Polarized redundant-baseline calibration for 21 cm cosmology without adding spectral structure

Joshua S. Dillon,[1]⋆† Saul A. Kohn,[2] Aaron R. Parsons,[1] James E. Aguirre,[2] Zaki S. Ali,[1] Gianni Bernardi,[3,4,5] Nicholas S. Kern,[1] Wenyang Li,[6] Adrian Liu,[1]‡ Chuneeta D. Nunhokee[1,4] and Jonathan C. Pober[6]

[1]*Department of Astronomy, University of California, Berkeley, Berkeley, CA, 94720, USA*
[2]*Department of Physics and Astronomy, University of Pennsylvania, Philadelphia, PA, 19104, USA*
[3]*INAF-Istituto di Radioastronomia, Bologna, I-40129, Italy*
[4]*Department of Physics and Electronics, Rhodes University, Grahamstown, 6140, South Africa*
[5]*Square Kilometre Array South Africa, Cape Town, 7405, South Africa*
[6]*Department of Physics, Brown University, Providence, RI, 02912, USA*

ABSTRACT

21 cm cosmology is a promising new probe of the evolution of visible matter in our universe, especially during the poorly constrained Cosmic Dawn and Epoch of Reionization. However, in order to separate the 21 cm signal from bright astrophysical foregrounds, we need an exquisite understanding of our telescopes so as to avoid adding spectral structure to spectrally smooth foregrounds. One powerful calibration method relies on repeated simultaneous measurements of the same interferometric baseline to solve for the sky signal and for instrumental parameters simultaneously. However, certain degrees of freedom are not constrained by asserting internal consistency between redundant measurements. In this paper, we review the origin of these *degeneracies* of redundant-baseline calibration and demonstrate how they can source unwanted spectral structure in our measurement and show how to eliminate that additional, artificial structure. We also generalize redundant calibration to dual-polarization instruments, derive the degeneracy structure, and explore the unique challenges to calibration and preserving spectral smoothness presented by a polarized measurement.

Key words: instrumentation: interferometers – dark ages, reionization, first stars.

1 INTRODUCTION

21 cm cosmology – the mapping of neutral hydrogen in the universe using redshifted 21 cm hyperfine emission or absorption – promises unprecedented insight into the intergalactic medium and its density, ionization, thermal, and velocity structure. By probing that evolution, we can better understand the astrophysics of the first stars, galaxies, and black holes that eventually led to the Epoch of Reionization (EoR) (Furlanetto, Oh & Briggs 2006; Morales & Wyithe 2010; Pritchard & Loeb 2012; Loeb & Furlanetto 2013; Zaroubi 2013) and precisely test our cosmological theories at both low (Chang et al. 2008; Wyithe, Loeb & Geil 2008) and high redshifts (McQuinn et al. 2006; Mao et al. 2008).

Realizing that promise requires overcoming two key challenges. The first is that 21 cm fluctuations are predicted to be extremely faint, requiring massive radio telescopes to detect. Masui et al. (2013) used the Green Bank Telescope to detect the signal in cross-correlation with a galaxy survey. The pioneering generation of EoR observatories – including the Low Frequency Array (LOFAR; Patil et al. 2017), the Giant Metrewave Radio Telescope (GMRT; Paciga et al. 2013), the Murchison Widefield Array (MWA; Dillon et al. 2015b; Beardsley et al. 2016; Trott et al. 2016), and the Donald C. Backer Precision Array for Probing the Epoch of Reionization (PAPER; Ali et al. 2015) – have put upper limits on the signal, but it is very difficult for them to achieve more than a tentative detection. That is why the next-generation telescopes have invested in vastly increased collecting area, including the Canadian Hydrogen Intensity Mapping Experiment (CHIME; Bandura et al. 2014) and the Hydrogen Intensity and Real-time Analysis eXperiment (HIRAX; Newburgh et al. 2016) at low *z* and the Hydrogen Epoch of Reionization Array (HERA; Dillon & Parsons 2016; DeBoer et al. 2017) and the Square Kilometre Array (SKA-low; Koopmans et al. 2015) at high *z*.

The second and perhaps greater challenge is avoiding the contamination of the cosmological signal by astrophysical foregrounds $\sim 10^5$ times brighter (Santos, Cooray & Knox 2005; Jelić et al. 2008; Bernardi et al. 2009; Ghosh et al. 2012; Kohn et al. 2016). While foregrounds are spectrally very smooth, the 21 cm brightness temperate fluctuations are expected to exhibit structure on fine Δz (and thus fine frequency) scales. The strategy of excising or downweighting foregrounds by simply excising a few low-$k_\parallel$ Fourier modes (Liu & Tegmark 2011; Chapman et al. 2013; Dillon, Liu & Tegmark 2013; Dillon et al. 2014; Bonaldi & Brown 2015) is problematized by the spectrally complex response of any interferometer, which generically features a position- and frequency-dependent point-spread function that is very difficult to invert accurately (Dillon et al. 2015a).

Fortunately, this effect on typical power-law foregrounds is limited to a region of 2DFourier space (i.e. $k_\parallel$ and $k_\perp$) known as the *wedge* (Datta, Bowman & Carilli 2010; Morales et al. 2012; Parsons et al. 2012b; Vedantham, Udaya Shankar & Subrahmanyan 2012; Hazelton, Morales & Sullivan 2013; Pober et al. 2013; Thyagarajan et al. 2013; Liu, Parsons & Trott 2014a,b). Avoiding foregrounds by working outside the wedge (in the so-called *EoR window*) can achieve more robust foreground isolation, but at the cost of losing considerable sensitivity (Pober et al. 2014). This separation works as long as foregrounds are intrinsically spectrally smooth and the instrument does not impart additional spectral structure.

The instrumental response to polarized foregrounds is another potential cause for concern. 21 cm experiments largely concentrate on unpolarized (Stokes *I*) measurements because the cosmological signal is essentially unpolarized (Mishra & Hirata 2017). However, all instruments have frequency- and direction-dependent responses to linear polarization (Stokes *Q* and Stokes *U*) and circular polarization (Stokes *V*) that can be difficult to disentangle from unpolarized emission without an exquisite instrument model (Jones 1941). Galactic Faraday rotation can turn foreground Stokes *Q* into Stokes *U* and back as a function of frequency – often many times over the band of interest. Since Stokes *Q* and Stokes *U* generally leak differently into our estimate of Stokes *I*, this can introduce spectral structure into the observed foregrounds that depends on the rotation measure. A *Q*- or *U*-contaminated estimate of Stokes *I* power spectrum will show leakage from low $k_\parallel$ (inside the wedge) to higher $k_\parallel$, potentially introducing a bias to measurements made only inside the EoR window (Jelić et al. 2010; Moore et al. 2013; Asad et al. 2015; Lenc et al. 2016; Nunhokee et al. 2017). Measurements made to quantify the magnitude of the effect (e.g. Kohn et al. 2016; Moore et al. 2017; Asad et al. 2018) generally show that, to current sensitivity levels, it is not a show-stopper for EoR science.

Fundamentally, any solution to the problem of foregrounds – polarized or unpolarized – relies on accurate instrumental knowledge. Knowing antenna locations and their polarized, frequency-dependent primary beam response functions is key to both foreground subtraction and to making Stokes-*I* measurements free from polarization leakage (Shaw et al. 2014, 2015).

Accurate instrumental knowledge requires an accurate calibration of our analogue signal chains. Complex antenna bandpasses with complex spectral structure, if not accurately measured and accounted for, can scatter foreground power outside the wedge far in excess of the cosmological signal. As an example: an uncalibrated ~ 1 per cent sinusoidal ripple in visibilities due to cable reflections can produce foreground contamination $\sim 10^6$ times larger than the nominal power spectrum signal, obliterating the EoR window (Ewall-Wice et al. 2016).

Traditionally, antenna bandpass calibration has depended upon producing an accurate sky model through iterative cycles of map-making, source extraction, and recalibration (Pearson & Readhead 1984; Yatawatta et al. 2013). However, any realistic level of error in one's sky or instrument model leads to gain calibration errors. These errors are chromatic because the visibilities used to create them have spectral structure that depends on their length (among other factors). Since the same antennas are involved in both long and short baselines, short baseline measurements inherit chromatic errors from long baselines with more intrinsic spectral structure. This leakage of spectral structure via the gain errors can dramatically restrict the size of the EoR window and thus the sensitivity to the 21 cm signal. Avoiding this effect requires either assumptions about the spectral and temporal smoothness of antenna bandpasses (Yatawatta 2016; Barry et al. 2016) or a reweighting of baselines (Ewall-Wice et al. 2017).

More recently, another approach has been employed to considerable success. Instead of calibrating by reference to a sky-model, arrays with highly redundant configurations can use internal consistency between repeated measurements to solve for most of the calibration degrees of freedom. The idea was developed in the present formalism by Liu et al. (2010), though it has its antecedents in older work (e.g. Wieringa 1992). This *redundant-baseline calibration* relies on a simple counting argument. Since baselines with the same separation between antennas measure an identical integral over the sky, if the number of instantaneous visibility measurements is significantly larger than the number of unique baselines, then it may be possible to solve for antenna-based calibration parameters and unique visibilities simultaneously.

A number of 21 cm arrays have been designed or were reconfigured to take advantage of redundant-baseline calibration. This is in part due to the synergy with the many short baselines measuring the same modes over and over, which helps more economically meet the extraordinary sensitivity requirements for 21 cm cosmology (Parsons et al. 2012a). The technology demonstrator MITEoR pioneered the approach for EoR science, showing for the first time calibration residuals largely consistent with thermal noise (Zheng et al. 2014). PAPER was reconfigured in a redundant configuration for its 32-element observing season (Parsons et al. 2014). The MWA was recently expanded to include a pair of 36-element redundant hexes (Li et al., in review). HERA was designed to be completely redundantly calibratable, including its outrigger antennas (Dillon & Parsons 2016; DeBoer et al. 2017).

That said, redundant-baseline calibration is not the end of the story. Like any calibration scheme based on internal consistency, it falls a bit short of a complete solution for every signal-chain calibration parameter.[1] Roughly speaking, redundant-baseline calibration reduces the calibration problem from one number per frequency channel per antenna element to just a few numbers per frequency for the whole array. Still, the power of redundant calibration is that it vastly reduces the number of degrees of freedom in the calibration problem, yielding precise relative calibration and reducing the complexity of any future *absolute calibration* referenced to the sky.

The linear combinations of gains and visibilities that redundant-baseline calibration cannot solve for, the *degeneracies* of its underlying system of equations, have been the source of much confusion. What exactly can redundant-baseline calibration solve for? And

---

[1]These are similar but not identical to the integer multiple of π phase ambiguities that arise in self-calibration (Yatawatta 2012; Lannes & Prieur 2014).

what are the degeneracies? In this paper, we present a pedagogical explanation of the degeneracies and show their importance by demonstrating the pitfalls they may present. For redundant-baseline calibration to be useful for 21 cm cosmology, it must preserve the spectral smoothness of the observation and it must be possible to understand and undo any modifications made to the degenerate part of the calibration solutions. In principle, any two solutions that differ only in their degeneracies can be transformed and be brought in line with a good absolute calibration by reference to the sky. In practice, this can be somewhat perilous – especially if one wants to use prior knowledge about the instrument to restrict the spectral degrees of freedom in bandpass calibration in a way that enforces relative smoothness and cannot account for discontinuities in frequency. Redundant-baseline calibration as a single modular analysis step has seen increasing use recently (e.g. Zheng et al. 2014, 2017; Ali et al. 2015) and therefore it is valuable to consider the subtleties of maintaining spectral smoothness.

In this paper, we show how the degeneracies be cleanly and self-consistently accounted for so that the resultant calibration solutions can later be combined with a sky-referenced absolute calibration. In Section 2, we review the mathematical formalism underpinning redundant-baseline calibration, explain the mathematical origin of the degeneracies, and show how careless handling of the degeneracies can lead to considerable spectral structure. Then, in Section 3, we generalize the derivation of degeneracies to observations with dual-polarized antennas and highlight new complications and analysis choices that arise.

2 A PEDAGOGICAL REVIEW OF REDUNDANT-BASELINE CALIBRATION

We begin this section with a review of the antenna calibration problem in the context of a single visibility polarization and how it can be addressed by taking advantage of redundant baselines (Section 2.1). Drawing heavily on Liu et al. (2010), we then present an iterative algorithm for minimizing the error in our solution (Section 2.2). Next we explain what exactly redundant-baseline calibration can and cannot solve for (Section 2.3) and how these unsolvable quantities can introduce spectral structure into our calibration solutions if we are not careful (Section 2.4). Finally we step back with an overview of the key assumptions that underlie the use of redundant-baseline calibration (Section 2.5).

2.1 The calibration problem

Fundamentally, the problem of calibration[2] boils down to one key equation:

$$V_{ij}^{\text{obs}}(\nu) = g_i(\nu)g_j^*(\nu)V_{ij}^{\text{true}}(\nu) + n_{ij}(\nu). \tag{1}$$

The observed visibility V_{ij} between antennas i and j at a given time and frequency is related to the true, underlying visibility by a pair of complex and frequency-dependent gain factors, g_i and g_j^*, along with Gaussian random noise n_{ij}. These gain factors incorporate the frequency response of the analogue signal chain, including amplifiers, attenuators, cable reflections, and of course the phase factor

$e^{-2\pi i\tau\nu}$ due to the light-travel time delay τ along the signal path. When we calibrate, we want to solve for both those gains and visibilities. In essence, we have a system of equations for all antenna pairs i and j given by

$$V_{ij}^{\text{obs}}(\nu) = g_i(\nu)g_j^*(\nu)V_{i-j}^{\text{sol}}(\nu), \tag{2}$$

and we want to find the optimal gain and visibility solutions that minimize the error in this system.

In equation (2), we write our visibility V_{i-j} as shorthand for $V(\boldsymbol{r}_i - \boldsymbol{r}_j)$, the visibility for the baseline vector $\boldsymbol{b}_{ij} \equiv \boldsymbol{r}_i - \boldsymbol{r}_j$. Two pairs of identical elements with precisely the same baseline separation between are sensitive to the same mode on the sky. The measured visibilities may be quite different due to differences in signal chains between the four elements, but that fact is incorporated into the gains. In a highly redundant array, there are many more measurements than there are unique baselines.

Consider the example of HERA, which is made of 14 m dishes with a core of 320 hexagonally packed elements. The simplest proposed configuration of a HERA-like instrument – a densely packed hexagonal array of 331 elements (11 on each edge) – has $331(331 − 1)/2 = 54,615$ baselines and thus measures 54615 visibilities. However, it only has 630 unique baseline separations (Dillon & Parsons 2016). That means we have a non-linear system of 54615 equations to determine 630 unique visibilities and the 331 complex gains; the system is vastly overdetermined.[3] With uncorrelated Gaussian noise and an overdetermined system of equations, it is useful to minimize χ^2, defined as

$$\chi^2(\nu) = \sum_{\text{all pairs } i,j} \frac{\left| V_{ij}^{\text{obs}}(\nu) - g_i(\nu)g_j^*(\nu)V_{i-j}^{\text{sol}}(\nu) \right|^2}{\sigma_{ij}^2(\nu)}, \tag{3}$$

where $\sigma_{ij}^2(\nu)$ is the variance of $n_{ij}(\nu)$.

If all one wishes to minimize is the difference between the calibration solution and the data, equation (3) is sufficient. However, one may also wish to impose additional prior information about the sky or the instrument. To enforce that gains are relatively spectrally smooth, as in the consensus optimization of Yatawatta (2016), one could add a penalty factor χ^2 for gain discontinuities between nearby frequency channels.[4] Likewise, Sievers (2017) proposes adding information about sky sources and their statistics into χ^2, as well as information about known deviations from redundancy between putatively identical baselines, to bridge the gap between sky-based calibration and redundant calibration. That said, our focus is the simplest and most straightforward approach to redundant-baseline calibration in which we aim to minimize the χ^2 in equation (3). This non-linear least-squares optimization can be done independently between frequencies (and times, for that matter), so we drop the explicit dependence on ν for now.

2.2 Linearizing and minimizing χ^2

There are many methods for finding set of gains and visibilities that minimize χ^2. They generally involve linearizing equation (2) in

---

[2] For the purposes of this work, we mean signal-chain calibration of antenna-based gains. This is sometimes called *direction-independent* calibration to contrast it with *direction-dependent* calibration that corrects for the shape and possible antenna-to-antenna variation of the primary beam response. The two effects can always be factored, though it can be useful to consider them together in the context of sky-based calibration.

[3] Ignoring for the moment the degeneracies in this system of equations that we will return to in Section 2.3.

[4] Alternatively, one can take the approach of Zheng et al. (2014) and use a Weiner filter to smooth the gains after redundant-baseline calibration based on an estimate of how much time-to-time and frequency-to-frequency variation is due to thermal noise. This will generally not produce the same result the Yatawatta (2016) method, since it smooths about a different χ^2 minimum – one without that penalty factor for non-smoothness in the gains.

order to find parameter solutions, often iteratively. Liu et al. (2010) suggests two. The first involves linearizing equation (2) by taking the logarithm of both sides and then solving independently for the real and imaginary parts of the gains and visibilities. Unfortunately, this method produces biased solutions that do not actually minimize χ^2.

That is why Liu et al. (2010) also advance a linearized method that implements the Gauss–Newton algorithm. Given some starting gains g_i^0 and visibilities V_{i-j}^0, we can rewrite equation (2) as

$$V_{ij}^{\text{obs}} = (g_i^0 + \Delta g_i)(g_j^0 + \Delta g_j)^*(V_{i-j}^0 + \Delta V_{i-j}). \tag{4}$$

Here, the Δ terms are solved for at each iteration, allowing us to update our guesses until we converge on the global minimum. If we assume they are all small because our initial guess is close, we drop the Δ^2 terms and get a linear system of equations for the Δ terms:

$$V_{ij}^{\text{obs}} - g_i^0 g_j^{0*} V_{i-j}^0 = \Delta g_i g_j^{0*} V_{i-j}^0 + g_i^0 \Delta g_j^* V_{i-j}^0 + g_i^0 g_j^{0*} \Delta V_{i-j}. \tag{5}$$

The complex conjugation of the system in equation (5) requires us to break it into real and imaginary parts in order to write the system as a matrix. This yields

$$\begin{aligned}
\mathbb{Re}\left[V_{ij}^{\text{obs}} - g_i^0 g_j^{0*} V_{i-j}^0\right] \\
= \mathbb{Re}[\Delta g_i]\,\mathbb{Re}\left[g_j^{0*} V_{i-j}^0\right] - \mathbb{Im}[\Delta g_i]\,\mathbb{Im}\left[g_j^{0*} V_{i-j}^0\right] \\
+ \mathbb{Re}\left[\Delta g_j\right]\mathbb{Re}\left[g_i^0 V_{i-j}^0\right] + \mathbb{Im}\left[\Delta g_j\right]\mathbb{Im}\left[g_i^0 V_{i-j}^0\right] \\
+ \mathbb{Re}\left[\Delta V_{i-j}\right]\mathbb{Re}\left[g_i^0 g_j^{0*}\right] - \mathbb{Im}\left[\Delta V_{i-j}\right]\mathbb{Im}\left[g_i^0 g_j^{0*}\right]
\end{aligned} \tag{6}$$

and

$$\begin{aligned}
\mathbb{Im}\left[V_{ij}^{\text{obs}} - g_i^0 g_j^{0*} V_{i-j}^0\right] \\
= \mathbb{Im}[\Delta g_i]\,\mathbb{Re}\left[g_j^{0*} V_{i-j}^0\right] + \mathbb{Re}[\Delta g_i]\,\mathbb{Im}\left[g_j^{0*} V_{i-j}^0\right] \\
+ -\mathbb{Im}\left[\Delta g_j\right]\mathbb{Re}\left[g_i^0 V_{i-j}^0\right] + \mathbb{Re}\left[\Delta g_j\right]\mathbb{Im}\left[g_i^0 V_{i-j}^0\right] \\
+ \mathbb{Im}\left[\Delta V_{i-j}\right]\mathbb{Re}\left[g_i^0 g_j^{0*}\right] + \mathbb{Re}\left[\Delta V_{i-j}\right]\mathbb{Im}\left[g_i^0 g_j^{0*}\right].
\end{aligned} \tag{7}$$

We can now write this set of equations compactly as

$$\boldsymbol{d} = \boldsymbol{A}\boldsymbol{x}. \tag{8}$$

Here, $\boldsymbol{d}$ contains both the real and imaginary components of the differences between our guesses and V_{ij}^{obs}. $\boldsymbol{x}$ contains the real and imaginary parts of our Δg_i and ΔV_{i-j} terms, and $\boldsymbol{A}$ contains all the coefficients in equations (6) and (7). $\boldsymbol{d}$ has a length equal to twice the number of observed visibilities while $\boldsymbol{x}$ has a length equal to twice the sum of the number of unique visibilities and the number of antennas.

If $\boldsymbol{N}$ is noise covariance between visibility measurements, then the optimal estimate of the Δ terms is given by

$$\widehat{\boldsymbol{x}} = \left(\boldsymbol{A}^\mathsf{T} \boldsymbol{N}^{-1} \boldsymbol{A}\right)^{-1} \boldsymbol{A}^\mathsf{T} \boldsymbol{N}^{-1} \boldsymbol{d}. \tag{9}$$

Since visibility measurements have uncorrelated noise, $\boldsymbol{N}$ is diagonal and has the form $N_{ij,kl} = \sigma_{ij}^2 \delta_{ik}\delta_{jl}$ where δ_{ik} is the Kronecker delta. However, if all baseline variances σ_{ij}^2 are identical, equation (9) reduces to

$$\widehat{\boldsymbol{x}} = (\boldsymbol{A}^\mathsf{T}\boldsymbol{A})^{-1}\boldsymbol{A}^\mathsf{T}\boldsymbol{d}. \tag{10}$$

Thus, to find the set of gains and visibilities that minimizes χ^2, we iteratively set up and solve equation (10) until we reach the desired level of convergence. In practice, this is complicated by the fact that $\boldsymbol{A}^\mathsf{T}\boldsymbol{A}$ is not invertible – a consequence of the inherent degeneracies in the system of equations that we will discuss at length in Section 2.3 – and therefore requires a modified inversion technique like the Moore–Penrose pseudo-inverse.

This method is sufficient for understanding how the structure of $\boldsymbol{A}^\mathsf{T}\boldsymbol{A}$ relates to χ^2, but we also want to emphasize that the Gauss–Newton method of Liu et al. (2010) is not the only approach to redundant-baseline calibration. Noorishad et al. (2012) advance the weighted alternating least-squares technique, which linearizes equation (2) by alternatingly holding gains and visibilities constant and zeroing-in on the minimum. Though this can take more steps to converge, it can also be much faster because it can cut the size of the matrix inversions, the rate-limiting step in most least-squares minimizations. The `omnical` code developed for MITEoR (Zheng et al. 2014, 2017) is based on this approach.[5] Ram Marthi & Chengalur (2013) explore another well-known approach, steepest descent, which is much faster per iteration (no matrix inversion is required) but can take many more steps to converge to high precision. The Levenberg–Marquardt algorithm would be a natural compromise between steepest descent and Gauss–Newton, but it and the many other non-linear least-squares minimization techniques are outside the scope of this work.

2.3 Degeneracies in redundant-baseline calibration

Regardless of the technique for solving for gains and visibilities and regardless of how overdetermined the system in equation (2) is, the structure of χ^2 (and thus $\boldsymbol{A}$) guarantees that there will always be a few terms that we cannot solve for with redundant-baseline calibration. These unsolvable quantities are the degeneracies of the system.[6] They manifest as changes that one can make to g_i or V_{i-j}^{sol} or both such that $g_i g_j^* V_{i-j}^{\text{sol}}$ is unchanged, leaving χ^2 also unchanged. Once χ^2 is minimized, any change in these degeneracies keeps χ^2 minimized. In single-polarization calibration, there are exactly four such degeneracies per frequency and per time. They are as follows:

(1) **The overall amplitude.** If $g_j \to A g_j$ and $V_{i-j}^{\text{sol}} \to V_{i-j}^{\text{sol}}/A^2$, then $g_i g_j^* V_{i-j}^{\text{sol}}$ is unchanged.

(2) **The overall phase.** If $g_j \to g_j e^{i\psi}$, the changes in g_i and g_j^* always exactly cancel out.

(3) **The x-phase gradient.** If $g_j \to g_j e^{i\Phi_x x_j}$ and $V_{i-j}^{\text{sol}} \to V_{i-j}^{\text{sol}} e^{-i\Phi_x \Delta x_{ij}}$, then $g_i g_j^* V_{i-j}^{\text{sol}}$ is unchanged for all baselines.

(4) **The y-phase gradient.** Likewise, if $g_j \to g_j e^{i\Phi_y y_i}$ and $V_{i-j}^{\text{sol}} \to V_{i-j}^{\text{sol}} e^{-i\Phi_y \Delta y_{ij}}$, then $g_i g_j^* V_{i-j}^{\text{sol}}$ is similarly unchanged.

The first two degeneracies are due to straightforward cancellations. The third and fourth are a bit more subtle. They rely on the fact one can add a linear phase gradient $\boldsymbol{\Phi} \equiv (\Phi_X, \Phi_Y)$ in antenna position $\boldsymbol{r}_j \equiv (x_j, y_j)$ which, due to the complex conjugation of g_j^*, produces a phase factor that depends only on baseline $\boldsymbol{b}_{ij} \equiv (\Delta x_{ij}, \Delta y_{ij}) = (x_i - x_j, y_i - y_j)$. This can be exactly cancelled by rephasing the unique visibility solutions, which can depend on baseline vector but not on absolute antenna position. The phase gradient terms are often referred to as the tip-tilt terms since they correspond to moving the phase centre and thus the apparent position of sources on the sky (Zheng et al. 2014, 2017).

---

[5]The most recent version of `hera_cal` has interfaces to both `omnical`, which is based on the alternating least-squares method, and `redcal`, which implements the full matrix-inversions of the Gauss–Newton method using the `linsolve` package. The former is much faster, but more vulnerable to false-minima (Li et al., in review).

[6]For more concise but less pedagogical discussions of these degeneracies, see Liu et al. (2010), Zheng et al. (2014), Zheng et al. (2017), or Li et al. (in review).

We can think of these four degeneracies as four different vectors in the solution space of the gains and visibilities. Movement along any of these directions does not affect χ^2, so there is no way to know the 'optimal' amount to move along these directions at each iterative step. This is manifested in the structure of $A^\mathsf{T}A$, which has four zero-eigenvalues and a nullspace spanned by the locally linearized versions of these vectors. That said, it is difficult to simply compute the eigenvectors and eigenvalues of $A^\mathsf{T}A$ and verify that they are the same four vectors by inspection. Eigenvalue decomposition algorithms will produce some unpredictable linear combination of the four degeneracies precisely because they are degeneracies. However, if one adds extra constraints to the system in equation (2) that fix the degeneracies, $A^\mathsf{T}A$ becomes full-rank. Since these degeneracies are independent between frequencies, nothing in the χ^2 minimization algorithm requires that they be spectrally smooth.

2.4 Avoiding spectral discontinuities in the degeneracies

The immediate question posed by these degeneracies is simply 'what should we do about them?' Ultimately, we need to know how bright the sky is and we need to know where our array is pointed. We will need to fix the degeneracies with a sky-referenced absolute calibration, either through a diffuse sky-model (Zheng et al. 2014, 2017) or via imaging and traditional self-calibration (Li et al., in review). However, uncertainties about both our sky-model and our beams can lead to enough spectral structure to bias the entire EoR window (Ewall-Wice et al. 2017). Though we need to calibrate four degeneracies per frequency (and time), we will likely want to restrict the degrees of freedom in our absolute calibration so as to avoid adding spectral structure to our overwhelmingly bright foregrounds. If we plan to constrain our absolute calibration in this way, we must ensure that redundant-baseline calibration does not add any spectral structure within the degenerate subspace that absolute calibration may not take out.

Fixing the amplitude degeneracy to achieve this goal is easy; all we have to do is ensure that the degeneracy is constant over frequency and time. There are several ways to do this. We could fix the amplitude of the gain of a particular reference antenna or we could fix the mean amplitude of all antennas. We choose to set the mean amplitude of the gain products (as they appear in equation 3) to 1. This preserves the mean visibility amplitude, though none of these choices make much difference and or risks adding any spectral structure.

Phase calibration is trickier. Naively approaching the three phase degeneracies in the same way – for example by setting the mean gain phases and gain phase slopes to zero – is problematized by phase-wraps. Antenna gain phases generically evolve with frequency. When one antenna's phase-wraps from π to $-\pi$, the ψ that we need to add to all phases to make the average 0 might be quite different from one frequency to the next. This creates a phase discontinuity in the gains and in the visibilities. Unless it is fixed by absolute calibration, such a discontinuity in frequency space will leak foreground power to all modes in Fourier space and will contaminate the EoR window. Working in terms of real and imaginary parts of the gains and visibilities does not solve the phase-wrap problem because the degeneracies are inherently degeneracies in amplitude and phase.

To understand how fixing the degeneracies to a constant value can produce this sort of spectral structure, let us consider separating antenna gains into the form

$$g_j(\nu) \equiv |g_j(\nu)|e^{i\phi_j(\nu)-2\pi i \tau_j \nu}. \tag{11}$$

τ_j is the delay for the jth antenna and ϕ_j is its residual phase. In practice, the phase structure of $g_j(\nu)$ is dominated by τ_j which comes from light-travel time delays along the cable and other delays in the analogue and digital signal chain. While ϕ_j generally does not phase-wrap over the band of interest, typical delays for EoR instruments like MWA, PAPER, and HERA of the order of 10s or 100s of ns over the typical ∼100 MHz bandwidth create many phase-wrappings for each antenna.

In Fig. 1, we show this effect more clearly for a noise-free simulation of a 7-element version of HERA (14.6 m separation between seven dishes arranged in a hexagon). We simulate $\phi_j(\nu)$ with 1 MHz channels over a fiducial 100 MHz bandwidth to be small and spectrally smooth (i.e. describable by only a few Fourier modes). We also simulate random delays, τ_j^{true}, but then pre-calibrate them out imperfectly with estimated delays τ_j^{est}:

$$V_{ij}^{\text{pre-cal}}(\nu) = V_{ij}^{\text{obs}}(\nu)\exp\left[2\pi i(\tau_i^{\text{est}} - \tau_j^{\text{est}})\right]. \tag{12}$$

To the extent that $\tau_j^{\text{true}} \approx \tau_j^{\text{est}}$, equation (12) reduces to

$$V_{ij}^{\text{pre-cal}}(\nu) \approx \left|g_i(\nu)g_j^*(\nu)\right|e^{i\phi_i(\nu)-i\phi_j(\nu)}V^{\text{true}}. \tag{13}$$

In our simulation, however, we intentionally introduce residual delay errors on each antenna at the ∼ 50, ∼ 5, or ∼0.5 ns levels. These realistically represent the cases where delays are unaccounted for, where delays are corrected for but poorly, and where delays are well-measured and taken out. We then perform redundant-baseline calibration, perfectly minimizing χ^2 in all three cases,[7] and show how our decision to fix the three phase degeneracies to zero creates spectral structure in the gain solutions.

Fig. 1 shows us that if we do not do a good job correcting the delays before redundant-baseline calibration and degeneracy-fixing, the difference between the true delays and the delay guesses will phase-wrap (Fig. 1, top row, left-hand and middle panels), causing phase jumps in all antenna gains (bottom row, left-hand and middle panels). While a few such jumps may be easily spotted, increasing the number of antennas or the delay error level makes this effect so complicated that it becomes difficult to perform sky-referenced absolute calibration.[8] Assuming that phases are zeroed at the centre of the band and phase-wraps occur at π and $-\pi$, then we need $|\Delta\tau_{\text{max}}| < (\nu_{\text{max}} - \nu_{\text{min}})^{-1}$. This yields a maximum delay error of 10 ns for 100 MHz of bandwidth. The requirement gets more stringent if instead phases have $\phi = 0$ at 0 MHz; in the above example we needed to restrict phase errors to $|\Delta\tau_{\text{max}}| < 2.5$ ns to avoid phase-wraps, which is conservative but not too terribly challenging.

We cannot just do anything – we must fix the degeneracies if we want to keep any later degeneracy-fixing to match the sky (i.e. absolute calibration) as simple and smooth as possible. But in doing so, we must make sure we avoid phase-wraps from relatively small delay errors. It follows then we must begin degeneracy-fixing with a smooth, approximate solution. A single delay per antenna is usually

---

[7]Technically, a noise-free simulation has undefined χ^2, but we simply mean that all the differences in the numerator of equation (3) are zero to within numerical precision.

[8]Zheng et al. (2017) propose an alternative degeneracy-fixing scheme where three antennas have their phases set to zero. This can ameliorate the problem seen in Fig. 1, which presents the worst-case scenario in the left column. However, the alternative approach relies on prior knowledge that the reference antennas are stable and working properly – which may not always be the case. It also is more likely to phase the array far from zenith, making absolute calibration somewhat more difficult.
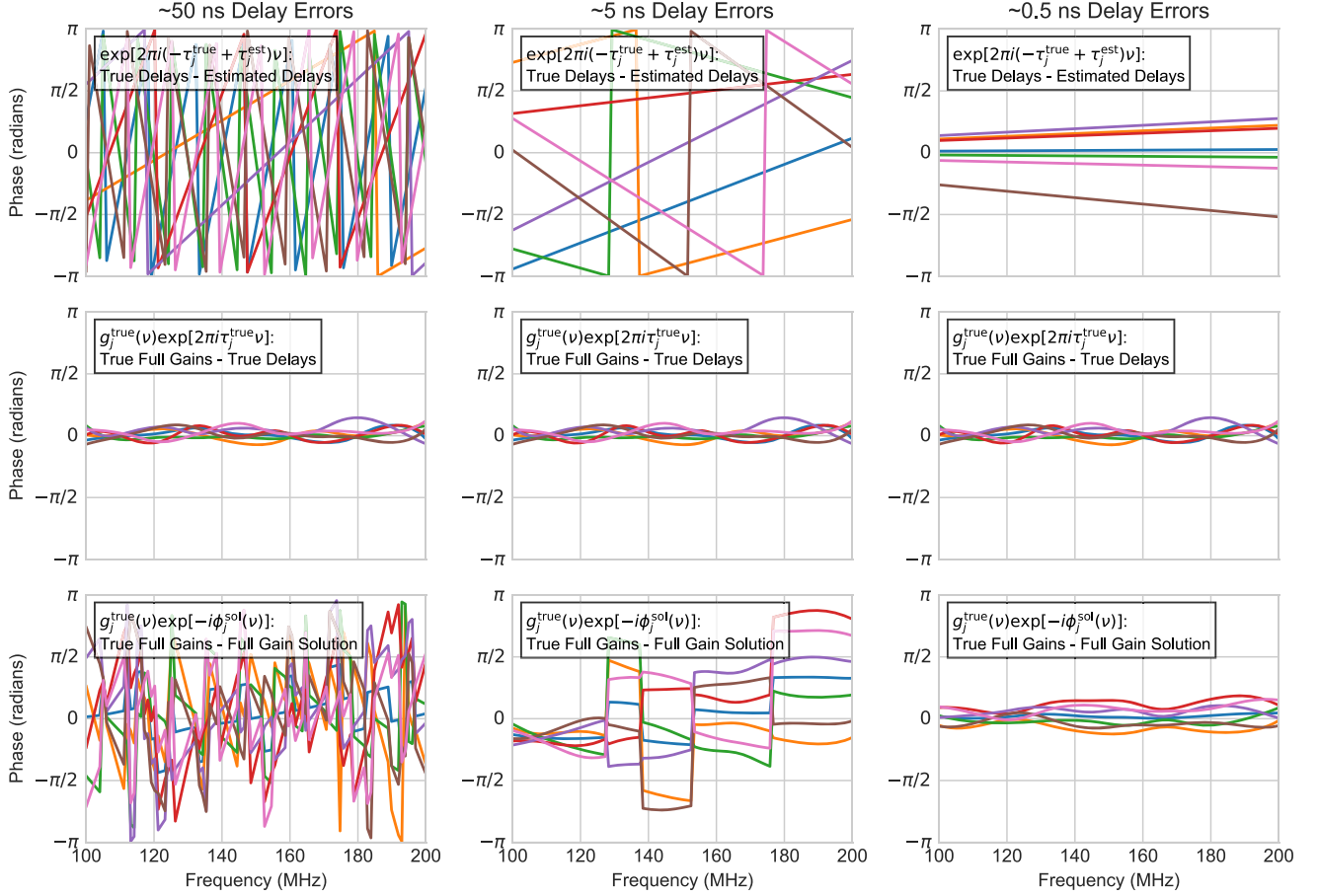
Figure 1. Getting accurate delays before performing redundant-baseline calibration is key to avoiding spectral structure introduced by degeneracy-fixing. Here, we show the phases of complex antenna gains as a function of frequency for each element in a simulated 7-element HERA observation, comparing the case where the errors in our estimate of the random delays are ~ 50, ~ 5, and ~ 0.5 ns. Each line is a different antenna. Different levels of delay errors produce different amounts of phase-wrapping, seen most clearly in the difference between the true delays, τ_j^{true}, and the estimated delays, τ_j^{est} (top row). When the true delays are taken out of the true full gains (middle row), the residual phase, ϕ_j from equation (11), is small and simulated to have little intrinsic spectral structure. However, when the true gains are pre-calibrated with differing levels of delay errors, the ensuing phase-wraps combine with our degeneracy-fixing method to produce discontinuities in the calibration solution if the delay errors are large enough (bottom row). In all three noise-free simulations, the gain and visibility solutions converge to perfect redundancy with the input simulation; redundant-baseline calibration does its job minimizing χ^2. Examining the 5 ns case, it is clear that individual antenna phase-wraps in the delay error (top row) correspond to and in fact cause dramatic phase jumps in the gain errors (bottom row). The 0.5 ns case, which does not phase-wrap, still shows some differences between the true gains and the gain solutions, but this is expected. While redundant-baseline calibration did not give exact the right answer inside the degenerate subspace, its errors are as spectrally smooth as the simulated gains and therefore much more easily removed with absolute calibration.

sufficient. Since antenna delays are generally stable over time, using the same delays for multiple integrations also gives us consistent degeneracy structure from integration to integration, making absolute calibration easier. Also, since we need a good starting point for linearized redundant-baseline calibration anyway, an accurate per-antenna delay kills two birds with one stone.

Thankfully, good methods for finding antenna delays already exist, including both a sky-based approach (Kern et al. 2017) and one using only redundant-baseline pairs instead of a sky model. This technique, called `firstcal`,[9] fits the ratio of two redundant visibilities to a single τ which is a combination of the four antenna τs involved in the ratio. By building up a system of equations much like that of redundant-baseline calibration, τ_i can be largely solved for. A similar method was used in Parsons et al. (2014), though a more complete and formal description can be found in Li et al.

(in review). This improves upon Zheng et al. (2014) which uses a sky and instrument model for so-called rough calibration. It also improves upon Zheng et al. (2017) which suggests a per-frequency technique that is invulnerable to phase-wrapping but can create much more temporal structure than the more physically motivated delay approach of `firstcal`. There are still three `firstcal` degeneracies, very similar to the ones in Section 2.3, and they affect whether source images appear in the right place on the sky. However, since these degeneracies do not source the kind of spectral structure in the gains that causes the phase-wrapping we are focused on in this work, we forgo a more detailed exploration of them here.

2.5 Assumptions of redundant-baseline calibration

The need for an accurate delay for each antenna in order to avoid spectral structure highlights the fact that redundant-baseline cali-

---

[9]Also available in the most recent version of `hera_cal`.

bration relies on a number of assumptions to work well. While a single delay per antenna has, in our experience, been close enough to restrict our calibration space to the $-\pi$ to π range, there is no a priori reason for this to be true. It is possible that we need to initialize our redundant calibration with a low-order polynomial in ν, instead just a linear phase slope created by a delay. As long as we can get within a phase-wrap, we do not have to worry about degeneracy-fixing.

That said, there are many other possible deviations from redundancy. It is beyond the scope of this work to quantify the impact of these on cosmological measurements. However, our pedagogical review would be incomplete with an explicit enumeration of the possible future issues with redundant-baseline calibration:

(i) As we already discussed above, it is essential that we can find a spectrally smooth starting solution, ideally a single delay per antenna, that lets redundant-baseline converge to the right answer and prevents degeneracy-fixing from introducing spectral structure into the gain and visibility solutions.

(ii) If antennas are not placed perfectly redundantly or if the array is not perfectly coplanar, we break the assumption that nominally redundant baselines are actually measuring the same mode on the sky. Since small model errors can create worrisome spectral structure (Ewall-Wice et al. 2017), this deserves further investigation. However, this effect can potentially be ameliorated by a couple of extensions to standard redundant-baseline calibration framework. If the deviations from the redundant grid are well-measured, one can add additional degrees of freedom that admit such non-redundancy self-consistently by Taylor expanding the visibility solutions as a function of baseline (Liu et al. 2010). Alternatively, the framework of Sievers (2017) incorporates partial redundancy statistically into χ^2, penalizing not-quite redundant baselines less for having differing visibility solutions. Both require further investigation outside the scope of this work.

(iii) While one generally seeks to make one's elements as identical as possible, some variance from primary beam to primary beam is inevitable. Unfortunately, no method we know of can fully absorb this effect into redundant calibration if the actual beam-to-beam variation is poorly characterized. This is likely to induce spectral structure into calibration as well, as was seen in simulations of sky-based calibration with primary beam errors (Ewall-Wice et al. 2017). This can also manifest as an antenna-to-antenna variation of polarized leakage, which we will discuss in detail in Section 3.3.

(iv) Correlated noise between two antennas, either due to transmission by one or due to communication between nearby inputs on a printed circuit board, leads to apparent signal that is not redundant between baselines (Kohn & Aguirre 2015). This crosstalk can be mitigated by good antenna design (DeBoer et al. 2017), in the electronics with Walsh modulation (Zheng et al. 2014), or by filtering signals that apparently do not change as the sky rotates (Parsons et al. 2016), but the exact magnitude of the effect remains an open question.

(v) Pairs of nominally redundant baselines that are separated from each other by several km see a different isoplanetic patch of the ionosphere (Vedantham & Koopmans 2015), leading to two slightly differently warped skies and thus different visibilities. At these distances we might also begin to worry about array non-coplanarity and pointing errors due to the curvature of the Earth. We include this effect for completeness; it is not expected to affect the redundant-baseline calibration of any operational or planned highly redundant interferometer.

3 INCORPORATING POLARIZATION IN REDUNDANT-BASELINE CALIBRATION

Having summarized the challenges presented by the challenges of redundant-baseline calibration, we now are ready to relax the simplification that we only need to calibrate a single antenna polarization. Most 21 cm interferometers (including LOFAR, GMRT, MWA, PAPER, CHIME, HIRAX, HERA, and SKA-low) are dual-polarization instruments. They simultaneously measure electromagnetic signals from two orthogonal antenna polarizations, n and e.[10] These are correlated to form four visibility polarizations, V_{ij}^{ee}, V_{ij}^{en}, V_{ij}^{ne}, and V_{ij}^{nn}. The cross-polarized visibilities, V_{ij}^{en} and V_{ij}^{ne}, will generally have much lower signal-to-noise (SNR) ratios than the parallel-polarized visibilities, V_{ij}^{ee} and V_{ij}^{nn}, though they have the same noise levels. This is because the cross-polarized visibilities are only sensitive to intrinsically linearly polarized emission (<1 per cent of the foregrounds) and leakage from unpolarized emission (~ 10 per cent of the foregrounds near the horizon) which usually dominates in 21 cm cosmology (Lenc et al. 2016; Kohn et al. 2016; Nunhokee et al. 2017). However, all four visibility polarizations are necessary to make images of each of the four Stokes parameters I, Q, U, and V (Thompson, Moran & Swenson 2017). The generalization to polarized redundant-baseline calibration has not been addressed in the literature and presents new and unique challenges.

Generalizing equation (2), we now model our observations as

$$V_{ij}^{ab,\text{obs}} = g_i^a g_j^{b*} V_{i-j}^{ab,\text{sol}} \tag{14}$$

where a and b stand in for either e or n for antennas i and j, respectively, and we have again omitted the explicit frequency dependence.[11] If we only want to calibrate the highest SNR visibilities – V_{ij}^{ee} and V_{ij}^{nn} – then this problem decouples into the independent calibration of the East–West visibilities with the East–West gains and the North–South visibilities with the North–South gains. Degeneracy-fixing is the same as in Section 2.3, effectively leaving the problem of relative calibration between e and n for later absolute calibration.

Though we have four times as many visibilities, we only have double the number of complex gains; this gives us a new way to connect together visibilities self-consistently via the gains. Ideally, we would solve for everything simultaneously, minimizing a χ^2 generalized from equation (3) to

$$\chi^2 = \sum_{a,b \in e,n} \left[\sum_{\text{all pairs } i,j} \frac{\left| V_{ij}^{ab,\text{obs}} - g_i^a g_j^{b*} V_{i-j}^{ab,\text{sol}} \right|^2}{\left(\sigma_{ij}^{ab} \right)^2} \right], \tag{15}$$

This couples all four visibility polarizations via the antenna gains into one large system of equations. It produces a very similar A matrix to that described in equations (6) and (7), but with four times as many equations, nearly four times as many parameters to solve for.[12]

---

[10]These are often referred to x- and y-polarizations (to highlight their orthogonality), but we use e and n (for East–West and North–South orientations) to avoid confusion with antenna position. Of course, there is no requirement that the polarizations must line up with the cardinal directions, but we are assuming that they are orthogonal to one another.

[11]Though for now we assume that a single complex gain per antenna is sufficient, we will relax this assumption in Section 3.3.

[12]An explicit derivation of A for full-polarization χ^2 minimization is not particularly illuminating and so we omit it here.

3.1 Degeneracies in polarized redundant-baseline calibration

Just as in the single polarization case, the form of χ^2 in equation (15) features several different ways to modify the gains and visibilities so as to leave χ^2 unchanged. Once again, these degeneracies are key to maintaining spectral smoothness and depend on precisely what analytical approach we take.

The polarized extension of redundant-baseline calibration where we only consider the high-SNR parallel-polarized visibilities, V_{ij}^{ee} and V_{ij}^{nn}, ignoring V_{ij}^{en} and V_{ij}^{ne}, is trivial. The e and n polarizations decouple and there are precisely eight degeneracies – the same four we saw in Section 2.3 for each of the two antenna polarizations. This approach, which we refer to as *2-pol* calibration (in contrast *1-pol* above or *4-pol* below), is simpler and computationally cheaper, but comes at the cost of throwing away some information. The cross-polarized visibilities are left to be calibrated later using the gains from the 2-pol solution.

The best approach in an ideal world would be to minimize the χ^2 for both gain polarizations and all four visibility polarizations simultaneously. Decomposing the associated A matrix reveals that the nullspace is spanned by six eigenvectors, not eight. Bringing in V_{ij}^{en} and V_{ij}^{ne} apparently allows us to solve for two of the eight degeneracies. In retrospect, this is not surprising as there are six independent ways that gains and visibilities can be modified without changing χ^2 in equation (15). Those 4-pol degeneracies are as follows:

(1) **The overall *e*-polarization amplitude.** All $g_j^e \longrightarrow A_e g_j^e$.

(2) **The overall *n*-polarization amplitude.** All $g_j^n \longrightarrow A_n g_j^n$. The combined effect of these two can be perfectly cancelled by transforming the visibility solutions as

$$V_{i-j}^{ab,\text{sol}} \rightarrow V_{i-j}^{ab,\text{sol}}/(A_a A_b) \tag{16}$$

where both a and b stand in for either e or n.

(3) **The *e*-polarization overall phase.** All $g_j^e \longrightarrow g_j^e e^{i\psi_e}$.

(4) **The *n*-polarization overall phase.** All $g_j^n \longrightarrow g_j^n e^{i\psi_n}$. Once again, these effects can be cancelled by transforming the visibility solutions as

$$V_{i-j}^{ab,\text{sol}} \rightarrow V_{i-j}^{ab,\text{sol}} e^{i(\psi_a - \psi_b)}. \tag{17}$$

Here V_{i-j}^{nn} and V_{i-j}^{ee} are unmodified, as in the 1-pol case.

(5) **The *x*-phase gradient.** Just as in the 1-pol case, if $g_j^a \rightarrow g_j^a e^{i\Phi_x x_j}$ and $V_{i-j}^{ab,\text{sol}} \rightarrow V_{i-j}^{ab,\text{sol}} e^{-i\Phi_x \Delta x_{ij}}$, then χ^2 is unchanged.

(6) **The *y*-phase gradient.** Likewise, if $g_j^a \rightarrow g_j^a e^{i\Phi_y y_j}$ and $V_{i-j}^{ab,\text{sol}} \rightarrow V_{i-j}^{ab,\text{sol}} e^{-i\Phi_y \Delta y_{ij}}$, then χ^2 is unchanged.

It is notable that the last two degeneracies are polarization-independent. By introducing V_{ij}^{en} and V_{ij}^{ne}, we have broken the two independent phase gradient degeneracies that we get in the 2-pol case. To understand this, consider the alternative case where $g_i^e \rightarrow g_i^e e^{i\Phi_x^e x_i}$ and . We would need to transform $V_{i-j}^{en} \longrightarrow V_{i-j}^{en} e^{-i(\Phi_x^e x_i - \Phi_{x}^n x_j)}$. If $\Phi_x^e \neq \Phi_x^n$, then the phase factor cannot be factored in the exponent and the new visibility becomes an explicit function of antenna position, not just the baseline vector. But, assuming antennas are identical and perfectly positioned, the unique visibility V_{i-j}^{en} cannot depend on position. This contradiction is only resolved only if $\Phi_x^e = \Phi_x^n = \Phi_x$ and $\Phi_y^e = \Phi_y^n = \Phi_y$.

3.2 Difficulties of 4-polarization redundant-baseline calibration

Breaking two additional degeneracies – and therefore having to rely on absolute calibration to solve for two fewer numbers per frequency
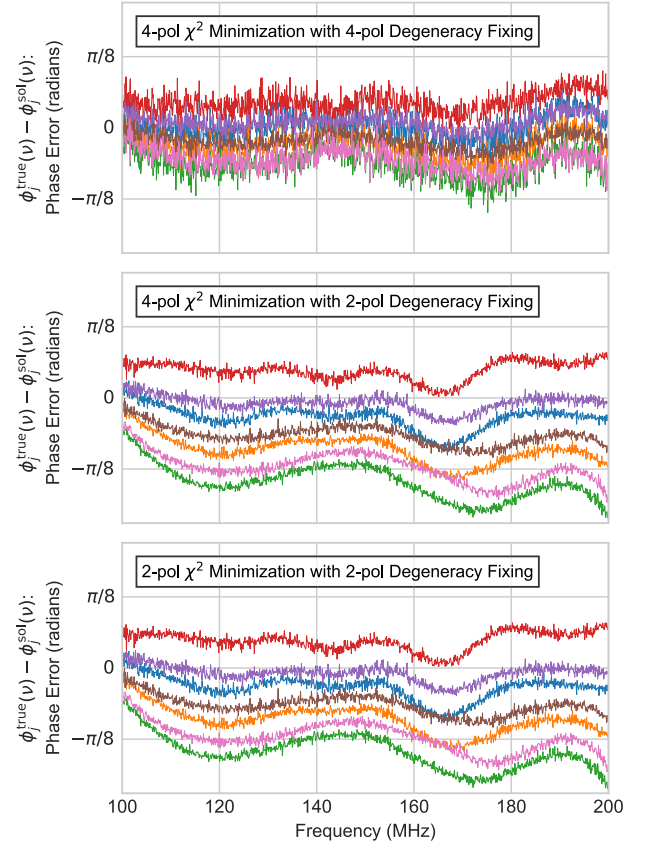
Figure 2. While 4-pol calibration has the potential to break two of the eight redundant-baseline calibration degeneracies, these are only broken by the addition of cross-polarized visibilities that tie e and n polarizations together. These visibilities generally have much less sky-signal, which means that they have low SNR. Fixing only the six true degeneracies described in Section 3.1 to 1 (for amplitudes) or 0 (for phases) gives smaller errors on the gains at the cost of much nosier gain solutions. Here, we plot the error in the phase of a subset of the antenna gains of a 19-element array with three different calibration techniques. First, we show the result of 4-pol redundant-baseline calibration after fixing the six degeneracies (top panel). Next we show the result of the same χ^2 minimization procedure but after fixing the eight degeneracies of 2-pol calibration (middle panel). This essentially throws out the information in those degenerate modes, saving the problem for later absolute calibration. It produces very similar results to having simply excluded the cross-polarized visibilities from both χ^2 minimization and degeneracy-fixing (bottom panel).

and time – sounds really useful. Unfortunately, it is not so simple. Cross-polarized visibilities, as we mentioned above, have much lower SNR than parallel-polarized visibilities. Since introducing V_{ij}^{en} and V_{ij}^{ne} broke two of the four phase gradient degeneracies, it follows that our constraint within that subspace comes only from cross-polarized observations. While we have information to break the degeneracy, it is very noisy compared to the rest of the data that informs our calibration.

The effect of this is shown most clearly in Fig. 2. We simulate spectrally smooth gains and visibilities for a 19-element HERA-like array, much like we did in Fig. 1 (without delay errors), but now for both e and n antenna polarizations and all four visibility polarizations. Then we add white noise at equal levels to all 'observed' visibilities, but only after decreasing the signal level in the cross-polarized visibilities by a factor of 10. This gives our parallel-

polarized visibilities an SNR of ∼10 and our cross-polarized visibilities an SNR of ∼1.

We then assess three different calibration techniques and plot the gain phase errors they produce. The first is standard 4-pol redundant-baseline calibration (top panel). The second is standard 2-pol calibration (bottom panel). The third is a hybrid approach where we minimize χ^2 as in the 4-pol approach with all four visibility polarizations, but then throw away V_{i-j}^{en} and V_{i-j}^{ne} and fix the eight degeneracies of the 2-pol case (middle panel).

The 4-pol approach has smaller phase errors than the other two, but it is much noisier. The additional constrained degrees of freedom in the 4-pol approach, the two phase gradient degeneracies, are only informed by the low SNR cross-polarized visibilities. No inverse noise variance weighting scheme can help us here. It is not that the high-SNR parallel-polarized visibilities are given too little weight relative to the low-SNR cross-polarized visibilities. The parallel-polarized visibilities get zero weight in the determination of these two modes, regardless of how much we up-weight them, because they contain no information about the modes. The rest of the χ^2 landscape is dominated by the parallel-polarized visibilities, even when they get equal noise weighting, because they have much larger signals. The extra noise effect goes away if cross-polarized and parallel polarized visibilities have the same SNR. Likewise, the improvement in the phase errors goes away as the signal level in the cross-polarized visibilities goes to zero.

The smaller errors are unlikely to be worth the additional cost. This extra noise gets spread through the gains to the calibrated parallel-polarization visibilities, giving them a much higher effective noise level that will integrate down more slowly with repeated observations. It seems better to effectively impose a strong prior that we trust sky-based absolute calibration for these modes – which ultimately correspond to the relative pointing of the array between n and e – in order to not introduce additional spectral structure into the gains and high-SNR visibilities.

It may be possible to extract a constrained or smoothed solution for the two additional degeneracies that has many fewer degrees of freedom and thus is less noisy. Doing so will require additional assumptions about true relative phase gradients and so we leave that investigation to future work.

Likewise, one possible instrumental solution would be to rotate some of the antenna feeds by ±45°, producing a dipole pattern that looks, for example, like + × + × from above. We would then have four dipole orientations, e, n, +, and −, and up to 16 visibility polarizations per baseline. While cross-polarized visibilities have very low SNR, visibilities like V_{ij}^{e+} should actually exhibit only a moderate hit to SNR. This could connect together all visibility in one redundant-calibration system through relatively high-SNR measurements. On the other hand, it increases the complexity of the calibration and mapping problems and reduces the number of simultaneously redundant baselines for any particular separation. Not all array configurations that are redundant with two dipole orientations polarizations are still redundant when some of the antennas are rotated. The possibility merits further investigation, though because no current or planned experiment features a mix of e/n and +/− orientated feeds, we again leave it for future work.

This all raises another question: is it worthwhile to bother including the cross-polarized visibilities in χ^2 minimization at all? It is certainly more computationally difficult to perform fully polarized redundant-baseline calibration. By doubling the number of gains and quadrupling the number of visibilities solved for simultaneously, we increase the cost of matrix inversion in equation (9) by up to a factor of 64, though if we only have to perform the inver-

sion half as often as in the 2-pol case. However, if our chief aim is to make a measurement of polarized foregrounds, then excluding cross-polarized visibilities from χ^2 will generally lead to calibrated cross-polarized visibilities that are less redundant with each other – effectively a noisier measurement. Even if all we care about is parallel-polarized visibilities for a PAPER-style power spectrum analysis (Parsons et al. 2014), we should expect that including cross-polarized visibilities leads to lower noise. More information is always better, even if it is noisy.

That decrease in noise is real, though difficult to see in Fig. 2. It turns out to be quite small in our simulation. Comparing the gains only in the non-degenerate subspace of 2-pol redundant calibration, we find that χ^2 minimization with all four visibility polarizations decreases the errors in the gains by 0.62 per cent. This improvement is small for a number of reasons: the number of data points doubled, but the new ones all had low SNR. The number of new variables also increased for the cross-polarized visibility solutions, though it did not quite double. The improvement in the parallel-polarized visibility solutions, only 0.0044 per cent, is roughly the fractional improvement squared because cross- and parallel-polarized visibilities are connected only indirectly in the linear system in equation (14). Apparently, as long as the assumption of redundancy holds, we need to choose between a small decrease in noise and a sizable increase in computational cost.

3.3 The effect of non-redundant polarization leakage

Another concern about 4-pol χ^2 minimization is the possibility that cross-polarized visibilities may be less redundantly calibratable than parallel-polarized visibilities. The sources of non-redundancy like position errors and beam-to-beam variation, which we discussed in Section 2.5, affect both types of visibilities in similar ways. This is not the case with antenna polarization leakage, often referred by its symbol as *D-terms*.[13] *D*-terms represent the complex response of the e or n antenna feed to the n- or e-polarized component, respectively, of the local electric field (Hamaker, Bregman & Sault 1996; Sault, Hamaker & Bregman 1996). They are also another way for Faraday rotation to produce spectral structure in our Stokes I estimate due to varying Stokes U and Stokes V leakage as a function of frequency.

Following Thompson et al. (2017), we can write the observed visibilities in terms of the 'true' visibilities as

$$\frac{V_{ij}^{ee,\text{obs}}}{g_i^e g_j^{e*}} = V_{ij}^{ee} + D_i^{n \to e} V_{ij}^{ne} + D_j^{n \to e*} V_{ij}^{en} + D_i^{n \to e} D_j^{n \to e*} V_{ij}^{nn}$$

$$\frac{V_{ij}^{en,\text{obs}}}{g_i^e g_j^{n*}} = V_{ij}^{en} + D_i^{n \to e} V_{ij}^{nn} + D_j^{e \to n*} V_{ij}^{ee} + D_i^{n \to e} D_j^{e \to n*} V_{ij}^{ne}$$

$$\frac{V_{ij}^{ne,\text{obs}}}{g_i^n g_j^{e*}} = V_{ij}^{ne} + D_i^{e \to n} V_{ij}^{ee} + D_j^{n \to e*} V_{ij}^{nn} + D_i^{e \to n} D_j^{n \to e*} V_{ij}^{en}$$

$$\frac{V_{ij}^{nn,\text{obs}}}{g_i^n g_j^{n*}} = V_{ij}^{nn} + D_i^{e \to n} V_{ij}^{en} + D_j^{e \to n*} V_{ij}^{ne} + D_i^{e \to n} D_j^{e \to n*} V_{ij}^{ee}, \quad (18)$$

---

[13]Two different effects are called 'polarization leakage,' and it is important to be clear what we mean. Our earlier discussions of visibility polarization leakage (Sections 1 and 3) were focused on leakage from Stokes I into cross-polarized visibilities and from Stokes Q and U into the parallel-polarized visibilities. These effects are largely due to beam geometry and are biggest near the horizon for a zenith-pointed instrument. They peak at ∼10 per cent for PAPER (Nunhokee et al. 2017), though they can vary substantially between feed and element designs. In this paper, we mean only the direction-independent polarization leakage terms described by equation (18).
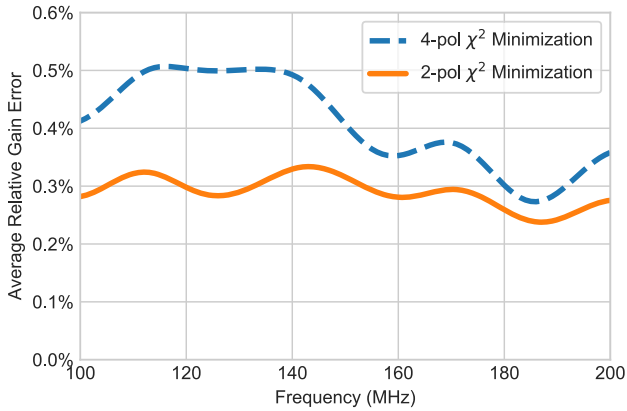
Figure 3. Antenna polarization leakage due to the ~ 1 per cent sensitivity of an antenna to a perpendicularly polarized electric field, known as D-terms, is a potentially important source of non-redundancy. Here, we see the resulting average relative gain errors (as defined in equation 19) from a simulation of a HERA-19-like array with smooth spectrum gains, visibilities, and D-terms, much like in Figs 1 and 2. We calibrate our simulation using only parallel-polarized visibilities (2-pol) and with both parallel- and cross-polarized visibilities (4-pol). Both calibrations produce gain errors because antenna-to-antenna variation in D-terms breaks the assumption of redundancy that underlies equation (14). However, because cross-polarized visibilities generally have lower magnitudes than parallel-polarized, D-term leakage biases cross-polarized visibilities more than parallel-polarized visibilities. This makes them less redundant, leading to larger errors in the 4-pol χ^2 minimization approach.

where, for example, $D_j^{n \to e}$ represents leakage from intrinsic n-polarized electric field to e-oriented antenna feed on the ith antenna. D-terms are generally small (~ 1 per cent), but their effect on cross-polarized visibilities can be important. They provide a mechanism by which parallel-polarized visibilities with ~ 10 times larger magnitudes can significantly affect the cross-polarized antennas.

While the presence of D-terms makes polarized mapmaking more challenging, it is not inherently a problem for redundant calibration. If the D-terms were all identical from antenna to antenna, we could simply redefine our 'true' visibilities to be the right-hand side of equation (18) and solve for model visibilities that are actually some unknown linear combination of all four visibility polarizations. In practice, that is not the case.

To explore the effect of D-terms that vary from antenna to antenna, we again simulate a noise-free HERA-like 19-element array with relatively smooth foregrounds, gains, and D-terms. Gains are set to have an average value of 1.0 with random, antenna-varying but spectrally smooth real and imaginary components at the 15 per cent level. Likewise, D-terms have an average of zero but have a random, smooth component that varies from antenna to antenna the at the 1 per cent level. Our cross-polarized visibilities have lower signal amplitudes by ~ 10 than our parallel polarized visibilities. To compare the results from 2-pol and 4-pol χ^2 minimization, we first take the gains, g_i^a coming out of both algorithms and fix all eight 2-pol degeneracies. Next we replace those degeneracies with the true simulated gains to produce $\bar{g}_j^a$ and compute gain errors for each antenna as a function of frequency. Due to our degeneracy-fixing and replacement, the errors are only in the non-degenerate subspace of 2-pol redundant-baseline calibration and are due solely to D-terms, not noise. Finally, in Fig. 3, we plot the average of those

errors defined as

$$\varepsilon = \frac{\sum_{a \in e,n} \sum_j \left| \bar{g}_j^a - g_j^{a,\text{true}} \right|}{\sum_{a \in e,n} \sum_j g_j^{a,\text{true}}}. \tag{19}$$

Both methods have identical degeneracy-fixing, substituting the degenerate subspace of the true gain solutions into the eight degenerate degrees of freedom, in order to make the errors directly comparable. Without D-terms, both techniques would produce the exact correct answer. As expected, introducing D-terms of the order of ~ 1 per cent produces comparable errors in the gain estimates, generally less than 0.5 per cent, independent of array size. The simple redundant calibration model from equation (14) that we are optimizing cannot possibly perfectly reproduce the observed visibilities from equation (18) when the D-terms vary. The precise magnitude of the error should not be taken too literally since the average magnitude of D-terms depends on antenna design choices, like using a dish or a dipole, and on the precision of element construction.

Interestingly, the effect is about 40 per cent worse when using 4-pol χ^2 minimization compared to 2-pol χ^2 minimization, even with the same degeneracy-fixing method. This makes sense: D-terms leak more signal from parallel-polarized visibilities into cross-polarized visibilities than vice versa precisely because there is more signal to leak. This means that D-terms make cross-polarized visibilities less redundant than their parallel-polarized counterparts and thus the assumptions underlying 4-pol redundant-baseline calibration weaker. Thankfully the errors do not introduce additional spectral structure beyond what was already in the simulation, though 4-pol calibration clearly magnifies the error associated with ignoring D-terms compared to the 2-pol approach.

It may be possible to simultaneously calibrate D-terms, gains, and visibilities through a generalized redundant-baseline calibration formalism based on equation (18). It is not clear if D-terms are the leading-order deviation from non-redundancy and therefore the best way to extend the degrees of freedom in the calibration model. It is also likely that such a generalization will introduce additional degeneracies which will need to fix in a frequency-smooth way. We leave those questions for future work.

4 DISCUSSION OF ANALYSIS CHOICES FOR REDUNDANT-BASELINE CALIBRATION

In this paper, we explored some of the subtleties of calibrating antenna gains and phases by taking advantage of repeated, supposedly identical measurements. Using redundant baselines, one can solve for most – but not all – of the calibration degrees of freedom. These remaining unsolved quantities, the *degeneracies* of redundant-baseline calibration, must be resolved using information about the sky. Redundant-baseline calibration makes the problem of subsequent absolute calibration much easier. Instead of calibrating one complex number per polarization per antenna per frequency, one only needs to calibrate six to eight numbers per frequency for the whole array, depending on one's analysis approach. Good calibration is key to preserving the separation in Fourier space of the 21 cm signal and the spectrally smooth foregrounds that are $\sim 10^5$ times brighter. Small errors can make a big difference.

Our strategy for dealing with the redundant calibration degeneracies focuses on simplifying the later absolute calibration that will ultimately resolve them. In principle, any solution that redundant calibration produces in the degenerate subspace can be removed and recalibrated. In practice, we will likely want to restrict the degrees of freedom that go into absolute calibration to avoid adding spectral structure on scales we have an a priori reason to believe

that the instrument's spectral response is smooth. Enabling these sorts of restrictions on absolute calibration means that we must be very careful to avoid adding extra spectral structure in the degeneracies. We saw in Section 2.4 how large delay errors could cause exactly such a problem and showed the accuracy necessary to avoid introducing spectral structure.

We also explored the generalization of redundant-baseline calibration to observations with both orthogonal antenna polarizations (and thus four visibility polarizations). We derived the new degeneracies in Section 3.1, showing how the expansion to two antenna polarizations doubles the number of degeneracies to eight per frequency, and how the inclusion of cross-polarized visibilities breaks two of those eight. However, because cross-polarized visibilities generally have much lower SNR than the parallel-polarized visibilities, that degeneracy breaking is noisy and risks adding spectral structure that will integrate down more slowly that the rest of the thermal noise (see Fig. 2).

This problem can be solved in two ways. Either we simply ignore the cross-polarized visibilities and calibrate the instrumental polarizations completely separately, the 2-pol approach, or we calibrate all four visibilities together (4-pol χ^2 minimization) but ignore the two weakly broken degeneracies (2-pol degeneracy-fixing). The former approach throws away information, raising noise slightly, and it can lead to less redundant calibrated cross-polarized visibilities. The later approach also solves the pressing problem of high noise across the entire band and produced slightly less noisy calibration solutions, but it introduces extra vulnerability to the antenna-to-antenna variation of polarization leakage D-terms. These non-redundancies break the assumptions of redundant-baseline calibration asymmetrically, biasing the cross-polarized visibilities more strongly than the parallel-polarized visibilities.

For cosmological measurements, the 2-pol approach seems to be most conservative approach for now. For polarized foreground studies, the case may be different. The calibration simulations in this work are physically motivated, but are still somewhat simplistic. Further simulation with realistic beams is needed to quantify the difference between the 2-pol and 4-pol approaches. It may be that the right answer is not the same for every instrument. The 2-pol approach also happens to be the cheapest, computationally.

That said, much work remains to make redundant-calibration robust to deviations from redundancy due to position errors, beam variations, D-terms, etc. Perhaps the question of polarization will be worth revisiting in the context of techniques for incorporating or mitigating non-redundancy (e.g. Liu et al. 2010 or Sievers 2017). As our instrumental knowledge and the sophistication of our calibration methods improves, we may be able to take advantage of the extra information and degeneracy breaking power of the cross-polarized visibilities that we lose by conservatively excluding them from redundant-baseline calibration.

REFERENCES

Ali Z. S. et al., 2015, ApJ, 809, 61
Asad K. M. B. et al., 2015, MNRAS, 451, 3709
Asad K. M. B., Koopmans L. V. E., Jelić V., de Bruyn A. G., Pandey V. N., Gehlot B. K., 2018, MNRAS, **476,** 3051
Bandura K. et al., 2014, in Stepp L. M., Gilmozzi R., Hall H. J., eds, SPIE Conf. Ser. Vol. 9145, Ground-based and Airborne Telescopes V. SPIE, Bellingham, p. 22
Barry N., Hazelton B., Sullivan I., Morales M. F., Pober J. C., 2016, MNRAS, 461, 3135
Beardsley A. P. et al., 2016, ApJ, 833, 102
Bernardi G. et al., 2009, A&A, 500, 965
Bonaldi A., Brown M. L., 2015, MNRAS, 447, 1973
Chang T., Pen U., Peterson J. B., McDonald P., 2008, Phys. Rev. Lett.., 100, 091303
Chapman E. et al., 2013, MNRAS, 429, 165
Datta A., Bowman J. D., Carilli C. L., 2010, ApJ, 724, 526
DeBoer D. R. et al., 2017, PASP, 129, 045001
Dillon J. S., Parsons A. R., 2016, ApJ, 826, 181
Dillon J. S., Liu A., Tegmark M., 2013, Phys. Rev. D, 87, 043005
Dillon J. S. et al., 2014, Phys. Rev. D, 89, 023002
Dillon J. S. et al., 2015a, Phys. Rev. D, 91, 023002
Dillon J. S. et al., 2015b, Phys. Rev. D, 91, 123011
Ewall-Wice A. et al., 2016, MNRAS, 460, 4320
Ewall-Wice A., Dillon J. S., Liu A., Hewitt J., 2017, MNRAS, 470, 1849
Furlanetto S. R., Oh S. P., Briggs F. H., 2006, Phys. Rep., 433, 181
Ghosh A., Prasad J., Bharadwaj S., Ali S. S., Chengalur J. N., 2012, MNRAS, 426, 3295
Hamaker J. P., Bregman J. D., Sault R. J., 1996, A&AS, 117, 137
Hazelton B. J., Morales M. F., Sullivan I. S., 2013, ApJ, 770, 156
Jelić V. et al., 2008, MNRAS, 389, 1319
Jelić V., Zaroubi S., Labropoulos P., Bernardi G., de Bruyn A. G., Koopmans L. V. E., 2010, MNRAS, 409, 1647
Jones R. C., 1941, J. Opt. Soc. Am. (1917–1983), 31, 488
Kern N. S., Carilli C. L., Kohn S. A., Parsons A. R., Dillon J. S., Ali Z. S., 2017, HERA Memo #30: Redundant Calibration Degeneracies with Four Polarizations
Kohn S. A., Aguirre J. E., 2015, HERA Memo #6: Crosstalk Measurements from PAPER-64
Kohn S. A. et al., 2016, ApJ, 823, 88
Koopmans L. et al., 2015, in Bourke T. L. et al., eds, Proc. Sci., The Cosmic Dawn and Epoch of Reionisation with SKA. SISSA, Trieste, PoS(AASKA14)001
Lannes A., Prieur J.-L., 2014, Astron. Nachr., 335, 198
Lenc E. et al., 2016, ApJ, 830, 38
Liu A., Tegmark M., 2011, Phys. Rev. D, 83, 103006
Liu A., Tegmark M., Morrison S., Lutomirski A., Zaldarriaga M., 2010, MNRAS, 408, 1029
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
Loeb A., Furlanetto S. R., 2013, The First Galaxies In The Universe. Princeton Univ. Press, Princeton, NJ

Mao Y., Tegmark M., McQuinn M., Zaldarriaga M., Zahn O., 2008, Phys. Rev. D, 78, 023529

Masui K. W. et al., 2013, ApJ, 763, L20

McQuinn M., Zahn O., Zaldarriaga M., Hernquist L., Furlanetto S. R., 2006, ApJ, 653, 815

Mishra A., Hirata C. M., 2017, preprint (arXiv:1707.03514)

Moore D. F., Aguirre J. E., Parsons A. R., Jacobs D. C., Pober J. C., 2013, ApJ, 769, 154

Moore D. F. et al., 2017, ApJ, 836, 154

Morales M. F., Wyithe J. S. B., 2010, ARA&A, 48, 127

Morales M. F., Hazelton B., Sullivan I., Beardsley A., 2012, ApJ, 752, 137

Newburgh L. B. et al., 2016, in Hall H. J., Gilmozzi R., Marshall H. K., eds, Proc. SPIE Conf. Ser. Vol. 9906, Ground-based and Airborne Telescopes VI. SPIE, Bellingham, p. 99065X

Noorishad P., Wijnholds S. J., van Ardenne A., van der Hulst J. M., 2012, A&A, 545, A108

Nunhokee C. D. et al., 2017, ApJ, 848, 47

Paciga G. et al., 2013, MNRAS, 433, 639

Parsons A., Pober J., McQuinn M., Jacobs D., Aguirre J., 2012a, ApJ, 753, 81

Parsons A. R., Pober J. C., Aguirre J. E., Carilli C. L., Jacobs D. C., Moore D. F., 2012b, ApJ, 756, 165

Parsons A. R. et al., 2014, ApJ, 788, 106

Parsons A. R., Liu A., Ali Z. S., Cheng C., 2016, ApJ, 820, 51

Patil A. H. et al., 2017, ApJ, 838, 65

Pearson T. J., Readhead A. C. S., 1984, ARA&A, 22, 97

Pober J. C. et al., 2013, ApJ, 768, L36

Pober J. C. et al., 2014, ApJ, 782, 66

Pritchard J. R., Loeb A., 2012, Rep. Prog. Phys., 75, 086901

Ram Marthi V., Chengalur J., 2013, preprint (arXiv:1310.1449)

Santos M. G., Cooray A., Knox L., 2005, ApJ, 625, 575

Sault R. J., Hamaker J. P., Bregman J. D., 1996, A&AS, 117, 149

Shaw J. R., Sigurdson K., Pen U.-L., Stebbins A., Sitwell M., 2014, ApJ, 781, 57

Shaw J. R., Sigurdson K., Sitwell M., Stebbins A., Pen U.-L., 2015, Phys. Rev. D, 91, 083514

Sievers J. L., 2017, preprint (arXiv:1701.01860)

Thompson A. R., Moran J. M., Swenson G. W., Jr, 2017, Interferometry and Synthesis in Radio Astronomy, 3rd edn. Springer International Publishing, Switzerland

Thyagarajan N. et al., 2013, ApJ, 776, 6

Trott C. M. et al., 2016, ApJ, 818, 139

Vedantham H. K., Koopmans L. V. E., 2015, MNRAS, 453, 925

Vedantham H., Udaya Shankar N., Subrahmanyan R., 2012, ApJ, 745, 176

Wieringa M. H., 1992, Exp. Astron., 2, 203

Wyithe J. S. B., Loeb A., Geil P. M., 2008, MNRAS, 383, 1195

Yatawatta S., 2012, Exp. Astron., 34, 89

Yatawatta S., 2016, preprint (arXiv:1605.09219)

Yatawatta S. et al., 2013, A&A, 550, A136

Zaroubi S., 2013, in Wiklind T., Mobasher B., Bromm V., eds, Astrophysics and Space Science Library, Vol. 396, The First Galaxies. Springer-Verlag, Berlin, p. 45

Zheng H. et al., 2014, MNRAS, 445, 1084

Zheng H. et al., 2017, MNRAS, 465, 2901

This paper has been typeset from a TEX/LATEX file prepared by the author.