Draft version November 17, 2016
Preprint typeset using LaTeX style emulateapj v. 12/16/11

PAPER-64 III: INFORMATION LOSS IN OPTIMAL QUADRATIC ESTIMATION OF THE 21CM POWER SPECTRUM

Matthew J. Kolopanis[1,2], Daniel C. Jacobs[1], Carina Cheng[3], Aaron R. Parsons[3,4], Zaki S. Ali[3], Haoxuan Zheng[5], Jonathan C. Pober[6], Adrian Liu[3,7], James E. Aguirre[8], Richard F. Bradley[9,10,11], Gianni Bernardi[12,13,14], Chris L. Carilli[15,16], David R. DeBoer[4], Matthew R. Dexter[4], Jasper Grobbelaar[12], Jasper Horrell[12], Judd Bowman[1], Pat Klima[10], David H. E. MacMahon[4], Matthys Maree[12], David F. Moore[8], Nima Razavi[16], Irina I. Stefan[16], William P. Walbrugh[12], Andre Walker[12]

Draft version November 17, 2016

ABSTRACT

The use of Optimal Quadratic Estimators in power spectrum estimation is a popular choice among recent radio arrays. Properly modeling and accounting for covariance among measurement channels allows for lossless power spectrum estimation with this technique; however the use of empirically estimated covariance matrices has been shown to cause over fitting of noise-like terms and as a result information loss. We show previous investigations have not properly identified and characterized this effect and offer a new method to account for and overcome information loss in power spectrum estimation. The application of this method to PAPER data is presented in our accompanying paper Kolopanis et al (in prep).

1. INTRODUCTION

The anticipated detection of the 21cm power spectrum will provide new information about the evolution of cosmic hydrogen through the dark ages and the first types of luminous bodies in our universe. Many radio experiments are attempting to detect this signal include the Precision Array for Probing the Epoch of Reionizaiton (PAPER[1]; Parsons et al. 2010), the Giant Metre-wave Radio Telescope (GMRT; Paciga et al. 2013), the Low Frequency Array (LOFAR [2]; Yatawatta et al. 2013), the Murchison Widefield Array (MWA[3]; Bowman et al. 2013), and the upcoming Hydrogen Epoch of Reionization Array (HERA[4]; Pober et al. 2014;).

The theoretical predictions for the foregrounds faced by these telescopes can be found in Morales & Wyithe (2010) and Pritchard & Loeb (2012). A review of the physics of reionization is found in Furlanetto et al. (2006).

A characteristic region of Fourier space is expected to be free of Foreground signals and ideal for EoR analysis. This "EoR window" consists of a region where spectrally smooth foregrounds do not contaminate the otherwise isotropic cosmological 21cm signal(Bowman et al. 2009; Morales et al. 2006).

The EoR Window is accompanied by an area in Fourier space of foreground leakage has been identified as a result of instrumental mode mixing (Parsons et al. 2012). This leakage sends some spectrally smooth foregrounds to high Fourier modes resulting in a "wedge" like shape (Pober et al. 2013; Thyagarajan et al. 2013, 2015a,b; Trott et al. 2012; Liu et al. 2014a,b; Hazelton et al. 2013; Morales et al. 2012; Vedantham et al. 2012; Datta et al. 2010). Foregrounds can extend beyond the horizon level as a result of two effects described in Pober et al. (2012): side lob power and the the extension of sources which are not entirely spectrally smooth beyond the physical horizon for a baseline.

A number of techniques are being investigated to remove foreground contamination from both point sources and diffuse structure in both image and Fourier domains (Beardsley et al. 2016; Pober et al. 2016; Trott et al. 2012) and Line et al. (in review). Aside from removing foreground contaminations, some analysis pipelines attempt to exploit the isolation of the foreground wedge and focus their analysis in the "EoR window" as a way to avoid foregrounds (Jacobs et al. 2015; Jacobs et al. 2016; Trott et al. 2016; Dillon et al. 2013, 2015; Trott 2014; Ali et al. 2015; Parsons et al. 2014).

Many so called "foreground avoidance" techniques rely on the use of an Optimal Quadratic Estimator of the 21cm power spectrum (Ali et al. 2015 hereafter A15, Liu & Tegmark 2011; Dillon et al. 2013, 2015; Liu et al. 2014a,b; Trott et al. 2012). These OQEs rely on empirical estimation of data's covariance and as a result may produce signal loss due to over fitting of noise (A15, Dillon et al. (2015); Switzer & Liu (2014)). Some techniques have been proposed to avoid information loss like the use of eigenmode projection (Dillon et al. 2015) or reducing temporal covariance through sub-optimal filter-

[1] School of Earth and Space Exploration, Arizona State U., Tempe AZ
[2] Department of Physics, Arizona State U., Tempe AZ
[3] Astronomy Dept., U. California, Berkeley CA
[4] Radio Astronomy Lab., U. California, Berkeley CA
[5] Dept. of Physics, Massachusetts Inst. of Tech., Cambridge MA
[6] Physics Dept. U. Washington, Seattle WA
[7] Berkeley Center for Cosmological Physics, Berkeley, CA
[8] Dept. of Physics and Astronomy, U. Penn., Philadelphia PA
[9] Dept. of Electrical and Computer Engineering, U. Virginia, Charlottesville VA
[10] National Radio Astronomy Obs., Charlottesville VA
[11] Dept. of Astronomy, U. Virginia, Charlottesville VA
[12] Square Kilometer Array, S. Africa, Cape Town South Africa
[13] Dept. of Physics and Electronics, Rhodes University
[14] Harvard-Smithsonian Cen. for Astrophysics, Cambridge MA
[15] National Radio Astronomy Obs., Socorro NM
[16] Cavendish Lab., Cambridge UK

[1] `eor.berkeley.edu`
[2] `www.lofar.org`
[3] `mwatelescope.org`
[4] `reionization.org`

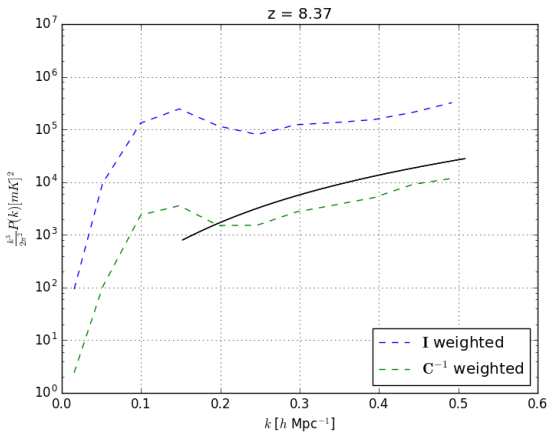

FIG. 1.— 2 sigma upper limits on the 21cm power spectrum using the A15 data. Blue line indicates the power spectrum estimate when no inverse covariance weighting is applied. Green indicates the power spectrum estimate with inverse covariance weighting applied. The expected 2-sigma noise level for each filter's total integration time is given in black. The presence of data below the theoretical lower noise bound may be indicative of information loss

ing in A15.

We find, however, the characterization and presence of information loss in these studies is insufficient and the effect is more prevalent that previously expected. Our investigation focuses on the technique and data used in A15. We offer a new method to characterize and account for these effects in power spectrum estimation using OQEs.

This paper is organized as follows: we provide an review of the mathematics of OQE's in section 2. Our investigation of information loss and new method to overcome this effect is described in section 3. We conclude with a discussion of these findings in section 4.

2. THE OPTIMAL QUADRATIC ESTIMATOR

The optimal quadratic estimator relies on the formalism of the work done by Tegmark (1997) and expanded to the 21cm signal in Liu & Tegmark (2011). The estimator is developed to help provide an improved estimate of the underlying 21cm power spectrum in the presence of foreground contaminations. Ideally the 21cm power spectrum (P_{21}) is defined:

$$\left\langle \tilde{T}_b(\mathbf{k})\tilde{T}_b^{\star}(\mathbf{k}')\right\rangle = (2\pi)^3\,\delta(\mathbf{k}-\mathbf{k}')P_{21}(\mathbf{k}) \tag{1}$$

where brackets denote ensemble averages, $\tilde{T}_b(\mathbf{k})$ is the Fourier transform of the 21cm brightness temperature and δ is the Dirac delta function. As in A15, the quadratic estimation of this power spectrum begins with the un-normalized bandpowers in the α-th k-bin:

$$\hat{q}_\alpha = \frac{1}{2}\mathbf{x}_1\mathbf{C}^{-1}\mathbf{Q}_\alpha\mathbf{C}^{-1}\mathbf{x}_2 \tag{2}$$

the subscripts 1 and 2 denote independent sky measurements, $\mathbf{C}$ is the empirical covariance matrix ($\mathbf{C} = \left\langle \mathbf{x}\mathbf{x}^\dagger\right\rangle$), and $\mathbf{Q}_\alpha$ is a matrix operator which performs the Fourier transform and bins the data into the α-th k-bin. $\mathbf{Q}_\alpha$ can formally be evaluated as $\mathbf{Q}_\alpha \equiv \frac{\delta\mathbf{C}}{\delta p_\alpha}$ or the derivative of the covariance with respect to the true bandpower in the k-bin.

The set of $\hat{q}_\alpha$ can be related to the properly normalized power spectrum estimates by a normalization matrix ($\mathbf{M}$)[5]:

$$\hat{\mathbf{p}} = \mathbf{M}\hat{\mathbf{q}} \tag{3}$$

The normalized bandpowers form our estimate of $P_{21}(\mathbf{k})$ and are related to the true power spectrum $\mathbf{p}$ via a windowing matrix ($\mathbf{W}$):

$$\hat{\mathbf{p}} = \mathbf{W}\mathbf{p} \tag{4}$$

In other words, the normalized power spectrum probes a linear combination of the true power spectrum estimates in each k-bin. To compute the normalization matrix, $\mathbf{M}$, Liu et al. (2014b) shows that:

$$\mathbf{W} = \mathbf{M}\mathbf{F} \tag{5}$$

where $\mathbf{F}$ is the Fischer information matrix and can be computed:

$$\mathbf{F}_{\alpha\beta} = \frac{1}{2}\left[\mathbf{C}^{-1}\mathbf{Q}_\alpha\mathbf{C}^{-1}\mathbf{Q}_\beta\right] \tag{6}$$

The normalization requires a choice of the windowing function, $\mathbf{W}$, to compute the desired power power spectrum. A15 chooses $\mathbf{W}$ such that every k-bin is a linear combination of k modes higher than itself. This restricts contaminating foregrounds to only the lowest k-bins without allowing them to contaminate high k-bins where the EoR signal may be stronger.

Estimating the uncertainties in $\hat{\mathbf{p}}$ is accomplished via the bandpower covariance matrix:

$$\mathbf{\Sigma} = \left\langle \hat{\mathbf{p}}\hat{\mathbf{p}}^\dagger\right\rangle - \left\langle\hat{\mathbf{p}}\right\rangle\left\langle\hat{\mathbf{p}}^\dagger\right\rangle = \mathbf{M}\mathbf{F}\mathbf{M}^\dagger \tag{7}$$

3. INFORMATION LOSS IN OQES

An OQE is traditionally regarded as a lossless power spectrum estimator (Tegmark 1997), but A15, Dillon et al. (2015) and Switzer & Liu (2014) describe how the use of this kind of estimator with empirically estimated covariance matrices can lead to information loss.

To summarize, lossless estimation requires *a priori* knowledge of the statistics of the covariance matrix $\mathbf{C}$ but estimating $\mathbf{C}$ empirically fundamentally violates this assumption. As a result, random fluctuations in the data from noise may cause statistically independent sky measurements to appear covariant and thus be down-weighted during power spectrum estimation.

Our investigation of potential signal loss in our power spectrum estimator begins by running data filtered by the FRF used in A15 and the optimal FRF through the power spectrum pipeline and comparing the relative amplitudes of the estimates with and without the inverse covariance weighting. The 2σ upper limits from both the inverse covariance weighted data (green) and non-weighted data (blue) from the A15 FRF is displayed in Figure 1. Each panel also contains the 2σ theoretical noise sensitivity for each FRF computed using the 21cmSense package (Pober et al. 2013).

Even with perfect foreground removal, a lossless power spectrum estimator would not be able to produce results

[5] In Liu & Tegmark (2011) and Liu et al. (2014b) Equation 2 is instead written as the normalized power ($\hat{p}_\alpha$) and relies on the estimation matrix $\mathbf{E}_\alpha$. We can relate the matrices $\mathbf{E}_\alpha = M_\alpha\mathbf{Q}_\alpha$ to compute the bandpowers directly.

below the thermal noise level. A possible solution is a miscalculation of the black curves in Figure 1, however we continue our investigation assuming the thermal noise calculations from Pober et al. (2013) are valid.

We then perform Monte-Carlo simulations of a Gaussian temperature field with a flat amplitude $P(k)$ over 60 sky realizations for all input signal levels in the range $[10^4, 10^{13}]$ $(mK)^2$. In A15, the covariance of the data + injected sky signal was only applied to the injected sky signal to determine what loss would occur to this signal if it existed in the data.

We apply the covariance matrix to the sum of data and injected signal to investigate the possibility of total power loss before and after the application of the OQE. For noise limited data being processed with a lossless power spectrum estimator, the total output power of data + injected sky signal should lay along the sum of the expected noise level (green dashed line) and the simulated sky signal in Figure 2.

The application of this method to the A15 data the the optimally FRF data both is is shown in Figure 2. This figure illustrate loss for high values of P_{inj} but at lower values of P_{in} it is difficult to discern whether it is indeed suffering from information loss. While the power spectrum estimates in this regime do reside below the non-covariance weighted (**I** weighted) power spectra for the data, they are still close to the theoretical noise level. As a result it is difficult to discern between this being a result of information loss or the proper down-weighting of foreground signals. To further address this potential issue, the simulation is repeated with simulated visibilities consisting of a white noise signal instead of the observed data. For white noise visibilities, **C** should converge to **I** multiplied by a scalar quantity (the input noise power). The expected result from applying the Monte-Carlo simulation described above would be identical input and output power for all injected sky signals. The disagreement between these quantities and the similar output shapes of Figure 2 and Figure 3 indicate the use of empirically estimated covariance matrices in this application of and OQE are causing signal loss.

To overcome the effects of signal loss in our power spectrum estimator, we interpret data points and their error bars for each k-bin and injected sky signal as probability distributions. We then compute the probability for each sky signal that the output power of the Monte-Carlo simulations would be observed above the 1σ upper bounds of the data's power spectrum in each k-bin. This is interpreted as the probability that an underlying 21cm background signal of level P_{inj} would have been detected if present in the data. An example of these probability curves for the A15 noise simulation is displayed in Figure 4.

A hyperbolic tangent is fit to the probability of detection as a function of P_{inj}, from this an arbitrary level of confidence that an underlying signal would be detected can be assigned. We interpret this confidence level as our signal loss corrected upper limits.

4. DISCUSSION

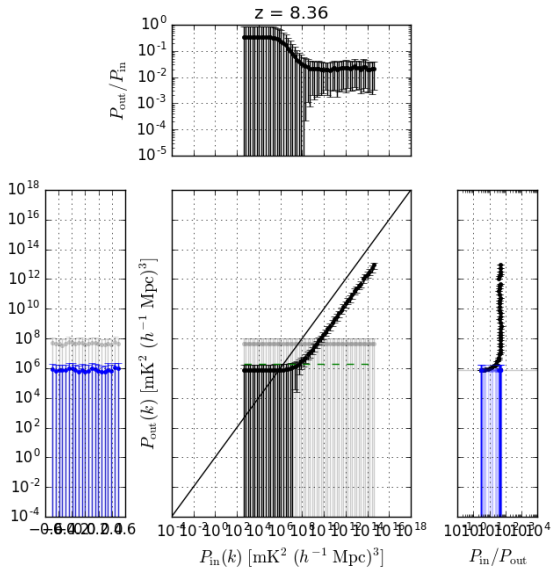

FIG. 2.— Left: A15 Loss curve. **these figures are awful but are mostly filling space currently until we make 'em pretty**

REFERENCES

Ali, Z. S., et al. 2015, Astrophys. J., 809, 61
Beardsley, A. P., Hazelton, B. J., Sullivan, I. S., et al. 2016, ArXiv e-prints, arXiv:1608.06281
Bowman, J. D., Morales, M. F., & Hewitt, J. N. 2009, ApJ, 695, 183
Bowman, J. D., Cairns, I., Kaplan, D. L., et al. 2013, PASA, 30, 31
Datta, A., Bowman, J. D., & Carilli, C. L. 2010, ApJ, 724, 526
Dillon, J. S., Liu, A., & Tegmark, M. 2013, Phys. Rev. D, 87, 043005
Dillon, J. S., Neben, A. R., Hewitt, J. N., et al. 2015, Phys. Rev. D, 91, 123011
Furlanetto, S. R., Oh, S. P., & Briggs, F. H. 2006, Phys. Rep., 433, 181
Hazelton, B. J., Morales, M. F., & Sullivan, I. S. 2013, ApJ, 770, 156
Jacobs, D. C., et al. 2015, Astrophys. J., 801, 51
Jacobs, D. C., Hazelton, B. J., Trott, C. M., et al. 2016, ApJ, 825, 114
Liu, A., Parsons, A. R., & Trott, C. M. 2014a, Phys. Rev. D, 90, 023018
—. 2014b, Phys. Rev. D, 90, 023019
Liu, A., & Tegmark, M. 2011, Phys. Rev. D, 83, 103006
Morales, M. F., Bowman, J. D., & Hewitt, J. N. 2006, ApJ, 648, 767
Morales, M. F., Hazelton, B., Sullivan, I., & Beardsley, A. 2012, ApJ, 752, 137
Morales, M. F., & Wyithe, J. S. B. 2010, ARA&A, 48, 127
Paciga, G., Albert, J. G., Bandura, K., et al. 2013, MNRAS, 433, 639
Parsons, A. R., Pober, J. C., Aguirre, J. E., et al. 2012, ApJ, 756, 165
Parsons, A. R., Backer, D. C., Foster, G. S., et al. 2010, AJ, 139, 1468
Parsons, A. R., Liu, A., Aguirre, J. E., et al. 2014, ApJ, 788, 106
Pober, J. C., Parsons, A. R., Jacobs, D. C., et al. 2012, AJ, 143, 53
Pober, J. C., Parsons, A. R., Aguirre, J. E., et al. 2013, ApJ, 768, L36
Pober, J. C., Liu, A., Dillon, J. S., et al. 2014, ApJ, 782, 66

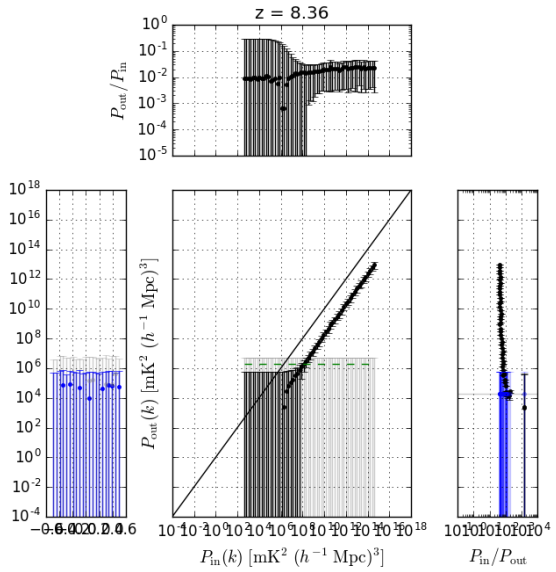

FIG. 3.— Left: A15 Loss curve noise. **these figures are awful but are mostly filling space currently until we make 'em pretty**

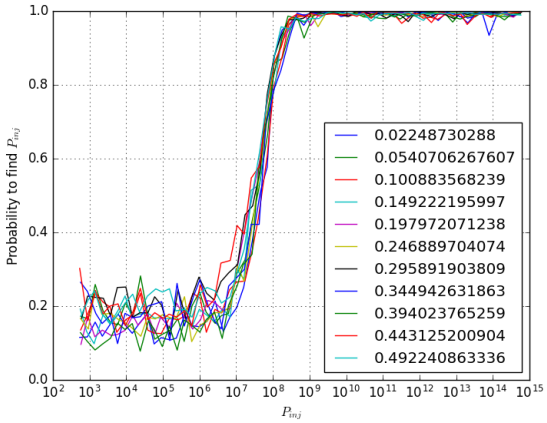

FIG. 4.— A15 probability curve **these figures are awful but are mostly filling space currently until we make 'em pretty**

Pober, J. C., Hazelton, B. J., Beardsley, A. P., et al. 2016, ApJ, 819, 8
Pritchard, J. R., & Loeb, A. 2012, Reports on Progress in Physics, 75, 086901
Switzer, E. R., & Liu, A. 2014, The Astrophysical Journal, 793, 102
Tegmark, M. 1997, Phys. Rev. D, 55, 5895
Thyagarajan, N., Udaya Shankar, N., Subrahmanyan, R., et al. 2013, ApJ, 776, 6
Thyagarajan, N., Jacobs, D. C., Bowman, J. D., et al. 2015a, ApJ, 807, L28
—. 2015b, ApJ, 804, 14
Trott, C. M. 2014, PASA, 31, e026
Trott, C. M., Wayth, R. B., & Tingay, S. J. 2012, ApJ, 757, 101
Trott, C. M., Pindor, B., Procopio, P., et al. 2016, The Astrophysical Journal, 818, 139
Vedantham, H., Udaya Shankar, N., & Subrahmanyan, R. 2012, ApJ, 745, 176
Yatawatta, S., de Bruyn, A. G., Brentjens, M. A., et al. 2013, A&A, 550, A136