The Bias and Uncertainty of Redundant and Sky-based Calibration Under Realistic Sky and Telescope Conditions

Ronny C. Joseph, Cathryn M. Trott, and Randall B. Wayth

Abstract

The advent of a new generation of low-frequency interferometers has opened a direct window into the Epoch of Reionization (EoR). However, key to a detection of the faint 21 cm signal, and reaching the sensitivity limits of these arrays, is a detailed understanding of the instruments and their calibration. In this work, we use simulations to investigate the bias and uncertainty of redundancy-based calibration. Specifically, we study the influence of the flux distribution of the radio sky and the impact of antenna position offsets on the complex calibration solutions. We find that the position offsets introduce a bias into phase component of the calibration solutions. This phase bias increases with the distance between bright radio sources and the pointing center, and with the flux density of these sources. This is potentially problematic for redundant calibration on Murchison Widefield Array (MWA) observations of EoR fields 1 and 2. EoR field 0, however, lacks such sources. We also compared the simulations with theoretical estimates for the bias and uncertainty in sky-model-based calibration on incomplete sky models for the redundant antenna tiles in the MWA. Our results indicate that redundant calibration outperforms sky-based calibration due to the high positional precision of the MWA antenna tiles.

Key words: dark ages, reionization, first stars – instrumentation: interferometers – methods: numerical – methods: statistical – techniques: interferometric

1. Introduction

Over the past few years, the latest generation of low-frequency interferometers has pushed up the upper limits of the 21 cm power spectrum of the Epoch of Reionization (EoR) (Beardsley et al. 2016; Patil et al. 2017). However, none of the current instruments, e.g., the Murchison Widefield Array (MWA; Tingay et al. 2013), the LOw Frequency ARray (LOFAR; van Haarlem et al. 2013), and the Precision Array for Probing the Epoch of Reionization (PAPER; Parsons et al. 2010), have detected a signal thus far. The signal, emitted by neutral hydrogen during the EoR, is a direct probe into the state of the Intergalactic Medium (IGM; Furlanetto et al. 2006) and allows us to directly study the conditions under which the first luminous objects were formed. For more in depth reviews, see Morales & Wyithe (2010), Pritchard & Loeb (2012), McQuinn (2016), and Furlanetto (2016).

Foreground sources and instrumental effects pose large challenges to the detection of this faint signal. The low-frequency foregrounds, e.g., the Milky Way and extragalactic sources, are expected to be 4–5 orders of magnitude stronger than the neutral hydrogen signal (Bowman et al. 2006; Furlanetto et al. 2006; Morales et al. 2006; Jelic et al. 2008; Pritchard & Loeb 2008). The removal of these foregrounds requires a detailed understanding of the instrument and its calibration, because the subtraction of these foregrounds, in particular bright compact sources, is sensitive to calibration errors (Datta et al. 2009). Failing to remove bright sources accurately leads to contamination of EoR data, causing the so-called “wedge” feature in the 2D-Power Spectrum. (Datta et al. 2010; Morales et al. 2012; Trott et al. 2012; Vedantham et al. 2012) This leakage of bright source residuals into the power spectrum, makes certain scales of the EoR signal inaccessible if not dealt with correctly.

Adequate removal of these foregrounds and extraction of the faint signal from the data puts stringent requirements on our calibration accuracy and precision. Standard calibration schemes correct the sky signal cross-correlations (or “visibilities”) measured by radio interferometers using sky models, hereby solving for the gain factors that cause the discrepancy between the modeled visibilities and the measured visibilities (see Rau et al. 2009, for a review). Sky-model-based calibration has undergone tremendous progress in the past years in order to overcome direction dependent calibration effects, e.g., varying antenna primary beam shapes, and ionospheric distortions, that limit this new generation of instruments. This progress resulted in a large variety of improved sky based calibration implementations, e.g., RTS (Mitchell et al. 2008), SAGEcal (Yatawatta et al. 2009; Kazemi et al. 2011), SPAM (Intema et al. 2009), FHD (Sullivan et al. 2012), and facet calibration (van Weeren et al. 2016) to name a few. Nevertheless, at the operating frequency of these new low-frequency interferometers (80–200 MHz), our limited understanding of the sky leads to incomplete models. Model-based calibration with incomplete sky models causes calibration errors that lead to image artifacts, which in turn limit the dynamic range of observations (Grobler et al. 2014; Wijnholds et al. 2016) and, more relevant to EoR science, contaminate the power spectrum (Barry et al. 2016; Ewail-Wice et al. 2016; Trott & Wayth 2017). Redundant calibration, however, allows us to escape our ignorance of the low-frequency sky because it does not require modeling (Wieringa 1992). Because of this reason redundant calibration is undergoing a renaissance, resulting in further studies by Noorishad et al. (2012), Liu et al. (2010), Ali et al. (2015), Dillon & Parsons (2016), showing the applicability and some limitations of redundant calibration in low-frequency radio telescopes. More recently, redundant calibration was compared to sky model based calibration by Li et al. (2018). Despite the inability of redundant calibration to solve for direction dependent effects, it
still remains an interesting alternative to calibrate a radio telescope to first order, where sky based calibration can resolve higher-order effects.

In this paper, we will study the theoretical performance of redundant calibration. We specifically look at how redundant calibration depends on the flux distribution of the sky and positional errors of the antennas. This allows us to determine which regions of the sky should be calibrated with sky-based calibration or redundant calibration, to yield the most accurate and precise result for a given antenna position precision of the array. We do this by running simulations of redundant calibration in which we calibrate a redundant array with ideal antenna responses on a realistic multi-source sky, while changing the flux and position of a single calibrator source. We compare the distribution of solutions we obtain from these simulations with a theoretical estimate of the sky model calibration bias and an uncertainty due to an incomplete calibration model.

The structure of the paper is as follows: Section 2 discusses sky model calibration and our analytic description of the impact of an incomplete sky model on the bias and uncertainty of the calibration solutions. Section 3 reviews redundant calibration, and describes the set up of the redundant calibration simulations. We discuss the influence of the sky flux distribution on redundant calibration solutions and the impact of position offsets, using a simple 5-element interferometer to demonstrate the fundamental issues of redundant calibration. We conclude our results with a comparison between the bias and uncertainty of redundant calibration, and sky model based calibration for the redundant MWA tiles in Section 4, and we discuss the implications for the MWA in Section 5.

2. Sky Model Calibration

In this section, we describe and derive the impact of an incomplete sky model on the calibration solutions in a sky model based approach. Earlier works studied the effect of calibration on incomplete sky models via analysis and simulations. Salvini & Wijnholds (2014) discuss the statistical performance, Barry et al. (2016) study the impact on EoR power spectrum estimation, and Grobler et al. (2014), Wijnholds et al. (2016) study its impact on the deconvolution of a 2-point source sky. In this work, we compare redundant calibration with theoretical estimates for the bias and uncertainty introduced by calibration on an incomplete sky model.

We can write the measured correlation for a pair of antennas $i$ and $j$ in the absence of noise $C_{ij}$ as a product of the antenna gain factors $g_i^*$ and $g_j$, and the true visibility $V_{ij}$

$$C_{ij} = g_i^* g_j V_{ij},$$

the superscript “$^*$” indicates complex conjugation. In sky model based calibration we solve for the gains $g$ by minimizing the difference between our modeled visibilities $M$ and the measured correlations $C$.

$$\min_g \| C - g g^\times M \|.$$  

Here, we write the minimization in the most general form, without explicitly choosing a matrix or vector notation for discussion purposes, as we will switch between those later on. The caveat of this approach is that the signals from unmodeled sources are absorbed into the calibration solutions. To understand how this impacts the solutions, we first derive the uncertainty of sky based calibration solutions due to a stochastic sky of point sources and thermal noise. We then use this result to derive the bias due to model incompleteness.

2.1. Model Incompleteness Uncertainty

To derive the minimum uncertainty on the estimated complex gain solutions $g$ we use the Cramér-Rao Lower Bound (CRLB) on the estimated gain parameters. Throughout this derivation, we assume the model used for calibration is a single point source with flux density $S(\nu)$ located at some location $l$ in the sky. The model visibility for a given baseline $u$ at frequency $\nu$ is then given by:

$$M(u, \nu) = S(\nu)A(l)\exp[-2\pi i u \cdot l],$$  

$A(l)$ is the antenna beam response, which we choose to be a Gaussian. We choose an unmodeled source background described by a broken power-law source count distribution $dN/dS$ (Gervasi et al. 2008; Intema et al. 2011; Franzen et al. 2016; Williams et al. 2016)

$$dN = \begin{cases} k_1 S^{-\gamma_1} & \text{if } S_{\text{low}} \leq S \leq S_{\text{mid}} \\ k_2 S^{-\gamma_2} & \text{if } S_{\text{mid}} \leq S \leq S_{\text{high}} \end{cases}$$

where $dN/dS$ gives the number of sources per Jy per steradian, and $S$ is the source flux in Jy. Throughout this paper, we will use $k_1 = 4100$, $\gamma_1 = 1.59$, $\gamma_2 = 2.5$, $S_{\text{low}} = 400$ mJy, $S_{\text{mid}} = 1$ Jy, and $S_{\text{high}} = 5$ Jy. To derive the CRLB on the estimated gain parameters we first compute the Fisher Information Matrix (FIM) $\mathcal{I}$ (Kay 1993). This takes the following form for a complex multivariate normal distribution with mean $M$ and gain independent data covariance $\Sigma_{\text{data}}$:

$$\mathcal{I}_{i,j} = 2 \Re \left( \frac{\partial M^H}{\partial g_i} \Sigma^{-1} \frac{\partial M}{\partial g_j} \right).$$

where the superscript “$^H$” denotes the Hermitian transpose, the superscript “$^{-1}$” denotes the matrix inverse, $M$ is a vector where each entry is the model visibility of a baseline pair $ij$ for a single frequency channel, and $\Sigma_{\text{data}}$ is the data covariance matrix. The covariance of the data is the sum of thermal noise variance $\Sigma_{\text{thermal}}$ and the variance of our stochastic background sky $\Sigma_{\text{sky}}$, as we assume the thermal noise is baseline independent and we ignore the compact Fourier beam kernel that creates correlations between closely spaced baselines. We describe the thermal noise as

$$\Sigma_{\text{thermal}} = \sqrt{\frac{\text{SEFD}}{B \nu t}},$$

where $B$ is the bandwidth of a single frequency channel, and $t$ is the integration time of the observation. Throughout this paper, we adopt the MWA EoR parameters unless stated otherwise, SEFD = $10^4$ Jy, $B = 40$ kHz, $t = 120$ s. For these parameters, the thermal noise is $\Sigma_{\text{thermal}} \sim 9$ Jy. We take the expression for the visibility variance for a baseline in a single frequency channel due to a stochastic sky $\Sigma_{\text{sky}}$, from Trott et al. (2016),
wherein we assume a flat spectral index of our sources within a single frequency channel, and \( \sigma \) is the frequency dependent beamwidth. Throughout this paper we assume a beamwidth of \( \sigma = 30^\circ \), similar to the MWA beam at 150 MHz, resulting in a sky variance of \( \Sigma_{\text{sky}} \sim 2.5 \times 10^3 \text{Jy} \). Because the noise variance \( \Sigma_{\text{thermal}} \) and sky variance \( \Sigma_{\text{sky}} \) are baseline independent, the total data covariance matrix \( \Sigma_{\text{data}} \) is a diagonal matrix. We can therefore rewrite the FIM elements as

\[
I_{i,j} = 2 \Re \left( \sum_n \frac{1}{\Sigma_{\text{data}}} \frac{\partial M_n^*}{\partial g_i} \frac{\partial M_n}{\partial g_j} \right)
\]

(8)

where we sum over the data index \( n \). For the CRLB we are only interested in the variance on a gain parameter \( \Sigma_{\text{g}} \), i.e., we only compute \( I_{m,n}^{-1} \) for \( m = n \), which reduces to

\[
\Sigma_{\text{g}} = \frac{\Sigma_{\text{thermal}} + \Sigma_{\text{sky}}}{2 |S(\nu)A(\nu)|^2 (N-1)}
\]

(9)

where \( N-1 \) is the number of baselines formed by an antenna in the array. We note that the variance scales inversely with the number of antennas in the array, and beam-weighted apparent flux density of the modeled source squared. We will use this expression to compare the uncertainty of redundant calibration with sky model based calibration.

### 2.2. Model Incompleteness Bias

To derive an expression for the bias, i.e., the mean deviation from the true solutions introduced by the model incompleteness, we follow Wijnholds et al. (2016) and reformulate Equation (2) explicitly in terms of visibility matrices and gain vectors

\[
\min_{\Delta g} \| C - gMg^H \|.
\]

(10)

\( C \) and \( M \) are matrices containing the measured and modeled visibilities, e.g., \( C_{ij} \) is the measured visibility between antenna \( i \) and \( j \), and the vector \( g \) contains the complex antenna gains. We ignore the auto-correlations, therefore, the diagonals of \( M \) and \( C \) are zero, and if we ignore the noise we can write the measurements \( C \) in terms of the modeled \( M \) and unmodeled \( U \) sky visibilities. We can also write the gain vector \( g \) as a sum of the true gains \( g_t \) and a deviation introduced by the calibration process \( \Delta g \).

\[
C = g_t(M + U)g_t^H
\]

\[
g = g_t + \Delta g.
\]

Furthermore, we can use the Hadamard product \( \odot \), i.e., the element-wise product, to rewrite Equation (10) into

\[
\min_{\Delta g} \| g_t g_t^H \odot U - (g_t \Delta g^H + \Delta gg_t^H) \odot M \|
\]

(12)

where we have dropped all higher-order terms of \( \Delta g \). Wijnholds et al. (2016) derive an approximate closed form solution for \( \Delta g \) by rewriting Equation (12) into a least squares form. We will take the solution as the conclusion of this short review, and point the interested reader to their work for the detailed derivation. The closed form solution takes the following form

\[
\begin{bmatrix} \Delta g^* \\ \Delta g \end{bmatrix} \approx \begin{bmatrix} A & B \\ C & D \end{bmatrix}^{-1} \begin{bmatrix} E_g^* \\ F_g \end{bmatrix}
\]

(13)

The block matrices are given by

\[
A = M^*G_t G_t^H M \odot I \\
B = M^*G_t \odot G_t M \\
C = B^* D = A^* \\
E = M^*G_t G_t^H U \odot I \\
F = G_t^H U^* \odot MG_t^H,
\]

(14)

where \( G_t = \text{diag}(g_t) \), and \( I \) is the identity matrix. Here, \( A \) and \( B \) encode the total modeled power summed over baselines, and the power in an individual baseline, respectively, whereas \( E \) and \( F \) are the equivalent expressions for the unmodeled power. Intuitively, these matrices describe the additional bias in the solutions from correlations between the model and the residual signal, and the overall power ratio of model to unmodeled sky. Minimising both of these bias terms is desirable for good sky-based calibration. We can use Equation (13) to derive the mean gain offset \( \langle \Delta g \rangle \) in the case that our sky model consists of a single point source in the presence of a more complicated sky.

\[
\begin{bmatrix} \Delta g^* \\ \Delta g \end{bmatrix} \approx \begin{bmatrix} A & B \\ C & D \end{bmatrix}^{-1} \begin{bmatrix} E_g^* \\ F_g \end{bmatrix}
\]

(15)

with

\[
\langle E \rangle = M^*G_t G_t^H \langle U \rangle^* \odot I \\
\langle F \rangle = G_t^H \langle U \rangle^* \odot MG_t^H.
\]

(16)

We can parameterize the mean unmodeled visibility contribution of our stochastic sky \( \langle V_s \rangle \) using the sky visibility variance \( \Sigma_{\text{sky}} \). If we consider the Fourier transform of each point source as a phasor in the complex plane (Re, Im), we can consider a stochastic sky of point sources as a random walk through this plane (see Figure 1). Each point source contributes a new complex phasor to our total unmodeled sky phasor. The path length of this random walk, i.e., the total amplitude of our unmodeled visibility, is on average given by the root mean square of distribution from which the phasors are drawn. In our
analysis we assume this to be a Gaussian distribution, therefore, the unmodeled visibility amplitude equates to the variance.

Now, we have yet to explore (are still left with) the orientation of the unmodeled visibility component, since every net orientation has equal probability. For our calculations we choose the unmodeled visibility to have an angle of 45°, with respect to the model visibility as a measure for some average offset introduced into our visibility amplitude and phase angle. Even though the true phase angle of the unmodeled visibility is uniformly distributed, we find that this approximation yields comparable results to Monte Carlo simulations with a distribution of phases. We will use these results in Section 4 where we compare the results from our redundant calibration simulations with the theoretical performance of sky model based calibration.

3. Redundant Baseline Calibration

In this section, we provide a short review of Wieringa (1992) and Liu et al. (2010) to highlight the key features of redundant calibration. Considering the simplest redundant case, which is a 5-element linear array with equal spacings Δx between the antennas, we have four unique baselines. Four baselines at the shortest spacing, three at 2Δx, two at 3Δx and only one at 4Δx. Using only the first two sets of baselines, we can create a determined system of linear equations, solving for the two unknown visibilities for each set of redundant baselines and five unknown antenna gains. One way of doing this is the logcal algorithm, in which we take the logarithm of Equation (1), while noting that every quantity is a complex number |g| exp[iφ] with amplitude |g| and phase φ. This yields two equations,

\[
\ln|C_{ij}| = \ln|g_i| + \ln|g_j| + \ln|V_{ij}|
\]

\[
\arg[C_{ij}] = \phi_j - \phi_i + \arg[V_{ij}]
\]

where |g_i| is the gain amplitude and φ_i is the gain phase of antenna i, ln|V_{ij}| is the true visibility amplitude and arg[V_{ij}] is the true visibility phase measured by a baseline pair i and j. Because the amplitude and phase decouple, we can rewrite this into two different matrix equations that can be solved independently, see Equation (18).

\[ c_\alpha = A_\alpha x_\alpha, \]

where the index α = {η, φ}, η for the amplitude equations and φ for the phase equations, are the measured correlation amplitude and phase vectors, x_α contains the gain amplitude ln|g_α| and visibility amplitude ln|V_{ij}|, x_α contains the gain phase φ_α and visibility phase arg[V_{ij}]. A is the matrix that maps the gain and visibility into the measured correlations. Equation (19) shows this explicitly for the phase,

\[
\begin{pmatrix}
\arg[c_{12}] \\
\arg[c_{23}] \\
\arg[c_{34}] \\
\arg[c_{45}] \\
\arg[c_{51}] \\
\arg[c_{24}] \\
\arg[c_{35}] \\
\end{pmatrix}
= 
\begin{pmatrix}
-1 & 0 & 0 & 0 & 0 & 1 & 0 \\
0 & -1 & 0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & -1 & 1 & 0 & 1 \\
0 & 0 & 0 & 0 & 0 & 1 & 0 \\
-1 & 0 & 1 & 0 & 0 & 0 & 0 \\
0 & -1 & 0 & 1 & 0 & 0 & 1 \\
0 & 0 & 0 & 1 & 0 & 1 & 0 \\
\end{pmatrix}
\begin{pmatrix}
\phi_1 \\
\phi_2 \\
\phi_3 \\
\phi_4 \\
\phi_5 \\
\phi_6 \\
\phi_7 \\
\end{pmatrix}
\]

arg[v_α] is the phase of the visibility measured by the Δx spacings and for the 2Δx spacings we have arg[v_α]. This system is, however, degenerate and needs to be constrained by setting a reference antenna for which the amplitude gain and phase gain are specified. In the specific case of phase calibration we need two additional constraints because a tilt in the array is equivalent to a rotation of the sky,

\[
0 = \ln|g_l|, \quad 0 = \phi_i
\]

Given the unmodeled visibility to have an angle of 45°, we determine system of linear equations, solving for the two unknown antenna gains. One way of doing this is the logcal algorithm, in which we take a Taylor expansion around solution guesses g_0 and v_0 of the true solutions g_i and v_i. This yields one single equation in which we solve for the gains in their complex forms,

\[ c_{ij} = g_i^0 g_j^0 v_i^0 v_j^0 \Delta g_i + g_i^0 v_i^0 \Delta v_j + g_i^0 g_j^0 \Delta v_j. \]

In lincal, we solve for the differences between true solutions and the guesses Δg_i and Δv_j. Allowing us to iteratively correct our guesses. Similarly to logcal we can rewrite this into a matrix equation containing the real and imaginary components of the gains and visibilities, for the details see the appendix of Li et al. (2018). Current implementations of redundant calibration use logcal to find an initial estimate and further refine the solutions with lincal (Zheng et al. 2014). In this work, we will do the same.

3.1. Simulating the Bias and Uncertainty

To estimate the bias and uncertainty of redundant calibration we simulate the calibration of the antennas in a (nearly) redundant array. In our simulations, we define a group of redundant baselines when they lie within 1/6A of each other in the υ-φ plane. This is well within the linear regime of a sinusoidal centered at exp[2πυD]. However, we will show that deviations from non-redundancy within this threshold impact the calibration accuracy and precision. We also assume a Gaussian beam, similar to our sky model derivation, and assume the beams are identical for each antenna. This is not strictly true for phased arrays (Wijnholds et al. 2010), however, it suffices as a first-order approximation.

We generate a background of radio sources with a flux distribution according to Equation (4), and uniformly distribute them over the sky. Finally, we add a source with arbitrary flux and location, similar to the calibrator source in Section 2. These sources are gridded onto an (l, m)-grid, and Fourier transformed to generate visibilities using powerbox (Murray 2018), a tool written to simulate EoR data sets and forward-model them to power spectra. We interpolate the visibilities to produce the measurements for each baseline. Finally, we assume Gaussian-distributed noise in the real and imaginary components with a variance according to Equation (6). These visibilities are then passed to our redundant calibration module, which is a direct
implementation of the algorithm described in Section 3. The code is publicly available (Joseph 2018).

3.2. The Sky Dependent Uncertainty

We first study the influence of the sky, and show how it affects the uncertainty of the estimated gain solutions. We start out with a simple sky model of a statistical background sky while moving a high flux density source with respect to phase center and trying to calibrate on each realization of the sky. This allows us to study the performance of redundant calibration in drift-scan mode, and simultaneously study the performance of redundant calibration in the MWA EoR fields, which depending on the field have strong in-beam sources. Figure 2 shows the dependence of the calibration solutions for an ideal interferometer, i.e., perfect gains $g = 1$ and perfect redundancy, as a function of source position in terms of the direction cosine $l$, the native interferometry sky coordinate. We show the results for a pure logcal-calibration, similar to Wieringa (1992), and for a logcal + lincal-calibration.

The results show that overall the solution variance for both the gain amplitude solutions behave better when the strong source is near the center of the beam, because the signal-to-noise ratio ($S/N$) is higher at the pointing center. We do note that our implementation of the lincal algorithm seems slightly biased in the presence of noise, the mean of the solution is $1\%$ below the true value. We filtered out $<1\%$ of the solution realizations due to bad convergence, i.e., solutions with unrealistically high gain amplitudes.

The gain phase solutions show a similar dependence with some additional structure in the variance due to a problem which is inherent to logcal: phase wrapping. The logcal implementation can only determine phases between $-\pi \leq \phi \leq \pi$, in which the arctan is defined. When a certain redundant set of baselines measures a visibility phase of $[\phi_i] = \pi$, due to the location of the dominant source on the sky, the solutions become very sensitive to noise. The visibility phase starts to “jump” between $-\pi$ and $\pi$ causing large variances in the phase calibration solutions. These phase wrapping points can be determined by solving $2\pi ul = \pi n$ for odd numbers of $n$, i.e., solving for the source coordinate $l$ when a given baseline with length $n$ measures a phase of $\pi$.

We can understand the effect of phase wrapping by adding a noise vector $n_{\phi}$ to the measurement equation, see Equation (22).

$$c_{\phi} = A_{\phi} x_{\phi} + n_{\phi}.$$  \hspace{1cm} (22)

The discrete set of solutions peaks is caused by phase wrapping in the set of three long baselines. Each peak represents a combination of phase wrapping baselines as marked above the peak.

The phase noise $n_{\phi} \propto N/S$ (Liu et al. 2010), however, when the measured visibility phase approaches $\pi$, this noise vector diverges $[n_{\phi}] \rightarrow 2\pi$. Because the noise of a single baseline is mixed into all solutions when estimating $\hat{x}$ we get large offsets in the calibration solutions:

$$\hat{x} = x_{\phi} + [A^T A]^{-1} A^T n_{\phi}.$$  \hspace{1cm} (23)

Looking at the logcal + lincal solutions in Figure 2 we see similar behavior in the phase solutions and amplitude solutions. We also note that our implementation of the algorithm seems to be very sensitive to noise. This results in large variations in the amplitude solutions when the dominant source moves away from phase center. Figure 2 shows the mean and the variance of the solutions, even though the solutions do not strictly follow a Gaussian distribution. Figure 3 shows the distribution of logcal solutions for antenna 2 at the first phase wrapping point. The distribution has seven distinct peaks, each peak represents a combination of
phase wrapping baselines. The first phase wrap occurs in the set of three long redundant baselines, and therefore there are at maximum $\sum_{k=1}^{n} n!/(k!(n-k)!)$ different combinations and solutions peaks possible. The actual spectrum depends on the array geometry, which is encapsulated in the matrix $A$. These solutions are the starting point for the logcal algorithm, and our implementation of it is not able to recover the true solutions when given a bad starting point.

We attempted to circumvent the phase wrapping of a specific baseline in a single channel by extending the logcal algorithm to incorporate the measurements of neighboring frequency channels while assuming the gain solutions remain the same. However, due to the same mixing that takes place in the single frequency channel implementation, phase wrapping will still remain a problem unless a clever selection of frequencies is used to circumvent phase wrapping. We discuss this in the Appendix. We also note that current implementation do apply a pre-calibration step to unwrap the visibility phases by averaging over baselines within a redundant group (Zheng et al. 2014), or by using the products of visibilities to construct a system equations to solve for the phases (Li et al. 2018).

![Figure 4](image.png)

**Figure 4.** From left to right: the logcal amplitude gain and phase gain solutions, the logcal + lincal amplitude gain and phase gain solutions for a single antenna in an nearly redundant 5-element interferometer as a function of strong source position $i$. The dark line represents the mean of the solutions, the shaded area indicates the 1σ solutions variance. The amplitude solution variance inversely follows the shape of the beam, i.e., the variance increases when the beam response decreases. The mean of the phase solutions generally fluctuates around $\phi = 0$ and deviates along with the variance at so-called phase wrapping points, which are further explained in the text. The additional position offset, which causes a phase offset from the ideal redundant phase, is absorbed into the solutions causing a slope.

![Figure 5](image.png)

**Figure 5.** The magnitude of the offset residual $\hat{b}$ for an offset in a given antenna. The first column shows the residual when offsetting the first antenna, i.e., the reference antenna, the second column shows the result for the second antenna etc. We see that offsetting the 3rd antenna, e.g., the middle antenna, impacts only its calibration solution. Its solutions have the strongest constraints due to its baseline participation. It participated in all baselines groups and has a high participation number in each group.

### 3.3. The Position Offset Bias

In the previous section, we described the results for an ideal radio interferometer. However, in reality all antennas will have slight position offsets from their perfectly redundant positions. To understand the impact of positions offset, we simulate redundant calibration under the same conditions as before, but now we offsetting one antenna in the $x$-direction by $\delta x = 20$ cm. The results are shown in Figure 4.

We can clearly see that both the amplitude and phase solutions are affected by the position offset. We can understand the oscillatory behavior of the amplitude solutions by returning to the complex plane. Imagine the complex visibility of the main calibrator as measured by a baseline as a phasor in this plane, we can think of the total sum of background sources as a similar phasor. Each redundant baseline should measure the same amplitude of the sum of these phasors. However, due to the non-redundancy introduced by position offsets the non-redundant pairs measure a different amplitude, this difference propagates through to the solutions. As the primary source moves across the sky, its phasor will rotate in the complex plane, constructively and destructively interfering with the background visibility creating this oscillatory behavior. These oscillations are damped as the primary source becomes attenuated as it moves outside of the primary beam. The behavior of the mean phase solutions can be explained in a similar fashion, using the phase of the phasors rather than the amplitude.

For low-N arrays similar to this 5-element toy model this error propagates to all antennas solutions, due to the coupling of all gain solutions to the visibilities. However, when increasing the number of antennas in the array the coupling becomes weaker increases as the number of measurements increases.

Liu et al. (2010) discusses the inclusion of antenna offsets as a linear perturbation to perfect redundancy, and how these offsets can be solved for by expanding Equation (19), however, implementations of redundant calibration do not take this into account (Zheng et al. 2014). We can describe the effect of baseline offsets by adding another vector $\hat{b}$ to Equation (22) which contains the visibility offsets due to deviations from redundancy. Note that this additional vector can contain any type of offset, e.g., antenna beam variations.

$$c_\alpha = A_\alpha x_\alpha + n_\alpha + b_\alpha.$$  

(24)

Inverting this equation using the standard least square solution, and not taking into account this extra term, leaves us with an additional residual. We can calculate these residuals for different antenna offsets. Figure 5 shows the magnitude of the offset residuals $\hat{b}$ in the phases of the estimated gain and
visibilities when offsetting different antennas by the same amount. These results show that offsetting the antenna with the highest baseline participation does not propagate to all antenna solutions and leaves the visibilities unaffected. Offsetting the reference antenna, in this case antenna 1, has the strongest impact on the solutions of all other antennas. This implies that the choice of reference antenna is not as arbitrary as one might think.

4. Comparing Sky and Redundancy Based Calibration

Having varied several parameters within our redundant calibration simulations we can now move forward and apply this formalism to the MWA “hexes”. The hexes contain 72 antenna tiles arranged in two hexagons, see Figure 6. The shortest baseline, defining the hexagonal lattice, has a length of 14 meters. The hexagons are also placed to be redundant with each other, i.e., they have the same orientation. This provides extra sensitivity on scales relevant for the EoR experiment, and adds redundancy for calibration purposes. Due to a lack of redundant baselines connecting one hex to the reference antenna in the other hex, we either need to invoke another degeneracy parameter to encapsulate a phase offset between the two or calibrate them separately. Li et al. (2018) calibrate the hexes simultaneously, however, for simplicity and speed we calibrate a single hex in our simulations. A single hex forms 630 baselines, of which 601 are redundant, organized in 71 redundant groups.3

Figure 6 also shows the non-redundancy of each antenna within the hexes, the antennas are placed with an accuracy on the order of centimetres. This is an order of magnitude below the redundant calibration threshold of 1/6λ at 150 MHz, i.e., ~30 cm.

3 Theoretically we can also include non-redundant antennas in the calibration, as long as the number of unknowns is lower than the number of measurements. For each redundant hex in the MWA we can add six non-redundant tiles before the system becomes unsolvable.

In our final set of redundant calibration simulations we will offset each tile in a single hex according to Gaussian distributions with mean $\mu_x = 0$, and $\sigma_x = 10^{-3} - 10$ cm. Now we calibrate while moving a 100 Jy source across the sky, and run a separate set of simulations where we fix the location of the source at 3° off-zenith while changing its flux density instead. Figure 7 shows the bias and uncertainty we derive from these simulations. To compute these from the distribution of solutions we obtain we calculate the median offset from the true gain solutions for the bias, i.e., $g = 1$ and we take the standard deviation for the uncertainty. All results are averaged over all antennas.

To compare with traditional sky-based calibration, we use the bias and uncertainty derived in Section 2. The contour lines in Figure 7 are the ratios between either the bias or uncertainty of redundant calibration and sky model calibration. To make the comparison slightly easier, we take two cuts through the plots in Figure 7 at an antenna position precision of $\sigma_x = 0.02$ m and $\sigma_x = 0.10$ m, these cuts are shown in Figure 8. From Figures 7 and 8, we can conclude the following statements. The amplitude bias depends strongly on the flux density of the primary source and its location on the sky. Redundancy-based calibration has a lower amplitude bias when the sky is dominated by a single point source, and it quickly reaches the accuracy of our implementation of the algorithm as the source moves out of the field of view.

The amplitude uncertainties of redundancy-based and sky-based calibration are comparable. For both redundant and sky model based calibration, they decrease comparably as a function of primary source flux density. However, as a function of source location redundant calibration quickly reaches the noise floor when the source is beyond the FHWM of the primary beam.

Interestingly, the phase bias increases with primary source flux density and distance of the bright source to phase center. The bias reaches a maximum when the source is at the FWHM of the primary beam, if the source moves beyond this the bias decreases again. When the sky is dominated by a single source...
that is off-center, the bias of redundant calibration becomes comparable or larger to that of sky-based calibration. The flux at which the two become comparable is dependent on the magnitude of the position offsets in the array.

The phase uncertainty depends strongly on the source flux density and its location on the sky. When the primary source brightness becomes comparable to the background sky the uncertainties of redundant become larger than that of sky model based calibration.

Figure 7. From left to right: the bias and uncertainty of the amplitude solutions $|g|$, and the bias and uncertainty of the phase solutions $|\phi|$ of redundant calibration on an MWA hex-like array. Top row: we vary the position deviations $\sigma_x$ and source peak flux $S_{\text{peak}}$ of a source located $87^\circ$ above the horizon. Bottom row we vary the position deviations $\sigma_x$ and source elevation of a 100 Jy source. All results are averaged over all antennas in the hex. The contour lines are the ratios between the bias and uncertainty of redundancy-based and sky model based calibration, e.g., an uncertainty contour line of 2.0 indicates the uncertainty is twice as large for redundant calibration as compared to sky model based calibration.

Figure 8. From left to right: the bias and uncertainty of the amplitude solutions $|g|$, and the bias and uncertainty of the phase solutions $|\phi|$ of redundant calibration on an hex array with a position precision $\sigma_x \sim 2$ cm. Top row: we vary the position deviations $\sigma_x$ and source peak flux $S_{\text{peak}}$ of a source located $87^\circ$ above the horizon. Bottom row we vary the position deviations $\sigma_x$ and source elevation of a 100 Jy source. All results are averaged over all antennas in the hex. The blue lines are the bias and uncertainties for an array with positional offsets $\sigma_x = 0.02$ m, in orange the bias and uncertainties for an array with positional offsets $\sigma_x = 0.1$ m, and in green the theoretical estimates for the bias and uncertainty of sky based calibration.
based calibration. Also note that the behavior of the uncertainty as a function of source elevation changes for different positional precisions. When the primary source is between phase center and the FWHM of the primary the uncertainties of redundant calibration are larger for arrays with large positional offsets. Particularly when the brightest source is at the FWHM of the primary, the uncertainties become larger than that of sky model based calibration. This increase in uncertainty can be explained by phase wrapping as discussed earlier. As the sources moves the measured visibilities phases start wrapping around $2\pi$ creating a spectrum of solutions that widens the distribution.

5. Discussion

The most notable results from our simulations are the results for the phase bias. Redundant calibration was proposed as an alternative to a sky model based approach because it is agnostic of the sky and therefore it does not suffer from the systematics introduced by an incomplete sky model. However, this work shows that systematics arise in a different way, because we impose the condition that our telescope is perfectly redundant. This manifests itself in systematic phase offsets in our calibration solutions because redundant calibration absorbs antenna position offsets into the calibration solutions. These phase offsets become more prominent when there is a high flux density source away from the pointing center.

Barry et al. (2016), Ewall-Wice et al. (2016), Trott & Wayth (2017) show that calibration on incomplete sky models causes contamination in the EoR power spectrum. Similarly redundant calibration can introduce contamination. The relative position offset changes as a function of wavelength. Therefore, the measured phase offset will therefore also vary as a function of frequency that can introduce a contamination to the EoR power spectrum.

We demonstrated the influence of the sky flux distribution on the performance of redundant calibration. Figure 9 shows a map of the radio sky at 408 MHz (Haslam et al. 1982) with the MWA EoR target fields. We can clearly see that these fields are not devoid of high flux density sources. EoR field 1 contains Fornax A and Pictor A, and EoR field 2 contains Hydra A among others. However, the results of the redundant calibration simulations for a single MWA hex show that position offsets at the position precision levels of the MWA are not a large source of bias and uncertainty for the phase solutions. Redundant calibration even outperforms sky based calibration on a single source. However, if a redundant array, such as HERA, has positional offsets in the order of 10 cm, careful consideration has to made on when to do redundant calibration. As demonstrated the phase bias can go up to an order of magnitude higher than that of the MWA-like array under these conditions. Fortunately for HERA the primary beam is narrower than that of the MWA, the latter suffers from significant side lobes, this and its large number of redundant baselines makes HERA somewhat robust against positional offsets (Liu et al. 2010). However, the exact trade off is still unclear.

We do note we have simplified the sky model based approach for analytic tractability. In reality, a sky model will contain more than one calibration source; therefore, the bias and uncertainty for a sky based approach will certainly be lower than presented here. But as a general lesson, we can conclude that for redundant calibration it is preferable to have strong sources like those present in the EoR1 and EoR2 fields either at the pointing center or at the edge of the beam. EoR field 0 would therefore be an excellent field for redundant calibration.

In this work, we have not considered the differences in the antenna response of different antenna tiles. Work by Line et al. (2018) shows that the tile beam differences are on the order of 10%. This poses most likely the largest hurdle for redundant calibration. Studying the effect of these beam differences and how it impacts the redundancy of the MWA hexes and other
radio telescopes is therefore crucial to understand the limitations of redundant calibration in realistic telescopes.

This is also where the true strength of sky based calibration methods comes into play. Because redundant calibration relies on the assumption that each antenna observes the same radio sky, it is also unable to solve for direction dependent effects introduced by different antenna responses and ionospheric distortions. The field of direction dependent calibration faced quite a number of challenges, e.g., diffuse emission detected by shorter baselines, solving for enough different directions to capture variations in the ionosphere or the primary beam responses, optimizing the calibration timescale to reduce computational load, and the observed curvature of the sky (w-correction) due to the wide FoVs of these new arrays. A large effort has gone into solving these issues, e.g., SAGEcal resolved the computational load of solving for a large number of directions by using the SAGE algorithm rather than traditional least squares optimization (Yatawatta et al. 2009), facet calibration divides the sky in facets to reduce the number of parameters to solve for simultaneously (van Weeren et al. 2016), RTS employs the MWA’s uv-coverage to perform snapshot imaging tackling the w-term problem (Mitchell et al. 2008), with the diffuse emission of the Milky Way remaining as a major challenge. We have only mentioned a few implementations available as each science case has its own goal accompanied with its own implementation of sky based calibration. But the result of this large effort are impressive high-fidelity images, required to either study foreground sources or to subtract them. The latter being the goal for the EoR experiment. Solving for these higher-order calibration features is, however, out of reach for redundant calibration. Furthermore, redundant calibration does not truly escape the need for a sky model, because the degenerate parameters in Equation (20) need to be constrained by external information, i.e sky-based calibration. Li et al. (2018) directly compare redundant calibration using OMNICAL and sky model calibration using FHD and find that they perform similarly on data from real MWA EoR observations, that include the position offsets and tile beam differences. Li et al. (2018) also investigate the complimentary nature of the two different calibration techniques and find that combining the two methods improves the sensitivity to the EoR power spectrum, demonstrating that “hybrid approaches” are the best way forward. However, this final step can also introduce errors with spectral structure due to an incomplete sky model. Redundant calibration is, at best, a way to add another constraint for a first-order calibration step. Higher order effects require a pure sky based calibration that include direction dependent effects. Nevertheless, redundant calibration can still add useful information if carefully applied. Sievers (2017) propose a calibration algorithm that sits in the middle ground between agnostic redundant calibration and pure sky model based calibration maximally using the information in the data of a generic radio telescopes. This methodology seems to be a promising avenue for the MWA and the future SKA.

6. Conclusion and Future Work

In this work, we use a rudimentary implementation of the logcal and lincal algorithm to understand the fundamental limitations of redundant calibration on nearly redundant telescopes. We simulate redundant calibration under different radio sky conditions and find that the phase solutions are systematically impacted by position offsets in a redundant telescope. Based on our simulations we conclude the following key statement: redundant calibration performs best when strong radio sources are either at field center or at the edge of the primary beam. We also compare redundant calibration to sky model based calibration and find that for the MWA redundant calibration of the redundant hexes performs better than a sky based approach. However, we require further work to understand the impact of non-redundancies introduced by differences in tile beam responses that may be of larger concern to the MWA EoR experiment. Moreover, we also assumed that the sky model consists of only a single source. More work would be required to understand the completeness threshold above which a sky based approach truly outperforms redundant calibration on a nearly redundant array. Finally, to optimally calibrate our radio telescopes we require a hybrid approach that bridges the gap between redundancy-based and sky based calibration, and we see this as the way forward in the calibration of the current and next generation of radio telescopes.

R.C.J. thanks Ruby Byrne for useful discussions, and Wenyang Li, Bart Pindor and Christopher Jordan for constructive feedback on this manuscript. This work was supported by the Centre for All Sky Astrophysics in 3 Dimensions (ASTRO 3D), an Australian Research Council Centre of Excellence, funded by grant CE170100013. This research has made use of NASA’s Astrophysics Data System. We acknowledge the International Centre for Radio Astronomy Research (ICRAR), a Joint Venture of Curtin University and The University of Western Australia, funded by the Western Australian State government. This scientific work makes use of the Murchison Radio-astronomy Observatory, operated by CSIRO. We acknowledge the Wajarri Yamatji people as the traditional owners of the Observatory site. Support for the operation of the MWA is provided by the Australian Government (NCRIS), under a contract to Curtin University administered by Astronomy Australia Limited.

Appendix

Multi Frequency Implementation

Redundant calibration is typically presented on a channel-by-channel basis in contrast with standard model based calibration schemes that operate over a range of frequencies. The multi-frequency approach uses all of the information available in other frequency channels. In this section we will discuss the multi-frequency implementation of logcal, allowing it to benefit from the multi-frequency information available in radio interferometry data. Another motivation is that a multi-frequency implementation has the prospect of resolving the phase wrapping problem. Earlier, we described that phase wrapping occurs when a specific baseline $u$ observes a source on a specific coordinate $l = 1/2u$, the phase becomes ill-defined at this point. That same baseline should measure a defined phase when observing that same source at a different frequency. This property has motivated us to extend the classical redundant calibration framework by solving for the gain and visibilities simultaneously at different frequency channels. We will assume that the gain is the same at those frequencies, as a first-order approximation, but the visibilities are different.
We extend Equation (18), by stacking the measurement vectors at different frequencies and adding the visibilities from the additional frequency channels. We then construct the matrix accordingly by realizing that \( A \) can be split into two components:

\[
\begin{pmatrix}
0 & 1 & 0 & 0 & 0 \\
0 & -1 & 1 & 0 & 0 \\
0 & 0 & -1 & 1 & 0 \\
0 & 0 & 0 & -1 & 1 \\
1 & 0 & 1 & 0 & 0 \\
0 & -1 & 0 & 1 & 0 \\
0 & 0 & -1 & 0 & 0
\end{pmatrix}
\quad \begin{pmatrix}
1 & 0 \\
1 & 0 \\
1 & 0 \\
1 & 0 \\
1 & 0 \\
0 & 1 \\
0 & 1
\end{pmatrix},
\]

where \( A_\ell \) maps the gains onto the measurements and \( A_v \) maps the visibilities onto the measurements. We can construct a multi-frequency matrix combining \( A_\ell \) and \( A_v \), e.g., for a two-channel solutions estimation.

\[
A = \begin{pmatrix} A_\ell & A_v \\ A_\ell & A_v \end{pmatrix}
\]

Using this extended version of redundant calibration we return to the 5-element interferometer, while varying the number of channels involved to calibrate our antennas. Figure 10 shows the results when we attempt to calibrate five antennas using two frequency channels spaced around the actual frequency channel we are interested in. We observe a significant change in the structure of the variance of the phase solutions. The peaks around the phase wrapping points have severely decreased in width. However, it has not resolved the phase wrapping point, which was the aim of this multi-frequency implementation.

To understand why a multi-frequency extension of redundant calibration does not solve the problems immediately we have to return to Equation (18). The matrix \( [A^T A]^{-1} A^T \) mixes the phase wrapping and the non-phase wrapping channels into the calibration solutions. Although adding extra frequency channels does adds another set of constraints to the calibration solutions, the solutions do not escape the impact of the phase wrapping channel. Adding more channels therefore would decrease variance, but not resolve a phase wrapping point. To really resolve phase wrapping in redundant calibration we require the inclusion of knowledge of the sky.

Figure 10. From left to right: the amplitude gain, the phase gain solutions for a single antenna in an ideal 5-element interferometer, the visibility amplitude and the visibility phase for one of the two redundant groups as a function of strong source position \( l \). The dark blue line represents the mean of the solutions, the shaded blue area indicates the 1\( \sigma \) solutions variance. The amplitude solution variance inversely follows the shape of the beam, i.e., the variance increases when the beam response decreases. The mean of the phase solutions generally fluctuates around \( \phi = 0 \) and deviates along with the variance at so-called phase wrapping points, which are further explained in the text. The additional position offset, which causes a phase offset from the ideal redundant phase, is absorbed into the solutions causing a slope.
