent some of the highest angular resolution full-sky maps ever produced at these frequencies. We anticipate additional ancillary science taking advantage of the combination of high fidelity and high resolution of these maps, including but not limited to studies of the cosmic ray emissivity at low frequencies, searches for giant radio galaxies, and constraining the galactic synchrotron spectrum. The maps will be made freely available online at the Legacy Archive for Microwave Background Data Analysis (LAMBDA)<sup>1</sup>.

<sup>1</sup> <https://lambda.gsfc.nasa.gov/>

The structure of this paper is as follows. In §2, we present Tikhonov-regularized  $m$ -mode analysis imaging, a new imaging technique that allows us to image the entire visible sky in one coherent synthesis imaging step with exact widefield corrections. In §3 we describe our observations with the Owens Valley Long Wavelength Array (OVRO-LWA). In §4 we present the sky maps, and in §5 we discuss some of the sources of error present in the maps. Finally in §6 the new sky maps are compared against other low-frequency sky maps, and in §7 we present our conclusions.

## 2. ALL-SKY IMAGING

The goal of all imaging algorithms is to estimate the brightness of the sky  $I_\nu(\hat{r})$  in the direction  $\hat{r}$  and frequency  $\nu$ . A radio interferometer measures the visibilities  $V_\nu^{ij}$  between pairs of antennas – here numbered  $i$  and  $j$  respectively. If the antennas are separated by the baseline  $\vec{b}_{ij}$ , and  $A_\nu(\hat{r})$  describes their response to the incident radiation, then

$$V_\nu^{ij} = \int_{\text{sky}} A_\nu(\hat{r}) I_\nu(\hat{r}) \exp\left(2\pi i \hat{r} \cdot \vec{b}_{ij}/\lambda\right) d\Omega. \quad (2)$$

Constructing an image from the output of a radio interferometer consists of estimating  $I_\nu(\hat{r})$  given the available measurements  $V_\nu^{ij}$ .

For later convenience we will define the baseline transfer function  $B_\nu^{ij}(\hat{r})$  such that

$$V_\nu^{ij} = \int_{\text{sky}} B_\nu^{ij}(\hat{r}) I_\nu(\hat{r}) d\Omega. \quad (3)$$

The baseline transfer function defines the response of a single baseline to the sky, and is a function of the antenna primary beam, and baseline length and orientation.

Naively one might attempt to solve Equation 2 by discretizing, and subsequently solving the resulting matrix equation. If the interferometer is composed of  $N_{\text{base}}$  baselines, and measures  $N_{\text{freq}}$  frequency channels over  $N_{\text{time}}$  integrations, then the entire data set consists of  $N_{\text{base}}N_{\text{freq}}N_{\text{time}}$  complex numbers. If the sky is discretized into  $N_{\text{pix}}$  pixels then the relevant matrix has dimensions of  $(N_{\text{base}}N_{\text{freq}}N_{\text{time}}) \times (N_{\text{pix}})$ . For making single-channel maps with the OVRO-LWA this becomes a 5 petabyte array (assuming each matrix element is a 64-bit complex floating point number). This matrix equation is therefore prohibitively large, and solving Equation 2 by means of discretization is usually impossible although Zheng et al. (2017b) demonstrate this technique with the MITEOR telescope.

Instead it is common to make mild assumptions that simplify Equation 2 and ease the computational burden in solving for  $I_\nu(\hat{r})$ . For example, when all of

the baselines  $\vec{b}_{ij}$  lie in a plane and the field-of-view is small, Equation 2 can be well-approximated by a two-dimensional Fourier transform (Thompson et al. 2001). The restriction on baseline co-planarity and field-of-view can be relaxed by using W-projection (Cornwell et al. 2008). Known primary beam effects can also be accounted for during imaging by using A-projection (Bhatnagar et al. 2013).

### 2.1. $m$ -Mode Analysis

On the other hand, transit telescopes can take advantage of a symmetry in Equation 2 that greatly reduces the amount of computer time required to image the full-sky with exact incorporation of widefield imaging effects. This technique, called  $m$ -mode analysis, also obviates the need for gridding and mosaicing. Instead the entire sky is imaged in one coherent synthesis imaging step.

In this context we will define a transit telescope as any interferometer where the response pattern of the individual elements does not change with respect to time. This may be an interferometer like the OVRO-LWA where the correlation elements are fixed dipoles, but it may also be an interferometer like LOFAR or the MWA if the steerable beams are held in a fixed position (not necessarily at zenith).

We will briefly summarize  $m$ -mode analysis below, but the interested reader should consult Shaw et al. (2014, 2015) for a complete derivation.

For a transit telescope, the visibilities  $V_\nu^{ij}$  are a periodic function of sidereal time.<sup>2</sup> Therefore it is a natural operation to compute the Fourier transform of the visibilities with respect to sidereal time  $\phi \in [0, 2\pi]$ .

$$V_{m,\nu}^{ij} = \int_0^{2\pi} V_\nu^{ij}(\phi) \exp\left(-im\phi\right) d\phi \quad (4)$$

The output of this Fourier transform is the set of  $m$ -modes  $V_{m,\nu}^{ij}$  where  $m = 0, \pm 1, \pm 2, \dots$  is the Fourier conjugate variable to the sidereal time. The  $m$ -mode corresponding to  $m = 0$  is a simple average of the visibilities over sidereal time. Similarly  $m = 1$  corresponds to the component of the visibilities that varies over half-day timescales. Larger values of  $m$  correspond to components that vary on quicker timescales.

Shaw et al. (2014, 2015) show that there is a discrete linear relationship between the measured  $m$ -modes  $V_{m,\nu}^{ij}$  and the spherical harmonic coefficients of the sky bright-

<sup>2</sup> This is not strictly true. Ionospheric fluctuations and non-sidereal sources (such as the Sun) will violate this assumption. This paper will, however, demonstrate that the impact on the final maps is mild.

ness  $a_{lm,\nu}$ .

$$V_{m,\nu}^{ij} = \sum_l B_{lm,\nu}^{ij} a_{lm,\nu}, \quad (5)$$

where the transfer coefficients  $B_{lm,\nu}^{ij}$  define the interferometer's response to the sky. These transfer coefficients are computed from the spherical harmonic transform of the baseline transfer function defined by Equation 3.

Equation 5 can be recognized as a matrix equation where the transfer matrix  $\mathbf{B}$  is block-diagonal.

$$\underbrace{\begin{pmatrix} \vdots \\ m\text{-modes} \\ \vdots \end{pmatrix}}_{\mathbf{v}} = \underbrace{\begin{pmatrix} \ddots & & \\ & \text{transfer matrix} & \\ & & \ddots \end{pmatrix}}_{\mathbf{B}} \underbrace{\begin{pmatrix} \vdots \\ a_{lm} \\ \vdots \end{pmatrix}}_{\mathbf{a}} \quad (6)$$

$$\mathbf{B} = \begin{pmatrix} m=0 & & & \\ & m=\pm 1 & & \\ & & m=\pm 2 & \\ & & & \ddots \end{pmatrix} \quad (7)$$

The vector  $\mathbf{v}$  contains the list of  $m$ -modes and the vector  $\mathbf{a}$  contains the list of spherical harmonic coefficients representing the sky brightness. In order to take advantage of the block-diagonal structure in  $\mathbf{B}$ ,  $\mathbf{v}$  and  $\mathbf{a}$  must be sorted by the absolute value of  $m$ . Positive and negative values of  $m$  are grouped together because the brightness of the sky is real-valued, and the spherical harmonic transform of a real-valued function has  $a_{l(-m)} = (-1)^m a_{lm}^*$ .

In practice we now need to pick the set of spherical harmonics we will use to represent the sky. For an interferometer like the OVRO-LWA with lots of short baselines, a sensible choice is to use all spherical harmonics with  $l \leq l_{\max}$  for some  $l_{\max}$ . The parameter  $l_{\max}$  is determined by the maximum baseline length of the interferometer. For an interferometer without short spacings, a minimum value for  $l$  might also be used. This  $l_{\min}$  parameter should be determined by the minimum baseline length. A rough estimate of  $l$  for a baseline of length  $b$  at frequency  $\nu$  is  $l \sim \pi b \nu / c$ . For the OVRO-LWA we therefore adapt  $l_{\min} = 0$  and  $l_{\max} = 1000$  across all frequencies. However this choice of  $l_{\max}$  actually limits the angular resolution above 55 MHz, and therefore future work will increase  $l_{\max}$  to obtain better angular resolution.

The interferometer's sensitivity to the monopole ( $a_{00}$ ) deserves special consideration. Venumadhav et al. (2016) prove – under fairly general assumptions – that a

baseline with nonzero sensitivity to  $a_{00}$  must also have some amount of cross-talk or common-mode noise. In fact the sensitivity to  $a_{00}$  is proportional to a sum of these effects. For example, one way a baseline can have nonzero sensitivity to  $a_{00}$  is if the baseline is extremely short. In this case the antennas are so close together that voltage fluctuations in one antenna can couple into the other antenna. In order to make an interferometric measurement of  $a_{00}$ , this coupling must be measured and calibrated. Consequently we set  $a_{00} = 0$  in our analysis. In the future this will be addressed with the inclusion of calibrated total power radiometry.

The size of a typical block in the transfer matrix is  $(2N_{\text{base}}N_{\text{freq}}) \times (l_{\max})$ . If each element of the matrix is stored as a 64-bit complex floating point number, a single block is 500 MB for the case of single-channel imaging with the OVRO-LWA, which a modern computer can easily store and manipulate in memory. The power of  $m$ -mode analysis is the block-diagonal structure of Equation 6. By breaking up the problem into  $N$  independent blocks, the computational complexity involved in inverting a single block is reduced by a factor of  $N^3$  relative to a dense matrix. However we must invert  $N$  blocks, so the overall computational savings is proportional to  $N^2$ . For the case of the OVRO-LWA the equation breaks up into  $\sim 10^3$  blocks and so we save a factor of  $\sim 10^6$  in processing time by using  $m$ -mode analysis.

## 2.2. $m$ -Mode Analysis Imaging

Imaging in  $m$ -mode analysis essentially amounts to inverting Equation 6 to solve for the spherical harmonic coefficients  $\mathbf{a}$ . The linear-least squares solution, which minimizes  $\|\mathbf{v} - \mathbf{Ba}\|^2$ , is given by

$$\hat{\mathbf{a}}_{\text{LLS}} = (\mathbf{B}^* \mathbf{B})^{-1} \mathbf{B}^* \mathbf{v}, \quad (8)$$

where  $*$  indicates the conjugate-transpose. However, usually one will find that  $\mathbf{B}$  is not full-rank and hence  $\mathbf{B}^* \mathbf{B}$  is not an invertible matrix. For example, an interferometer located in the northern hemisphere will never see a region of the southern sky centered on the southern celestial pole. The  $m$ -modes contained in the vector  $\mathbf{v}$  must contain no information about the sky around the southern celestial pole, and therefore the act of multiplying by  $\mathbf{B}$  must destroy some information about the sky. The consequence of this fact is that  $\mathbf{B}$  must have at least one singular value that is equal to zero. It is then trivial to show that  $\mathbf{B}^* \mathbf{B}$  must have at least one eigenvalue that is equal to zero, which means it is not an invertible matrix.

Another way of looking at the problem is that because the interferometer is not sensitive to part of the southern

hemisphere, there are infinitely many possible solutions to Equation 6 that will fit the measured data equally well. So we need to regularize the problem and apply an additional constraint that prefers a unique solution. For example, you may prefer that in the absence of any information the sky should be set to zero or you may prefer that the sky should be set to some prior expectation.

### 2.2.1. Tikhonov Regularization

The process of Tikhonov regularization minimizes  $\|\mathbf{v} - \mathbf{Ba}\|^2 + \varepsilon\|\mathbf{a}\|^2$  for some arbitrary value of  $\varepsilon > 0$  chosen by the observer. The solution that minimizes this expression is given by

$$\hat{\mathbf{a}}_{\text{Tikhonov}} = (\mathbf{B}^* \mathbf{B} + \varepsilon \mathbf{I})^{-1} \mathbf{B}^* \mathbf{v}. \quad (9)$$

Tikhonov regularization adds a small value  $\varepsilon$  to the diagonal of  $\mathbf{B}^* \mathbf{B}$ , fixing the matrix's singularity. We can see this by using the singular value decomposition (SVD) of the matrix  $\mathbf{B} = \mathbf{U} \Sigma \mathbf{V}^*$ . Equation 9 becomes

$$\hat{\mathbf{a}}_{\text{Tikhonov}} = \mathbf{V} (\Sigma^2 + \varepsilon \mathbf{I})^{-1} \Sigma \mathbf{U}^* \mathbf{v}, \quad (10)$$

where

$$\Sigma = \begin{pmatrix} \sigma_1 & & \\ & \sigma_2 & \\ & & \ddots \end{pmatrix}.$$

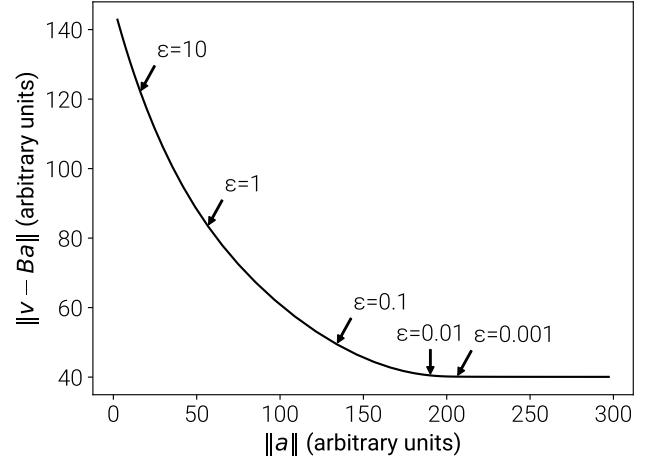
The diagonal elements of  $\Sigma$  are the singular values of  $\mathbf{B}$  and so the contribution of each singular component to the Tikhonov-regularized solution is scaled by  $\sigma_i / (\sigma_i^2 + \varepsilon)$ , where  $\sigma_i$  is the singular value for the  $i$ th singular component. Tikhonov regularization therefore acts to suppress any component for which  $\sigma_i \lesssim \sqrt{\varepsilon}$ . If  $\sigma_i = 0$ , the component is set to zero.

In practice the measurement  $\mathbf{v}$  is corrupted by noise with covariance  $\mathbf{N}$ . For illustrative purposes we will assume that  $\mathbf{N} = n\mathbf{I}$  for some scalar  $n$ . In this case the covariance of the Tikhonov-regularized spherical harmonic coefficients is

$$\mathbf{C} = n\mathbf{V} (\Sigma^2 + \varepsilon \mathbf{I})^{-2} \Sigma^2 \mathbf{V}^*. \quad (11)$$

Each singular component is scaled by a factor of  $\sigma_i^2 / (\sigma_i^2 + \varepsilon)^2$ . In the absence of Tikhonov regularization ( $\varepsilon = 0$ ) singular components with the smallest singular values – the ones that the interferometer is the least sensitive to – actually come to dominate the covariance of the measured spherical harmonic coefficients. Tikhonov regularization improves this situation by downweighting these components.

While Tikhonov regularization will force unmeasured modes to zero, if a prior map of the sky already exists, it will be preferable to instead minimize  $\|\mathbf{v} - \mathbf{Ba}\|^2 + \varepsilon\|\mathbf{a} - \mathbf{a}_{\text{prior}}\|^2$ .



**Figure 1.** An example L-curve using OVRO-LWA data at 36.528 MHz. The  $x$ -axis is the norm of the solution (in this case the spherical harmonic coefficients) given in arbitrary units. The  $y$ -axis is the least-squares norm given in arbitrary units.  $\varepsilon$  is the regularization parameter. When the regularization parameter is small, the norm of the solution grows rapidly. When the regularization parameter is large, the least-squares norm grows rapidly.

### 2.2.2. L-Curves

Tikhonov regularization requires the observer to pick the value of  $\varepsilon$ . If  $\varepsilon$  is too large then too much importance is placed on minimizing the norm of the solution and the least-squares residuals will suffer. However if  $\varepsilon$  is too small then the problem will be poorly regularized and the resulting sky map may not represent the true sky. Picking the value of  $\varepsilon$  therefore requires understanding the trade-off between the two norms.

This trade-off can be analyzed quantitatively by trialing several values of  $\varepsilon$ , and computing  $\|\mathbf{v} - \mathbf{Ba}\|^2$  and  $\|\mathbf{a}\|^2$  for each trial. An example is shown in Figure 1. The shape of this curve has a characteristic L-shape, and as a result this type of plot is called an L-curve. The ideal value of  $\varepsilon$  lies near the turning point of the plot. At this point a small decrease in  $\varepsilon$  will lead to an undesired rapid increase in  $\|\mathbf{a}\|^2$ , and a small increase in  $\varepsilon$  will lead to an undesired rapid increase in  $\|\mathbf{v} - \mathbf{Ba}\|^2$ .

In practice, the L-curve should be used as a guide to estimate a reasonable value of  $\varepsilon$ . However better results can often be obtained by tuning the value of  $\varepsilon$ . For instance increasing the value of  $\varepsilon$  can improve the noise properties of the map by down-weighting noisy modes. Decreasing the value of  $\varepsilon$  can improve the resolution in the map by up-weighting the contribution of longer baselines, which are likely fewer in number. In this respect choosing the value of  $\varepsilon$  is analogous to picking the weighting scheme in traditional imaging where robust

weighting schemes can be tuned to similar effect (Briggs 1995).

### 2.2.3. The Moore-Penrose Pseudoinverse

The Moore-Penrose pseudoinverse (denoted in this paper with a superscript  $\dagger$ ), is commonly applied to find the minimum-norm linear-least squares solution to a set of linear equations. This can be used in place of Tikhonov regularization as

$$\hat{\mathbf{a}}_{\text{Moore-Penrose}} = \mathbf{B}^\dagger \mathbf{v}. \quad (12)$$

Much like Tikhonov regularization, the Moore-Penrose pseudoinverse sets components with small singular values (below some user-defined threshold) to zero. Components with large singular values (above the user-defined threshold) are included in the calculation at their full amplitude with no down-weighting of modes near the threshold. The essential difference between using the Moore-Penrose pseudoinverse and Tikhonov regularization is that the pseudoinverse defines a hard transition from on to off. Modes are either set to zero or included in the map at their full amplitude. On the other hand Tikhonov regularization smoothly interpolates between these behaviors. Because of this, Tikhonov regularization tends to produce better results in practical applications.

## 2.3. CLEAN

In traditional radio astronomy imaging, the CLEAN algorithm (Högbom 1974) is used to interpolate between measured visibilities on the  $uv$ -plane. In the absence of this interpolation, gaps in the interferometer's  $uv$ -coverage are assumed to be zero, and – in the image plane – sources are convolved with a point spread function (PSF) that is characteristic of the  $uv$ -coverage. Fundamentally, the interferometer's PSF is determined by which modes were assumed to be zero in the initial imaging process.

In  $m$ -mode analysis imaging, we assumed modes were zero in two separate ways:

1. We selected a set of spherical harmonic coefficients  $a_{lm}$  to describe the sky brightness distribution. All modes with  $l > l_{\max}$  are neglected and assumed to be zero.
2. Tikhonov regularization forces linear combinations of spherical harmonic coefficients with  $\sigma_i \lesssim \sqrt{\varepsilon}$  towards zero.

As a consequence the final map of the sky is not assembled from a complete set of spherical harmonics.

Therefore just as in traditional imaging,  $m$ -mode analysis imaging produces dirty maps where sources are convolved with a PSF. This PSF can be improved by increasing the number and variety of baselines, which increases the number of modes where  $\sigma_i \gg \sqrt{\varepsilon}$ . Alternatively by collecting more data, the signal-noise ratio of the measured  $m$ -modes increases, which allows the observer to lower the value of  $\varepsilon$  without increasing the noise in the maps. Finally, the CLEAN algorithm can be applied to interpolate some of the missing information that was assumed to be zero.

The PSF of a dirty  $m$ -mode analysis map may be computed with

$$\mathbf{a}_{\text{PSF}}(\theta, \phi) = (\mathbf{B}^* \mathbf{B} + \varepsilon \mathbf{I})^{-1} \mathbf{B}^* \mathbf{B} \mathbf{a}_{\text{PS}}(\theta, \phi), \quad (13)$$

where  $\mathbf{a}_{\text{PSF}}(\theta, \phi)$  is the vector of spherical harmonic coefficients representing the PSF at the spherical coordinates  $(\theta, \phi)$ , and  $\mathbf{a}_{\text{PS}}(\theta, \phi)$  is the vector of spherical harmonic coefficients for a point source at  $(\theta, \phi)$  given by

$$\mathbf{a}_{\text{PS}}(\theta, \phi) = \begin{pmatrix} \vdots \\ Y_{lm}^*(\theta, \phi) \\ \vdots \end{pmatrix} = \begin{pmatrix} \vdots \\ Y_{lm}^*(\theta, 0) \times e^{im\phi} \\ \vdots \end{pmatrix}. \quad (14)$$

In general the PSF can be a function of the right ascension and declination. However point sources at the same declination take the same track through the sky and therefore (barring any ionospheric effects) will have the same PSF. The PSF is therefore only a function of the declination. For example, sources at low elevations will tend to have an extended PSF along the north-south axis due to baseline foreshortening. Some examples PSFs are shown in Figure 3. Adapting CLEAN for  $m$ -mode analysis requires either pre-computing Equation 13 at a grid of declinations, or a method for rapidly evaluating Equation 13 on the fly.

For an interferometer with more baselines than spherical harmonics used in the maps (eg. the OVRO-LWA),  $\mathbf{B}^* \mathbf{B}$  can be a much smaller matrix than the full transfer matrix  $\mathbf{B}$ . Therefore pre-computing  $\mathbf{B}^* \mathbf{B}$  can allow the entire matrix to fit into memory on a single machine. This greatly reduces the amount of disk I/O necessary for solving Equation 13.

Additionally we can pre-compute the Cholesky decomposition of  $\mathbf{B}^* \mathbf{B} + \varepsilon \mathbf{I} = \mathbf{U}^* \mathbf{U}$ , where  $\mathbf{U}$  is an upper-triangular matrix. Inverting an upper triangular matrix is an  $\mathcal{O}(N^2)$  operation (instead of  $\mathcal{O}(N^3)$  for a general

matrix inverse).<sup>3</sup> Equation 13 can then be rapidly evaluated from right to left as:

$$\mathbf{a}_{\text{PSF}} = \mathbf{U}^{-1} (\mathbf{U}^*)^{-1} (\mathbf{B}^* \mathbf{B}) \mathbf{a}_{\text{PS}}. \quad (15)$$

Furthermore, Equation 15 does not need to be separately evaluated for each CLEAN component. Instead we can identify  $N$  CLEAN components, accumulate  $\mathbf{a}_{\text{PS}}$  for each component, and evaluate Equation 15 on the accumulation. This can greatly reduce the number of times this equation needs to be evaluated, but care must be taken to ensure that the  $N$  components are not so close together that sidelobes from one may interact with another.

Altogether the adaptation of CLEAN applied to the maps presented in this paper is summarized below.

**Precondition:**  $\mathbf{a}$  is the solution to Equation 9

```

1: function CLEAN( $\mathbf{a}$ )
2:    $\mathbf{M} \leftarrow \mathbf{B}^* \mathbf{B}$ 
3:    $\mathbf{U} \leftarrow \text{chol}(\mathbf{M} + \varepsilon \mathbf{I})$   $\triangleright$  Cholesky decomposition
4:   while noise in map > threshold do
5:     find  $N$  pixels with the largest residual flux
6:      $\mathbf{x} \leftarrow \sum_{i=1}^N (\text{pixel flux}) \times \mathbf{a}_{\text{PS}}(\theta_i, \phi_i)$ 
7:      $\mathbf{y} \leftarrow \mathbf{U}^{-1} (\mathbf{U}^*)^{-1} \mathbf{M} \mathbf{x}$ 
8:      $\mathbf{a} \leftarrow \mathbf{a} - (\text{loop gain}) \times \mathbf{y}$ 
9:     record subtracted components
10:     $\mathbf{a} \leftarrow \mathbf{a} + (\text{restored components})$ 
11:   return  $\mathbf{a}$ 
```

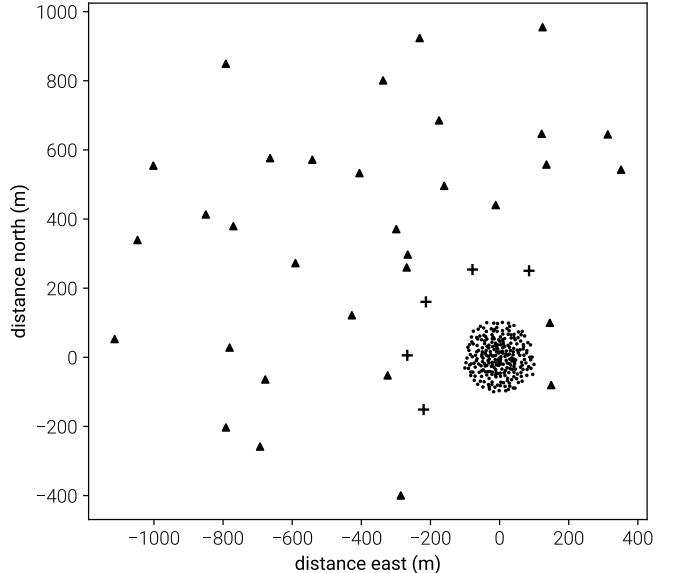
In summary, Tikhonov-regularized  $m$ -mode analysis imaging constructs a widefield synthesis image of the sky from a complete earth rotation, and with exact treatment of widefield-effects. This is accomplished by solving a regularized block-diagonal matrix equation (Equation 9). The solution to this equation generates a map where sources are convolved with a PSF characteristic of the interferometer (a function of the frequency, antenna response, and baseline distribution with a full Earth rotation). The CLEAN algorithm is adopted to deconvolve the PSF and produce the final sky maps.

### 3. OBSERVATIONS

#### 3.1. The Owens Valley Long Wavelength Array

The Owens Valley Long Wavelength Array (OVRO-LWA) is a 288-element interferometer located at the Owens Valley Radio Observatory near Big Pine, California (Hallinan et al. in prep.). The OVRO-LWA is a low-frequency instrument with instantaneous bandwidth covering 27.384 MHz to 84.912 MHz and with 24

<sup>3</sup> Inverting a matrix is numerically unstable and should be avoided. Instead of computing  $\mathbf{A}^{-1}$ , the linear equation  $\mathbf{Ax} = \mathbf{b}$  is solved each time the matrix inverse is needed.

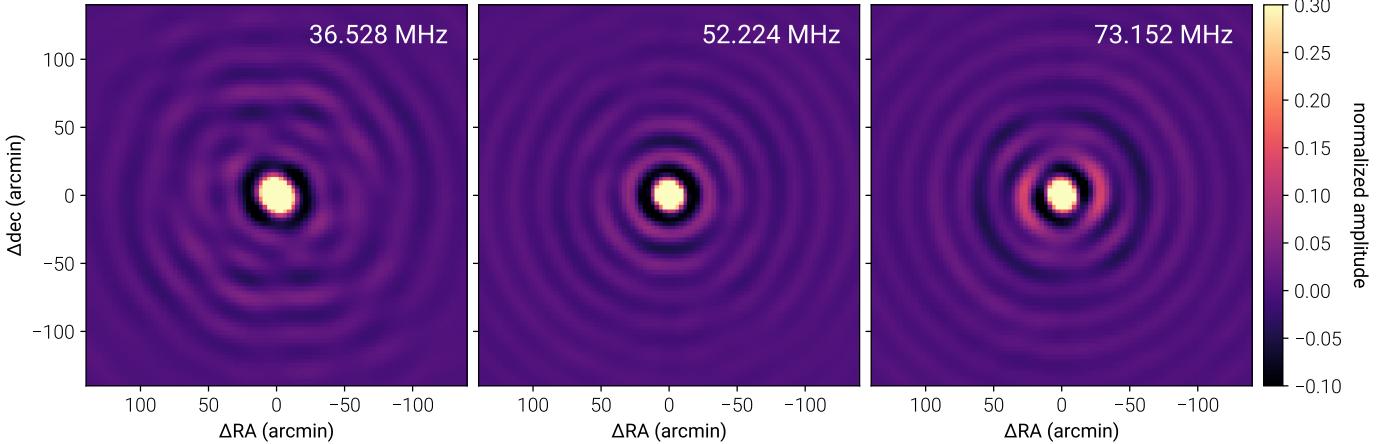


**Figure 2.** This figure shows the antenna layout for the OVRO-LWA. Black dots correspond to antennas within the 200 m diameter core of the array. The 32 triangles are the expansion antennas built in early 2016 in order to increase the longest baseline to  $\sim 1.5$  km. The 5 crosses are antennas equipped with noise-switched front ends.

kHz channelization. Each antenna stand hosts two perpendicular broadband dipoles so that there are  $288 \times 2$  signal paths in total. These signal paths feed into the 512-input LEDA correlator (Kocz et al. 2015), which allows the OVRO-LWA to capture the entire visible hemisphere in a single snapshot image. In the current configuration 32 antennas (64 signal paths) are unused.

The 288 antennas are arranged in a pseudo-random configuration optimized to minimize sidelobes in snapshot imaging (see Figure 2). 251 of the antennas are contained within a 200 m diameter core. 32 antennas are placed outside of the core in order to extend the maximum baseline length to  $\sim 1.5$  km. The final 5 antennas are equipped with noise-switched front ends for calibrated total power measurements of the global sky brightness. These antennas are used as part of the LEDA experiment (Price et al. in prep.) to measure the global signal of 21 cm absorption from the cosmic dawn. A final stage of construction will involve 64 antennas installed on long baselines out to a maximum length of 2.6 km.

The dataset used in this paper spans 28 consecutive hours beginning at 2017-02-17 12:00:00 UTC time. During this time the OVRO-LWA operated as a zenith-pointing drift scanning interferometer. The correlator dump time was selected to be 13 seconds such that the correlator output evenly divides a sidereal day.



**Figure 3.** This figure shows the  $m$ -mode analysis imaging PSF at  $\delta = +45^\circ$  at 3 separate frequencies: 36.528 MHz (left panel), 52.224 MHz (middle panel), and 73.152 MHz (right panel). The PSF is computed by evaluating Equation 13.

### 3.2. Gain Calibration

Antenna gain calibration is accomplished using alternating least squares otherwise known as “stefcal” (Mitchell et al. 2008; Salvini & Wijnholds 2014). Mitchell et al. (2008) and Salvini & Wijnholds (2014). The calibration routine is written in the Julia programming language (Bezanson et al. 2017), and is publicly available online<sup>4</sup> under an open source license (GPLv3 or any later version).

Cyg A and Cas A are – by an order of magnitude – the brightest point-like radio sources in the northern hemisphere. Therefore the optimal time to solve for the interferometer’s gain calibration is when these sources are at high elevations. The antenna complex gains are measured from a 22 minute track of data from when Cyg A and Cas are at high elevations. The gains measured in this way are then used to calibrate the entire 28 hour dataset. The calibration sky model consists only of Cyg A and Cas A. The model flux of Cyg A is set to the Baars et al. (1977) spectrum while the flux of Cas A is measured from the data itself (using a preliminary calibration). Residual errors in the astrometry are corrected post-imaging by registering the images with respect to all VLSSr (Lane et al. 2014) sources brighter than 30 Jy. Additional cuts are made to ensure that sources aren’t too close to each other or other bright sources, and that the VLSSr flux density matches is consistent with the flux density measured by the OVRO-LWA.

Temperature fluctuations within the electronics shelter generate 0.1 dB sawtooth oscillations in the analog gain. These oscillations occur with a variable 15 to 17 minute period. The amplitude of these gain fluctuations is calibrated by smoothing the autocorrelation

amplitudes on 45 minute timescales. The ratio of the measured auto-correlation power to the smoothed auto-correlation power defines a per-antenna amplitude correction that is then applied to the cross-correlations.

### 3.3. Primary Beam Measurements

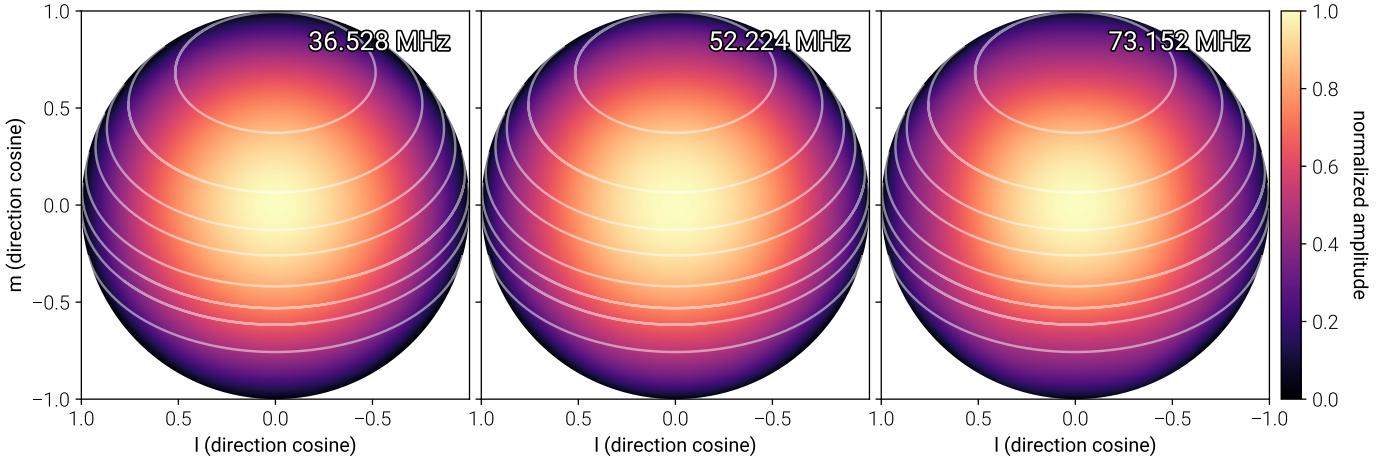
In order to generate wide-field images of the sky, the response of the antenna to the sky must be known. Drift-scanning interferometers like the OVRO-LWA can empirically measure their primary beam under a mild set of symmetry assumptions (Pober et al. 2012). In this work we assume that the primary beam is invariant under north-south and east-west flips, and additionally that the  $x$ - and  $y$ -dipoles have the same response to the sky after rotating one by  $90^\circ$ . These symmetries are apparent in the antenna design, but real-world defects and coupling with nearby antennas will contribute towards breaking these symmetries at some level. However some amount of symmetry must be assumed in order to break the degeneracy between source flux and beam amplitude when the flux of a source is unknown.

We measure the flux of several bright sources (Cyg A, Cas A, Tau A, Vir A, Her A, Hya A, Per B, and 3C 353) as they pass through the sky and then fit a beam model composed of Zernike polynomials to those flux measurements. We select the basis functions to have the desired symmetry ( $Z_0^0$ ,  $Z_2^0$ ,  $Z_4^0$ ,  $Z_4^4$ ,  $Z_6^0$ ,  $Z_6^4$ ,  $Z_8^0$ ,  $Z_8^4$ , and  $Z_8^8$ ) and the beam amplitude at zenith is constrained to be unity. See Figure 4 for an illustration of a fitted beam model at several frequencies. This process is repeated for each frequency channel.

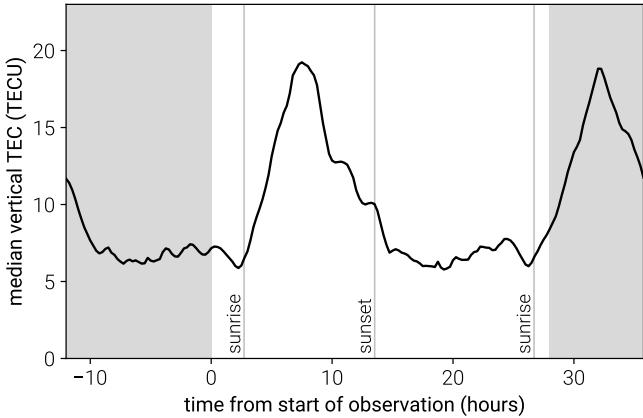
### 3.4. Ionospheric Conditions

The geomagnetic conditions during this time were mild. The Disturbance storm time (Dst) index was  $> -30$  nT during the entirety of the observing pe-

<sup>4</sup> <https://github.com/mweastwood/TTCal.jl>



**Figure 4.** This figure shows the empirical fits to the OVRO-LWA Stokes-I primary beam (the response of the  $x$ - and  $y$ -dipoles has been summed) at three frequencies: 36.528 MHz (left panel), 52.224 MHz (middle panel), and 73.152 MHz (right panel). The source tracks used to measure the beam model are overlaid. From north to south, these tracks correspond to Cas A, Cyg A, Per B, Tau A, Vir A, Her A, 3C 353, and Hya A. The fitting process is described in §3.3, and residuals for Cyg A and Cas A can be seen in Figure 6.



**Figure 5.** The median vertical TEC within 200 km of OVRO during the time of the observation. The gray shaded regions indicate times outside of the observing period. The gray vertical lines indicate sunrise and sunset (as labeled).

riod.<sup>5</sup> Following the classification scheme of Kintner et al. (2008), a weak geomagnetic storm has  $Dst < -30$  nT. Stronger geomagnetic storms have  $Dst \ll -30$  nT.

Despite the mild conditions, low frequency interferometric observations are still affected by the index of refraction in the ionosphere. Figure 5 shows the median vertical total electron content (TEC) above OVRO measured from GPS (Iijima et al. 1999). The median is computed over all GPS measurements within 200 km of the observatory. Over the observing period the TEC smoothly varies from 20 TECU at midday to 5 TECU

<sup>5</sup> The Dst index was obtained from the World Data Center for Geomagnetism, Kyoto University (<http://swdcwww.kugi.kyoto-u.ac.jp/>). Accessed 2017-07-25.

during the night. However, this measurement is only sensitive to large scale fluctuations in the ionosphere and doesn't capture small scale fluctuations.

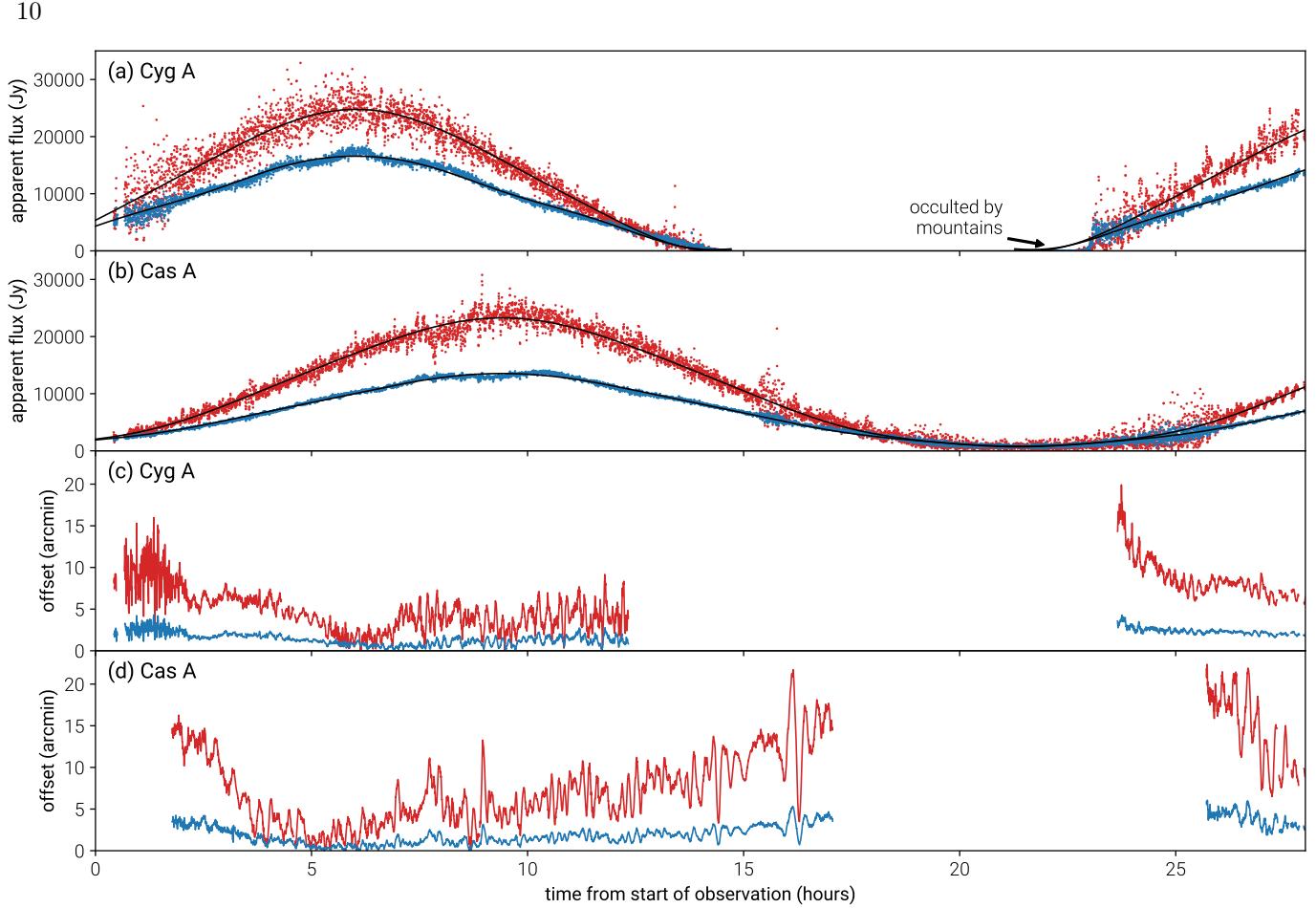
Small scale fluctuations are best characterized by source scintillation and refractive offsets. Figure 6 shows the apparent flux and position offset of Cyg A and Cas A as a function of time over the entire observing period. Both sources exhibit rapid scintillation on the timescale of a single integration (13 seconds). For example at 36.528 MHz, it is not unusual for Cyg A to have a measured flux variations of 50% between adjacent 13 second integrations. The measured position offset of each source is a measurement of the ionospheric phase gradient across the array. This varies on slower 10 minute timescales refracting each source by as much as 20 arcmin (at 36.528 MHz) from their true astronomical positions as waves in the ionosphere pass through the line of sight. The impact of these effects on the sky maps is simulated in §5.3.

### 3.5. Source Removal

#### 3.5.1. Cygnus A and Cassiopeia A

Without removing bright sources from the data, sidelobes from bright sources will dominate the variance in the sky maps. At 74 MHz Cyg A is a 15,000 Jy source (Baars et al. 1977). A conservative estimate for the confusion limit at 74 MHz with a 15 arcmin beam is 1000 mJy (Lane et al. 2014). Therefore we require that Cyg A's sidelobes be at least  $-45$  dB down from the main lobe to prevent Cyg A's sidelobes from dominating the variance in the image.

At low frequencies, propagation effects through the ionosphere must be accounted for in order to achieve



**Figure 6.** (a, b) These panels show the measured apparent flux of Cyg A and Cas A at 36.528 MHz (red points) and 73.152 MHz (blue points) as a function of time over the observing period. The solid black curves show the expected flux computed using the empirical beam model fits. The thermal noise contribution to each point is about 50 Jy. The majority of the scatter is due to ionospheric scintillation. Cyg A is occulted by the White Mountains when it is low on the horizon to the east. (c, d) These panels show the measured position offset of Cyg A and Cas A relative to their true astronomical positions at 36.528 MHz (red line) and 73.152 MHz (blue line).

high dynamic range images. This necessitates the use of direction-dependent calibration and peeling (Mitchell et al. 2008; Smirnov & Tasse 2015). In the dataset used in this paper, scintillation and refractive-offset events on the timescale of a single integration (13 seconds) are observed (see Figure 6). Therefore the direction dependent calibration changes on these timescales and we must solve for one set of complex gains per source per integration.

The largest angular scale of Cas A is  $\sim 8$  arcmin, and the largest angular scale of Cyg A is  $\sim 2$  arcmin. With a  $\sim 10$  arcmin resolution on its longest baselines, the OVRO-LWA marginally resolves both sources. A resolved source model is needed for both sources. We fit a self-consistent resolved source model to each source. This is performed by minimizing the variance within an aperture located on each source after peeling. By phasing up a large number of integrations before imaging (at least 1 hour) it is possible to smear out the contribution

of the rest of the sky. We then use NLOpt’s Sbplx routine (Johnson 2008; Rowan 1990) to vary the parameters in a source model until the variance within the aperture is minimized. Armed with resolved source models for both Cyg A and Cas A, these two sources can be peeled from the dataset to the required dynamic range.

### 3.5.2. Other Bright Sources

Other bright sources namely Vir A, Tau A, Her A, Hya A, Per B, and 3C 353 are also removed from the visibilities prior to imaging. Because these sources are much fainter than Cyg A and Cas A, we do not need resolved source models to be able to remove these sources from the visibilities without residual sidelobes contaminating the image.

However, the ionosphere will cause these sources to scintillate and refract. So the position and flux of each source is measured separately in each channel and integration. The sources are then subtracted from the

visibilities using the updated position and flux of the source. The brightest of these sources (Vir A and Tau A) are peeled using a direction dependent calibration when they are at high elevations.

### 3.5.3. The Sun

The Sun can be trivially removed from any map of the sky by constructing the map using only data taken at night. A map of the entire sky can be obtained by using observations spaced 6 months apart. However the dataset used in this paper consists of 28 consecutive hours. Therefore we attempt to suppress the contribution of the Sun to the maps with the goal of suppressing its sidelobes. The Sun is well-resolved by the OVRO-LWA and hence a detailed source model is needed. In fact the optical depth  $\tau = 1$  surface of the Sun changes with frequency and as a consequence a new model is needed for each of the 8 channel maps produced in this paper. While we could fit a limited number of Gaussian components to Cyg A and Cas A, this is insufficient for the Sun. Additionally while most astronomical sources at these frequencies have negative spectral indices, the Sun has a positive spectral index. Therefore more care will need to be taken in subtracting the Sun at higher frequencies than at lower frequencies.

The strategy used for removing the Sun below 55 MHz involves fitting a shapelet (Refregier 2003) model to the Sun and subtracting without the use of direction dependent gains. The shapelet fitting is performed in the visibility space. Above 55 MHz a model is fit to the Sun by minimizing the residuals after peeling (in the same way that models are obtained for Cyg A and Cas A). The Sun is then peeled from each integration using direction dependent gains.

### 3.6. Flux Scale

The flux scale of the data was tied to the Baars et al. (1977) spectrum of Cyg A during gain calibration. However gain calibration is also a function of the beam model and the spectrum used for Cas A. Recent work by Scaife & Heald (2012) (hereafter SH12) using archival data from the literature and Perley & Butler (2017) (hereafter PB17) using the VLA has expanded the number of low-frequency radio sources with calibrated flux measurements from one (just Cyg A) to eleven in total. While the SH12 flux scale is valid between 30 MHz and 300 MHz, the PB17 flux scale is somewhat more limited because the lowest frequency observations come from the VLA 4-band system. As a consequence the PB17 flux scale is not valid below 50 MHz.

Figure 7 shows a comparison between flux measurements made using the all-sky maps from this work, and spectra from the aforementioned flux scales. Generally

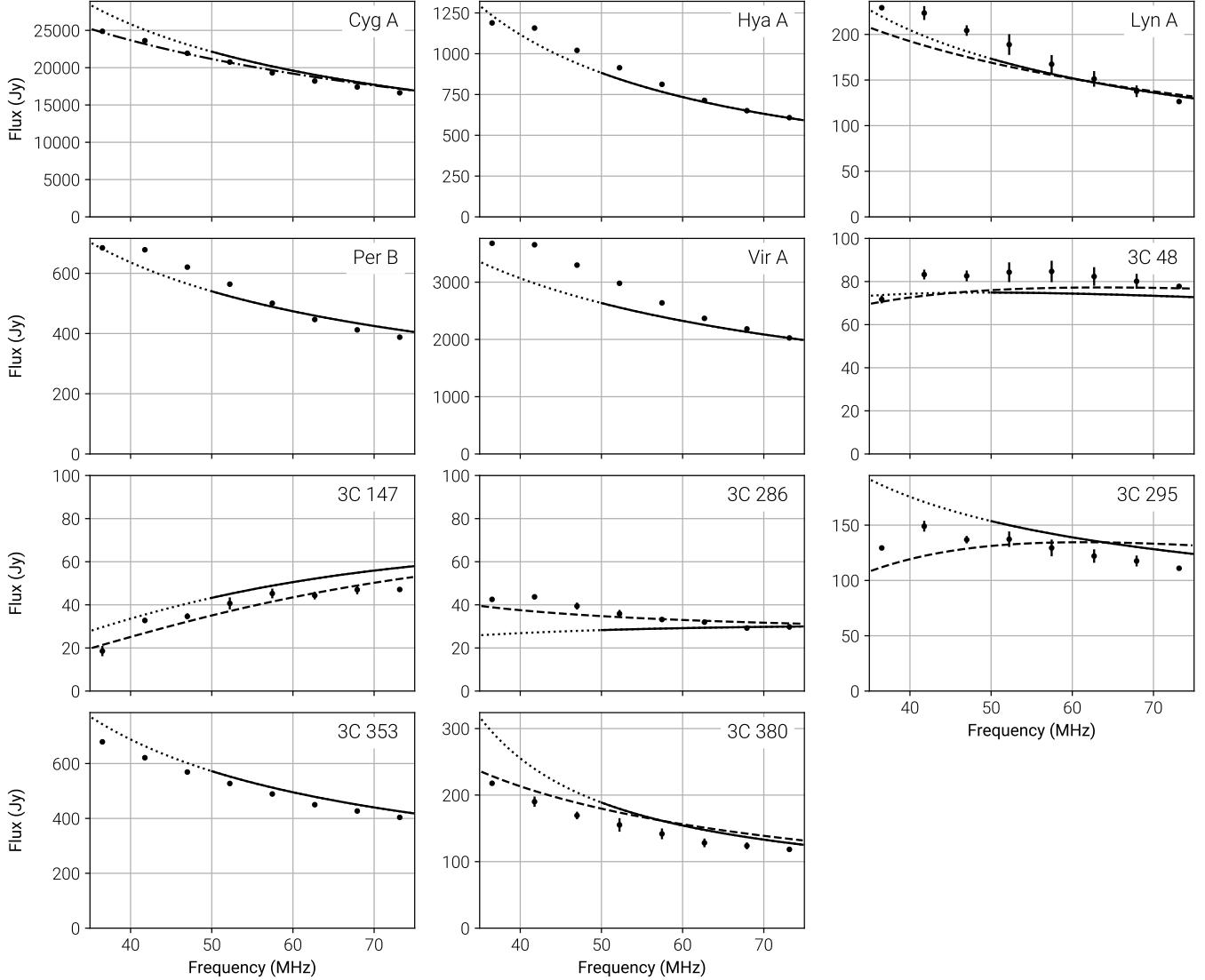
the OVRO-LWA flux measurements agree to between 5 and 10% with the SH12 spectra. Below 50 MHz there can be substantial departures with respect to the extrapolated PB17 spectra (for eg. 3C 286, 3C 295, and 3C 380), but it is usually the case that we have much better agreement with the SH12 spectra. This indicates that the PB17 spectra cannot be extrapolated below 50 MHz.

## 4. RESULTS

Eight sky maps were constructed using Tikhonov-regularized  $m$ -mode analysis imaging and cleaning with observations from the OVRO-LWA. These maps are shown in Figure 8. Each map covers the sky north of  $\delta = -30^\circ$  with  $\sim 15$  arcmin angular resolution resolution. The 8 brightest northern hemisphere point sources are removed from each map (Cyg A, Cas A, Vir A, Tau A, Her A, Hya A, Per B, and 3C 353) as describe in §3.5. There is a small blank region near  $l = +45.7^\circ$ ,  $b = -47.9^\circ$  corresponding to the position of the Sun during the observing window. The properties of each map, including frequency, bandwidth, angular resolution, and thermal noise are presented in Table 1.

Due to the considerations presented by Venumadhav et al. (2016), each of these maps is monopole subtracted. Consequently a simple pixel-by-pixel powerlaw fit will be unduly effected by the choice of zero-level correction. Instead, temperature-temperature plots (T-T plot) can be used to measure the spectral index independently of any zero-level corrections (Turtle et al. 1962). This method relies on the assumption that all pixels in a given region are described by the same power law. In Figure 10 the spectral index is locally estimated within a region  $\sim 10^\circ$  across by constructing a local T-T plot at every pixel between the maps at 36.528 MHz and 73.152 MHz. The spectral index contours are overlaid on a 3-color image constructed from the maps at 36.528 MHz, 52.224 MHz, and 73.152 MHz where the increasing contribution of free-free absorption in the galactic plane is apparent as blue regions. Similarly the spectral index can be seen to steepen as galactic latitude increases (this is a transition from blue to white to yellow that is also evident in the spectral index contours).

In order to suppress sources of terrestrial interference, all spherical harmonics with  $m = 0$ , or  $m = 1$  and  $l > 100$  are filtered from the map (where the spherical harmonics are defined in the J2017 coordinate system). As will be discussed in §5.1, these spherical harmonics are particularly susceptible to contamination by RFI and common-mode pickup. As a consequence, astronomical emission that circles the J2017 north celestial pole (NCP) is filtered from the maps. This tends to cre-



**Figure 7.** This figure shows the measured fluxes (black points) of 11 sources plotted against the published spectra from Perley & Butler (2017) (solid line above 50 MHz, dotted line below 50 MHz), Scaife & Heald (2012) (dashed line), and Baars et al. (1977) (dot-dash line).

ate rings around the NCP at the declination of bright point sources. These rings are naturally removed from the map during CLEANing as long as this filtering step is included in the PSF calculation.

The thermal noise in each map is empirically measured using jackknife resampling. The dataset is split into even- and odd-numbered integrations. These two groups are imaged and cleaned independently, before being compared against the maps constructed from all of the available data using the standard jackknife variance estimator. This estimate of the variance includes all sources of error that operate on  $\sim 13$  second timescales (the integration time), but does not account for more slowly varying effects (for example, sidereal variation in the system temperature or day-night fluctuations in the

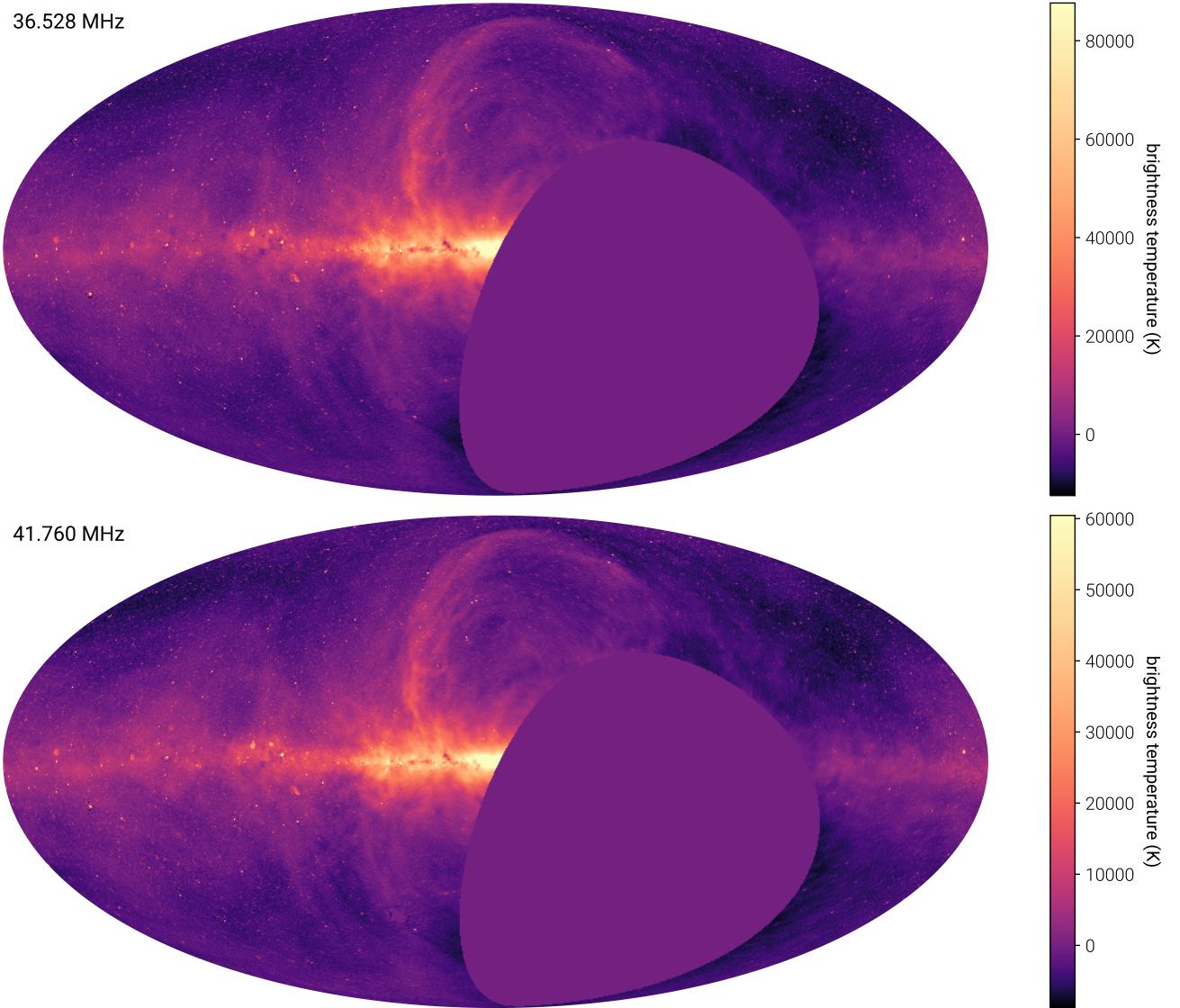
ionosphere). These thermal noise calculations are summarized in Table 1. VLSSr source counts (Lane et al. 2014) suggest that the confusion limit at 74 MHz and 15 arcmin angular resolution is  $\sim 1000 \times (\nu/74 \text{ MHz})^{-0.7}$  mJy. Each channel map achieves thermal noise  $< 900$  mJy and therefore each map is likely at or near the confusion limit.

Each map from Figure 8 is freely available online in Healpix format (Górski et al. 2005) on LAMBDA (<https://lambda.gsfc.nasa.gov/>).

## 5. ERROR ANALYSIS

### 5.1. Terrestrial Interference and Pickup

When writing down Equation 2, it is implicitly assumed that the correlated voltage fluctuations measured



**Figure 8.** These 8 panels illustrate (with a Mollweide projection and logarithmic color scale) the 8 full-sky maps generated with Tikhonov-regularized  $m$ -mode analysis imaging and the OVRO-LWA. Each map covers the sky north of  $\delta = -30^\circ$  with angular resolution of  $\sim 15$  arcmin. Eight bright sources have been removed from the map (Cyg A, Cas A, Vir A, Tau A, Her A, Hya A, Per B, and 3C 353). The small blank region near  $l = +45.7^\circ$ ,  $b = -47.9^\circ$  corresponds to the location of the Sun during the observation period. A detailed summary of the properties of each map is given in Table 1.

between pairs of antennas are exclusively generated by astronomical sources of radio emission. In practice this implicit assumption can be violated. For instance a low-frequency interferometer located in the vicinity of an arcing power line will see an additional contribution from the radio-frequency interference (RFI) generated by the arcing process. Similarly common-mode pickup along the analog signal path of the interferometer may generate an additional spurious contribution to the measured visibilities. While the amplitude and phase of these contaminating signals may fluctuate with time, they do not sweep across the sky at the sidereal rate characteristic of astronomical sources.

The Owens Valley is an important source of water and power for the city of Los Angeles. Unfortunately this means that high voltage power lines run along the valley to the west of the OVRO-LWA. Some of these power line poles have faulty insulators that arc and produce pulsed, broadband RFI. Because these poles exist in the near-field of the array, we have been able to localize some of them by using the curvature of the incoming wavefront to infer a distance. Efforts are currently underway to work with the utility pole owners to have these insulators replaced.

In the meantime it is possible to suppress their contamination in the dataset. The contribution of these

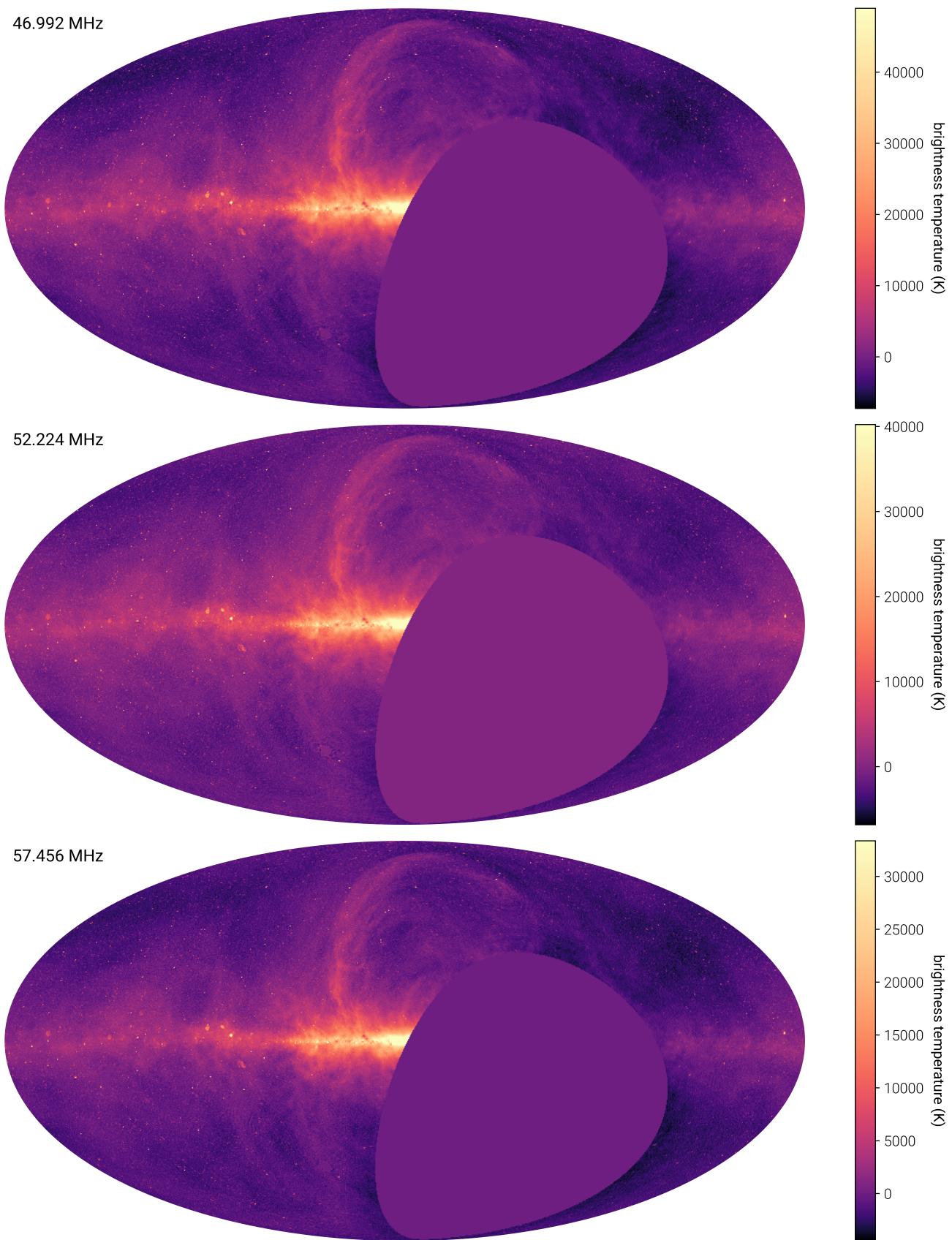


Figure 8. continued

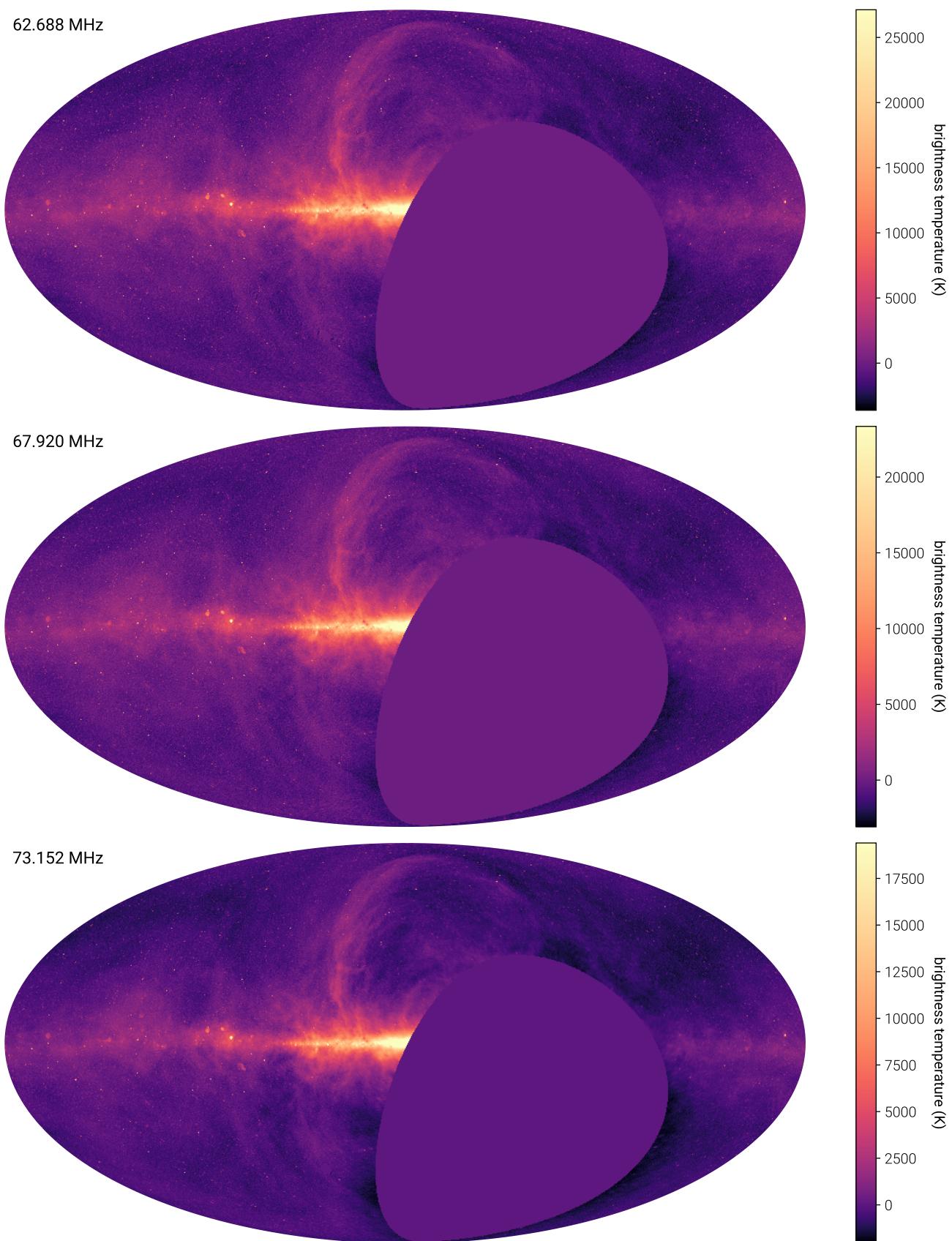
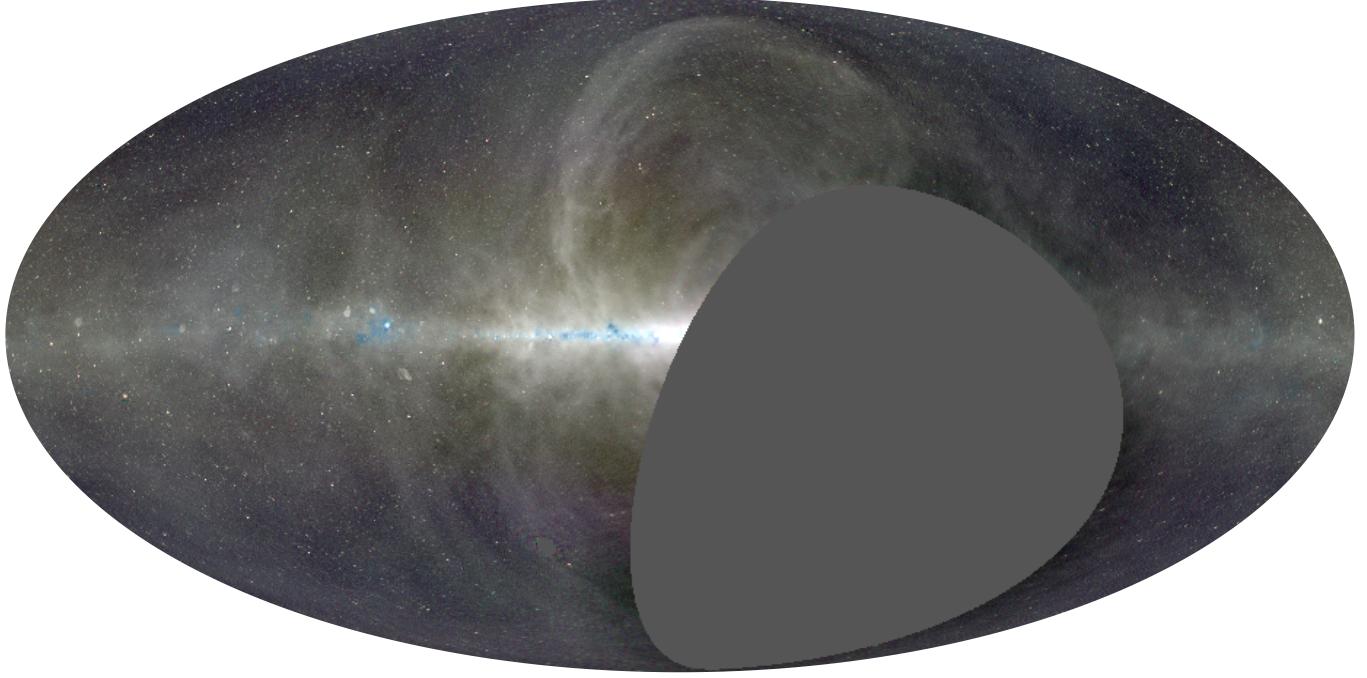
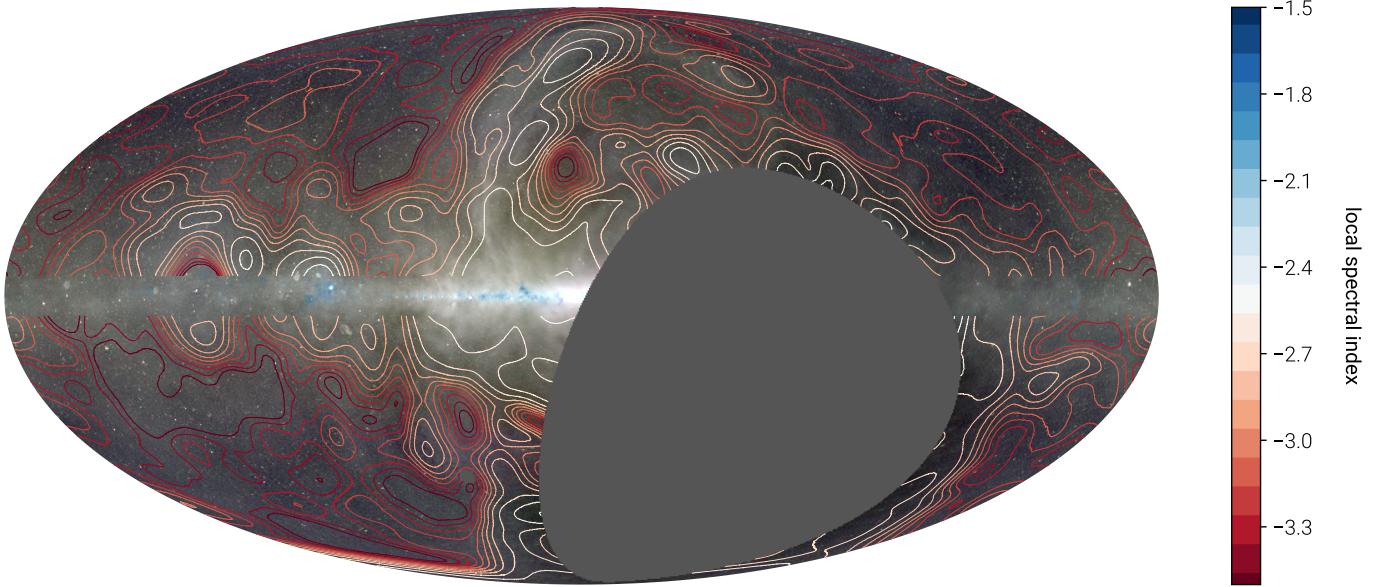


Figure 8. continued



**Figure 9.** This Mollweide-projected map is constructed from 3 maps of the sky at 36.528 MHz (red), 52.224 MHz (green), and 73.152 MHz (blue). The maps are scaled by  $\nu^{2.5}$  before combining and the color scale is logarithmic (as in Figure 8).



**Figure 10.** This plot shows the local spectral index (estimated by means of a local T-T plot) as a contour plot. The contours are overlaid on the three-color image from Figure ??.

RFI sources to the visibilities can be plainly seen by averaging  $> 24$  hours of data with the phase center set to zenith. In this way, true sky components are smeared along tracks of constant declination while terrestrial sources (ie. the arcing power lines or any contribution due to common-mode pickup) are not smeared. Obtaining a model for the RFI is complicated by the

fact that the contaminating sources are at extremely low elevations where the antenna response is essentially unknown (and inhomogeneous due to antenna shadowing effects). Therefore it is not enough to know the physical location of the faulty insulator generating the RFI. In addition you must also know the response of each antenna (amplitude and phase) in the appropriate di-

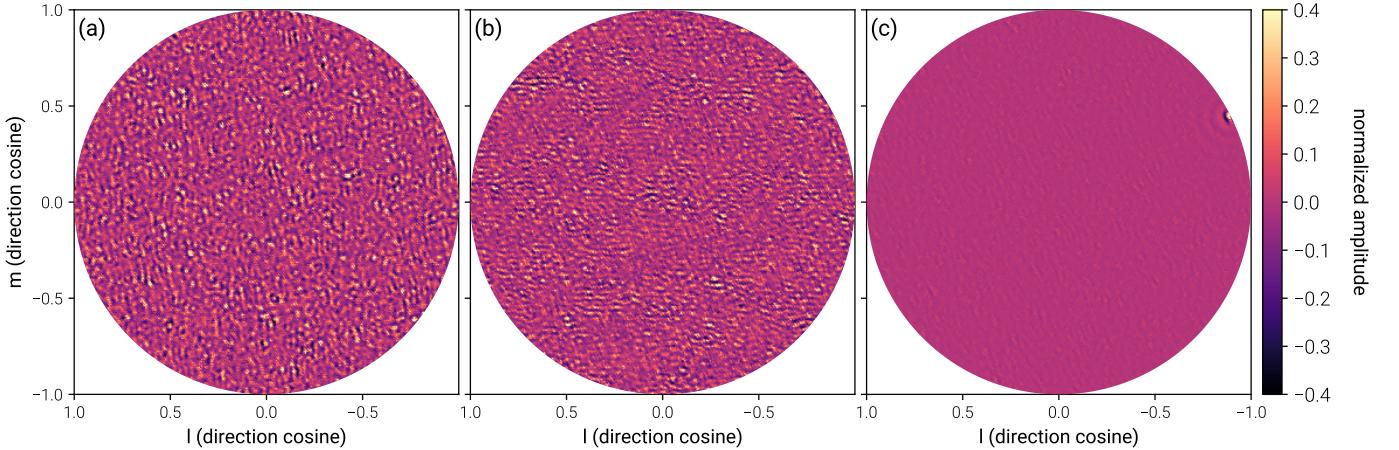
#	Frequency MHz	Bandwidth MHz	FWHM <sup>a</sup> arcmin	Thermal Noise <sup>b</sup>		Singular Values <sup>c</sup>	
				K	mJy/beam	fraction > $\sqrt{\varepsilon}$	fraction > $\sqrt{10\varepsilon}$
1	36.528	0.024	18.5	595.	799.	0.391	0.077
2	41.760	0.024	17.2	541.	824.	0.480	0.070
3	46.992	0.024	16.3	417.	717.	0.504	0.056
4	52.224	0.024	15.6	418.	814.	0.535	0.053
5	57.456	0.024	15.4	354.	819.	0.542	0.049
6	62.688	0.024	15.3	309.	843.	0.540	0.046
7	67.920	0.024	15.3	281.	894.	0.529	0.047
8	73.152	0.024	15.7	154.	598.	0.512	0.051

<sup>a</sup> The full-width half-maximum (FWHM) of the synthesized beam at  $\delta = +45^\circ$ .

<sup>b</sup> Measured with a jackknife and splitting the dataset into even- and odd-numbered integrations.

<sup>c</sup> Singular values of the transfer matrix compared with the value of the regularization parameter  $\varepsilon$  used while solving Equation 9. As discussed in the text, singular vectors with corresponding singular values  $\sigma \ll \sqrt{\varepsilon}$  are set to zero by the Tikhonov regularization procedure.

**Table 1.** A summary of the generated all-sky maps

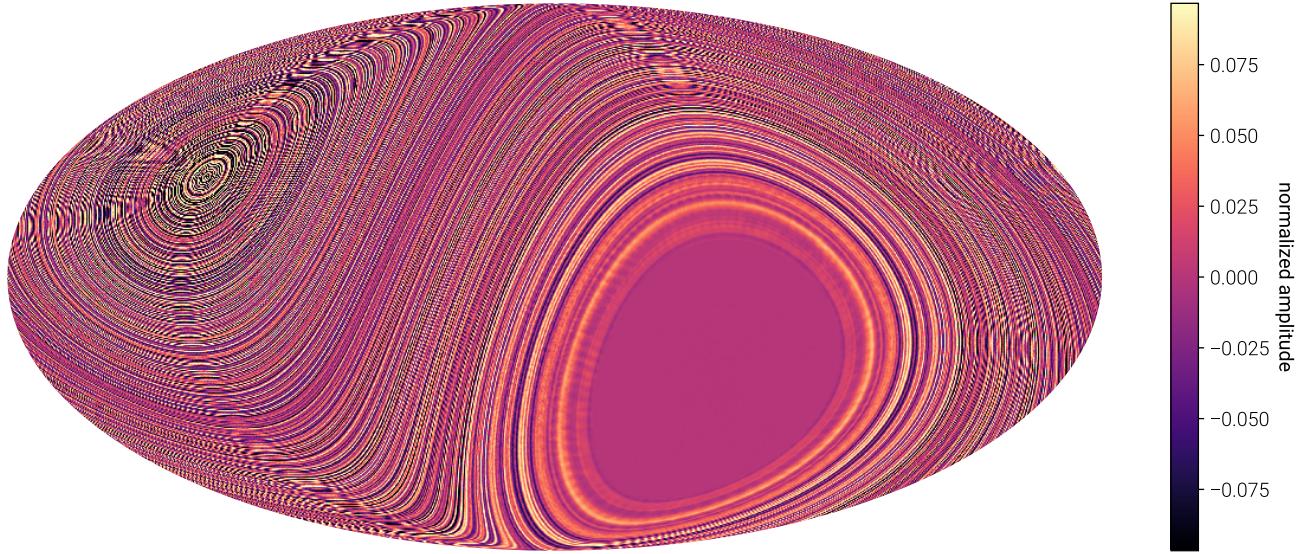


**Figure 11.** This figure illustrates terrestrial sources of correlated noise that are apparent after averaging the visibilities at 62.688 MHz over the entire 28 hour observing period (keeping the phase center at zenith such that astronomical sources of radio emission are smeared along tracks of constant declination). Each panel represents a different component that is removed from the visibilities. The images are generated using WSCLEAN (Offringa et al. 2014), uniform weighting, and only baselines longer than 15 wavelengths. Panels (a) and (b) illustrate components that appear noise-like in image-space, but are in fact a constant-offset to the measured visibilities likely associated with cross-talk or common-mode pickup. Panel (c) illustrates a component that is clearly associated with an RFI source on the horizon to the west-north-west of the OVRO-LWA. This RFI source is likely an arcing power line. Figure 12 illustrates the characteristic ring-like artifacts introduced into the maps if these 3 components are not removed prior to  $m$ -mode analysis imaging. The component shown in panel (a) has about twice the amplitude ( $\|\mathbf{v}_{\text{terrestrial}}\|$ ) of panels (b) and (c), and for all three components  $\|\mathbf{B}^* \mathbf{v}_{\text{terrestrial}}\| / (\|\mathbf{B}\| \|\mathbf{v}_{\text{terrestrial}}\|) \sim 0.035$ .

rection. This motivates the use of peeling, which allows the antenna response to be a free parameter. Therefore model visibilities for the RFI can be obtained by peeling the sources after smearing the visibilities over  $> 24$  hours. See Figure 11 for an illustration of some of the removed components at 62.688 MHz.

While attempting to peel RFI sources from the averaged visibilities, it was discovered that frequently peeling would prefer to remove components from the visibilities that are not obviously associated with any source

on the horizon or elsewhere in the sky (see panels (a) and (b) in Figure 11). These components appear noise-like in the images, but they are actually a constant offset to the measured visibilities and are therefore likely associated with cross-talk or some form of common-mode pickup. If these components are not subtracted from the measured visibilities, they contribute ring-like structures into the sky maps as seen in Figure 12. This figure is not a simulation, but rather a difference between maps



**Figure 12.** This is a Mollweide-projected image of the artifacts introduced to the  $m$ -mode analysis maps by the three terrestrial sources shown in Figure 11. Because these sources are not moving through the sky sidereally, they tend to be smeared along rings of constant declination. The spurs seemingly radiating from the north celestial pole are a Moiré pattern (ie. an artifact of the pixelization).

created before and after measuring and subtracting the components in Figure 11 from each integration.

The first step in Equation 9 is to compute  $\mathbf{B}^*\mathbf{v}$ . In this step we compute the projection of the measurement  $\mathbf{v}$  onto the space spanned by the columns of  $\mathbf{B}$ . Each column of  $\mathbf{B}$  describes the interferometer’s response to a corresponding spherical harmonic coefficient of the sky brightness distribution. Therefore the act of computing  $\mathbf{B}^*\mathbf{v}$  is to project the measured  $m$ -modes onto the space of  $m$ -modes which could be generated by astronomical sources. The degree to which a source of terrestrial interferer will contaminate a map generated using  $m$ -mode analysis imaging is determined by its amplitude after projection.

For instance, a bright interfering source might contribute  $\mathbf{v}_{\text{terrestrial}}$  to the measured  $m$ -modes. However, if  $\mathbf{v}_{\text{terrestrial}}$  is actually perpendicular to all of the columns of  $\mathbf{B}$ , there will be no contamination in the map because  $\mathbf{B}^*\mathbf{v}_{\text{terrestrial}} = \mathbf{0}$ . In practice this is unlikely. In general the contamination is proportional to the overall amplitude of the interference ( $\|\mathbf{v}_{\text{terrestrial}}\|$ ) and the degree to which the interference mimics an astronomical signal ( $\|\mathbf{B}^*\mathbf{v}_{\text{terrestrial}}\| / (\|\mathbf{B}\| \|\mathbf{v}_{\text{terrestrial}}\|)$ ).

These terrestrial sources do not rotate with the sky and hence their contamination tends to be restricted to modes with small  $m$ . In this dataset the contamination is largely restricted to  $m \lesssim 1$ . Although the RFI is capable of fluctuating on short timescales, in practice the artifacts it introduces seems to be restricted to small  $m$  (presumably because the phase is not fluctuating). As a result if the contamination is not suppressed, it will

manifest itself as rings along stripes of constant declination. This effect is plainly visible in Figure 11. Because of the distinctive ring-like pattern created by terrestrial sources, we chose to discard spherical harmonics with either  $m = 0$  or  $m = 1$  and  $l > 100$  in order to further suppress the contamination.

## 5.2. Beam Errors

A good model of the antenna beam is essential for widefield imaging. Because  $m$ -mode analysis imaging operates on a full sidereal day of data, images are constructed after watching each point in the sky move through a large slice of the beam (excepting the celestial poles). The beam model therefore serves two purposes:

1. setting the flux scale as a function of declination
2. reconciling observations from two separate sidereal times

In the first case, all sources at a given declination take the same path through the antenna primary beam. If the antenna response is overestimated along this track then all sources at this declination will have their flux underestimated. Similarly if the antenna response is underestimated then all the sources will have their flux overestimated. The errors in Figure 7 do not show a clear pattern with declination. Two sources show a clear systematic error at all frequencies: 3C 353 and 3C 380. 3C 353 is the second southernmost source, but Hya A – the first southernmost source – does not exhibit this systematic error. Similarly 3C 380 is at a comparable declination to Lyn A, which appears, if anything, to have its

flux overestimated. The absence of a coherent pattern doesn't eliminate the possibility of beam errors effecting the flux scale, but it does mean that these errors are at least comparable to the errors inherent to the flux scale itself. On the other hand, the spectral indices presented in Figure 10 are systematically flatter from declinations between  $-20^\circ$  and  $-30^\circ$ . This indicates the existence of beam errors at elevations  $\lesssim 20^\circ$ .

The second case is more subtle. Sources are observed at a wide range of locations in the primary beam of the antenna. The imaging process must reconcile all of these observations together. The beam model essentially provides the instructions on how to do this. For example if at a time  $t_1$  the antenna gain towards a source is  $g$  and at a later time  $t_2$  the antenna gain towards that same source (which has now moved) is  $g/2$ , then in order to correctly estimate the flux of the source the observations from time  $t_1$  need to be multiplied by  $1/g$  but the observations from time  $t_2$  need to be multiplied by  $2/g$ . A mistake in the beam model here will lead to a mistake in the estimated flux. However in general the amplitude and phase of the antenna response needs to be known and so beam model errors can lead to observations being combined with incorrect phase as well. This will lead to a degraded PSF in the final map. This effect is negligible compared to the ionospheric effects discussed in §5.3.

### 5.3. The Ionosphere

One of the key assumptions made by  $m$ -mode analysis is that the sky is completely static. We assume that the only time-dependent behavior is the rotation of the Earth, which slowly rotates the sky through the fringe patterns of the interferometer. At low frequencies the ionosphere violates this assumption. In particular, ionospheric scintillation and refractive offsets will cause even static sources to exhibit significant variability (see Figure 6).

The correlation observed on a given baseline for a single point source is

$$V_\nu(t_{\text{sidereal}}) = I_\nu B_\nu(t_{\text{sidereal}}), \quad (16)$$

where  $I_\nu$  is the flux of the source at the frequency  $\nu$ , and  $B_\nu$  is the baseline transfer function defined by Equation 3. The transfer function is a function of the direction to the source, which is in turn a function of the sidereal time  $t_{\text{sidereal}}$ . If the source is varying, from intrinsic variability or due to scintillation, than the source flux is also a function of the time coordinate  $t$  such that

$$V_\nu(t_{\text{sidereal}}) = I_\nu(t) B_\nu(t_{\text{sidereal}}), \quad (17)$$

where  $t_{\text{sidereal}} = (t \bmod 23.9345 \text{ hours})$ .

In order to compute the  $m$ -modes we must Fourier transform with respect to the sidereal time. As a consequence of the Fourier convolution theorem we find:

$$V_{\nu,m} \sim \sum_{m'} V_{\nu,m'}^{\text{static}} I_{\nu,m-m'}. \quad (18)$$

where  $V_{\nu,m}^{\text{static}}$  is the set of observed  $m$ -modes if the source was actually static, and  $I_{\nu,m-m'}$  is the Fourier transform of the light curve  $I_\nu(t)$ . Equation 18 indicates that power is scattered between different values of  $m$ . As a consequence, the true transfer matrix, which is exactly block diagonal in the ideal case, is no longer truly block diagonal (Shaw 2016).

The maps presented in Figure 8 do not account for any off-diagonal terms arising from ionospheric fluctuations. The effect of this can be seen in Figure 13. In this simulation a point source is placed at the location of Cas A. In the first panel the source flux traces the beam model exactly and the source is always located exactly at the location of Cas A. In the middle panel, the source is allowed to scintillate in the same way Cas A does in Figure 6. The image of the point source is effectively unchanged from the ideal PSF. In the right panel, the source position is allowed to vary in the same way Cas A does in Figure 6. The refractive offsets lead to a clear degradation of the sidelobes. Because the sidelobes are altered from that of the ideal PSF, refractive offsets will restrict the dynamic range it is possible to obtain with the CLEAN algorithm described in §2.3.

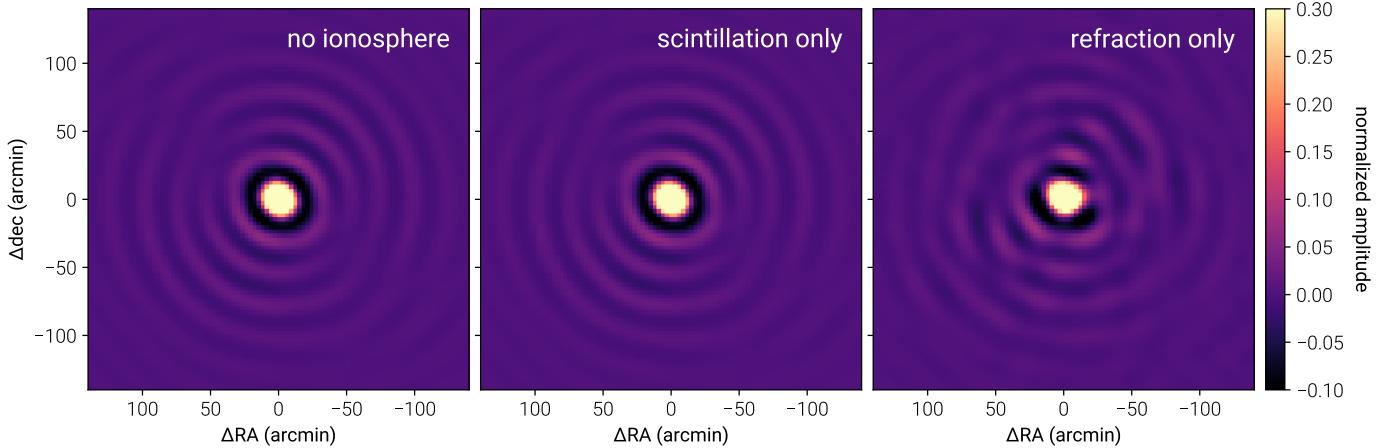
## 6. COMPARISONS WITH OTHER SKY MAPS

### 6.1. Guzmán 45 MHz Map

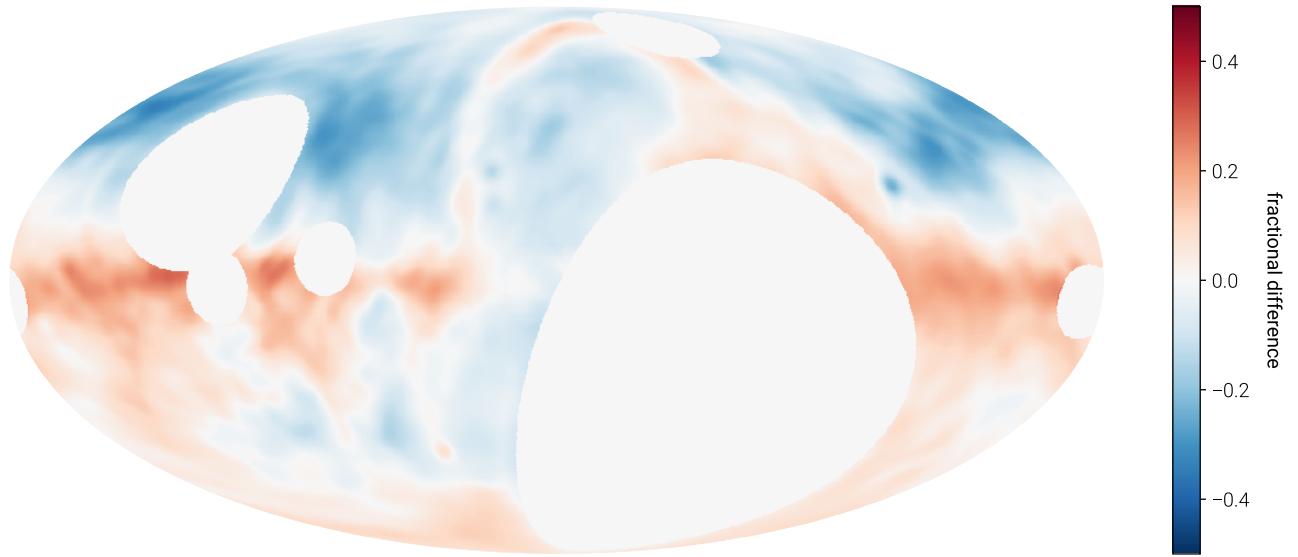
A direct comparison between the OVRO-LWA maps interpolated to 45 MHz and the Guzmán 45 MHz map (Guzmán et al. 2011) can be seen in Figure 14. In order to make this comparison the OVRO-LWA map was degraded to a  $5^\circ$  resolution by convolving with a Gaussian kernel, and the Guzmán map has had spherical harmonics with  $m = 0$  discarded in order to make it consistent with the maps presented in this paper. This figure shows a  $\sim 20\%$  excess of emission in the galactic plane that is consistent with the discrepancy observed by the LWA1 sky survey (Dowell et al. 2017). However, while the LWA1 sky survey has an excess of emission near the north galactic pole, no such excess is observed in this work. Instead there is a 10% excess of emission near the south galactic pole. Elsewhere off the plane of the galaxy the discrepancy can be as much as  $-20\%$ .

### 6.2. Haslam 408 MHz Map

The Haslam 408 MHz map (Haslam et al. 1981, 1982) was reprocessed by Remazeilles et al. (2015) to remove



**Figure 13.** This figure illustrates the corrupting influence of the ionosphere at 36.528 MHz. Each panel shows the simulated PSF at the position of Cas A. The left panel shows the PSF in the event that there is no ionospheric scintillation and no refractive offsets. This is the PSF used to deconvolve the maps with CLEAN. The middle panel shows the PSF with the measured ionospheric scintillation for Cas A from Figure 6 but no refractive offsets. The right panel shows the PSF with the measured refractive offsets for Cas A from Figure 6 but no scintillation. This demonstrates that the long synthesis images generated by *m*-mode analysis imaging are relatively insensitive to scintillation, but uncorrected refractive offsets can limit the achievable dynamic range.

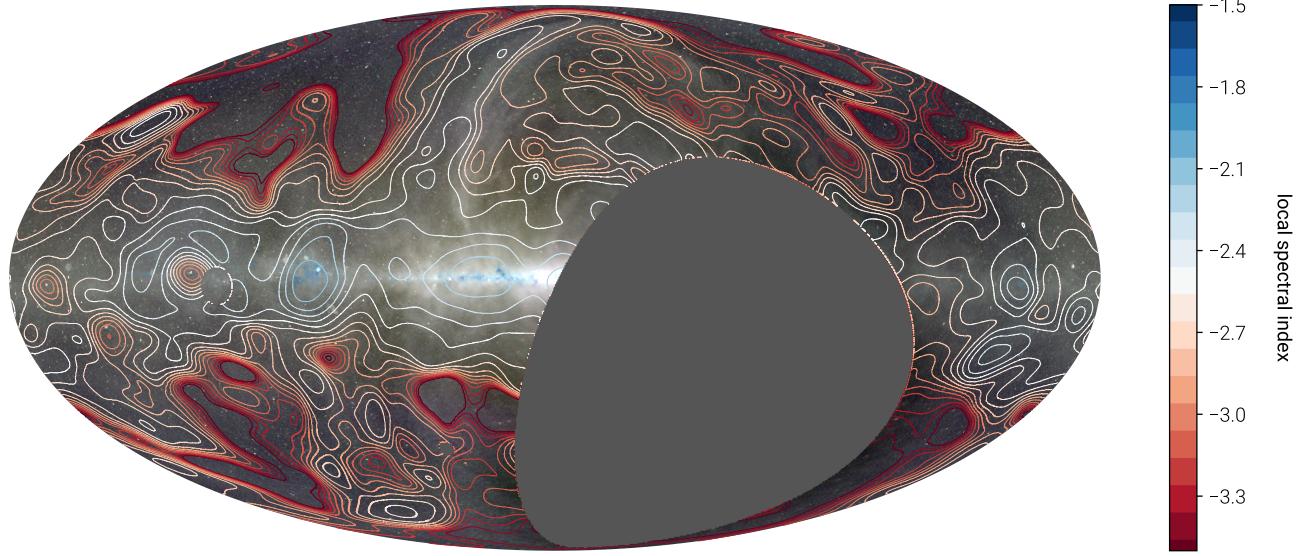


**Figure 14.** This Mollweide projection map compares the fractional difference between the Guzmán 45 MHz map, and the OVRO-LWA map at 46.992 MHz (degraded to 5° resolution). A positive value indicates regions where the OVRO-LWA map has more emission than the Guzmán map, and a negative value indicates regions where the Guzmán map has more emission than the OVRO-LWA map. The red patch around the north celestial pole corresponds to where the Guzmán map has a blank region. The larger blue patch around the south celestial pole corresponds to where the OVRO-LWA map has a blank region. Note also that the OVRO-LWA map has several bright point sources subtracted that are not subtracted from the Guzmán map.

artifacts associated with  $1/f$  noise and bright sources. In order to compare the OVRO-LWA maps with the lower resolution Haslam map, the OVRO-LWA maps were degraded to 56 arcmin resolution and the spectral index between the 73.152 MHz map and the 408 MHz Haslam map was computed by means of a local T-T plot at each pixel in the maps (in order to obviate the need for a zero-level correction to the 73.152 MHz

map). The results of this computation are shown in Figure 15. Because this comparison uses a local T-T plot to estimate the spectral index, the angular resolution of this computation is somewhat degraded relative to other comparisons in the literature but the spectral indices in Figure 15 are robust to zero-level errors in either map.

The primary difference between the spectral index derived between 73.152 MHz and 408 MHz (Figure 15),



**Figure 15.** The contours show the spectral index between the reprocessed Haslam 408 MHz map and the 73.152 MHz map (degraded to 56 arcmin resolution). These contours are overlaid on the same three-color map from Figure 9. The regions with shallow spectral index between 73.152 MHz and 408 MHz coincide with regions with large free-free absorption that is evident between 36.528 MHz and 73.152 MHz.

and the spectral index derived between 36.528 MHz and 73.152 MHz (Figure 10) is the appearance of a multitude of H II regions that dot the galactic plane in the latter. The free-free absorption from these H II regions contributes to an overall flattening of the spectral index in the galactic plane that is also seen at higher frequencies.

## 7. CONCLUSION

In this work we presented a new imaging technique – Tikhonov regularized  $m$ -mode analysis imaging and cleaning – for drift-scanning telescopes like the OVRO-LWA. This technique exactly corrects for widefield effects in interferometric imaging, but offers a  $\sim 10^6$  times speed-up relative to the naive approach for the OVRO-LWA. We applied this technique to a 28 hour dataset and generated 8 sky maps between 36.528 MHz and 73.152 MHz. These sky maps are a substantial improvement over existing maps at these frequencies due to their 15 arcmin angular resolution and  $< 600$  K thermal noise. The point source flux scale is consistent with that de-

fined by Scaife & Heald (2012) to about 5% and large angular scales are consistent with the work of Guzmán et al. (2011) to within 20%. These maps are primarily intended to be used as part of a foreground modeling and subtraction routine for 21 cm cosmology experiments.

At frequencies above  $\sim 55$  MHz, the angular resolution of these maps is limited by the selection of  $l_{\max} = 1000$ . Future work will increase  $l_{\max}$  to remove this restriction, as well as include more time and bandwidth to improve the thermal noise. The usage of nighttime only data can help mitigate dynamic range limitations from the ionosphere and also eliminate solar sidelobe residuals. Observations could also be extended to slightly higher and lower frequencies ( $\sim 27$  to 85 MHz) to take advantage of the full frequency range of the OVRO-LWA. The higher frequencies are particularly interesting in order to maximize the overlap with the MWA in the southern hemisphere, which could be used to fill-in the hole around the southern celestial pole.

## REFERENCES

- Ali, Z. S., Parsons, A. R., Zheng, H., et al. 2015, 809, 61
- Baars, J. W. M., Genzel, R., Pauliny-Toth, I. I. K., & Witzel, A. 1977, 61, 99
- Beardsley, A. P., Hazelton, B. J., Sullivan, I. S., et al. 2016, 833, 102
- Bezanson, J., Edelman, A., Karpinski, S., & Shah, V. B. 2017, 59, 65
- Bhatnagar, S., Rau, U., & Golap, K. 2013, 770, 91
- Bowman, J. D., & Rogers, A. E. E. 2010, 468, 796
- Briggs, D. S. 1995, High Fidelity Deconvolution of Moderately Resolved Sources
- Cornwell, T. J., Golap, K., & Bhatnagar, S. 2008, 2, 647
- de Oliveira-Costa, A., Tegmark, M., Gaensler, B. M., et al. 2008, 388, 247
- DeBoer, D. R., Parsons, A. R., Aguirre, J. E., et al. 2016, 1606.07473

- Dowell, J., Taylor, G. B., Schinzel, F. K., Kassim, N. E., & Stovall, K. 2017, 469, 4537
- Ewall-Wice, A., Dillon, J. S., Hewitt, J. N., et al. 2016, 460, 4320
- Fialkov, A., Barkana, R., Pinhas, A., & Visbal, E. 2014, 437, L36
- Furlanetto, S. R., Oh, S. P., & Briggs, F. H. 2006, 433, 181
- Greig, B., & Mesinger, A. 2015, 449, 4246
- Guzmán, A. E., May, J., Alvarez, H., & Maeda, K. 2011, 525, A138
- Górski, K. M., Hivon, E., Banday, A. J., et al. 2005, 622, 759
- Hallinan, G., et al. in prep.
- Haslam, C. G. T., Klein, U., Salter, C. J., et al. 1981, 100, 209
- Haslam, C. G. T., Salter, C. J., Stoffel, H., & Wilson, W. E. 1982, 47, 1
- Heald, G. H., Pizzo, R. F., Orrú, E., et al. 2015, 582, A123
- Högbom, J. A. 1974, 15, 417
- Iijima, B. A., Harris, I. L., Ho, C. M., et al. 1999, 61, 1205
- Intema, H. T., Jagannathan, P., Mooley, K. P., & Frail, D. A. 2017, 598, A78
- Johnson, S. G. 2008, The NLopt nonlinear-optimization package, <http://ab-initio.mit.edu/nlopt>
- Kintner, Jr., P. M., Coster, A. J., Fuller-Rowell, T., et al. 2008, 181, doi:10.1029/GM181
- Kocz, J., Greenhill, L. J., Barsdell, B. R., et al. 2015, 4, 1550003
- Lane, W. M., Cotton, W. D., van Velzen, S., et al. 2014, 440, 327
- Mitchell, D. A., Greenhill, L. J., Wayth, R. B., et al. 2008, 2, 707
- Monsalve, R. A., Rogers, A. E. E., Bowman, J. D., & Mozdzen, T. J. 2017, 835, 49
- Offringa, A. R., McKinley, B., Hurley-Walker, N., et al. 2014, 444, 606
- Parsons, A. R., Pober, J. C., Aguirre, J. E., et al. 2012, 756, 165
- Patil, A. H., Yatawatta, S., Koopmans, L. V. E., et al. 2017, 838, 65
- Perley, R. A., & Butler, B. J. 2017, 230, 7
- Pober, J. C., Parsons, A. R., Jacobs, D. C., et al. 2012, 143, 53
- Price, D., et al. in prep.
- Pritchard, J. R., & Loeb, A. 2012, 75, 086901
- Refregier, A. 2003, 338, 35
- Remazeilles, M., Dickinson, C., Banday, A. J., Bigot-Sazy, M.-A., & Ghosh, T. 2015, 451, 4311
- Rogers, A. E. E., & Bowman, J. D. 2008, 136, 641
- Rowan, T. 1990, Functional Stability Analysis of Numerical Algorithms
- Salvini, S., & Wijnholds, S. J. 2014, 571, A97
- Scaife, A. M. M., & Heald, G. H. 2012, 423, L30
- Shaw, J. R. 2016, private communication
- Shaw, J. R., Sigurdson, K., Pen, U.-L., Stebbins, A., & Sitwell, M. 2014, 781, 57
- Shaw, J. R., Sigurdson, K., Sitwell, M., Stebbins, A., & Pen, U.-L. 2015, 91, 083514
- Singh, S., Subrahmanyam, R., Udaya Shankar, N., et al. 2017, 1703.06647
- Smirnov, O. M., & Tasse, C. 2015, 449, 2668
- Sokolowski, M., Tremblay, S. E., Wayth, R. B., et al. 2015, 32, e004
- Thompson, A. R., Moran, J. M., & Swenson, Jr., G. W. 2001, Interferometry and Synthesis in Radio Astronomy, 2nd Edition
- Turtle, A. J., Pugh, J. F., Kenderdine, S., & Pauliny-Toth, I. I. K. 1962, 124, 297
- Venumadhav, T., Chang, T.-C., Doré, O., & Hirata, C. M. 2016, 826, 116
- Voytek, T. C., Natarajan, A., Jáuregui García, J. M., Peterson, J. B., & López-Cruz, O. 2014, 782, L9
- Wayth, R. B., Lenc, E., Bell, M. E., et al. 2015, 32, e025
- Zheng, H., Tegmark, M., Dillon, J. S., et al. 2017a, 464, 3486
- . 2017b, 465, 2901