University Library

MINERVA ACCESS

A gateway to Melbourne's research publications

Minerva Access is the Institutional Repository of The University of Melbourne

Author/s:
Dillon, JS;Neben, AR;Hewitt, JN;Tegmark, M;Barry, N;Beardsley, AP;Bowman, JD;Briggs, F;Carroll, P;De Oliveira-Costa, A;Ewall-Wice, A;Feng, L;Greenhill, LJ;Hazelton, BJ;Hernquist, L;Hurley-Walker, N;Jacobs, DC;Kim, HS;Kittiwisit, P;Lenc, E;Line, J;Loeb, A;McKinley, B;Mitchell, DA;Morales, MF;Offringa, AR;Paul, S;Pindor, B;Pober, JC;Procopio, P;Riding, J;Sethi, S;Shankar, NU;Subrahmanyan, R;Sullivan, I;Thyagarajan, N;Tingay, SJ;Trott, C;Wayth, RB;Webster, RL;Wyithe, S;Bernardi, G;Cappallo, RJ;Deshpande, AA;Johnston-Hollitt, M;Kaplan, DL;Lonsdale, CJ;McWhirter, SR;Morgan, E;Oberoi, D;Ord, SM;Prabu, T;Srivani, KS;Williams, A;Williams, CL

Title:
Empirical covariance modeling for 21 cm power spectrum estimation: A method demonstration and new limits from early Murchison Widefield Array 128-tile data EMPIRICAL COVARIANCE MODELING for 21 CM POWER .. JOSHUA S. DILLON et al.

Date:
2015-06-19

Citation:
Dillon, J. S., Neben, A. R., Hewitt, J. N., Tegmark, M., Barry, N., Beardsley, A. P., Bowman, J. D., Briggs, F., Carroll, P., De Oliveira-Costa, A., Ewall-Wice, A., Feng, L., Greenhill, L. J., Hazelton, B. J., Hernquist, L., Hurley-Walker, N., Jacobs, D. C., Kim, H. S., Kittiwisit, P. ,... Williams, C. L. (2015). Empirical covariance modeling for 21 cm power spectrum estimation: A method demonstration and new limits from early Murchison Widefield Array 128-tile data EMPIRICAL COVARIANCE MODELING for 21 CM POWER .. JOSHUA S. DILLON et al.. Physical Review D Particles Fields Gravitation and Cosmology, 91 (12), https://doi.org/10.1103/PhysRevD.91.123011.

Persistent Link:
https://hdl.handle.net/11343/116546

Empirical covariance modeling for 21 cm power spectrum estimation: A method demonstration and new limits from early Murchison Widefield Array 128-tile data

Joshua S. Dillon,[1,2,*] Abraham R. Neben,[1,2] Jacqueline N. Hewitt,[1,2] Max Tegmark,[1,2] N. Barry,[3] A. P. Beardsley,[3] J. D. Bowman,[4] F. Briggs,[5,6] P. Carroll,[3] A. de Oliveira-Costa,[1,2] A. Ewall-Wice,[1,2] L. Feng,[1,2] L. J. Greenhill,[7] B. J. Hazelton,[3] L. Hernquist,[7] N. Hurley-Walker,[8] D. C. Jacobs,[4] H. S. Kim,[9,6] P. Kittiwisit,[4] E. Lenc,[10,6] J. Line,[9,6] A. Loeb,[7] B. McKinley,[9,6] D. A. Mitchell,[11,6] M. F. Morales,[3] A. R. Offringa,[5,6] S. Paul,[12] B. Pindor,[9,6] J. C. Pober,[3] P. Procopio,[9,6] J. Riding,[9,6] S. Sethi,[12] N. Udaya Shankar,[12] R. Subrahmanyan,[12,6] I. Sullivan,[3] Nithyanandan Thyagarajan,[4] S. J. Tingay,[8,6] C. Trott,[8,6] R. B. Wayth,[8,6] R. L. Webster,[9,6] S. Wyithe,[9,6] G. Bernardi,[13,14] R. J. Cappallo,[15] A. A. Deshpande,[12] M. Johnston-Hollitt,[16] D. L. Kaplan,[17] C. J. Lonsdale,[15] S. R. McWhirter,[15] E. Morgan,[2] D. Oberoi,[18] S. M. Ord,[8,6] T. Prabu,[12] K. S. Srivani,[12] A. Williams,[8] and C. L. Williams[1,2]

[1]*Department of Physics, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA*
[2]*MIT Kavli Institute, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA*
[3]*Department of Physics, University of Washington, Seattle, Washington 98195, USA*
[4]*School of Earth and Space Exploration, Arizona State University, Tempe, Arizona 85287, USA*
[5]*Research School of Astronomy and Astrophysics, Australian National University, Canberra, Australian Capital Territory 2611, Australia*
[6]*ARC Centre of Excellence for All-sky Astrophysics (CAASTRO)*
[7]*Harvard-Smithsonian Center for Astrophysics, Cambridge, Massachusetts 02138, USA*
[8]*International Centre for Radio Astronomy Research, Curtin University, Bentley, Western Australia 6102, Australia*
[9]*School of Physics, The University of Melbourne, Parkville, Victoria 3010, Australia*
[10]*Sydney Institute for Astronomy, School of Physics, The University of Sydney, New South Wales 2006, Australia*
[11]*CSIRO Astronomy and Space Science (CASS), PO Box 76, Epping, New South Wales 1710, Australia*
[12]*Raman Research Institute, Bangalore 560080, India*
[13]*Department of Physics and Electronics, Rhodes University, Grahamstown, 6140, South Africa*
[14]*Square Kilometre Array South Africa (SKA SA), Cape Town 7405, South Africa*
[15]*MIT Haystack Observatory, Westford, Massachusetts 01886, USA*
[16]*School of Chemical & Physical Sciences, Victoria University of Wellington, Wellington 6140, New Zealand*
[17]*Department of Physics, University of Wisconsin–Milwaukee, Milwaukee, Wisconsin 53201, USA*
[18]*National Centre for Radio Astrophysics, Tata Institute for Fundamental Research, Pune 411007, India*

(Received 9 March 2015; published 19 June 2015)

The separation of the faint cosmological background signal from bright astrophysical foregrounds remains one of the most daunting challenges of mapping the high-redshift intergalactic medium with the redshifted 21 cm line of neutral hydrogen. Advances in mapping and modeling of diffuse and point source foregrounds have improved subtraction accuracy, but no subtraction scheme is perfect. Precisely quantifying the errors and error correlations due to missubtracted foregrounds allows for both the rigorous analysis of the 21 cm power spectrum and for the maximal isolation of the "EoR window" from foreground contamination. We present a method to infer the covariance of foreground residuals from the data itself in contrast to previous attempts at *a priori* modeling. We demonstrate our method by setting limits on the power spectrum using a 3 h integration from the 128-tile Murchison Widefield Array. Observing between 167 and 198 MHz, we find at 95% confidence a best limit of $\Delta^2(k) < 3.7 \times 10^4$ mK^2 at comoving scale $k = 0.18\ h\,\mathrm{Mpc}^{-1}$ and at $z = 6.8$, consistent with existing limits.

DOI: 10.1103/PhysRevD.91.123011 PACS numbers: 95.75.-z, 95.75.Kk, 98.80.-k, 98.80.Es

*jsdillon@mit.edu

 © 2015 American Physical Society

I. INTRODUCTION

Tomographic mapping of neutral hydrogen using its 21 cm hyperfine transition has the potential to directly probe the density, temperature, and ionization of the intergalactic medium (IGM), from redshift 50 (and possibly earlier) through the end of reionization at $z \sim 6$. This unprecedented view of the so-called "cosmic dawn" can tightly constrain models of the first stars and galaxies [1–4] and eventually yield an order of magnitude more precise test of the standard cosmological model ΛCDM than current probes [5].

Over the past few years, first generation instruments have made considerable progress toward the detection of the power spectrum of the 21 cm emission during the epoch of reionization (EoR). Telescopes such as the Low Frequency Array [6], the Donald C. Backer Precision Array for Probing the Epoch of Reionization (PAPER [7]), the Giant Metrewave Radio Telescope (GMRT [8]), and the Murchison Widefield Array (MWA [9–11]) are now operating, and have begun to set limits on the power spectrum. GMRT set some of the earliest limits [8], and both PAPER [12] and the MWA [13] have presented upper limits across multiple redshifts using small prototype arrays. PAPER has translated its results into a constraint on the heating of the IGM by the first generation of x-ray binaries and miniquasars [14] and has placed the tightest constraints so far on the power spectrum [15] and the thermal history of the IGM [16].

Despite recent advances, considerable analysis challenges remain. Extracting the subtle cosmological signal from the noise is expected to require thousand hour observations across a range of redshifts [17–22]. Even more daunting is the fact that the 21 cm signal is probably at least 4 orders of magnitude dimmer than the astrophysical foregrounds—due to synchrotron radiation both from our Galaxy and from other galaxies [23–28].

Recently, simulations and analytical calculations have established the existence of a region in cylindrical Fourier space—in which three-dimensional (3D) Fourier modes $\vec{k}$ are binned into $k_\parallel$ modes along the line of sight and $k_\perp$ modes perpendicular to it—called the "EoR window" that should be fairly free of foreground contamination [22,29–36]. Observations of the EoR window confirm that it is largely foreground-free [13,26] up to the sensitivity limits of current experiments. The boundary of the EoR window is determined by the volume and resolution of the observation, the intrinsic spectral structure of the foregrounds, and the so-called "wedge."

Physically, the wedge arises from the frequency dependence of the point spread function (PSF) of any interferometer, which can create spectral structure from spectrally smooth foregrounds in our 3D maps (see [35] for a rigorous derivation). Fortunately, in $k_\parallel - k_\perp$ space, instrumental chromaticity from flat-spectrum sources is restricted to the region below

$$k_\parallel = \theta_0 \frac{D_M(z)E(z)}{D_H(1+z)} k_\perp, \tag{1}$$

where $D_H \equiv c/H_0$, $E(z) \equiv \sqrt{\Omega_m(1+z)^3 + \Omega_\Lambda}$, and $D_m(z) \equiv \int_0^z dz'/E(z')$ with cosmological parameters from [37]. The size of the region is determined by θ_0, the angle from zenith beyond which foregrounds do not significantly contribute. While most of the foreground emission we observe should appear inside the main lobe of the primary beam, foreground contamination from sources in the sidelobes are also significant compared to the signal [38,39]. A conservative choice of θ_0 is therefore $\pi/2$, which reflects the fact that the maximum possible delay a baseline can measure corresponds to a source at the horizon [30]. Still, this foreground isolation is not foolproof and can be easily corrupted by miscalibration and imperfect data reduction. Further, slowly varying spectral modes just outside the wedge are also affected when the foreground residuals have spectral structure beyond that imprinted by the chromaticity of the interferometer.

To confidently detect the 21 cm EoR power spectrum, we need rigorous statistical techniques that incorporate models of the cosmological signal, the foregrounds, the instrument, the instrumental noise, and the exact mapmaking procedure. With this information, one may use estimators that preserve as much cosmological information as possible and thoroughly propagate errors due to noise and foregrounds through the analysis pipeline.

The development of such statistical techniques has progressed rapidly over the past few years. The quadratic estimator formalism was adapted [40] from previous work on the cosmic microwave background [41] and galaxy surveys [42]. It was accelerated to meet the data volume challenges of 21 cm tomography [43] and refined to overcome some of the difficulties of working with real data [13]. Further, recent work has shown how to rigorously incorporate the interferometric effects that create the wedge [35,36,44], though they rely on precision instrument modeling, including exact per-frequency and per-antenna primary beams and complex gains. A similar technique designed for drift-scanning telescopes using spherical harmonic modes was developed in [45,46], which also demonstrated the need for a precise understanding of one's instrument.

However, at this early stage in the development of 21 cm cosmology, precision instrument characterization remains an active area of research [47–50]. We thus pursue a more cautious approach to foreground modeling that reflects our incomplete knowledge of the instrument by modeling the residual foreground covariance from the data itself. As we will show, this mitigates systematics such as calibration errors that would otherwise impart spectral structure onto the foregrounds, corrupting the EoR window. While not a fully Bayesian approach like those of [51] and [52], our technique discovers both the statistics of the foregrounds and the power spectrum from the data. Our foreground models are subject to certain prior assumptions but are allowed to be data

motivated in a restricted space. However, by working in the context of the quadratic estimator formalism, we can benefit from the computational speedups of [43]. This work is meant to build on those techniques and make them more easily applied to real and imperfect data.

This paper is organized into two main parts. In Sec. II we discuss the problem of covariance modeling in the context of power spectrum estimation and present a method for the empirical estimation of that foreground model, using MWA data to illustrate the procedure. Then, in Sec. III, we explain how these data were taken and reduced into maps and present the results of our power spectrum estimation procedure on a few hours of MWA observation, including limits on the 21 cm power spectrum.

II. EMPIRICAL COVARIANCE MODELING

Before presenting our method of empirically modeling the statistics of residual foregrounds in our maps, we need to review the importance of these covariances to power spectrum estimation. We begin in Sec. II A with a brief review of the quadratic estimator formalism for optimal power spectrum estimation and rigorous error quantification. We then discuss in Sec. II B the problem of covariance modeling in greater detail, highlighting exactly which unknowns we are trying to model with the data. Next we present in Sec. II C our empirical method of estimating the covariance of foreground residuals, illustrated with an application to MWA data. Lastly, we review in Sec. II D the assumptions and caveats that we make or inherit from previous power spectrum estimation work.

A. Quadratic power spectrum estimator review

The fundamental goal of power spectrum estimation is to reduce the volume of data by exploiting statistical symmetries while retaining as much information as possible about the cosmological power spectrum [41]. We seek to estimate a set of band powers $\mathbf{p}$ using the approximation that

$$P(\mathbf{k}) \approx \sum_\alpha p_\alpha \chi_\alpha(\mathbf{k}), \tag{2}$$

where $P(\mathbf{k})$ is the power spectrum as a function of wave vector $\mathbf{k}$ and χ_α is an indicator function that equals 1 wherever we are approximating $P(\mathbf{k})$ by p_α and vanishes elsewhere.

Following [13,40,43], we estimate power spectra from a "data cube"—a set of sky maps of brightness temperature at many closely spaced frequencies—which we represent as a single vector $\hat{\mathbf{x}}$ whose index iterates over both position and frequency. From $\hat{\mathbf{x}}$, we estimate each band power as

$$\hat{p}_\alpha = \frac{1}{2} M_{\alpha\beta} (\hat{\mathbf{x}}_1 - \boldsymbol{\mu})^\mathsf{T} \mathbf{C}^{-1} \mathbf{C}_{,\beta} \mathbf{C}^{-1} (\hat{\mathbf{x}}_2 - \boldsymbol{\mu}) - b_\alpha. \tag{3}$$

Here $\boldsymbol{\mu} = \langle \hat{\mathbf{x}} \rangle$, the ensemble average of our map over many different realizations of the observation, and $\mathbf{C}$ is the covariance of our map,

$$\mathbf{C} = \langle \hat{\mathbf{x}} \hat{\mathbf{x}}^\mathsf{T} \rangle - \langle \hat{\mathbf{x}} \rangle \langle \hat{\mathbf{x}} \rangle^\mathsf{T}. \tag{4}$$

$\mathbf{C}_{,\beta}$ is a matrix that encodes the response of the covariance to changes in the true, underlying band powers; roughly speaking, it performs the Fourier transforming, squaring, and binning steps one normally associates with computing power spectra.[1] Additionally, $\mathbf{M}$ is an invertible normalization matrix and b_α is the power spectrum bias from nonsignal contaminants in $\hat{\mathbf{x}}$. In this work, we follow [13] and choose a form of $\mathbf{M}$ such that $\Sigma \equiv \mathrm{Cov}(\hat{\mathbf{p}})$ is diagonal, decorrelating errors in the power spectrum and thus reducing foreground leakage into the EoR window. In order to calculate $\mathbf{M}$ and Σ quickly, we use the fast method of [43] which uses fast Fourier transforms and Monte Carlo simulations to approximate these matrices.

Finally, temporally interleaving the input data into two cubes $\hat{\mathbf{x}}_1$ and $\hat{\mathbf{x}}_2$ with the same sky signal but independent noise avoids a noise contribution to the bias b_α as in [13]. Again following [13], we abstain from subtracting a foreground residual bias in order to avoid any signal loss (as discussed in Sec. II C 3).

B. What does our covariance model represent?

Our brightness temperature data cubes are made up of contributions from three statistically independent sources: the cosmological signal, $\hat{\mathbf{x}}^S$; the astrophysical foregrounds, $\hat{\mathbf{x}}^{FG}$; and the instrumental noise $\hat{\mathbf{x}}^N$. It follows that the covariance matrix is equal to the sum of their separate covariances:

$$\mathbf{C} = \mathbf{C}^S + \mathbf{C}^{FG} + \mathbf{C}^N. \tag{5}$$

Hidden in the statistical description of the different contributions to our measurement is an important subtlety. Each of these components is taken to be a particular instantiation of a random process, described by a mean and covariance. In the case of the cosmological signal, it is the underlying statistics—the mean and covariance—which encode information about the cosmology and astrophysics. However, we can only learn about those statistics by assuming statistical isotropy and homogeneity and by assuming that spatial averages can stand in for ensemble averages in large volumes. In the case of the instrumental noise, we usually think of the particular instantiation of the noise that we see as the result of a random trial.

The foregrounds are different. There is only one set of foregrounds, and they are not random. If we knew exactly how the foregrounds appear in our observations, we would subtract them from our maps and then ignore them in this analysis. We know that we do not know the foregrounds exactly, and so we choose to model them with our best guess, $\boldsymbol{\mu}^{FG}$. If we define the cosmological signal to consist

[1]For a derivation of an explicit form of $\mathbf{C}_{,\beta}$, see [40] or [43].

only of fluctuations from the brightness temperature of the global 21 cm signal, then the signal and the noise both have $\boldsymbol{\mu}^S = \boldsymbol{\mu}^N = 0$. Therefore, we start our power spectrum estimation using Eq. (3) by subtracting off our best guess as to the foreground contamination in our map. But how wrong are we?

The short answer is that we do not really know that either. But, if we want to take advantage of the quadratic estimator formalism to give the highest weight to the modes we are most confident in, then we must model the statistics of our foreground residuals. If we assume that our error is drawn from some correlated Gaussian distribution, then we should use that *foreground uncertainty covariance* as the proper $\mathbf{C}^{FG}$ in Eq. (3).

So what do we know about the residual foregrounds in our maps? In theory, our dirty maps are related to the true sky by a set of point spread functions that depend on both position and frequency [44]. This is the result of both the way our interferometer turns the sky into measured visibilities and the way we make maps to turn those visibilities into $\hat{\mathbf{x}}$. In other words, there exists some matrix of PSFs, $\mathbf{P}$ such that

$$\langle \hat{\mathbf{x}} \rangle = \mathbf{P}\mathbf{x}^{\text{true}}. \tag{6}$$

The spectral structure in our maps that creates the wedge feature in the power spectrum is a result of $\mathbf{P}$.

We can describe our uncertainty about the true sky—about the positions, fluxes, and spectral indices of both diffuse foregrounds and points sources—with a covariance matrix $\mathbf{C}^{FG,\text{true}}$ [40,43], so that

$$\mathbf{C}^{FG} = \mathbf{P}\mathbf{C}^{FG,\text{true}}\mathbf{P}^{\mathsf{T}}. \tag{7}$$

This equation presents us with two ways of modeling the foregrounds. If we feel that we know the relationship between our dirty maps and the true sky precisely, then we can propagate our uncertainty about a relatively small number of foreground parameters, as discussed by [40] and [43], through the $\mathbf{P}$ matrix to get $\mathbf{C}^{FG}$. This technique, suggested by [44], relies on precise knowledge of $\mathbf{P}$. Of course, the relationship between the true sky and our visibility data depends both on the design of our instrument and on its calibration. If our calibration is very good—if we really understand our antenna gains and phases, our primary beams, and our bandpasses—then we can accurately model $\mathbf{P}$.

If we are worried about systematics (and at this early stage of 21 cm tomography with low frequency radio interferometers, we certainly are), then we need a complementary approach to modeling $\mathbf{C}^{FG}$ directly, one that we can use both for power spectrum estimation and for comparison to the results of a more theoretically motivated technique. This is the main goal of this work.

C. Empirical covariance modeling technique

The idea of using empirically motivated covariance matrices in the quadratic estimator formalism has some history in the field. Previous MWA power spectrum analysis [13] used the difference between time-interleaved data cubes to estimate the overall level of noise, empirically calibrating T_{sys}, the system temperature of the elements. PAPER's power spectrum analysis relies on using observed covariances to suppress systematic errors [14] and on bootstrapped error bars [12,14]. A similar technique was developed contemporaneously with this work and was used by [15] to estimate covariances.

$\mathbf{C}^{FG}$ has far more elements than we have measured voxels—our cubes have about 2×10^5 voxels, meaning that $\mathbf{C}^{FG}$ has up to 2×10^{10} unique elements. Therefore, any estimate of $\mathbf{C}^{FG}$ from the data needs to make some assumptions about the structure of the covariance. Since foregrounds have intrinsically smooth spectra, and since one generally attempts to model and subtract smooth spectrum foregrounds, it follows that foreground residuals will be highly correlated along the line of sight. After all, if we are undersubtracting foregrounds at one frequency, we are probably undersubtracting at nearby frequencies too. We therefore choose to focus on empirically constructing the part of $\mathbf{C}^{FG}$ that corresponds to the frequency-frequency covariance—the covariance along the line of sight. If there are n_f frequency channels, then that covariance matrix is only $n_f \times n_f$ elements and is likely dominated by a relatively small number of modes.

In this section, we will present an approach to solving this problem in a way that faithfully reflects the complex spectral structure introduced by an (imperfectly calibrated) interferometer on the bright astrophysical foregrounds. As a worked example, we use data from a short observation with the MWA which we will describe in detail in Sec. III. We begin with a uniformly weighted map of the sky at each frequency, a model for both point sources and diffuse emission imaged from simulated visibilities, and a model for the noise in each uv cell as a function of frequency.

The idea to model $\mathbf{C}^{FG}$ empirically was put forward by [53]. He attempted to model each line of sight as statistically independent and made no effort to separate $\mathbf{C}^{FG}$ from $\mathbf{C}^{N}$ or to reduce the residual noisiness of the frequency-frequency covariance.

Our approach centers on the idea that the covariance matrix can be approximated as block diagonal in the uv basis of Fourier modes perpendicular to the line of sight. In other words, we are looking to express $\mathbf{C}^{FG}$ as

$$C^{FG}_{uu'vv'ff'} \approx \delta_{uu'}\delta_{vv'}\hat{C}_{ff'}(k_\perp), \tag{8}$$

where $k_\perp$ is a function of $\sqrt{u^2+v^2}$. This is the tensor product of our best guess of the frequency-frequency covariance $\hat{\mathbf{C}}$ and the identity in both Fourier coordinates perpendicular to the line of sight. In this way, we can model

different frequency-frequency covariances as a function of $|u|$ or equivalently, $k_\perp$, reflecting that fact that the wedge results from greater leakage of power up from low $k_\parallel$ as one goes to higher $k_\perp$. This method also has the advantage that $\mathbf{C}$ becomes efficient to both write down and invert directly, removing the need for the preconditioned conjugate gradient algorithm employed by [43].

This approximation is equivalent to the assumption that the residuals in every line of sight are statistically independent of position. This is generally a pretty accurate assumption as long as the primary beam does not change very much over the map from which we estimate the power spectrum. However, because $\hat{C}_{ff'}(k_\perp)$ depends on the angular scale, we are still modeling correlations that depend only on the distance between points in the map.

While we might expect that the largest residual voxels correspond to errors in subtracting the brightest sources, the voxels in the residual data cube (the map minus the model) are only weakly correlated with the best-guess model of the foregrounds (we find a correlation coefficient $\rho = 0.116$, which suggests that sources are removed to roughly the 10% level, assuming that undersubtraction dominates). As we improve our best guess of the model foregrounds through better deconvolution, we expect ρ to go down, improving the assumption that foregrounds are block diagonal in the uv basis. We will now present the technique we have devised in four steps, employing MWA data as a method demonstration.

1. Compute sample covariances in uv annuli

We begin our empirical covariance calculation by taking the residual data cubes, defined as

$$\hat{\mathbf{x}}^{\rm res} \equiv \hat{\mathbf{x}}_1/2 + \hat{\mathbf{x}}_2/2 - \boldsymbol{\mu}, \tag{9}$$

and performing a discrete Fourier transform[2] at each frequency independently to get $\tilde{\mathbf{x}}^{\rm res}$. This yields $n_x \times n_y$ sample "lines of sight" (uv cells for all frequencies), as many as we have pixels in the map. As a first step toward estimating $\hat{\mathbf{C}}$, we use the unbiased sample covariance estimator from these residual lines of sight. However, instead of calculating a single frequency-frequency covariance, we want to calculate many different $\hat{\mathbf{C}}^{\rm res}$ matrices to reflect the evolution of the spectral structure with $k_\perp$ along the wedge. We therefore break the uv plane into concentric annuli of equal width and calculate $\hat{\mathbf{C}}^{\rm res}_{uv}$ for each uv cell as the sample covariance of the $N^{\rm LOS} - 2$ lines of sight in that annulus, excluding the cell considered and its complex conjugate. Since the covariance is assumed to be block diagonal, this eliminates a potential bias that comes from downweighting a uv cell using information about that cell. Thus,

$$\hat{C}^{\rm res}_{uv,ff'} = \sum_{\substack{\text{other } u',v' \\ \text{in annulus}}} \frac{(\tilde{x}^{\rm res}_{u'v'f} - \langle \tilde{x}^{\rm res}_f \rangle)(\tilde{x}^{\rm res}_{u'v'f'} - \langle \tilde{x}^{\rm res}_{f'} \rangle)^*}{N^{\rm LOS} - 2 - 1}, \tag{10}$$

where $\langle \tilde{x}^{\rm res}_f \rangle$ is an average over all u' and v' in the annulus. We expect this procedure to be particularly effective in our case because the uv coverage of the MWA after rotation synthesis is relatively symmetric.

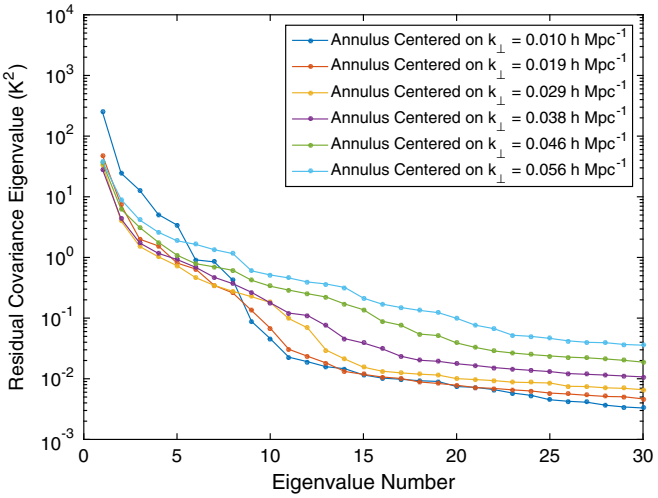


FIG. 1 (color online). The evolution of the wedge with $k_\perp$ motivates us to model foregrounds separately for discrete values of $k_\perp$. In this plot of the 30 largest eigenvalues of the observed residual covariance (which should include both noise and foregrounds) sampled in six concentric annuli, we see steeper declines toward a noise floor for the inner annuli than the outer annuli. This is consistent with the expected effect of the wedge—higher $k_\perp$ modes should be foreground contaminated at higher $k_\parallel$.

As a sense check on these covariances, we plot their largest 30 eigenvalues in Fig. 1. We see that as $|u|$ (and thus $k_\perp$) increases, the eigenspectra become shallower. At high $k_\perp$, the effect of the wedge is to leak power to a range of $k_\parallel$ values. The eigenspectrum of intrinsically smooth foregrounds should be declining exponentially [54]. The wedge softens that decline. These trends are in line with our expectations and further motivate our strategy of forming covariance matrices for each annulus independently.

Because we seek only to estimate the foreground portion of the covariance, the formal rank deficiency of $\hat{\mathbf{C}}^{\rm res}_{uv}$ is not a problem.[3] All we require is that the largest (and thus more foreground-dominated) modes be well measured. In this analysis, we used six concentric annuli to create six different frequency-frequency foreground covariances. Using more annuli allows for better modeling of the evolution of the wedge with $k_\perp$ at the expense of each

[2]For simplicity, we used the unitary discrete Fourier transform for these calculations and ignore any factors of length or inverse length that might come into these calculations only to be canceled out at a later step.

[3]In fact, the rank of $\hat{\mathbf{C}}^{\rm res}_{uv}$ is $N_{\rm LOS} - 3$ if $N^{\rm LOS} - 2 \leq n_f$.

estimate being more susceptible to noise and rank deficiency.

2. Subtract the properly projected noise covariance

The covariances computed from these uv lines of sight include contributions from the 21 cm signal and instrumental noise as well as foregrounds. We can safely ignore the signal covariance for now as we are far from the regime where sample variance is significant. We already have a theoretically motivated model for the noise (based on the uv sampling) that has been empirically rescaled to the observed noise in the difference of time-interleaved data (the same basic procedure as in [13]). We would like an empirical estimate of the residual foreground covariance alone to use in $\mathbf{C}^{FG}$ and thus must subtract off the part of our measurement we think is due to noise variance.

To get to $\hat{\mathbf{C}}^{FG}_{uv}$ from $\hat{\mathbf{C}}^{\rm res}_{uv}$, we subtract our best guess of the portion of $\hat{\mathbf{C}}^{\rm res}_{uv}$ that is due to noise, which we approximate by averaging the noise model variances in all the other uv cells in the annulus at that given frequency, yielding

$$\hat{C}^{\rm N}_{uv,ff'} = \frac{1}{N^{\rm LOS}} \sum_{\substack{\text{other } u',v' \\ \text{in annulus}}} \delta_{uu'}\delta_{vv'}\delta_{ff'} C^{N}_{uu'vv'ff'}. \tag{11}$$

Note, however, that $\hat{\mathbf{C}}^{\rm N}_{uv}$ is full rank while $\hat{\mathbf{C}}^{\rm res}_{uv}$ is typically rank deficient. Thus a naive subtraction would oversubtract the noise variance in the part of the subspace of $\hat{\mathbf{C}}^{\rm N}_{uv}$ where $\hat{\mathbf{C}}^{\rm res}_{uv}$ is identically zero. Instead, the proper procedure is to find the projection matrices Π_{uv} that discard all eigenmodes outside the subspace where $\hat{\mathbf{C}}^{\rm res}_{uv}$ is full rank. Each should have eigenvalues equal to zero or one only and have the property that

$$\Pi_{uv}\hat{\mathbf{C}}^{\rm res}_{uv}\Pi^{\dagger}_{uv} = \hat{\mathbf{C}}^{\rm res}_{uv}. \tag{12}$$

Only after projecting out the part of $\hat{\mathbf{C}}^{N}_{uv}$ inside the unsampled subspace can we self-consistently subtract our best guess of the noise contribution to the subspace in which we seek to estimate foregrounds. In other words, we estimate $\hat{\mathbf{C}}^{FG}_{uv}$ as

$$\hat{\mathbf{C}}^{FG}_{uv} = \hat{\mathbf{C}}^{\rm res}_{uv} - \Pi_{uv}\hat{\mathbf{C}}^{\rm N}_{uv}\Pi^{\dagger}_{uv}. \tag{13}$$

We demonstrate the effectiveness of this technique in Fig. 2 by plotting the diagonal elements of the Fourier transform of $\hat{\mathbf{C}}^{\rm res}_{uv}$ and $\hat{\mathbf{C}}^{FG}_{uv}$ along the line of sight. Subtracting the noise covariance indeed eliminates the majority of the power in the noise dominated modes at high $k_\parallel$; thus we expect it also to fare well in the transition region near the edge of the wedge where foreground and noise contributions are comparable.

3. Perform a $k_\parallel$ filter on the covariance

Despite the relatively clean separation of foreground and noise eigenvalues, inspection of some of the foreground-dominated modes in the top panel of Fig. 3 reveals residual

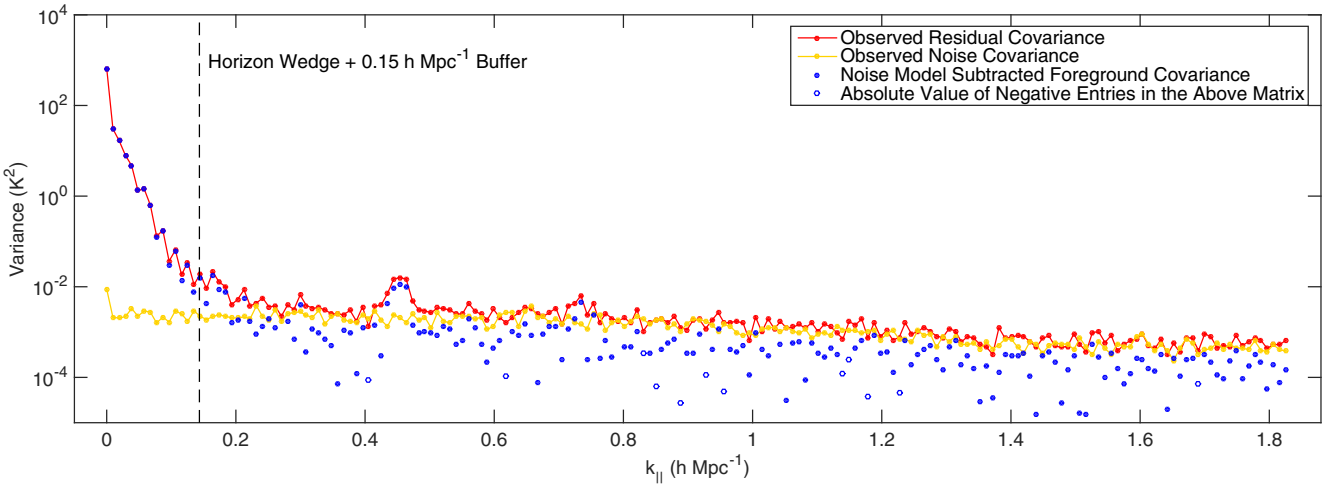


FIG. 2 (color online). Examining the diagonal elements of the observed residual and inferred foreground covariance matrices in Fourier space reveals the effectiveness of subtracting model for the noise covariance. In red, we plot the observed residual covariance, which contains both foregrounds and noise. As a function of $k_\parallel$, the two separate relatively cleanly—there is a steeply declining foreground portion on the left followed by a relatively flat noise floor on the right. The theory that the right-hand portion is dominated by noise is borne out by the fact that it so closely matches the observed noise covariance, inferred lines of sight of $\mathbf{x}_1 - \mathbf{x}_2$, which should have only noise and no sky signal at all. The regions where they differ significantly, for example at $k_\parallel \sim 0.45\ h\,\mathrm{Mpc}^{-1}$, are attributable to systematic effects like the MWA's coarse band structure that have not been perfectly calibrated out. For the example covariances shown here (which correspond to a mode in the annulus at $k_\perp \approx 0.010\ h\,\mathrm{Mpc}^{-1}$), we can see that subtracting a properly projected noise covariance removes most of the power from the noise-dominated region, leaving only residual noise that appears both as negative power (open blue circle) and as positive power (closed blue circles) at considerably lower magnitude.

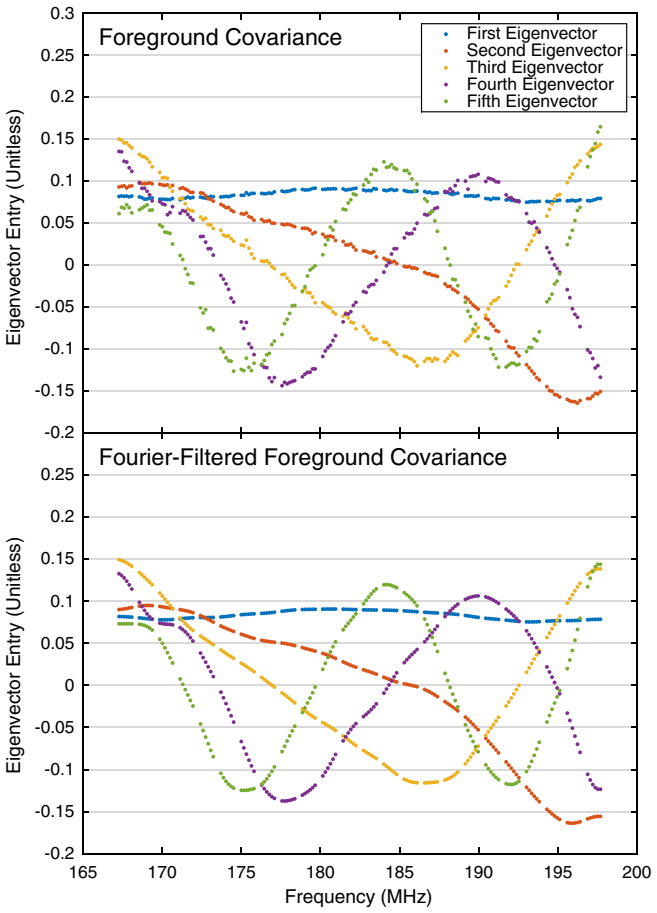


FIG. 3 (color online). The foreground covariance we estimate from our limited data set is still very noisy, and we run the risk of overfitting the noise in our measurements if we take it at face value. In the top panel, we plot the eigenvectors corresponding to the five largest eigenvalues of $\hat{\mathbf{C}}^{FG}$ for a mode in the annulus centered on $k_\perp \approx 0.010\ h\,\mathrm{Mpc}^{-1}$. In the bottom panel, we show dominant eigenvectors of the Fourier-filtered covariance. As expected, they resemble the first five Fourier modes. The missing data every 1.28 MHz are due to channels flagged at the edge of the coarse bandpass of the MWA's polyphase filter bank—the most difficult part of the band to calibrate.

noise. Using a foreground covariance constructed from these noisy foreground eigenmodes to downweight the data during power spectrum estimation would errantly downweight some high $k_\parallel$ modes in addition to the low $k_\parallel$ foreground-dominated modes. To avoid this double counting of the noise, we allow the foreground covariance to include only certain $k_\parallel$ modes by filtering $\hat{\mathbf{C}}^{FG}_{uv}$ in Fourier space to get $\hat{\mathbf{C}}^{FG,\mathrm{filtered}}_{uv}$. Put another way, we are imposing a prior on which Fourier modes we think have foreground power in them. The resulting noise filtered eigenmodes are shown in the bottom panel of Fig. 3.

In practice, implementing this filter is subtle. We interpolate $\hat{\mathbf{C}}^{FG}$ over the flagged frequency channels using a cubic spline, then symmetrically pad the covariance matrix, forcing its boundary condition to be periodic. We then Fourier transform, filter, inverse Fourier transform, remove the padding, and then rezero the flagged channels.

Selecting a filter to use is also a subtle choice. We first keep modes inside the horizon wedge with an added buffer. For each annulus, we calculate a mean value of $k_\perp$, and then use Eq. (1) to calculate the $k_\parallel$ value of the horizon wedge, using $\theta_0 = \pi/2$. Although the literature suggests a 0.1 to 0.15 $h\,\mathrm{Mpc}^{-1}$ buffer for "suprahorizon emission" due to some combination of intrinsic spectral structure of foregrounds, primary beam chromaticity, and finite bandwidth [26,55], we pick a conservative 0.5 $h\,\mathrm{Mpc}^{-1}$. Then we examine the diagonal of $\hat{\mathbf{C}}^{FG}$ (Fig. 2) to identify additional foreground modes, this time in the EoR window, due to imperfect bandpass calibration appearing as spikes. One example is the peak at $k_\parallel \sim 0.45\ h\,\mathrm{Mpc}^{-1}$. Such modes contribute errant power to the EoR window at constant $k_\parallel$. Since these modes result from the convolution of the foregrounds with our instrument, they also should be modeled in $\mathbf{C}^{FG}$ in order to minimize their leakage into the rest of the EoR window.

One might be concerned that cosmological signal and foregrounds theoretically both appear in the estimate of $\mathbf{C}^{FG}$ that we have constructed, especially with our conservative 0.5 $h\,\mathrm{Mpc}^{-1}$ buffer that allows foregrounds to be discovered well into the EoR window. For the purposes of calculating $\mathbf{C}^{-1}(\hat{\mathbf{x}} - \boldsymbol{\mu})$ in the quadratic estimator in Eq. (3), that is fine since its effect is to partially relax the assumption that sample variance can be ignored. However, the calculation of the bias depends on being able to differentiate signal from contaminants [40,41,43].

The noise contribution to the bias can be eliminated by cross-correlating maps made from interleaved time steps [13]. However, we cannot use our inferred $\mathbf{C}^{FG}$ to subtract a foreground bias without signal loss. That said, we can still set an upper limit on the 21 cm signal. By following the data and allowing the foreground covariance to have power inside the EoR window, we are minimizing the leakage of foregrounds into uncontaminated regions and we are accurately marking those regions as having high variance. As calibration and the control of systematic effects improves, we should be able to isolate foregrounds to outside the EoR window, impose a more aggressive Fourier filter on $\mathbf{C}^{FG}$, and make a detection of the 21 cm signal by employing foreground avoidance.

4. Cut out modes attributable to noise

After suppressing the noisiest modes with our Fourier filter, we must select a cutoff beyond which the foreground modes are irrecoverably buried under noise. We do this by inspecting the eigenspectrum of $\hat{\mathbf{C}}^{FG,\mathrm{filtered}}_{uv}$. The true $\mathbf{C}^{FG}$, by definition, admits only positive eigenvalues (though some of them should be vanishingly small).

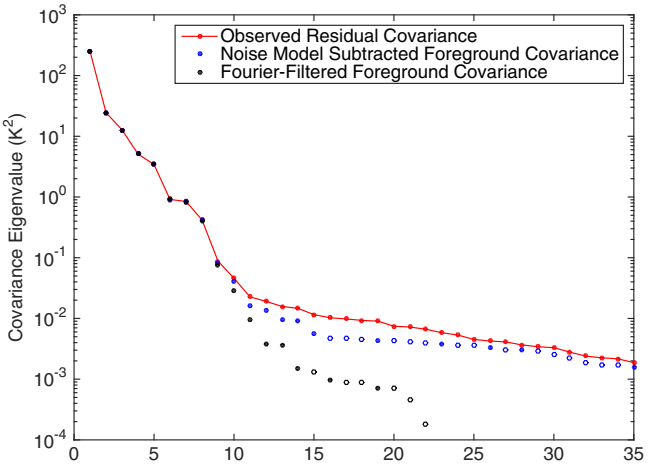


FIG. 4 (color online). The evolution of the eigenvalues of our estimated foreground covariance matrix for a mode in the annulus corresponding to $k_\perp \approx 0.010\ h\,\mathrm{Mpc}^{-1}$ at each of the first three stages of covariance estimation. First we calculate a sample covariance matrix from the residual data cubes (shown in red). Next we subtract our best guess as to the part of the diagonal of that matrix that originates from instrumental noise, leaving the blue dots (open circles are absolute values of negative eigenvalues). Then we filter out modes in Fourier space along the line of sight that we think should be noise dominated, leaving the black dots. Finally, we project out the eigenvectors associated with eigenvalues whose magnitude is smaller than the largest negative eigenvalue, since those are likely due to residual noise. What remains is our best guess at the foreground covariance in an annulus and incorporates as well as possible our prior beliefs about its structure.

By limiting the number of eigenvalues and eigenvectors we ultimately associate with foregrounds, we also limit the potential for signal loss by allowing a large portion of the free parameters to get absorbed into the contaminant model [15,56]. When measuring the power spectrum inside the EoR window, we can be confident that signal loss is minimal compared to foreground bias and other errors.

We plot in Fig. 4 the eigenspectra of $\hat{\mathbf{C}}_{uv}^{\mathrm{res}}$, $\hat{\mathbf{C}}_{uv}^{FG}$, and $\hat{\mathbf{C}}_{uv}^{FG,\mathrm{filtered}}$, sorted by absolute value. There are two distinct regions—the sharply declining foreground-dominated region and a flatter region with many negative eigenvalues. We excise eigenvectors whose eigenvalues are smaller in absolute value than the most negative eigenvalue. This incurs a slight risk of retaining a few noise dominated modes, albeit strongly suppressed by our noise variance subtraction and our Fourier filtering. Finally we are able to construct the full covariance $\hat{\mathbf{C}}$ using Eq. (8).

D. Review of assumptions and caveats

Before proceeding to demonstrate the effectiveness of our empirical covariance modeling method, it is useful to review and summarize the assumptions made about mapmaking and covariance modeling. Some are inherited from the previous application of quadratic power spectrum estimation to the MWA [13], while others are necessitated by our new, more faithful foreground covariance. Relaxing these assumptions in a computationally efficient manner remains a challenge we leave for future work.

(i) We adopt the flat sky approximation as in [13,43], allowing us to use the fast Fourier transform to quickly compute power spectra. The error incurred from this approximation on the power spectrum is expected to be smaller than 1% [13].

(ii) We assume the expectation value of our uniformly weighted map is the true sky (i.e., $\langle\hat{\mathbf{x}}\rangle = \mathbf{x}^{\mathrm{true}}$) when calculating $\mathbf{C}_{,\beta}$ in Eq. (3), again following [13]. In general $\langle\hat{\mathbf{x}}\rangle$ is related to $\mathbf{x}^{\mathrm{true}}$ by $\mathbf{P}$, the matrix of point spread functions [44]. Here we effectively approximate the PSF as position independent. Relaxing this approximation necessitates the full mapmaking theory presented in [44] which has yet to be integrated into a power spectrum estimation pipeline.

(iii) We approximate the foreground covariance as uncorrelated between different uv cells (and thus block diagonal). At some level there likely are correlations in uv, though those along the line of sight are far stronger. It may be possible to attempt to calculate these correlations empirically, but it would be very difficult considering relative strength of line-of-sight correlations. It may also be possible to use a nonempirical model, though that has the potential to make the computational speedups of [43] more difficult to attain.

(iv) We approximate the frequency-frequency foreground covariance as constant within each annulus, estimating our covariance for each uv cell only from other cells in the same annulus. In principle, even if the foreground residuals were isotropic, there should be radial evolution within each annulus which we ignore for this analysis.

(v) The Fourier filter is a nontrivial data analysis choice balancing risk of noise double counting against that of insufficiently aggressive foreground downweighting.

(vi) In order to detect the 21 cm signal, we assume that foregrounds can be avoided by working within the EoR window. Out of fear of losing signal, we make no effort to subtract a residual foreground bias from the window. This makes a detection inside the wedge impossible, and it risks confusing foreground contamination in the window for a signal. Only analysis of the dependence of the measurement on z, k, $k_\parallel$, and $k_\perp$ can distinguish between systematics and the true signal.

III. RESULTS

We can now demonstrate the statistical techniques we have motivated and developed in Sec. II on the problem of

estimating power spectra from a 3 h observation with the 128-antenna MWA. We begin with a discussion of the instrument and the observations in Sec. III A. In Sec. III B we detail the data processing from raw visibilities to calibrated maps from which we estimate both the foreground residual covariance matrix and the power spectrum. Finally, in Sec. III we present our results and discuss lessons learned looking toward a detection of the 21 cm signal.

A. Observation summary

The 128-antenna Murchison Widefield Array began deep EoR observations in mid-2013. We describe here the salient features of the array and refer to [10] for a more detailed description. The antennas are laid out over a region of radius 1.5 km in a quasirandom, centrally concentrated distribution which achieves approximately complete uv coverage at each frequency over several hours of rotation synthesis [57]. Each antenna element is a phased array of 16 wideband dipole antennas whose phased sum forms a discretely steerable 25° beams (full width at half maximum) at 150 MHz with frequency-dependent, percent level sidelobes [48]. We repoint the beam to our field center on a 30 min cadence to correct for earth rotation, effectively acquiring a series of drift scans over this field.

We observe the MWA "EOR0" deep integration field, centered at R.A.$(\mathrm{J2000}) = 0^{\mathrm{h}}0^{\mathrm{m}}0^{\mathrm{s}}$ and decl.$(\mathrm{J200}) = -30°0'0''$. It features a near-zenith position, a high Galactic latitude, minimal Galactic emission [58], and an absence of bright extended sources. This last property greatly facilitates calibration in comparison to the "EOR2" field—a field dominated by the slightly resolved radio galaxy Hydra A at its center—which was used by [59] and [13]. A nominal 3 h set of EOR0 observations was selected during the first weeks of observing to use for refining and comparing data processing, imaging, and power spectra pipelines [60]. In this work, we use the "high band," near-zenith subset of these observations with 30.72 MHz of bandwidth and center frequency of 182 MHz, recorded on Aug 23, 2013, between 16:47:28 and 19:56:32 UTC (22.712 and 1.872 h LST).

B. Calibration and mapmaking summary

Preliminary processing, including radio frequency interference (RFI) agging followed by time and frequency averaging, was performed with the COTTER package [61] on raw correlator data. These data were collected at 40 kHz resolution with an integration time of 0.5 s, and averaged to 80 kHz resolution with a 2 s integration time to reduce the data volume. Additionally, 80 kHz at the upper and lower edges of each of 24 coarse channels (each of width 1.28 MHz) of the polyphase filter bank is flagged due to known aliasing.

As in [13], we undertake snapshot-based processing in which each minute-scale integration is calibrated and imaged independently. Snapshots are combined only at the last step in forming a Stokes I image cube, allowing us to properly align and weight them despite different primary beams due to sky rotation and periodic repointing. While sources are forward modeled for calibration and foreground subtraction using the full position dependent PSF (i.e., the synthesized beam), we continue to approximate it as position independent (and equal to that of a point source at the field center) during application of uniform weighting and computation of the noise covariance.

We use the calibration, foreground modeling, and first stage image products produced by the Fast Holographic Deconvolution[4] (FHD) pipeline as described by [60]. The calibration implemented in the FHD package is an adaptation of the fast algorithm presented by [63] with a baseline cutoff of $b > 50\lambda$. In this data reduction, the point source catalogs discussed below are taken as the sky model for calibration. Solutions are first obtained per antenna and per frequency before being constrained to linear phase slopes and quadratic amplitude functions after correcting for a median antenna-independent amplitude bandpass. The foreground model used for subtraction includes models both of diffuse radio emission [64] and point sources. In detail, the point source catalog is the union of a deep MWA point source survey within 20° of the field center [65], the shallower but wider MWA commissioning point source survey [66], and the Culgoora catalog [67]. Note that calibration and foreground subtraction of off-zenith observations are complicated by Galactic emission picked up by primary beam sidelobes, and are active topics of investigation [38,39,68]. During these observations a single antenna was flagged due to known hardware problems, and 1–5 more were flagged for any given snapshot due to poor calibration solutions.

These calibration, foreground modeling, and imaging steps constitute notable improvements over [13]. In that work, the presence of the slightly resolved Hydra A in their EOR2 field likely limited calibration and subtraction fidelity as only a point source sky model was used. In contrast, the EOR0 field analyzed here lacks any such nearby radio sources. Our foreground model contains ∼2500 point sources within the main lobe and several thousand more in the primary beam sidelobes in addition to the aforementioned diffuse map. A last improvement in the imaging is the more frequent interleaving of time steps for the cross power spectrum, which we performed at the integration scale (2 s) as opposed to the snapshot scale (a few minutes). This ensures that both $\hat{\mathbf{x}}_1$ and $\hat{\mathbf{x}}_2$ have identical sky responses and thus allows us to accurately estimate the noise in the array from difference cubes. Assuming that the system temperature contains both an instrumental noise temperature and a frequency dependent sky noise temperature that scales as $\nu^{-2.55}$, the observed

[4]For a theoretical discussion of the algorithm see [62]. The code is available at https://github.com/miguelfmorales/FHD.

residual root-mean-square brightness temperature is consistent with $T_{\rm sys}$ ranging from 450 K at 167 MHz to 310 K at 198 MHz, in line with expectations [57].

As discussed in [60] and [69], FHD produces naturally weighted sky, foreground model, "weights," and "variances" cubes, as well as beam-squared cubes. All are saved in image space using the HEALPix format [70] with $N_{\rm side} = 1024$. Note that these image cubes are crops of full-sky image cubes to a $16° \times 16°$ square field of view, as discussed below. The sky, foreground model, and weights cubes are image space representations of the measured visibilities, model visibilities, and sampling function, respectively, all originally gridded in uv space using the primary beam as the gridding kernel. The variances cube is similar to the weights cube, except the gridding kernel is the square of the uv space primary beam. It represents the proper quadrature summation of independent noise in different visibilities when they contribute to the same uv cell, and will ultimately become our diagonal noise covariance model. The FHD cubes from all ninety-four 112 s snapshots are optimally combined in this "holographic" frame in which the true sky is weighted by two factors of the primary beam, as in [13].

We perform a series of steps to convert the image cube output of FHD into uniformly weighted Stokes I cubes accompanied by appropriate uv coverage information for our noise model. We first map these data cubes onto a rectilinear grid, invoking the flat sky approximation. We do this by rotating the (RA, Dec) HEALPix coordinates of the EOR0 field to the north pole (0°, 90°), and then projecting and gridding onto the xy plane with $0.2° \times 0.2°$ resolution over a $16° \times 16°$ square field of view. To reduce the data volume while maintaining cosmological sensitivity, we coarse grid to approximately 0.5° resolution by Fourier transforming and cropping these cubes in the uv plane at each frequency. We form a uniformly weighted Stokes I cube $I_{\rm uni}(\vec{\theta})$ by first summing the XX and YY data cubes, resulting in a naturally weighted, holographic Stokes I cube $I_{{\rm nat},h}(\vec{\theta}) = I_{XX,h}(\vec{\theta}) + I_{YY,h}(\vec{\theta})$. Then we divide out the holographic weights cube $W_h(\vec{\theta})$ in uv space, which applies uniform weighting and removes one image space factor of the beam, and lastly divide out the second beam factor $B(\vec{\theta})$: $I_{\rm uni}(\vec{\theta}) = \mathcal{F}^{-1}[\mathcal{F}I_{{\rm nat},h}(\vec{\theta})/\mathcal{F}W_h(\vec{\theta})]/B(\vec{\theta})$, where $\mathcal{F}$ represents a Fourier transform and $B(\vec{\theta}) = [B^2_{XX}(\vec{\theta}) + B^2_{YY}(\vec{\theta})]^{1/2}$. Consistent treatment of the variances cube requires uv space division of *two* factors of the weights cube followed by image space division of *two* factors of the beam.

Lastly, we frequency average from 80 kHz to 160 kHz, flagging a single 160 kHz channel the edge of each 1.28 MHz coarse channel due to polyphase filter bank attenuation and aliasing, which make these channels difficult to reliably calibrate. Following [13], we also flag poorly observed uv cells and uv cells whose observation times vary widely between frequencies. In all cases, we formally set the variance in flagged channels and uv cells in $\mathbf{C}^N$ to infinity and use the pseudoinverse to project out flagged modes [13].

C. Power spectrum results

We can now present the results of our method applied to 3 h of MWA-128T data. We first study cylindrically averaged, two-dimensional (2D) power spectra and their statistics, since they are useful for seeing the effects of foregrounds and systematic errors on the power spectrum. We form these power spectra with the full 30.72 MHz instrument bandwidth to achieve maximal $k_\parallel$ resolution.

We begin with the 2D power spectrum itself (Fig. 5) in which several important features can be observed. First, the

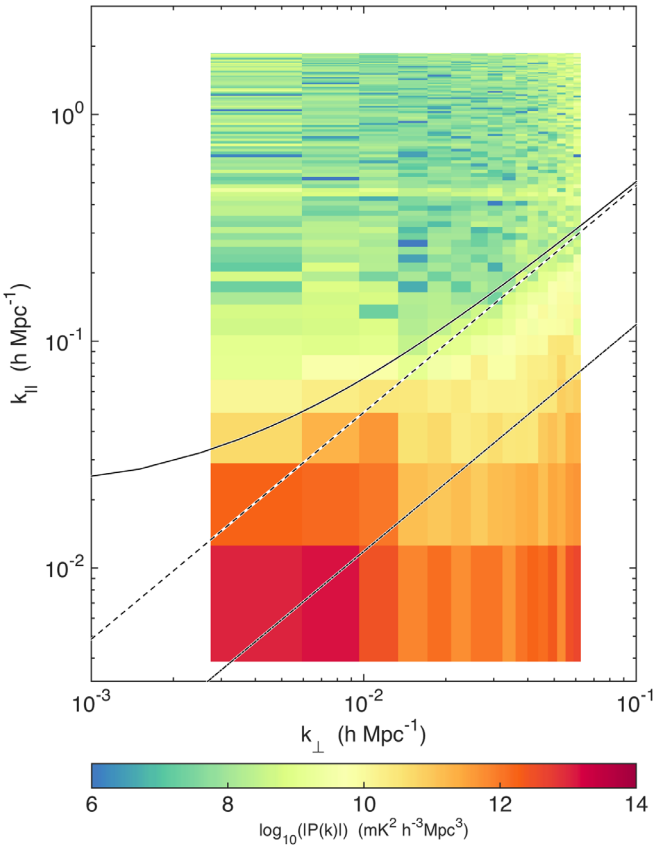


FIG. 5 (color online). Our power spectrum clearly exhibits the typical EoR window structure with orders-of-magnitude suppression of foregrounds in the EoR window. Here we plot our estimates for $|P(k_\perp, k_\parallel)|$ for the full instrumental bandwidth, equivalent to the range $z = 6.2$ to $z = 7.5$. Overplotted is the wedge from Eq. (1) corresponding to the first null in the primary beam (dash-dotted line), the horizon (dashed line), and the horizon with a relatively aggressive 0.02 $h\,{\rm Mpc}^{-1}$ buffer (solid curve). In addition to typical foreground structure, we also see the effect of noise at high and very low $k_\perp$ where baseline coverage is poor. We also clearly see a line of power at constant $k_\parallel \approx 0.45\ h\,{\rm Mpc}^{-1}$, attributable to miscalibration of the instrument's bandpass and cable reflections [69].

wedge and EoR window are clearly distinguishable, with foregrounds suppressed by at least 5 orders of magnitude across most of the EoR window. At high $k_\perp$, the edge of the wedge is set by the horizon while at low $k_\perp$ the cutoff is less clear. There appears to be some level of suprahorizon emission, which was also observed with PAPER in [26] and further explained by [35]. Consistent with Fig. 1 we see the strongest foreground residual power at low $k_\perp$, meaning that there still remains a very large contribution from diffuse emission from our Galaxy—potentially from sidelobes of the primary beam affecting the shortest baselines [39,68].

We also see evidence for less-than-ideal behavior. Through we identified spectral structure appearing at $k_\parallel \sim 0.45\ h\,\mathrm{Mpc}^{-1}$ in Fig. 2 and included it in our foreground residual covariance, that contamination still appears here as a horizontal line. By including it in the foreground residual model, we increase the variance we associate with those modes and we decrease the leakage out of those modes, isolating the effect to only a few $k_\parallel$ bins.

While Fig. 5 shows the magnitude of the 2D power spectrum, Fig. 6 shows its sign using a split color scale, providing another way to assess foreground contamination in the EoR window. Because we are taking the cross power spectrum between two cubes with identical sky signal but independent noise realizations, the noise dominated regions should be positive or negative with equal probability. This is made possible by our use of a power spectrum estimator normalized such that $\Sigma \equiv \mathrm{Cov}(\hat{\mathbf{p}})$ is a diagonal matrix [43]. This choice limits leakage of foreground residuals from the wedge into the EoR window [13].

By this metric, the EoR window is observed to be noise dominated with only two notable exceptions. The first is the region just outside the wedge at low $k_\perp$ attributable to suprahorizon emission due to some combination of intrinsic foreground spectral structure, beam chromaticity, and finite bandwidth. This suggests our aggressive $0.02\ h\,\mathrm{Mpc}^{-1}$ cut beyond the horizon will leave in some foreground contamination when we bin to form one-dimensional (1D) power spectra. As long as we are only claiming an upper limit on the power spectrum, this is fine. A detection of foregrounds is also an upper limit on the cosmological signal. More subtle is the line of positive power at $k_\parallel \sim 0.45\ h\,\mathrm{Mpc}^{-1}$, confirming our hypothesis that the spike observed in Fig. 5 is indeed an instrumental systematic since it behaves the same way in both time-interleaved data cubes. There is also a hint of a similar effect at $k_\parallel \sim 0.75\ h\,\mathrm{Mpc}^{-1}$, possibly visible in Fig. 2 as well. We attribute both to bandpass miscalibration due to cable reflections, complicated at these frequency scales by the imperfect channelization of the MWA's two-stage polyphase filter, as well as slight antenna dependence of the bandpass due to cable length variation [69].

Additionally, the quadratic estimator formalism relates our covariance models of residual foregrounds and noise to the expected variance in each band power [13,40,43], which we plot in Fig. 7. As we have chosen our power spectrum normalization $\mathbf{M}$ such that $\Sigma \equiv \mathrm{Cov}(\hat{\mathbf{p}})$ is diagonal, it is sufficient to plot the diagonal of $\Sigma^{1/2}$, the standard deviation of each band power. The EoR window is seen clearly here as well. There is high variance at low and high $k_\perp$ where the uv coverage is poor, and also in the wedge due to foreground residuals. It is particularly pronounced in the bottom left corner, which is dominated by residual diffuse foregrounds.

As our error covariance represents the error due to both noise and foregrounds we expect to make in an estimate of the 21 cm signal, it is interesting to examine the "signal to error ratio" in Fig. 8—the ratio of Fig. 5 to Fig. 7. The ratio

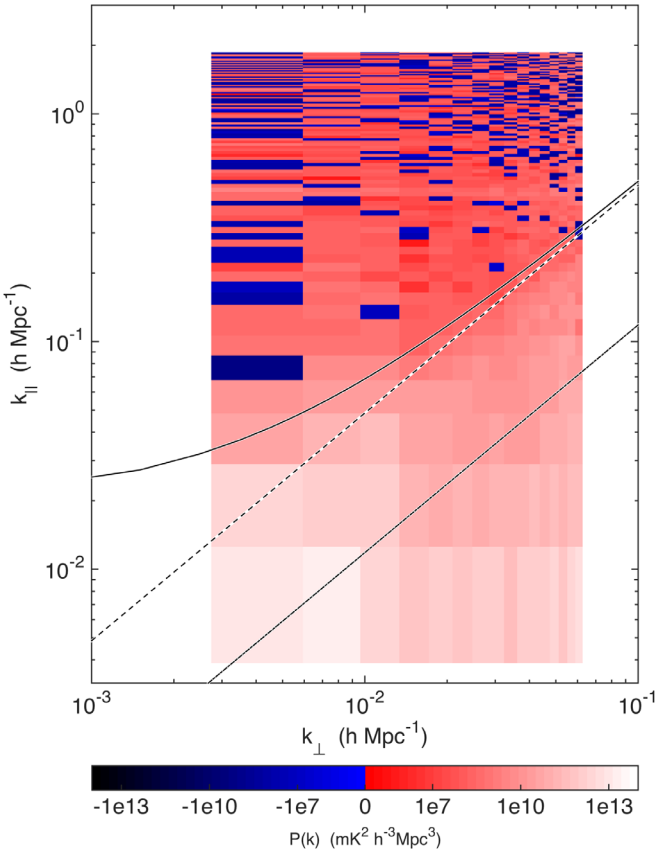


FIG. 6 (color online). By using an estimator of the power spectrum with uncorrelated errors between bins, we can see that most of the EoR window is noise dominated in our power spectum measurement. Here we show the inverse hyperbolic sine of the power spectum, which behaves linearly near zero and logarithmically at large magnitudes. Because we are taking a cross power spectrum between two data cubes with uncorrelated noise, noise dominated regions are equally likely to have positive power as negative power. Since we do not attempt to subtract a foreground bias, foreground contaminated regions show up as strongly positive. That includes the wedge, the bandpass line at $k_\parallel \approx 0.45\ h\,\mathrm{Mpc}^{-1}$ (see Fig. 5), and some of the EoR window at low $k_\perp$ and relatively low $k_\parallel$, consistent with the suprahorizon emission seen in [26].

is of order unity in noise dominated regions—though it is slightly lower than what we might naively expect due to our conservative estimate of Σ [13]. That explains the number of modes with very small values in Fig. 7. In the wedge and just above it, however, the missubtracted foreground bias is clear, appearing as a high significance "detection" of the foreground wedge in the residual foregrounds. The bandpass miscalibration line at $k_\parallel \sim 0.45\ h\,\mathrm{Mpc}^{-1}$ also appears clearly due to both foreground bias and possibly an underestimation of the errors. Hedging against this concern, we simply project out this line from our estimator that bins 2D power spectra into 1D power spectra by setting the variance of those bins to infinity.

Though useful for the careful evaluation of our techniques and of the instrument, the large bandwidth data cubes used to make Figs. 5 and 6 encompass long periods of cosmic time over which the 21 cm power spectrum is expected to evolve. The cutoff is usually taken to be $\Delta z \lesssim 0.5$ [5]. These large data cubes also violate the assumption in [43] that channels of equal width in frequency correspond to equal comoving distances, justifying the use of the fast Fourier transform. Therefore, we break the full bandwidth into three 10.24 MHz segments before forming spherically averaged power spectra, and estimate the foreground residual covariance and power spectrum independently from each. We bin our 2D power spectra into 1D power spectra using the optimal estimator formalism of [13]. In our case, since we have chosen $\mathbf{M}$ such that Σ is

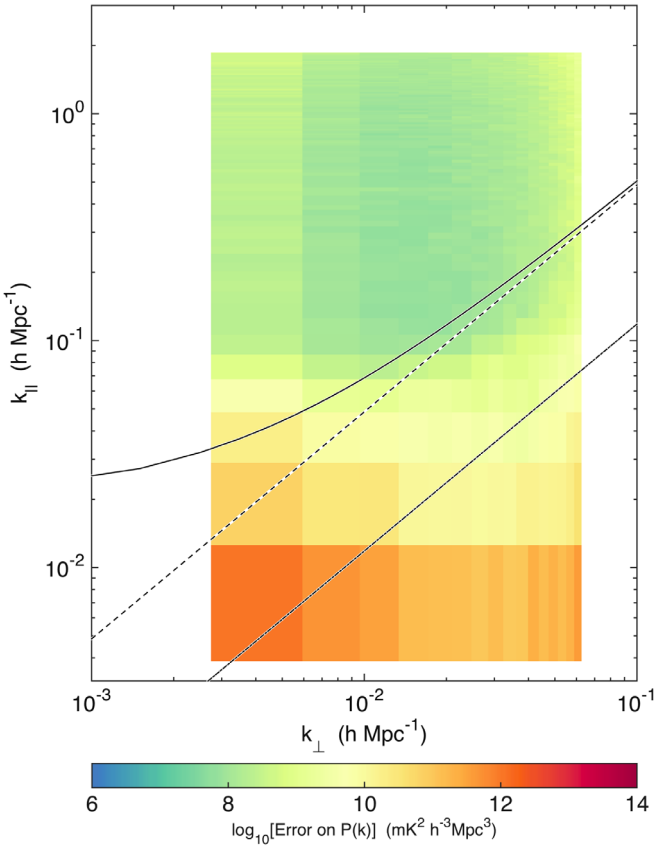


FIG. 7 (color online). By including both residual foregrounds and noise in $\mathbf{C}$, our model for the covariance, we can calculate the expected variance on each band power in $\hat{\mathbf{p}}$, which we show here. We see more variance at high (and also very low) $k_\perp$ where we have few baselines. We also see high variance at low $k_\parallel$ consistent with foregrounds. We see the strongest foregrounds at low $k_\perp$, which implies that the residual foregrounds have a very strong diffuse component that we have much to gain from better diffuse models to subtract. We also see that foreground-associated variance extends to higher $k_\parallel$ at high $k_\perp$, which is exactly the expected effect from the wedge. Both these observations are consistent with the structure of the eigenmodes we saw in Fig. 1. Because we have chosen a normalization of $\hat{\mathbf{p}}$ such that the $\mathrm{Cov}(\hat{\mathbf{p}})$ is diagonal, this is a complete description of our errors. Furthermore, it means that the band powers form a mutually exclusive and collectively exhaustive set of measurements.

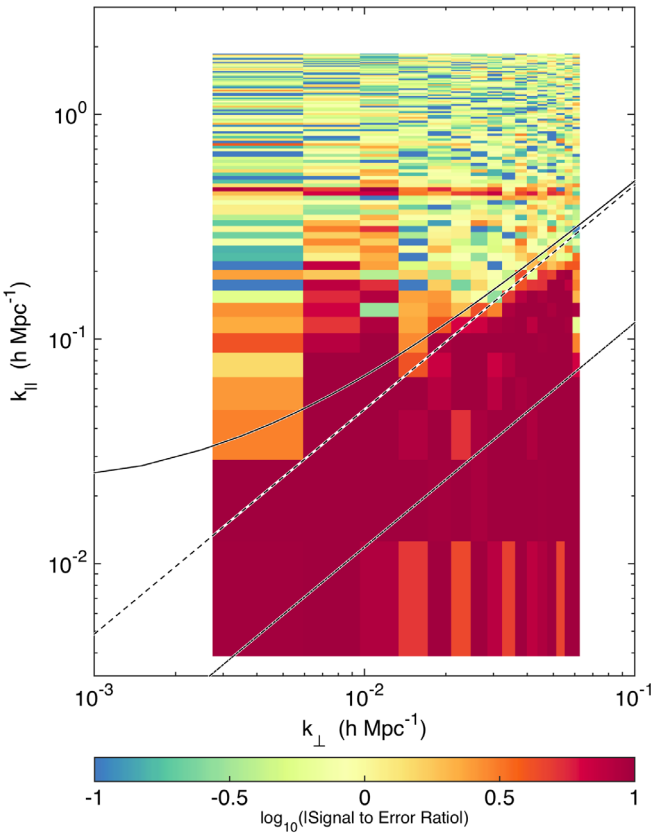


FIG. 8 (color online). The foregrounds' wedge structure is particularly clear when looking at the ratio of our measured power spectrum to the modeled variance, shown here. Though the variance in foreground residual dominated parts of the $k_\perp$-$k_\parallel$ plane are elevated (see Fig. 7), we still expect regions with signal to error ratios greater than one. This is largely due to the fact that we choose not to subtract a foreground bias for fear of signal loss. This figure shows us most clearly where the foregrounds are important and, as with Fig. 6, it shows where we can hope to do better with more integration time and where we need better calibration and foreground modeling to further integrate down.

diagonal, this reduces to simple inverse variance weighting with the variance on modes outside the EoR window or in the $k_\parallel \sim 0.45\ h\,\mathrm{Mpc}^{-1}$ line set to infinity.

In Fig. 9 we show the result of that calculation as a "dimensionless" power spectra $\Delta^2(k) \equiv k^3 P(k)/2\pi^2$. We choose our binning such that the window functions (calculated as in [13] from our covariance model) were slightly overlapping.

Our results are largely consistent with noise. Since noise is independent of $k_\parallel$ and $k \approx k_\parallel$ for most modes we measure, the noise in $\Delta^2(k)$ scales as k^3. We see deviations from that trend at low k where modes are dominated by residual foreground emission beyond the horizon wedge and thus show elevated variance and bias in comparison to modes at higher k. Since we do not subtract a bias, even these "detections" are upper limits on the cosmological signal.

A number of barely significant detections are observed at higher k. Though we excise bins associated with the $k_\parallel \sim 0.45\ h\,\mathrm{Mpc}^{-1}$ line, the slight detections may be due to leakage from that line. At higher z, the feature may be due to reflections from cables of a different length, though some may be plausibly attributable to noise. Deeper integration is required to investigate further.

Our best upper limit at 95% confidence is $\Delta^2(k) < 3.7 \times 10^4\ \mathrm{mK}^2$ at $k = 0.18\ h\,\mathrm{Mpc}^{-1}$ around $z = 6.8$. Our absolute lowest limit is about 2 times lower than the best limit in [13], though the latter was obtained at substantially higher redshift and lower k, making the two somewhat incomparable. Our best limit is roughly 3 orders of magnitude better than the best limit of [13] over the same redshift range, and the overall noise level (as measured by the part of the power spectrum that scales as k^3) is more than 2 orders of magnitude smaller. This cannot be explained by more antenna tiles alone; it is likely that the noise level was overestimated in [13] due to insufficiently rapid time interleaving of the data cubes used to infer the overall noise level.

Although one cannot directly compare limits at different values of k and z, our limit is similar to the GMRT limit [8], $6.2 \times 10^4\ \mathrm{mK}^2$ at $k = 0.50\ h\,\mathrm{Mpc}^{-1}$ and $z = 8.6$ with 40 h of observation, and remains higher than the best PAPER

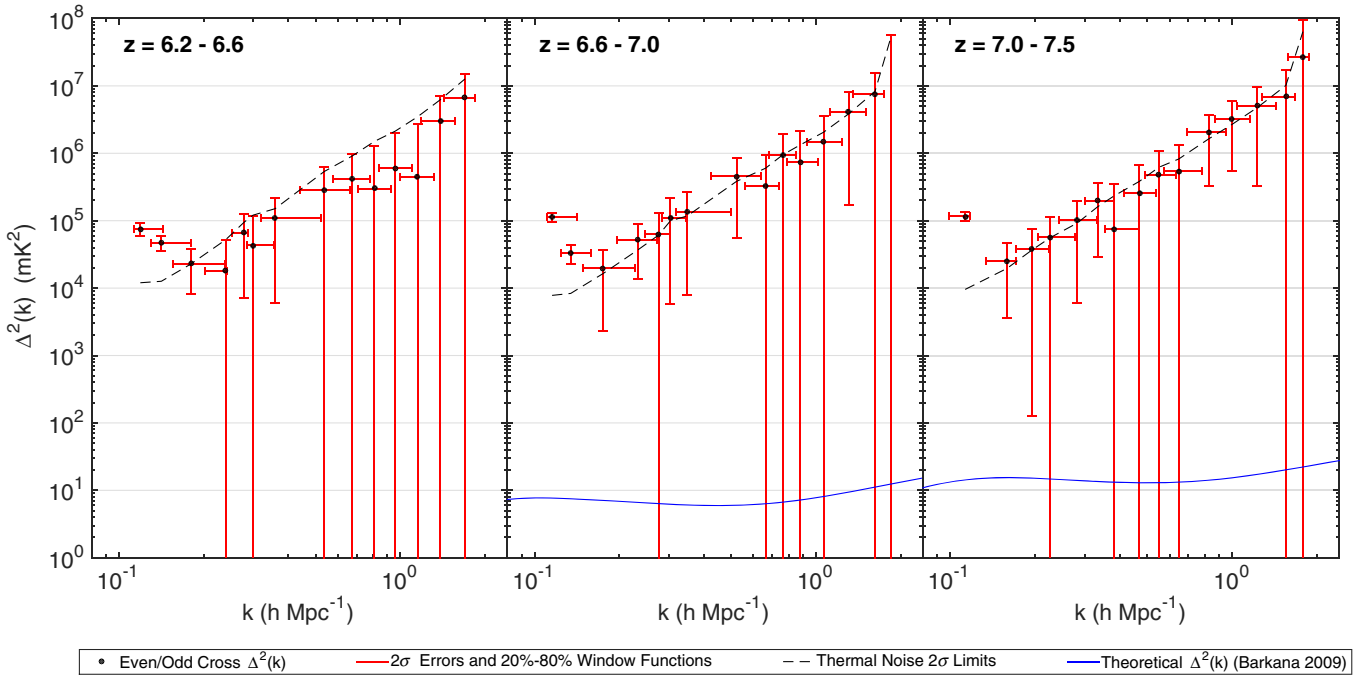


FIG. 9 (color online). Finally, we can set confident limits on the 21 cm power spectrum at three redshifts by splitting our simultaneous bandwidth into three 10.24 MHz data cubes. The lowest k bins show the strongest "detections," though they are attributable to suprahorizon emission [26] that we expect to appear because we only cut out the wedge and a small buffer ($0.02\ h\,\mathrm{Mpc}^{-1}$) past it. We also see marginal "detections" at higher k which are likely due to subtle bandpass calibration effects like cable reflections. The largest such error, which occurs at bins around $k_\parallel \sim 0.45\ h\,\mathrm{Mpc}^{-1}$ and can be seen most clearly in Fig. 8, has been flagged and removed from all three of these plots. Our absolute lowest limit requires $\Delta^2(k) < 3.7 \times 10^4\ \mathrm{mK}^2$ at 95% confidence at comoving scale $k = 0.18\ h\,\mathrm{Mpc}^{-1}$ and $z = 6.8$, which is consistent with published limits [8,12–15]. We also include a simplistic thermal noise calculation (dashed line), based on our observed system temperature. Though it is not directly comparable to our measurements, since it has different window functions, it does show that most of our measurements are consistent with thermal noise. For comparison, we also show the theoretical model of [71] (which predicts that reionization ends before $z = 6.4$) at the central redshift of each bin. While we are still orders of magnitude away from the fiducial model, recall that the noise in the power spectrum scales inversely with the integration time, not the square root.

limit [15] of 502 mK^2 between $k = 0.15\ h\,\mathrm{Mpc}^{-1}$ and $k = 0.50\ h\,\mathrm{Mpc}^{-1}$ and $z = 8.4$ with 4.5 months of observation.

In Fig. 9 we also plot a theoretical model from [71] predicting that reionization has ended by the lowest redshift bin we measure. We remain more than 3 orders of magnitude (in mK^2) from being able to detect that particular reionization model, naively indicating that roughly 3000 h of data are required for its detection. This appears much larger than what previous sensitivity estimates have predicted for the MWA (e.g. [57]) in the case of idealized foreground subtraction.

However, much of this variance is due to the residual foregrounds and systematics in the EoR window identified by our empirical covariance modeling method, not thermal noise (see Fig. 7). More integration will not improve those modes unless it allows for a better understanding of our instrument, better calibration, and better foreground models—especially of diffuse emission which might contaminate the highly sensitive bottom left corner of the EoR window. Eliminating this apparent "suprahorizon" emission, seen most clearly as detections in Fig. 8 below $k \approx 0.2\ h\,\mathrm{Mpc}^{-1}$, is essential to achieving the forecast sensitivity of the MWA [57]. If we can do so, we may still be able to detect the EoR with 1000 h or fewer. This is especially true if we can improve the subtraction of foregrounds to the point where we can work within the wedge, which can vastly increase the sensitivity of the instrument [55,57]. On the other hand, more data may reveal more systematics lurking beneath the noise which could further diminish our sensitivity.

IV. SUMMARY AND FUTURE DIRECTIONS

In this work, we developed and demonstrated a method for empirically deriving the covariance of residual foreground contamination, $\mathbf{C}^{FG}$, in observations designed to measure the 21 cm cosmological signal. Understanding the statistics of residual foregrounds allows us to use the quadratic estimator formalism to quantify the error associated with missubtracted foregrounds and their leakage into the rest of the EoR window. Because of the complicated interaction between the instrument and the foregrounds, we know that the residual foregrounds will have complicated spectral structure, especially if the instrument is not perfectly calibrated. By deriving our model for $\mathbf{C}^{FG}$ empirically, we could capture those effects faithfully and thus mitigate the effects of foregrounds in our measurement (subject to certain caveats which we recounted in Sec. II D).

Our strategy originated from the assumption that the frequency-frequency covariance, modeled as a function of $|u|$, is the most important component of the foreground residual covariance. We therefore used sample covariances taken in annuli in Fourier space as the starting point of our covariance model. These models were adjusted to avoid double counting the noise variance and filtered in Fourier space to minimize the effect of noise in the empirically estimated covariances. Put another way, we combined our prior beliefs about the structure of the residual foregrounds with their observed statistics in order to build our models.

We demonstrated this strategy through the power spectrum analysis of a 3 h preliminary MWA data set. We saw the expected wedge structure in both our power spectra and our variances. We saw that most of the EoR window was consistent with noise, and we understand why residual foregrounds and systematics affect the regions that they do. We were also able to set new MWA limits on the 21 cm power spectrum from $z = 6.2$ to 7.5, with an absolute best 95% confidence limit of $\Delta^2(k) < 3.7 \times 10^4\ \mathrm{mK}^2$ at $k = 0.18\ h\,\mathrm{Mpc}^{-1}$ and $z = 6.8$, consistent with published limits [12,14].

This work suggests a number of avenues for future research. Of course, improved calibration and mapmaking fidelity—especially better maps of diffuse Galactic structure—will improve power spectrum estimates and allow deeper integrations without running up against foregrounds or systematics. Relaxing some of the mapmaking and power spectrum assumptions discussed in Sec. II D may further mitigate these effects. A starting point is to integrate the mapmaking and statistical techniques of [44] with the fast algorithms of [43]. The present work is based on the idea that it is simpler to estimate $\mathbf{C}^{FG}$ from the data than from models of the instrument and the foregrounds. But if we can eliminate systematics to the point where we really understand $\mathbf{P}$, the relationship between the true sky and our dirty maps, then perhaps we can refocus our residual foreground covariance modeling effort on the statistics of the true sky residuals using the fact that $\mathbf{C}^{FG} = \mathbf{P}\mathbf{C}^{FG,\mathrm{true}}\mathbf{P}^{\mathsf{T}}$. Obtaining such a complete understanding of the instrument will be challenging, but it may be the most rigorous way to quantify the errors introduced by missubtracted foregrounds and thus to confidently detect the 21 cm power spectrum from the epoch of reionization.

ACKNOWLEDGMENTS

We would like to thank Adrian Liu, Aaron Parsons, and Jeff Zheng for helpful discussions. We also acknowledge an anonymous referee whose insightful questions led to the refinement of method in order to a bias that occurs when a uv cell is used to estimate its own covariance and the ultimate form of Eq. (10). We would also like to thank Rennan Barkana for the theoretical power spectra in Fig. 9. This work was supported by NSF Grants No. AST-0457585, No. AST-0821321, No. AST-1105835, No. AST-1410719, No. AST-1410484, No. AST-1411622, and No. AST-1440343, by the MIT School of Science, by the Marble Astrophysics Fund, and by generous donations from Jonathan Rothberg and an anonymous donor. D. C. J. would like to acknowledge NSF support under Grant No. AST-1401708. This scientific

work makes use of the Murchison Radio-astronomy Observatory, operated by CSIRO. We acknowledge the Wajarri Yamatji people as the traditional owners of the Observatory site. Support for the MWA comes from the U.S. National Science Foundation (Grants No. AST-0457585, No. PHY-0835713, No. CAREER-0847753, and No. AST-0908884), the Australian Research Council (LIEF Grants No. LE0775621 and No. LE0882938), the U.S. Air Force Office of Scientific Research (Grant No. FA9550-0510247), and the Centre for All-sky Astrophysics (an Australian Research Council Centre of Excellence funded by Grant No. CE110001020). Support is also provided by the Smithsonian Astrophysical Observatory, the MIT School of Science, the Raman Research Institute, the Australian National University, and the Victoria University of Wellington (via Grant No. MED-E1799 from the New Zealand Ministry of Economic Development and an IBM Shared University Research Grant). The Australian Federal government provides additional support via the Commonwealth Scientific and Industrial Research Organisation (CSIRO), National Collaborative Research Infrastructure Strategy, Education Investment Fund, and the Australia India Strategic Research Fund, and Astronomy Australia Limited, under contract to Curtin University. We acknowledge the iVEC Petabyte Data Store, the Initiative in Innovative Computing, and the CUDA Center for Excellence sponsored by NVIDIA at Harvard University, and the International Centre for Radio Astronomy Research (ICRAR), a Joint Venture of Curtin University and The University of Western Australia, funded by the Western Australian State government.

[1] S. R. Furlanetto, S. P. Oh, and F. H. Briggs, Phys. Rep. **433**, 181 (2006).
[2] M. F. Morales and J. S. B. Wyithe, Annu. Rev. Astron. Astrophys. **48**, 127 (2010).
[3] J. R. Pritchard and A. Loeb, Rep. Prog. Phys. **75**, 086901 (2012).
[4] A. Loeb and S. R. Furlanetto, *The First Galaxies in the Universe* (Princeton University Press, Princeton, NJ, 2013).
[5] Y. Mao, M. Tegmark, M. McQuinn, M. Zaldarriaga, and O. Zahn, Phys. Rev. D **78**, 023529 (2008).
[6] M. A. Garrett, arXiv:0909.3147.
[7] A. R. Parsons *et al.*, Astrophys. J. **139**, 1468 (2010).
[8] G. Paciga *et al.*, Mon. Not. R. Astron. Soc. **433**, 639 (2013).
[9] C. J. Lonsdale *et al.*, Proc. IEEE **97**, 1497 (2009).
[10] S. J. Tingay *et al.*, Pub. Astron. Soc. Aust. **30**, e007 (2013).
[11] J. D. Bowman *et al.*, Pub. Astron. Soc. Aust. **30**, e031 (2013).
[12] D. C. Jacobs *et al.*, Astrophys. J. **801**, 51 (2015).
[13] J. S. Dillon *et al.*, Phys. Rev. D **89**, 023002 (2014).
[14] A. R. Parsons *et al.*, Astrophys. J. **788**, 106 (2014).
[15] Z. S. Ali *et al.*, arXiv:1502.06016.
[16] J. C. Pober *et al.*, arXiv:1503.00045.
[17] M. F. Morales, Astrophys. J. **619**, 678 (2005).
[18] J. D. Bowman, M. F. Morales, and J. N. Hewitt, Astrophys. J. **638**, 20 (2006).
[19] A. Lidz, O. Zahn, M. McQuinn, M. Zaldarriaga, and L. Hernquist, Astrophys. J. **680**, 962 (2008).
[20] G. Harker *et al.*, Mon. Not. R. Astron. Soc. **405**, 2492 (2010).
[21] A. Parsons, J. Pober, M. McQuinn, D. Jacobs, and J. Aguirre, Astrophys. J. **753**, 81 (2012).
[22] N. Thyagarajan *et al.*, Astrophys. J. **776**, 6 (2013).
[23] A. de Oliveira-Costa, M. Tegmark, B. M. Gaensler, J. Jonas, T. L. Landecker, and P. Reich, Mon. Not. R. Astron. Soc. **388**, 247 (2008).
[24] V. Jelić *et al.*, Mon. Not. R. Astron. Soc. **389**, 1319 (2008).
[25] G. Bernardi *et al.*, Astron. Astrophys. **500**, 965 (2009).
[26] J. C. Pober *et al.*, Astrophys. J. **768**, L36 (2013).
[27] S. Yatawatta *et al.*, Astron. Astrophys. **550**, A136 (2013).
[28] V. Jelić *et al.*, Astron. Astrophys. **568**, A101 (2014).
[29] A. Datta, J. D. Bowman, and C. L. Carilli, Astrophys. J. **724**, 526 (2010).
[30] A. R. Parsons, J. C. Pober, J. E. Aguirre, C. L. Carilli, D. C. Jacobs, and D. F. Moore, Astrophys. J. **756**, 165 (2012).
[31] H. Vedantham, N. Udaya Shankar, and R. Subrahmanyan, Astrophys. J. **745**, 176 (2012).
[32] M. F. Morales, B. Hazelton, I. Sullivan, and A. Beardsley, Astrophys. J. **752**, 137 (2012).
[33] B. J. Hazelton, M. F. Morales, and I. S. Sullivan, Astrophys. J. **770**, 156 (2013).
[34] C. M. Trott, R. B. Wayth, and S. J. Tingay, Astrophys. J. **757**, 101 (2012).
[35] A. Liu, A. R. Parsons, and C. M. Trott, Phys. Rev. D **90**, 023018 (2014).
[36] A. Liu, A. R. Parsons, and C. M. Trott, Phys. Rev. D **90**, 023019 (2014).
[37] P. A. R. Ade *et al.* (Planck Collaboration), Astron. Astrophys. **571**, A16 (2014).
[38] J. Pober *et al.* (to be published).
[39] N. Thyagarajan *et al.*, Astrophys. J. **804**, 14 (2015).
[40] A. Liu and M. Tegmark, Phys. Rev. D **83**, 103006 (2011).
[41] M. Tegmark, Phys. Rev. D **55**, 5895 (1997).
[42] M. Tegmark, A. J. S. Hamilton, M. A. Strauss, M. S. Vogeley, and A. S. Szalay, Astrophys. J. **499**, 555 (1998).
[43] J. S. Dillon, A. Liu, and M. Tegmark, Phys. Rev. D **87**, 043005 (2013).
[44] J. S. Dillon, M. Tegmark, A. Liu, A. Ewall-Wice, J. N. Hewitt, M. F. Morales, A. R. Neben, A. R. Parsons, and H. Zheng, Phys. Rev. D **91**, 023002 (2015).
[45] J. R. Shaw, K. Sigurdson, U.-L. Pen, A. Stebbins, and M. Sitwell, Astrophys. J. **781**, 57 (2014).
[46] J. R. Shaw, K. Sigurdson, M. Sitwell, A. Stebbins, and U.-L. Pen, Phys. Rev. D **91**, 083514 (2015).

[47] A. Sutinjo, J. O'Sullivan, E. Lenc, R. B. Wayth, S. Padhi, P. Hall, and S. J. Tingay, arXiv:1412.4466.
[48] A. Neben *et al.*, arXiv:1505.07114.
[49] L. B. Newburgh *et al.*, Proc. SPIE Int. Soc. Opt. Eng. **9145**, 4 (2014).
[50] J. C. Pober, A. R. Parsons, D. C. Jacobs, J. E. Aguirre, R. F. Bradley, C. L. Carilli, N. E. Gugliucci, D. F. Moore, and C. R. Parashare, Astrophys. J. **143**, 53 (2012).
[51] P. M. Sutter, B. D. Wandelt, J. D. McEwen, E. F. Bunn, A. Karakci, A. Korotkov, P. Timbie, G. S. Tucker, and L. Zhang, Mon. Not. R. Astron. Soc. **438**, 768 (2014).
[52] P. M. Sutter, B. D. Wandelt, and S. S. Malu, Astrophys. J. Suppl. Ser. **202**, 9 (2012).
[53] A. Liu, Ph. D. thesis, Massachusetts Institute of Technology, 2012.
[54] A. Liu and M. Tegmark, Mon. Not. R. Astron. Soc. **419**, 3491 (2012).
[55] J. C. Pober *et al.*, Astrophys. J. **782**, 66 (2014).
[56] E. R. Switzer and A. Liu, Astrophys. J. **793**, 102 (2014).
[57] A. P. Beardsley *et al.*, Mon. Not. R. Astron. Soc. **429**, 5 (2013).
[58] A. de Oliveira-Costa, M. Tegmark, B. M. Gaensler, J. Jonas, T. L. Landecker, and P. Reich, Mon. Not. R. Astron. Soc. **388**, 247 (2008).
[59] C. L. Williams *et al.*, Astrophys. J. **755**, 47 (2012).
[60] D. Jacobs *et al.* (to be published).
[61] A. R. Offringa *et al.*, Pub. Astron. Soc. Aust. **32**, e008 (2015).
[62] I. S. Sullivan *et al.*, Astrophys. J. **759**, 17 (2012).
[63] S. Salvini and S. J. Wijnholds, Astron. Astrophys. **571**, A97 (2014).
[64] A. Beardsley *et al.* (to be published).
[65] P. Carroll *et al.* (to be published).
[66] N. Hurley-Walker, Pub. Astron. Soc. Aust. **31**, e045 (2014).
[67] O. B. Slee, Aust. J. Phys. **48**, 143 (1995).
[68] N. Thyagarajan *et al.* (to be published).
[69] P. Carroll *et al.* (to be published).
[70] K. M. Górski, E. Hivon, A. J. Banday, B. D. Wandelt, F. K. Hansen, M. Reinecke, and M. Bartelmann, Astrophys. J. **622**, 759 (2005).
[71] R. Barkana, Mon. Not. R. Astron. Soc. **397**, 1454 (2009).