

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

MICHAEL W. EASTWOOD, GREGG HALLINAN

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). Following the epoch of recombination the universe continued to expand and cool. Small inhomogeneities observed in the CMB grew larger and galaxies formed. The first stars formed and flooded the universe with light during the cosmic dawn, bringing the dark ages to an end. Eventually the cumulative influence of these galaxies completed a phase change in the intergalactic medium (IGM), and the universe returned to an ionized state following the conclusion of the epoch of reionization (EoR). Low frequency radio telescopes are poised to provide the first direct detection of the IGM from the cosmic dawn and the EoR using the redshifted hyperfine structure line of neutral hydrogen.

These low-frequency 21-cm cosmology experiments are limited by the dynamic range they can achieve against the low-redshift sources of emission. Depending on the redshift, length scale, and model the expected amplitude of the cosmological signal is 10 to 100 mK. On the other hand the galactic synchrotron emission is 1000 to 10000 K. **TODO: citation** Therefore experiments conservatively need to achieve at least 4 orders of dynamic range. In theory this kind of dynamic range can be obtained by exploiting the fact that the foreground emission is predominantly synchrotron emission that has a smooth power law spectrum. On the other hand the frequency structure of the cosmological signal is expected to be un-smooth because different frequencies trace different redshifts. This is the basic assumption made by all published upper limits on the cosmological 21 cm power spectrum. In practice the interferometer itself imparts frequency structure into the measurements and therefore an experiment’s ability to separate the foreground signal from the cosmological signal can be limited by the instrumental spectral smoothness.

Ideally the foreground removal strategy should be informed by the measured spatial and frequency structure

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 a data-driven interpolation of various maps between 10 MHz and 100 GHz, but below 408 MHz, the interpolation is largely driven by the Haslam 408 MHz map (Haslam et al. 1981, 1982).

TODO: TGSS, MSSS, GLEAM

At lower frequencies, free-free absorption and synchrotron self-absorption becomes important.

Galactic structure.

Planets?

2. OBSERVATIONS

2.1. 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 24.768 MHz to 82.296 MHz **TODO: double check this frequency range, I think it's wrong**. Each antenna stand hosts two perpendicular broadband dipoles so that there are 288×2 signal paths in total. These signal paths feed into the 512-input LEDA correlator, 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 maximally non-redundant configuration optimized for minimum sidelobes in snapshot imaging. 251 of these antennas are contained within a 200 m diameter core that is independently optimized for imaging sidelobe levels. 32 antennas are placed outside of the core in order to extend the maximum baseline length out to ~ 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

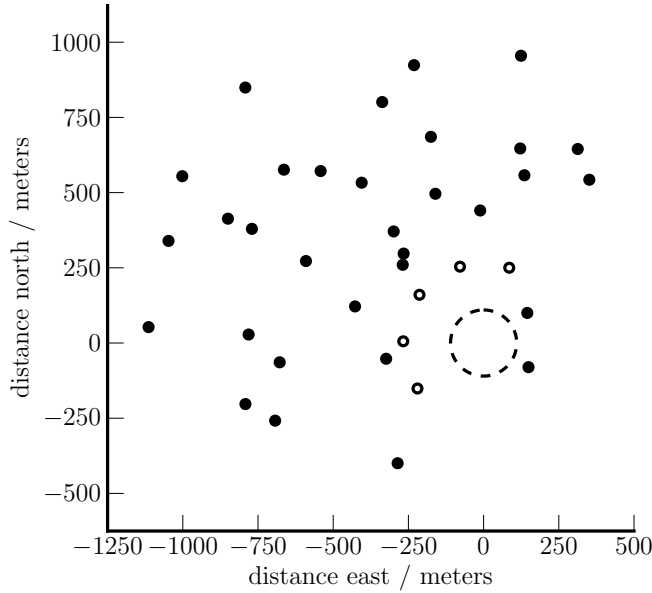


Figure 1. This figure shows the antenna layout for the OVRO LWA. The 200 meter diameter area indicated by the dashed line contains 251 antennas distributed in a pseudo-random configuration. The 32 filled circles are the expansion antennas built in early 2016 to increase the longest baseline to 1.5 km. The 5 open circles are antennas equipped with noise-switched front ends.

LEDA experiment to measure the global signal of 21 cm absorption from the cosmic dawn. **TODO: cite a leda paper** Calibrated total power measurements from these antennas will be used to constrain sky modes with $l \lesssim 2$ including the monopole. **TODO: What values of l actually contribute to the total power measurements?**

2.2. The OVRO LWA 100 Hour Dataset

Because the correlator only accepts 512 inputs (256 antennas with 2 polarizations for each antenna), a subset of the antennas must be selected for correlation.

Include plot of antenna layout.

The integration time is 13 seconds. This was chosen because the remainder of one sidereal day divided into 13 second intervals is only 90 milliseconds. Finer control of the integration time is not possible due to technical constraints.

2.3. Antenna 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. 2014), and is publicly available online¹ under an open source license (GPLv3+).

The antenna complex gains are updated once per day when Cas A and Cyg A are near the meridian. 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 for > 24 hours. **TODO: quantify this? how much does the next day’s calibration differ?**

The flux scale is tied to the work of Perley & Butler (2016).

TODO: sawtooth pattern removal

2.4. Source Removal

2.4.1. 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 2016). 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 changes on these timescales and the we must solve for one set of complex gains per source per integration.

The largest angular scale of Cas A is ~ 8 arcminutes, while the largest angular scale of Cyg A is ~ 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 Sbplx 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

2.4.2. The Sun

The Sun can be trivially removed from any map of the sky by constructing the map using only data taken

¹ <https://github.com/mwestwood/TTCal.jl>

at night. A map of the entire sky can be obtained by using observations spaced 6 months apart. However the OVRO LWA 100 hour dataset consists of 100 consecutive hours of data, and hence roughly half of the data is from the day.

The Sun is a time variable source in radio.

TODO: does the previous routine work on the Sun too? worth a shot

2.4.3. Sparking Power Lines

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. 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 path difference from a point source of RFI located at the position \vec{r} to two antennas located at positions \vec{x}_1 and \vec{x}_2 respectively is given by

$$\delta = \|\vec{r} - \vec{x}_1\| - \|\vec{r} - \vec{x}_2\|, \quad (1)$$

where $\|\cdot\|$ denotes the Euclidean norm. The contribution to the visibility observed on the baseline between these two antennas is therefore

$$V_{\nu}^{\text{RFI}} = F_{\nu} \exp\left(2\pi i \frac{\delta}{\lambda}\right), \quad (2)$$

where F_{ν} is the flux of the RFI at the frequency ν (wavelength λ). At first glance it might seem to be sufficient to know the location of the RFI. Then by fitting for the value of F_{ν} in equation 2 one can subtract the contribution of the RFI from each visibility. This approach is insufficient because the RFI originates from extremely low elevations where antenna shadowing effects are large. In practice this means that the response of each antenna to a source near or on the horizon is unknown and inhomogeneous between antennas. Therefore a free parameter for the response of each antenna towards the source of RFI must also be included as a free parameter when removing the RFI.

TODO: can we take advantage of the fact that the antenna response shouldn't be changing very much between integrations?

If the RFI from the horizon is not adequately removed from the data set, it will introduce systematic errors into the map of the sky.

2.5. Fitting the Beam

The OVRO LWA is composed of dipole antennas that are not combined into a beam prior to cross-correlation. As a result

2.6. Ionospheric Conditions

3. 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 ν . 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

$$V_{\nu}^{ij} = \int_{\text{sky}} A_{\nu}(\hat{r}) I_{\nu}(\hat{r}) \exp\left(2\pi i \hat{r} \cdot \vec{b}_{ij} / \lambda\right) d\Omega. \quad (3)$$

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

Naively one might attempt to solve equation 3 by discretizing, and subsequently solving the resulting matrix equation. If the interferometer is composed of N_{base} baselines, and measures N_{freq} frequency channels over N_{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_{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 Owens Valley 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 3 by means of discretization is usually impossible **TODO: cite Jayce's paper (has he published that yet?) where he does this, but explain why his work is the exception instead of the rule.**

Instead it is common to make mild assumptions that simplify equation 3 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 3 can be well-approximated by a two-dimensional Fourier transform.

TODO: discuss w-projection

TODO: discuss a-projection

Zheng et al. (2017b)

3.1. m -Mode Analysis

On the other hand, transit telescopes can take advantage of a symmetry in equation 3 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 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_{ν}^{ij} are a periodic function of sidereal time². Therefore it is a very natural operation to compute the Fourier transform of the visibilities with respect to sidereal time $\phi \in (0, 2\pi]$.

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

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.

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

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

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

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

$$\mathbf{B} = \begin{pmatrix} m=0 & & & \\ & m=\pm 1 & & \\ & & m=\pm 2 & \\ & & & \ddots \end{pmatrix} \quad (6)$$

The vector \mathbf{v} contains the list of m -modes and the vector \mathbf{a} contains the list of spherical harmonic coefficients rep-

resenting 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 .

$$\overbrace{\begin{pmatrix} \vdots \\ m\text{-modes} \\ \vdots \end{pmatrix}}^{\mathbf{v}} = \overbrace{\begin{pmatrix} \ddots & & \\ & \text{transfer matrix} & \\ & & \ddots \end{pmatrix}}^{\mathbf{B}} \overbrace{\begin{pmatrix} \vdots \\ a_{lm} \\ \vdots \end{pmatrix}}^{\mathbf{a}} \quad (7)$$

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 $l_{\max} = 1000$. **TODO: exclude $l = 0$ and cite tejas' paper**

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. 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 7. 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.

Imaging in m -mode analysis essentially amounts to inverting equation 7 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

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

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 $\mathbf{B}^* \mathbf{B}$ must have at least one zero eigenvalue, which means it is not invertible.

Another way of looking at the problem is that because

² 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.

the interferometer is not sensitive to part of the southern hemisphere, there are infinitely many possible solutions to equation 7 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. The process of Tikhonov regularization minimizes $\|\mathbf{v} - \mathbf{B}\mathbf{a}\|^2 + \epsilon\|\mathbf{a}\|^2$ for some arbitrary value of $\epsilon > 0$ chosen by the observer. The solution that minimizes this expression is given by

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

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

TODO: describe how Tikhonov regularization improves the noise properties of the map

TODO: describe the Moore-Penrose pseudoinverse and why I have elected not to use it

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

Tikhonov regularization requires the observer to pick the value of ϵ . If ϵ 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 ϵ is too small then the problem will be poorly regularized and the resulting sky map may not represent the true sky.

A tool that can be used to address this problem is the L-Curve.

3.2. *m*-Mode Analysis with the OVRO LWA

3.3. CLEAN

4. ERRORS

4.1. Stationary Correlated Noise

However it is instructive to analyze the contaminating effect the RFI will have on *m*-mode analysis maps of the sky. The correlation observed on a given baseline and frequency channel for a single source of RFI is

$$V(t) = I(t)B^{\text{near-field}} \quad (10)$$

However we need a strategy to mitigate the RFI before imaging. A constant source attached to the horizon has a fringe rate of zero. The only possible astronomical origins for a source with zero fringe rate are:

1. a point source located at the north or south celestial poles, and
2. a ring of constant declination.

Therefore while snapshot images will show the RFI source on the horizon, extended integrations will show

the RFI as rings of constant declination. We can take advantage of this discrepancy to separate astronomical emission from terrestrial RFI.

4.2. Ionospheric Scintillation

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

$$V_\nu(t_{\text{sidereal}}) = I_\nu B_\nu(t_{\text{sidereal}}), \quad (11)$$

where I_ν is the flux of the source at the frequency ν , and B_ν 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_{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

$$V_\nu(t_{\text{sidereal}}) = I_\nu(t)B_\nu(t_{\text{sidereal}}), \quad (12)$$

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**

$$V_{\nu,m} = \sum_t V_\nu(t)e^{-imt} \quad (13)$$

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

$$V_{\nu,m} = \sum_{m'} V_{m'}^{\text{static}} I_{\nu,m-m'}. \quad (14)$$

This will tend to scatter power between *m*-modes. Simulate images of point sources with this

4.3. Ionospheric Refraction

4.4. Beam Errors

5. RESULTS

5.1. Sky Maps

5.2. Error Maps

TODO: Cite Abhilash, Eastwood 2016

6. DISCUSSION

#	ν / MHz	$\Delta\nu$ / MHz
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

REFERENCES

- Bezanson, J., Edelman, A., Karpinski, S., & Shah, V. B. 2014, 1411.1607
- de Oliveira-Costa, A., Tegmark, M., Gaensler, B. M., et al. 2008, MNRAS, 388, 247
- Haslam, C. G. T., Klein, U., Salter, C. J., et al. 1981, A&A, 100, 209
- Haslam, C. G. T., Salter, C. J., Stoffel, H., & Wilson, W. E. 1982, A&AS, 47, 1
- Hinshaw, G., Larson, D., Komatsu, E., et al. 2013, ApJS, 208, 19
- Johnson, S. G. 2008, The NLOpt nonlinear-optimization package, <http://ab-initio.mit.edu/nlopt>
- Lane, W. M., Cotton, W. D., Helmboldt, J. F., & Kassim, N. E. 2012, Radio Science, 47, RS0K04
- Mitchell, D. A., Greenhill, L. J., Wayth, R. B., et al. 2008, IEEE Journal of Selected Topics in Signal Processing, 2, 707
- Perley, R. A., & Butler, B. J. 2016, ArXiv e-prints, arXiv:1609.05940
- Rowan, T. 1990, PhD thesis, Department of Computer Sciences, University of Texas at Austin
- Salvini, S., & Wijnholds, S. J. 2014, A&A, 571, A97
- Shaw, J. R., Sigurdson, K., Pen, U.-L., Stebbins, A., & Sitwell, M. 2014, ApJ, 781, 57
- Shaw, J. R., Sigurdson, K., Sitwell, M., Stebbins, A., & Pen, U.-L. 2015, PhRvD, 91, 083514
- Smirnov, O. M., & Tasse, C. 2015, MNRAS, 449, 2668
- Zheng, H., Tegmark, M., Dillon, J. S., et al. 2017a, MNRAS, 464, 3486
- . 2017b, MNRAS, 465, 2901