DUAL POLARIZATION GAIN ESTIMATION FOR RADIO TELESCOPE ARRAYS

Albert-Jan Boonstra[1,2] *and Alle-Jan van der Veen*[2]

[1]ASTRON, PO Box 2, 7990 AA Dwingeloo, The Netherlands
[2]Delft Univ. of Technology, Dept. Electrical Eng./DIMES, Mekelweg 4, 2628 CD Delft, The Netherlands

The complex receiver gains and sensor noise powers of radio telescope arrays are initially unknown and need to be calibrated. This can be accomplished by observing a relatively strong astronomical point source. Here we consider calibration algorithms for the case of dual polarized arrays and formulate new Least Squares algorithms based on factor analysis, parallel factor analysis, and eigenvalue decompositions. We show that at least three sources with different polarization states are needed to obtain a unique solution for the gains. We also propose a closed form solution which has a performance comparable to the iterative parallel factor method.

1. INTRODUCTION

For unpolarized telescope arrays, a standard calibration procedure is to point the telescopes to a strong astronomical source, and to estimate a covariance matrix $\hat{\mathbf{R}}$, containing all correlation products between the telescope output signals. Asymptotically, $\hat{\mathbf{R}}$ converges to its expected value $\mathbf{R}$ which has the model $\mathbf{R} = \mathbf{g}\sigma_s^2\mathbf{g}^H + \mathbf{D}$. Here, σ_s^2 is the known source flux, $\mathbf{g}$ is a vector containing the complex gains to be estimated, and $\mathbf{D}$ is a diagonal matrix containing the unknown noise powers per antenna element (it is assumed that the noise power is uncorrelated from one antenna to another). This is essentially the model considered by [1] [2]. Improved estimation algorithms using iterative and closed form least squares techniques have recently been derived [3]: by incorporating proper weighting, these methods are proved to be asymptotically statistically efficient [4].

For dual polarized telescope arrays, much less is known. In 1995, Hamaker et al. [5] [6] developed a matrix formalism in which the polarization properties of the astronomical signals and their propagation through the ionosphere and the astronomical receiving instrument are efficiently incorporated. An iterative procedure similar to SelfCal [7] is used to estimate the polarization gain coefficients, but it is not known to what solution it will converge.

In [8], the scalar gain calibration methods of [4] [3] was extended to polarized arrays. In this paper we follow the notation and derivations of [8] and [6], and extend it with more efficient algorithms to obtain Least Squares solutions for the gain factors. We will also verify that at least three sky sources with different polarization states are needed to find the gain factors.

Notation ^{H} is the complex conjugate (Hermitian) transpose, ^{t} the matrix transpose, overbar $^-$ the complex conjugate, and † the matrix pseudo inverse (Moore-Penrose inverse). $\mathcal{E}\{\cdot\}$ is the expectation operator. $\mathbf{I}$ is the identity matrix.

This research was partly supported by the NOEMI project of the STW under contract no. DEL77-4476.

2. DATA MODEL

2.1. Coherency

In aperture synthesis radio astronomy, the output of the interferometers is the correlation of the field strengths at the different telescopes, also known as coherencies [7]. The electric field at the location of an antenna element can be described by two linear polarization components, stacked in a 2×1 vector: $\mathbf{e}_i = [e_{ix},\ e_{iy}]^t$. The correlation between two different telescopes i and j is a 2×2 interferometer coherency matrix $\boldsymbol{E}_{ij} = \mathcal{E}\{\mathbf{e}_i\mathbf{e}_j^H\}$. If there are p telescopes, each with two polarizations, then the $2p$ observed electric fields can similarly be stacked in one vector: $\mathbf{e} = \left(\mathbf{e}_1^t, \cdots, \mathbf{e}_p^t\right)^t$. The $2p \times 2p$ Hermitian coherency matrix $\mathbf{E}$ is defined by $\mathbf{E} = \mathcal{E}\{\mathbf{e}\mathbf{e}^H\}$ which can be written in terms of interferometer coherency matrices $\boldsymbol{E}_{ij}$ as

$$\boldsymbol{E}_{i,j} = \begin{bmatrix} \mathcal{E}\{e_{ix}\overline{e}_{jx}\} & \mathcal{E}\{e_{ix}\overline{e}_{jy}\} \\ \mathcal{E}\{e_{iy}\overline{e}_{jx}\} & \mathcal{E}\{e_{iy}\overline{e}_{jy}\} \end{bmatrix}, \quad \mathbf{E} = \begin{bmatrix} \boldsymbol{E}_{11} & \cdots & \boldsymbol{E}_{1p} \\ \vdots & \ddots & \vdots \\ \boldsymbol{E}_{p1} & \cdots & \boldsymbol{E}_{pp} \end{bmatrix}$$

$\mathbf{E}$ is dependent on frequency and time, but for our analysis we assume that we work in a narrow subband and estimate the coherencies at sufficiently short time scales.

2.2. Observed covariance matrix

Instead of the field strengths, each telescope measures a voltage vector $\mathbf{v}_i$. Their relation is given by $\mathbf{v}_i = \boldsymbol{J}_i\mathbf{e}_i$, where $\boldsymbol{J}_i$ is a 2×2 matrix called the Jones matrix. It also incorporates the various ionospheric and atmospheric distortions, gain phase rotations and antenna feed polarization leakage. Hence, the $\boldsymbol{J}_i$ are unknown and have to be estimated.

The observed voltages of the dual polarization output signals of the telescopes i and j are cross-correlated into covariance matrices $\boldsymbol{R}_{ij}$, for which $\boldsymbol{R}_{ij} := \mathcal{E}\{\mathbf{v}_i\mathbf{v}_j^H\} = \boldsymbol{J}_i\boldsymbol{E}_{ij}\boldsymbol{J}_j^H$. Stacking the telescope output voltages $\mathbf{v}_i$ into a $2p$-dimensional vector $\mathbf{v} = [\mathbf{v}_1^t, \cdots, \mathbf{v}_p^t]^t$, and defining

$$\mathbf{J} = \begin{bmatrix} \boldsymbol{J}_1 & & 0 \\ & \ddots & \\ 0 & & \boldsymbol{J}_p \end{bmatrix}, \qquad \mathbf{R} = \begin{bmatrix} \boldsymbol{R}_{11} & \cdots & \boldsymbol{R}_{1p} \\ \vdots & \ddots & \vdots \\ \boldsymbol{R}_{p1} & \cdots & \boldsymbol{R}_{pp} \end{bmatrix}$$

it follows that the $2p \times 2p$ covariance matrix $\mathbf{R}$ is given by $\mathbf{R} = \mathbf{J}\mathbf{E}\mathbf{J}^H$.

In practice the observations are distorted by noise. The system noise signals of each of the two polarization channels, $\mathbf{n}_i = [n_{ix}, n_{iy}]^t$ are stacked into a vector: $\mathbf{n} =$

$\left(\mathbf{n}_1^t, \cdots, \mathbf{n}_p^t\right)^t$. The noise signals are uncorrelated between the telescopes, and up to a certain level also uncorrelated between the two polarizations of a telescope. In our analysis we assume that this is the case. Then the noise matrix $\mathbf{D} = \mathcal{E}\{\mathbf{n}\mathbf{n}^H\}$ is diagonal: $\mathbf{D} = \mathrm{diag}(\sigma_{1x}^2, \sigma_{1y}^2, \cdots, \sigma_{px}^2, \sigma_{py}^2)$. The system noise can be considered additive, so that the covariance matrix of the received data can be written as $\mathbf{R} = \mathbf{J}\,\mathbf{E}\,\mathbf{J}^H + \mathbf{D}$.

2.3. Point source model

Under certain conditions, the electric field can be modeled as the contributions of a finite number of point sources:

$$\mathbf{R} \;=\; \sum_{\ell} \mathbf{J}_\ell\, \mathbf{E}_\ell\, \mathbf{J}_\ell^H \;+\; \mathbf{D}$$

where ℓ is the source direction and $\mathbf{E}_\ell$ is the coherency due to a single source from direction ℓ. Suppose that the source has sky brightness $\mathbf{B}_\ell$ (a 2×2 matrix determined by the source flux polarization components or Stokes parameters). The relation of $\mathbf{B}_\ell$ to $\mathbf{E}_\ell$ can be written as $\boldsymbol{E}_{ij,\ell} = w_{ij,\ell}\, \mathbf{B}_\ell$ where $w_{ij,\ell}$ is the phase shift due to the geometric delay in an interferometer pair i-j [7]. Note that $w_{ij,\ell} = w_{i,\ell}\overline{w_{j,\ell}}$, where $w_{i,\ell}$ is the phase shift at a single telescope. Note that it is the same for the x and the y polarization of this telescope. Thus define

$$\boldsymbol{W}_{i,\ell} = \begin{bmatrix} w_{i,\ell} & 0 \\ 0 & w_{i,\ell} \end{bmatrix}, \qquad \mathbf{W}_\ell = [\boldsymbol{W}_{1,\ell}^t, \cdots, \boldsymbol{W}_{p,\ell}^t]^t,$$

then $\boldsymbol{E}_{ij,\ell} = \boldsymbol{W}_{i,\ell}\; \mathbf{B}_\ell\; \boldsymbol{W}_{j,\ell}^H$, and $\mathbf{E}_\ell = \mathbf{W}_\ell\; \mathbf{B}_\ell\; \mathbf{W}_\ell^H$. The overall observed point source model thus becomes

$$\mathbf{R} \;=\; \sum_{\ell} \mathbf{J}_\ell \mathbf{W}_\ell \mathbf{B}_\ell \mathbf{W}_\ell^H \mathbf{J}_\ell^H \;+\; \mathbf{D}. \qquad (1)$$

3. GAIN CALIBRATION OBSERVATIONS

During a calibration observation, the telescopes are pointed at a single dominant point source in the sky, with known sky brightness. The sum in equation (1) is reduced to a single term. Because the geometry of the telescope array is known, the delay matrix $\mathbf{W}_\ell$ is known as well. We thus obtain the observation model

$$\mathbf{R} = \mathbf{G}\mathbf{B}\mathbf{G}^H + \mathbf{D} \qquad (2)$$

where we defined the $2p \times 2$ gain matrix $\mathbf{G}$ by $\mathbf{G} = \mathbf{J}\mathbf{W}$. Our objective is to estimate $\mathbf{G}$ and $\mathbf{D}$, assuming that an estimate of $\mathbf{R}$ and $\mathbf{B}$ are available. Since $\mathbf{W}_\ell$ is known, $\mathbf{J}$ is easily determined from $\mathbf{G}$. Alternatively, $\mathbf{R}$ can be corrected in advance for $\mathbf{W}_\ell$, after which we can assume without loss of generality that $\mathbf{W}_\ell = \mathbf{I}$ and that $\mathbf{G} = \mathbf{J}$ is direction-independent.

$\mathbf{R}$ is estimated by an observation covariance matrix $\widehat{\mathbf{R}}$, obtained by cross-correlation of N samples $\mathbf{x}_n$ of the telescope output signal vector, $\widehat{\mathbf{R}} = \frac{1}{N}\sum_{n=1}^{N} \mathbf{x}_n \mathbf{x}_n^H$.

4. GAIN ESTIMATION: MAXIMUM LIKELIHOOD AND LEAST SQUARES

As will be shown later in this paper, there are observations required of three astronomical sources which must have different polarization states ($\mathbf{R}_m = \mathbf{G}\mathbf{B}_m\mathbf{G}^H + \mathbf{D}$, $m = 1, \cdots, 3$). As these observations are independent, the joint pdf can be written as a product of the pfd's belonging to the three individual observations. This leads to the following negative logarithmic maximum likelihood formula [9].

$$l(\mathbf{G}, \mathbf{D}) = \sum_{m=1}^{3} \left(\, \log|\mathbf{R}_m| + \mathrm{tr}(\mathbf{R}_m^{-1}\hat{\mathbf{R}}_m)\, \right)$$

Finding the maximum of this formula involves complicated complex derivatives, and a solution probably can be found only in interative form. We therefore formulate the following Least Squares problem for finding $\mathbf{G}$ and $\mathbf{D}$.

$$\{\hat{\mathbf{G}}, \hat{\mathbf{D}}\} = \underset{\mathbf{G},\mathbf{D}}{\mathrm{argmin}} \;\sum_{m=1}^{3} \|\, \widehat{\mathbf{R}}_m - (\mathbf{G}\mathbf{B}_m\mathbf{G}^H + \mathbf{D})\, \|_F^2$$

We assume that the noise power $\mathbf{D}$ is identical for all three observations.

5. FACTOR ANALYSIS ALGORITHMS

In this section we describe factor analysis algorithms which are needed for the polarization gain estimations. The gain estimation algorithms themselves are described in the next section. We consider the factor model [10] $\mathbf{R} = \mathbf{A}\mathbf{A}^H + \mathbf{D}$, where the factor $\mathbf{A}$ is rank-two, and we present two computationally efficient techniques.

5.1. Alternating Least Squares

A straightforward technique to try to optimize a cost function over many parameters is to alternatingly minimize over a subset, keeping the remaining parameters fixed. In our case, assume at the k-th iteration that we have an estimate $\hat{\mathbf{D}}[k]$. The next step is to minimize the LS cost function with respect to the gain vector only:

$$\hat{\mathbf{A}}[k] \;=\; \underset{\mathbf{A}}{\mathrm{argmin}}\; \|\, \hat{\mathbf{R}} - \mathbf{A}\mathbf{A}^H - \hat{\mathbf{D}}[k]\, \|_F^2 \qquad (3)$$

The minimum is found from the eigenvalue decomposition $\hat{\mathbf{R}} - \hat{\mathbf{D}}[k] = \mathbf{U}\boldsymbol{\Lambda}\mathbf{U}^H$, where the matrix $\mathbf{U} = [\mathbf{u}_1\,,\; \cdots\;,\; \mathbf{u}_{2p}]$ contains the eigenvectors $\mathbf{u}_i$, and $\boldsymbol{\Lambda}$ is a diagonal matrix containing the eigenvalues λ_i, sorting in descending order. The factor minimizing (3) is given by $\hat{\mathbf{A}}[k] = [\mathbf{u}_1\lambda_1^{1/2} \quad \mathbf{u}_2\lambda_2^{1/2}]$. The second step is minimizing with respect to the system noise matrix $\mathbf{D}$, keeping the gain vector fixed:

$$\hat{\mathbf{D}}[k+1] \;=\; \underset{\mathbf{D}}{\mathrm{argmin}}\; \|\, \hat{\mathbf{R}} - \hat{\mathbf{A}}[k]\hat{\mathbf{A}}[k]^H - \mathbf{D}\, \|_F^2 \qquad (4)$$

where $\mathbf{D}$ is constrained to be diagonal with nonnegative entries. The minimum is obtained by subtracting $\hat{\mathbf{A}}[k]\hat{\mathbf{A}}[k]^H$ from $\hat{\mathbf{R}}$ and discarding all off-diagonal elements. Since each of the minimizing steps in the iteration loop reduces the model error, we obtain monotonic convergence to a local minimum. Simulations indicate that in the absence of a reasonable initial point, convergence can be very slow.

5.2. Closed form algorithm

The crux of this method is the observation that the off-diagonal entries of $\mathbf{A}\mathbf{A}^H$ are equal to those of $\mathbf{R}$, and known, so that we only need to reconstruct the diagonal entries of $\mathbf{A}\mathbf{A}^H$. We further note that $\mathbf{A}\mathbf{A}^H$ is rank 2, so any submatrix of $\mathbf{R}$ that does not contain elements from the main diagonal is also rank 2. This property can be used to estimate the ratio between any triplet of columns of $\mathbf{R}$ away from the diagonal, and subsequently to estimate how the main diagonal of $\mathbf{R}$ has to be changed so that the resulting $\mathbf{R}'$ is rank 2, or $\mathbf{R}' = \mathbf{A}\mathbf{A}^H$. The gain factor $\mathbf{A}$ can then be extracted by an eigenvalue decomposition.

To illustrate the idea, let (i, j, k) be a triplet of column indices, and let $\mathbf{M}$ be a submatrix of $\mathbf{R}$ consisting of columns (i, j, k), and all rows with indices unequal to i, j, k. Then $\mathbf{M}$ has 3 columns, and rank 2, so that there exists a vector $\mathbf{v} = [v_1, v_2, v_3]^t$ such that $\mathbf{M}\mathbf{v} = 0$. The vector can be found from an SVD of $\mathbf{M}$. It follows that $[r'_{ii}, r_{ij}, r_{ik}]\mathbf{v} = 0$, so that $\hat{r}'_{ii} = -(r_{ij}v_2 + r_{ik}v_3)/v_1$. This estimate can be improved by considering all possible triplets containing i, and combining the ratios. After filling in all diagonal entries of $\mathbf{R}'$ in this way, a rank-2 factorization of $\mathbf{R}' = \mathbf{A}\mathbf{A}^H$ provides an estimate for the factor $\mathbf{A}$. An estimate for $\mathbf{D}$ is subsequently found from $\mathbf{R} - \mathbf{A}\mathbf{A}^H$.

6. POLARIZATION GAIN ESTIMATION ALGORITHMS

6.1. Closed form algorithm

One reference source
Consider a single source, $\mathbf{R} = \mathbf{G}\mathbf{B}\mathbf{G}^H + \mathbf{D}$, where $\mathbf{R}$ has been estimated and $\mathbf{B}$ is known from sky tables. Using factor analysis, we can find $\mathbf{D}$ and a factor $\mathbf{A}$ such that $\mathbf{R} = \mathbf{A}\mathbf{A}^H + \mathbf{D}$. However, $\mathbf{A}$ is not unique: for any 2×2 unitary matrix $\mathbf{Q}$, we have $\mathbf{A}\mathbf{A}^H = (\mathbf{A}\mathbf{Q})(\mathbf{Q}^H\mathbf{A}^H)$. Hence, we can estimate $\mathbf{A}$ only up to a unitary factor. It follows that $\mathbf{G} = \mathbf{A}\mathbf{Q}\mathbf{B}^{-1/2}$, where $\mathbf{Q}$ is unknown. It is not possible to estimate $\mathbf{G}$ in more detail using only a single reference source.

Two reference sources
With two reference sources, we have

$$\begin{aligned}\mathbf{R}_1 &= \mathbf{G}\mathbf{B}_1\mathbf{G}^H + \mathbf{D}\\ \mathbf{R}_2 &= \mathbf{G}\mathbf{B}_2\mathbf{G}^H + \mathbf{D}\end{aligned}$$

$\mathbf{R}_1$ and $\mathbf{R}_2$ are observed, $\mathbf{B}_1$ and $\mathbf{B}_2$ are the known polarization matrices from the reference sources, and $\mathbf{D}$ is known from a factor analysis. Again, they are unique only up to unknown 2×2 unitary factors $\mathbf{Q}_1$, $\mathbf{Q}_2$.

A generalized eigenvalue decomposition of the pair $(\mathbf{B}_1, \mathbf{B}_2)$ provides the factorizations

$$\mathbf{B}_1 = \mathbf{M}\boldsymbol{\Lambda}_1\mathbf{M}^H, \quad \mathbf{B}_2 = \mathbf{M}\boldsymbol{\Lambda}_2\mathbf{M}^H\,, \tag{5}$$

where $\mathbf{M}$ is a square invertible matrix and $\boldsymbol{\Lambda}_1, \boldsymbol{\Lambda}_2$ are positive diagonal matrices. It is assumed that the generalized eigenvalues are distinct.

The same decomposition on $(\mathbf{R}_1 - \mathbf{D}, \mathbf{R}_2 - \mathbf{D})$ gives

$$\begin{aligned}\mathbf{R}_1 &= \mathbf{A}\boldsymbol{\Lambda}_1\mathbf{A}^H + \mathbf{D}\\ \mathbf{R}_2 &= \mathbf{A}\boldsymbol{\Lambda}_2\mathbf{A}^H + \mathbf{D}\end{aligned} \tag{6}$$

Note that the $\boldsymbol{\Lambda}_1$ in both cases are the same, and also $\boldsymbol{\Lambda}_2$. This is because

$$(\mathbf{R}_1 - \mathbf{D}) - \lambda(\mathbf{R}_2 - \mathbf{D}) = \mathbf{G}(\mathbf{B}_1 - \lambda\mathbf{B}_2)\mathbf{G}^H$$

so that $(\mathbf{B}_1, \mathbf{B}_2)$ and $(\mathbf{R}_1, \mathbf{R}_2)$ have the same rank reducing numbers λ.

Combining the two equations, we immediately obtain

$$\mathbf{A} = \mathbf{G}\mathbf{M}\bar{\boldsymbol{\Phi}} \quad \Rightarrow \quad \mathbf{G} = \mathbf{A}\boldsymbol{\Phi}\mathbf{M}^{-1} \tag{7}$$

where $\boldsymbol{\Phi} = \mathrm{diag}(\boldsymbol{\phi})$ is an unknown diagonal matrix with unimodular diagonal entries, representing phase ambiguities. Without further information, these cannot be further resolved.

An alternative computation that leads to (7) would go via factor analysis of $\mathbf{R}_1$ and $\mathbf{R}_2$ separately:

$$\begin{aligned}\mathbf{R}_1 &= \mathbf{A}_1\mathbf{A}_1^H + \mathbf{D}_1 = (\mathbf{A}_1\mathbf{Q}_1)(\mathbf{Q}_1^H\mathbf{A}_1^H) + \mathbf{D}_1\\ \mathbf{R}_2 &= \mathbf{A}_2\mathbf{A}_2^H + \mathbf{D}_2 = (\mathbf{A}_2\mathbf{Q}_2)(\mathbf{Q}_2^H\mathbf{A}_2^H) + \mathbf{D}_2\end{aligned}$$

where $\mathbf{Q}_1$ and $\mathbf{Q}_2$ are unknown unitary matrices. Setting $\mathbf{D}_1 = \mathbf{D}_2 = \mathbf{D}$, comparing the two equations with the model (6), and inserting (5), we obtain

$$\mathbf{G} \;=\; \mathbf{A}_1\mathbf{Q}_1\boldsymbol{\Lambda}_1^{-1/2}\mathbf{M}^{-1} \;=\; \mathbf{A}_2\mathbf{Q}_2\boldsymbol{\Lambda}_2^{-1/2}\mathbf{M}^{-1}$$

The latter equality relates $\mathbf{Q}_1$ to $\mathbf{Q}_2$ as

$$\mathbf{A}_1^{\dagger}\mathbf{A}_2 = \mathbf{Q}_1(\boldsymbol{\Lambda}_1^{-1/2}\boldsymbol{\Lambda}_2^{1/2})\mathbf{Q}_2^H\,.$$

This has the form of an SVD, and $\mathbf{Q}_1$ and $\mathbf{Q}_2$ can be computed as the left and right singular vectors of $\mathbf{A}_1^{\dagger}\mathbf{A}_2$. However, these are unique only up to an unknown diagonal phase matrix $\boldsymbol{\Phi}$. Hence we obtain

$$\mathbf{G} \;=\; (\mathbf{A}_1\mathbf{Q}_1)\boldsymbol{\Phi}(\boldsymbol{\Lambda}_1^{-1/2}\mathbf{M}^{-1})$$

where only $\boldsymbol{\Phi}$ is unknown. This is of the same form as (7).

Thus, we can estimate $\mathbf{G}$ only up to two unknown phases. With some effort, this can be converted to a more convenient normalization: if we define a normalized $\mathbf{G}$ to have positive real entries on its first row, then the normalized $\mathbf{G}$ is unique. This is the best that can be expected using two reference sources.

Three reference sources
If a third observation of a point source is available, $\mathbf{R}_3 = \mathbf{G}\mathbf{B}_3\mathbf{G}^H + \mathbf{D}$, then the ambiguity can be further reduced to a single common phase. Indeed, after a similar generalized eigenvalue decomposition on the pair $(\mathbf{R}_1, \mathbf{R}_3)$, we have available two sets of equations,

$$\mathbf{G} = \mathbf{A}_1\boldsymbol{\Phi}_1\mathbf{M}_1^{-1} = \mathbf{A}_2\boldsymbol{\Phi}_2\mathbf{M}_2^{-1}$$

or, denoting by $\circ$ the Khatri-Rao (column-wise Kronecker) product,

$$\mathrm{vec}(\mathbf{G}) = (\mathbf{M}_1^{-T} \circ \mathbf{A}_1)\boldsymbol{\phi}_1 = (\mathbf{M}_2^{-T} \circ \mathbf{A}_2)\boldsymbol{\phi}_2 \tag{8}$$

or

$$[(\mathbf{M}_1^{-T} \circ \mathbf{A}_1) \quad -(\mathbf{M}_2^{-T} \circ \mathbf{A}_2)] \begin{bmatrix}\boldsymbol{\phi}_1\\ \boldsymbol{\phi}_2\end{bmatrix} = 0\,.$$

Thus, $[\boldsymbol{\phi}_1^T,\ \boldsymbol{\phi}_2^T]^T$ is the unique solution in the null space, and determined up to a scaling ϕ. Inserting in (8), we obtain $\mathbf{G}$ up to an unknown scaling by ϕ.

6.2. Parallel factor analysis

In [11] [12] [13] parallel factor analysis solutions are given for the model $||\hat{\mathbf{R}} - \mathbf{A}\mathbf{B}\mathbf{C}||_F^2$. Our model, which is based on a combination of three sources $\mathbf{B}_1$, $\mathbf{B}_2$, and $\mathbf{B}_3$, is slightly different. Let $\tilde{\mathbf{R}} = \left[\hat{\mathbf{R}}_1^t\ \hat{\mathbf{R}}_2^t\ \hat{\mathbf{R}}_3^t\right]^t$, and let $\tilde{\mathbf{D}} = \left[\mathbf{D}^t\ \mathbf{D}^t\ \mathbf{D}^t\right]^t$. Define also $\tilde{\mathbf{G}}_{\mathbf{B}} = \left[(\mathbf{G}\mathbf{B}_1)^t\ (\mathbf{G}\mathbf{B}_2)^t(\mathbf{G}\mathbf{B}_3)^t\right]^t$, then, assuming that $\mathbf{D}$ is found by using one of the factor analysis algorithms describes earlier, the Least Squares cost function to minimize is given by:

$$\{\hat{\mathbf{G}}\} = \underset{\mathbf{G}}{\mathrm{argmin}}\ \ ||\,\tilde{\mathbf{R}} - \tilde{\mathbf{D}} - (\tilde{\mathbf{G}}_{\mathbf{B}}\mathbf{G}^H)\,||_F^2 \tag{9}$$

Our model differs from [13] [11] in the sense that the $\mathbf{B}_k$ matrices in general are not diagonal, also our model has more structure than in [12] as the two $\mathbf{G}$ matrices are identical. A solution for the cost function above can be found by iteratively estimating the gain matrix $\hat{\mathbf{G}}$ by

$\hat{\mathbf{G}} = (\tilde{\mathbf{G}}_{\mathbf{B}})^{\dagger}(\tilde{\mathbf{R}} - \tilde{\mathbf{D}}))^{H}$ and inserting it in $\tilde{\mathbf{G}}_{\mathbf{B}}$. For this algorithm, a good initial point is needed; one of the two proposed closed form algorithms can be used for this purpose.

7. SIMULATIONS

The performance of the closed form and parafac dual polarization gain estimation methods is studied by applying them on three generated covariance matrices $\hat{\mathbf{R}}_k$, based on the "true" values of $\mathbf{G}$, $\mathbf{D}$, and $\mathbf{B}_k$. The performance is quantified by calculating the model error $||\tilde{\mathbf{R}} - \tilde{\mathbf{D}} - \tilde{\mathbf{G}}_{\mathbf{B}}\hat{\mathbf{G}}||_F$. Here, the noise matrices $\hat{\mathbf{D}}$ or $\tilde{\mathbf{D}}$ are estimated using the closed form factor analysis algorithms. In the simulations, the $\mathbf{B}_k$ matrices used are:

$$\mathbf{B}_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \ \mathbf{B}_2 = \begin{bmatrix} 1.2 & 0.1 \\ 0.1 & 0.8 \end{bmatrix}, \ \mathbf{B}_3 = \begin{bmatrix} 1 & -0.1\imath \\ 0.1\imath & 1 \end{bmatrix}$$

The gain magnitudes are nominally one, with gain magnitude variations (over the 2p telescope channels) up to 10%, and phase variations up to 2π radians. Figure 1 shows the

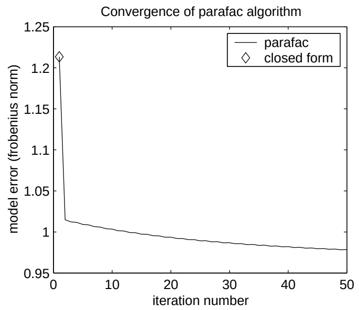

Figure 1.
Convergence of the parafac algorithm, using the closed form algorithm as a starting point (Nsam = 1000, $p = 5$, $\mathbf{D} = \mathbf{I}$, $N_{it} = 256$).

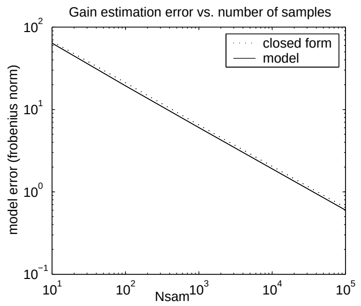

Figure 2.
Performance of the closed form algorithm, compared to the model error using the true gains. ($p = 5$, $\mathbf{D} = 10\ \mathbf{I}$, $N_{it} = 128$)

convergence of the parafac method, using the closed from method as initial point. The improvement of the parafac method w.r.t. the closed form method is not dramatic. Figure 2 shows the model error of the closed from method compared to the model error of the true gains. The closed form model error apparently is very close to the model error of the true gains.

8. CONCLUSIONS

We have shown that observations of at least three astronomical sources are required to fully solve the dual polarization gain estimation problem of radio telescope arrays. We have presented several closed form and iterative solutions for the dual polarization gain estimation problem. These solutions are not optimal as these do not fully exploit the structure which is present in the model. However, initial simulations show that, for these solutions, the model error is close to the model error using the true gain values.

References

[1] H. van Someren-Greve, "Logarithmic Least Square Gain Decomposition Algorithm for the WSRT," tech. rep., 1980. IWOS software documentation, ASTRON internal document.

[2] T. Cornwell and P.N.Wilkinson, "A New Method for Making Maps with Unstable Radio Interferometers," *Mon. Not.R. Astron. Soc.*, vol. 196, pp. 1067–1086, 1981.

[3] A.J.Boonstra and A. van der Veen, "Gain Decomposition Methods for Radio Telescope Arrays," in *IEEE Workshop on Statistical Signal Processing*, IEEE, August 2001.

[4] A.J. Boonstra ,and A.J. van der Veen, "Gain Calibration Methods for Radio Telescope Arrays," *IEEE Transactions on Signal Processing*, scheduled for the January issue 2003.

[5] J. P. Hamaker, J. D. Bregman, and R. J. Sault, "Understanding Radio Polarimetry. I. Mathematical Foundations.," *Astronomy and Astrophysics Supplement Series*, vol. 117, pp. 137–147, May 1996.

[6] J. P. Hamaker, "Understanding Radio Polarimetry. IV. The Full-Coherency Analogue of Scalar Self-Calibration: Self-Alignment, Dynamic Range and Polarimetric Fidelity," *Astronomy and Astrophysics Supplement Series*, vol. 143, pp. 515–534, May 2000.

[7] A.R. Thompson, J.R. Moran, and G.W. Swenson, *Interferometry and Synthesis in Radio Astronomy.* New York: John Wiley & Sons, first ed., 1986.

[8] A. Boonstra and A. van der Veen, "Gain Estimation Methods for Polarized Radio Telescope Arrays," in *URSI XXVII th General Assembly*, URSI, August 17 -24 2002.

[9] S.Haykin and J.Litva and T.L. Shepherd (Eds.), ed., *Radar Array Processing*, vol. 25 of *Springer Series in Infromation Sciences.* Berlin: Springer Verlag, 1993.

[10] D.N. Lawley and A.E. Maxwell, *Factor Analysis as a Statistical Method.* Butterworth & Co, London, 1971.

[11] A.J. van der Veen, "Joint Diagonalisation Via Subspace Fitting Techniques," in *ICASSP*, IEEE, May 7-11 2001.

[12] X. Liu, and N.D. Siridopoulis, "Parafac Methods For Blind Beamforming: Multilinear ALS Performance and CRB," in *ICASSP*, Istanbul, Turkey, June 2000.

[13] A. Yeredor, "Non-Orthogonal Joint Diagonalisation in the Least-Squeares Sense With Application in Blind Source Separation," *IEEE Transactions on Signal Processing*, vol. 50, p. 15451553, July 7 2002.