GAIN DECOMPOSITION METHODS FOR RADIO TELESCOPE ARRAYS

A.J. Boonstra[1,2] *and A.J. van der Veen*[2]

[1]ASTRON, Netherlands Foundation for Research in Astronomy
Oude Hoogeveensedijk 4, 7991 PD Dwingeloo, The Netherlands
Tel: +31(0)521 595100, fax: +31(0)521 597332, email: boonstra@astron.nl

[2] Department of Electrical Engineering, Delft University of Technology
Mekelweg 4, 2628 CD Delft, The Netherlands
Tel: +31(0)15 2781372, fax: +31(0)15 2786190, email: allejan@cas.et.tudelft.nl

ABSTRACT

In radio telescope arrays, the complex receiver gains and sensor noise powers are initially unknown and have to be calibrated. Gain calibration enhances the quality of astronomical sky images and moreover, improve the effectiveness of certain radio telescope phased-array data processing techniques, such as radio interference (RFI) mitigation and beamforming. In this paper we present several closed form and iterative complex gain estimation methods. These methods are analyzed and compared to the Cramer-Rao lower bound for the variance of the estimated gain. The models are tested both on simulated data and on observed telescope data.

Keywords: applications in radio astronomy, sensor array processing.

1. INTRODUCTION

Gain calibration techniques for radio telescope systems exist already for a long time [1][2]. However, since studies started for a next generation of radio telescopes (the Square Kilometer Array radio telescope or SKA [3]), phased array beamforming issues and radio frequency interference (RFI) suppression techniques received renewed interest [4] in radio astronomy. For RFI suppression, and for phased array beamforming, gain calibration of the telescope array is an important factor. Maximum likelihood techniques exist for estimation of the gain and phase of signals impinging on the telescope arrray [5] and for estimation of the direction of arrival of the impinging signals [6]. For computational reasons (SKA will have many sensor elements) and for robustness reasons (iterative maximum likelihood techniques depend on a good initial point) we investigated several closed form and iterative complex gain estimation methods and found that these techniques perform well.

The complex gains and noise powers of individual telescopes of a telescope array (figure 1) can be estimated by observing a strong astronomical source in the centre of the field of view of the telescopes. In most cases, single point sources can be found which dominate the field of view of a radio telescope. A telescope output signal is the sum of the telescope system noise (uncorrelated among the telescopes) and the astronomical source flux, which is correlated, multiplied by the telescope gain. The source flux is the same for each of the telescopes, but the telescope gains and noise powers usually are not. The gains consist of the combined effect of atmospheric disturbances, telescope geometry, receiver characteristics, and electronic (amplifier) gains, whereas the system noise powers can differ by several dB's.

A.J. Boonstra was supported by the NOEMI project of the STW under contract no. DEL77-4476.

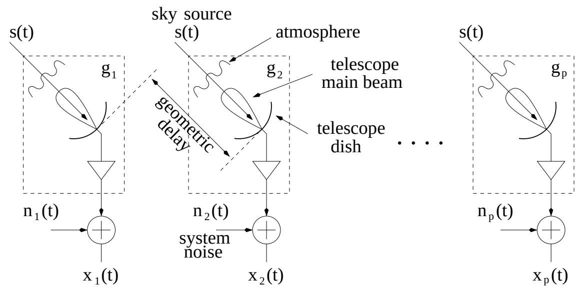

Fig. 1. Radio telescope array

The output of the backend processing is a sequence of covariance matrices formed by cross correlation of all the telescope outputs x_i. The aim in this paper is to estimate the complex gain factors and the system noise powers from an observed covariance matrix, assuming that the astronomical source flux is known from tables. We present three algorithms to extract these parameters.

2. DATA MODEL

Assume that during the calibration observations the telescopes are pointed at a single radio source in the sky. For a telescope array (figure 1) the output x_i of element i at a certain time t can be modeled (using the narrow band assumption) as

$$x_i(t) = g_i a_i s(t) + n_i(t) \tag{1}$$

where g_i is the complex gain of the sensor, n_i is the system noise of channel i, a_i is the narrow band phase offset due to the geometric delay, and $s(t)$ is the flux of the impinging external source. For the gain calibration observation, the sky source is located in the centre of the field of view. The geometry and look direction of a telescope is known, so the narrow band phase offset due to the geometric delay is known as well and can be compensated for. Hence without loss of generality we may assume in our model that the phase offsets are $a_i = 1$.

In radio astronomy, the sensor array output

$$\mathbf{x}(t) = [x_1(t)\,,\ \cdots\ ,\ x_p(t)]^t \qquad (2)$$

is usually correlated with itself to form a covariance matrix. Here the superscript t means the transpose operator, and p is the number of telescopes. The true covariance matrix $\mathbf{R}$ and estimate $\widehat{\mathbf{R}}$ based on N samples, assuming stationarity over this interval, are given by

$$\mathbf{R} \ = \ E\{\mathbf{x}(t)\mathbf{x}(t)^H\} \qquad (3)$$

$$\widehat{\mathbf{R}} \ = \ \frac{1}{N}\sum_{n=0}^{N-1}\mathbf{x}(t+nT)\mathbf{x}(t+nT)^H \qquad (4)$$

where superscript H denotes the complex conjugate transpose.

The signal power $\sigma_s^2 = E\{|s(t)|^2\}$ is known from tables, hence without loss of generality we may model it as $\sigma_s^2 = 1$. The covariance matrix can now be written as

$$\mathbf{R} = \mathbf{g}\mathbf{g}^H + \mathbf{D} \qquad (5)$$

where $\mathbf{D}$ is a diagonal matrix containing the system noise contributions, $d_i = E\{|n_i(t)|^2\} \geq 0$. The gain vector $\mathbf{g}$ can be written as a product of a gain magnitude $\boldsymbol{\gamma} = [\gamma_i, \cdots, \gamma_p]^t$ $(\gamma_i > 0)$ and a phasor $\mathbf{z} = [e^{\jmath\phi_1}, \cdots, e^{\jmath\phi_p}]^t$; i.e. $\mathbf{g} = \boldsymbol{\gamma}\odot\mathbf{z}$, where $\odot$ is the Schur-Hadamard (elementwise) matrix product. The ij-th element of $\mathbf{R}$ is thus given by

$$r_{ij} = \gamma_i\gamma_j e^{\jmath(\phi_i - \phi_j)} + d_i\delta_{ij} \qquad (6)$$

Since the phases are underdetermined, we define without loss of generality the phase of the first sensor to be zero: $\phi_1 = 0$. The objective at this point is, given $\widehat{\mathbf{R}}$, estimate $\mathbf{g}$ and $\mathbf{D}$ according to the model (5).

3. GAIN DECOMPOSITION ALGORITHMS

3.1. Alternating least squares gain estimation (ALS)

The covariance matrix in equation (5) is composed of a rank-one matrix $\mathbf{g}\mathbf{g}^H$ and a diagonal matrix $\mathbf{D}$. The gain extraction procedure is based on minimizing the model error:

$$\{\widehat{\mathbf{g}}, \widehat{\mathbf{D}}\} = \arg\min_{\mathbf{g},\mathbf{D}\geq 0} \|\widehat{\mathbf{R}} - \mathbf{D} - \mathbf{g}\mathbf{g}^H\|_F^2 \qquad (7)$$

where $\|\cdot\|_F$ denotes the Frobenius norm. In the ALS technique, we alternatingly minimize over one component, keeping the other component fixed. In particular, assume at the k-th iteration that we have an estimate $\widehat{\mathbf{D}}^{(k)}$. The next step is to minimize equation (7) with respect to the gain vector only:

$$\widehat{\mathbf{g}}^{(k)} = \arg\min_{\mathbf{g}} \|\widehat{\mathbf{R}} - \mathbf{g}\mathbf{g}^H - \widehat{\mathbf{D}}^{(k)}\|_F^2 \qquad (8)$$

The minimum is found from the eigenvalue decomposition of $\widehat{\mathbf{R}} - \widehat{\mathbf{D}}^{(k)}$,

$$\widehat{\mathbf{R}} - \widehat{\mathbf{D}}^{(k)} = \mathbf{U}^{(k)}\boldsymbol{\Lambda}^{(k)}\mathbf{U}^{(k)H} \qquad (9)$$

where the matrix $\mathbf{U} = [\mathbf{u}_1 \cdots \mathbf{u}_p]$ contains the eigenvectors $\mathbf{u}_i$, and $\boldsymbol{\Lambda}$ is a diagonal matrix containing the eigenvalues λ_i. The gain estimate minimizing (8) is given by

$$\widehat{\mathbf{g}}^{(k)} = \mathbf{u}_1^{(k)}\sqrt{\lambda_1^{(k)}} \qquad (10)$$

where λ_1 is the largest eigenvalue, and $\mathbf{u}_1$ is the corresponding eigenvector. The second step is to minimize (7) with respect to the system noise matrix $\mathbf{D}$, keeping the gain vector fixed. The minimum is obtained by subtracting $\widehat{\mathbf{g}}^{(k)}\widehat{\mathbf{g}}^{(k)H}$ from $\widehat{\mathbf{R}}$ and discarding all off-diagonal elements. The condition that the diagonal elements of $\widehat{\mathbf{D}}^{(k+1)}$ should be positive is implemented by subsequently setting $\widehat{d_i}^{(k+1)} = \max(\widehat{d_i}^{(k+1)}, 0)$. The two minimizations steps are repeated until the model error (7) converges. Since each of the minimizing steps in the iteration loop reduces the model error, we obtain monotonic convergence to a local minimum. Although the iteration is very simple to implement, simulations indicate that convergence usually is very slow, especially in the absence of a reasonable initial point.

3.2. Column ratio gain estimation (COL)

We now set out to find a closed form estimate of $\mathbf{g}$, which recovers $\mathbf{g}$ exactly when applied to $\mathbf{R}$ (hence asymptotic for $\widehat{\mathbf{R}}$). The crux of this method is the observation that the off-diagonal entries of $\mathbf{g}\mathbf{g}^H$ are equal to those of $\mathbf{R}$, hence known, so that we only need to reconstruct the diagonal entries of $\mathbf{g}\mathbf{g}^H$. This can be done in closed form by estimating the column ratios of $\mathbf{R}$ away from the diagonal as discussed below. The diagonal of the covariance matrix $\mathbf{R}$ is then replaced with the estimate producing a matrix of the form $\mathbf{R}' = \mathbf{g}\mathbf{g}^H$. The gain vector $\mathbf{g}$ can then be extracted by an eigenvalue decomposition of $\mathbf{R}'$.

The ratio α_{ij} of two of elements g_i and g_j of the complex gain vector $\mathbf{g}$ can be estimated from the data $\mathbf{R}$ by solving

$$\mathbf{c}_i = \alpha_{ij}\,\mathbf{c}_j \qquad (11)$$

where $\mathbf{c}_i$ and $\mathbf{c}_j$ are the i-th and j-th column of the matrix $\mathbf{R}$, not including the entries r_{ii}, r_{ij}, r_{ji} and r_{jj} because r_{ii} and r_{jj} contain also the unknown system noise contributions d_i. Solving for α_{ij} in the Least Squares sense gives

$$\alpha_{ij} = (\mathbf{c}_i^H\mathbf{c}_i)^{-1}\mathbf{c}_i^H\mathbf{c}_j = \frac{\sum_{k\neq i,j} r_{ki}^* r_{kj}}{\sum_{k\neq i,j} r_{ki}^* r_{ki}} \qquad (12)$$

We can subsequently estimate $|g_i|^2$ as $|g_i|^2 = \Re(\alpha_{ij}r_{ij}^*)$, for any choice of j. This estimate can be improved if all $(p-1)$ available column ratios are used. The next step is to form $\mathbf{R}'$ equal to $\mathbf{R}$ but with the diagonal entries replaced by the estimates of $|g_i|^2$ obtained above. The resulting matrix $\mathbf{R}'$ is an estimate of $\mathbf{g}\mathbf{g}^H$, and $\mathbf{g}$ is found from an eigenvalue decomposition of $\mathbf{R}' = \mathbf{U}\boldsymbol{\Lambda}\mathbf{U}^H$, similarly as in (9), (10) before. In the actual algorithm, we follow the same procedure but replace $\mathbf{R}$ by the sample estimate $\widehat{\mathbf{R}}$.

3.3. Logarithmic least square gain estimation (LOGLS)

An alternative closed form estimate (as in use at the Westerbork Sythesis Radio Telescope, WSRT [1] since 1980) is obtained by minimizing the mean square error of the *logarithms* of (6). Taking the logarithm has the effect that the equations become linear as the product of gains become sums. We start by taking the logarithm of the off-diagonal elements ($\forall\ i \neq j$) of equation (6) and define the logarithmic model errors Δ_{ij} ($\forall\, i \neq j$) as

$$\Delta_{ij} \equiv \ln(\widehat{r}_{ij}) - \ln(\gamma_i) - \ln(\gamma_j) - \jmath(\phi_i - \phi_j) \bmod 2\pi\jmath \quad (13)$$

Minimization in the least square sense of the sum-squared error $\sum |\Delta_{ij}|^2$ over the real gains and and phases is obtained by setting

$$\frac{\partial}{\partial \ln(\gamma_k)} \sum_{\substack{i,j=1 \\ i\neq j}}^{p} |\Delta_{ij}|^2 = 0\ , \quad \frac{\partial}{\partial \ln(e^{\jmath\phi_k})} \sum_{\substack{i,j=1 \\ i\neq j}}^{p} |\Delta_{ij}|^2 = 0 \quad (14)$$

After some manipulations the equation for the gain magnitude (14) becomes:

$$\sum_{\substack{i=1 \\ i\neq k}}^{p} \Big(\Re\{\ln(\widehat{r}_{ik})\} - \ln(\gamma_i) - \ln(\gamma_k) \Big) = 0 \quad (15)$$

for $k = 1, \cdots, p$. This equation can easily be written in matrix form and solved in closed form using Woodbury's identity. The same procedure leads to a closed form solution for the phase. In this method, phase unwrapping is necessary. This is done by using a simple phase quadrant estimation procedure.

4. GAIN ESTIMATION SIMULATIONS

4.1. Method

The aim of the simulations is to evaluate the gain estimation accuracy as a function of signal to noise ratio ($\text{SNR}_i = \mathbf{g}^H\mathbf{g}/d_i$), i.e. the ratio of the astronomical source power (normalized here to unity) and array gain to the noise power in the i-th channel.

In the simulations we use eight telescope channels. The gain magnitude was kept fixed during the simulations, and was taken as a nominal value plus a (uniformly selected) random deviation of 10% of the nominal value . The gain phase was randomly distributed in the interval $[0, 2\pi]$ and also kept fixed during the simulations. In the presentation of the results, we split the gain estimates in a magnitude and a phase, since they might have different accuracies, and since the Cramer-Rao bounds are based on these (real) parameters.

4.2. Cramer-Rao lower bound of the gain estimates

The Cramer-Rao Bound (CRB) gives a lower bound on the variance of any unbiased estimator [7]. In our situation, we assume that the source signal and the channel noise are independent Gaussian distributed with zero mean, and satisfy the model in equation (5). Define the parameter vector

$\boldsymbol{\theta} \equiv [\gamma_1, \cdots, \gamma_p, \phi_2, \cdots, \phi_p, d_1, \cdots, d_p]^t$ (Note that the phase ϕ_1 of the first sensor is not a parameter.) The CRB is then given by [7]

$$\text{var}\,(\widehat{\theta}_i(\mathbf{X})|\boldsymbol{\theta}) \geq \left[\mathbf{I}_F^{-1}\right]_{ii} \quad (16)$$

where $\mathbf{I}_F$ is the Fisher information matrix, where $\mathbf{X}$ is defined as $\mathbf{X} \equiv (\mathbf{x}[1]\cdots\mathbf{x}[N])$, and N is the number of samples. Following standard techniques [7], the Fisher information matrix can be written as

$$\mathbf{I}_{F,ij}(\boldsymbol{\theta}) = N\text{tr}\Big(\mathbf{R}^{-1}\frac{\partial\mathbf{R}}{\partial\theta_i}\mathbf{R}^{-1}\frac{\partial\mathbf{R}}{\partial\theta_j}\Big) \quad (17)$$

Inserting the model (5), the components in the Fisher information matrix can easily be found as

$$\mathbf{R}^{-1} = \mathbf{D}^{-1}\left(\mathbf{I} - \frac{\mathbf{g}\mathbf{g}^H\mathbf{D}^{-1}}{1+\mathbf{g}^H\mathbf{D}^{-1}\mathbf{g}}\right) \quad (18)$$

$$\partial\mathbf{R}/\partial\gamma_i = (\mathbf{e}_i \odot \mathbf{z})\,\mathbf{g}^H + \mathbf{g}\,(\mathbf{e}_i \odot \mathbf{z}^H) \quad (19)$$

$$\partial\mathbf{R}/\partial\phi_i = \jmath(\ (\mathbf{g}\odot\mathbf{e}_i)\mathbf{g}^H - \mathbf{g}(\mathbf{g}\odot\mathbf{e}_i)^H) \quad (20)$$

$$\partial\mathbf{R}/\partial d_i = \mathbf{e}_i\mathbf{e}_i^t \quad (21)$$

where $\mathbf{e}_i$ denotes the i-th unit vector. The estimation variance of the model parameters is calculated by evaluating equation (16).

4.3. Comparison of the gain decomposition methods: simulation results

For a typical online gain calibration measurement at a radio observatory, astronomical sources are used with noise powers in the range of 0.1 to 10% of the telescope system noise power. The integration time of the correlation data can be several seconds to a few minutes.

Figure 2 shows the results of a gain estimation simulation in which the gain estimation variance is plotted versus SNR for a fixed number of samples. The three models are plotted together with the Cramer-Rao lower bound. In the -10 to 0 dB SNR range, the gain estimation errors lie very close to the CRB (for 16 k samples) and the estimators are unbiased. Below an SNR of -15 dB the gain estimation starts to deviate from the bound. The ALS method breaks down at higher SNR than the other two methods.

The phase estimation tends to break down earlier than the gain magnitude estimation. The phase breakdown point is observed around -16 dB, and is a bit lower for the LOGLS method. At low SNR some of the curves drop below the Cramer-Rao bound. Here, the estimators are biased.

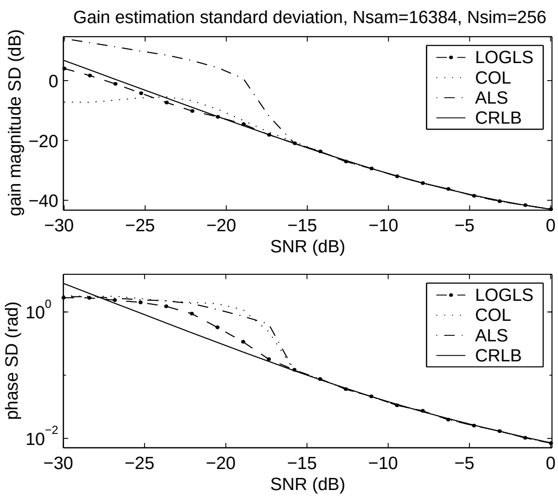

Fig. 2. Gain estimation standard deviation versus SNR

In figure 3 the gain estimates are plotted as a function of the number of observed time samples for a fixed SNR. Note that also here, the phase estimators break down earlier than the gain magnitude estimators.

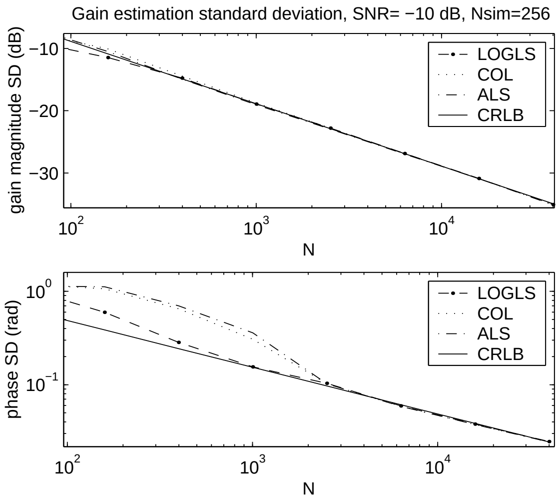

Fig. 3. Gain estimation standard deviation vs. number of samples

5. EXPERIMENTAL RESULTS

The gain estimation methods were tested on real telescope data. An eight channel datarecorder was connected to the Westerbork Synthesis Radio Telescope, which was pointing at the astronomical source 3C48. Baseband signals were recorded corresponding to a sky frequency of 1420 MHz. The SNR of the source relative to the system noise is -13 dB. A narrow band is selected (by means of an FFT) and covariance matrices are derived by cross correlation of the input sequences. The observed correlation coefficient is 0.055 with a spread of about 10% due to the different gains of the telescopes. The gain decomposition algorithms are applied to the covariance matrices. Figure 4 shows the observed gain magnitude estimation standard deviation and the CRLB. The entire dataset is

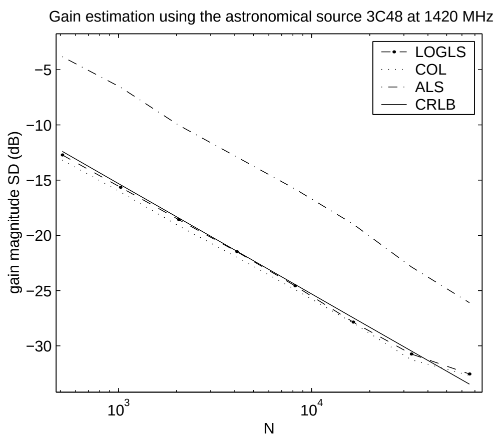

Fig. 4. Gain estimation standard deviation from observations with the astronomical source 3C48.

used to obtain a fair estimate of the complex gains, these numbers are used for the calculation of the CRLB. The curves for the LOGLS and COLS gain estimations lie very close to the CRLB, just as is the case with the simulations.The ALS however, performs not too well for SNR's in the range below about -15 dB; the ALS estimates are biased. The small deviation of the LOGLS and COLS curves from the CRLB curve could be caused by the fact that for the CRLB calculations not the true gains were used (as they are not known) but the estimated gains.

6. CONCLUSIONS

In our simulations the three gain estimation methods do not differ much in performance. The main difference is that the ALS method for gain magnitude estimation breaks down a bit earlier than the two other methods. Also, the phase estimation seems to break down earlier than the gain magnitude estimation. For 16 k samples and for SNRs higher than -15 dB, the estimators are unbiased (for the gain distribution used).

In general the measurement results support the conclusions from the simulations. However, the ALS gain magnitude estimates deviate 8 dB from the CRLB. The ALS estimator is biased for the SNR of the observation.

Further research will focus on other methods, like the Gauss-Newton iterative method [8], on processing efficiency, and the methods will be extended to multiple sources.

7. REFERENCES

[1] H. van Someren-Greve. Logarithmic least square gain decomposition algorithm for the WSRT. IWOS software documentation, ASTRON internal document, 1980.

[2] R.A. Perley F.R. Schwab and A.H.Bridle. Synthesis imaging in radio astronomy. *Astronomical Society of the Pacific Conference Series*, 6, 1994.

[3] A.B.Smolders and M.P. van Haarlem, editors. *Perspectives on Radio Astronomy: Technologies for Large Antenna Arrays*. ASTRON, conference proceedings edition, April 1999.

[4] A. Leshem A.J. van der Veen and A.J. Boonstra. Multichannel interference mitigation techniques in radio astronomy. *Astrophysical Journal Supplements*, 131(1):355–374, November 2000.

[5] Q. Cheng Y.Hua and Petre Stoica. Asymptotic performance of optimal gain and phase estimators of sensor arrays. *IEEE Transactions on Signal Processing*, 48(12):3587–3590, December 2000.

[6] M. Pesavento and A.B. Gershman. Array processing in the presence of unknown nonuniform sensor noise: a maximum likelihood direction finding algorithm and cramer-rao bounds. *IEEE Wokshop on Statistical Signal Processing*, 2000.

[7] S.M. Kay. *Fundamentals of Statistical Signal Processing: Estimation Theory*, volume 1. Prentice Hall, Inc., New Jersey, 1993.

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