

Sky-based Calibration Procedures for FHD Using a Comprehensive Signal Path Framework

N. Barry and the UW team

20 September 2017

ABSTRACT

Precision calibration is needed for Epoch of Reionization (EoR) analysis, and many different varieties of calibration have been studied. However, in order to facilitate community discussions, there must be a common language across multiple calibration procedures. We present a signal path framework that is both simple and comprehensive which can be used by many different calibration philosophies. We then use this framework to detail all assumptions, approximations, and areas of focus for sky-based calibration used by Fast Holographic Deconvolution (FHD), a pipeline created for Murchison Widefield Array (MWA) EoR analysis.

1 INTRODUCTION

Widefield interferometry has required new calibration techniques that differ from typical radio astronomy methods. Developing these techniques is ongoing, and almost every EoR limit paper discusses new methodologies (Paciga et al. 2013; Dillon et al. 2014; Parsons et al. 2014; Dillon et al. 2015; Ali et al. 2015; Jacobs et al. 2016; Beardsley et al. 2016; Patil et al. 2017). However, the large breadth of calibration development has resulted in complicated and diverse mathematical formulation among the community.

We present a simple framework that can be used to describe many different calibration techniques, following a scheme used in Hamaker et al. (1996) that was very successful in describing instrument beams. This is then applied to FHD sky calibration techniques as an example, and we hope to generate the community’s interest in making a memo series using this new framework.

FHD sky calibration was created with the first iteration of the MWA in mind. There was no redundancy in the placement of the MWA antennas, thus all calibration needed to come from an external source. However, typical isolated calibrator sources did not provide the needed accuracy, and communication satellites were too sparse and unreliable. Therefore an in-situ calibration procedure was developed, where *all* sources were used to calibrate in the same field of view and at the same time as the measurement.

Sky calibration calculates instrumental parameters from comparisons between the measured sky and a model of the sky in visibility space. The sky model consists of all reliable sources seen by instrument, including sources in the sidelobes. We explain how we use this sky model to calibrate using a generalized framework. We describe our assumptions, our methodology, and our limitations in hopes that a common language can be developed between the different calibration procedures.

2 SIGNAL PATH

The sky signal is modified during the journey from when it was emitted to when it was recorded. Whether or not a particular con-

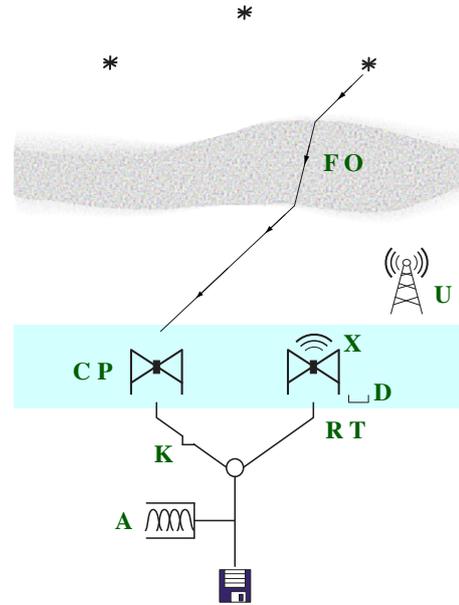


Figure 1. The signal path through the instrument. There are three categories of signal modification: before antenna, at antenna (colored light blue), and after antenna. We describe all effects labeled for FHD sky-calibration of the MWA, but we focus mainly on effects that occur after the antenna in this memo.

tribution along the signal path is described as part of the calibration process depends on the pipeline, so for clarity we describe all effects here. We define all potential known signal modifications to build a mathematical framework.

There are three main types of signal modification: those that occur before, during, and after interaction with the antenna elements. We adhere to notation from Hamaker et al. (1996) whenever possible in our brief catalog of interactions.

I: Before antenna

- Faraday rotation from interaction with the ionosphere, **F**.
- Source position offsets **O** due to variation in ionospheric thickness.

Type	Variables	Definition
I	F	Faraday rotation
	O	Source position offsets
	U	RFI time correlations
S	P	Parallactic rotation
	X	Cross-talk antenna correlations
	C	Nominal configuration
	D	Errors in the nominal configuration.
E	R	Electronic gain amplitude and phase
	T	Temperature amplitude changes
	K	Cable reflections
	A	Frequency correlations
V	\mathbf{V}^{true}	Emitted visibilities
	\mathbf{V}^{meas}	Measured visibilities
	noise	Independent thermal noise

Table 1. Brief definitions of the variables used within the signal path framework, organized by type. There are four types, **I**) interactions that occur before the antenna elements, **S**) interactions that occur at the antenna elements, **E**) interactions that occur after the antenna elements, and **V**) visibility-related variables.

- Time correlations caused by unchecked RFI, **U**.

S: At antenna

- Parallax rotation **P** between the rotating basis of the sky and the basis of the antenna elements.
- Antenna correlations from cross-talk between antennas, **X**.
- Nominal configuration of the antenna elements and the interferometer as a whole, **C**, usually referred to as the beam.
- Errors in the expected nominal configuration, **D**¹.

E: After antenna

- Electronic gain amplitude and phase **R** from a typical response of each antenna.
- Gain amplitude changes from temperature effects **T** on amplifiers.
- Gain amplitude and phase oscillations **K** due to cable reflections, both at the end of the cables and at locations where the cable is kinked.
- Frequency correlations **A** caused by aliasing in polyphase filter banks.

Each contribution can be modeled as a matrix which depends on $[f, t, PP', AB]$: frequency, time, antenna polarization products, and antenna cross-correlations. The expected contributions along the signal path are thus $\mathbf{I} = \mathbf{UOF}$, $\mathbf{S} = \mathbf{DCXP}$, and $\mathbf{E} = \mathbf{AKTR}$. The measurement equation takes the form

$$\mathbf{V}^{\text{meas}} \sim \mathbf{E} \mathbf{S} \mathbf{I} \mathbf{V}^{\text{true}} + \text{noise}, \quad (1)$$

where \mathbf{V}^{meas} are the measured visibilities, \mathbf{V}^{true} are the emitted visibilities, **I** are contributions that occur between emission to the ground, **S** are contributions from the antenna configurations, and **E** are contributions from the electronic response. We have included all known contributions and modifications to the signal. However, there are most likely contributions that we did not include, hence we describe Equation 1 as an approximation.

¹ We deviate somewhat from Hamaker et al. (1996) here. Any polarization correlation from Hamaker et al. (1996) was counted as an error in the nominal configuration, whereas we expect and model polarization correlation in the Jones matrices.

3 ASSUMPTIONS

We will use our generalized framework described in §2 to detail the assumptions and sky-based calibration methods for the FHD pipeline on the MWA. There are two classes of assumptions: those made on the model sky (incorporating **I** and **S**) and those made on the electronic response (**E**). Our calibration procedure focuses mostly on the electronic response. However, we briefly describe our model-based assumptions for completeness.

3.1 Model-based Assumptions

We build model visibilities from our knowledge of the sky to estimate the true visibilities in Equation 1. This includes as many reliable sources as possible in the primary lobe and sidelobes, along with configuration effects like the beam. Our model visibilities capture some effects in the signal path, essentially grouping the model as $\mathbf{M} \sim \mathbf{S} \mathbf{I} \mathbf{V}^{\text{true}}$. Even though we don't classify this as calibration within our own pipeline, we briefly describe below the known contributions to the model visibilities using the signal path framework.

F: We currently calibrate only on Stokes I sources which are unaffected by Faraday rotation **F**.

O: Deconvolution builds a model of the sky given calibrated visibilities and beam models, which can be fed back into a loop where the generated model is used in lieu of the pre-defined calibration catalog. This will help capture source position offsets **O** caused by varying ionosphere. For faster processing, a pre-defined catalog is usually used, where observations with substantial **O** effects from varying ionosphere are flagged.

U: We use the package `AOFLAGGER`² to flag line-like RFI contributions (Offringa et al. 2010, 2012). There have been indications of TV-related broadband RFI contributions, and we currently remove them by flagging the entire observation.

P: We utilize basis transformations to account for parallactic rotation.

X: There have been no indications of cross-talk between the antenna tiles for the MWA. We assume that it does not contribute.

C: The nominal configuration is modelled in depth. We currently use an antenna model that includes mutual coupling and embedded element effects. See Sutinjo et al. (2015) for more information.

D: The main error to the nominal configuration that we have found is the tendency for dipoles to short out due to the acidic and pervasive nature of Murchison dust in combination with rain. We can model antenna tiles with the exclusion of dead dipoles, though we have found that it is not a significant contribution yet. We do not model delay-line errors, which would contribute to polarization correlation.

3.2 Electronic Response Assumptions

We are left with one type of modification to the signal that has not been accounted for by the model visibilities: **E**, the electronic response. This is what we classify as our calibration, and we will go into much greater detail on how the electronic response is estimated.

The electronic response **E** varies slowly with time, and thus does not change significantly over a two minute observation. Due to our model-based assumptions, we do not have any time correlations, antenna correlations, or unknown polarization correlations. The

² <https://sourceforge.net/p/aoflagger/wiki/Home/>

Type	Variables	Definition
Observation Parameters	f	Measured frequencies of the observation.
	P	Instrumental polarizations of the array $\{X, Y\}$. The product PP' is $\{XX, YY, XY, YX\}$. This observation parameter is also a set.
	a	Antennas in the array.
	t	Time steps within an observation.
Sets	$A = \{a_0, a_1, a_2, \dots, a_{127}\}$ $B = \{a_0, a_1, a_2, \dots, a_{127}\}$	All antennas of the array. A and B can be iterated separately to form cross-correlated antenna pairs.
	$C = \{c_2, c_1, c_0, \phi_1, \phi_0\}$	Coefficients of a 2 nd order amplitude polynomial and a linear phase fit across the frequency band.
	$L = \{l_{90}, l_{150}, l_{230}, l_{320}, l_{400}, l_{524}\}$	Sets of antennas associated with cable lengths from LNA to beamformer, listed in the subscripts in meters. There are 19, 31, 23, 8, 17, and 30 antennas of each type, respectively.
	$D = \{c, \tau, \phi\}$	Amplitude, mode, and phase of a cable reflection fit across the frequency band. Currently only used for l_{150} antennas.
	$T = \{\rho_{-2}, \rho_{-1}, \rho_0, \rho_{+1}, \rho_{+2}\}$	Observation timing sets based off of pointings (ρ) away from zenith, where -2 is two pointings before zenith, -1 is one pointing before zenith, etc.
	$[f, t]$	A combined set of all frequencies and times.
	Groups, Matrices, & Vectors	α_L
θ_T		A grouping of observation times, where parameters are per timing group rather than per observation time.
$\mathbf{G}_{A,P}(f_o, f_i)$		The full gain matrix for each antenna in group A per P where input and output frequencies are correlated.
$\mathbf{g}_{A,P}(f)$		A vectorized approximation of the gains \mathbf{G} for each antenna in the group A per P over frequency.
$\mathbf{m}_{AB,PP'}([f, t])$		A vector of the simulated model visibilities from a model sky with frequency-dependent beam effects for each antenna pair AB and polarization product PP' over the set $[f, t]$.
$\mathbf{n}_{AB,PP'}([f, t])$		A vector of the thermal noise for each antenna pair AB and polarization product PP' over the set $[f, t]$.
$\mathbf{v}_{AB,PP'}([f, t])$		A vector of the uncalibrated data visibilities for each antenna pair AB and polarization product PP' over the set $[f, t]$.
Functions	$\mathcal{P}()$	A polynomial fit as a function of frequency of the input.
	$\mathcal{R}\langle \rangle$	A resistant mean of the input vector over an antenna set L (and optionally a time group θ_T). Outliers beyond 2σ are excluded in the average.
DFT	k	The Nyquist frequency index of the Fourier dual of frequency.
	κ	The hyperfine sub-Nyquist frequency index of the Fourier dual of frequency, with resolution at $\frac{1}{20}$ th of k .
	n	The index of the frequency.
	N	The total number of frequency channels.
	τ_κ	The Fourier dual of frequency: a timing delay in the detection of the waveform between one antenna and another. The κ index indicates it runs over the hyperfine index.

Table 2. Definitions of the variables used within this memo, organized by type.

electronic response is thus simply a time-independent gain \mathbf{G} per observation which is independent per antenna and polarization.

We begin by rewriting Equation 1 using these assumptions. The measured cross-correlated visibilities are a function of frequency, time, and polarization. The sets of antennas $A = \{a_0, a_1, a_2, \dots, a_{127}\}$ and $B = \{a_0, a_1, a_2, \dots, a_{127}\}$ are iterated through independently to calculate these cross-correlations. A visibility is measured for each linear polarization $P = \{X, Y\}$ and $P' = \{X, Y\}$, creating the polarization product, $PP' = \{XX, XY, YX, YY\}$.

The resulting relation between the measured visibilities and the model visibilities is

$$\mathbf{v}_{AB,PP'}([f_o, t]) \sim \mathbf{G}_{A,P}(f_o, f_i) \mathbf{G}_{B,P'}^*(f_o, f_i) \mathbf{m}_{AB,PP'}([f_i, t]) + \mathbf{n}_{AB,PP'}([f_o, t]), \quad (2)$$

where $\mathbf{v}_{AB,PP'}([f_o, t])$ are the measured visibilities and $\mathbf{m}_{AB,PP'}([f_i, t])$ are the model visibilities. Both are frequency and time vectors $[f, t]$ of the visibilities over all A and B antenna pairs and over all P and P' polarization products. $\mathbf{G}_{A,P}(f_o, f_i)$ is a frequency matrix of gains given input frequencies f_i which affect multiple output frequencies f_o for instrumental polarization P for antenna A (and likewise for B). Thermal noise \mathbf{n} is independent for each visibility.

Our notation has been specifically chosen. Naturally discrete variables, antenna pairs and polarization products, are described in the subscripts. Naturally continuous variables, frequency and time, are function arguments. We group frequency and time into a set $[f, t]$ for the visibilities to create vectors. Since frequency and time are independent, this notation is more compact. In contrast, the gain matrices \mathbf{G} are not independent in frequency. A full matrix

must be used to accurately capture correlation due to aliasing in the polyphase filter banks.

However, we make the assumption the frequencies are independent (contribution $\mathbf{A} \rightarrow \mathbb{I}$) to reduce Equation 2 significantly

$$\mathbf{G}_{A,P}(f_o, f_i) \rightarrow \text{diag}(\mathbf{g}_{A,P}(f)). \quad (3)$$

The instrumental gains \mathbf{g} are now an independent vector of frequencies for antenna A (and likewise for B) per instrumental polarization P . We flag frequency channels which are most affected by this aliasing to make this assumption viable. However, this does introduce contamination in confined areas in the power spectrum due to the regularity of the flagging.

We can fully vectorize the variables in Equation 2:

$$\begin{aligned} \mathbf{v}_{AB,PP'}([f, t]) \sim \\ \text{diag}(\mathbf{g}_{A,P}(f)) \text{diag}(\mathbf{g}_{B,P'}^*(f)) \mathbf{m}_{AB,PP'}([f, t]) \\ + \mathbf{n}_{AB,PP'}([f, t]). \end{aligned} \quad (4)$$

The gains in Equation 2 and Equation 4 will encode systematics between the model visibilities and the true visibilities due to the imperfection of the model. This is not instrumental, thus the calibration solutions will be contaminated. We attempt to remove some of this effect in later sections based off of work done in Barry et al. (2016). Our model currently does not include large-scale diffuse emission, so we constrain our calibration to be calculated only from visibilities larger than 50 wavelengths to reduce overfitting effects (Patil et al. 2016).

4 PER FREQUENCY FIT

Equation 4 can be used to solve for the instrumental gains for all frequencies and polarizations independently. This allows the use of Alternating Direction Implicit (ADI) methods for fast and efficient solving of $\mathcal{O}(N^2)$ (Salvini & Wijnholds 2014). Due to this independence, parallelization can also be applied. Noise is ignored during the ADI for simplicity; true thermal noise is Gaussian and will average out during the ADI. However, if any noise is non-Gaussian, it will contribute to the instrumental gains.

We begin solving Equation 4 by estimating an initial solution for $\mathbf{g}_{B,P'}^*(f)$. This choice must force the gains into a region with a local minimum for the ADI method. A good choice is the average gain expected across all antennas, and in practice a set of 1's has been sufficient for the MWA. For a more complicated method, scaled autocorrelations have also been used.

With an input for $\mathbf{g}_{B,P'}^*(f)$, Equation 4 can then become a linear least-squares problem.

$$\begin{aligned} \chi_{A,P}^2([f, t]) = \sum_B \left| \mathbf{v}_{AB,PP'}([f, t]) \right. \\ \left. - \text{diag}(\mathbf{g}_{A,P}(f)) \text{diag}(\mathbf{g}_{B,P'}^*(f)) \mathbf{m}_{AB,PP'}([f, t]) \right|^2, \end{aligned} \quad (5)$$

where $\mathbf{g}_{A,P}(f)$ is found given a minimization of $\chi_{A,P}^2([f, t])$ for each antenna A and instrumental polarization P . All time steps are used to find the temporally-constant gains over the observation. We have also assumed $PP' \rightarrow PP$ (resulting in $\{XX, YY\}$ only) for computation efficiency since these contributions are most significant. A full polarization treatment is being developed.

The current estimation of $\mathbf{g}_{B,P}^*(f)$ is then updated with knowledge from $\mathbf{g}_{A,P}(f)$ by adding together the current and new estimation and dividing by two. By updating in partial steps, a smooth

convergence is ensured. The linear least-squares process is then repeated with an updated $\mathbf{g}_{B,P}^*(f)$ until convergence is reached.³

5 BANDPASS

The resulting gains from the least-squares iteration process are completely independent in frequency, antenna, time, and polarization. However, this is not a completely accurate representation of the gains. It was necessary to make this assumption for the efficient solving technique in §4, but we can incorporate our prior knowledge of the nature of the instrument and its spectral structure *ex post facto*.

For example, we did not account for noise contributions during the per-frequency ADI fit; this adds spurious deviations from the gain's true value with mean of zero. In addition, we did not account for imperfections in the model; this adds spectral structure deviations which depend on LST and the UV plane. We reduce both by averaging the normalized full band structure of the gains.

For the MWA, sets of antennas experience the same attenuation as a function of frequency due to cable types, cable lengths, and whitening filters. We group these antennas into the set $L = \{l_{90}, l_{150}, l_{230}, l_{320}, l_{400}, l_{524}\}$ where the subscript of l is the length of the cable type. Noise and spectral structure from unmodelled sources and their point spread functions will differ antenna-to-antenna, and thus the average will decrease these unwanted effects. We get

$$|\mathbf{g}_{L,P}(f; \alpha_L)| = \mathcal{R} \langle |\mathbf{g}_{A \in L, P}(f)|, 2\sigma \rangle, \quad (6)$$

where \mathcal{R} is the resistant mean function that calculates the distribution of the amplitudes of a similar antenna set α_L for each frequency and polarization and then calculates the mean of that distribution after Gaussian 2σ outliers have been excluded. We choose the resistant mean because outlier contributions are more reliably reduced than median calculations. The variable change $\mathbf{a} \rightarrow \alpha$ indicates one parameter per group of antennas.

If more observations are available, we follow a similar averaging process over time. The MWA is very stable in time, and a normalized bandpass per antenna should be nearly identical from one time to the next, excluding potential Van Vleck corrections and noise contributions. Gains from different LSTs will have different spectral structure from unmodelled sources, and thus an average will remove even more of this effect.

We create a time set of $T = \{\rho_{-2}, \rho_{-1}, \rho_0, \rho_{+1}, \rho_{+2}\}$ where times are grouped by how many pointings⁴ they are away from zenith. Whenever possible, we use

$$|\mathbf{g}_{L,P}(f; \alpha_L, \theta_T)| = \mathcal{R} \langle |\mathbf{g}_{A \in L, P}(f)|, 2\sigma, \theta_T \rangle, \quad (7)$$

where θ_T runs over observations within a pointing and over as many days as possible. The variable change $t \rightarrow \theta$ indicates one parameter per group of times.

6 LOW ORDER POLYNOMIALS

An overall amplitude still must be accounted for within the gains. These vary slowly per antenna and differ from day-to-day, therefore

³ We have found that allowing the first 10 iterations to only update the phase of $\mathbf{g}_{B,P}^*(f)$ helps to converge faster.

⁴ A pointing defines a group of observations with the same electronic delay. As the sky rotates throughout the night, different electronic delays are used to roughly point the instrument to the same location in the sky.

they cannot be included in the average bandpass. We have found that temperature is a main cause in the amplitude variations due to the temperature-dependence of the amplifiers, which we have described as \mathbf{T} in our signal path framework. This dependence varies slightly as a function of frequency, and is easily characterized with a low-order polynomial.

In addition to fitting polynomials to the amplitude as a function of frequency, we also fit the phase. The phase is extremely linear for the MWA. We have found that using a per-antenna linear fit as the calibration phase solution is a good estimate. However, averaging similar to §5 could theoretically be done. Characterizing small fluctuations in the phase will be addressed when that level of accuracy is achieved.

We fit polynomials as a function of frequency for each observation to account for these overall variations. For the amplitude, we fit

$$c_2 f^2 + c_1 f + c_0 = \mathcal{P} \left(\frac{|\mathbf{g}_{A,P}(f)|}{|\mathbf{g}_{L,P}(f; \boldsymbol{\alpha}_L, \theta_T)|} \right), \quad (8)$$

where the bandpass contribution, $|\mathbf{g}_{L,P}(f; \boldsymbol{\alpha}_L, \theta_T)|$, is removed before the fit and c_2 , c_1 , and c_0 are the resulting coefficients. For the phase, we fit

$$\phi_1 f + \phi_0 = \mathcal{P} \left(\arg \frac{\mathbf{g}_{A,P}(f)}{|\mathbf{g}_{L,P}(f; \boldsymbol{\alpha}_L, \theta_T)|} \right), \quad (9)$$

where the polynomial fit is done over the phase of the residual and ϕ_1 and ϕ_0 are the resulting coefficients. Due to phase jumps between $-\pi$ and π , special care is taken to ensure the function is continuous across the π boundary⁵. We can create a set of these coefficients, $C = \{c_2, c_1, c_0, \phi_1, \phi_0\}$, for easy reference.

7 CABLE REFLECTIONS

Reflections due to a mismatched impedance between cables must also be accounted for within the gains (contribution \mathbf{K} in the signal path framework). The MWA has cables of various lengths, and while they all meet engineering specifications, the reflection signal is still orders of magnitude above the EoR and is different for each antenna. Averaging the gains across antennas and time in §5 effectively removed cable reflections from the solutions, thus we must specifically incorporate them. Currently we only fit for the cable reflection for the l_{150} antennas because of its prime location within the EoR window in power spectrum space.

We find the theoretical location of the potential mode using the nominal cable length and the specified light travel time of the cable. We then perform a hyperfine DFT of the gain for each observation around the theoretical mode

$$\mathbf{g}_{A,P}(\tau_\kappa) = \sum_{n=0}^{N-1} \left(\frac{\mathbf{g}_{A,P}(f_n)}{|\mathbf{g}_{L,P}(f_n; \boldsymbol{\alpha}_L, \theta_T)|} - (c_2 f_n^2 + c_1 f_n + c_0) e^{i(\phi_1 f_n + \phi_0)} \right) e^{-2\pi i \kappa \frac{n}{N}}, \quad (10)$$

where τ_κ is the delay, $[n, N] \in \mathbb{Z}$, and κ is the hyperfine index component. Typically, we set the range of κ to be $[k_{\tau_o} - \frac{1}{20}k, k_{\tau_o} +$

⁵ We "unwrap" the phase to account for this, where we take the Riemann sheets and create a new continuous plane. We then solve, and "re-wrap." There can be ambiguity in which Riemann sheet to place the phase if the phase varies quickly, but this is not an issue with the MWA.

Contribution	Variables	DoF	Total DoF
ADI method only (§4)	$f \times P \times t \times a$	101,154,816	
Bandpass (§5)	$f \times P \times T \times L$	23,040	1,531,554
Polynomials (§6)	$C \times P \times t \times a$	1,317,120	
Cable reflections (§7)	$D \times P \times t \times l_{150}$	191,394	

Table 3. The degrees of freedom (DoF) for each contribution to our calibration solutions over the 2013 season. The variables listed show each independent parameter per calculation. Table 2 defines the variables. The total DoF reduction between the ADI method and our various extra calibration procedures is over a factor of 66.

$\frac{1}{20}k]$, where k_{τ_o} is the index of the theoretical mode and k is the normal DFT index in the range of $[0, N - 1]$ (Beardsley 2015).

The maximum $|\mathbf{g}_{A,P}(\tau_\kappa)|$ around k_{τ_o} is chosen as the experimental cable reflection. The associated amplitude c , phase ϕ , and mode τ are then calculated to generate the experimental cable reflection contribution $c e^{-2\pi i \tau f + i \phi}$ to the gain for each observation. We can create a set of these coefficients, $D = \{c, \tau, \phi\}$, for easy reference.

8 CONCLUSIONS

We built a signal path framework to catalog all known contributions to the emitted signal, and then separated them into modifications on the sky model and the electronic response. This framework defines a common language to describe assumptions and methods for other pipelines and procedures. With this foundation, we then described our calibration procedure.

Our final calibration solution is

$$\mathbf{g}_{A,P}(f; \boldsymbol{\alpha}_L, \theta_T, [C], [D]) = \underbrace{|\mathbf{g}_{L,P}(f; \boldsymbol{\alpha}_L, \theta_T)|}_{\text{bandpass §5}} \underbrace{\left((c_2 f^2 + c_1 f + c_0) e^{i(\phi_1 f + \phi_0)} \right)}_{\text{per antenna gain and phase §6}} + \underbrace{c e^{-2\pi i \tau f + i \phi}}_{\text{cable reflection §7}}, \quad (11)$$

for antennas with 150 m cables, and

$$\mathbf{g}_{A,P}(f; \boldsymbol{\alpha}_L, \theta_T, [C]) = \underbrace{|\mathbf{g}_{L,P}(f; \boldsymbol{\alpha}_L, \theta_T)|}_{\text{bandpass §5}} \underbrace{(c_2 f^2 + c_1 f + c_0) e^{i(\phi_1 f + \phi_0)}}_{\text{per antenna gain and phase §6}}, \quad (12)$$

for all other antennas. The bandpass amplitude solution generated in §5 is over a set of antennas with the same cable/attenuation properties ($\boldsymbol{\alpha}_L$) and includes many observations within a pointing set (θ_T) covering many days. The same bandpass is applied to all antennas of the appropriate type and all observing times from the respective pointing. In contrast, the polynomials (§6) and the cable reflection (§7) are fit independently for each observation and per antenna. We divide the data visibilities by the applicable form of $\mathbf{g}_A \mathbf{g}_B^*$ to form our final, calibrated visibilities.

The methodology behind our calibration methods has two main pillars: efficient yet realistic. We gain efficiency by calculating gains through a linear least-squares ADI method. However, we incorporate

as many instrumental features as possible afterwards. We remove noise and unmodelled spectral structure through averaging the normalized bandpass in time and antenna sets. We account for a slowly changing gain with a 2nd order polynomial in the amplitude and a linear polynomial in phase across the frequency band. Finally, we account for an experimentally-determined cable reflection for certain antennas.

We can also represent our methodology as a reduction in the number of degrees of freedom used to calculate the gains. Table 3 shows the separate degrees of freedom used for each contribution to our gain solutions for the season of 2013, and the final total. To compare, we also show how many degrees of freedom are used to calculate the gains with just the per-frequency calibration solutions.

In our simplification of the calibration problem in §3, we created independence in the output set of $[f, P, a, t]$ that was unrealistic. Using our calibration procedures, we introduced priors in the set of $[f, P, a, t]$ that was motivated by our instrument. On-going research continues to improve our calibration as we integrate more observations and understand the finer details of our instrument.

APPENDIX: FUTURE DIRECTIONS

Active development on calibration can progress in a variety of different ways. We document them here to highlight the future directions that calibration might take.

- We currently do not include XY or YX visibilities in calculating the χ^2 between data and model visibilities. The XY or YX contributions are small in Stokes I, but including them would result in more accurate results.

- The bandpass is calculated with the resistant mean by cable type (L) and by time (T), but we might resort to averaging only in time. Currently, per-antenna spectral characteristics are being reduced during cable averaging. With the addition of more observations, these per-antenna spectral characteristics can be kept while still reducing noise and unmodelled source structure by only averaging in time.

- The amplitude polynomial fits are a strong function of temperature. The number of degrees of freedom could be reduced by calculating polynomials as a function of frequency and temperature over a set of observations, rather than calculate a new polynomial for every observation. However, we have seen that AC unit servicing will change this function of temperature. There is also little curvature across the frequency band in the amplitude, so reducing the 2nd order polynomial to a linear fit might be applicable. There also might be reason to fit linear amplitude functions separately for post- and pre-digital gain jumps due to a change in bit statistics.

- The cable reflection modes are relatively stable in time. Averaging could be done as a function of time to reduce the number of degrees of freedom. In the past, we had not seen improvement in the power spectrum. Finding the cable reflections for other cable lengths other than only the antennas within l_{150} is also most likely necessary.

- Autocorrelations can help with finding instrumental effects, but the noise contribution is unavoidable. If the noise contribution is understood, along with the associated bit statistics, autocorrelations could be used to model antennas.

- There is a lot of research being done on weighting schemes. Currently, we only calibrate on baselines longer than 50 wavelengths, but this can be made to be more sophisticated. See Ewall-Wice et al. (2016) for more information on baseline weighting schemes.

REFERENCES

- Ali Z. S., et al., 2015, *The Astrophysical Journal*, 809, 61
- Barry N., Hazelton B., Sullivan I., Morales M. F., Pober J. C., 2016, *Monthly Notices of the Royal Astronomical Society*, 461, 3135
- Beardsley A., 2015, PhD thesis, University of Washington
- Beardsley A. P., et al., 2016, *The Astrophysical Journal*, 833, 102
- Dillon J. S., et al., 2014, *Physical Review D*, 89, 023002
- Dillon J. S., et al., 2015, *Physical Review D*, 91, 123011
- Ewall-Wice A., Dillon J. S., Liu A., Hewitt J., 2016, arXiv:1610.02689 [astro-ph]
- Hamaker J. P., Bregman J. D., Sault R. J., 1996, *Astronomy and Astrophysics Supplement Series*, 117, 137
- Jacobs D. C., et al., 2016, *The Astrophysical Journal*, 825, 114
- Offringa A. R., Bruyn D., G A., Biehl M., Zaroubi S., Bernardi G., Pandey V. N., 2010, *Monthly Notices of the Royal Astronomical Society*, 405, 155
- Offringa A. R., van de Gronde J. J., Roerdink J. B. T. M., 2012, *Astronomy & Astrophysics*, 539, A95
- Paciga G., et al., 2013, *Monthly Notices of the Royal Astronomical Society*, 433, 639
- Parsons A. R., et al., 2014, *The Astrophysical Journal*, 788, 106
- Patil A. H., et al., 2016, *Monthly Notices of the Royal Astronomical Society*, 463, 4317
- Patil A. H., et al., 2017, *The Astrophysical Journal*, 838, 65
- Salvini S., Wijnholds S. J., 2014, *Astronomy & Astrophysics*, 571, A97
- Sutinjo A., O’Sullivan J., Lenc E., Wayth R. B., Padhi S., Hall P., Tingay S. J., 2015, *Radio Science*, 50, 2014RS005517