UNLOCKING SENSITIVITY WITH A TWO-BASELINE POWER SPECTRUM ESTIMATOR FOR 21 CM INTEFEROMETRIC PROBES OF REIONIZATION

Yunfan Gerry Zhang [1], Adrian Liu [1, 2], Aaron R. Parsons[1, 2]

Draft version July 12, 2017

ABSTRACT

Measuring the cosmological 21 cm power spectrum with radio interferometers requires high sensitivity. Many modern low-frequency interferometers are thus designed with antennas placed on regular grid to maximize redundancy. Current visibility-based power spectrum pipelines, though shown to ease control of systematics, lack the ability to include partial redundancy. We introduce a method to include partial redundancy in such power spectrum pipelines of drift-scan arrays. Our method pairs baselines to cross-multiply at a time lag, and quantifies the sensitivity contributions of each pair of baselines. Using the 128-element configurations and beams of the Donald C. Backer Precision Array for Probing the Epoch of Reionization (PAPER-128), and 4 planned versions of the Hydrogen Epoch of Reionization Array (HERA), we illustrate how our method applies to different arrays and predict the sensitivity improvements of including each baseline pairs. We show that inclusion of partially redundant baselines would account for 20% to 60% of the sensitivity of PAPER-128 and different configurations of HERA.

1. INTRODUCTION

The epoch of reionization represents the last key stage of our Universe's early evolution. Study of this event stands at the intersection of cosmology and astrophysics. Understanding this event not only serves as a scientific goal of its own, but also as a gateway to crucial information regarding fundamental physics of inflation, neutrino mass and phenomenology of the first stars and galaxies (e.g. Liu et al. 2016; Liu & Parsons 2016; Mao et al. 2008; Chen 2015; Bull et al. 2015; Oyama et al. 2013).

Observational studies of reionization, including Gunn-Peterson measurements of quasi-stellar objects (Fan et al. 2006) and cosmic microwave background temperature and anisotropy measurements (CMB;Planck Collaboration et al. 2016), the Kinetic Sunyaev-Zeldo'vich effect (Smith & Ferraro 2016) and Lyman alpha emitter clustering (McQuinn et al. 2007) have given us indications of the rough time-frame of reionization, but only limited constraints on the finer spatial and temporal structures. A surge of recent radio-astronomical experiments of reionization focus on measuring the "spin-flip" transition of neutral hydrogen of characteristic wavelength 21 cm (Furlanetto et al. 2006; Pritchard & Loeb 2012). The 21 cm brightness temperature is a direct tracer of neutral hydrogen through the epoch of reionization, and thus tomography of the 21 cm line is a direct measurement of the full temporal and spatial variations of this event. Before realizing of full-scale 21 cm tomography, many current radio interferometric efforts aim to measure the spatial power spectrum of 21 cm brightness temperature fluctuations. Current-generation instruments include the Donald C. Backer Precision Array for Probing the Epoch of Reionization (PAPER; Ali et al. 2015; Parsons et al. 2014), Murchison Widefield Array (MWA; Bowman et al. 2013; Tingay et al. 2013), Low Frequency Array (LOFAR; van Haarlem, M. P. et al. 2013). Next-generation instruments include the Hydrogen Epoch of

Reionization Array (HERA; e.g. DeBoer et al. 2016; Dillon & Parsons 2016a; Neben et al. 2016; Ewall-Wice et al. 2016) under construction, and the Square Kilometer Array Low (SKA-low; e.g. Mesinger et al. 2015) currently in planning stages.

The highly redshifted 21 cm signal is faint and diffuse, in contrast to the localized bright sources targeted by many traditional radio telescopes. Current 21 cm experiments are sensitivity-starved, with the sensitivity requirement further increased considering the foreground contaminations 5 orders of magnitude brighter than the cosmological signal of interest. Low-frequency radio interferometers aiming to measure the 21 cm signal are thus designed differently from traditional instruments. Specially designed arrays such as PAPER and HERA feature multiple copies of the same baselines to repeatedly measure the same Fourier signal to increase sensitivity (Parsons et al. 2012a). To satisfy the sensitivity need, modern arrays are large. Ranging upwards from 100 elements, these arrays are typically of the drift-scan type to limit cost.

Analysis pipelines for the 21 cm power spectrum typically fall into two categories. In the first, images are formed in Fourier domain through rotation synthesis, followed by a foreground mitigation step to construct a power spectrum. An alternative technique works directly with visibilities from baselines, delay-transforming and cross-multiplying them to form the power spectrum. This technique avoids many systematics associated with combining data from different baselines and tracks the native sampling of the interferometer. An example of the visibility-based pipeline was presented in Ali et al. (2015), which provides the newest upper limit to the power spectrum measurements with the 64-element version of PAPER (henceforth as PAPER-64). However, one disadvantage of existing visibility-based pipelines is their lack of use of partial redundancy. Baselines of different lengths and orientations contain partially redundant information. While imaging based power-spectrum pipelines naturally includes all redundancy information,

[1] Astronomy Dept., U. California, Berkeley, CA
[2] Radio Astronomy Lab., U. California, Berkeley, CA

visibility-based pipelines so far only cross multiply fully redundant baselines, i.e. baselines of the same length and orientation.

Visibility-based pipelines are not fundamentally limited from using partially redundant baselines. In fact, most sensitivity forecasts to date do include partial redundancy (Pober et al. 2014; DeBoer et al. 2016; Dillon & Parsons 2016b). Recently, Paul et al. (2016) proposed a visibility-based approach to extract power spectrum from partially redundant baselines in sky-tracking measurements. Arrays capable of such tracking include MWA and LOFAR. Our work introduces a two-baseline power spectrum estimator that can be applied to both tracking capable arrays and drift-scan only telescopes, such as PAPER, HERA and potentially SKA-low.

The Earth's rotation causes the baselines in a drift-scan array to pick up different modes of the sky with time. Rotation-synthesis makes use of the rotation-induced uv coverage map to form image. In a visibility based pipeline, we can extract the same rotation-induced redundancy. In this point of view, baselines that are slightly different in length and orientation "rotate into" each other at a time delay. We can thus cross-multiply time-shifted visibilities, with the proper weighting, to form power spectra. Due to the large number of elements of modern arrays, the task of cross-multiplying every baseline against every other, scaling as number of array elements to the fourth power, can be computationally formidable, and many pairs of baselines provide only negligible redundancy information. Our contribution is thus twofold. First we introduce a formalism to estimate the power spectrum from pairs of partially-redundant baselines in a visibility pipeline. Secondly we show how to use the two-baseline estimator to automatically pre-select baseline pairs and time offsets, making the problem computationally efficient. More precisely, our formalism allows one to simultaneously identify the baselines that give good redundancy, find the time offset that corresponds to maximal redundancy for a given pair of baselines, and quantify the sensitivity associated with cross multiplying such a pair of partially redundant baselines, which in turns is used as weight to combine measurements in a power spectrum pipeline.

The rest of this paper is organized as follows. In section 2 we introduce some terminology and notation used in the rest of the paper. In section 3 we introduce the formalism for weighting partially redundant baselines. In section 4 we present numerical tests of this technique as well as the expected sensitivity improvement this method provides for HERA and PAPER-128 pipelines, and with section 5 we conclude.

2. NOTATION AND TERMINOLOGY

In order to avoid confusion and ambiguity for the rest of this paper, we introduce some terminology that may differ from what is commonly found in the literature.

We make the distinction between a *baseline*, which corresponds to two specific antennas, and a *class of baselines*, which refers to all baselines of the same length and orientation in a given array. Baselines of the same class are traditionally called "redundant baselines", because they measure the same Fourier mode in the sky. We shall call baselines in the same class *equivalent baselines*, and reserve the word *redundancy* of two baselines

in reference to a variable function of the relative time-offsets of their visibility time series. With this terminology, two equivalent baselines are fully redundant with each other simultaneously at all times. Non-equivalent baselines also have partial redundancy, and the redundancy can be maximized if their respected time series are shifted with respect to one another by some delay.
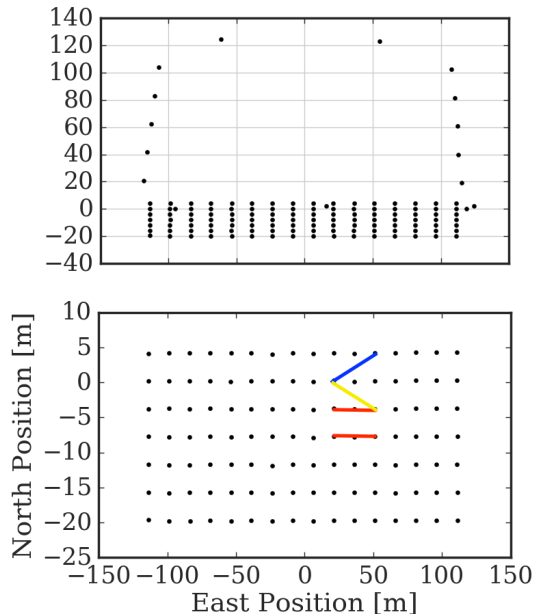

Fig. 1.— The PAPER128 layout. Each dot corresponds to the location of an antenna. The top panel shows the antenna positions drawn to scale; the bottom panel show the antenna labels and distances of the 112-element grid, in other words excluding the outrigger antennas. The numbering of the antennas in the bottom panel are original labels during instrument assembly and does not bear significant meaning. In the bottom panel, the two baselines marked with red segments are example of an equivalent pair, each with antennas separated by 2 units east and 0 unit north. Similarly, the baselines marked in blue and yellow are examples of classes that , due to the small North-South separation within the grid (see top panel), are expected to be nearly equivalent to the red baselines.

We shall use the 128-element PAPER array (henceforth as PAPER-128) to motivate our formalism and demonstrate our method, and extend our results to several HERA configurations in Sec 4.5. The PAPER array is located in the Karoo desert in South Africa (30:43:17.5 S, 21:25:41.8 E). The layout pattern with antenna labels is shown in Fig. 1. The array consists of a 112-element core is a rectangular grid, and 16 "outriggers" used primarily to aid calibration. In the bottom panel, the two baselines marked in red, are an example of equivalent pair. We denote a equivalency-class of baselines in the PAPER grid by their separations, in this case {2,0}, for the antennas are separated by 1 unit east and 0 unit north. Similarly, the baselines marked in blue and yellow are respectively examples of {2,1} and {2, -1}. Note {2,0} and {-2, 0} for example are the same class and should not be counted twice. Antennas in purely north-south baselines are close (4m), and hence these baselines are not suitable for cross-multiplication due to cross-coupling. On the other hand, the close North-South sep-

aration means that classes such as {2,0} and {2,1} are expected to be near-equivalent. The PAPER-64 analysis of Ali et al. (2015) used three classes of baselines, the PAPER-128 equivalent of which are {2, 0}, {2, 1} and {2, -1}. There each of these classes of baselines are cross multiplied within itself. This paper provides the method for inter-class multiplications. To do so we will use the short hand notation $\{m, n\} : \{m', n'\}$ to denote a pair of baseline classes to be cross-multiplied.

3. METHOD

In this section we introduce our method to cross-multiply nearly equivalent baseline classes.

3.1. *uvw tracks*

Radio interferometric observations are often described in the coordinates uvw define as:

$$(u, v, w) = \frac{\nu}{c}\boldsymbol{b}, \qquad (1)$$

where $\boldsymbol{b}$ is the baseline vector in Cartesian coordinates, with first and second coordinates pointing East and North, respectively. ν is the frequency of observation, and c is the speed of light. Relative to a phase center on the sky, each baseline maps to a point in uvw space. As the Earth rotates, the points trace out tracks in the uvw space. We show in Fig. 2 uvw tracks of the 3 PAPER-128 baselines colored in Fig. 1, projected onto uv and vw planes. The tracks are traced over 12 sidereal hours, at 0.15GHz, relative to a phase center that passes through zenith.

Equivalent baselines follow identical uvw tracks. Traditionally, we can identify redundancy of nearly equivalent baselines as crossings of the uv tracks, a 2 dimensional projection of Fig. 2. As we see in the top panel of Fig. 2, there would be many such crossings. However, there are several reasons that uv track-crossing do not imply perfect redundancy. In fact, in our case uv track-crossings are not accurate enough for time offset determination, nor can it give estimate of the degree of redundancy. The most obvious reason is that the 3 dimensional uvw tracks do not actually cross, as is evident from the vw projection in the bottom panel of Fig. 2. For drift-scan arrays, even a hypothetical crossing in uvw space would not imply perfect redundancy. This shall become evident in the next section, after we develop a more general formalism that accounts for the point spread function of the finite beams, and estimate the degree of redundancy for general combination of baselines at a general time-offset. Then, relation of track-crossing and redundancy for both drift-scan and tracking measurements will be explored in Appendix A.

3.2. *Formalism*

In this section we formulate a power spectrum estimator from the product of delay-transformed visibilities from two arbitrary baselines.

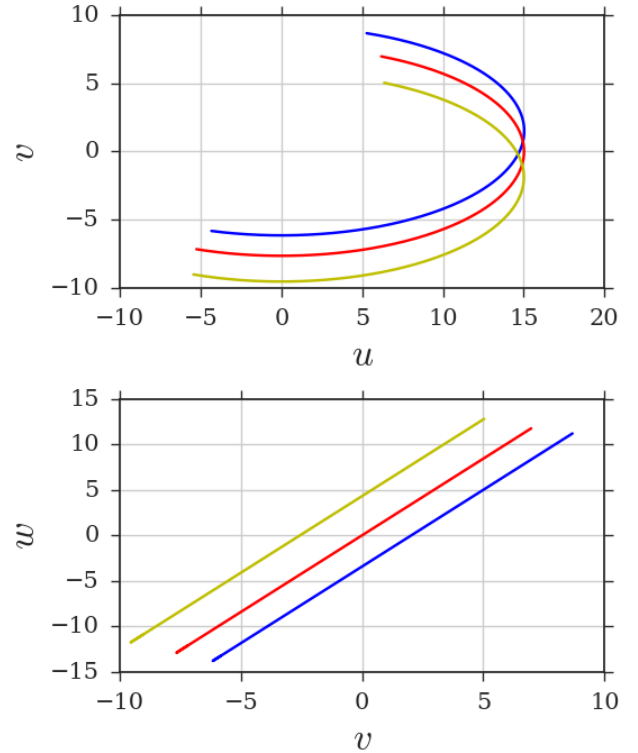We begin with the visibility as commonly defined in the literature (e.g. Thomson et al. 2017; Parsons et al.

FIG. 2.— Tracks of 3 PAPER-128 baseline classes shown in Fig. 1, here with the same respective colors. Tracks are traced out over half a sidereal day for frequency $\nu = 0.15$GHz and a phase center that passes through zenith. The top panel shows uv tracks with crossings among each pair of baseline classes. Projection to the vw plane in the bottom panel shows that tracks in uvw space do not cross.

2012a):

$$V_\nu(\boldsymbol{b}) = \int d\Omega A(\hat{\boldsymbol{s}}, \nu)\phi(\nu)I_\nu(\hat{\boldsymbol{s}}) \exp\left(-2\pi i \frac{\nu}{c}\boldsymbol{b}\cdot\hat{\boldsymbol{s}}\right),$$

$$\approx \frac{2k_B}{\lambda^2} \int d\Omega A(\hat{\boldsymbol{s}}, \nu)\phi(\nu)T(\hat{\boldsymbol{s}}) \exp\left(-2\pi i \frac{\nu}{c}\boldsymbol{b}\cdot\hat{\boldsymbol{s}}\right),$$

$$(2)$$

where k_B is Boltzmann's constant, λ is the mean wavelength, $\hat{\boldsymbol{s}}$ and Ω are a direction in the sky and its corresponding solid angle. Inside the integral we have $\phi(\nu)$ the frequency bandpass profile, $A(\hat{\boldsymbol{s}}, \nu)$ the (frequency-dependent) primary beam, and I the specific intensity, which has been related to T, the brightness temperature in the Rayleigh-Jeans limit. The beam power pattern A is dimensionless, normalized to 1 at its peak (zenith), and we assume it to be the same for all baselines.

Power-spectrum measurements are typically taken from ~ 10MHz centered around the corresponding redshift of interest (e.g. 150 MHz for z=9.5).

Delayed-transformed visibility has gained popularity in recent years since it has been shown that foregrounds are isolated in delay space. We define the delay-transformed visibility by Fourier transforming visibility along the fre-

quency axis (Parsons et al. 2012b) as:

$$V(\boldsymbol{b}, \tau) = \int d\nu V_\nu(\boldsymbol{b})\phi(\nu) \exp\left(2\pi i\nu\tau\right),$$

$$= \int d\Omega d\nu B(\hat{\boldsymbol{s}}, \nu) T(\hat{\boldsymbol{s}}, \nu) \exp\left[-2\pi i\nu\left(\frac{\boldsymbol{b}\cdot\hat{\boldsymbol{s}}}{c} - \tau\right)\right]. \tag{3}$$

Here the delay τ is the Fourier dual of ν. Eq. (3) expresses the delay-transformed visibility as an integral over observational coordinates $\hat{\boldsymbol{s}}$ and ν. For notational simplicity here and in the rest of this paper, we have defined the quantity

$$B(\hat{\boldsymbol{s}}, \nu) \equiv \frac{2k_B}{\lambda^2}\phi(\nu)A(\hat{\boldsymbol{s}}, \nu). \tag{4}$$

Ultimately, we would like to relate the data, collected with coordinates $\hat{\boldsymbol{s}}$ and ν, to the power spectrum, written with $\boldsymbol{r}$ and $\boldsymbol{k}$, the cosmological position coordinate and wavenumber. To do so, we start by noticing that

$$r = \frac{c}{H_0}\int_0^z \frac{dz'}{E(z')},$$
$$\approx \frac{c}{H_0}\int_0^{z_0} \frac{dz'}{E(z')} - \frac{c(1+z)^2}{\nu_{21}H_0 E(z)}\left(\nu - \nu_0\right), \tag{5}$$
$$\equiv X - Y\Delta\nu,$$

where $\nu_{21} = 1420\text{MHz}$ is the $21\,\text{cm}$ transition rest frequency, ν_0 a reference central frequency with corresponding redshift z_0, and

$$E(z) = \sqrt{\Omega_m(1+z)^3 + \Omega_\Lambda}. \tag{6}$$

Inverting for ν:

$$\nu = \frac{X - r}{Y} + \nu_{21}. \tag{7}$$

In the thin-shell limit, we can write:

$$d^2r = X^2 d\Omega. \tag{8}$$

Note Eq. (8) does not require the flat-sky approximation; the angular integral is still performed over the curved sky.

We can rewrite the delayed-transformed visibility as

$$V(\boldsymbol{b}, \tau) = \int \frac{d^3r}{X^2Y}B(\boldsymbol{r})T(\boldsymbol{r})\exp\left[-2\pi i\left(\frac{\boldsymbol{b}}{c}\cdot\hat{\boldsymbol{r}} - \tau\right)\nu_r\right], \tag{9}$$

where $d\nu = -dr/Y$ and $d^3r = -X^2Y d\Omega d\nu$. We have written ν_r as a reminder that ν and r are related by Eq. (7).

Existing visibility based power-spectrum pipelines for redundant drift-scan arrays relate the power-spectrum to the conjugate square of the visibilities (Parsons et al. 2012b, 2014; Ali et al. 2015). We would like to generalize such relations by relating the power-spectrum to the product of two visibilities from two arbitrary baselines and time-offsets. The beam pattern of a baseline shifts relative to the sky as the Earth rotates. Here we choose to fix the sky, and denote the rotated coordinates in the topocentric frame with the 3 dimensional rotation operator $\boldsymbol{\Gamma}$:

$$V_\psi(\boldsymbol{b}', \tau) = \int \frac{d^3r}{X^2Y}B(\boldsymbol{\Gamma r})T(\boldsymbol{r})$$
$$\exp\left[-2\pi i\left(\frac{\boldsymbol{b}}{c}\cdot\boldsymbol{\Gamma}\hat{\boldsymbol{r}} - \tau\right)\nu_r - i\psi_\nu\right], \tag{10}$$

where we introduced a frequency-dependent phase ψ_ν, which to first order corresponds to phasing the two visibilities to the same phase center. [3]

With implicit bounds of integrals from $-\infty$ to ∞, we have:

$$\langle V^*(\boldsymbol{b}, \tau)V_\psi(\boldsymbol{b}', \tau)\rangle = \int \frac{d^3r d^3r'}{(X^2Y)^2}\left(\langle T^*(\boldsymbol{r})T(\boldsymbol{r}')\rangle\right)B^*(\boldsymbol{r})B(\boldsymbol{\Gamma r'})\Phi(\boldsymbol{r}, \boldsymbol{\Gamma r'}),$$

$$= \int \frac{d^3r d^3r'}{(X^2Y)^2}\left(\int \frac{d^3\kappa}{(2\pi)^3}\frac{d^3\kappa'}{(2\pi)^3}\langle T^*(\boldsymbol{\kappa})T(\boldsymbol{\kappa}')\rangle e^{-i(\boldsymbol{\kappa}\cdot\boldsymbol{r} - \boldsymbol{\kappa}'\cdot\boldsymbol{r}')}\right)B^*(\boldsymbol{r})B(\boldsymbol{\Gamma r'})\Phi(\boldsymbol{r}, \boldsymbol{\Gamma r'}),$$

$$= \int \frac{d^3r d^3r'}{(X^2Y)^2}\left(\int \frac{d^3\kappa}{(2\pi)^3}P(\kappa)e^{-i\boldsymbol{\kappa}\cdot(\boldsymbol{r} - \boldsymbol{r}')}\right)B^*(\boldsymbol{r})B(\boldsymbol{\Gamma r'})\Phi(\boldsymbol{r}, \boldsymbol{\Gamma r'}),$$

$$\approx P(k_{b,\tau})\int \frac{d^3r d^3r'}{(X^2Y)^2}\delta_D^{(3)}(\boldsymbol{r} - \boldsymbol{r}')B^*(\boldsymbol{r})B(\boldsymbol{\Gamma r'})(\boldsymbol{r}, \boldsymbol{\Gamma r'}), \tag{11}$$

$$= P(k_{b,\tau})\int \frac{d^3r}{(X^2Y)^2}B^*(\boldsymbol{r})B(\boldsymbol{\Gamma r})\exp\left[-i2\pi\frac{\nu_r}{c}(\hat{\boldsymbol{r}}\cdot\boldsymbol{b} - \boldsymbol{\Gamma}\hat{\boldsymbol{r}}\cdot\boldsymbol{b}') - i\psi_\nu\right],$$

$$= P(k_{b,\tau})\int \frac{d\Omega d\nu}{X^2Y}B^*(\hat{\boldsymbol{s}}, \nu)B(\boldsymbol{\Gamma}\hat{\boldsymbol{s}}, \nu)\exp\left[-i2\pi\frac{\nu}{c}(\hat{\boldsymbol{s}}\cdot\boldsymbol{b} - \boldsymbol{\Gamma}\hat{\boldsymbol{s}}\cdot\boldsymbol{b}') - i\psi_\nu\right],$$

with

$$\Phi(\boldsymbol{r}, \boldsymbol{\Gamma r'}) \equiv \exp\left[i\frac{2\pi}{c}\left(\boldsymbol{b}\cdot\nu_r\hat{\boldsymbol{r}} - \boldsymbol{b}'\cdot\nu_{r'}\boldsymbol{\Gamma}\hat{\boldsymbol{r}'}\right)\right]\exp\left[-i2\pi\tau\left(\nu_r - \nu_{r'}\right) - i\psi_\nu\right]. \tag{12}$$

[3] Turns out for many cases, the first-order interpretation is largely sufficient, but we still keep the term to reserve the option to determine its exact form numerically.

Above δ_D is the Dirac delta-function and in transition from cosmological coordinates back to observing coordinates we have written $\hat{\boldsymbol{r}} \equiv \hat{\boldsymbol{s}}$. The third equality of Eq.(11) follows from assumption of translational invariance of statistics of the 21 cm field. For the inequality in the fourth line we have assumed that the 3D power spectrum varies negligibly over the k-space of interest and thus pulled out the integral a power spectrum centered at

$$k_{\boldsymbol{b},\tau} \equiv 2\pi\sqrt{\left(\frac{\tau}{Y}\right)^2 + \left(\frac{b}{\lambda X}\right)^2}. \qquad (13)$$

In Section 4.3 we will examine the exact form of the window function probed as well as its effect on foreground isolation. Since $\boldsymbol{\Gamma}$ is a sky rotation, it does not affect ν, hence we have taken ν_r outside the parenthesis in the phase term $\exp\left[-i2\pi\frac{\nu_r}{c}\left(\hat{\boldsymbol{r}}\cdot\boldsymbol{b} - \boldsymbol{\Gamma}\hat{\boldsymbol{r}}\cdot\boldsymbol{b}'\right)\right]$. Notice that the phase factor $\exp\left[-i2\pi\tau\left(\nu - \nu'\right)\right]$ drops out in the end. This means that the location and magnitude of the correlation peak does not depend on delay τ.

Since the beam pattern and bandpass are given in $\hat{\boldsymbol{s}}$ and ν, we convert the integral back to these coordinates to get the general relation between the delay-transformed visibilities and the power spectrum:

$$\langle V^*(\boldsymbol{b},\tau)V_\psi(\boldsymbol{b}',\tau)\rangle$$
$$\approx P(k_{b,\tau})\int\frac{d\Omega d\nu}{X^2 Y}B^*(\hat{\boldsymbol{s}},\nu)B(\boldsymbol{\Gamma}\hat{\boldsymbol{s}},\nu)$$
$$\exp\left[i2\pi\frac{\nu}{c}\left(\hat{\boldsymbol{s}}\cdot\boldsymbol{b} - \boldsymbol{\Gamma}\hat{\boldsymbol{s}}\cdot\boldsymbol{b}'\right) - i\psi_\nu\right]. \qquad (14)$$

We can therefore form the power spectrum estimator for the baseline pair $\{\boldsymbol{b},\boldsymbol{b}'\}$:

$$\hat{P}(k_{b,\tau}) \equiv \frac{V^*(\boldsymbol{b},\tau)V_\psi(\boldsymbol{b}',\tau)}{\Theta}, \qquad (15)$$

where the weight is defined as

$$\Theta \equiv \int d\nu\,\Theta_\nu, \qquad (16)$$

with

$$\Theta_\nu \equiv e^{-i\psi_\nu}\int\frac{d\Omega}{X^2 Y}B^*(\hat{\boldsymbol{s}},\nu)B(\boldsymbol{\Gamma}\hat{\boldsymbol{s}},\nu)e^{i2\pi\frac{\nu}{c}\left(\hat{\boldsymbol{s}}\cdot\boldsymbol{b} - \boldsymbol{\Gamma}\hat{\boldsymbol{s}}\cdot\boldsymbol{b}'\right)}. \qquad (17)$$

Notice that Θ has no dependence on τ. In Eq. (17) we introduced an extra phase ψ, the origin of which we shall explain in 3.3. We point out that although all our derivations focused on drift-scan telescopes, we can get the analogous result for tracking measurements simply by noticing that for a tracking primary beam, $\boldsymbol{\Gamma}$ becomes a rotation around zenith, and so:

$$\Theta_\nu \equiv e^{-i\psi_\nu}\int\frac{d\Omega}{X^2 Y}B^*(\hat{\boldsymbol{s}},\nu)B(\hat{\boldsymbol{s}},\nu)e^{i2\pi\frac{\nu}{c}\left(\hat{\boldsymbol{s}}\cdot\boldsymbol{b} - \boldsymbol{\Gamma}\hat{\boldsymbol{s}}\cdot\boldsymbol{b}'\right)}. \qquad (18)$$

Roughly speaking, Eqs. (15) through (17) tell us that the cross multiplications of visibilities at a time delay is proportional to the power spectrum times the Fourier transform of the cross multiplied beam patterns. As a check, when applied to equivalent baselines, $\boldsymbol{b} = \boldsymbol{b}'$,

$\hat{\boldsymbol{s}} = \boldsymbol{\Gamma}\hat{\boldsymbol{s}}$, and Eq.(14) reduces to Eq.(B9) of Parsons et al. (2014). With Eq. (15) and Eq. (16) we can, for any given pair of baseline classes and time delay, estimate the degree of redundancy, here represented by Θ. This allows us to achieve our goals stated in the introduction: to identify candidate baseline pairs with significant redundancy, to find the time offset that maximizes redundancy, and to quantify the degree of such redundancy. We can do all the above simply by computing the weight Θ from Eq.(16) for various time offsets.

3.3. *Rephasing*

If $\psi_\nu = 0$ in Eq. (17), Θ_ν at the peak of correlation is in general complex, and often far from real. Furthermore, this phase of peak correlation is inevitably frequency dependent. This frequency dependence of the phase would lead to destructive interference when we integrate over frequency, unless we correct Θ_ν by an extra phase ψ_ν.

To first order, the physical origin of ψ_ν lies in the two visibilities having different phase centers. By default the correlators of a drift-scan array phase the two visibilities both to zenith at the same time. When they are cross-multiplied with a time lag, the visibilities must be rephased before delay transform to the account for the movement of the zenith. The effect of the phase thus roughly corresponds in a shift of the delay mode measured, and we should expect ψ_ν to be in first order a linear function of ν:

$$\psi_\nu \sim 2\pi\Delta\tau\nu. \qquad (19)$$

In Fig. 3 we show the peak phases for for a given baseline pair of PAPER128 ($\{2,0\}$:$\{2,1\}$), comparing the drift-scan phase dependence with that of the same baseline with hypothetical tracking-elements (Eq. (17) and Eq. (18)). In the drift-scan case, we indeed see a linear relation corresponding to a delay of $\Delta\tau \approx 15.7$ns. In the tracking case, because of minimal zenith movement, only second order effects are observed. The origin of the second order effects can be seen as due to the w-term, or more precisely the fact that uvw tracks do not cross when their 2 dimensional projections do. We refer the reader to Appendix A for further explanation. In both cases, the full effects are encapsulated in the phase of Θ_ν and can be thus determined empirically without any additional computation.

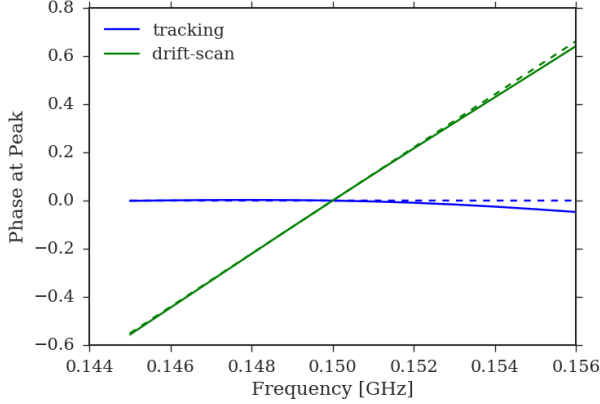FIG. 3.— Frequency dependent peak phases for PAPER-128 baseline pair {2,0}:{2,1}. Both the drift-scan case and a hypothetical tracking baseline are shown in solid lines. The first-order, linear effects are shown in dashed lines. We have also fixed the global phase in both cases to 0 at 0.15GHz. The drift-scan case exhibits linear behavior and conforms well to the first order effect of a shift in delay space due to the movement of zenith.
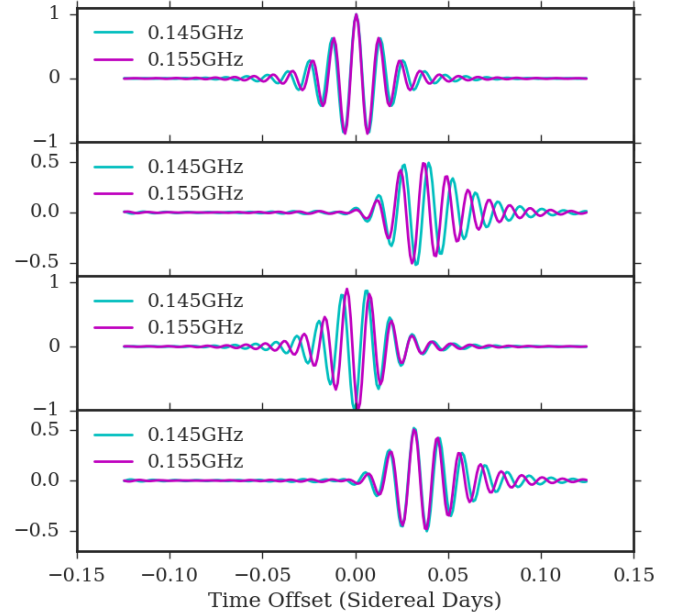

FIG. 4.— Comparisons of the peak phases of two different frequencies. Specifically shown are the real parts of Θ_ν. First and third panels shows equivalent baselines, second and fourth show a pair of nearly equivalents. The top two panels have zero phase shift ($\psi_\nu = 0$), and the bottom two are rephased to a time offset of 0.0325 sidereal days. The first and last panels thus show a coherently rephased series that would add constructively at the respective peaks.

4. ANALYSIS

In this section we delve into visualizing and testing our method, and explore sensitivity contributions of various baseline classes for PAPER and HERA array configurations.

4.1. Visualization

So far for clarity and generality we have avoided the traditional formulation of radio astronomy in terms of the uv-plane. To gain some visual intuition of the formalism in Section 3, we show explicit beam fringe patterns of the two baselines and their interactions. For this section, we relax into the flat sky approximation so that $(l, m, \sqrt{1 - l^2 - m^2}) \equiv \hat{s}$ and examine a single frequency of 150MHz.

At the time-offset of maximum redundancy, we expect two beams to have to same fringe pattern in both frequency and phase. Due to the time delay, however, the beam centers would be slightly shifted with respect to each other. This we show in the top panels of Fig. 5. The top left and middle panels show the beam-fringe patterns (real parts) for baselines {2,0} and {2,1}, delayed by 0.0325 sidereal days. The product of the beam fringe patterns in top right shows that the fringes cancel out as we expect. On the bottom of Fig. 5 we show the instantaneous uv coverage of two baselines and their product. We see from the bottom right panel that the cross-multiplied baselines recover power concentrated at $u \approx \frac{30\mathrm{m}}{c/1.5 \times 10^8 \mathrm{Hz}} \approx 15$. Note the bottom right panel is not the Fourier transform of the top right panel.

To illustrate the effect of rephasing, we compare in Fig. 4 real parts of Θ_ν of both equivalent and nearly equivalent baseline pairs for two channels: 0.145 GHz and 0.155 GHz. The top two panels have zero rephasing, and the bottom two are linearly rephased to a differential delay of $\Delta\tau \approx 15.7$ns, as we determined above. The first and third panels show the equivalent baseline pairs {2,0}:{2,0}, and second and fourth panels show {2,0}:{2,1}. Summing over frequency without proper rephasing leads to destructive interference and signal loss. The wider the frequency profile, the more destructive the interference would be. Only after rephasing to the correct time-offset for each baseline pairs can we constructively combine the frequency channels.

We see from Fig. 4 that although the amplitude of correlations match up for all time offsets, the phase would only locally match. The need for rephasing can thus be understood as a symptom of the underlying spatial decoherence; coherence at the phase center would not extend to the entire beam pattern (Fig. 5). Thus while integrating over frequencies and sky-direction in Eq. (16) there are necessarily sky-directions that do not add coherently. This decoherence across the beam pattern cannot be removed and plays a fundamental role in determining how much sensitivity one can recover from nearly equivalent baselines.
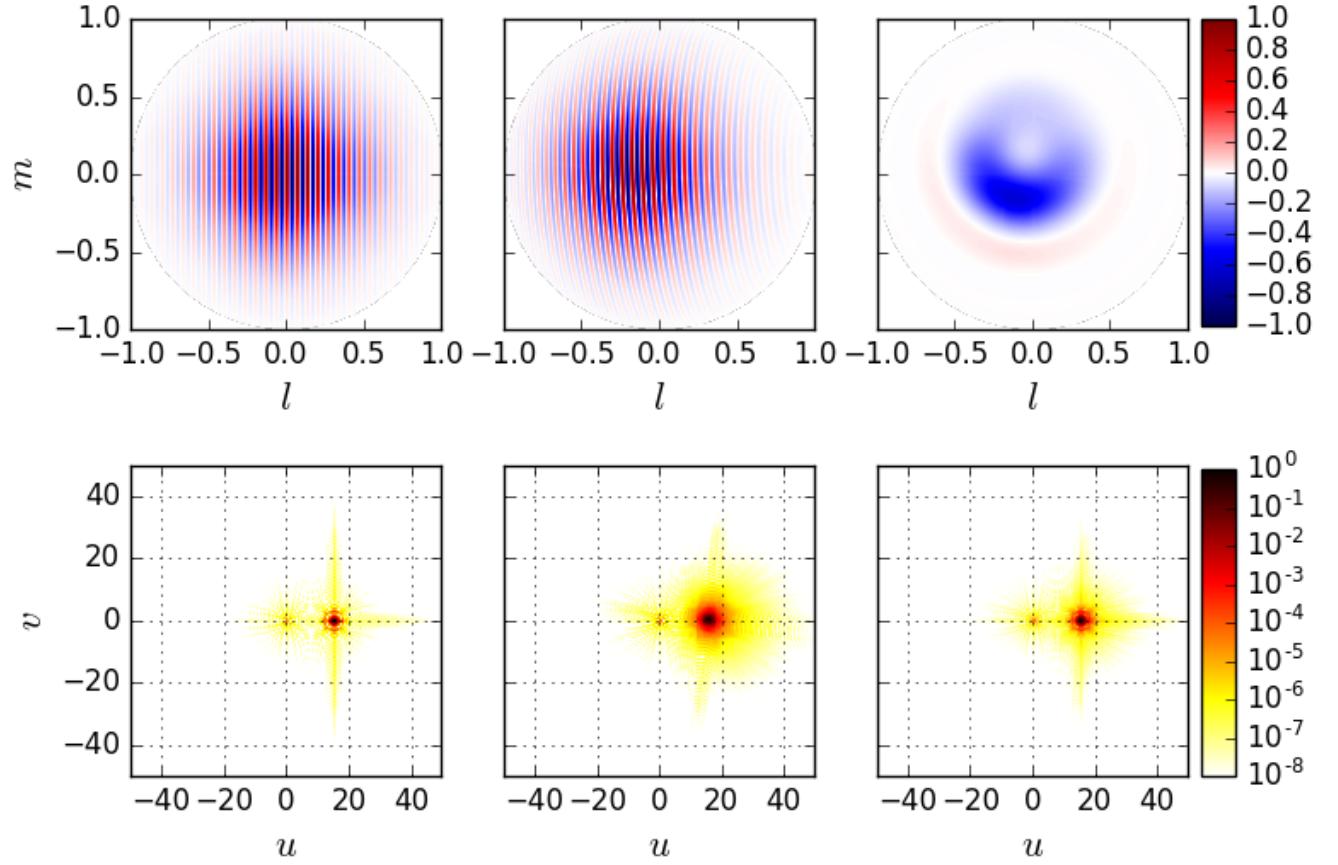
FIG. 5.— PAPER128 beam-fringe patterns and their point-spread functions on the uv-plane for a frequency of 150MHz. Selected baselines have 30 meter East-West separation, and 0 and 4 meter North-South separation, respectively. The top panels show the real parts of the beam-fringe patterns of the baselines (left and middle), and their conjugate product (right). The beam-fringe values are normalized such that the peak of the original beam is unity. The bottom left and middle panels show the peak-normalized uv beam power of the two beam-fringe patterns. The product of the beam spreads displayed in the bottom right panel shows that power concentrated at $u \approx 15$ is recovered.

4.2. Numerical Test

Bringing together the discussion from Section 3, we present a numerical check of Eq. (14), including rephasing. To do so, we need to compare the amplitude and phase of the integral weight Θ for a pair of baselines with products of simulated visibilities of those baselines. We use 10 frequency channels evenly spaced from 145MHz to 155MHz for the comparison. We rephase one of the simulated visibilities, as well as computed Θ_ν to the same delay as computed from the peak phase of the latter.
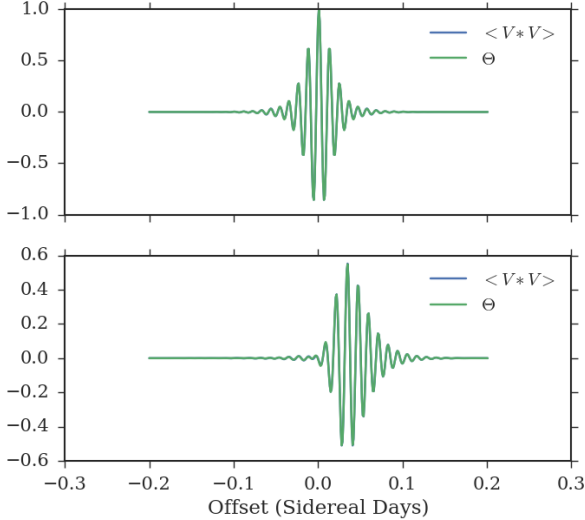

FIG. 6.— Numerical comparisons of the visibility correlation peaks to the Ω factor in Eq.(14). We generated a Gaussian random sky on healpix maps, computed visibilities and cross correlated them to find the correlation. Top panel shows the equivalent baseline pairs {2,0} against {2,0}, bottom panel shows {2,0} against {2,1}. Evaluations of weight Θ are shown green, and the visibility correlations of simulated random sky is given in blue. In both cases only the real parts are shown. We see that in both cases the theory and simulation line up in both amplitude and phase.

Fig. 6 shows the real parts of the resulting degree of redundancy as a function of time offset, for equivalent baseline pairs ({2,0}:{2,0}) in the top panel and nearly equivalent ({2,0}:{2,1}) in the bottom. Shown in green for the two cases are the computed values of Θ, both normalized to the peak of the equivalent case. The plot on the bottom is rephased with a delay of $\Delta\tau \approx 15.7$ns as outlined in Section 3.3. Since both y-axes are normalized to the same scale, we see the equivalent baseline reach maximal redundancy with 0 time-offset, while the nearly equivalent pair reaches maximum redundancy of $\Theta \approx 0.5$ at $dT \approx 0.325$days.

Almost completely overlapping with the green curves are the cross-multiplied visibilities from a simple simulation, shown in blue. For the simulation, we populate random values of brightness temperature on a healpix map (Gorski et al. 2005; Gorski et al. 1999) [4]. We rotate the baseline positions with the appropriate rotation matrix, multiplying the sky by the primary beam to get

the visibilities, for each baseline [5]. For the nearly equivalent case, we rephase the delayed visibility ({2,0}) by the same delay $\Delta\tau \approx 17$ns, followed by delay-transform. The resulting delay-space visibilities for the two baselines are then convolved via the Fourier convolution theorem, to obtain values of the cross correlation as a function of time-offset. We do this for both the equivalent ({2,0}:{2,0}) and nearly equivalent ({2,0}:{2,1}) case. Since the blue and green curves overlap, we have verified that in this case Eq. (14) is valid and the computed Θ can be used as weight to obtain the power spectrum.

4.3. Chromaticity

A crucial concern when measuring the 21 cm power spectrum is the possibility of astrophysical foreground contamination. Being orders of magnitude larger than the cosmological signal, even small residuals from an imperfect isolation of these foregrounds can masquerade as a false detection. Recent literature treatments of this problem have shown that if one combines the mathematical properties of interferometric measurement with the empirical fact that foregrounds are spectrally smooth, a clean separation of foregrounds and cosmological signal can in principle be achieved in harmonic space (Liu et al. 2014a,b). These studies suggest that foregrounds preferentially appear in harmonic modes of the sky corresponding to fluctuations that are angularly fine but spectrally smooth, a region in harmonic space colloquially known as "the wedge". One's foreground mitigation strategy can then consist mainly of avoiding those modes by working in the complementary region known as "the EoR window". However, for this strategy to be viable, it is necessary to show that one's data analysis pipeline does not corrupt the separation between the wedge and the EoR window. We will now do so for our proposed two-baseline power spectrum estimator.

Traditionally, spectral and angular dimensions are cylindrically binned with the coordinates $k_\perp$ and $k_\parallel$, where the angular direction $k_\perp$ is probed by the beam and line-of-sight direction $k_\parallel$ is probed by the frequency spectrum. Despite their usefulness, however, $(k_\perp, k_\parallel)$ are local coordinates that are well defined only in the narrow field-of-view limit. As discussed in previous sections, the combination of information from nearly equivalent baselines for wide-field telescopes is inherently a three-dimensional problem. Although the estimator of Section 3 does not have such limitations, to discuss foreground signature with possible curved sky effects, we must employ a set of basis functions where angular fluctuations are encoded by spherical harmonics, and radial fluctuations by spherical Bessel functions.

Liu et al. (2016) discusses in detail power spectrum analyses using spherical harmonics. In such as basis, modes in the sky are indexed by the magnitude k of the wavevector $\mathbf{k}$ and spherical harmonic indices (ℓ, m). In

---

[4] We use functionalities in the python package AIPY for healpix mapping as well as coordinate transforms.

[5] There are two obvious ways to achieve the rotation. One can either fix the sky and rotate the baselines, or the other way around. We found however, that we must not physically rotate the sky map, for the numerical round-offs due to finite resolutions of the map turns out to be significant. Thus we let the sky, represented by the healpix map, be fixed, and rotate the baselines.

other words, one defines

$$\overline{T}_{\ell m}(k) \equiv \sqrt{\frac{2}{\pi}} \int d\Omega dr r^2 j_\ell(kr) Y_{\ell m}^*(\hat{\mathbf{r}}) T(\hat{\mathbf{r}}), \qquad (20)$$

where $Y_{\ell m}$ is a spherical harmonic function and j_ℓ is the ℓth-order spherical Bessel function of the first kind. The power spectrum is then related to these spherical harmonic Bessel modes via the relation

$$\langle \overline{T}_{\ell m}(k) \overline{T}_{\ell' m'}^*(k') \rangle = \frac{\delta^D(k-k')}{k^2} \delta_{\ell\ell'} \delta_{mm'} P(k). \quad (21)$$

A delay-transformed visibility $V(\mathbf{b}, \tau)$ is related to the spherical Fourier-Bessel modes by

$$V(\mathbf{b}, \tau) = \sqrt{\frac{2}{\pi}} \sum_{\ell m} \int dk\, k^2 g_{\ell m}(k; \mathbf{b}, \tau) \overline{T}_{\ell m}(k), \quad (22)$$

where

$$g_{\ell m}(k; \mathbf{b}, \tau) = \int d\Omega d\nu B(\hat{\mathbf{r}}, \nu) Y_{\ell m}(\hat{\mathbf{r}}) j_\ell(kr) e^{i2\pi\nu(\tau - \mathbf{b}\cdot\hat{\mathbf{r}}/c)}$$
$$(23)$$

and the analogous quantity for V_ψ is given by

$$g_{\ell m}^\psi(k; \mathbf{b}, \tau) = \int d\Omega d\nu B(\mathbf{\Gamma}\hat{\mathbf{r}}, \nu) Y_{\ell m}(\hat{\mathbf{r}}) j_\ell(kr)$$
$$\times e^{i2\pi\nu(\tau - \mathbf{b}\cdot\mathbf{\Gamma}\hat{\mathbf{r}}/c) + i\psi_\nu}. \quad (24)$$

Note that unlike the primary beam and the fringe pattern, the spherical harmonic function does *not* rotate, since it originated from a spherical harmonic expansion of the sky temperature, which is fixed. With these expressions and a little algebra combining Eqs. (21) through (24), we arrive at a form of the two-baseline estimator (compare with Eq. (14)):

$$\hat{P}(k) = \frac{\langle V^*(\mathbf{b}, \tau) V_\psi(\mathbf{b'}, \tau) \rangle}{\Theta}$$
$$= \sum_\ell \int dk\, W_\ell(k; \mathbf{b}, \mathbf{b'}, \tau) P(k), \qquad (25)$$

where we have defined the *window function*

$$W_\ell(k; \mathbf{b}, \mathbf{b'}, \tau) = \frac{2k^2}{\pi\Theta} \sum_m g_{\ell m}(k; \mathbf{b}, \tau) g_{\ell m}^{\psi*}(k; \mathbf{b'}, \tau). \quad (26)$$

By construction, the window function sums to unity when integrated over all k and summed over all ℓ. We may therefore interpret our estimator $\hat{P}$ of the power spectrum to be a weighted average over all possible ℓ and k modes, with the window function providing the weights of this average. This means that window functions can serve as indicators of foreground leakage into the EoR window. One writes down the estimator for a hypothetical measurement of a power spectrum mode within the EoR window, and additionally computes the window function for that estimator. Ideally, the window function for $\hat{P}(k)$ will be sharply peaked around k. In general, however, window functions will have wings that encroach upon other modes. If the window function is substantially non-zero at low-k modes—which is where foreground emission typically resides, given its spectral smoothness—then the EoR window will be contaminated by foregrounds.

Figure 7 shows four example window functions. The two window functions in the top plot are for equivalent baseline power spectrum estimators, i.e., where one squares the delay spectrum from a single baseline and normalizes the result. The window function localized at low ℓ is for the baseline {2,0} of around 30 m in length, while the window function localized at high ℓ is for {12,0}, roughly 180 m in length. The window functions for near-equivalent baseline estimates are shown in the bottom plot, for {2,0}:{2,1} and {12,0}:{12,1} at low- and high-l, respectively. In all cases, the window functions that we show are for $\tau = 703$ ns, and were generated using primary beam models for PAPER. For numerical convenience we assumed that the bandpass ϕ_ν (see Eq. (4)) takes the form of a half-period sine curve that peaks at 150 MHz and goes to zero 5 MHz on either side of the peak.

One sees that all four window functions are fairly localized around specific values of k, which to a good approximation are described by Eq. (13). This suggests that both equivalent- and near-equivalent-baseline delay-spectrum estimators are good estimators of the power spectrum. At higher ℓ, the window functions for both estimators become elongated in k, with stronger tails towards the low k modes where foregrounds reside. This is the phenomenology of the wedge, where at fine angular scales the foreground leakage to higher k modes are more pronounced. Importantly, however, we note that the window functions for the near-equivalent-baseline estimator are no more elongated than for the equivalent-baseline estimator. One may thus conclude that the our proposed near-equivalent-baseline estimator does not result in extra foreground contamination of the EoR window.
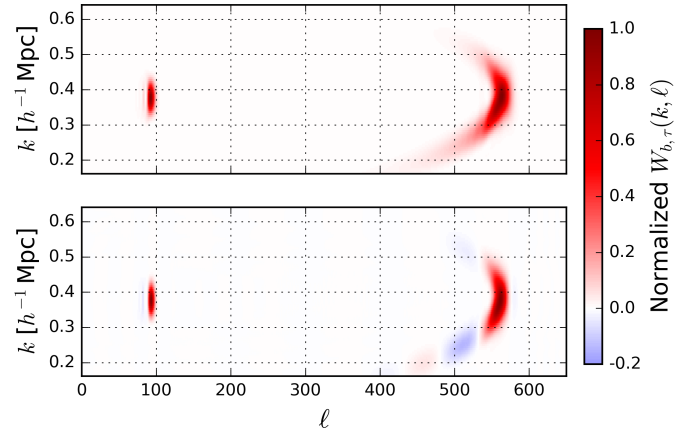

FIG. 7.— Sample window functions for delay-spectrum estimates of the power spectrum explored in this paper. The top plot shows window functions for a traditional equivalent-baseline delay-spectrum estimate, where a single baseline's delay spectrum is squared. The bottom plot shows the window functions resulting from nearly equivalent baselines. In each plot, the window function on the right corresponds to the window function for baselines of length $b \sim 180$ m in the East West direction, while the window function on the left is for $b \sim 30$ m. The general extent of the window functions for the two-baseline estimator are seen to be roughly the same as for the single-baseline estimator, suggesting that the former is just as good as the latter in keeping the EoR window free of foregrounds.
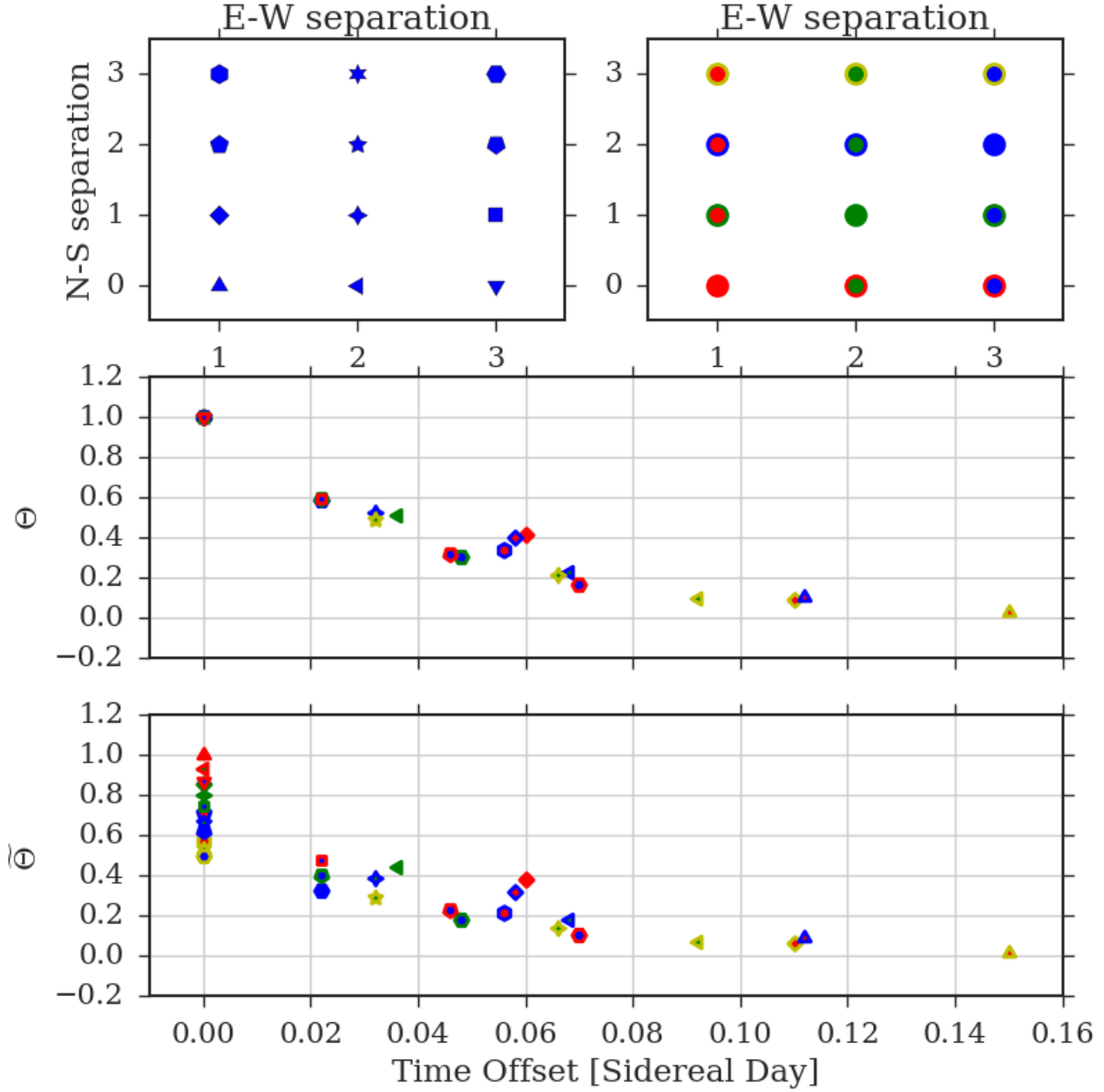
FIG. 8.— Relative sensitivity contributions of selected baseline combinations in PAPER-128. Each labeled point in the middle and bottom panels correspond to a cross-correlated baseline pair {m,n}:{p,q}. The shape of the symbol encodes the first baseline {m,n}, as displayed in the top left legend panel. The edge and face colors encode the second baseline {p,q}, as displayed in the top right legend panel. The middle panel shows the peak height (Θ) of each baseline combination, while the bottom panel multiplies the heights by the corresponding multiplicities as in Eq. (27). In both the middle and bottom panels, we have chosen to fix the value of {1,0}:{1,0} to unity.
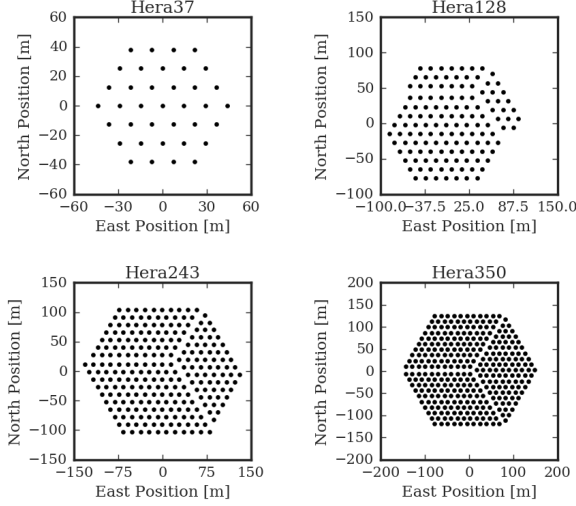
FIG. 9.— Planned Hydrogen Epoch of Reionization Array antenna configurations. HERA 37 is expected to complete and start collecting data in summer of 2017, and the other three are planned configurations in the next phases. For HERA350, only the 320 elements in the core are shown.

4.4. Sensitivity

In this section we discuss the sensitivity contributions of a variety of baseline pairs from PAPER128. Recall the sensitivity contributions of a particular baseline pair is given by Θ (Eq. (16)).

In the middle panel of Fig. 8 we show the peak heights Θ and locations ΔT for a variety of baseline combinations. Baseline pairs that are mirror images of each other give the same amount of redundancies (Θ), with the opposite time offset, as expected from symmetry. For example, {1,0}:{1,1} is mirror image of {1,0}:{1,-1} and these two pairs of baselines give the same sensibility contribution. Thus we only show a subset of representative baseline pairs to illustrate the contributions. For a more complete result see Fig. 10.

Baseline pairs that have smaller optimal time delays tend to have higher correlations. In other words, correlation peaks that are closer to zero time offset are higher. This is expected for two reasons; number one is that the longer the time delay that produces redundancy, the more the sky has moved with respect to the beams and hence the less overlaps in patch of sky surveyed. The other reason is that smaller optimal time-offset corresponds to smaller differences in orientation and length of the pair of baselines and hence better redundancy.

To determine that actual relative contribution to sensitivity of these baseline pairs, we have to account for the multiplicities of these baseline classes, in other words the number of antenna pairs with the same length and orientation. We would like to estimate an effective weight $\widetilde{\Theta}$ that accounts for both the peak height Θ and the multiplicities of the baseline classes. From Fig. 8 we see for example {1,0} has higher multiplicity than {2,0}, or {1,1}. The latest release of PAPER64 data uses the PAPER128-equivalent baselines {2,1}, {2,0} and {2,-1} (Ali et al. 2015). There, the three sets of equivalent base-

line classes are only cross multiplied with itself. Assuming that each equivalent baseline delivers the same Θ, the relative contribution to sensitivity can be estimated.

First we can average the visibilities of the equivalent baselines. Since the core of PAPER-128 has 16 by 7 antenna configuration, there are $M \equiv (16 - |m|) \times (7 - |n|)$ copies of the baseline class $\{m, n\}$. This means that if we add visibility measurements of all the equivalent baselines, we get a factor of $\sqrt{M}$ reduction in noise level σ_N of the visibility. The sensitivity contribution of $\{m, n\}$, cross multiplied with $\{m', n'\}$ thus roughly scales as $\sqrt{((16 - |m|)(7 - |n|)(16 - |m'|)(7 - |n'|))} = \sqrt{MM'}$. For cross-multiplications of nearly equivalent baselines of types $\{m, n\}$ and $\{m', n'\}$, we get an effective weight:

$$\widetilde{\Theta}_{bb'} \propto \Theta_{bb'} \times \sqrt{MM'}. \qquad (27)$$

Shown in the bottom panel of Fig. 8 are the peak heights weighted by the multiplicity factor. Points that have zero time delay are the equivalent baseline pairs and their weighted correlation values simply reflect the multiplicity factor. For clarity of presentation we have "folded over" the negative time delays and combined baseline pairs that are identical modulus parity. We point out that the data points shown here are not all the cases of highest correlation.

With $\widetilde{\Theta}$, we can estimate the power spectrum by inverse covariance weighting:

$$\begin{aligned} P(k_\tau) &= \frac{\sum_{bb'} P(k_{b,\tau})/\sigma_P^2(bb')}{\sum_{bb'} \sigma_P^2(bb')}, \\ &= \frac{\sum_{bb'} P(k_{b,\tau})\widetilde{\Theta}_{bb'}^2}{\sum_{bb'} \widetilde{\Theta}_{bb'}^2}, \end{aligned} \qquad (28)$$

where the sum is over classes of baseline pairs. We define the estimator sensitivity to be the inverse of the power spectrum noise variance:

$$\rho \propto \frac{1}{\sigma_P^2} \propto \rho_0^2 \sum_{bb'}^{N} \widetilde{\Theta}_{bb'}^2, \qquad (29)$$

where, if σ_S^2 and σ_N^2 are the characteristic signal and noise levels of a single-baseline visibility, $\rho_0 \equiv \sigma_S^2/\sigma_N^2$ is the signal to noise ratio.

The scaling in Eq. (27) was rough for simplicity of motivation. As we derive in Appendix B, this weight should be corrected by a factor proportional ρ_0:

$$\widetilde{\Theta}_{bb'} = \frac{\Theta_{bb'}\sqrt{M_b M_{b'}}}{\sqrt{1 + \rho_0 (M_b + M_{b'})}}. \qquad (30)$$

For a given ρ_0 Eq. (30) quantifies the relative sensitivity contribution of a baseline pair bb'. Assuming a reionization signal of $\Delta_{21cm}^2 \sim 30mK^2$, observation centered at 150MHz ($z = 8.5$), and 120 days of integration with PAPER antennas, we have roughly (See Eq.(20) in Parsons et al. (2012a))

$$\rho_0 \sim 0.001 \left[\frac{L}{40m}\right] \left[\frac{0.1hMpc^{-1}}{k}\right]^3, \qquad (31)$$

where L is the average baseline length between the pair.

As expected, baseline-pairs that have smaller $\widetilde{\Theta}$ contribute less to the sensitivity.

4.5. *Array Configuration Comparisons*

We run our algorithm over all possible baseline-pairs of PAPER128, HERA37, HERA128, HERA243 and HERA350. The HERA antenna configurations are shown in Fig. 4.3. The hexagonal design is the densest pattern of antenna-packing. The larger arrays are designed with a "gap" dividing the antennas into three different groups. The gaps are designed so as to improve uv coverage and ease calibration without compromising sensitivity, but also produces many more nearly equivalent baselines than a pure hexagonal layout. The motivations behind the designs are explained in Dillon & Parsons (2016a). Compared to PAPER128, the hexagonal pattern of HERA lack short baselines oriented close to each other, and the smaller beam means that we expect to see only longer nearly equivalent baselines. The lower multiplicities per class of baselines is compensated by the larger number of classes of baseline-pairs, especially given the gap in the larger versions.
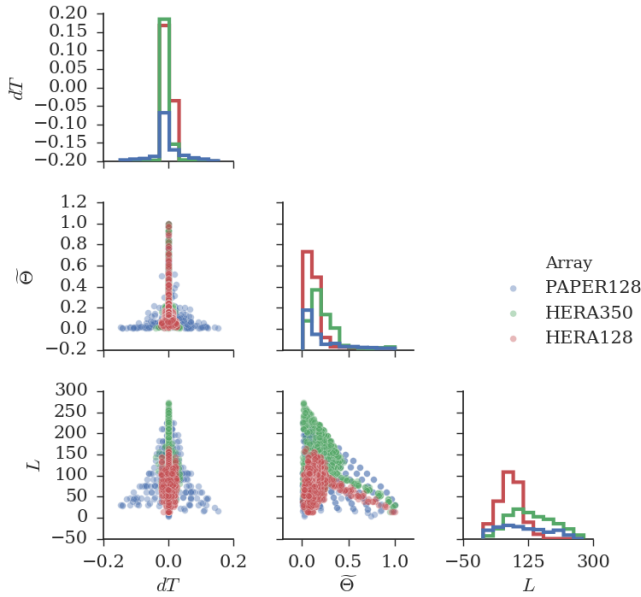

FIG. 10.— Pairplots of the top contributing baseline pairs in three arrays. Plotted properties are optimal time delay dT, effective weight $\widetilde{\Theta}$ and baseline length L. Only those points with $\widetilde{\Theta} > 0.01$ are shown (the weight for the top class of equivalent pair is normalized to 1). Only one baseline length is shown since all top contributing pairs have very similar lengths as expected. The scatter plots are shown with transparency so that darker regions indicate degeneracies. Scatter points of the 3 arrays overlap, in order indicated by the legend.

In Fig. 10 we present pair-distributions of 3 different properties for baselines that contribute well to the sensitivity ($\widetilde{\Theta} > 0.01$, where again the weight for the top class of equivalent pair is normalized to 1). The properties shown are effective weight $\widetilde{\Theta}$, optimal time offset dT, and average baseline length L. We only show 3 of the mentioned arrays for visual clarity. The other two results are similar barring intuitive differences. We discuss each

cross-distribution plot:

- dT vs. $\widetilde{\Theta}$: This relation is familiar from Fig. 8. The points at $dT = 0$ are the equivalent baselines. Note PAPER128 has more data points with high dT.

- dT vs. L: PAPER128 shows the trend that longer baselines correspond to lower dT. HERA arrays do not exhibit this trend here because of a selection effect. Only baseline pairs with $\widetilde{\Theta} > 0.01$ are shown. Shorter HERA baselines require much longer dT to overlap, partly because of the hexagonal structure requires $60°$ of rotation to overlap baselines, partly because of the smaller primary beam. Thus most short HERA baseline pairs are not shown in this figure.

- $\widetilde{\Theta}$ vs. L: The general trend to be seen here is that longer baselines tend to have lower $\widetilde{\Theta}$. This is due to the lower multiplicity of longer baselines. The trend is particularly obvious in the linear structure for the HERA arrays, which are the equivalent baselines. Since they all have the unity Θ by definition, the linearly decreasing trend of $\widetilde{\Theta}$ is a direct measure of the baseline multiplicity structures of the HERA array configurations.

- The top nearly equivalent pairs in HERA the same Θ (not shown) as in PAPER128, but much lower $\widetilde{\Theta}$. This is because they are longer baselines with lower multiplicity. In the end these baseline classes still lead to high contributions to total sensitivity (Fig. 11) because there are a lot more such baseline pairs for HERA.

Having quantified the sensitivity from a given pair of baselines, we study the cumulative sensitivity of the array depending on which baseline pairs we include. Evidently we should prefer the pairs with larger $\widetilde{\Theta}$. In Fig. 11 we plot ρ against the minimum $\widetilde{\Theta}$. $\rho(\widetilde{\Theta}_{min})$ is the sensitivity of the array when baseline-pairs that have $\widetilde{\Theta} > \widetilde{\Theta}_{min}$ are included. The dashed lines represent the values when only the equivalent baseline-pairs are used. We see as expected that in all cases using the nearly equivalent baselines lead to more and more significant improvements with lower $\widetilde{\Theta}_{min}$, or in other words when more baseline pairs are used. The small HERA37, with no gap (like in HERA350) or short nearly equivalent baselines (like in PAPER 128), will not benefit much from the nearly equivalent baselines. The maximum benefits for other cases are expected to be around 20% to 60%. PAPER128 is designed with highly redundant nearly equivalent baselines, and thus these baselines start contributing at higher $\widetilde{\Theta}_{min}$, but the gapped HERA configurations will benefit even more from nearly equivalent baselines at low $\widetilde{\Theta}_{min}$ due to the presence of more classes of such pairs. Note that here we normalized ρ such that the contribution of the top equivalent baseline pair, such as are 1. This plot therefore does not compare the absolute sensitivity across the different arrays. The stepwise pattern is characteristic of a regular grid; as we

step to lower $\widetilde{\Theta}$ large groups of baseline pair classes get included in "batch".
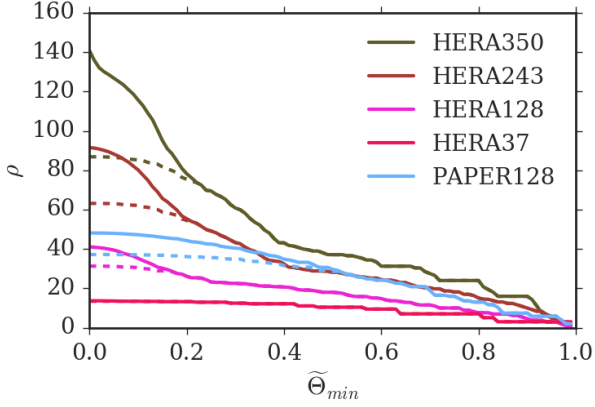

FIG. 11.— Sensitivity of redundant arrays as a function of the minimum effective weight. Dashed lines represent when only equivalent baseline pairs are used, while solid lines indicate use of both the equivalent and nearly equivalent baselines are used. The y-axis is normalized independently for each array such that the contribution of the top single equivalent pair class is unity, and thus does not indicate a comparison of absolute sensitivity across the different arrays shown.

5. CONCLUSION

Redundant arrays are designed to maximize sensitivity. Current generations of redundant radio arrays, such as those probing the power spectrum of the epoch of reionization could benefit from data analysis techniques that improve the sensitivity. We present a visibility and delay-transform based method of cross-multiplying baselines that are close in length and orientation to each other, thereby extracting the sensitivity contained in such partial redundancy. Our discussion focuses on drift-scan arrays but also provide the analogous expressions for tracking measurements. Our method relies on cross-multiplications of visibilities at a time-offset. For drift-scan arrays, the movement of zenith during this time-offset requires a frequency-dependent rephasing prior to delay transform. We show that such zenith-rephasing lead to manageable chromatic leakage. Given an antenna array configuration, our method identifies the best baseline pairs to cross-multiply and predict the optimal time-offset ΔT, weight Θ and rephase delay $\Delta\tau$. With the predicted results one can incorporate partial redundancy into existing delay-tranform based power-spectrum pipelines through a few steps. 1). Rephase the visibilities prior to delay transforming by $\Delta\tau$. 2). Shift the visibilities in time by ΔT. 3) Cross multiply the visibilities of the two baselines to form the power spectrum. 4) combine the different baseline pairs by appropriate inverse-variance weighting that takes into account the predicted sensitivity contributions of each case. We showed that incorporation of partial redundancy accounts for 20% to 60% sensitivity for various configurations of PAPER and HERA.

APPENDIX

A.
TRACK-CROSSING AND W-TERM

In this section we link our results from Section 3 to the traditional views of rotation synthesis and in particular uv tracks. Typical convention of rotation synthesis has rotation operates on $\boldsymbol{b}'$ instead of $\hat{\boldsymbol{s}}$, in which case Eq. (16) becomes

$$\Theta \equiv \int \frac{d\Omega d\nu}{X^2 Y} B^*(\hat{\boldsymbol{s}}, \nu) B(\boldsymbol{\Gamma}\hat{\boldsymbol{s}}, \nu) e^{-i2\pi\frac{\nu}{c}\hat{\boldsymbol{s}}\cdot(\boldsymbol{b}-\widetilde{\boldsymbol{\Gamma}}\boldsymbol{b}')}, \tag{A1}$$

where $\widetilde{\boldsymbol{\Gamma}}$ is the inverse of $\boldsymbol{\Gamma}$. For tracking elements, $\boldsymbol{\Gamma}$ becomes a rotation around zenith:

$$\Theta \equiv \int \frac{d\Omega d\nu}{X^2 Y} B^*(\hat{\boldsymbol{s}}, \nu) B(\hat{\boldsymbol{s}}, \nu) e^{-i2\pi\frac{\nu}{c}\hat{\boldsymbol{s}}\cdot(\boldsymbol{b}-\widetilde{\boldsymbol{\Gamma}}\boldsymbol{b}')}, \tag{A2}$$

Track-crossing corresponds to vanishing of the exponent

$$\boldsymbol{b} - \widetilde{\boldsymbol{\Gamma}}\boldsymbol{b}' = 0. \tag{A3}$$

If exponent vanishes, Θ_ν is real, and one would not observe any non-zero phase at the peak. However, with inclusion of the w-term, we see that track crossings in uv-plane do not actually imply crossing in the uvw space, therefore the exponent Eq. (A3) does not vanish and we observe non-zero, frequency dependent peak phases even in the case of tracking arrays, as shown in Fig. 3.

Furthermore, in the case of drift-scan arrays, even a crossing in uvw space does not necessarily correspond to the maximum of correlation. From Eq. (A1), we see the track-crossing condition Eq. (A3) maximizes Θ if and only if no other term in the integral depends on $\boldsymbol{\Gamma}$. We see that this is true for the tracking case (Eq. (A2)), but not the drift-scan case (Eq. (A1)). Only in the special case of a tracking measurement, where the w-term happens to be negligible, do uv-track crossings correspond exactly to maxima of correlations. Incidentally, at the crossing points the exponent in Eq. (A3) vanishes and ψ_ν is zero for all ν. In general, second order effects due to the presence of w is observed for the tracking measurements as in Fig. 3.

B.

DERIVATION OF NOISE COVARIANCE

In this appendix we give a brief derivation of the effective weight quoted in 4.4. We combine the power spectrum measurements from distinct baseline classese by inverse variance weighting. We separate the visibility and power spectrum into signal and noise contributions:

$$V = V_S + V_N,$$
$$P = P_S + P_N. \tag{B1}$$

We write the noise variance of power spectrum and visibility as:

$$\sigma_V^2 = \langle |V_N|^2 \rangle,$$
$$\sigma_P^2 = \langle P_N^2 \rangle. \tag{B2}$$

One may notice that we have used a single covariance for the complex visibility. It is straightforward to show that the same result holds if a separate real and imaginary components are used, as long as they are independent of each other. In fact, for simplicity and without loss of generality we shall treat the visibility as a real quantity in the rest of this derivation. Note that though we can assume $\langle V_N^{odd-power} \rangle = 0$, the same is not true for P_N.

The variance of P constructed with visibilities V_1 and V_2 from two baseline classes can be estimated[6]:

$$\sigma_P^2 = \langle P^2 \rangle - \langle P \rangle^2,$$
$$\propto \langle \frac{(V_{1S} + V_{1N})^2 (V_{2S} + V_{2N})^2}{\Theta^2} \rangle - \langle \frac{(V_{1S} + V_{1N})(V_{2S} + V_{2N})}{\Theta} \rangle^2,$$
$$= \frac{1}{\Theta^2} \left(V_{1S}^2 \sigma_{V2}^2 + V_{2S}^2 \sigma_{V1}^2 + \langle V_{1N}^2 V_{2N}^2 \rangle \right),$$
$$= \frac{1}{\Theta^2} \left[V_S^2 (\sigma_{V2}^2 + \sigma_{V1}^2) + \sigma_{V1}^2 \sigma_{V2}^2 \right], \tag{B3}$$

where in the second last line we have substituted in the visibility noise variance. In the final line we used Wick's theorem and the fact that the signal from the two visibilities are equal.

Recall from the discussion on multiplicities we can write

$$\sigma_V^2 = \frac{\sigma_0^2}{M}, \tag{B4}$$

where σ_0 is the single-baseline noise level. Letting $\rho_0 = V_S^2/\sigma_0^2$ be the signal to noise ratio for a single baseline, we can write

$$\sigma_P^2 \propto \frac{\sigma_0^4}{\Theta^2} \left[\rho_0 \left(\frac{1}{M_1} + \frac{1}{M_2} \right) + \frac{1}{M_1 M_2} \right],$$
$$\propto \frac{1}{\widetilde{\Theta}_{12}^2}. \tag{B5}$$

Thus we have defined a slightly modified version of the effective weight in Eq. (30):

$$\widetilde{\Theta}_{12} = \frac{\Theta_{12} \sqrt{M_1 M_2}}{\sqrt{1 + \rho_0 (M_1 + M_2)}}. \tag{B6}$$

REFERENCES

Ali, Z. S., et al. 2015, The Astrophysical Journal, 809, 61
Bowman, J. D., et al. 2013, Publications of the Astronomical Society of Australia, 30
Bull, P., Ferreira, P. G., Patel, P., & Santos, M. G. 2015, The Astrophysical Journal, 803, 21
Chen, X. 2015, International Journal of Modern Physics A, 30, 1545011
DeBoer, D. R., et al. 2016, ArXiv e-prints
Dillon, J. S., & Parsons, A. R. 2016a, ArXiv e-prints
—. 2016b, ApJ, 826, 181
Ewall-Wice, A., et al. 2016, ArXiv e-prints
Fan, X., Carilli, C. L., & Keating, B. 2006, ARA&A, 44, 415
Furlanetto, S. R., Oh, S. P., & Briggs, F. H. 2006, Physics Reports, 433, 181

Gorski, K. M., Hivon, E., Banday, A. J., Wandelt, B. D., Hansen, F. K., Reinecke, M., & Bartelmann, M. 2005, The Astrophysical Journal, 622, 759
Gorski, K. M., Wandelt, B. D., Hansen, F. K., Hivon, E., & Banday, A. J. 1999, ArXiv Astrophysics e-prints
Liu, A., & Parsons, A. R. 2016, Monthly Notices of the Royal Astronomical Society, 457, 1864
Liu, A., Parsons, A. R., & Trott, C. M. 2014a, Phys. Rev. D, 90, 023018
—. 2014b, Phys. Rev. D, 90, 023019
Liu, A., Pritchard, J. R., Allison, R., Parsons, A. R., Seljak, U. c. v., & Sherwin, B. D. 2016, Phys. Rev. D, 93, 043013
Liu, A., Zhang, Y., & Parsons, A. R. 2016, ApJ, 833, 242
Mao, Y., Tegmark, M., McQuinn, M., Zaldarriaga, M., & Zahn, O. 2008, Phys. Rev. D, 78, 023529
McQuinn, M., Hernquist, L., Zaldarriaga, M., & Dutta, S. 2007, MNRAS, 381, 75
Mesinger, A., Ferrara, A., Greig, B., Iliev, I., Mellema, G., Pritchard, J., & Santos, M. 2015, Advancing Astrophysics with the Square Kilometre Array (AASKA14), 11

[6] We assume all noise terms to be independent for simplicity, in practice the correlation of different measurements ifrom equivalent baselines are alleviated by grouping the baselines in the class and the days of observation, as in Ali et al. (2015)

Neben, A. R., et al. 2016, ArXiv e-prints
Oyama, Y., Shimizu, A., & Kohri, K. 2013, Physics Letters B,
    718, 1186
Parsons, A., Pober, J., McQuinn, M., Jacobs, D., & Aguirre, J.
    2012a, The Astrophysical Journal, 753, 81
Parsons, A. R., Pober, J. C., Aguirre, J. E., Carilli, C. L., Jacobs,
    D. C., & Moore, D. F. 2012b, The Astrophysical Journal, 756,
    165
Parsons, A. R., et al. 2014, The Astrophysical Journal, 788, 106
Paul, S., et al. 2016, ApJ, 833, 213
Planck Collaboration et al. 2016, A&A, 596, A108

Pober, J. C., et al. 2014, ApJ, 782, 66
Pritchard, J. R., & Loeb, A. 2012, Reports on Progress in
    Physics, 75, 086901
Smith, K. M., & Ferraro, S. 2016, ArXiv e-prints
Thomson, A., Moran, J., & Swenson, G. 2017, Interferometry and
    Synthesis in Radio Astronomy, 3rd edn. (Springer International
    Publishing)
Tingay, S. J., et al. 2013, Publications of the Astronomical
    Society of Australia, 30
van Haarlem, M. P. et al. 2013, A&A, 556, A2