EPOCH OF REIONIZATION POWER SPECTRUM RESULTS FROM PAPER-128

Carina Cheng[1], et al.

[1]Astronomy Dept., U. California, Berkeley, CA

ABSTRACT

1. INTRODUCTION

[Need to cite papers...]

The Epoch of Reionization (EoR) represents an uncharted phase in cosmic history when the first stars and galaxies ionized the neutral hydrogen present in the early Universe. During this time, it is believed that baryonic matter fell into dark matter potential wells, gravitationally collapsing into the first luminous structures. These sources, possibly quasars or young dwarf galaxies, emit radiation that ionized the abundant hydrogen in the intergalactic medium (IGM). The subsequent formation of a web of galaxies across an ionized IGM is thought to represent the large scale structure that we observe today.

The EoR has not yet been directly detected. However, constraints have been made on its timing and duration through different methods. As reionization progressed, the number of free electrons susceptible to Thompson scattering increased, resulting in scattered cosmic microwave background (CMB) photons whose imprint has been left on CMB observations. In particular, the Thompson scattering optical depth parameter τ is a useful metric for characterizing the duration of reionization. Recent *Planck* measurements of τ suggest a late reionization history, occurring between redshifts of 7.7 and 8.7.

Galaxy observations have also begun to place constraints on the EoR. With their high luminosities, quasars make for powerful probes of the state of the IGM through absorption line studies. As photons from a quasar are redshifted, they may be absorbed by neutral hydrogen at the Lyman-α line, producing Gunn-Peterson trough features in quasar spectra. Observations of these spectra in high redshift galaxies have signalled the end of reionization by $z \sim 6$, and experiments are continuously pushing galaxy observations to higher redshifts.

A third powerful probe of the EoR makes use of the 21 cm wavelength emission produced by hydrogen due to its two quantum states. Because hydrogen is so pervasive in the Universe and a transition from its spin up state to spin down state gives off a well-defined energy difference in the form of a 21 cm wavelength photon, observations targeting this signal have the potential to trace the EoR from start to finish. However, 21 cm experiments face major challenges, including those from bright foregrounds and instrumental systematics. Nevertheless, the observations and analyses of data from current radio telescope arrays have already begun to constrain the 21 cm signal and make statements about the nature of the IGM during the EoR.

Major existing 21 cm experiments include both those targeting to measure the sky-averaged 'global' 21 cm signal and those employing interferometric elements and the use of cross-correlations and statistical power-spectral measurements. Examples of the former are EDGES (Bowman & Rogers 2010), the LWA (Ellingson et al. 2009), LEDA (Greenhill & Bernardi 2012), DARE (Burns et al. 2012), SciHi (Voytek et al. 2014), BIGHORNS (Sokolowski et al. 2015), and SARAS (Patra et al. 2015). Major interferometers include the GMRT, LOFAR, the MWA, the 21CMA, and PAPER.

In this paper we present results from the 128 antenna configuration of the Donald C. Baker Precision Array for Probing the Epoch of Reionization (PAPER). PAPER is a highly redundant array consisting of dual-polarization dipole antennas. Its configuration is optimal for boosting power spectrum sensitivity and the analysis used relies on a foreground avoidance approach where we measure power spectrum values in a region of delay space that is uncontaminated by foregrounds. Previous upper limits on the EoR power spectrum include those made by PAPER-32, which used its redundant configuration to its advantage to obtain an upper limit of $\Delta^2(k) \leq (41\mathrm{mK})^2$ at $k = 0.28 h\mathrm{Mpc}^{-2}$. Doubling in array size and employing an updated analysis pipeline that included fringe-rate filtering and an optimal quadratic estimator framework for the power spectrum analysis achieved an upper limit on $\Delta^2(k)$ of $(22.4\mathrm{mK}))^2$ in the range $0.15 < k < 0.5 h\mathrm{Mpc}^{-1}$ at $z = 8.4$. This result from PAPER-64 was the best published upper limit thus far, with PAPER-128 achieving a factor of ?? in improvement.

PAPER-128 has been able to use much of the same technical analysis framework from previous PAPER analyses, while improving on the limits set by PAPER-64 with a longer observing time, double the number of antennas, and the development of new data quality tests, signal loss simulations, and inverse covariance weighting techniques.

The organization of the paper is as follows. ...
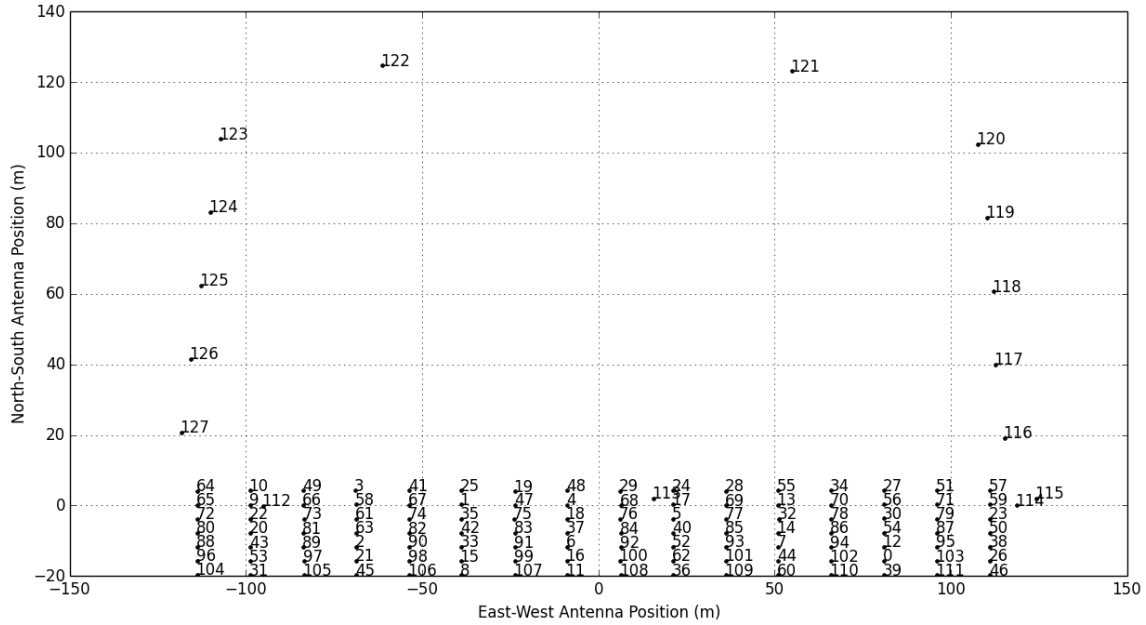
2. OBSERVATIONS

Figure 1. The antenna layout of PAPER-128. Core grid antennas are labeled 1 through 111 and outrigger antennas are labelled 112 through 127.

PAPER is located in the Karoo desert in South Africa at a latitude of $-30:43:18$ and longitude of $21:25:42$. The PAPER-128 antenna configuration is shown in 1. It consists of a 7x16 core grid of antennas and 16 outrigger antennas, some of which are located inside the core grid and some placed farther away to establish longer baselines. The analysis carried out in this paper makes use of only the core grid of antennas, with the East-West direction shown in 1 as to the right and left. Adjacent East-West baselines have a length of $15m$ and adjacent North-South baselines have a length of $4m$. The baselines used in our analysis are...

The PAPER signal chain consists of drift-scan antenna elements... **[mention Casper and instrumentation things]** The PAPER beam is...

The PAPER-128 observing season began in November, 2013, with nightly data taken for approximately 4 months. The second observing season started up again in June, 2014 and lasted until February, 2015. Both these seasons encompass part or all of the months corresponding to our optimal observing season for which the Galaxy is below the horizon. Collectively, PAPER-128 logged 233 nights of raw data. In this analysis, we use ?? nights of data from season ??. We exclude nights of known hardware malfunctioning (power failures, overheating of equipment, etc.), as well as outlier nights whose data did not pass our quality checks (see: quality paper). For our analysis, we use an LST range of ?? which corresponds to observations of a 'cold patch' of the night sky. All polarization products are saved, but we only make use of the XX and YY polarizations in this paper.

3. DATA REDUCTION

Raw PAPER data consists of 1024 frequency channels sampled every 97.6kHz. Every 10 minute file then goes through our compression pipeline, which consists of a two-step radio frequency interference (RFI) flagger and averaging along both the frequency (delay filtering) and time (fringe rate filtering) axes. Final compressed data consists of 203 frequency channels covering a bandwidth of $100 - 200$MHz. Each 10 minute file contains data sampled at integrations of 31.65s. The amplitudes of one night of raw and compressed data is shown in 2 as waterfall plots of frequency versus time.

3.1. *Redundant Calibration*

The sought-after 21 cm signal lies beneath bright foregrounds and various instrumental systematics that contaminate our data. Because foregrounds are predicted to be $4-5$ orders of magnitude brighter than the cosmological signal, 21 cm experiments require high dynamic

Figure 2. Waterfall plots (in log Jy) of one night of raw PAPER-128 data (left) and the same data after compression (right). Compression downsizes files by averaging over time and frequency and flags radio frequency interference.

ranges and considerable sensitivities. PAPER is designed to boost its sensitivity by containing many short, repeated baselines of the same length and orientation in its array. Thus, PAPER is able to measure the same Fourier-mode on the sky multiple times while beating down the amount of noise with each successive measurement.

Calibration of antennas in a redundant array is therefore crucial in order to maximize sensitivity. In our raw data, visibility measurements from two identical baseline types are not necessarily equal. This is due to changing system temperatures and different signal chains as data flows from each antenna into the correlator. However, PAPER's redundant layout can be used to solve for complex antenna gains, which when applied to the data, bring all visibility measurements for a specific baseline type into alignment.

Redundant calibration has proven to be a powerful tool. PAPER-64 relied on redundant calibration to achieve its power spectrum limits, and here we update our method and tools and apply it to PAPER-128. As a review, let's start with the visibility measurement equation:

$$V'_{ij} = g_i^* g_j V_{ij} + n_{ij} \qquad (1)$$
$$g_j = \eta_j + i\phi_j \qquad (2)$$

Here V'_{ij} represents a measured visibility for a baseline with antennas i and j, V_{ij} represents the 'true' sky visibility for that same baseline, g_i and g_j are the complex gain factors for each antenna, and n_{ij} is noise. Each gain number consists of an amplitude η and phase ϕ. The visibility equation can be written down for all baselines and - without using array redundancy - would not be least squares solvable because the number of unknowns $(2N + N(N-1)/2)$ is greater than the number of measurements $(N(N-1)/2)$, where N is the number of antennas . However, if V_{ij}, the 'true', or 'model' visibility, is constrained to be the same for all identical baselines, the system of equations becomes solvable.

For our analysis, we use the package OMNICAL for redundant calibration. Similarly to the analysis of PAPER-64, OMNICAL comprises of three steps. The first is `FirstCal`, which uses all baseline redundancies to generate a static gain solution for each antenna that will unwrap any phase wrapping between two identical baselines. We perform `FirstCal` because the next stage of OMNICAL cannot tell the difference between a phase of 0 and 2π, for example. The second step is `LogCal`, which takes the log of all the visibility equations and separates the real and imaginary components into two matrices. Coarse solutions are determined for both the antenna gains and 'model' visibilities (one for each baseline type) simultaneously. The final step of OMNICAL is `LinCal`, which applies small perturbations to the `LogCal`

solutions in an iterative fashion, honing in on the optimal solutions.

We perform redundant calibration for each epoch of data used in our analysis, as each epoch has to be separately calibrated since they are on different overall gain scales. We divide out the antenna gains for each of our measured visibilities, bringing them into alignment with each other (for identical baseline types) and the 'model' visibility for that baseline type. We also include the subtraction of crosstalk in our dataset. Crosstalk shows up as a constant signal in time in our visibilities and is attributed to signals from one antenna being picked up by another. OMNICAL handles crosstalk by subtracting off an averaged residual (differences between measured visibilities and the 'model' visibility, averaged over baseline type) for each visibility measurement.

For PAPER-128, we update the OMNICAL software package, originally written in C, with a user-friendly interface written in Python. Additionally, we employ a chi-square metric on a per antenna basis in order to exclude dead or faulty antennas and improve our calibration solutions. **[Add plots to show stability of OMNICAL solutions and chi-square examples]**

It is important to note that while OMNICAL is powerful for ensuring array redundancy, it is not able to solve for 4 calibration parameters - namely, to correct for the overall gain, phase, and tip/tilt of the array. In order to do this, we turn to...

3.2. *Bandpass Calibration*
3.3. *Delay-Domain Foreground Removal*

Foregrounds pose a significant challenge to EoR experiments. Some experiments, like LOFAR and the MWA, rely on modelling and foreground subtraction to create a clean EoR 'window' in which to make their power spectrum measurements. Previous PAPER results, including this one, instead focus on a foreground avoidance approach.

In order to filter foregrounds, we exploit their spectrally-smooth nature. By Fourier-transforming our visibilities along the frequency axis, the delay domain provides a natural space in which to isolate foregrounds. Physically, a time delay can be thought of as the extra time light from a source takes to reach one antenna compared to another one (due to slightly different travelling distances). The maximum delay a source on the sky can have is if it is located on the horizon. In this case, the delay is determined by the length of the baseline. Consequently, all spectrally-smooth foreground signals show up within a maximum delay (or 'horizon limit') given a baseline length. The EoR signal, however, is believed to mimic noise which has a random frequency structure and is not confined to be within the horizon limit. A simple way to avoid foregrounds without removing the cosmological signal, therefore, is to make our power spectrum

measurements outside of the horizon limit that contains the bright foregrounds.

In practice, we first flag specific frequency channels that contain known RFI. If this is not done prior to foreground filtering, the EoR window may become contaminated by leakage from the Fourier transform of sharp RFI features. In addition to flagging entire channels with obvious RFI (such as channels occupied by ORBCOMM and the International Space Station) we also flag some with lower levels of RFI caused by radio and other Earth-origin activity.

Next, we filter our calibrated visibilities by deconvolving out our sampling function (which contains flags due to RFI) from our delay-domain visibilities using a CLEAN-like algorithm and restricting our clean components to inside the horizon limit (set by the baseline length) plus a 15ns buffer. The Fourier-transformed clean components are then subtracted from our visibilities. This filtering process is performed on a per-baseline, per-integration basis, and we achieve a brightness suppression of 2 orders of magnitude in our visibilities.

3.4. *RFI Removal, LST-Binning, and Absolute Calibration*

We perform a final round of RFI-removal by flagging visibilities that lie more than 4σ above the mean on a time, frequency, and baseline basis. Finally, we stack our data in LST into two datasets, alternating between even and odd Julian Dates to create an 'even' and 'odd' LST-binned dataset.

Recalling that redundant calibration through OMNICAL does not solve for the overall calibration parameters of our array, we calibrate these final unknowns by matching our LST-binned datasets with that of PAPER-64. For the PAPER-64 analysis, absolute calibration was carried out by standard radio astronomy self-calibration. Overall phase terms were fitted using Pictor A, Fornax A, and the Crab Nebula, and the overall flux scale was determined using Pictor A as a calibrator source. PAPER-64's analysis produced even and odd LST-binned data covering an LST range of $0 - 8$, fully encompassing the LST ranges we use in our analysis.

We match the overall gain scale of our data to that of PAPER-64 by fitting an 8th order polynomial to the time-averaged ratio of the two LST-aligned foreground-containing datasets and then applying that polynomial to our foreground-filtered LST-binned datasets. Similarly, we match the overall phase by fitting a linear line to the phase difference between PAPER-128 and PAPER-64.

3.5. *Fringe-Rate Filtering*

In order to further suppress foregrounds and instrumental noise, we weight our visibilities on a temporal basis in a technique known as fringe-rate filtering. We take the Fourier transform along time to enter fringe-rate space, in which our data spans a range of fringe-rates determined by our integration time. In the fringe-rate domain, it is a natural space in which to filter out noise that is not associated with the rotating sky. This is because signals attached to the sky have a maximum fringe-rate set by the rotation velocity of the Earth (i.e. sources at the equator appear the move the fastest), and therefore any signals that exceed this fringe-rate cannot be the cosmological signal and can safely be filtered without any loss of information.

It was shown in Parsons et al. (2016) that filtering in fringe-rate space is equivalent to weighting data by the primary beam. Because of this, we determine an optimal fringe-rate filter (Figure 3) that up-weights parts of the sky that our beam is most sensitive to, while down-weighting the opposite. We also filter out the zero fringe-rate bin, which contains signals that are constant in time, such as crosstalk. We then apply the filter on a per-baseline basis as a convolution in the time-domain, which has the affect of averaging our data in time.

Using an optimal fringe-rate filter comes at the cost of signal loss due to a reduction of the number of independent time samples in our dataset. This is described in more detail in section 4.2.

4. POWER SPECTRUM ANALYSIS

With our final visibilities in hand, we are ready to estimate the 21 cm power spectrum, $\hat{\mathbf{p}}(\mathbf{k})$. The power spectrum measurement is a function of $\mathbf{k} = (k_x, k_y, k_z)$, where each component of k represents the Fourier-dual to spatial coordinates (x, y, z). Our visibilities have already been Fourier-transformed along two of these axes, namely, along $k_\perp = \sqrt{k_x^2 + k_y^2}$. This perpendicular component of k is analogous to the spatial Fourier-modes in the plane of the sky, while the parallel component of k, $k_\parallel$, is the Fourier-transform of our visibilities along the frequency axis as sampled by our interferometer. It is important to distinguish $k_\parallel$, which represents delay-modes, from η, which is the perfect Fourier transform along frequency. Because of the chromaticity of our instrument, taking the Fourier transform of our data along frequency produces $k_\parallel$ modes, and there exists a wedge in k-space which is contaminated by foregrounds. In other words, $k_{perp} \propto \mathbf{b}/\lambda$, and as our baseline length increases, the horizon limit in delay space as described in section 3.3 also increases, creating a foreground dominated wedge.

Our visibilities measure the sky temperature in units of Jys. Recalling that we are interested in the power of the cosmological signal on different k-scales, we can map our data to a power spectral measurement using the standard definition

$$\langle \tilde{T}_b(\mathbf{k})^* \tilde{T}_b(\mathbf{k'}) \rangle = (2\pi)^3 \delta^D(\mathbf{k} - \mathbf{k'})\hat{\mathbf{p}}(\mathbf{k}) \qquad (3)$$

where $\tilde{T}_b(\mathbf{k})$ is the spatial Fourier-transform of the brightness temperature of the sky, δ^D is the Dirac delta function, and $\langle \rangle$ is an ensemble average. As a simplistic explanation, we need to take our visibility temper-

Figure 3. Top: The optimal fringe-rate profile (weighting in fringe-rate space) for our I polarization beam for an E/W baseline separation of $30m$. Bottom: The time-domain convolution kernel corresponding to the curve in the top panel.

ature measurements, Fourier-transform them along the frequency axis and convert from units of Jys to Kelvin, square, and bin in constant contours of k. This can be represented mathematically by

$$\hat{\mathbf{p}} = \left(\frac{\lambda^2}{2k_B}\right)\frac{X^2Y}{\Omega^{pp}B}|\tilde{V}(\mathbf{u}, \eta)|^2 \qquad (4)$$

where k_B is the Boltzmann constant, X and Y are factors that relate the cosmological Fourier coordinates to interferometric coordinates by $(Xk_x, Xk_y, Yk_z) \equiv 2\pi(u, v, \eta)$, Ω_{pp} is the integral of the primary beam squared over frequency and angle, and $\tilde{V}$ is the Fourier transformed visibilities.

In practice, however, we use linear algebra techniques as explained in the following section in order to obtain both $\hat{\mathbf{p}}(\mathbf{k})$ and its associated error bars.

4.1. *Optimal Quadratic Estimators*

We follow the optimal quadratic estimator (OQE) formalism used by Ali et al. (2015) to form estimates of our power spectrum $\hat{\mathbf{p}}(\mathbf{k})$. As a brief review, we recap the method here as applied to our data.

We begin with the data vector $\mathbf{x}$, which consists of our visibility measurements as a function of time and frequency for each baseline. We divide our data into separate datasets by baseline type, and apply the OQE formalism to each separately, forming multiple power spectra. In our analysis, we use [??] types of baselines. After forming the power spectra, we fold them into our final power spectrum by [??].

Recalling that our goal is to measure power spectrum values as a function of $\mathbf{k}$, the Fourier dual to the spatial directions, the next step is to form the set $\hat{q}_\alpha$. This quantity, defined as

$$\hat{q}_\alpha = \frac{1}{2}\mathbf{x}^\dagger\mathbf{C}^{-1}\mathbf{Q}_\alpha\mathbf{C}^{-1}\mathbf{x} - b_\alpha \qquad (5)$$

represents power spectrum quantities (un-normalized) for each k bin indexed by α. Here we see two copies of our data vector $\mathbf{x}$ and two copies of $\mathbf{C}^{-1}$ where

$$\mathbf{C} \equiv \langle\mathbf{x}\mathbf{x}^\dagger\rangle, \qquad (6)$$

the true covariance matrix of $\mathbf{x}$. The operator $\mathbf{Q}_\alpha$ transforms our visibilities into power spectrum space by performing a Fourier transform along the frequency axis, and b_α is a noise bias that is subtracted off.

Our data $\mathbf{x}$ has dimensions of time versus frequency and exists for every baseline (in a given redundant group), and group (such as an 'even' or 'odd' LST-binned set). Notice that there is a copy of $\mathbf{C}^{-1}\mathbf{x}$ on each side of the expression for $\hat{q}_\alpha$. When performing this computation, we do not allow the cross-multiplication to occur if both quantities are derived for the same baseline, or the same group. For example, we do not allow $(\mathbf{C}^{-1}\mathbf{x})_{even,A}$ to be multiplied with $(\mathbf{C}^{-1}\mathbf{x})_{even,B}$

since both quantities are formed from the 'even' dataset, nor do we allow $(\mathbf{C}^{-1}\mathbf{x})_{even,A}$ to be multiplied with $(\mathbf{C}^{-1}\mathbf{x})_{odd,A}$, since both quantities are formed for baseline A. We do not include these cases in order to eliminate any biases from allowing auto power from identical baselines or groups to remain.

For faster computation (due to less cross-multiplications), we divide our baselines into 5 subgroups and sum up the quantities $\mathbf{C}^{-1}\mathbf{x}$ for each. We still do not allow $\hat{q}_\alpha$ to be computed if the same baseline appears in both subgroups that are being cross-multiplied.

Once we have formed our $\hat{q}_\alpha$s, we can normalize them using the matrix $\mathbf{M}$. Namely,

$$\hat{\mathbf{p}} = \mathbf{M}\hat{\mathbf{q}} \qquad (7)$$

where $\hat{\mathbf{p}}$ is the normalized estimate of the true power spectrum $\mathbf{p}$. Before we compute $\mathbf{M}$, we define the Fisher matrix $\mathbf{F}$ to be

$$\mathbf{F}_{\alpha\beta} = \frac{1}{2}tr(\mathbf{C}^{-1}\mathbf{Q}_\alpha\mathbf{C}^{-1}\mathbf{Q}_\beta) \qquad (8)$$

where α and β represent different bandpowers or k bins. The Fisher matrix encodes information about [??]. It is related to the normalization matrix $\mathbf{M}$ by the window function matrix $\mathbf{W}$, which relates the true power spectrum $\mathbf{p}$ with our estimates $\hat{\mathbf{p}}$:

$$\mathbf{W} = \mathbf{M}\mathbf{F} \qquad (9)$$
$$\hat{\mathbf{p}} = \mathbf{W}\mathbf{p} \qquad (10)$$

The window function is constructed so that each row sums to unity. This ensures proper normalization of the power spectrum estimates. In practice, we have a choice of $\mathbf{M}$ as described in Liu et al. (2014) and in our analysis choose to use a method where it is determined from the Cholesky decomposition of $\mathbf{F}$. Namely, we decompose the Fisher matrix such that $\mathbf{F} = \mathbf{L}\mathbf{L}^\dagger$, where $\mathbf{L}$ is a lower triangular matrix. Next, we construct $\mathbf{M} = \mathbf{D}\mathbf{L}^{-1}$, where $\mathbf{D}$ is a diagonal matrix. Hence, $\mathbf{W} = \mathbf{D}\mathbf{L}^\dagger$ and is an upper triangular matrix. This window function was constructed in a way to prevent the leakage of foreground power from low k to high k modes. Specifically, we order the elements in $\mathbf{F}$ in such a way so that power can leak from high k modes to low k modes, but not vice versa. Since most foreground power shows up at low k's, this method ensures a window function that retains clean, noise-dominated measurements.

4.2. *Signal Loss*

Both the fringe-rate filtering and inverse covariance weighting of our visibilities produces data that is susceptible to signal loss. This is due to the reduction of independent time samples from fringe-rate filtering which

leads to a more noisy dataset. Consequently, when computing $\mathbf{C}^{-1}$ from the data, the result from the inverse covariance weighting of that same data is to over-fit the noise. This can produce power spectra values below the thermal noise limit that is predicted based on observation parameters.

We recognize that the permutation of fringe-rate filtering and then applying the OQE framework, as well as our choice for $\mathbf{C}^{-1}$, play a large role in the amount of signal loss that results. In order to quantify signal loss, we run simulations where various levels of an EoR-like signal are added on top of our data. This injected signal, $\mathbf{e}$, resembles the predicted EoR signal by mimicking random noise and by being common to all baselines. It is fringe-rate filtered in the same way as our data, representing what could be the cosmological signal contained within our data. We then compare the input and output power spectrum of $\mathbf{e}$ as our full OQE framework is applied.

Mathematically, suppose the addition of the injected signal onto our data is

$$\mathbf{x} + \mathbf{e} = \mathbf{r}. \tag{11}$$

The input power spectrum of $\mathbf{e}$ is computed as the power spectrum that results when an identity matrix $\mathbf{I}$ is used to replace any covariance matrices used in our pipeline. Ignoring normalization factors, this is written as

$$P_{in} \propto \mathbf{e}^{\dagger}\mathbf{I}^{-1}\mathbf{Q}\mathbf{I}^{-1}\mathbf{e} \tag{12}$$

and there is no signal loss associated with this computation. The expression for the output power spectrum, P_{out}, is comprised of two terms: the covariance treated power spectrum associated with $\mathbf{r}$, and that of $\mathbf{x}$ alone.

$$\begin{aligned} \mathbf{p}_x + \mathbf{p}_e &= \mathbf{p}_r \\ P_{in} \equiv \mathbf{p}_e &= \mathbf{p}_r - \mathbf{p}_x \\ &\propto \mathbf{r}^{\dagger}\mathbf{C}_r^{-1}\mathbf{Q}\mathbf{C}_r^{-1}\mathbf{r} - \mathbf{x}^{\dagger}\mathbf{C}_x^{-1}\mathbf{Q}\mathbf{C}_x^{-1}\mathbf{x} \end{aligned} \tag{13}$$

The explicit signal loss calculation described here is

new to the PAPER-128 analysis. We test the expression in (13) by running several simulations without data (i.e. pure noise) in order to verify its output. [**reference signal loss paper?**]

The ratio of P_{in} to P_{out} changes as the level of the injected signal $\mathbf{e}$ is increased. As shown in [**reference figure here**], when the injected signal level is low compared to the data $\mathbf{x}$, the recovered output power spectrum is the same as the input. However, when $\mathbf{e}$ begins to dominate over $\mathbf{x}$, we move to a regime where the input power spectrum is greater than the output, implying signal loss due to the over-fitting of noise during inverse covariance weighting.

Our power spectrum for the data alone is equivalent to the second term in (13), barring normalization factors. We choose to apply a conservative signal loss factor - that is, the factor that we multiply our power spectrum points by is the ratio P_{in}/P_{out} that corresponds to a P_{out} value equal to our highest measured power spectrum value (the point that produces the most loss). We apply this correction factor to our entire power spectrum.

4.3. *Jack-Knifes*

5. RESULTS

Our final 21 cm power spectrum measurements from PAPER-128 are for a redshift of $z = $ [**??**]. The data used for this result incorporates [**??**] days and [**??**] channels, centered around [**??**]. We present both the power spectrum points from $\hat{\mathbf{p}}(\mathbf{k})$ in units of $mk^2(h^{-1}Mpc)^3$ and the corresponding dimensionless power spectrum points, $\Delta^2(\mathbf{k})$, which is related to $\hat{\mathbf{p}}(\mathbf{k})$ by

$$\Delta^2(\mathbf{k}) = \frac{k^3}{2\pi^2}\hat{\mathbf{p}}(\mathbf{k}). \tag{14}$$

6. CONCLUSIONS

7. ACKNOWLEDGEMENTS

REFERENCES

Ali, Z. S., et al. 2015, ApJ, 809, 61
Bowman, J. D., & Rogers, A. E. E. 2010, Nature, 468, 796
Burns, J. O., et al. 2012, Advances in Space Research, 49, 433
Ellingson, S. W., Clarke, T. E., Cohen, A., Craig, J., Kassim, N. E., Pihlstrom, Y., Rickard, L. J., & Taylor, G. B. 2009, IEEE Proceedings, 97, 1421
Greenhill, L. J., & Bernardi, G. 2012, ArXiv e-prints
Liu, A., Parsons, A. R., & Trott, C. M. 2014, PhRvD, 90, 023019

Parsons, A. R., Liu, A., Ali, Z. S., & Cheng, C. 2016, ApJ, 820, 51
Patra, N., Subrahmanyan, R., Sethi, S., Udaya Shankar, N., & Raghunathan, A. 2015, ApJ, 801, 138
Sokolowski, M., et al. 2015, PASA, 32, e004
Voytek, T. C., Natarajan, A., Jáuregui García, J. M., Peterson, J. B., & López-Cruz, O. 2014, ApJL, 782, L9