ADDING SENSITIVITY TO 21CM INTEFEROMETRIC PROBES OF REIONIZATION BY OPTIMIZING CHOICE OF BASELINES

Yunfan Gerry Zhang [1], Adrian C. Liu [1, 2], Aaron R. Parsons[1, 2]
Draft version March 30, 2017

ABSTRACT

The observational effort to measure the primordial 21cm power spectrum with radio interferometers require high sensitivity, or signal to noise ratio. We introduce a straightforward and accurate method to extract additional sensitivity from data of redundant radio arrays. Our method efficiently finds optimal baselines to cross correlate, including baselines that are slightly different in length and orientation, and quantifies the sensitivity contributions of given baselines. Using the configurations and beams of the latest and final 128-element version of Precision Array for Probing the Epoch of Reionization (PAPER-128), as well as the planned 37, 128, 240 and 350 element versions of the Hydrogen Epoch of Reionization Array (HERA) now under development, we illustrate how our method applies to different arrays and predict the sensitivity improvements of including each baseline pairs. We show that inclusion of non-perfectly redundant baselines would increase the sensitivity of PAPER-128 and different configurations of HERA by about to 20% to 60%, with HERA-350 benefitting the most from the technique.

1. INTRODUCTION

The epoch of reionization, or cosmic dawn, 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 is important not only because it is a scientific goal of its own, but also because it can provide crucial information to 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 etc.).

Arguably the most promising observational probe of the epoch of reionization comes from measurement of the "spin-flip" transition of neutral hydrogen of characteristic wavelength 21cm (Furlanetto et al. 2006; Pritchard & Loeb 2012). While probes such as quasar spectra rely on emission and scattering of free electrons, and are thus limited to the lower redshifts towards the end of reionization, the 21cm line directly probes the abundance and distribution of neutral hydrogen, and thus would potentially shed light on all stages of reionization. Radio interferometric efforts to measure the 21cm power spectrum has been a top priority in recent years of astronomy. Current generation instruments include the 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 such as the Hydrogen Epoch of Reionization Array (HERA) (e.g. DeBoer et al. 2016; Dillon & Parsons 2016; Neben et al. 2016; Ewall-Wice et al. 2016) currently under construction and the Square Kilometer Array Low (SKA-low) (e.g. Mesinger et al. 2015) in planning.

One of the main challenges of observations of cosmic 21cm line, interferometric or global, is foreground contamination. Both sources within our galaxy, and to a lesser extent distant sources outside of our galaxy emit radio contamination (via for example synchrotron processes) up to five orders of magnitude stronger than the reionization signal. There are two common methods to deal with the foreground contamination in power spectrum measurements. In the first, sometimes called the foreground removal technique, individual sources are identified and removed in the image domain. The other technique, commonly referred to as "foreground avoidance", makes use of the fact that most common foreground contaminants have smooth spectra, and thus is constrained to a Fourier domain "foreground wedge" (Liu et al. 2014a,b). Thus, contamination can be avoided by restricting our observation to "outside the wedge", where theoretically foreground contamination is limited and noise becomes dominant. The main challenge of using the avoidance technique is thus the sensitivity, or the signal to noise ratio. This is the motivation for the design of the maximum redundancy arrays such as PAPER and HERA. Traditional imaging arrays are designed to image localized bright sources and thus favor non-redundant patterns for uv coverage. The redundant arrays, on the other hand, focus on high-sensitivity measurements of the same Fourier mode with many tightly packed antennas at equal spacing. Since baselines of the same length and orientation measure the same Fourier modes on the sky, a maximum redundancy array is able to increase the signal to noise ratio by averaging over measurements of the same baseline. Ali et al. (2015) provides the newest upper limit to the power spectrum measurements with the 64-element version of PAPER (henceforth referred to as PAPER-64). From some models we roughly expect that the sensitivity required for detection is a factor of 10 away.

This paper explores a new technique to add sensitivity to the power spectrum analysis of redundant arrays, by making use of the earth's rotation. The earth's rotation causes the baselines to pick up different modes of the sky with time. This effect is used extensively to improve uv coverage in imaging with minimum redun-

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

dancy arrays. In a maximally redundant array, we can make use of the fact that baselines that are slightly different rotate into each other at a time delay to extract power-spectrum from their respective visibilities. Here we present a method to improve the sensitivity of the maximum redundancy arrays by including the uv redundancy of different baselines at a time lag. Recently, Paul et al. (2016) proposed an algorithmic approach to extract information from near-equivalent baselines, by carefully treating the w-term in the fringe pattern to obtain weights for cross-correlation. The algorithm makes use of a time dependent coordinate system ('moving grid') and is thus computationally expensive. We here present a formalism that allows us to compute such weights by simply performing a numerical integral. This allows us to identify good baselines to cross-correlate and predict the weight of sensitivity with speed and accuracy, thereby making the power spectrum estimation from a large number of baseline pairs computationally tractable.

Baselines of the same length and orientation are traditionally called "redundant baselines", because they measure the same Fourier mode in the sky. In order to eliminate confusion and ambiguity, we shall introduce slightly different terminology. We shall call baselines that are the same length and orientation "equivalent baselines", inspired by the mathematical notion of equivalency classes. Two equivalent baselines will be redundant with each other simultaneously at all times. Non-equivalent baselines can also be partially redundant if their respected time series is shifted with respect to one another by some delay. In other words, one baseline can be "rotated into" another. We call these "near-equivalent baselines". Exactly how "near" is "near-enough" depends on the array configuration. All non-equivalent baselines contribute non-zero sensitivity to the power spectrum analysis It is our goal to a) identify the near-equivalent baselines that give good redundancy, b) to find the optimal time offset[3] for a given pair of baselines, and c) to quantify the sensitivity improvement associated with cross multiplying such a pair of near-equivalent baselines.

The rest of this paper is organized as follows. In section 2 we explain the theoretical basis for this cross multiplication. In section 3 we present numerical tests of this technique as well as the expected sensitivity improvement with this method for HERA and PAPER-128 and with section 4 we conclude.

2. METHOD

2.1. *Rough Idea*

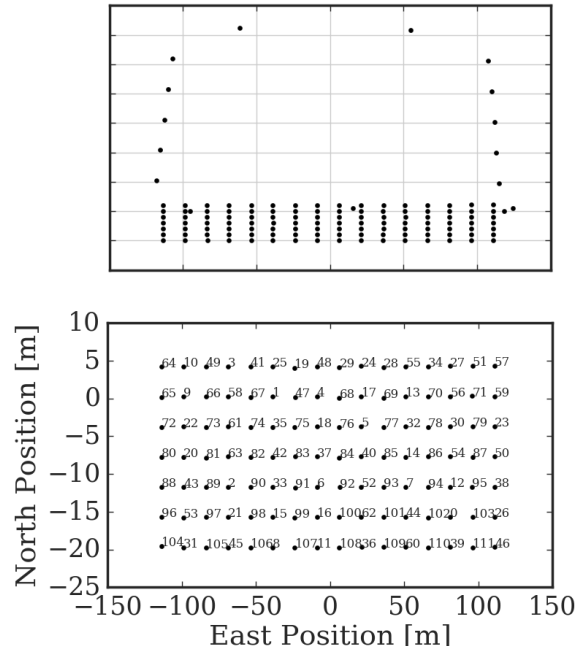[3] By optimal we mean the time offset that corresponds to maximal redundancy.

FIG. 1.— The PAPER-128 layout. Each blue dot corresponds to the location of an antenna. Top panel shows the antenna positions drawn to scale, bottom panel show the antenna labels and distances, excluding the outrigger antennas. The numbering of the antennas the bottom panel are original labels during instrument assembly and does not bear significant meaning. Baselines corresponding to the same separations are called equivalent. In the bottom panel, the two baselines (5,77) and (40,85) are an example of equivalent pair, with separation denoted by sep1,0, for the antennas are separated by 1 unit east and 0 unit north. Similarly, the baselines (5,69) and (5,85) are examples of sep1,1 and sep1,-1, respectively. Note sep1,0 and sep-1,0 for example are the same baselines and should not be counted twice.

We shall use the 128-element PAPER array to demonstrate our method, and extend our results to several HERA configurations in Sec 3.3. 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 are show in Fig. 1. We see the antenna spacing in North-South directions are comparatively close (4m), so that baselines such as 0_44 and 0_7 are very close to equivalent. In the bottom panel, the two baselines (5,77) and (40,85) are an example of equivalent pair, with separation denoted by sep1,0, for the antennas are separated by 1 unit east and 0 unit north. Similarly, the baselines (5,69) and (5,85) are examples of sep1,1 and sep1,-1, respectively. Note sep1,0 and sep-1,0 for example are the same baselines and should not be counted twice. Antennas in purely north-south baselines are close enough to induce cross-talk, and hence are not suitable for use. The original PAPER-64 analysis (Ali et al. 2015) used three classes of baselines, here equivalent to sep2,0, sep2,1 and sep2,-1 Ali et al. (2015) of PAPER-128. There each of these classes of baselines are cross multiplied to itself. We shall see that in addition these baseline classes can be cross multiplied with a time offset.

Given a point source on the sky, each baseline maps the source to a point in uv plane. As the earth rotates with respect to the source, the point traces out tracks in

the uv plane. We show in Fig. 2 uv tracks of PAPER-128 over 4.8 hours, at 0.15GHz, for a source that passes through zenith.

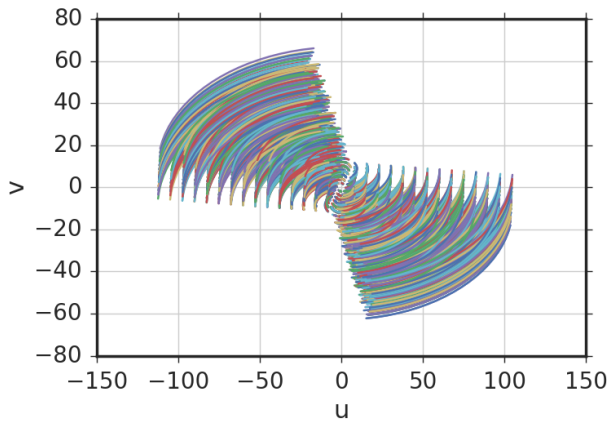Fig. 3.— Sample beam response of HERA (left) and PAPER (right) antennas, both for frequency of $\nu = 150$MHz and Stokes I polarization. Notice A in this paper is a "baseline's beam", equivalent to squares of the antenna voltage beams shown here. The circles centered around zenith (center of beam) here are spaced 10 degrees apart. The finite size of the beams limits the accuracy of the uv track-crossing as a method of redundancy search.

Fig. 2.— Tracks of PAPER-128 (grid only, excluding outriggers) for requency $\nu = 0.15$GHz and a hypothetical source that passes through zenith. These tracks are traced out over 0.2 sidereal days, or roughly 4.8 hours. Color represents different baselines. As the earth rotates, tracks are traced out counterclockwise.

Roughly speaking, we can identify redundancy of near-equivalent baselines as crossings of the uv tracks. As we see in Fig. 2, there are many such crossings. Crossing tracks, however does not imply perfect redundancy. The reason lies in the finite size of the beams. We show sample beams of HERA and PAPER antennas in Fig. 3 for reference. The inaccuracy is expected to be even more significant for the smaller HERA beam, which has a larger corresponding uv point spread function.

Thus although it is a valid quick method to identify some baselines to cross-multiply, track-crossing is not accurate enough for time offset determination, nor can give estimate of the degree of redundancy. To determine the time-offset corresponding to maximal redundancy, we must develop a more general formalism that accounts for the point spread function of the finite beams, and estimates the degree of redundancy for general combination of baselines at a general time-offset, as we do in the next section.
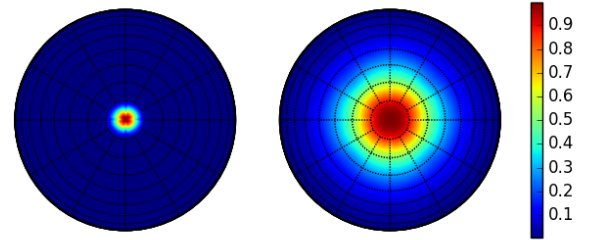
2.2. Formalism

Below we shall derive theoretical expectations of cross multiplications of two near-equivalent baselines. More precisely we shall relate the product of visibilities of two different baselines to the power spectrum.

We take the visibility as commonly defined in the literature (e.g. Parsons et al. 2012a):

$$
V_\nu(\boldsymbol{b}) = \int d\Omega A_\nu(\hat{\boldsymbol{s}})\phi(\nu)I_\nu(\hat{\boldsymbol{s}}) \exp\left[-2\pi i\frac{\nu}{c}\boldsymbol{b}\cdot\hat{\boldsymbol{s}}\right],
$$
$$
\approx \frac{2k_B}{\lambda^2} \int d\Omega A_\nu(\hat{\boldsymbol{s}})\phi(\nu)T(\hat{\boldsymbol{s}}) \exp\left[-2\pi i\frac{\nu}{c}\boldsymbol{b}\cdot\hat{\boldsymbol{s}}\right],
$$
(1)

Here λ is a mean wavelength, $\boldsymbol{b}$ is the baseline length, $\hat{\boldsymbol{s}}$ and Ω are a direction in the sky and its corresponding solid angle. A_ν is the (frequency dependent) beam, and I is the specific intensity, which has been related to T, the brightness temperature in the Rayleigh-Jeans limit. k_B is Boltzmann's constant. Note that A here is the baseline beam, i.e. product of two antenna voltage beams, examples of which are shown in Fig. 3. $\phi(\nu)$ is the frequency bandpass profile. In practice, power spectrum measurements are typically taken from a few ten MHz centered around the corresponding redshift of interest (e.g. 150 MHz for z=9.5). We take a moment to note that in Eq. 1, and our subsequent discussion, no flat-sky approximation has been made. The angular integral is performed over the dome $d\Omega$.

We define the delay-transformed visibility (Parsons

et al. 2012b):

$$V(\boldsymbol{b}, \tau) = \int d\nu V_\nu(\boldsymbol{b}) \phi(\nu) \exp\left[-2\pi i \nu \tau\right],$$
$$= \frac{2k_B}{\lambda^2} \int d\Omega d\nu A(\hat{\boldsymbol{s}}, \nu) \phi(\nu) T(\hat{\boldsymbol{s}}, \nu) \exp\left[-2\pi i \nu \left(\frac{\boldsymbol{b} \cdot \hat{\boldsymbol{s}}}{c} + \tau\right)\right] \tag{2}$$

Eq. (2) expresses the delay-transformed visibility as an integral over observation coordinates $\hat{\boldsymbol{s}}$ and ν. Ultimately, we would like to relate the data, collected with coordinates $\hat{\boldsymbol{s}}$ and ν, to the power spectrum, written with cosmological coordinates $\boldsymbol{r}$ and $\boldsymbol{k}$. We start by noticing that

$$r = \frac{c}{H_0} \int_0^z \frac{dz'}{E(z')},$$
$$\approx \frac{c}{H_0} \int_0^{z_0} \frac{dz'}{E(z')} - \frac{c(1+z)^2}{\nu_{21} H_0 E(z)} (\nu - \nu_0), \tag{3}$$
$$\equiv D_c - Y\Delta\nu,$$

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

$$E(z) = \sqrt{\Omega_m(1+z)^3 + \Omega_\Lambda}. \tag{4}$$

Inverting for ν:

$$\nu = \frac{D_c - r}{Y} + \nu_{21}. \tag{5}$$

We can thus rewrite the delayed transformed visibility as

$$V(\boldsymbol{b}, \tau) = \frac{2k_B}{\lambda^2} \int_H \frac{d^3r}{X^2Y} A(\boldsymbol{r})\phi(\boldsymbol{r})T(\boldsymbol{r}) \exp\left[-2\pi i \left(\frac{\boldsymbol{b}}{c} \cdot \hat{\boldsymbol{r}} + \tau\right)\nu_r\right]. \tag{6}$$

Here $(r_x, r_y, r_z) = (X\hat{s}_x, X\hat{s}_y, Y\nu)$, and so $d\nu = -dr/Y$ and $d^3r = -X^2Y d\Omega d\nu$. $(Xk_x, Xk_y, Yk_z) = \frac{2\pi}{c}(b_x, b_y, \tau)$ relate the cosmological coordinates $\boldsymbol{r}$ and $\boldsymbol{k}$ to the measured coordinates. We have written ν_r to remind us that ν and r are related by Eq. (5). The beam reception pattern A is dimensionless, normalized to 1 at its peak (zenith), and we assume it to be the same for all baselines.

With a time offset, the beam pattern has moved relative to the sky. Here we choose to fix the sky, and denote the rotated coordinates of the beam pattern with the 3 dimensional rotation operator Γ. With implicit bounds of integrals from $-\infty$ to ∞, we have:

$$\langle V^*(\boldsymbol{b}, \tau)V(\boldsymbol{b}', \tau')\rangle$$
$$= \left(\frac{2k_B}{X^2Y\lambda^2}\right)^2 \int d^3r d^3r' \left(\langle T^*(\boldsymbol{r})T(\boldsymbol{r}')\rangle\right) A^*(\boldsymbol{r})A(\Gamma r')\Phi_{b,\tau}(\boldsymbol{r}, \Gamma r'),$$
$$= \left(\frac{2k_B}{X^2Y\lambda^2}\right)^2 \int d^3r d^3r' \left(\int \frac{d^3\kappa}{(2\pi)^3} \frac{d^3\kappa'}{(2\pi)^3} \langle T^*(\boldsymbol{\kappa})T(\boldsymbol{\kappa}')\rangle e^{-i(\boldsymbol{\kappa}\cdot\boldsymbol{r} - \boldsymbol{\kappa}'\cdot\boldsymbol{r}')}\right) A^*(\boldsymbol{r})A(\Gamma r')\Phi_{b,\tau}(\boldsymbol{r}, \Gamma r'),$$
$$= \left(\frac{2k_B}{X^2Y\lambda^2}\right)^2 \int d^3r d^3r' \left(\int \frac{d^3\kappa}{(2\pi)^3} P(\kappa) e^{-i\boldsymbol{\kappa}\cdot(\boldsymbol{r}-\boldsymbol{r}')}\right) A^*(\boldsymbol{r})A(\Gamma r')\Phi_{b,\tau}(\boldsymbol{r}, \Gamma r'),$$
$$\approx \left(\frac{2k_B}{X^2Y\lambda^2}\right)^2 P(k_{b,\tau}) \int d^3r d^3r' \delta_D^{(3)}(\boldsymbol{r} - \boldsymbol{r}') A^*(\boldsymbol{r})A(\Gamma r')\Phi_{b,\tau}(\boldsymbol{r}, \Gamma r'), \tag{7}$$
$$= \left(\frac{2k_B}{X^2Y\lambda^2}\right)^2 P(k_{b,\tau}) \int d^3r |A^*(\boldsymbol{r})A(\Gamma r)||\phi(\nu_r)|^2 \exp\left[-i2\pi\nu_r \left(\hat{\boldsymbol{r}} \cdot \frac{\boldsymbol{b}}{c} - \hat{\Gamma}r \cdot \frac{\boldsymbol{b}'}{c}\right)\right],$$
$$= \left(\frac{2k_B}{\lambda^2}\right)^2 P(k_{b,\tau}) \int \frac{d\Omega d\nu}{X^2Y} |A^*(\hat{\boldsymbol{s}}, \nu)A(\hat{\Gamma}s, \nu)||\phi(\nu)|^2 \exp\left[-i2\pi\nu \left(\hat{\boldsymbol{s}} \cdot \frac{\boldsymbol{b}}{c} - \Gamma\hat{\boldsymbol{s}} \cdot \frac{\boldsymbol{b}'}{c}\right)\right],$$

where in transition from cosmological coordinates back to observing coordinates we have written $\hat{\boldsymbol{r}} \equiv \hat{\boldsymbol{s}}$, and

$$\Phi_{b,\tau}(\boldsymbol{r}, \Gamma r') = |\phi^*(\nu_r)\phi(\nu_{r'})| \exp\left[-i\frac{2\pi}{c}\left(\boldsymbol{b} \cdot \nu_r \hat{\boldsymbol{r}} - \boldsymbol{b}' \cdot \nu_{r'}\Gamma\hat{r}'\right)\right] \exp\left[-i2\pi\tau(\nu_r - \nu_{r'})\right]. \tag{8}$$

The second to third line of Eq.(7) follows from assumption of Gaussian random sky, and the third to fourth line follows from the assumption that the 3D power spectrum varies negligibly over the k-space of interest so that $\hat{P}_{21}(k + k_2) \approx \hat{P}_{21}(k)$. Since Γ is a sky rotation, it does not affect ν, hence we have taken ν_r outside the parenthesis. Notice that the phase factor $\exp\left[-i2\pi\tau(\nu - \nu')\right]$ drops out in the end. This means that correlation is peaked at the same time-offset for all delay channels.

Finally, since the beam pattern and bandwidth 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:

$$\langle V^*(\boldsymbol{b}, \tau) V(\boldsymbol{b'}, \tau') \rangle$$

$$= \left(\frac{2k_B}{\lambda^2} \right)^2 P(k_{b,\tau}) \int \frac{d\Omega d\nu}{X^2 Y} |A^*(\hat{\boldsymbol{s}}, \nu) A(\Gamma\hat{\boldsymbol{s}}, \nu)||\phi(\nu)|^2$$

$$\exp \left[-i2\pi\nu \left(\hat{\boldsymbol{s}} \cdot \frac{\boldsymbol{b}}{c} - \Gamma\hat{\boldsymbol{s}} \cdot \frac{\boldsymbol{b'}}{c} \right) \right]. \tag{9}$$

In other words the power spectrum estimate from visibilities of a baseline pair is given by

$$P(k_{b,\tau}) = \left(\frac{\lambda^2}{2k_B} \right)^2 \frac{\langle V^*(\boldsymbol{b}, \tau) V(\boldsymbol{b'}, \tau') \rangle}{\Theta}, \tag{10}$$

where the weight

$$\Theta = \int \frac{d\Omega d\nu}{X^2 Y} |A^*(\hat{\boldsymbol{s}}, \nu) A(\Gamma\hat{\boldsymbol{s}}, \nu)||\phi(\nu)|^2 e^{-2\pi i \nu \left(\hat{\boldsymbol{s}} \cdot \frac{\boldsymbol{b}}{c} - \Gamma\hat{\boldsymbol{s}} \cdot \frac{\boldsymbol{b'}}{c} \right)}. \tag{11}$$

So roughly speaking the cross multiplications of visibilities at a time delay in uv-space is proportional to the power spectrum times the Fourier transform of the cross multiplied beam pattern. We can then in principle combine information from different baseline pairs if we correct for the phase and normalization. As a check, when applied to equivalent baselines, $\boldsymbol{b} = \boldsymbol{b'}$, $\hat{\boldsymbol{s}} = \Gamma\hat{\boldsymbol{s}}$, and Eq.(9) reduces to Eq.(B9) of Parsons et al. (2014).

With Eq. 10 and Eq. 11 we can for any given baseline pair and time delay, estimate the degree of redundancy, here represented by Θ, thereby achieving all our goals stated in the introduction, i.e. to identify candidate baseline pairs with good 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.(11) for various time offsets, without having to actually cross-multiply visibilities at different offsets. This makes the task computationally tractable.

2.3. Rephasing

Before we discuss tests, there is one subtlety when applying Eq. 9. The optimal time offset is given by the Γ that maximizes the absolute value of the weight Θ. However, Θ is in general complex. First integrating over the spatial dimension Ω, we see the phase term at peak time-offset is inevitably frequency dependent. This frequency dependence of the phase would lead to destructive interference when we integrate over frequency.

The physical origin of this frequency-dependent phase lies in the two visibilities having different phase centers. By default the two visibilities are both phased to zenith at the same time. When they are cross-multiplied with a time lag, they must be rephased to the account for the movement of the zenith. Thus in practice we must rephase the visibilities before delay transform.

In Fig. 4 we compare the cross-multiplied visibilities of both equivalent and near-equivalent baseline pairs for two channels: 0.16 GHz and 0.17 GHz. The top two panels have zero rephasing, and the bottom two are rephased to a time offset of 0.055 sidereal days, the optimal time offset for the given baseline pair. The first and third panels show the equivalent baseline pairs sep1,0 against itself, and second and fourth panels show sep1,0 against

sep1,1. We see that in the un-rephased case in the second panel, although the magnitude of correlations match up for the two frequencies, the phases do not. This means summing over frequency leads to destructive interference and signal loss. The wider the frequency profile, the more destructive the interference would be. In the rephased case, the phases of the near-equivalent case match up and can be added without compromising sensitivity. We should thus rephase the data separately for each set of baseline pairs.
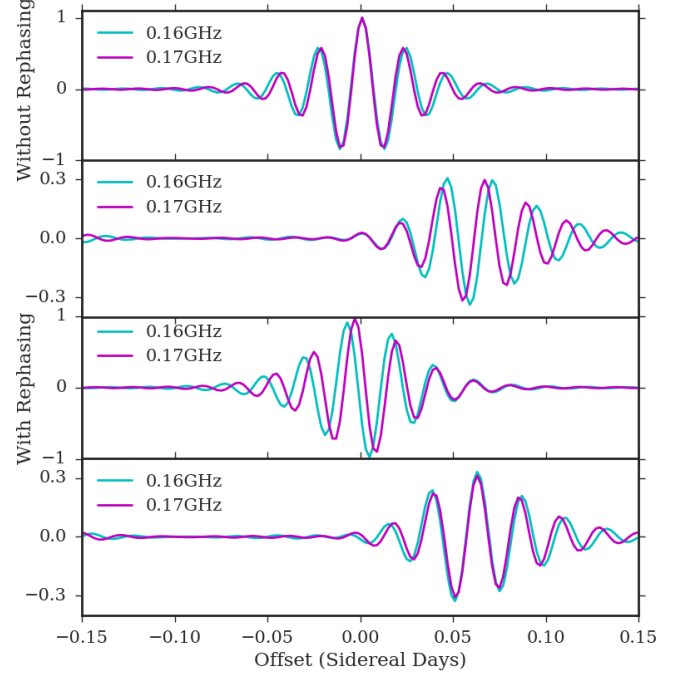

FIG. 4.— Comparisons of the peak phases of two different frequencies. First and third panels shows equivalent baselines, second and fourth show a pair of near-equivalents. For visual simplicity only the real parts are shown.

3. ANALYSIS

3.1. Numerical Test

First we present some numerical tests of our formalism to verify the validity of using Θ to estimate the degree of redundancy as a function of time lag. 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. For computational simplicity, we shall use a single frequency channel of 150MHz for the comparison. We may use this simplification without compromising generality because Eq. (9) is valid for a range of frequency channels if and only if it's valid for each single channels, provided that different channels are combined with rephasing as described in Sec. 2.3.

For the simulation, we take N=2500 random realizations of the sky on a healpix map (Grski et al. 2005; Gorski et al. 1999) [4], and pick two baselines sep1,0 and

---

[4] We use functionalities in the python package AIPY for healpix

sep1,1. For each realization, each pixel is given a Gaussian random value of brightness temperature. We then rotate the baseline positions with the appropriate rotation matrix, multiplying the sky by the primary beam to get the visibilities, for each baseline[5]. The resulting 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. The peak of the curve then corresponds to maximum redundancy. We do this for both the equivalent (sep1,0 against sep1,0) and near-equivalent (sep1,0 against sep1,1) case. The accuracy of this result is limited by (simulated) cosmic variance and finite spacial resolution, and hence can be beat down by averaging over a large number of universes. A numerical estimate of the error of the peak height with $N = 1$ is $\lesssim 20\%$, and thus with $N = 2500$, we achieve an error of peak height $< 0.5\%$.
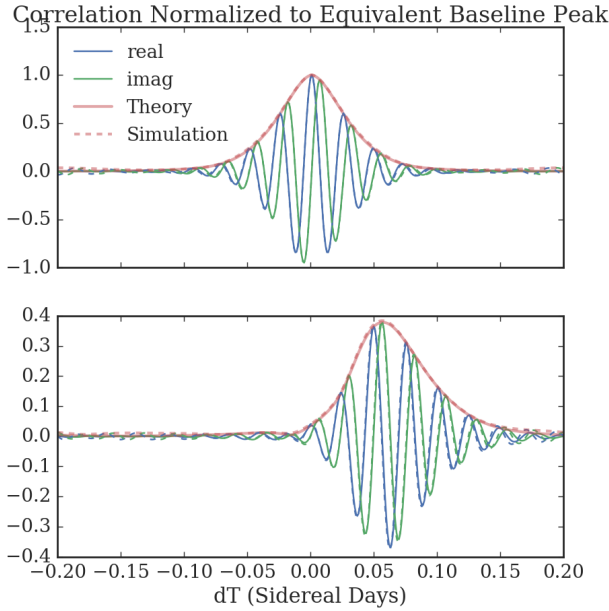

FIG. 5.— Numerical comparisons of the visibility correlation peaks to the Ω factor in Eq.(9). We generated 2400 instances of Gaussian random sky on healpix maps, computed visibilities and cross correlated them to find the correlation. Top panel shows the equivalent baseline pairs sep1,0 against sep1,0, bottom panel shows set1,0 against sep1,1. Evaluations of weight Θ are shown with solid curves, and the visibility correlations of simulated random sky is given with dashed lines. In both cases blue denotes real part, green the imaginary part, and red the magnitude. We see that in both cases the theory and simulation line up in both amplitude and phase.

The comparison is show in Fig. 5. Sensitivity contribution is inferred from height of the peak of the cross-multiplied visibilities as a function of time. We first compare the simulation to the analytical weight Θ for a pair of equivalent baselines (sep1,0, sep1,0), then nor-

malize the peaks of the near-equivalent baselines (sep1,0, sep1,1) to that of the equivalent ones. In Fig. 5 we show on the top comparison of the convolution of an equivalent baseline of PAPER-128, normalized to 1 at the peak, whose location is at zero time offset as expected. On the bottom we show comparison of a near-equivalent baseline pair, of classes sep1,0:1,1 and sep1,1:1,1. The peaks are normalized with the same factor as the peak height on the top, i.e. the two plots have thus the same scale on the vertical axis. The height of the plot thus quantifies the added sensitivity. We see at a peak of around 0.055 sidereal days, or 1.32 hours, the two baselines are maximally redundant.

At the optimal time separation, the integral in Eq. (9) is maximized. Thus as another check we expect the two beams to have to same fringe pattern (frequency and phase). Due to the time delay, however, the beam center would be slightly shifted with respect to each other. This we show in Fig. 6. The left and middle panels show the beam fringe patterns for baselines sep2,0 and sep2,1, delayed by 0.0325 sidereal days, and the right panel shows their cross product. The fringe pattern indeed cancels out as we expect.

3.2. *Sensitivity*

Having verified Eq. (9), we can thus predict the sensitivity contributions of a particular baseline pair simply by computing the integral Eq. İntuitively we expect the sensitivity to depend on both the uv coverage of the baseline and the patch of sky inside the beam. A larger beam like that of PAPER would tolerate larger time-offsets because more sky area can coincide in the two beams [6] Having computed all of the baseline pairs, we find that baseline pairs that are mirror images of each other give the same amount of redundancies (peak height), with the opposite time offset, as expected from symmetry. For example, sep1,0:1,1 is mirror image of sep1,0:1,-1 and these two baseline pairs give the same sensibility contribution. Thus we shall only show a subset of representative baseline pairs to illustrate the contributions from different classes of baseline pairs. For a more complete result see Fig. 10.

In the top panel of Fig. 7 we show the peak heights and locations for a variety of baseline combinations. We see that baseline pairs that have crossings at a smaller time delay tend to have higher correlations. In other words, correlation peaks that are closer to zero time lag are higher. This is expected since a) the longer the time delay, the more the antennas have moved with respect to the sky and hence the less overlaps in patch of sky surveyed, b) smaller optimal time-offset corresponds to smaller differences in orientation, and hence in length of baselines in PAPER.

To determine that actual relative contribution to sensitivity of these baseline pairs, we have to take into account of the multiplicities of these baselines. By these we mean how many physical antenna pairs have the same length and orientation. Looking at Fig. 7 we see for example sep1,0 will have higher multiplicity than sep2,0, or sep1,1. The latest release of PAPER-64 data uses the

---

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.

[6] Larger beams also imply smaller spread in uv space. This could lead to either larger or smaller redundancy, which depends on the overlap of two such point spread functions.
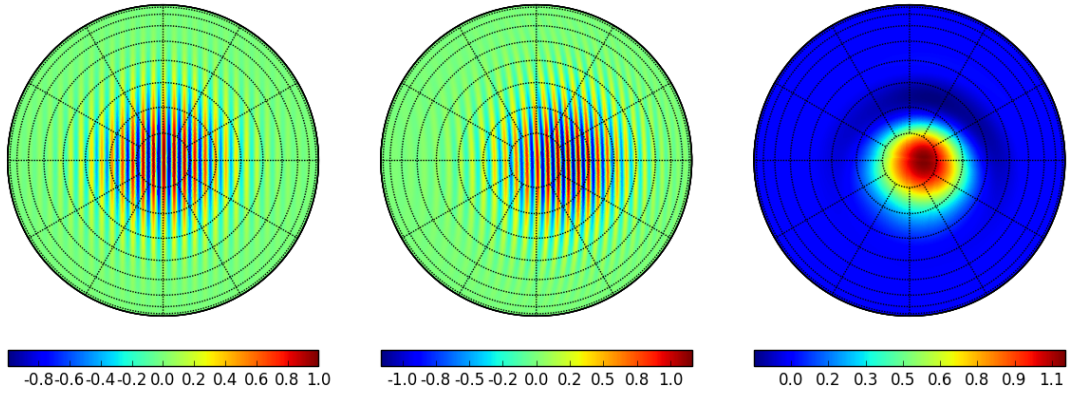
FIG. 6.— Beam fringe pattern of sep2,0(left), sep2,1 at a time delay (middle), and their conjugate product (right). Frequency of $\nu = 0.15$GHz is chosen and only the real components are shown. The colorbar values are normalized such that the peak of the original beam is 1.

128-equivalent baselines sep2,1, sep2,0 and sep2,-1 (Ali et al. 2015), and achieved a 2σ upper limit of $(22.4\text{mK})^2$. There, the three sets of equivalent baselines are only cross multiplied by itself. Assuming that each baseline delivers the same quality of data (meaning they have the same height of correlation peaks, which is in our normalization equal to unity), the relative contribution to sensitivity can be estimated.

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

$$\widetilde{\Theta}_{bb'} \propto \Theta_{bb'} \times \sqrt{MM'}. \qquad (12)$$

Shown in the bottom panel of Fig. 7 is 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, or in other words are mirror images of each other , by summing over them to get an extra factor of $\sqrt{2}$. For example, instead of plotting 1,1:1,1 and 1,-1:1,-1 separately, we plot 1,1:1,1 with twice the multiplicity. Similarly, baseline pairs such as 1,0:1,1 also get the factor of 2 because they are identical to 1,0:1,-1. Baseline pairs such as 1,0:1,0, or 1,1:1,-1 do not get the factor of $\sqrt{2}$ because their mirror images are themselves.

Having defined the modified weight $\widetilde{\Theta}$, we can estimate the power spectrum by inverse covariance weighting:

$$\begin{aligned} P(k_\tau) &= \frac{\sum_{bb'} P(k_{b,\tau})/\sigma_P^2(bb')}{\sum_{bb'} \sigma_P^2(bb')}, \\ &= \frac{\sum_{bb'} P(k_{b,\tau})\widetilde{\Theta}_{bb'}^2}{\sum_{bb'} \widetilde{\Theta}_{bb'}^2}, \end{aligned} \qquad (13)$$

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

$$\rho \propto \frac{1}{\sigma_P^2} \propto \frac{1}{\sigma_N^4} \sum_{bb'}^{N} \widetilde{\Theta}_{bb'}^2, \qquad (14)$$

where σ_N is a characteristic single-baseline noise level.

The scaling in Eq. 12 was rough for simplicity of motivation. As we derive in Appendix A, this weight should be corrected by a factor proportional to the signal to noise ratio of a single-baseline visibility $\rho_0 = \sigma_S^2/\sigma_N^2$:

$$\widetilde{\Theta}_{bb'} = \frac{\Theta_{bb'}\sqrt{M_b M_{b'}}}{\sqrt{1 + \rho_0 (M_b + M_{b'})}}. \qquad (15)$$

To move forward we must give an estimate of ρ_0 in Eq.15. Assuming a reionization signal of $\Delta_{21cm}^2 \sim$

$30mK^2$, observation at 150MHz ($z = 8.5$), 120 days of observation with PAPER antennas, we have roughly (See Eq.(20) in Parsons et al. (2012a))

$$\rho_0 \sim 0.001 \left[\frac{L}{40m}\right] \left[\frac{0.1hMpc^{-1}}{k}\right]^3, \qquad (16)$$

where L is the baseline length. We only need a single characteristic L even in the near-equivalent case because only baselines of nearly equal length would have high redundancy.As expected, baseline-pairs that have smaller $\widetilde{\Theta}$ contribute less to the sensitivity.

3.3. 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. 3.3. The hexagonal design is the densest pattern of antenna-packing. The larger arrays are designed with a "gap" dividing the antennas into three different regions. The gaps are designed so as to improve uv coverage and ease calibration without compromising sensitivity, but also produces many more near-equivalent baselines than the versions without the gap. The motivations behind the designs are explained in Dillon & Parsons (2016). The lack of short baselines that are visually close to each other, as well as the smaller beam (Fig. 3) means that we expect to see only longer near-equivalent baselines. The lower multiplicities per class of baselines is made up with the larger number of classes of baseline-pairs, especially given the gap in the larger versions.

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. 8 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 that in all cases using the near-equivalent baselines lead to more and more significant improvements with lower $\widetilde{\Theta}_{min}$, or in other words when worse baseline pairs are used. The small HERA37, with no gap (like in HERA350) or short near-equivalent baselines (like in PAPER 128), will not benefit much from the near-equivalent baselines. The maximum benefits for other cases are expected to be around 20% to 60%. PAPER128 is designed with highly redundant near-equivalent baselines, and thus these baselines start contributing at higher $\widetilde{\Theta}_{min}$, but the gapped HERA configurations will benefit even more from near-equivalent baselines at low $\widetilde{\Theta}_{min}$ due to there being more classes of such pairs. Note that here we normalized ρ such that the contribution of the top equivalent baseline pair, such as (sep0,1:0,1 in the PAPER128 case) 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".
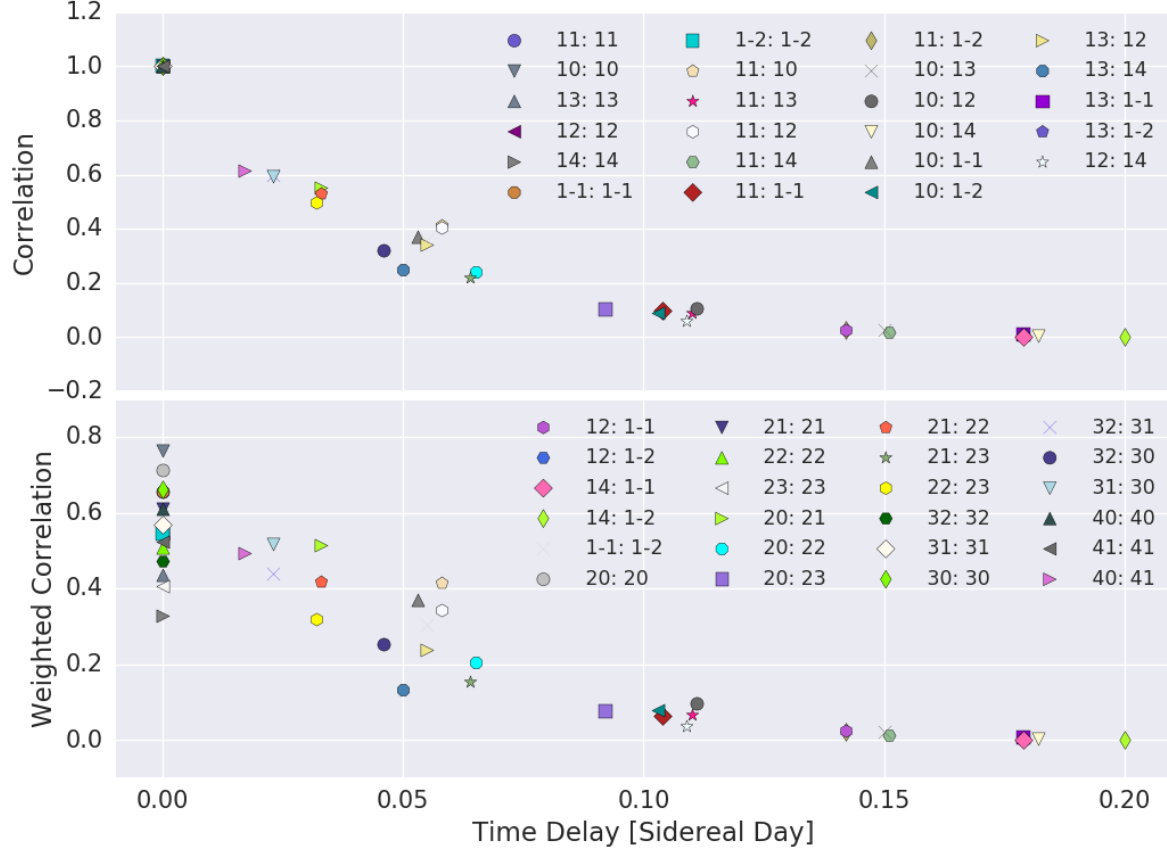
FIG. 7.— Relative sensitivity contributions of selected baseline combinations in PAPER-128. In the legend m,n:p,q denotes cross multiplying PAPER-128 baselines of east-west, north-south separations (m,n) and (p,q) respectively. The top panel shows the peak height (degree of correlation) of each baseline combination, while the bottom panel multiplies the heights by the corresponding multiplicities as in Eq. (12), and in some cases an extra factor of $\sqrt{2}$ (explained in text). In weighting by the multiplicity (bottom), we have chosen to fix the sensitivity contribution of 1,0:1,0 to unity. The same symbols and colors correspond in the top and bottom panel, with half of the legend shown in each.
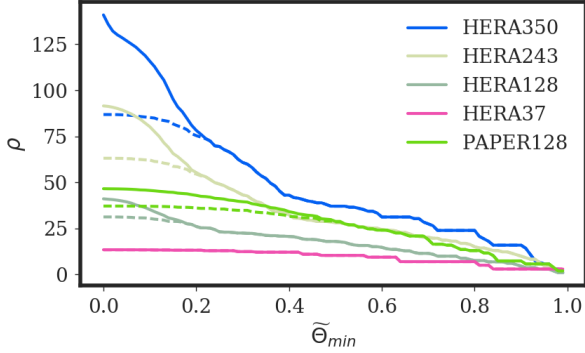
FIG. 8.— Sensitivity of redundant arrays as a function of the minimum effective weight. Dashed lines represent when only equivalent baseline pairs are used, while solid lines indicate use of both the equivalent and near-equivalent baselines are used. The y-axis is normalized independently for each array 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.

In Fig. 10 we present a comparison of 4 different properties for baselines that contribute well to the sensitivity ($\widetilde{\Theta} > 0.05$, where again the weight for the top class of equivalent pair is normalized to 1). The dT vs. Θ and dT vs. $\widetilde{\Theta}$ plots are familiar from Fig. 7. We only show 3 of the mentioned arrays for visual clarity. The other two results are similar barring intuitive differences. We make a few observations:

- Mapping of equivalent baseline pairs: the equivalent baselines are located in key points in the plots. They are the vertical bars in the lower left four plots and the highest peak in the first three histograms, the upper "diagonal edge" in the $\widetilde{\Theta}$ vs. dT and the dot of $\Theta = 1$.

- Longer baselines roughly peak at lower time-offset, and lower multiplicity. This is intuitive for PAPER since for the same north-south separation the rotation angle is smaller with longer east-west separation.

- Interestingly for HERA the near-equivalent baselines form multiple "U" shapes in dT, Θ plane, somewhat counter-intuitive especially given out last point. This is characteristic of the hexagonal layout. Baselines of equal length are more redundant at the right time delay than ones that are slightly different in length, though the latter has a shorter optimal time-delay.

- The top near-equivalent pairs in HERA the same Θ 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. 8) because there are a lot more such baseline pairs for HERA.

4. 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 an intuitive analysis of cross-multiplying baselines that are close in length and orientation to each other. Given an antenna array configuration, our method quickly identifies the best baseline pairs to cross-multiply and predict the expected sensitivity contribution. With the predicted result one can improve existing power-spectrum pipelines through 3 simple steps. 1). Rephase the visibilities prior to delay transforming by the zenith displacement (or more accurately by the phase predicted by our numerical analysis). 2). Shift the visibilities in time. 3) Cross multiply the visibilities of the two baselines to form the power-spectrum. 4) Finally combine the different baseline pairs by appropriate inverse-variance weighting that takes into account the predicted sensitivity contributions of each case. We showed that such techniques could lead to 20% to 60% increase in sensitivity for PAPER and HERA.
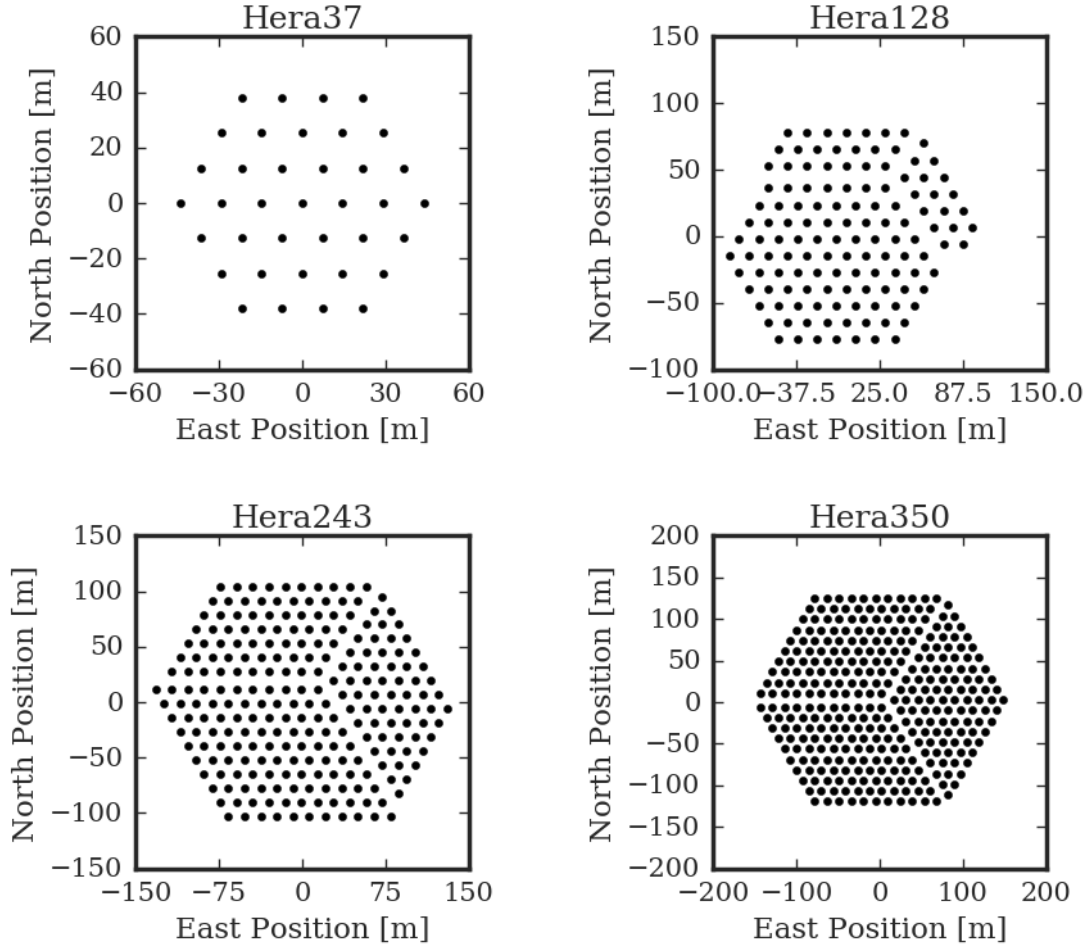
FIG. 9.— Planned Hydrogen Epoch of Reionization Array antenna configurations. HERA 37 is expected to complete and start collecting data in summer of 2017, and the other three are planned configurations in the next phases. For HERA350, only the 320 elements in the core are shown.
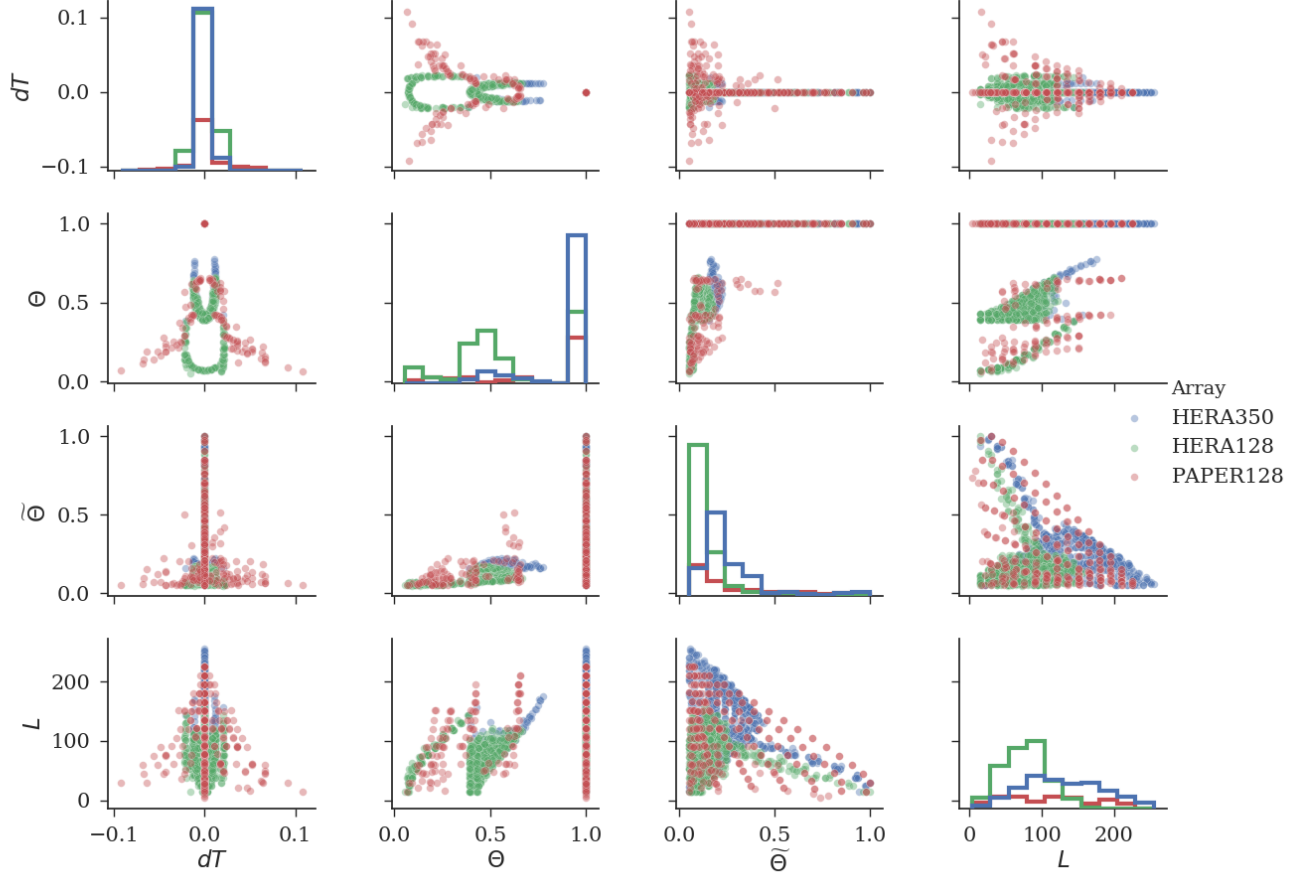
Fig. 10.— Pairplots of the top contributing baseline pairs in three arrays. Plotted properties are optimal time delay dT, peak height Θ, effective weight $\widetilde{\Theta}$ and baseline length L. Only those points with $\widetilde{\Theta} > 0.05$ are shown (the weight for the top class of equivalent 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.

APPENDIX

A.
DERIVATION OF NOISE COVARIANCE

In this appendix we give a brief derivation of the effective weight quoted in 3.2. We combine the different power spectrum measurements by inverse variance weighting [7]. We shall separate the visibility and power spectrum into signal and noise contributions:

$$V = V_S + V_N,$$
$$P = P_S + P_N. \tag{A1}$$

We shall denote the noise variance of power spectrum and visibility

$$\sigma_V^2 = \langle |V_N|^2 \rangle,$$
$$\sigma_P^2 = \langle P_N^2 \rangle. \tag{A2}$$

One may notice that we have used a single covariance for the complex quantity visibility. It's simple to show that the same result holds if we use a separate real and imaginary components, as long as they are independent of each other.

---

[7] In practice techniques such as bootstrapping is often used, see for example (Ali et al. 2015)

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.

Then the variance of P constructed with visibilities V_1 and V_2 from two baseline classes can be estimated[8]:

$$
\begin{aligned}
\sigma_P^2 &= \langle P^2 \rangle - \langle P \rangle^2, \\
&\propto \langle \frac{(V_{1S} + V_{1N})^2(V_{2S} + V_{2N})^2}{\Theta^2} \rangle - \langle \frac{(V_{1S} + V_{1N})(V_{2S} + V_{2N})}{\Theta} \rangle^2, \\
&= \frac{1}{\Theta^2} \left(V_{1S}^2 \sigma_{V2}^2 + V_{2S}^2 \sigma_{V1}^2 + \langle V_{1N}^2 V_{2N}^2 \rangle \right), \\
&= \frac{1}{\Theta^2} \left[V_S^2 (\sigma_{V2}^2 + \sigma_{V1}^2) + \sigma_{V1}^2 \sigma_{V2}^2 \right],
\end{aligned}
\tag{A3}
$$

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

Recall from the discussion on multiplicities we can write

$$
\sigma_V^2 = \frac{\sigma_0^2}{M},
\tag{A4}
$$

where σ_0 is some 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

$$
\begin{aligned}
\sigma_P^2 &\propto \frac{\sigma_0^4}{\Theta^2} \left[\rho_0 \left(\frac{1}{M_1} + \frac{1}{M_2} \right) + \frac{1}{M_1 M_2} \right], \\
&\propto \frac{1}{\widetilde{\Theta}_{12}^2},
\end{aligned}
\tag{A5}
$$

where we have defined a slightly modified version of the effective weight (compare with Eq. 12):

$$
\widetilde{\Theta}_{12} = \frac{\Theta_{12}\sqrt{M_1 M_2}}{\sqrt{1 + \rho_0 (M_1 + M_2)}}.
\tag{A6}
$$

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. 2016, ArXiv e-prints
Ewall-Wice, A., et al. 2016, ArXiv e-prints
Furlanetto, S. R., Oh, S. P., & Briggs, F. H. 2006, Physics Reports, 433, 181
Gorski, K. M., Wandelt, B. D., Hansen, F. K., Hivon, E., & Banday, A. J. 1999, ArXiv Astrophysics e-prints
Grski, K. M., Hivon, E., Banday, A. J., Wandelt, B. D., Hansen, F. K., Reinecke, M., & Bartelmann, M. 2005, The Astrophysical Journal, 622, 759
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
Liu, A., Parsons, A. R., & Trott, C. M. 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

Mao, Y., Tegmark, M., McQuinn, M., Zaldarriaga, M., & Zahn, O. 2008, Phys. Rev. D, 78, 023529
Mesinger, A., Ferrara, A., Greig, B., Iliev, I., Mellema, G., Pritchard, J., & Santos, M. 2015, Advancing Astrophysics with the Square Kilometre Array (AASKA14), 11
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., et al. 2014, The Astrophysical Journal, 788, 106
Parsons, A. R., Pober, J. C., Aguirre, J. E., Carilli, C. L., Jacobs, D. C., & Moore, D. F. 2012b, The Astrophysical Journal, 756, 165
Paul, S., et al. 2016, ApJ, 833, 213
Pritchard, J. R., & Loeb, A. 2012, Reports on Progress in Physics, 75, 086901
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

---

[8] 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)