

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

MICHAEL W. EASTWOOD, GREGG HALLINAN, TWILIGHT SPARKLE, RAINBOW DASH

## ABSTRACT

I present all-sky maps of the sky from the Owens Valley Long Wavelength Array. These maps are created from the application of *m*-mode analysis.

## 1. INTRODUCTION

Studies of the cosmic microwave background (CMB) have given us an unprecedented understanding of the universe during the epoch of recombination (Hinshaw et al. 2013; Planck Collaboration et al. 2014, 2016b). Before arriving at the Earth, the light from the CMB propagates through the intergalactic medium (IGM), intervening galaxies and galaxy clusters, and the Milky Way’s interstellar medium (ISM). As photons from the CMB propagate through the intervening material scattering, absorption, or additional emission can perturb the thermal CMB spectrum. In order to measure the pristine CMB, these foreground components (eg. synchrotron emission, free-free emission, thermal dust emission) are a nuisance and careful modeling and removal of these components is essential (Planck Collaboration et al. 2016a,c).

However in many cases the study of CMB foregrounds is scientifically interesting in its own right. Measurements of dust-correlated anomalous microwave emission (Leitch et al. 1997) led to the discovery of spinning dust emission (Draine & Lazarian 1998). Thomson scattering in the IGM attenuates the CMB spectrum, but the optical depth implies a mean redshift to the Epoch of Reionization (EoR) of  $8.8^{+1.7}_{-1.4}$  (Planck Collaboration et al. 2016b). The thermal Sunyaev-Zel’dovich effect (Sunyaev & Zel’dovich 1970, 1972) has been used to discover new galaxy clusters (Vanderlinde et al. 2010) and measure the distance to galaxy clusters when combined with X-ray measurements (Bonamente et al. 2006).

**TODO: kinetic SZ effect puts an upper limit on the duration of reionization**

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 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 ve-

locity of the gas.

**TODO: discuss 21cm ionizing efficiency, photon mean free path, minimum halo mass TODO: discuss what can be learned from global experiments**

The spatial power spectrum of this perturbation is expected 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  $z \sim 7$  **TODO: citation**.

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, and power spectrum experiments that incorporate angular information to attempt to measure the spatial power spectrum of cosmological 21 cm perturbations. Ongoing global experiments include EDGES (Bowman & Rogers 2010; Monsalve et al. 2017), LEDA **TODO: cite Price et al. 2017**, 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).

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 at least 4 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 (after convolving with the instrumental response) is spectrally smooth. However the low-frequency radio sky is composed of several components: galactic synchrotron emission, free-free emission and ab-

sorption, supernova remnants, radio galaxies, and a confusing background of point sources. Ideally a foreground removal strategy should be informed by the measured spatial and frequency structure of all foreground components. However, 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 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 higher frequencies where the majority of the input maps are from. Below 408 MHz, the interpolation is largely driven by the Haslam 408 MHz map (Haslam et al. 1981, 1982). At lower frequencies, free-free absorption and synchrotron self-absorption becomes increasingly important and hence spectral indices derived at higher frequencies need corrections at lower frequencies.

Guzmán et al. (2011)

TODO: TGSS, MSSS, GLEAM

## 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  $V_{ij,\nu}$  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

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

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

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 base-

lines  $\vec{b}_{ij}$  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, mosaicing, and multi-scale deconvolution. 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>1</sup>. Therefore it is a very natural operation to compute the Fourier transform of the visibilities with respect to sidereal time  $\phi \in [0, 2\pi)$ .

$$V_{m,\nu}^{ij} = \int_0^{2\pi} V_\nu^{ij}(\phi) \exp\left(-im\phi\right) d\phi \quad (3)$$

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.

TODO: introduce the baseline transfer function because I will want to use it in later sections

It can be shown that there is a discrete linear relationship between the  $m$ -modes  $V_{m,\nu}^{ij}$  and the spherical

<sup>1</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.

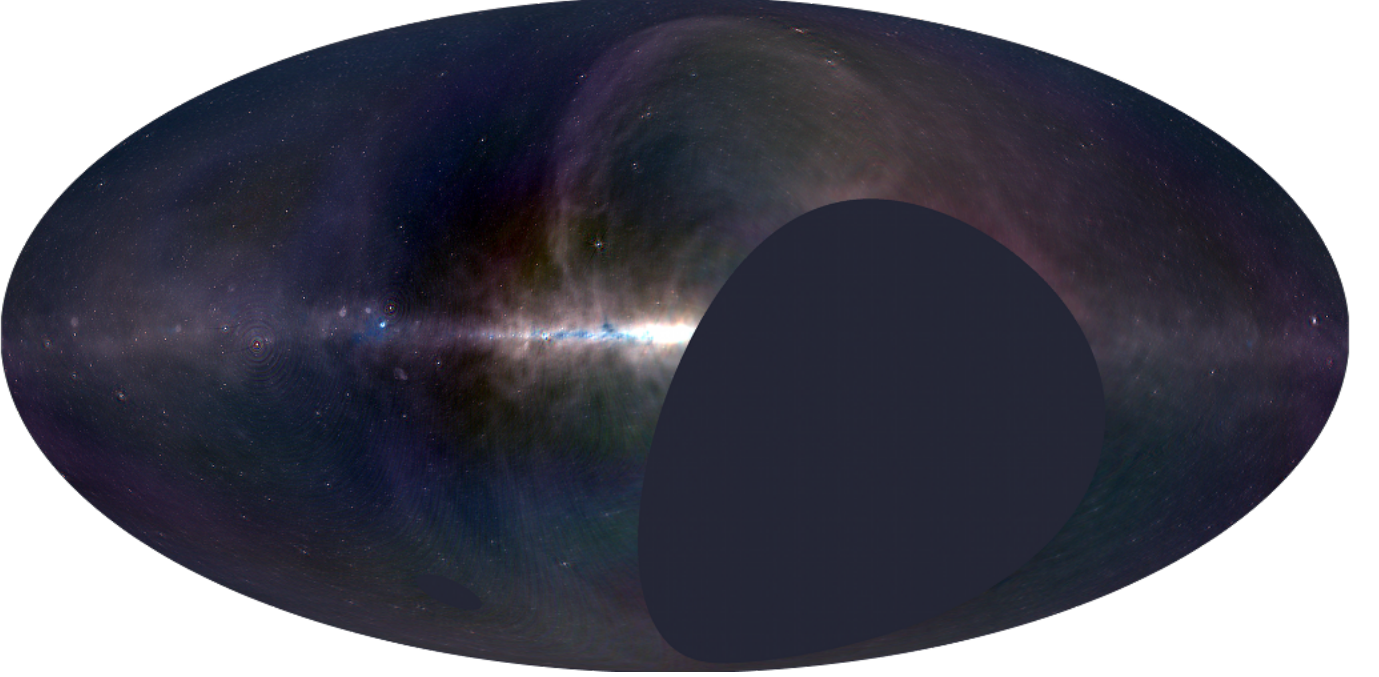


Figure 1. three color map

harmonic coefficients of the sky brightness  $a_{lm,\nu}$ .

$$V_{m,\nu}^{ij} = \sum_l B_{lm,\nu}^{ij} a_{lm,\nu}, \quad (4)$$

where the transfer coefficients  $B_{lm,\nu}^{ij}$  define the interferometer's response to the sky. For example, the transfer coefficients are a function of the baseline and antenna primary beam pattern.

Equation 4 can be recognized as a matrix equation where the transfer matrix  $\mathbf{B}$  is block-diagonal.

$$\mathbf{B} = \begin{pmatrix} m=0 & & & \\ & m=\pm 1 & & \\ & & m=\pm 2 & \\ & & & \ddots \end{pmatrix} \quad (5)$$

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 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  $a_{l(-m)} = (-1)^m a_{lm}^*$ .

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

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. When creating the maps presented in this paper, we use  $l_{\min} = 0$  and  $l_{\max} = 1000$ .

The interferometer's sensitivity to  $l = 0$ , however, deserves special consideration. Venumadhav et al. (2016) prove – under fairly general assumptions – that an interferometer is only sensitive to the monopole of the sky brightness if there exists some form of cross-talk or common-mode noise. In fact, the sensitivity of the interferometer is proportional to the amplitude of these effects. For consistency we will include  $a_{00}$  while solving Equation 6 for the vector  $\mathbf{a}$ , but we set  $a_{00} = 0$  afterwards because we do not have a measurement of the amplitude of cross-talk or common-mode noise, which would be needed to correctly calibrate  $a_{00}$ .

The size of a typical block in the transfer matrix is  $(2N_{\text{base}}N_{\text{freq}}) \times (l_{\text{max}})$ . 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. Compare this number with the 5 PB required for the naive approach. 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 the equation is reduced by a factor  $N^3$ . For the case of the OVRO-LWA the equation breaks up into  $\sim 10^3$  blocks and so we save a factor of  $\sim 10^9$  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  $\|\mathbf{v} - \mathbf{B}\mathbf{a}\|^2$ , is given by

$$\hat{\mathbf{a}}_{\text{LLS}} = (\mathbf{B}^*\mathbf{B})^{-1}\mathbf{B}^*\mathbf{v}, \quad (7)$$

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

$$\hat{\mathbf{a}}_{\text{Tikhonov}} = (\mathbf{B}^*\mathbf{B} + \varepsilon\mathbf{I})^{-1}\mathbf{B}^*\mathbf{v}. \quad (8)$$

Tikhonov regularization adds a small value  $\varepsilon$  to the diagonal of  $\mathbf{B}^*\mathbf{B}$ , fixing the matrix's singularity. We can see

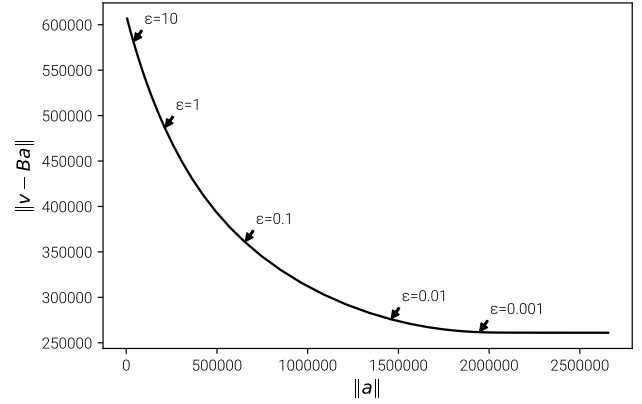


Figure 2. Example l-curve (from spw04)

this by using the singular value decomposition (SVD) of the matrix  $\mathbf{B} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^*$ . Equation 8 becomes

$$\hat{\mathbf{a}}_{\text{Tikhonov}} = \mathbf{V}(\mathbf{\Sigma}^2 + \varepsilon\mathbf{I})^{-1}\mathbf{\Sigma}\mathbf{U}^*\mathbf{v}, \quad (9)$$

where

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

The diagonal elements of  $\mathbf{\Sigma}$  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

$$\mathbf{C} = n\mathbf{V}(\mathbf{\Sigma}^2 + \varepsilon\mathbf{I})^{-2}\mathbf{\Sigma}^2\mathbf{V}^*. \quad (10)$$

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  $\|\mathbf{v} - \mathbf{B}\mathbf{a}\|^2 + \varepsilon\|\mathbf{a} - \mathbf{a}_{\text{prior}}\|^2$ .

### 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  $\|\mathbf{v} - \mathbf{B}\mathbf{a}\|^2$  and  $\|\mathbf{a}\|^2$  for each trial. An example is shown in Figure [TODO: create a figure](#). The shape of this curve has a characteristic L-shape, and as a result this type of plot is called an L-curve [TODO: reference](#). 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  $\|\mathbf{v} - \mathbf{B}\mathbf{a}\|^2$ .

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](#)).

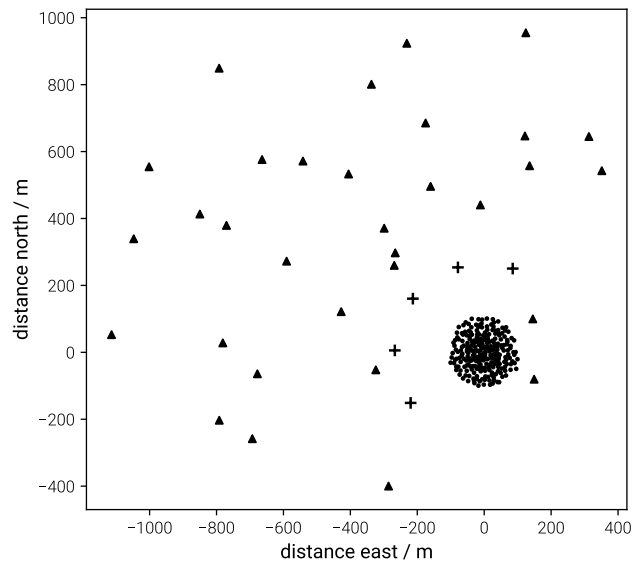
### 2.2.3. The Moore-Penrose Pseudoinverse

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

$$\hat{\mathbf{a}}_{\text{Moore-Penrose}} = \mathbf{B}^\dagger \mathbf{v}. \quad (11)$$

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. *m*-Mode Analysis Image Deconvolution



**Figure 3.** 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 open circles are the expansion antennas built in early 2016 in order to increase the longest baseline to 1.5 km. The 5 crosses are antennas equipped with noise-switched front ends.

While *m*-mode analysis imaging automatically deconvolves large angular scales, point sources do not get deconvolved by the imaging process. This is a limitation of the set of basis functions we have chosen to represent the sky – the spherical harmonics. Point sources carry power to large values of  $l$ , but the regularization will eventually begin to downweight spherical harmonics with  $l$  larger than what the interferometer is sensitive to. This consequence of this is that point sources will appear in the maps with some characteristic point spread function. This point spread function may be computed with

$$\mathbf{a}_{\text{PSF}}(\theta, \phi) = (\mathbf{B}^* \mathbf{B} + \varepsilon \mathbf{I})^{-1} \mathbf{B}^* \mathbf{B} \mathbf{a}_{\text{PS}}(\theta, \phi), \quad (12)$$

where  $\mathbf{a}_{\text{PSF}}(\theta, \phi)$  is the vector of spherical harmonic coefficients representing the point spread function 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

$$a_{lm, \text{PS}}(\theta, \phi) = Y_{lm}^*(\theta, \phi). \quad (13)$$

[TODO: include a few images of the computed PSF for my maps](#)

[TODO: describe how we incorporate this into major and minor iterations](#)

## 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 (OVRO) near Big Pine, California. The OVRO-LWA is a low-frequency instrument with instantaneous bandwidth covering 27.384 MHz to 84.912 MHz and 24 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. 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 out 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 **TODO: cite Price et al. 2017** to measure the global signal of 21 cm absorption from the cosmic dawn. Figure 3 is a diagram of the antenna configuration.

**TODO: cite Gregg’s instrument paper**

Beginning at 2017-02-17 12:00:00 UTC time, 28 consecutive hours of data was collected. This time was chosen based on the fact that it was raining at OVRO and rain tends to improve the low-frequency RFI environment considerably. 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.

**TODO: describe the ionospheric conditions at the time**

### 3.2. Gain Calibration

Antenna gain calibration is accomplished using an iterative method independently developed by 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>2</sup> under an open source license (GPLv3+).

The antenna complex gains are updated once per day from a 9 hour track of data when Cas A and Cyg A are at high elevations **TODO: the rainy day dataset actually uses a shorter track of data, we should try using the 9 hour track again**. These two sources are – by an order of magnitude – the brightest radio sources in the northern hemisphere. Therefore the optimal time to solve for the interferometer’s gain calibration is when Cas A and

Cyg A are at high elevations. A calibration obtained in this manner is good consistent with the following day’s calibration to within a 1% error in the flux scale and 1’ error in the astrometry **TODO: double check these numbers**. These residual errors are likely driven by differing ionospheric conditions.

The flux scale is tied to the work of Baars et al. (1977); Scaife & Heald (2012); Perley & Butler (2017).

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. Fortunately drift-scanning interferometers like the OVRO-LWA can empirically measure their primary beam under a mild set of symmetry assumptions (Pofer 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  $90^\circ$ . We measure the flux of several point sources **TODO: which ones??** 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.

### 3.4. Source Removal

#### 3.4.1. Cassiopeia A and Cygnus A

Without removing bright sources from the data, sidelobes from bright sources will dominate the variance in the image. At 74 MHz Cyg A is a 15,000 Jy source (Perley & Butler 2017). A conservative estimate for the confusion limit at 74 MHz with a 10 arcminute beam is 500 mJy (Lane et al. 2012). Therefore we require that Cyg A’s sidelobes be at most of  $-45$  dB down from its peak flux to prevent Cyg A’s sidelobes from dominating the variance in the image. **TODO: Measure Cyg A’s sidelobe levels if it is not subtracted**

At low frequencies, propagation effects through the ionosphere must be accounted for in order to achieve 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 diffraction events on the timescale of a single integration (13 seconds) are observed. Therefore the direction dependent calibration

<sup>2</sup> <https://github.com/mwestwood/TTCal.jl>

changes on these timescales and then we must solve for one set of complex gains per source per integration.

The largest angular scale of Cas A is  $\sim 8$  arcminutes, while the largest angular scale of Cyg A is  $\sim 2$  arcminutes. With a 10 arcminute resolution, 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 over 1 hour) it is possible to smear out the contribution of the rest of the sky. We then use NLOpt's Sbx routine (Johnson 2008; Rowan 1990) to vary the parameters in a source model until the variance within the aperture is minimized.

**TODO: Show a before and after with the updated source models**

**TODO: Give a plot with the updated source models**

#### 3.4.2. 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. Additionally while most astronomical sources at these frequencies have negative spectral indices, the Sun has a large positive spectral index. Therefore more care will need to be taken in subtracting the Sun at higher frequencies than at lower frequencies.

**TODO: describe sun removal – this isn't entirely settled yet**

## 4. ERROR ANALYSIS

### 4.1. Terrestrial Interference and Pickup

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

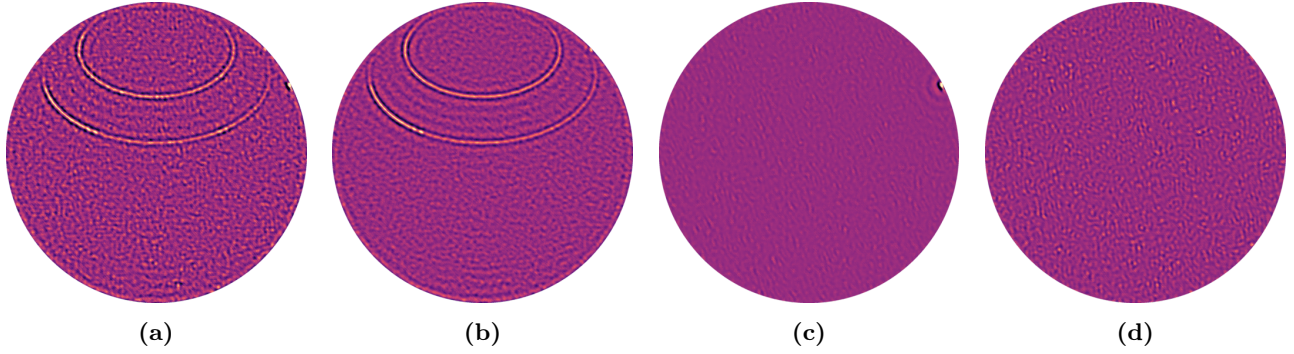
The Owens Valley is an important source of water and power for the city of Los Angeles. Unfortunately this means that high voltage power lines run along the valley

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 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 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 direction. This motivates the use of peeling (Mitchell et al. 2008; Smirnov & Tasse 2015), 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 4 for an illustration of this process.

However 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 panel (d) in Figure 4). In many cases the amplitude of these unassociated components exhibit the same sawtooth oscillations indicative of gain fluctuations in the analog electronics even after these gain fluctuations have been calibrated. This seems to imply that these unassociated components originate as common-mode pickup in the analog signal chain. That is, if the pickup occurs somewhere in the middle of the analog signal chain, it will not see the same gain as the astronomical signal. Therefore correcting the temperature dependence of the gains for the astronomical signal does not correct the temperature dependence of the pick-up and hence it will have a sawtooth pattern with respect to time.

The first step in equation 8 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



**Figure 4.** This figure illustrates the process of fitting for terrestrial sources of correlated noise. Each image is of the entire visible hemisphere above the OVRO-LWA but the data has been averaged over the entire 28 hour observing period. Because of this Cas A and Cyg A are smeared along tracks of constant declination. These images are generated using WSCLEAN (Offringa et al. 2014), uniform weighting, and baselines shorter than 15 wavelengths are flagged. (a) This image shows the initial data before any sources have been removed. (b) This image shows the final state after 3 terrestrial sources of correlated noise have been removed. (c) This image illustrates the contribution of an RFI source towards the north-west that was removed by peeling. (d) This removed component is not associated with a dot on the horizon. Instead it is likely associated with common-mode pickup in the analog signal chain.

$\mathbf{B}^* \mathbf{v}$  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 ( $\|\mathbf{B}^* \mathbf{v}_{\text{terrestrial}}\| / \|\mathbf{v}_{\text{terrestrial}}\|$ ).

**TODO: compute  $\mathbf{B}^* \mathbf{v}$  for panels (c) and (d) to show that the pick-up is pretty bad. my expectation is that it is actually worse than the RFI, but I don't actually know for sure**

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  $m \lesssim 3$ . 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 **TODO: make it**.

However because these rings are so distinctive in the final maps, it is possible to construct a Wiener filter that removes them at the cost of losing a small amount of information about astronomical sources.

**TODO: describe how the Wiener filter works**

#### 4.2. 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 con-

structed after watching each point in the sky move through large slices through 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.

**TODO: discuss the possibility of declination dependent flux-scale errors in these maps**

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. **TODO: quick simulation of this?**



### 4.3. 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 will cause even static sources to exhibit significant variability.

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

$$V_\nu(t_{\text{sidereal}}) = I_\nu B_\nu(t_{\text{sidereal}}), \quad (14)$$

where  $I_\nu$  is the flux of the source at the frequency  $\nu$ , and  $B_\nu$  is the baseline transfer function. 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, then the source flux is also a function of the time coordinate  $t$  such that

$$V_\nu(t_{\text{sidereal}}) = I_\nu(t) B_\nu(t_{\text{sidereal}}), \quad (15)$$

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

Now assume we have observed with our interferometer for a single sidereal day. In order to compute the  $m$ -modes we must Fourier transform with respect to sidereal time. In a real measurement this is a discrete Fourier transform of the observed correlation with respect to time (where the sum over time is restricted to a sidereal day). **TODO: fix normalization**

$$V_{\nu,m} = \sum_t V_\nu(t) e^{-imt} \quad (16)$$

Define  $V_{\nu,m}^{\text{static}}$  to be the observed  $m$ -modes if the source was actually static ( $I_\nu(t) \equiv I_{\nu,0}$ ). Then as a consequence of the Fourier convolution theorem

$$V_{\nu,m} = \sum_{m'} V_{m'}^{\text{static}} I_{\nu,m-m'}. \quad (17)$$

This will tend to scatter power between  $m$ -modes.

**TODO: cite richard for this**

**TODO: Simulate images of point sources with this**

**TODO: I think the conclusion is sort of exactly what you would expect: refraction broadens the PSF while scintillation effects the flux scale, but honestly not by much because we have such large averages that unless the scintillation is Cauchy distributed it is not that bad..**

### 4.4. Thermal Noise

**TODO: describe jack-knife tests**

## 5. RESULTS

### 5.1. Sky Maps

### 5.2. Error Maps

## 6. A NEARBY SUPERNOVA REMNANT????

## 7. DISCUSSION

These maps are pretty cool. Any questions?

**TODO: Cite Abhilash, Eastwood 2017**

## 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, *A&A*, 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
- Bonamente, M., Joy, M. K., LaRoque, S. J., et al. 2006, 647, 25
- 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
- Draine, B. T., & Lazarian, A. 1998, 508, 157
- Fialkov, A., Barkana, R., Pinhas, A., & Visbal, E. 2014, 437, L36
- Furlanetto, S. R., Oh, S. P., & Briggs, F. H. 2006, 433, 181
- Guzmán, A. E., May, J., Alvarez, H., & Maeda, K. 2011, 525, A138
- 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
- Hinshaw, G., Larson, D., Komatsu, E., et al. 2013, 208, 19
- Johnson, S. G. 2008, The NLOpt nonlinear-optimization package, <http://ab-initio.mit.edu/nlopt>
- Kocz, J., Greenhill, L. J., Barsdell, B. R., et al. 2015, 4, 1550003
- Lane, W. M., Cotton, W. D., Helmboldt, J. F., & Kassim, N. E. 2012, 47, RSOK04
- Leitch, E. M., Readhead, A. C. S., Pearson, T. J., & Myers, S. T. 1997, 486, L23
- 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
- Planck Collaboration, Ade, P. A. R., Aghanim, N., et al. 2014, 571, A16
- Planck Collaboration, Adam, R., Ade, P. A. R., et al. 2016a, 594, A10

#	$\nu$ / MHz	$\Delta\nu$ / MHz	$\theta$ / arcmin	noise / Jy
1	36.528	0.024		
2	41.760	0.024		
3	46.992	0.024		
4	52.224	0.024		
5	57.456	0.024		
6	62.688	0.024		
7	67.920	0.024		
8	73.152	0.024		

**Table 1.** Summary of the produced maps

- Planck Collaboration, Ade, P. A. R., Aghanim, N., et al. 2016b, 594, A13  
—, 2016c, 594, A25  
Poher, J. C., Parsons, A. R., Jacobs, D. C., et al. 2012, 143, 53  
Pritchard, J. R., & Loeb, A. 2012, 75, 086901  
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., 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  
Sunyaev, R. A., & Zel'dovich, Y. B. 1970, 2, 66  
—, 1972, 4, 173  
Thompson, A. R., Moran, J. M., & Swenson, Jr., G. W. 2001, Interferometry and Synthesis in Radio Astronomy, 2nd Edition  
Vanderlinde, K., Crawford, T. M., de Haan, T., et al. 2010, 722, 1180  
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  
Zheng, H., Tegmark, M., Dillon, J. S., et al. 2017a, 464, 3486  
—, 2017b, 465, 2901