Redundant and quasi-redundant calibration with point sources

February 10, 2021

1 Logarithmic Calibration

The amplitude calibration is written in matrix form as

$$\mathbf{d} = \mathbf{A}\,\mathbf{x} + \mathrm{Re}(\mathbf{w}), \tag{1}$$

where $\mathbf{d}$ represents the known $\ln|c_{ij}|$ values (c_{ij} are the measured visibilities for dishes i and j), $\mathbf{A}$ is the array configuration matrix and $\mathbf{x}$ is the matrix of unknowns consisting of η and $\ln|y_{i-j}|$. These matrices are given as follows,

$$\begin{pmatrix} \ln|c_{0,1}| \\ \ln|c_{1,2}| \\ \ln|c_{3,4}| \\ \ln|c_{4,5}| \\ \ln|c_{0,2}| \\ \ln|c_{3,5}| \\ \cdot \\ \cdot \\ \ln|c_{0,5}| \\ \ln|c_{2,3}| \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ & & & \cdot & & & \cdot & & & \cdot & & & \\ & & & \cdot & & & \cdot & & & \cdot & & & \\ 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \eta_0 \\ \cdot \\ \cdot \\ \eta_5 \\ y_{6,N} \\ y_{12} \\ y_{6,E} \\ y_{6\sqrt{2},NE} \\ y_{6\sqrt{2},SW} \\ y_{6\sqrt{5},NE} \\ y_{6\sqrt{5},SW} \end{pmatrix} + \begin{pmatrix} \mathrm{Re}(w_{0,1}) \\ \mathrm{Re}(w_{1,2}) \\ \mathrm{Re}(w_{3,4}) \\ \mathrm{Re}(w_{4,5}) \\ \mathrm{Re}(w_{0,2}) \\ \mathrm{Re}(w_{3,5}) \\ \cdot \\ \cdot \\ \mathrm{Re}(w_{0,5}) \\ \mathrm{Re}(w_{2,3}) \end{pmatrix}. \tag{2}$$

Note that $\mathbf{A}$ would be slightly different for the phase calibration, with the first 1 in each row changed to a negative 1.

This set of equations has no unique solution even in the noiseless case, due to a degeneracy which is broken by constraining the sum of the gain amplitudes.

It is worth noting that redundant calibration only solves for relative calibration up to the degeneracies. After solving for the relative complex gains of the antennas using redundant calibration, the absolute calibration is unknown and may be determined by other methods, such as observations of sources with known flux densities, if necessary.

The sum constraint applies to phase calibration as well. However, there are two additional constraints required for phase calibration. If we were to rotate the sky clockwise by a degree, ϕ_{rot}, and rotate the antennas by the same degree anti-clockwise, the gain phase would remain the same. We thus constrain the product of the phase and antenna position, for both x and y positions.

The least squares equation is then used to solve for the set of equations and is presented here in its familiar form,

$$\hat{\mathbf{x}} = [\mathbf{A}^t\mathbf{N}^{-1}\mathbf{A}]^{-1}\mathbf{A}^t\mathbf{N}^{-1}\mathbf{d}, \tag{3}$$

where $\hat{\mathbf{x}}$ contains the solutions to the gain amplitude and phase, which we refer to as the recovered gains. The noise covariance matrix is expressed as $\mathbf{N} = \langle \mathrm{Re}(\mathbf{w})\mathrm{Re}(\mathbf{w}^t) \rangle$ for the amplitude calibration, with the imaginary components taking the same form for the phase calibration. The constraints required to break degeneracies are then added as extra rows in the **A**, **N** and **d** matrices. We maintain diagonality of the noise matrix when adding the constraints by setting the additional rows of the matrix to be 1 along the diagonal, in order to weight the constraints.

2 Correlation Calibration

We can write a generalised form for the χ^2 which does not assume the true visibilities within a redundant block to be equal. Since we are now accounting for array perturbations, we can refer to the blocks as *quasi-redundant* blocks and write the visibility covariance as $\langle y_\alpha^* y_\beta \rangle = \mathcal{C}_{\alpha\beta}$ for α, β within a block, such that the visibility covariance matrix, $\mathbf{C}$, is calculated separately for each block. The full covariance matrix, $\mathbf{C}^{full}$ would then be a set of diagonal quasi-redundant blocks with zero covariance between visibilities in separate blocks. This is a valid assumption if the blocks are constructed correctly.

The noise is now written as $\mathbf{N} = \mathbf{N}_{vis} + \mathbf{C}$, which gives the equation for χ^2, written with gain factors,

$$\chi^2 = \mathbf{c}^T(\mathbf{N} + \mathbf{H}^T\mathbf{C}\mathbf{H})^{-1}\mathbf{c}, \tag{4}$$

where $\mathbf{H} = \mathbf{G}^{-1}$ for $G_{ij} = g_i^* g_j$. We have once more relabeled $\mathbf{N}_{vis} = \mathbf{N}$. We can rewrite the visibility covariance as $\mathbf{C} = \mathbf{S}\mathbf{S}^T + \mathbf{R}\mathbf{R}^T$ where $\mathbf{S}$ is the source vector and $\mathbf{R}$ is the quasi-redundant vector. The source vector, $\mathbf{S}$ can be used to account for very bright point sources with known positions which we are observing, and is especially useful for phase calibration. We leave the inclusion of bright sources for future work and set $\mathbf{S} = \mathbf{0}$ for the remainder of the thesis. The vector $\mathbf{R}$ is necessary for computational efficiency. Instead of including entire covariance matrices, we decompose the covariances into their eigenvalues and eigenvectors, so that $\mathbf{R} = \lambda^{1/2}\mathbf{v}$. If we wish to construct $\mathbf{R}$ in the traditional redundant calibration case, we can then use $\mathbf{C} = (\gamma\mathbf{1})^T(\gamma\mathbf{1})$ with $\gamma \to \infty$, for all redundant blocks. Thus, in the redundant CorrCal case, $\mathbf{R}$ would consist only of uncorrelated real and imaginary components: $R_r = \gamma[1, 0, 1, 0, ...]$ and $R_i = \gamma[0, 1, 0, 1, ...]$, with γ set to a large value.

The gradient of χ^2 is then taken with respect to antenna gains as

$$\nabla\chi^2 = \mathbf{c}^T(\mathbf{N} + \mathbf{H}^T\mathbf{C}\mathbf{H})^{-1}(\mathbf{H}'^T\mathbf{C}\mathbf{H} + \mathbf{H}^T\mathbf{C}\mathbf{H}')(\mathbf{N} + \mathbf{H}^T\mathbf{C}\mathbf{H})^{-1}\mathbf{c}. \quad (5)$$

Having a form for the χ^2 gradient allows for the use of efficient solvers for the gains, such as the conjugate gradient solver.

Several inputs are required for the use of CorrCal: (i) The visibility data as a single array which is separated according to the quasi-redundant blocks, and thereafter expressed in its real and imaginary components as $[r_1, i_1, r_2, i_2..]$, (ii) the diagonal elements of the noise covariance, (iii) the indices which separate the different components into their blocks, (iv) the vector $\mathbf{R}$ containing the eigenvalues and eigenvectors of the visibility covariance, (v) the source vector $\mathbf{S}$ and (vi) the antennas indices separated according to the different blocks.

In order to use CorrCal to test antennas gain recovery, we must also input simulated gains. The real and imaginary components of the gains, g are $e^{\eta} cos\phi$ and $e^{\eta} sin\phi$, respectively. We assume that $\eta, \phi = 0$ so that the simulated gains have real components of 1 and imaginary components of 0. The vector is again arranged as $[r_1, i_1, r_2, i_2..]$, with the gain recovery solution then returned in the same form. Since the simulated gains are set to unity, the gains have real/imaginary components that represent the amplitude/phase of the gains.

The χ^2 and $\nabla\chi^2$ returned from the CorrCal code are then used to obtain gain solutions using the conjugate gradient solver provided by Scipy

[1]. Another requirement for the use of the solver is an initial guess which provides a necessary starting point, after which the results are iterated over until an optimal solution is found.

3 Visibilities due to a Poisson distribution of dim and unresolved point sources

The visibility is given without explicit dependence on frequency, as

$$V(\boldsymbol{u}) = \int A(\boldsymbol{\theta}) I(\boldsymbol{\theta}) e^{-2\pi i \boldsymbol{u}.\boldsymbol{\theta}} d^2\boldsymbol{\theta}, \tag{6}$$

where $\boldsymbol{u} = \frac{\boldsymbol{b}}{\lambda}$, $A(\boldsymbol{\theta})$ is the antenna beam and the sky intensity is $I(\boldsymbol{\theta})$.

For point sources, we can use the equation for the sky brightness given as [1],

$$I = \int_0^{S_{max}} S \frac{dN}{dS} dS \quad [\text{Jy sr}^{-1}], \tag{7}$$

where S is the source flux expressed as a simple power law in frequency [1],

$$S_\nu = S_{\nu_0} \left(\frac{\nu}{\nu_0} \right)^{-\gamma} \quad [\text{Jy}] \tag{8}$$

and the source count distribution is [2]

$$\frac{dN}{dS}(\nu) = \alpha \left(\frac{S}{1 \text{ Jy}} \right)^{-\beta} \quad [\text{Jy}^{-1}\text{sr}^{-1}]. \tag{9}$$

The variables in equations 8 and 9 are obtained from point source surveys conducted at a particular frequency. The values used are: $\nu_0 \sim 150$ MHz, $\gamma = 0.8$, $\alpha_{\nu_0} = 4000$ $[\text{Jy}^{-1}\ \text{sr}^{-1}]$ and $\beta = 1.75$. In equation 7 we set an upper limit, S_{max} on the flux density of sources such that sources above this limit must be omitted and treated individually.

The value chosen for S_{max} depends on the sensitivity of HIRAX which is $\sim 12\ \mu$Jy for the full 1024 element HIRAX array in a single day [3]. Assuming a 5σ detection of point sources is required to treat these bright sources individually, we can set the $S_{max} = 5 \times 12\ \mu$Jy. For a scaled down version

[1] `https://docs.scipy.org/doc/scipy/reference/generated/scipy.optimize.fmin_cg.html/`

of the full array, we can multiply this value by $\sqrt{\frac{1024}{\mathrm{N}_{dish}}}$.

We can rewrite the visibility in equation 6 by substituting equation 7 such that

$$V(\boldsymbol{u}) = \int_{S_{min}}^{S_{max}} S\frac{dN}{dS}dS \int A(\boldsymbol{\theta})e^{-2\pi i \boldsymbol{u}.\boldsymbol{\theta}}d^2\boldsymbol{\theta} \quad [\mathrm{Jy}], \tag{10}$$

from which we can see that the two integrals can be computed separately. We can evaluate the dS integral by substituting equation 9 and then simplifying to obtain a constant value of $\frac{\alpha}{2-\beta}(S_{max}^{2-\beta} - S_{min}^{2-\beta})$.

To evaluate the second integral, let us write the expression for the Gaussian beam of antenna i as, $A_i(\boldsymbol{\theta}) = e^{-\frac{\theta_{sep}^2}{2\sigma^2}}$, where θ_{sep} is a 1D vector in radians, calculated as the separation of each pixel in a HEALPix map, with 2D angular vector, $\boldsymbol{\theta}$, from the zenith which has a 2D angular vector, $\boldsymbol{\theta}^{zenith}$. The quantity, σ, is obtained using the FWHM given as λ/D_{dish}. Assuming equal beams for all antennas, we use $A(\boldsymbol{\theta}) = A_i(\boldsymbol{\theta})A_j(\boldsymbol{\theta})$. Using this, we can write the visibility as a summation over the number of pixels, n_{pix}, as

$$V(\boldsymbol{u}) = \frac{\alpha}{2-\beta}(S_{max}^{2-\beta} - S_{min}^{2-\beta}) \left(\frac{4\pi}{n_{pix}} \sum_{n=1}^{n_{pix}} e^{-\frac{\theta_{sep,n}^2}{\sigma^2}} e^{-2\pi i \boldsymbol{u}.\boldsymbol{\theta}_n} \right). \tag{11}$$

4 Visibility covariance

We can now compute the visibility covariance for baselines a and b.

$$\begin{aligned} C_{ab} &= \langle V(\boldsymbol{u}_a)V^*(\boldsymbol{u}_b)\rangle \\ &= \langle I_\nu^2 \rangle \int A^2(\boldsymbol{\theta})e^{-2\pi i(\boldsymbol{u}_a - \boldsymbol{u}_b).\boldsymbol{\theta}}d^2\boldsymbol{\theta} \quad [\mathrm{Jy}]^2. \end{aligned} \tag{12}$$

The power spectrum for point sources following a Poisson distribution is written as

$$C^{PS,Poisson} = \langle I_\nu^2 \rangle = \int_0^{S_{max}} S^2\frac{dN}{dS}dS \quad [\mathrm{Jy}^2\ \mathrm{sr}^{-1}], \tag{13}$$

which simplifies to a constant value of $\frac{\alpha}{3-\beta}(S_{max}^{3-\beta} - S_{min}^{3-\beta})$. This is substituted into equation 12.

Writing the integral as a summation over n_{pix} again, we have

$$C_{ab} = \frac{\alpha}{3-\beta}\left(S_{max}^{3-\beta} - S_{min}^{3-\beta}\right)\left(\frac{4\pi}{n_{pix}}\sum_{n=1}^{n_{pix}} e^{-\frac{2\theta_{sep,n}^2}{\sigma^2}} e^{-2\pi i(\boldsymbol{u}_a - \boldsymbol{u}_b).\boldsymbol{\theta}_n}\right), \tag{14}$$

for a, b within a redundant block.

4.1 Visibilities due to bright and resolved point sources

We can write the visibility equation for the bright sources with flux density, $S > S_{max}$, that can be resolved by the telescope. We start off with a single source, whose sky brightness can be written as a Dirac delta function as

$$\begin{aligned} V_{singlesource}(\boldsymbol{u}) &= S_{source}\int A(\boldsymbol{\theta})\delta_D^2(\boldsymbol{\theta} - \boldsymbol{\theta}_s)e^{-2\pi i\boldsymbol{u}.\boldsymbol{\theta}}d^2\boldsymbol{\theta} \\ &= S_{source}A(\boldsymbol{\theta}_s)e^{-2\pi i\boldsymbol{u}.\boldsymbol{\theta}_s} \end{aligned} \tag{15}$$

where $\boldsymbol{\theta}_s$ is the position of the source. The absolute magnitude of all visibilities will be equal for a single source. Specifically, if we consider the source to be at the centre of the beam, such that $\boldsymbol{\theta}_s = \boldsymbol{\theta}_{zenith}$, then $A(\boldsymbol{\theta}) = 1$. The absolute magnitude of all visibilities will then be $V_{singlesource}(\boldsymbol{u}) = S_{source}$. Such a source would provide no phase information.

We code this up as

$$V_{brightsources} = \sum_{n=1}^{n_{pix}} I_n e^{-\frac{\theta_{sep,n}^2}{\sigma^2}} e^{-2\pi i\boldsymbol{u}.\boldsymbol{\theta}_n}, \tag{16}$$

where I_n is the sky brightness of each pixel, which depends on whether or not a source is contained in the pixel, such that $I_n = 0$ for pixels with no bright sources. All bright sources are simulated to occupy a single pixel so that each source individually can be approximated as a Dirac delta function, thereby returning the result from the analytic expression for a single source, given in equation 15.

References

[1] S. G. Murray, C. M. Trott, and C. H. Jordan. An improved statistical point-source foreground model for the epoch of reionization. *The Astrophysical Journal*, 845(1):7, aug 2017.

[2] Mario G. Santos, Asantha Cooray, and Lloyd Knox. Multifrequency analysis of 21 centimeter fluctuations from the era of reionization. *The Astrophysical Journal*, 625(2):575–587, Jun 2005.

[3] L. B. Newburgh, K. Bandura, M. A. Bucher, T.-C. Chang, H. C. Chiang, J.F. Cliche, R. Davé, M. Dobbs, C. Clarkson, K. M. Ganga, and et al. Hirax: a probe of dark energy and radio transients. *Ground-based and Airborne Telescopes VI*, Aug 2016.