An Efficient Feedback Calibration Algorithm for Direct Imaging Radio Telescopes

Adam P. Beardsley,[1]* Nithyanandan Thyagarajan,[1] Judd D. Bowman[1]
and Miguel F. Morales[2]

[1] *Arizona State University, School of Earth and Space Exploration, Tempe, AZ 85287, USA*
[2] *University of Washington, Department of Physics, Seattle, WA 98195, USA*

ABSTRACT

We present the E-field Parallel Imaging Calibration (EPICal) algorithm, which addresses the need for a fast calibration method for direct imaging radio astronomy correlators. Direct imaging involves a spatial fast Fourier transform of antenna signals, alleviating an $\mathcal{O}(N_a^2)$ computational bottleneck typical in radio correlators, and yielding a more gentle $\mathcal{O}(N_g \log_2 N_g)$ scaling, where N_a is the number of antennas in the array and N_g is the number of grid points in the imaging analysis. This can save orders of magnitude in computation cost for next generation arrays consisting of hundreds or thousands of antennas. However, because antenna signals are mixed in the imaging correlator without creating visibilities, gain correction must be applied prior to imaging, rather than on visibilities post-correlation. We develop the EPICal algorithm to form gain solutions quickly and without ever forming visibilities. This method scales as the number of antennas, and produces results comparable to those from visibilities. We use simulations to demonstrate the EPICal technique and study the noise properties of our gain solutions, showing they are similar to visibility based solutions in realistic situations. By applying EPICal to two seconds of Long Wavelength Array data we achieve a 65% dynamic range improvement compared to uncalibrated images, showing this algorithm is a promising solution for next generation instruments.

Key words: instrumentation: interferometers – techniques: image processing – techniques: interferometric

1 INTRODUCTION

In order to satisfy the survey speeds required for precision cosmology as well as searches for fast radio transients, radio astronomy is undergoing a paradigm shift toward interferometers consisting of hundreds to thousands of small, widefield antennas. Many arrays with this design are already built or under construction including the Hydrogen Epoch of Reionization Array[1] (HERA), the Murchison Widefield Array (MWA; Tingay et al. 2013; Bowman et al. 2013), the Donald C. Backer Precision Array for Probing the Epoch of Reionization (PAPER; Parsons et al. 2010), the LOw Frequency ARray (LOFAR; van Haarlem et al. 2013), the Canadian Hydrogen Intensity Mapping Experiment (CHIME; Bandura et al. 2014), the Long Wavelength Array (LWA; Ellingson et al. 2013), and the low frequency Square Kilometer Array (SKA1- Low; Mellema et al. 2013).

The most common radio correlator designs of today, the FX and lag correlators (Romney 1999), cross-multiply the signals from all pairs of antennas. This computation scales as the number of antennas squared, $\mathcal{O}(N_a^2)$ (Bunton 2004). As the number of elements in future arrays grows, the computational cost will become prohibitively expensive, and exploring efficient correlator schemes is essential to enable next generation instruments (Lonsdale et al. 2000). Meanwhile, radio transient monitoring requires access to high time and frequency resolution data to identify and characterize events such as fast radio bursts (FRBs, Lorimer et al. 2007), or to follow up gravitational wave candidates with radio observations (Abbott et al. 2016a; Abbott et al. 2016b). FRBs are relatively unexplored at low frequencies (< 1 GHz) (Rowlinson et al. 2016; Tingay et al. 2015; Trott et al. 2013), but are expected to occur on timescales $\Delta t \sim 1$–10 ms (Thornton et al. 2013). Recording the full visibility matrix for $N_a \gtrsim 10^3$ arrays at this timescale leads to extremely high data rates.

* E-mail: Adam.Beardsley@asu.edu
[1] http://reionization.org

Direct imaging correlators are a new variety of radio correlator which aim to alleviate both the computational strain of forming N_a^2 correlations and the high data throughput associated with short timescale science. This is done by performing a spatial fast Fourier transform (FFT) to image the antenna signals, then squaring and averaging in time. This process scales as $\mathcal{O}(N_g \log_2 N_g)$, where N_g is the number of grid points in the FFT (Morales 2011; Tegmark & Zaldarriaga 2009; Tegmark & Zaldarriaga 2010). The choice of grid points depends on the specific algorithm and implementation, and ultimately determines the extent and pixel size of the output image. For certain classes of telescopes, significantly those envisioned for next generation cosmology experiments, the computational scaling is a large improvement over the N_a^2 scaling of traditional methods. Furthermore, for the same temporal resolution, the output bandwidth will be lowered in cases where $N_g < N_a^2$.

A handful of prototype direct imaging correlators have been tested on arrays including the Basic Element for SKA Training II (BEST-2) array (Foster et al. 2014), the Omniscope (Zheng et al. 2014), and an earlier pulsar timing experiment at GHz frequencies (Otobe et al. 1994; Daishido et al. 2000). Each of these are examples of so-called FFT correlators – a subclass of direct imaging correlators which rely on identical antennas with restricted placement, which allows the FFT to be performed without gridding. We recently released the E-field Parallel Imaging Correlator (EPIC; Thyagarajan et al. 2017), which is a software implementation of the Modular Optimal Frequency Fourier (MOFF; Morales 2011) imaging algorithm. This architecture leverages the software holography/A-transpose framework to grid electric field data streams before performing the spatial FFT, allowing for an optimal map without placing constraints on array layout or requiring identical antennas (Morales & Matejek 2009; Bhatnagar et al. 2008; Tegmark 1997b).

A challenge common to all direct imaging algorithms is calibration of the antenna gains. With a traditional cross correlator, visibilities are averaged in time, reducing the data volume, then used to calibrate after observing and before further processing such as imaging. However, a direct imaging correlator mixes the signals from all antennas before averaging in time, making calibration a requirement at the front end, before any averaging and imaging. Previous solutions have involved applying calibration solutions generated from a parallel FX correlator (Zheng et al. 2014; Foster et al. 2014), or integrating a dedicated FX correlator which periodically formed the full visibility matrix to solve for gains (Wijnholds & van der Veen 2009; de Vos et al. 2009). While these solutions were sufficient to enable the exploration of FFT correlators and beamformers, they will not scale to future arrays with $N_a \gtrsim 10^3$.

Here we present the E-field Parallel Imaging Calibration (EPICal) algorithm – a novel solution to the calibration problem, which can be integrated into direct imaging correlators and scales only as the number of antennas, $\mathcal{O}(N_a)$ (Section 2.2). This method uses a correlation of the uncalibrated antenna signal stream with an output image pixel from the backend of the correlator to solve for the complex gains of the antennas. Because the calibration must be applied before gridding and imaging, our solution requires an iterative approach where the data from one time series is used to update the gains which are applied to the following
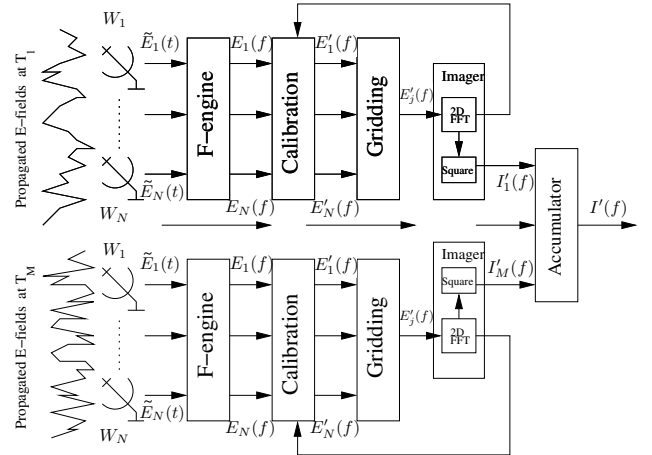
Figure 1. Data flow for the MOFF algorithm, reproduced from Thyagarajan et al. 2017 (their Fig. 1). The time domain electric field measurements are collected by the antenna elements then frequency transformed in the F-engine. The calibration is applied, then the electric field measurements are gridded using the antenna aperture illumination patterns. The imager performs a spatial 2D FFT on the gridded data before squaring and averaging. The accumulated images are then written to disk.

time series. An example implementation of the algorithm is available with the EPIC software package[2].

This work is a direct extension of Thyagarajan et al. 2017. We recommend the interested reader refer back to that work for details about the MOFF algorithm itself and the EPIC implementation. In this manuscript we review the MOFF algorithm and derive the calibration algorithm in §2. We then demonstrate the algorithm in simulations in §3, and discuss noise trends in §4. We apply the algorithm to a sample LWA data set in §5. Finally we conclude and discuss potential extensions to the algorithm in §6.

2   MATHEMATICAL FRAMEWORK

2.1   Review of the MOFF algorithm

We begin by reviewing the data flow of the MOFF algorithm, highlighting aspects relevant to the calibration. The interested reader is encouraged to refer to Morales 2011 and Thyagarajan et al. 2017 for a more thorough discussion. Figure 1 is reproduced from Thyagarajan et al. 2017 (their Fig. 1) to illustrate the various steps of the algorithm.

We begin by considering the electric field vector incident on the ground as a function of position and time, $\widehat{\mathbf{E}}(\mathbf{r}, \tau)$. Following Clark 1999, we make a few assumptions in order to clarify notation and simplify the problem. First we note that the electric field over a finite time interval, Δt, can be expressed as a Fourier series with coefficients corresponding to frequencies f. The Fourier coefficients are related to the time-domain electric field by[3] (e.g. Thompson et al. 2001,

---

[3]   A digital system will discretize the time-domain electric field and estimate equation 1 as a sum for a finite number of frequency

Appendix 3.1)

$$\mathbf{E}(\mathbf{r}, f_j, t) = \frac{1}{\Delta t} \int_{t-\Delta t/2}^{t+\Delta t/2} \widetilde{\mathbf{E}}(\mathbf{r}, \tau) e^{-2\pi i f_j \tau} d\tau. \tag{1}$$

Because we are considering a finite time interval, the coefficients are independent for discrete frequencies, $f_j = j/\Delta t$, for all integers j. Likewise, the variable t represents discrete times separated by Δt. The time-domain electric field over the same time interval, $t - \Delta t/2 \leq \tau \leq t + \Delta t/2$, can be recovered by an infinite sum,

$$\widetilde{\mathbf{E}}(\mathbf{r}, \tau) = \sum_{j=-\infty}^{\infty} \mathbf{E}(\mathbf{r}, f_j, t) e^{2\pi i f_j \tau}. \tag{2}$$

By analyzing the quasi-monochromatic components of the electric field spectrum, $\mathbf{E}(\mathbf{r}, f_j, t)$, we can treat frequency channels independently. For brevity, henceforth we will omit the subscript j from the frequency variable.

We will assume all antennas lie in a plane, reducing the problem to two dimensions. The MOFF algorithm in principle can account for non-coplanar arrays through the W-projection algorithm (Cornwell et al. 2008) by using the electric field Fresnel diffraction pattern to project each antenna measurement to a common plane. Because the EPIC code does not currently support this, we leave it to future work.

Next we treat the electric field for each polarization separately. This allows us to express the electric field for each polarization as a scalar, $E(\mathbf{r}, f, t)$, and treat subsequent field propagation and measurements as scalar rather than tensor multiplications. A full polarization treatment is required to construct polarized sky images, however the calibration will only require a sky model expressed in the instrumental polarization basis. We can therefore treat each instrument polarization independently.

Finally we assume the sources of the electric field are far away, allowing us to describe the electric fields propagating from sky sources as originating from an imaginary celestial sphere located a large distance from the ground. We express this intermediary electric field as an angular distribution, $\mathcal{E}(\hat{\mathbf{s}}, f, t)$, where $\hat{\mathbf{s}}$ is the unit vector in the direction of a patch of the celestial sphere. We further assume the space between the celestial sphere and the ground is empty, allowing us to express the electric field incident on the ground as a Fourier integral of the propagated fields across the hemisphere above the horizon.

$$E(\mathbf{r}, f, t) = \int_{\text{sky}} \mathcal{E}(\hat{\mathbf{s}}, f, t) e^{-2\pi i f \mathbf{r} \cdot \hat{\mathbf{s}}/c} d\Omega \tag{3}$$

Here $d\Omega$ is the differential solid angle. By making the latter assumptions we have left treatment of the ionosphere and terrestrial emission such as local interference or thermal radiation from the ground for future work.

Next we turn to the measurement performed by an antenna with frequency dependent far-field radiation pattern $\mathcal{W}_a(\hat{\mathbf{s}}, f)$, where the subscript a denotes a specific antenna. The far-field radiation pattern of the antenna gives the direction-dependent response of the antenna, and the total signal can be taken as the sum over all directions (e.g.

Kraus 1986). This allows us to express the electric field measurement that can be recovered from the induced voltage on an ideal antenna as a weighted integral over the sky (e.g. Tegmark & Zaldarriaga 2010).

$$E_a^I(f, t) = \int_{\text{sky}} \mathcal{W}_a(\hat{\mathbf{s}}, f) \mathcal{E}(\hat{\mathbf{s}}, f, t) e^{-2\pi i f \mathbf{r}_a \cdot \hat{\mathbf{s}}/c} d\Omega. \tag{4}$$

Here we have adopted the common convention of normalizing the radiation pattern to have response of unity at the center of the main lobe (e.g. Napier 1999).

Equivalently, we can employ properties of the Fourier transform to express the effect of the antenna radiation pattern in the plane of the aperture.

$$E_a^I(f, t) = \int_{\text{aperture}} W_a(\mathbf{r} - \mathbf{r}_a, f) E(\mathbf{r}, f, t) d^2\mathbf{r} \tag{5}$$

Here we used the Fourier relationship between the antenna's far-field radiation pattern and the aperture illumination pattern (sometimes referred to as the aperture distribution),

$$W_a(\mathbf{r} - \mathbf{r}_a, f) = \frac{f^2}{c^2} \int_{\text{sky}} \mathcal{W}_a(\hat{\mathbf{s}}, f) e^{2\pi i f (\mathbf{r} - \mathbf{r}_a) \cdot \hat{\mathbf{s}}/c} d\Omega. \tag{6}$$

Intuitively this is the response function by which the antenna weights the electric fields over the aperture, and sums at the feed. While it is possible to define each antenna aperture illumination pattern in a way to absorb the relative position of the antenna, we keep the $\mathbf{r}_a$ coordinate to retain the separable concepts of the primary beam (the product of two antenna radiation patterns) and the array point spread function (square of the Fourier transform of the array layout).

Thus far we have considered the perfectly measured electric field distribution according to the antenna far-field radiation patterns. Here we introduce a multiplicative complex gain as well as an additive noise term which corrupt the measured signals at each of the antennas.

$$E_a(f, t) = g_a(f, t) E_a^I(f, t) + n_a(f, t) \tag{7}$$

We have carried the frequency dependence to this point to make clear the complex gains are frequency dependent. However, each subsequent step treats each frequency channel independently, so we will drop the f to simplify notation. Furthermore, while antenna gains may vary with time, we will assume they are constant over the period of our calibration, and thus drop their time dependence from the notation.

The first step of the MOFF algorithm is to digitize the antenna signals and channelize the data over a finite time interval, effectively estimating the electric field spectrum in equation 1 over a finite frequency bandwidth. This operation is often referred to as the F-engine of a correlator. The MOFF shares the same F-engine design as an FX cross correlator which performs the frequency transform (F-engine) before cross multiplying (X-engine). For a recent review of spectrometers used in radio astronomy, please see Price 2016.

The MOFF algorithm next calls for a calibration of the electric field measurements. The goal of our new calibration method will be to form an estimate of the gains, g_a', where the prime represents an estimate. For now we will assume we have formed an estimate to proceed with the MOFF pipeline. We correct the incoming electric field data stream

channels. At this stage we are considering the true electric field and keep the integral.

using our current estimate of the gains.

$$E'_a(t) = E_a(t)/g'_a \tag{8}$$

The next step is to grid the calibrated electric field measurements to a finite set of regularly spaced positions on the plane of the ground, $\mathbf{r}_j$. The extent of the grid is chosen to encompass the footprint of the antenna array. The spacing of grid points should be chosen to be smaller than the antenna size to achieve sufficient sampling, and a typical choice is several grid points per linear dimension of the antennas. The size of the grid spacing determines the field of view of the image (through an inverse relationship), so this choice is typically limited on one end by the antenna size (instrumental field of view), and on the other by half the observation wavelength (full sky). The EPIC implementation currently only supports a rectangular grid, but in principle the type of grid can be extended to any arrangement which can be Fourier transformed with $\mathcal{O}(N_g \log_2 N_g)$ computational complexity including hexagonal grids (Mersereau 1979) or arbitrary hierarchies of rectangular grids with shears and rotations at each level (Tegmark & Zaldarriaga 2010).

MOFF uses the antenna aperture illumination patterns as the gridding kernel according to the software holography/A-transpose technique (Morales & Matejek 2009; Bhatnagar et al. 2008). Mathematically we can express the gridding operation as

$$E'_j = \sum_a W_a(\mathbf{r}_j - \mathbf{r}_a) E'_a(t) \tag{9}$$

Formally the value at each pixel is calculated by summing the antenna measurements over all antennas, each weighted by its respective aperture illumination pattern. However the aperture illumination patterns are typically compact and non-overlapping, so most pixels will only contain contribution from one antenna, allowing equation 9 to be calculated through a sparse matrix multiplication. This gridding places the data on a regular grid and achieves an optimally weighted map, analogous to widely used Cosmic Microwave Background analysis techniques (Tegmark 1997a). Unlike traditional correlators which grid visibilities, the MOFF grids the electric field measurements directly. This operation must be performed for every time interval, before any averaging of the data.

After gridding, the imaging portion of the MOFF performs a 2D spatial FFT to form instantaneous electric field images of the sky at fixed locations $\hat{\mathbf{s}}_k$. The FFT is possible because we have gridded to regularly spaced samples in a 2D plane. The Fourier dual coordinates to $\mathbf{r}f/c$ are (l, m), where $l = \sin\theta\cos\phi$, $m = \sin\theta\sin\phi$, and (θ, ϕ) are the zenith and azimuthal angles[4]. The full unit vector towards the sky is $\hat{\mathbf{s}} = (l, m, \sqrt{1 - l^2 - m^2})$. Being direction cosines, grid locations with $l^2 + m^2 > 1$ are non-physical and should be ignored.

The algorithm up to this point can be summarized

---

[4] Observations pointed away from zenith can be imaged by applying a phase to the electric fields in the same way one would phase visibilities, and redefining (θ, ϕ) coordinates relative to the phase center.

through the following equation.

$$\mathcal{E}'(\hat{\mathbf{s}}_k, t) = \frac{1}{N_a} \underbrace{\sum_j e^{2\pi i f \mathbf{r}_j \cdot \hat{\mathbf{s}}_k/c}}_{\text{2D FFT}} \underbrace{\sum_a W_a(\mathbf{r}_j - \mathbf{r}_a)}_{\text{Gridding}} \tag{10}$$
$$\times \underbrace{\frac{1}{g'_a}\left(g_a E_a^I(t) + n_a(t)\right)}_{\text{Calibration}}$$

The sum over antennas uses the antenna aperture illumination pattern to grid the calibrated electric field measurements onto regular gridpoints, $\mathbf{r}_j$. The sum over j denotes the 2D FFT to sky coordinates, resulting in an estimate for the instantaneous electric field image. The sums over a and j combine the antenna signals, which is the primary reason calibration must be performed beforehand where individual antennas retain their identities.

As an illustrative step, we can simplify the above expression for two regimes: where the grid size is significantly smaller than the antenna size, or all antennas are identical and pseudo-randomly placed (or both). In these cases, after exchanging the sums, we can use the discrete Fourier transform to approximate the Fourier relationship between the antenna aperture illumination pattern and the far-field radiation pattern.

$$\mathcal{E}'(\hat{\mathbf{s}}_k, t) = \frac{1}{N_a} \sum_a \frac{1}{g'_a}\left(g_a E_a^I(t) + n_a(t)\right) e^{2\pi i f \hat{\mathbf{s}}_k \cdot \mathbf{r}_a/c}$$
$$\times \sum_j W_a(\mathbf{r}_j - \mathbf{r}_a) e^{2\pi i f \hat{\mathbf{s}}_k \cdot (\mathbf{r}_j - \mathbf{r}_a)/c}$$
$$\approx \frac{1}{N_a} \sum_a \frac{1}{g'_a}\left(g_a E_a^I(t) + n_a(t)\right) e^{2\pi i f \hat{\mathbf{s}}_k \cdot \mathbf{r}_a/c} \mathcal{W}_a(\hat{\mathbf{s}}_k) \tag{11}$$

While this step is not necessary (and in fact is not a part of the imaging algorithm), the approximation holds for most current and planned cosmology experiments, so we will proceed with the more simple expression. In this form we see that the effect of gridding with the antenna aperture illumination pattern is to attenuate the image by a factor of the far-field radiation response, $\mathcal{W}_a(\hat{\mathbf{s}})$. This added factor of the instrument response is sometimes referred to as the *holographic frame* in the map-making literature (e.g. Morales & Matejek 2009; Sullivan et al. 2012).

Finally, the images are squared and averaged in time to form dirty images, $I'(\hat{\mathbf{s}}_k)$, then written to disk. While operations have been reordered, the resulting image is equivalent to one produced by gridding and imaging visibilities, up to a precision set by gridding coarseness. By shifting from cross-correlation of all pairs of antennas to a spatial FFT, the computational cost is significantly lowered for certain classes of arrays, particularly those with densely-filled apertures of small antennas. For these arrays, the output bandwidth is also greatly reduced as visibilities are replaced with gridded images. These gridding, computational, and data bandwidth differences were explored in Thyagarajan et al. 2017 (sections 4.3, 6.1, and 6.2 respectively).

2.2 Derivation of calibration

We saw above that the MOFF algorithm (in general any direct imaging architecture) combines the signals from all
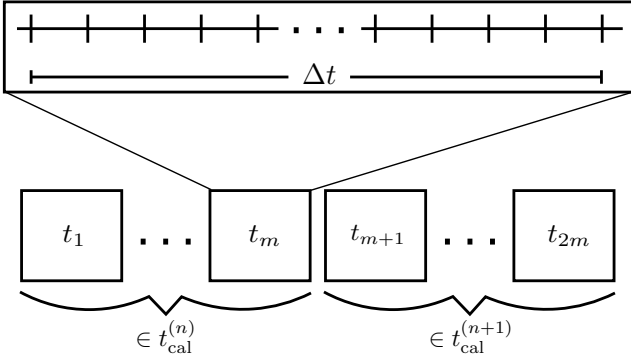
Figure 2. Schematic of timescales involved in the calibration algorithm. The top box depicts several digitized time samples of the observation. It is common that real time samples are separated by the Nyquist interval of $1/2B$ (or $1/B$ for complex samplers), where B is the bandwidth of the observation. The antenna electric field spectra, $E_a(f,t)$, are computed over the time interval Δt for consecutive times as shown. A finite stream of these intervals, labeled by the set of times $t_{\mathrm{cal}}^{(n)}$, will be used to generate an nth estimate of the gains, which are applied to the following stream, labeled by $t_{\mathrm{cal}}^{(n+1)}$. The gains are assumed to be constant between the two time streams $t_{\mathrm{cal}}^{(n)}$ and $t_{\mathrm{cal}}^{(n+1)}$.

antennas, and thus requires the calibration to be applied before imaging. Furthermore, visibilities are never formed, which are traditionally the basic measurement used to form calibration solutions. Here we derive an alternative method to estimate the antenna gains quickly, using the data products of the MOFF.

For some applications, the electric field measurements can be recorded directly, and gains may be determined in post-processing, such as in our demonstration in Section 5. However, with the large-N_{a}, high bandwidth, and high duty-cycle instruments required for future cosmological experiments, it is only practical to record time averaged quantities. Therefore we will assume a requirement that gains must be determined on relatively short timescales equal to or less than the stability timescale of the instrument. We will adopt a strategy to use a finite stream of data to form a gain estimate, which we will apply to the subsequent stream of data. This inherently assumes that the changes in gains are negligible over adjacent time streams. Figure 2 illustrates the several timescales involved in the calibration algorithm.

In the case of exceptionally stable instruments, where gains are constant over long timescales, the requirement for a fast calibration scheme is alleviated and more sophisticated fitting strategies can be undertaken, as we briefly discuss in Section 6. Even in this ideal case, we will avoid adding to the computational strain of the correlator by using quantities easily accessible to a direct imaging system.

Because data from one interval is used to update the gain solutions for the next, the process of finding a solution is an iterative one. We will use parenthetical superscripts to denote the calibration loop number. For example, we will assume we have already formed an estimate from n loops, $g_a^{(n)}$. We will use these estimates to calibrate the next stream of data in order to form an updated estimate, $g_a^{(n+1)}$.

As a starting point, we consider the feedback calibration

outlined in Morales 2011. There it was suggested to form a correlation of the uncalibrated antenna measurements with the complex conjugate of an image pixel from the output of the correlator. We will see below that this quantity can be related to the sum of visibilities involving the antenna used in the correlation – exactly the sum needed to calibrate in the case of a single point source sky. However, we aim for a more generalized solution for arbitrarily complex sky models. We therefore study the full expression for the antenna-pixel correlation,

$$\mathcal{K}_{a,\hat{\mathbf{s}}_0}^{(n)} \equiv \left\langle E_a(t)\mathcal{E}'^*(\hat{\mathbf{s}}_0,t) \right\rangle_{t_{\mathrm{cal}}^{(n)}} , \tag{12}$$

where the superscript n again represents the quantity formed in the nth calibration loop, and we use $t_{\mathrm{cal}}^{(n)}$ outside the brackets on the right-hand side to indicate that averaging is over all t within the nth calibration loop (see figure 2). The superscript * indicates complex conjugation, and $\hat{\mathbf{s}}_0$ is the pixel center nearest a bright calibrator of interest. The following will hold for any chosen pixel, $\hat{\mathbf{s}}_0$, though it is advantageous to choose a pixel which contains a bright source to achieve a high signal to noise ratio.

Plugging equation 11 into equation 12, and substituting the $g_b^{(n)}$ for g_b' to indicate the specific loop number, we find,

$$\begin{aligned}
\mathcal{K}_{a,\hat{\mathbf{s}}_0}^{(n)} = &\left\langle \left(g_a E_a^I(t) + n_a(t) \right) \frac{1}{N_{\mathrm{a}}} \right. \\
&\times \sum_b \frac{1}{g_b^{*(n)}} \left(g_b^* E_b^{I*}(t) + n_b^*(t) \right) \\
&\left. \times e^{-2\pi i f \hat{\mathbf{s}}_0 \cdot \mathbf{r}_b/c} \mathcal{W}_b^*(\hat{\mathbf{s}}_0) \right\rangle_{t_{\mathrm{cal}}^{(n)}} \\
= &\frac{1}{N_{\mathrm{a}}} \sum_b \frac{1}{g_b^{*(n)}} \mathcal{W}_b^*(\hat{\mathbf{s}}_0) e^{-2\pi i f \hat{\mathbf{s}}_0 \cdot \mathbf{r}_b/c} \\
&\times \left\langle \left(g_a E_a^I(t) + n_a(t) \right) \left(g_b^* E_b^{I*} + n_b^*(t) \right) \right\rangle_{t_{\mathrm{cal}}^{(n)}} \\
= &\frac{1}{N_{\mathrm{a}}} \sum_b \frac{1}{g_b^{*(n)}} \mathcal{W}_b^*(\hat{\mathbf{s}}_0) e^{-2\pi i f \hat{\mathbf{s}}_0 \cdot \mathbf{r}_b/c} \\
&\times \left(g_a g_b^* \left\langle E_a^I(t) E_b^{*I}(t) \right\rangle_{t_{\mathrm{cal}}^{(n)}} + g_a \left\langle E_a^I(t) n_b^*(t) \right\rangle_{t_{\mathrm{cal}}^{(n)}} \right. \\
&\left. + g_b^* \left\langle n_a(t) E_b^{I*}(t) \right\rangle_{t_{\mathrm{cal}}^{(n)}} + \left\langle n_a(t) n_b^*(t) \right\rangle_{t_{\mathrm{cal}}^{(n)}} \right) \tag{13}
\end{aligned}$$

where in the second step we group time-dependent terms, and in the third step we expanded the product within the time average. Because the noise does not correlate with the ideal antenna measurements we can drop the middle terms. We will also assume that the cross-correlation noise ($a \neq b$) is zero mean. Then by defining an ideal visibility as $V_{ab}^I \equiv \left\langle E_a^I(t) E_b^{*I}(t) \right\rangle_t$, we can simplify to

$$\begin{aligned}
\mathcal{K}_{a,\hat{\mathbf{s}}_0}^{(n)} = &\frac{1}{N_{\mathrm{a}}} \sum_b \frac{1}{g_b^{*(n)}} \mathcal{W}_b^*(\hat{\mathbf{s}}_0) e^{-2\pi i f \hat{\mathbf{s}}_0 \cdot \mathbf{r}_b/c} \\
&\times \left(g_a g_b^* V_{ab}^I + \delta_{ab} \left\langle |n_a(t)|^2 \right\rangle_{t_{\mathrm{cal}}^{(n)}} \right) \tag{14}
\end{aligned}$$

where δ_{ab} is the Kronecker-delta function. To be clear, visibilities are never directly formed from the data. Rather it is convenient to express $\mathcal{K}_{a,\hat{\mathbf{s}}_0}^{(n)}$ in terms of the ideal visibilities, which we can model when solving for the gains below.

We will ultimately wish to model the right side of equation 14 in order to solve for the antenna gains. However, the self-correlated noise, $\left\langle |n_a(t)|^2 \right\rangle$, can often be significantly larger than the cross correlation power, and difficult

to model precisely. It is therefore beneficial to subtract the self-correlation directly from $\mathcal{K}_{a,\hat{\mathbf{s}}_0}^{(n)}$. By properly accounting for the various weighting and phasing terms, we can use self-correlations of the uncalibrated antenna electric field measurements to remove the corresponding terms from equation 14:

$$\mathcal{C}_{a,\hat{\mathbf{s}}_0}^{(n)} \equiv \mathcal{K}_{a,\hat{\mathbf{s}}_0}^{(n)} - \frac{1}{N_a g_a^{*(n)}} \mathcal{W}_a^*(\hat{\mathbf{s}}_0) e^{-2\pi i f \hat{\mathbf{s}}_0 \cdot \mathbf{r}_a / c} \left\langle |E_a|^2 \right\rangle_{t_{\text{cal}}^{(n)}}$$

$$= \frac{g_a}{N_a} \sum_{b \neq a} \frac{g_b^*}{g_b^{*(n)}} \mathcal{W}_b^*(\hat{\mathbf{s}}_0) e^{-2\pi i f \hat{\mathbf{s}}_0 \cdot \mathbf{r}_b / c} V_{ab}^I. \quad (15)$$

Due to the above subtraction, the sum is now over $b \neq a$, explicitly removing the dependence on the self-correlated antenna signals.

The next step is to use this antenna-pixel correlation to update our gain solution. We will model the right hand side of equation 15 by assuming our current gain estimates are approximately correct, $g_b^{(n)} \approx g_b$, and replacing the ideal visibilities with a set of model visibilities, V_{ab}^M. These model visibilities rely on knowledge of the sky, and should be precomputed before observing. We can then solve equation 15 for g_a to achieve an updated estimate.

$$g_a^{(n+1)} = \frac{\mathcal{C}_{a,\hat{\mathbf{s}}_0}^{(n)} N_a}{\sum_{b \neq a} \mathcal{W}_b^*(\hat{\mathbf{s}}_0) e^{-2\pi i f \hat{\mathbf{s}}_0 \cdot \mathbf{r}_b / c} V_{ab}^M} \quad (16)$$

This equation is our prescription for estimating the antenna gains of a direct imaging array. The computation complexity of the part of the calculation that must be performed for each time interval scales as $\mathcal{O}(N_a)$ since we form an antenna-pixel correlation, $\mathcal{C}_{a,\hat{\mathbf{s}}_0}^{(n)}$, for each antenna.

We show schematically the process of calibrating a direct imaging correlator in Fig. 3. Computationally expensive steps that must be performed for each time interval are shown inside the gray box. The uncalibrated antenna signals are tapped out after the F-engine and correlated against the output image pixel of interest. The correlated values are then used to estimate the gains using equation 16, and additional fitting if desired (see Section 6). This step does not have the stringent cadence requirement of the steps in the gray box because the antenna-pixel correlations have been averaged in time. The gains are then passed back to the correlator to update the calibration for subsequent integration intervals.

While testing we found equation 16 resulted in oscillatory gain solutions over time, as is often the case in iterative methods. To mitigate this we introduce a damping factor, $0 \leq \gamma < 1$, which is used to attenuate the gain update, effectively giving the solutions memory of previous iterations.

$$g_a^{(n+1)} = \frac{(1-\gamma)\mathcal{C}_{a,\hat{\mathbf{s}}_0}^{(n)} N_a}{\sum_{b \neq a} \mathcal{W}_b^*(\hat{\mathbf{s}}_0) e^{-2\pi i f \hat{\mathbf{s}}_0 \cdot \mathbf{r}_b / c} V_{ab}^M} + \gamma g_a^{(n)} \quad (17)$$

We found that while equation 16 does indeed converge on good solutions, the process is made faster by tuning the damping factor. Once the loop converges the damped version is essentially a weighted average over the past several iterations, giving the solutions a longer effective integration time.

We conclude this section by connecting our calibration expression to that found in a visibility framework. In the special case of a single bright calibrating source at zenith, we can greatly simplify equations 15 and 16. We will assume
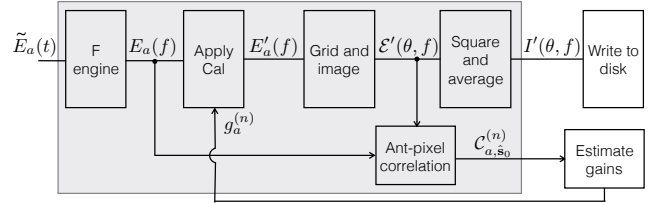
Figure 3. The general data flow of the MOFF correlator, with a feedback calibration loop. A pixel from the (unsquared) image is tapped out and correlated against the input antenna electric field measurements to form $\mathcal{C}_{a,\hat{\mathbf{s}}_0}^{(n)}$ coefficients (equation 15). These coefficients are then used to update the gain estimates using equations 16 and 17, and any additional fitting (Section 6). The gray box shows operations which must be done at high speed (before averaging in time), while the white boxes show operations which can be performed at lower speeds because they involve time averaged quantities, and need to be updated on slower timescales.

the radiation patterns are normalized such that $\mathcal{W}(0) = 1$. We can further drop the exponential phase terms because $\hat{\mathbf{s}}_0$ is perpendicular to our planar antenna array. We then absorb the true gains into the true visibilities in equation 15 to express as a sum of measured, uncalibrated, visibilities.

$$\mathcal{C}_{a,0}^{(n)} \rightarrow \frac{1}{N_a} \sum_{b \neq a} \frac{1}{g_b^{*(n)}} V_{ab} \quad (18)$$

We next plug this expression into equation 16 to find our simplified calibration solution for a single bright point source. Because our sky is a single bright point source, the model visibilities are simply the flux of the source, S_{src}.

$$g_a^{(n+1)} \rightarrow \frac{\sum_{b \neq a} V_{ab}/g_b^{*(n)}}{N_a S_{\text{src}}} \quad (19)$$

This is simply a gain-weighted sum of the measured visibilities over the flux of the source, which is indeed the limiting result from a visibility approach, for example seen in Mitchell et al. 2008. The ability to recover the equivalent expression despite not actually forming the visibilities is a result of the fact that only sums over visibilities come into the visibility-based solution, as was described in Morales 2011. We have confirmed the limiting case equivalence here, and will explore the more general case in more detail in the following sections.

3   SIMULATION

We first demonstrate our calibration method through a controlled simulation. A complex gain is created for each antenna with random phase and amplitude, which is used to corrupt the simulated data stream, then we attempt to estimate the gains using our calibration routine. We use the simulation software included in the EPIC package to create stochastic antenna measurements of a sky model. In this and following sections we restrict ourselves to arrays of identical antennas, and leave a demonstration with heterogenous arrays to future work.

Our simulated antenna array consists of the inner 51 antennas of the MWA layout (Beardsley et al. 2012), within a
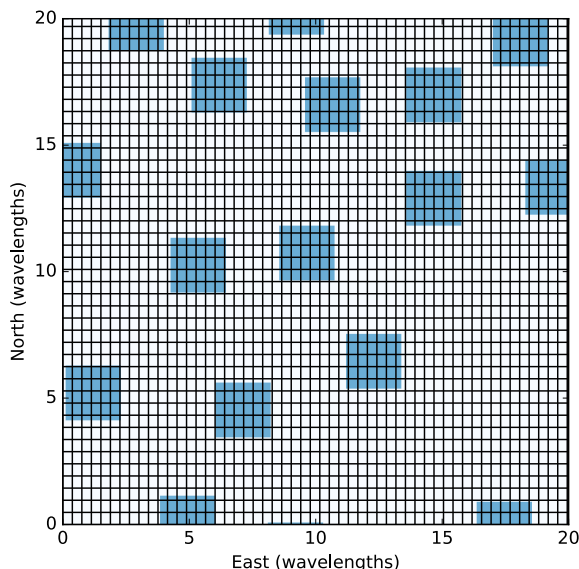
Figure 4. A subsection of the simulated array in Section 3, with overlaid grid. The blue squares show the locations of some of the antennas simulated. Grid points used to grid the electric field measurements (equation 9) are located at the intersections of black lines. Only a subsection of the array is shown so the grid points can be seen. We specified that the grid spacing be less than 0.5 wavelengths, and EPIC finds a grid which encompasses the array with the number of pixels in each dimension a power of two. This results in a grid spacing of 0.376 in the East-West direction, and 0.48 in the North-South direction, and a total grid size of 256×256 pixels.

bounding box of 150 m. The physical parameters of our simulation can be arbitrarily scaled to other cases. However, we provide typical units for an MWA observation as an example. The antenna aperture illumination pattern used is a 4.4 m square tophat; a very rough approximation to the square shaped tiles of the MWA. Because our algorithm treats frequency channels independently, we simulate only one channel. For context we treat this as a channel at 150 MHz with 40 kHz bandwidth, and adopt a sampling period of 25 μs. A subsection of the array is shown in Figure 4, along with lines depicting the grid points used to grid the electric field measurements (equation 9). The simulated signal consists of 10 random point sources with flux densities 0.5 Jy $\lesssim S \lesssim 1$ Jy within the main lobe of the primary beam.

For our unknown gains, we create a set of random complex numbers where the amplitude is drawn from a gaussian distribution centered around unity with width 0.25, and truncated at zero. The gain phases are drawn from a uniform distribution in the range $[-\pi, \pi]$. These are our "true gains", and we apply them to the frequency domain simulated antenna electric field measurements as in equation 7. Our analysis is blind to these values until the end of the process to check accuracy. The gain estimates are initialized with unity, $g_a^{(0)} = 1$.

We next process and image 400 time samples (10 ms). We also form the antenna-pixel correlations, $\mathcal{C}_{a,\hat{\mathbf{s}}_0}^{(0)}$, used in our calibration loop. The pixel used for the correlation is the source with the largest apparent flux (intrinsic flux attenuated by the primary beam). These correlation values are
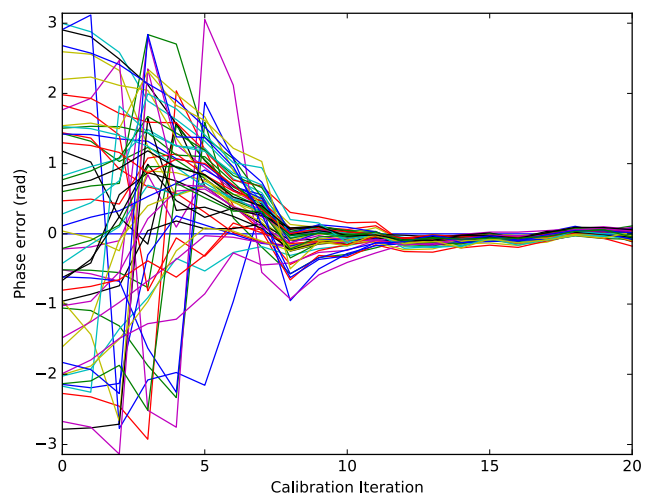

Figure 5. Phase error of gain estimates as a function of iteration for simulated calibration. The gains were initialized with random phases, but the calibration loop was able to recover the correct phases after about 12 iterations. Each line represents an antenna in our 51 MWA antenna sample.

used to update the gain estimates. We create perfect model visibilities of our 10 simulated sources to be used in equation 17. The updated gain estimates are used to calibrate the following 400 time samples. Through trial and error we found a damping factor of $\gamma = 0.35$ resulted in the quickest convergence in this simulation.

The calibration loop continues by updating the gain estimates every 400 time samples. The phase errors of our gain estimates are shown in Fig. 5 for 20 such iterations. The phase error plotted is the phase relative to the true gain for each antenna (various colored lines). One antenna was used as a reference to fix the absolute phase, so has zero phase error. The other 50 antennas are shown to have error spanning the range $[-\pi, \pi)$ initially, and after about 10 iterations lock into a solution, settling down to noise levels around iteration 12 (0.12 s). We stop the simulation when the updated gains trace the thermal noise of the simulated sources, which can be seen by the coherence of the 50 antenna gains after iteration 12.

The estimated gain amplitudes for the simulations are shown in Fig. 6. The quantity plotted is the magnitude of the estimated gains over the true gains, $\left| g_a^{(n)}/g_a \right|$, which places all antennas on the same scale. We can see the amplitudes converge toward their true values around the same time as the phases (iteration $\sim$12). At the beginning of calibration we can see the importance of the damping factor. At $n = 0$, a couple of gains are shown to have abnormally high amplitude estimates, notably one about 3.3 times its true value (red line). These unbalanced high estimates caused the entire set of gains to be under estimated at $n = 1$, even with a damping factor of 0.35. By $n = 5$ the unbalanced amplitudes have been damped out and the calibration continues. Without the damping factor, the oscillation seen in the first couple iterations would have been significantly larger and taken much longer to fade out.

After the gains have converged, we see both the phases
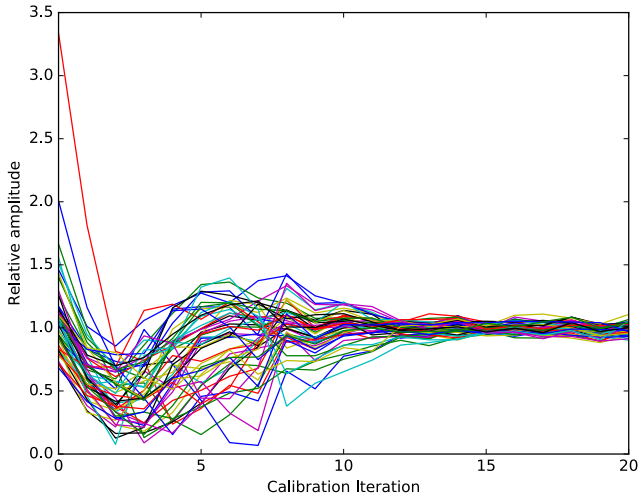
Figure 6. Same as Fig. 5, but for gain amplitudes. Each line represents an antenna in the 51 MWA antenna sample. The relative amplitude plotted is the magnitude of the estimated gains divided by the true gains. After about 12 iterations we see the calibration loop has settled around the correct values, with only noise remaining.

and amplitudes continue to fluctuate coherently. This is due to the stochastic fluctuations of the sources themselves, which are simulated as zero mean random processes with variance proportional to their flux density. For this simulation we restricted each calibration iteration to only 10 ms integration time. In principle a calibration implementation could use short integration times to allow the gains to converge, then increase the integration time to reduce this noise.

Images created at the beginning of calibration and at the end are shown in Fig. 7. Each image is obtained over a 400 time sample (10 ms) integration, corresponding to all snapshot images created with a given set of gain estimates. The left panel shows the image produced with our initialized unity gains. Because the phases are completely random, the image is essentially noise with the primary beam evident. After 20 calibration iterations, the image is far more clear, shown in the middle panel. The ten sources are visible, along with rumble throughout the image. By comparing with a perfectly calibrated image (right panel), we see that the rumble and the shapes of the point sources are matched between the two images. We therefore conclude that the artifacts visible are dominated by the point spread function, rather than calibration errors.

4   NOISE TRENDS

Next we study the effect of noise in our system, and the consequences of an incomplete sky model. We run a suite of simulations varying the receiver noise and integration times, while forming antenna-pixel correlations, $\mathcal{C}_{a,\hat{\mathbf{s}}_0}^{(n)}$, for EPICal gain solutions, and simultaneously forming visibilities from the same streams of measured electric fields to find visibility-based gain solutions for comparison.

We use a simulated sky consisting of a 5 Jy calibrator source, and 49 other random sources with apparent flux den-

sities 0.2 to 0.5 Jy (total sky power $\approx$ 23 Jy). We generated these sources randomly, but kept them fixed for each run. We simulate the same antenna layout as in Section 3. However, we increase the number of frequency channels simulated to 64, while keeping the flux of all sources constant across frequency. Because our calibration loop treats frequency channels independently, each channel can be treated as a separate trial of the simulation, and is used to better estimate the statistics.

To simulate receiver noise, we add a gaussian distributed complex random number to each antenna electric field measurement at each time sample according to equation 7. The level of the receiver noise, $\sigma_r = \langle |n_a(f,t)|^2 \rangle$, is varied in different simulation runs. We include two limiting cases, one where the receiver noise power is subdominant to the sky power ($\sigma_r = 10.0$ Jy), and the other where the receiver noise dominates the total power ($\sigma_r = 100.0$ Jy).

For each simulation run we find gain solutions using four methods:

i. EPICal, using a full sky model
ii. EPICal, using a single point source model
iii. Visibility-based, using a full sky model
iv. Visibility-based, using a single point source model.

In the cases where we use the full sky model, the model visibilities, V_{ab}^M, used in the calibration loop are created using all point sources in the simulated sky. For the single point source model, we only include the bright calibrator source in the model visibilities. While we vary the model used for calibration, the simulated "true sky" remains fixed with all 50 point sources.

For EPICal, we initialize our gain estimates with the true values ($g_a^{(0)} = g_a = 1$), and allow the estimate to be corrupted by the noise through ten iterations of the calibration loop. This allows us to bypass the initial convergence time and quickly study the effect of the noise over many simuations. We again adopt a damping factor $\gamma = 0.35$. Because EPICal updates the gain estimate at each calibration loop, t_{cal}, but retains a memory of previous iterations through the damping factor, the total integration time is not straightforward. We define an effective integration time by considering the relative weights of each previous $\mathcal{C}_{a,\hat{\mathbf{s}}_0}^{(n)}$ contributing to the current $g_a^{(n)}$ estimate, each with integration time t_{cal}. In the limit $n \to \infty$, the geometric series converges to an effective integration time of

$$t_{\mathrm{eff}} = t_{\mathrm{cal}} \times \frac{1+\gamma}{1-\gamma}. \tag{20}$$

With the damping factor we adopted here, $t_{\mathrm{eff}} \approx 2.1 \times t_{\mathrm{cal}}$.

We simultaneously form simulated visibilities, V_{ab}, by correlating all pairs of antenna electric field measurements (including the receiver noise). The electric field measurements are correlated for a duration equal to the effective integration time of the EPICal loop for comparison. We then find the visibility-based gain estimates by minimizing

$$\chi^2 = \sum_a \sum_{b \neq a} \left| V_{ab} - g_a g_b^* V_{ab}^M \right|^2, \tag{21}$$

using both versions of the model visibilities described above.

For each version of calibration, we observe the error in the gain estimates by averaging over both antennas and
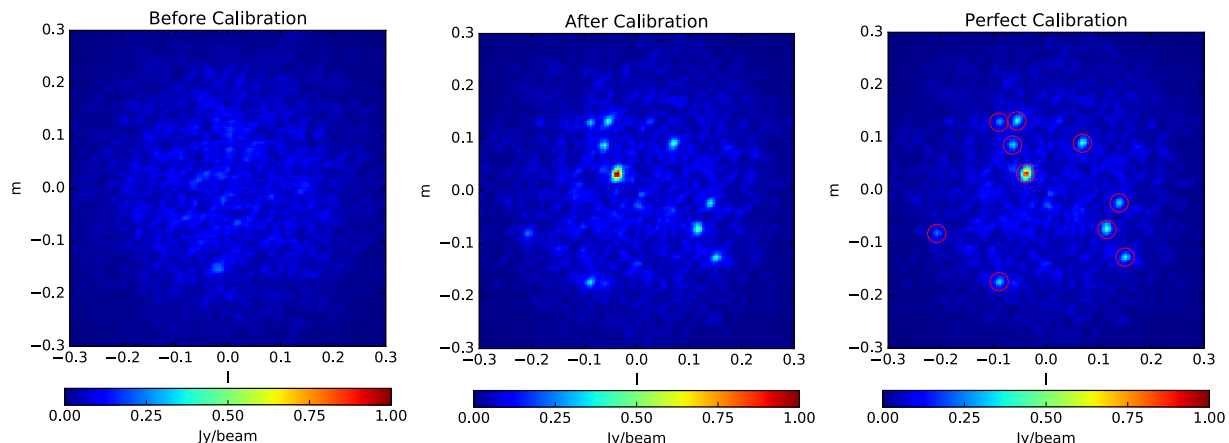
Figure 7. Images formed during simulated calibration. *Left:* An image generated from a single frequency channel and 400 time sample (10 ms) integration after the gain estimates are randomized. As expected with random phases, the image is completely noisy, with the shape of the primary beam evident. *Middle:* An image formed after calibration, again with a single frequency channel and 400 time sample integration. Now the simulated point sources are visible in the image. *Right:* An image generated from the same simulated data as the middle image, but with perfect calibration correction. Many of the features seen in the middle image are also seen here, confirming that they are point spread function artifacts. The ten simulated (and modeled) point sources are highlighted with red circles.

frequency channels.

$$\sigma_g = \left[\frac{1}{N_f N_a} \sum_f \sum_a \frac{\left| g_a^{(n)}(f) - g_a(f) \right|^2}{|g_a(f)|^2} \right]^{1/2} \qquad (22)$$

The results of our simulations for all four calibration methods are shown in Fig. 8. In the case of the full sky model, we see the errors trend downward with longer integration times, as expected. The EPICal errors are slightly higher than the visibility-based errors, on average about 23% difference. This is not surprising as EPICal only uses the information in a single pixel, while the visibilities use all information from the full sky model. However, we see that with this perfect model the errors do behave like noise in the sense that they integrate down with longer integrations with the expected $t^{-1/2}$ trend. With longer integration time (or larger damping factor), EPICal can achieve the same level of gain error as visibility based solutions.

The exact ratio of noise in gains from EPICal and the visibility-based estimator depends on the fraction of the total sky power contained in our calibrator source. In the example here, our calibrator accounted for 22% of the total sky power. We repeated the experiment in the regime where the calibrator dominated the sky and found that the difference in EPICal and visibility-based error goes to zero, as expected. For a typical HERA observation, the sky temperature is expected to by about 180 K (Jacobs et al. 2015), and a bright calibrator source could be about 10% of this power. In this regime we found that EPICal gain noise was about 60% higher than visibility-based gain noise for the same effective integration time.

In a more realistic situation, the observer will not have a perfectly complete model of the sky. A common method is to instead model the brightest sources in the field, or even a single dominant source. When using the single point source model, all calibration methods in Fig. 8 trend downward until they reach an error floor due to the confusion sources that were not modeled. The EPICal and visibility-based solu-
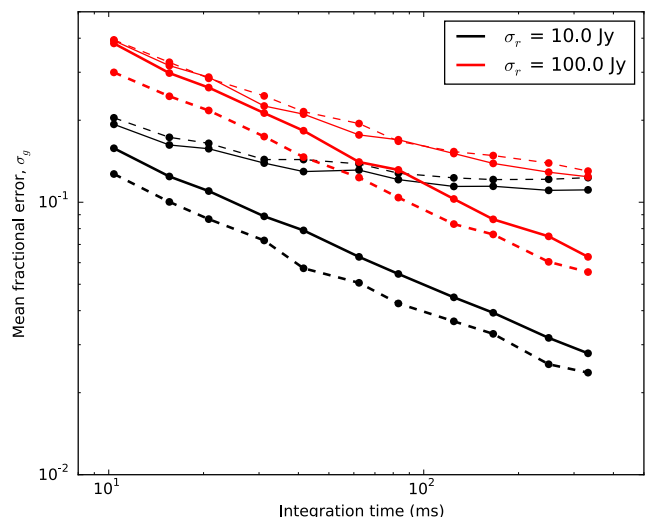
Figure 8. Gain estimate errors as a function of integration time, receiver noise (line color), and calibration method. The solid lines represent the error on the EPICal derived gain estimates, while the dashed lines represent the error on the visibility based estimates. The thicker lines were derived using a full sky model in the calibration loops, and the thinner lines used only a single point source model. In the full sky model case we see all methods trend down with longer integration time as expected. For a given integration time, EPICal derived errors are on average 23% higher than those of visibility based calibration. In the case of a single source model, the errors bottom out due to the flux in the measurements that is not modeled. In this regime the EPICal method performs slightly better than the visibility-based method due to effectively beamforming to the sky pixel where the incomplete model is most accurate. On average EPICal errors are 5% lower in this regime. The results from the high and low receiver noise regimes are similar, with both approaching the same floor in the single source model case.

tions reach a similar floor, but EPICal achieves a marginally lower level (on average 5% lower). This can be attributed to the same reason EPICal underperformed with the full sky model: EPICal only used a single pixel in the sky to form its solutions. When using the full sky model, it was under utilizing the information in the rest of the sky. But when the sky model only contains the bright point source, the single pixel used in EPICal is the location where this model is most accurate, effectively down weighting the pixels with incomplete sky model. In contrast, the visibility-based calibration method does not have any preference to a certain direction, and weights the sky equally (beyond the inherent response of the instrument). This results in a more harsh penalty for missing sources in the model. In the regime typical of HERA discussed above, EPICal's error floor was 10% lower than that of visibility-based solutions.

Any realistic sky model will lie somewhere between the two extremes explored here. In the case where the sky is modeled by a single source, as is often the case for an initial calibration, EPICal actually achieves smaller gain errors compared to the traditional visibility-based calibration. As the sky model improves, both visibility-based and EPICal gains improve, though the former quicker than the latter. In the limit of a perfect model, both calibration methods produce noise-like errors in their gains which scale down with more integration time.

Comparing the two regimes of receiver noise levels (red versus black lines in Fig. 8), we see that the gains errors behave as one might expect given our above interpretations. In the case of a perfect sky model, the higher receiver noise results in larger gain errors for a given integration time, but follows the same $t^{-1/2}$ noise trend. The imperfect model lines begin exhibit the same offset at low integration time because the errors are dominated by the thermal noise. But as integration time increases, they approach the same floor due to the same level of power from unmodeled sources.

5 APPLICATION TO LWA DATA

We next demonstrate our calibration algorithm using an observation from the LWA station in New Mexico (Ellingson et al. 2013). The data is from the LWA narrow-band transient buffer (TBN), with time domain voltage data from the 255 core antennas within a 110 m x 100 m ellipse. The central frequency is 74.03 MHz, with a total observing bandwidth of 100 kHz, analyzed by a filter bank into 512 channels of bandwidth 195.3125 Hz each and with a post-channelization sampling period of 5.12 ms. For this demonstration we limit ourselves to a single polarization.

For the gridding step, we adopt a grid with extent 120 m to encompass the footprint of the array, and a grid spacing of half wavelength (≈ 2 m). The result is a 64×64 pixel grid, making the MOFF computational complexity of $N_g \log_2 N_g$ on the same order as the N_a^2 computations for a cross correlator. Our gridding kernel is a square tophat function with sides of 3 m, corresponding to the size of the LWA ground screens.

After correcting for geometric cable delays, the instrument is sufficiently calibrated to produce recognizable images of the sky, as was seen in the demonstration of the EPIC imager in Thyagarajan et al. 2017 (their section 4.4).

However, we will aim to further improve on this calibration using our algorithm.

We proceed by forming model visibilities. We model only two bright objects as point sources: Cyg A with flux 20,539 Jy, and Cas A with flux 19,328 Jy (Lane et al. 2012). Because the raw data are attenuated by the primary beam of the instrument, we also account for this in our model using beam values consistent with Hicks et al. (2012).

We made several choices while studying the behavior of the LWA data to improve our calibration. Through our previous imaging work, we noted that the flux scale of uncalibrated images was consistent with average gain amplitudes of roughly 0.25. To allow the calibration to converge quickly we initialized our gain estimates at this level. We also found a boost in signal to noise ratio is achieved easily by assuming the gains are constant across a range of frequencies, and averaging solutions across channels. Here we average solutions across the central 300 channels, or about 58.6 kHz, and discard the remaining channels which are near the edge of the observational band and have significantly different gain amplitude. With a fractional bandwidth $B/f_0 = 7.9 \times 10^{-4} \ll 1$, we assume a smooth bandpass across the band. A damping factor of $\gamma = 0.7$ was adopted. We found a larger damping factor, relative to what was used in simulations, was beneficial with the LWA data. Several factors likely contribute to the difference in damping factors, including increased thermal noise in the LWA data compared to the simulations, as well as different antenna designs, array configurations, and sky signals. A larger damping factor yields a larger effective integration time and thus lower noise on each calibration solution (Eq. 20). We save for future work an investigation of robust determination of optimal damping factors.

Finally, we found that seven antennas[5] produced unstable gain solutions, and in fact corrupted the entire array. We therefore omitted these antennas from our analysis, resulting in a total of 248 antennas to calibrate. In each calibration loop, we form $C_{i,\hat{s}_0}, i = 1, 2, \ldots N_a$ and update our gain estimates over 10 samples (51.2 ms). We iterate the loop 30 times for a total of 1.536 seconds of data processed. The results of this calibration experiment are shown in Figs. 9 – 11.

Figure 9 shows the phase of our gain estimates over 30 calibration iterations, again with each colored line representing a different antenna. Given the quality of uncalibrated image demonstrated in Thyagarajan et al. (2017), we had anticipated that the phase solutions would not exhibit much variation between antennas. Hence the relatively large variation was unexpected. However, the phases are relatively flat after about 15 iterations (modulo noise), and exhibit a central "trunk" where the majority of phases are congregated. This behavior is suggestive that while the uncalibrated data were able to produce a viable image, the minor changes from our solutions will focus the image and improve the quality. The actual location of the "trunk" (slightly negative) is simply determined by the reference antenna chosen to have identically zero phase, but happens to be slightly more positive than the bulk of antennas.

---

[5] LWA antenna IDs 48, 85, 124, 148, 203, 217, and 244 were omitted.
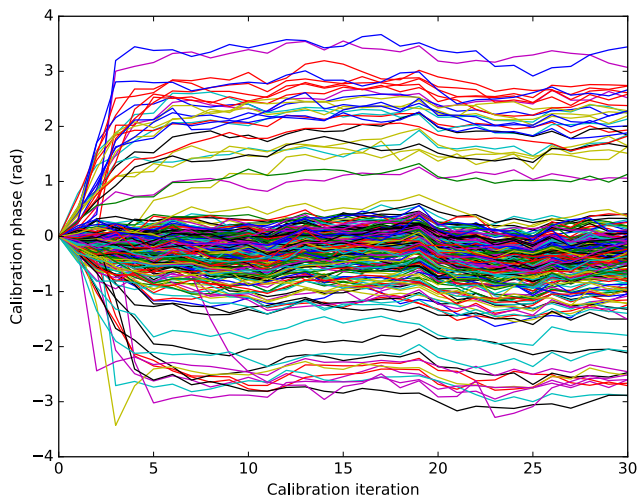
Figure 9. Gain phase solutions as a function of calibration iteration for an LWA TBN observation. The gain estimates are initialized with zero phase, but quickly span a 2π range, and settle into relatively flat, albeit noisy, solutions. The majority of phases congregate near zero, which is not surprising given the fairly good quality image produced from uncalibrated data. For plotting clarity, we unwrapped phases resulting in phases that appear to exceed $\pm\pi$.
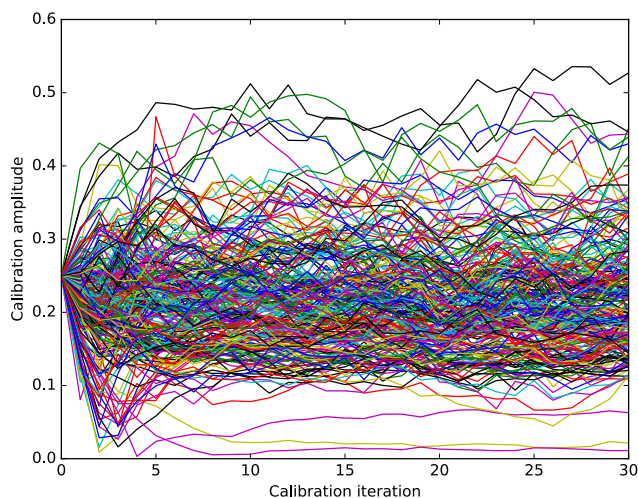
Figure 10. Same as Fig. 9, but for gain amplitudes. The gains estimates were initialized with amplitude 0.25 after inspection of the electric field values compared to our model sky. The solutions are noisy, but flat. The range in amplitudes is due to the non-uniformity of cables between the LWA antennas and receivers.

The gain amplitudes as a function of calibration iteration are shown in Fig. 10. To ensure the gain solutions and noise levels are reasonable, we compared to solutions found with a visibility-based approach. We first form visibilities by directly correlating antenna electric field measurements within each calibration iteration time interval. We then solve for visibility-based gains for each antenna and calibration iteration by minimizing χ^2 in equation 21. Because the field

has significant diffuse structure from the Galactic plane, we omit baselines shorter than 10 wavelengths. We then use an initial gain estimate of 0.25 (as we did above for EPICal), and apply the damping factor at each successive calibration iteration to achieve a weighted average analogous to the EPICal solutions. The resulting gain phases spanned the full 2π range, exhibiting the same "trunk" we see in Fig. 9. The gain amplitudes spanned a wide range similar to our EPICal results (from 0.004 to 0.68).

We next compare the noise levels between our EPICal solutions and visibility-based solutions. We estimate the fractional gain noise for each method by calculating the standard deviation of the gain estimates after iteration 15 for each antenna separately, and dividing by the respective mean gain amplitudes. For the EPICal solutions we found a fractional gain noise of 0.15, and for the visibility-based solutions we found 0.12. We conclude that in this demonstration, the fluctuations seen in Figures 9 and 10 are about 25% larger than we see for a visibility-based method, which is consistent with the noise trends we saw in Section 4 when noise dominated the error.

Figure 11 shows the improvement in the images due to our calibration. The left panel shows the uncalibrated image integrated over 51.2 ms, 58.6 kHz. Cyg A is prominent near the center of the image, and Cas A is also clearly visible in the upper right. The middle panel shows the image produced after calibration with identical integration time and bandwidth. We show each of these images with a color scale which saturates at half the maximum value in the respective image data. We find that this choice allows us to qualitatively compare the quality of the images, independent of the absolute scale which changed during the calibration. The sidelobes throughout the calibrated image are significantly suppressed, and the galactic plane is much more evident, despite only modeling Cyg A and Cas A. We also note that the feature just to the right of Cyg A is dimmer in the calibrated image, better matching the expected flux from the GSM. For reference we show the GSM in the right panel of Fig. 11, convolved by the LWA point spread function and weighted by two factors of the primary beam to match the holographic frame of the data images.

To compare the image qualities quantitatively we compute the dynamic range, defined as the peak of the image over the noise level. We estimate the noise level as the median of the absolute deviation of the image. With this metric we find the uncalibrated and calibrated images to have dynamic ranges of 80.4 and 133.1, respectively, a 65% improvement.

6 DISCUSSION

Through simulations and application to real data, we have shown that the EPICal algorithm is a viable solution for calibrating direct imaging arrays in real time. The computation necessary only scales with the number of antenna elements, making it a sub-dominant cost factor when designing the correlator. This strategy will enable fast read-out for arrays with many thousands of antennas, which will be necessary for future radio transient and cosmology experiments.

EPICal can be further improved through several extensions. Here we name a few potential considerations for fur-
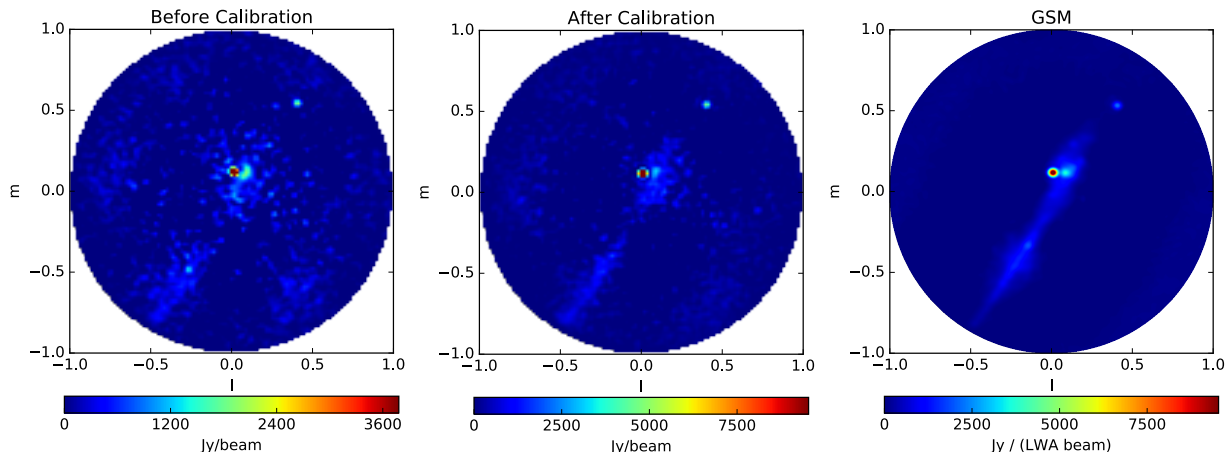
Figure 11. Images produced before (left) and after (middle) calibrating LWA data. These images were produced with 58.6 kHz bandwidth and 51.2 ms integration. The color scales are saturated at half the peak of the respective images. The calibrated image shows significant reduction in rumble throughout, while retaining the prominent Cyg A and Cas A sources. The galactic plane is also substantially more evident after calibration. For reference, the GSM is shown in the right panel convolved with the LWA point spread function and weighted by two factors of the primary beam to place it in the same holographic frame as the data images. (One beam factor is from the instrument measurement, and the other is from the gridding step of the analysis.) While much of the GSM is visible in the calibrated image, we note that the model used to calibrate was only the two bright sources, Cyg A and Cas A.

ther study. These topics can be thought of as extensions to the white "estimate gains" box in Fig. 3, which can be performed at a much slower rate than the actual antenna-pixel correlations.

Multiple pixel correlations. As was seen in section 4, the single pixel correlation demonstrated in this paper can underperform in the presence of a complex sky, when compared to visibility-based gains with perfect knowledge of the sky. These errors can be mitigated by using correlations of multiple image pixels to incorporate a higher fraction of the total sky power into the calibration loop. Of course this increases the computational cost to a scaling of $\mathcal{O}(N_\mathrm{a} N_\mathrm{pix})$, which will typically be much lower than the $\mathcal{O}(N_\mathrm{g} \log_2 N_\mathrm{g})$ of the correlator itself. These additional correlations would also enable direction dependent gain solutions. Each correlation could be used to independently solve for gains in each pixel direction, then fit to a beam model on the sky. This updated beam pattern would then feed into the gridding step of the correlator, allowing the imager to convolve the signals with the effective beam pattern in the ground plane.

Fitting gain models. With some knowledge of the instrumental bandpass, the noise of the gain solutions can be greatly reduced by fitting a model to the per-frequency solutions derived here. We used this method at a rudimentary level in our demonstration to LWA data by assuming the gains were constant over a narrow bandwidth. One could easily improve on this by extending the bandwidth and fitting for a low order polynomial in phase and amplitude. Additionally, with knowledge of any filters applied to the data stream in the receiver chain, we could include channels closer to the edge of the band.

Improving sky model. In this work we used a perfect sky model (for simulation), or a very simple sky model (for LWA data). In principle the images produced by the EPIC correlator can be used to improve the sky model used in calibration – similar to a major loop in self-calibration. This can be especially useful for compact, widefield arrays

which require a model of both compact and diffuse sources over a large patch of sky. As with any sky-based calibration scheme, EPICal requires that the sky model be attenuated by the primary beam, which can be difficult to measure at the precision necessary (e.g. Neben et al. 2015; Virone et al. 2014; Thyagarajan et al. 2015). However, the direct imaging correlator provides exactly the product of the sky and beam necessary, and can be iterated over to improve the images and gain solutions, leaving final beam correction to post-processing as in traditional processing.

Dynamic parameters. In our controlled experiments we fine-tuned a number of parameters based on our testing (e.g. damping factor, integration time, frequency averaging). A deployed system will require robust determination of these parameters to operate continuously. The specifics will be heavily dependent on the stability of the instrument, the frequency of observation, and the sky. For example, a stable instrument may be able to use short integrations to determine a rough estimate of the gains before switching to much longer integration (on order seconds to minutes) to highly increase signal to noise. At low frequencies or high imaging resolution, the dynamics of the ionosphere are important, and will likely drive the limit of time integration allowed.

The work here will serve as a foundation for further development. We have shown that the EPICal algorithm produces reliable calibration solutions, and have identified several aspects to increase the scope. With next generation instruments in the planning and development stages, EPICal is poised to enable new design spaces. The software is integrated into the EPIC package and freely available to use in simulations or post-processing of data. Work is underway to port the code to GPU systems for deployment.

ACKNOWLEDGEMENTS

REFERENCES

Abbott B. P., et al., 2016a, Phys. Rev. Lett., 116, 061102
Abbott B. P., et al., 2016b, ApJ, 826, L13
Bandura K., et al., 2014, Proc. SPIE, 9145, 914522
Beardsley A. P., et al., 2012, MNRAS, 425, 1781
Bhatnagar S., Cornwell, T. J. Golap, K. Uson, J. M. 2008, A&A, 487, 419
Bowman J. D., et al., 2013, Publ. Astron. Soc. Australia, 30, 31
Bunton J. D., 2004, Experimental Astronomy, 17, 251
Clark B. G., 1999, in Taylor G. B., Carilli C. L., Perley R. A., eds, Astronomical Society of the Pacific Conference Series Vol. 180, Synthesis Imaging in Radio Astronomy II. p. 1
Cornwell T. J., Golap K., Bhatnagar S., 2008, IEEE Journal of Selected Topics in Signal Processing, 2, 647
Daishido T., et al., 2000, Proc. SPIE, 4015, 73
Ellingson S. W., et al., 2013, IEEE Transactions on Antennas and Propagation, 61, 2540
Foster G., Hickish J., Magro A., Price D., Zarb Adami K., 2014, Monthly Notices of the Royal Astronomical Society, 439, 3180
Hicks B. C., et al., 2012, Publications of the Astronomical Society of the Pacific, 124, pp. 1090
Jacobs D. C., Beardsley A. P., Hazelton B. J., Thyagarajan N., Trott C., 2015, HERA Memo 10, System Temperatures of 21cm Arrays
Kraus J. D., 1986, Radio Astronomy, 2 edn. Cygnus-Quasar Books
Lane W. M., Cotton W. D., Helmboldt J. F., Kassim N. E., 2012, Radio Science, 47, RS0K04
Lonsdale C. J., Doeleman S. S., Cappallo R. J., Hewitt J. N., Whitney A. R., 2000, in Butcher H. R., ed., Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series Vol. 4015, Radio Telescopes. pp 126–134
Lorimer D. R., Bailes M., McLaughlin M. A., Narkevic D. J., Crawford F., 2007, Science, 318, 777
Mellema G., et al., 2013, Experimental Astronomy, 36, 235
Mersereau R. M., 1979, IEEE Proceedings, 67, 930
Mitchell D., Greenhill L., Wayth R., Sault R., Lonsdale C., Cappallo R., Morales M., Ord S., 2008, Selected Topics in Signal Processing, IEEE Journal of, 2, 707
Morales M. F., 2011, PASP, 123, 1265
Morales M. F., Matejek M., 2009, MNRAS, 400, 1814
Napier P. J., 1999, in Taylor G. B., Carilli C. L., Perley R. A., eds, Astronomical Society of the Pacific Conference Series Vol. 180, Synthesis Imaging in Radio Astronomy II. p. 37
Neben A. R., et al., 2015, Radio Science, 50, 614
Otobe E., et al., 1994, PASJ, 46, 503
Parsons A. R., et al., 2010, The Astronomical Journal, 139, 1468
Price D. C., 2016, preprint, (arXiv:1607.03579)
Romney J. D., 1999, in Taylor G. B., Carilli C. L., Perley R. A., eds, Astronomical Society of the Pacific Conference Series Vol. 180, Synthesis Imaging in Radio Astronomy II. p. 57
Rowlinson A., et al., 2016, MNRAS, 458, 3506
Sullivan I. S., et al., 2012, The Astrophysical Journal, 759, 17
Tegmark M., 1997a, Phys. Rev. D, 55, 5895
Tegmark M., 1997b, ApJ, 480, L87
Tegmark M., Zaldarriaga M., 2009, Phys. Rev. D, 79, 083530
Tegmark M., Zaldarriaga M., 2010, Phys. Rev. D, 82, 103501
Thompson A. R., Moran J. M., Swenson Jr. G. W., 2001, Interferometry and Synthesis in Radio Astronomy, 2nd Edition. Wiley
Thornton D., et al., 2013, Science, 341, 53
Thyagarajan N., et al., 2015, ApJ, 807, L28
Thyagarajan N., Beardsley A. P., Bowman J. D., Morales M. F., 2017, Monthly Notices of the Royal Astronomical Society, 467, 715
Tingay S. J., et al., 2013, PASA - Publications of the Astronomical Society of Australia, 30
Tingay S. J., et al., 2015, AJ, 150, 199
Trott C. M., Tingay S. J., Wayth R. B., 2013, ApJ, 776, L16
Virone G., et al., 2014, Antennas and Wireless Propagation Letters, IEEE, 13, 169
Wijnholds S., van der Veen A.-J., 2009, Signal Processing, IEEE Transactions on, 57, 3512
Zheng H., et al., 2014, MNRAS, 445, 1084
de Vos M., Gunst A., Nijboer R., 2009, Proceedings of the IEEE, 97, 1431
van Haarlem M. P., et al., 2013, A&A, 556, A2

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