Multisource Self-Calibration for Sensor Arrays

Stefan J. Wijnholds, *Student Member, IEEE*, and Alle-Jan van der Veen, *Fellow, IEEE*

***Abstract*—Calibration of a sensor array is more involved if the antennas have direction dependent gains and multiple calibrator sources are simultaneously present. We study this case for a sensor array with arbitrary geometry but identical elements, i.e., elements with the same direction dependent gain pattern. A weighted alternating least squares (WALS) algorithm is derived that iteratively solves for the direction independent complex gains of the array elements, their noise powers and their gains in the direction of the calibrator sources. An extension of the problem is the case where the apparent calibrator source locations are unknown, e.g., due to refractive propagation paths. For this case, the WALS method is supplemented with weighted subspace fitting (WSF) direction finding techniques. Using Monte Carlo simulations we demonstrate that both methods are asymptotically statistically efficient and converge within two iterations even in cases of low SNR.**

***Index Terms*—Array signal processing, radio astronomy, self-calibration.**

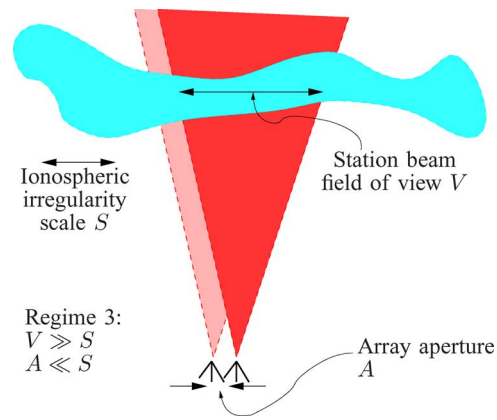

Fig. 1. Graphical representation of regime 3: all lines of sight towards a single source sample the same ionosphere, but the ionospheric delay differs per source due to the large FOV and short baselines (after [4]).

I. Introduction

In this paper we study the calibration of the direction dependent response of sensor array antennas, excited by simultaneously present calibrator sources. The antenna array has arbitrary geometry but identical antennas. The calibration involves the complex gains of the antenna elements towards each source. The source powers are known but we will allow for small deviations in apparent source locations to account for, e.g., refractive propagation paths. This problem is one of the main challenges currently faced in the field of radio astronomy. For low-frequency observations (<300 MHz) this community is building or developing a number of new instruments, for example the low frequency array (LOFAR) [1], the Murchison wide field array (MWA) [2] and the primeval structure telescope (PaST) [3], which are all large irregular phased arrays. These difficulties arise due to the influence of the ionosphere on the propagation of radio waves, which can qualitatively be categorized into the following four regimes depending on the field of view (FOV) of the individual receptors and the baselines between them [4].

1) All antennas and all lines of sight sample the same ionospheric delays, thus the ionosphere causes no distortion of the array manifold (small FOV, short baselines).
2) Lines of sight from different antennas sample different ionospheric patches, but all sources in the antenna FOV experience the same delay; the ionosphere thus causes an antenna based gain effect (small FOV, long baselines).
3) All lines of sight towards a single source sample the same ionosphere but the delay differs per source. This regime requires source based direction dependent calibration (large FOV, short baselines, see Fig. 1).
4) The ionospheric delays differ per station and per source (large FOV, long baselines).

The first two regimes can be handled under the self-calibration assumption used in astronomy that all errors are telescope based thereby reducing the estimation problem to a single direction independent gain factor per telescope [5]. The problem of estimating these direction independent gains based on a single calibrator source is treated in [6], while the multiple source case has been discussed in [7]. The problem of finding a complex gain per antenna per source, the fourth regime, is not tractable without further assumptions which allow to parameterize the behavior of the gains over space, time and/or frequency [8]. In this paper we will focus on the third regime, which is sketched in Fig. 1, thus filling the gap in the available literature. It allows for the calibration of individual closely packed groups of antennas such as a LOFAR station or a subarray of the MWA and PaST telescopes and thus forms a valuable step towards the calibration of the whole array.

We will state our problem in general terms, since the problem plays a role in a range of applications varying from underwater acoustics to antenna arrays. This explains the persistent interest in on-line calibration, autocalibration, or self-calibration of sensor arrays [8]–[12]. In most applications the main driver for studies on array calibration is to improve the DOA estimation accuracy. Many studies in this field, therefore, try to solve for

Manuscript received July 14, 2008; accepted March 27, 2009. First published May 12, 2009; current version published August 12, 2009. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Mats Bengtsson. This work was supported by the Netherlands Institute for Radio Astronomy (ASTRON) and by NWO-STW under the VICI programme (DTC.5893).

S. J. Wijnholds is with ASTRON, Dwingeloo, Dwingeloo NL-7991, The Netherlands (e-mail: wijnholds@astron.nl).

A.-J. van der Veen is with the Delft University of Technology, Delft, Delft 2628, The Netherlands (e-mail: a.j.vanderveen@tudelft.nl).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TSP.2009.2022894

1053-587X/$26.00 © 2009 IEEE

Authorized licensed use limited to: Technische Universiteit Delft. Downloaded on August 20, 2009 at 05:56 from IEEE Xplore. Restrictions apply.

the DOAs and a number of array parameters. In [9], [13], and [14] self-calibration schemes are presented which solve for the direction independent gains and the sensor positions, i.e., the directional response of the sensors is assumed known. Other studies assume a controlled environment to calibrate the array by measuring the calibrator sources one at a time [15], [16] or exploit the array geometry, e.g., a uniform linear array (ULA) having a Toeplitz matrix as array covariance matrix [10], [17]. Weiss and Friedlander [18] have presented a technique for almost fully blind signal estimation. Their work, however, focuses on estimation and separation of source signals, not on characterizing the array itself. We thus feel that the problem at hand also forms an interesting addition to the literature available on sensor array calibration in general.

In the next section, we introduce the data model and provide a mathematical formulation of the problem. In Section III the Cramér-Rao bound for this estimation problem is discussed. Section IV presents an alternating least squares (ALS) and a weighted alternating least squares (WALS) approach optimizing subsets of parameters iteratively. These algorithms are validated using Monte Carlo simulations in Section V. The conclusions are drawn in Section VI.

Notation: Overbar $\overline{(\cdot)}$ denotes conjugation, the transpose operator is denoted by ^{T}, the complex conjugate (Hermitian) transpose by ^{H} and the pseudoinverse by †. An estimated value is denoted by $\mathcal{E}\{\cdot\}$. $\odot$ is the element-wise matrix multiplication (Hadamard product), $\oslash$ is the element-wise matrix division, $\otimes$ denotes the Kronecker product and $\circ$ is used to denote the Khatri–Rao or column-wise Kronecker product of two matrices. $\mathrm{diag}(\cdot)$ converts a vector to a diagonal matrix with the vector placed on the main diagonal, $\mathrm{vecdiag}(\cdot)$ produces a vector from the elements of the main diagonal of its argument and $\mathrm{vec}(\cdot)$ converts a matrix to a vector by stacking the columns of the matrix. We exploit many properties of Kronecker products, including the following (for matrices and vectors of compatible dimensions):

$$\mathrm{vec}(\mathbf{ABC}) = (\mathbf{C}^T \otimes \mathbf{A})\mathrm{vec}(\mathbf{B}) \tag{1}$$

$$\mathrm{vec}\left(\mathbf{A}\mathrm{diag}(\mathbf{b})\mathbf{C}\right) = (\mathbf{C}^T \circ \mathbf{A})\mathbf{b} \tag{2}$$

$$(\mathbf{A} \circ \mathbf{B})^H(\mathbf{C} \circ \mathbf{D}) = \mathbf{A}^H\mathbf{C} \odot \mathbf{B}^H\mathbf{D}. \tag{3}$$

II. Data Model

We consider an array of p elements located at a single site ("Regime 3"). Denote the complex baseband output signal of the ith array element by $x_i(t)$ and define the array signal vector $\mathbf{x}(t) = [x_1(t), x_2(t), \cdots, x_p(t)]^T$. We assume the presence of q mutually independent i.i.d. Gaussian signals $s_k(t)$ impinging on the array, which are stacked in a $q \times 1$ vector $\mathbf{s}(t)$. Likewise the sensor noise signals $n_i(t)$ are assumed to be mutually independent i.i.d. Gaussian signals and are stacked in a $p \times 1$ vector $\mathbf{n}(t)$. If the narrow band condition holds [19], we can define the q spatial signature vectors $\mathbf{a}_k(p \times 1)$, which describe for the kth source the phase delays at each antenna due only to the geometry.

The q sources are considered calibrator sources, thus we assume that their powers, their nominal positions, and the locations of the antennas (hence $\mathbf{a}_k$) are known. Refractive effects caused by ionospheric phase gradients may shift the apparent locations of the sources, thus we will also consider cases where the $\mathbf{a}_k$ are only parametrically known.

The sensors are assumed to have the same direction dependent gain behavior towards the q source signals received by the array. This can include the antenna pattern and ionospheric phase effects. They are described by gain factors g_{0k} and are collected in a matrix $\mathbf{G_0} = \mathrm{diag}([g_{01}, g_{02}, \cdots, g_{0q}])$. The direction independent gains and phases (individual receiver gains) can be described as $\boldsymbol{\gamma} = [\gamma_1, \gamma_2, \cdots, \gamma_p]^T$ and $\boldsymbol{\phi} = [\mathrm{e}^{\mathrm{j}\phi_1}, \mathrm{e}^{\mathrm{j}\phi_2}, \cdots, \mathrm{e}^{\mathrm{j}\phi_p}]^T$, respectively, with corresponding diagonal matrix forms $\boldsymbol{\Gamma} = \mathrm{diag}(\boldsymbol{\gamma})$ and $\boldsymbol{\Phi} = \mathrm{diag}(\boldsymbol{\phi})$. With these definitions, the array signal vector can be described as

$$\mathbf{x}(t) = \boldsymbol{\Gamma\Phi}\left(\sum_{k=1}^{q} \mathbf{a}_k g_{0k} s_k(t)\right) + \mathbf{n}(t) = \mathbf{GAG}_0\mathbf{s}(t) + \mathbf{n}(t) \tag{4}$$

where $\mathbf{A} = [\mathbf{a}_1, \cdots, \mathbf{a}_k]$ (size $p \times q$) and $\mathbf{G} = \boldsymbol{\Gamma\Phi}$; for later use we also define $\mathbf{g} = \boldsymbol{\gamma} \odot \boldsymbol{\phi}$.

The signal is sampled with period T and N sample vectors are stacked into a data matrix $\mathbf{X} = [\mathbf{x}(T), \mathbf{x}(2T), \cdots, \mathbf{x}(NT)]$. The covariance matrix of $\mathbf{x}(t)$ is $\mathbf{R} = \mathcal{E}\{\mathbf{x}(t)\mathbf{x}^H(t)\}$ and is estimated by $\widehat{\mathbf{R}} = N^{-1}\mathbf{XX}^H$. Likewise, the source signal covariance $\boldsymbol{\Sigma}_s = \mathrm{diag}(\boldsymbol{\sigma}_s)$ where $\boldsymbol{\sigma}_s = [\sigma_{s1}^2, \sigma_{s2}^2, \cdots, \sigma_{sq}^2]^T$ and the noise covariance matrix is $\boldsymbol{\Sigma}_n = \mathrm{diag}(\boldsymbol{\sigma}_n)$ where $\boldsymbol{\sigma}_n = [\sigma_{n1}^2, \sigma_{n2}^2, \cdots, \sigma_{np}^2]^T$. Then the model for $\mathbf{R}$ based on (4) is

$$\mathbf{R} = \mathbf{GAG}_0\boldsymbol{\Sigma}_s\mathbf{G}_0^H\mathbf{A}^H\mathbf{G}^H + \boldsymbol{\Sigma}_n. \tag{5}$$

In this model, $\boldsymbol{\Sigma}_s$ is known. Since $\mathbf{G}_0$ and $\boldsymbol{\Sigma}_s$ are diagonal matrices, we can introduce

$$\begin{aligned}\boldsymbol{\Sigma} &= \mathbf{G}_0\boldsymbol{\Sigma}_s\mathbf{G}_0^H \\ &= \mathrm{diag}\left(\left[|g_{01}|^2\sigma_{s1}^2, \cdots, |g_{0q}|^2\sigma_{sq}^2\right]\right) = \mathrm{diag}(\boldsymbol{\sigma}).\end{aligned} \tag{6}$$

Since the direction dependent gains g_{0k} are not known, the real valued elements of $\boldsymbol{\sigma} = [\sigma_1^2, \sigma_2^2, \cdots, \sigma_q^2]^T$ are not known either. This implies that our problem is identical to estimating an unknown diagonal signal covariance matrix without any DOA dependent errors. We may thus restate (5) as

$$\mathbf{R} = \mathbf{GA\Sigma A}^H\mathbf{G}^H + \boldsymbol{\Sigma}_n \tag{7}$$

and solve for $\mathbf{G}$, $\boldsymbol{\Sigma}$ and $\boldsymbol{\Sigma}_n$ under the assumption that $\mathbf{A}$ is known (or parametrically known).

The data model described by (7) is commonly used in papers on sensor array calibration (e.g., [6] and [20]). Flanagan and Bell [9], Weiss and Friedlander [13], and See [14] effectively use the same model, but focus on position calibration of the array elements and are therefore more explicit on the form of $\mathbf{A}$. If the source positions and locations of the sensors within the array are known, an explicit formula for $\mathbf{a}_k$ can be used to compute the nominal spatial signature vectors. Estimation of source locations is required to account for the refractive effects produced by an ionospheric phase gradient.

The ith element of the array is located at $\mathbf{r}_i = [x_i, y_i, z_i]^T$. These positions can be stacked in a matrix $\mathcal{R} = [\mathbf{r}_1, \mathbf{r}_2, \cdots, \mathbf{r}_p]^T$ (size $p \times 3$). The position of the kth source can be denoted as $\mathbf{l}_k = [l_k, m_k, n_k]^T$. The source

Authorized licensed use limited to: Technische Universiteit Delft. Downloaded on August 20, 2009 at 05:56 from IEEE Xplore. Restrictions apply.

positions can be stacked in a matrix $\mathcal{L} = [\mathbf{l}_1, \mathbf{l}_2, \cdots, \mathbf{l}_q]^T$ (size $q \times 3$). The spatial signature matrix $\mathbf{A}$ can thus be described by

$$\mathbf{A} = \exp\left(-\mathrm{j}\frac{2\pi}{\lambda}\mathcal{R}\mathcal{L}^T\right) \tag{8}$$

where the exponential function is applied element-wise to its argument. In the remainder of this paper we will specialize to a planar array having $z_i = 0$ for convenience of presentation but without loss of generality.

From (7) we observe that $\mathbf{G}$ and $\mathbf{\Sigma}$ share a common scalar factor. Therefore we may impose the boundary condition $\sigma_1^2 = 1$. To solve the ambiguity in the phase solution of $\mathbf{G}$ we will take the first element as phase reference, i.e., $\phi_1 = 0$ is imposed.[1] When solving for source locations a similar problem occurs: a single rotation of all DOA vectors can be compensated by the direction independent gain phase solution. We will therefore fix the position of the first source.

In this paper we will address four related sensor array calibration problems based on this data model. These are summarized below where the parameter vectors adhering to the aforementioned boundary conditions are stated explicitly.

1) The sensor noise powers are the same for all elements, i.e., $\sigma_{n1}^2 = \sigma_{n2}^2 = \cdots = \sigma_{nq}^2$ such that $\mathbf{\Sigma}_n = \sigma_n^2\mathbf{I}$ where $\mathbf{I}$ is the identity matrix. In this scenario the parameter vector to be estimated is $\boldsymbol{\theta} = [\boldsymbol{\gamma}^T, \phi_2, \cdots \phi_p, \sigma_2^2, \cdots \sigma_q^2, \sigma_n^2]^T$.
2) The sensor noise powers are allowed to differ from one element to the other, i.e., $\mathbf{\Sigma}_n = \mathrm{diag}(\boldsymbol{\sigma}_n)$. In this case the parameter vector is $\boldsymbol{\theta} = [\boldsymbol{\gamma}^T, \phi_2, \cdots, \phi_p, \sigma_2^2, \cdots \sigma_q^2, \sigma_{n1}^2, \cdots \sigma_{np}^2]^T$.
3) $\mathbf{\Sigma}_n = \sigma_n^2\mathbf{I}$ and $\mathbf{A} = \mathbf{A}(\mathcal{L})$, i.e., similar to the first scenario but with unknown source locations. In this case $\boldsymbol{\theta} = [\boldsymbol{\gamma}^T, \phi_2, \cdots, \phi_p, \sigma_2^2, \cdots, \sigma_q^2, \sigma_n^2, \mathbf{l}_2^T, \cdots, \mathbf{l}_q^T]^T$.
4) $\mathbf{\Sigma}_n = \mathrm{diag}(\boldsymbol{\sigma}_n)$ and $\mathbf{A} = \mathbf{A}(\mathcal{L})$, giving $\boldsymbol{\theta} = [\boldsymbol{\gamma}^T, \phi_2, \cdots, \phi_p, \sigma_2^2, \cdots, \sigma_q^2, \sigma_{n1}^2, \cdots, \sigma_{np}^2, \mathbf{l}_2^T, \cdots, \mathbf{l}_q^T]^T$.

III. Cramér–Rao bound (CRB) Analysis

The CRB on the error variance for any unbiased estimator states that [22]

$$\mathbf{C} = \mathcal{E}\left\{(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta})(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta})^T\right\} \geq \frac{1}{N}\mathbf{J}^{-1} \tag{9}$$

where $\mathbf{C}$ is the covariance matrix of $\boldsymbol{\theta}$, and $\mathbf{C}$ is the Fisher Information Matrix (FIM). For Gaussian signals the FIM can be expressed as

$$\mathbf{J} = \mathbf{F}^H(\overline{\mathbf{R}}^{-1} \otimes \mathbf{R}^{-1})\mathbf{F} \tag{10}$$

where $\mathbf{F}$ is the Jacobian evaluated at the true values of the parameters, i.e.

$$\mathbf{F} = \left.\frac{\delta \mathrm{vec}(\mathbf{R})}{\delta \boldsymbol{\theta}^T}\right|_{\boldsymbol{\theta}}. \tag{11}$$

[1]In [21] it is shown that $\sum_{i=1}^p \phi_i = 0$ is the optimal constraint for this problem. This constraint has the disadvantage that the location of the phase reference is not well defined. Furthermore, the choice for the constraint used here simplifies our analysis in combination with the constraints required to uniquely identify the source locations and the apparent source powers.

For the calibration problem defined by the first scenario, the Jacobian can be partitioned in four parts following the structure in $\boldsymbol{\theta}$:

$$\mathbf{F} = [\mathbf{F}_\gamma, \mathbf{F}_\phi, \mathbf{F}_\sigma, \mathbf{F}_{\sigma_n}] \tag{12}$$

where it can be shown that

$$\mathbf{F}_\gamma = (\overline{\mathbf{G}\mathbf{R}_0}\mathbf{\Phi}) \circ \mathbf{I} + \mathbf{I} \circ (\mathbf{G}\mathbf{R}_0\overline{\mathbf{\Phi}}) \tag{13}$$

$$\mathbf{F}_\phi = \mathrm{j}\left((\overline{\mathbf{G}\mathbf{R}_0}\mathbf{G}) \circ \mathbf{I} - \mathbf{I} \circ (\mathbf{G}\mathbf{R}_0\overline{\mathbf{G}})\right)\mathbf{I}_s \tag{14}$$

$$\mathbf{F}_\sigma = \left((\overline{\mathbf{G}\mathbf{A}}) \circ (\mathbf{G}\mathbf{A})\right)\mathbf{I}_s \tag{15}$$

$$\mathbf{F}_{\sigma_n} = \mathrm{vec}(\mathbf{I}). \tag{16}$$

$\mathbf{I}_s$ is a selection matrix of appropriate size equal to the identity matrix with its first column removed so that the derivatives with respect to ϕ_1 and σ_1^2 are omitted.

In the second scenario the Jacobian can be partitioned in a similar way, the difference being that the expression for $\mathbf{F}_{\sigma_n}$ given by (16) should be replaced by

$$\mathbf{F}_{\sigma_n} = \mathbf{I} \circ \mathbf{I}. \tag{17}$$

In the third and fourth scenarios an additional component $\mathbf{F}_\mathbf{l}$ is added to the Jacobian containing the derivatives of $\mathrm{vec}(\mathbf{R})$ with respect to the source position coordinates. This component can be partitioned as $\mathbf{F}_\mathbf{l} = [\mathbf{F}_l, \mathbf{F}_m]$. Introducing

$$\mathbf{G}_x = \mathrm{diag}\left([x_1, x_2, \cdots x_p]^T\right)\mathbf{G} \tag{18}$$

$$\mathbf{G}_y = \mathrm{diag}\left([y_1, y_2, \cdots y_p]^T\right)\mathbf{G} \tag{19}$$

these components can be conveniently written as

$$\mathbf{F}_l = -\mathrm{j}\frac{2\pi}{\lambda}(\overline{\mathbf{G}\mathbf{A}} \circ \mathbf{G}_x\mathbf{A} - \overline{\mathbf{G}_x\mathbf{A}} \circ \mathbf{G}\mathbf{A})\mathbf{\Sigma}\mathbf{I}_s \tag{20}$$

$$\mathbf{F}_m = -\mathrm{j}\frac{2\pi}{\lambda}(\overline{\mathbf{G}\mathbf{A}} \circ \mathbf{G}_y\mathbf{A} - \overline{\mathbf{G}_y\mathbf{A}} \circ \mathbf{G}\mathbf{A})\mathbf{\Sigma}\mathbf{I}_s. \tag{21}$$

These equations show that the entries of the Jacobian related to derivatives with respect to the l- and m-coordinates of the sources are proportional to the x- and y-coordinates of the array elements respectively. The physical interpretation of this relation is that a plane wave propagating along the coordinate axis of the coordinate to be estimated provides a more useful test signal to estimate the source location than a signal propagating perpendicular to this axis.

IV. Algorithms

A. Generalized Least Squares Formulation

An asymptotically efficient estimate of the model parameters $\boldsymbol{\theta}$ can be obtained via the ML formulation. Since all signals are assumed to be i.i.d. Gaussian signals, the derivation is standard and ML parameter estimates for N independent samples are obtained by minimizing the negative log-likelihood function [23]

$$\widehat{\boldsymbol{\theta}} = \arg\min_{\boldsymbol{\theta}}\left(\ln|\mathbf{R}(\boldsymbol{\theta})| + \mathrm{tr}\left(\mathbf{R}^{-1}(\boldsymbol{\theta})\widehat{\mathbf{R}}\right)\right) \tag{22}$$

where $\mathbf{R}(\boldsymbol{\theta})$ is the model covariance matrix as function of $\boldsymbol{\theta}$ and $\widehat{\mathbf{R}}$ is the sample covariance matrix $N^{-1}\mathbf{X}\mathbf{X}^H$.

Authorized licensed use limited to: Technische Universiteit Delft. Downloaded on August 20, 2009 at 05:56 from IEEE Xplore. Restrictions apply.

It does not seem possible to solve this minimization problem in closed form. As discussed in [23] a weighted least squares covariance matching approach is known to lead to estimates that are, for a large number of samples, equivalent to ML estimates and are therefore asymptotically efficient and reach the CRB.

The least squares covariance model fitting problem can be defined as

$$\widehat{\boldsymbol{\theta}} = \arg\min_{\boldsymbol{\theta}} \left\| \widehat{\mathbf{R}} - \mathbf{R}(\boldsymbol{\theta}) \right\|_F^2. \tag{23}$$

Equivalently, we consider minimization of the cost function

$$\kappa(\boldsymbol{\theta}) = \left\| \widehat{\mathbf{R}} - \mathbf{R}(\boldsymbol{\theta}) \right\|_F^2 = \mathbf{f}^H(\boldsymbol{\theta})\mathbf{f}(\boldsymbol{\theta}) \tag{24}$$

where $\mathbf{f}(\boldsymbol{\theta}) = \mathrm{vec}(\widehat{\mathbf{R}} - \mathbf{R}(\boldsymbol{\theta}))$. The more general weighted least squares problem is obtained by introducing a weighting matrix $\mathbf{W}$ and optimizing

$$\kappa_{\mathbf{W}}(\boldsymbol{\theta}) = \mathbf{f}^H(\boldsymbol{\theta})\mathbf{W}\mathbf{f}(\boldsymbol{\theta}). \tag{25}$$

The optimal weight is known to be the inverse of the asymptotic covariance of the residuals, $\mathcal{E}\{\mathbf{f}(\boldsymbol{\theta}_0)\mathbf{f}(\boldsymbol{\theta}_0)^H\}$, where $\boldsymbol{\theta}_0$ is the true value of the parameters [23]. The optimal weight for Gaussian sources is thus

$$\mathbf{W}_{opt} = (\overline{\mathbf{R}} \otimes \mathbf{R})^{-1} = \overline{\mathbf{R}}^{-1} \otimes \mathbf{R}^{-1}. \tag{26}$$

The Kronecker structure of $\mathbf{W}_{opt}$ allows to introduce $\mathbf{W}_c = \mathbf{R}^{-(1/2)}$ and write the weighted least squares (WLS) cost function as

$$\kappa_{\mathbf{W}}(\boldsymbol{\theta}) = \left\| \mathbf{W}_c \left(\widehat{\mathbf{R}} - \mathbf{R}(\boldsymbol{\theta}) \right) \mathbf{W}_c \right\|_F^2. \tag{27}$$

As mentioned, this estimator is asymptotically unbiased and asymptotically efficient [23].

We propose to solve the least squares problems for the four cases defined in Section II by alternating between least squares solutions to subsets of parameters. The subsets are chosen such that the solution is either available in the literature or can be derived analytically. We will have four subsets of parameters: the complex direction independent gains of the elements $\mathbf{g}$, the apparent source powers $\boldsymbol{\sigma}$, the receiver noise powers $\boldsymbol{\sigma}_n$, and the source locations $\mathcal{L}$. In the following subsections we will develop the unweighted and weighted least squares solutions for these subsets before putting them together to form the alternating least squares (ALS) or weighted alternating least squares (WALS) method respectively.

B. Estimation of Direction Independent Gains

The least squares problem to find the omnidirectional gains $\widehat{\mathbf{g}}$ based on the weighted cost function $\kappa_{\mathbf{W}}(\boldsymbol{\theta})$ can be formulated as

$$\begin{aligned}\widehat{\mathbf{g}} = \arg\min_{\mathbf{g}} \Big\| &(\overline{\mathbf{W}}_c \otimes \mathbf{W}_c)\mathrm{vec}(\widehat{\mathbf{R}} - \boldsymbol{\Sigma}_n) + \\ &-(\overline{\mathbf{W}}_c \otimes \mathbf{W}_c)\mathrm{diag}\left(\mathrm{vec}(\mathbf{R}_0)\right)(\overline{\mathbf{g}} \otimes \mathbf{g}) \Big\|_F^2\end{aligned} \tag{28}$$

where we have introduced $\mathbf{R}_0 = \mathbf{A}\boldsymbol{\Sigma}\mathbf{A}^H$; in this subsection, $\mathbf{R}_0$ is assumed to be known. This problem can be solved using standard techniques by regarding $\mathbf{g}$ and $\overline{\mathbf{g}}$ as independent vector parameters and alternatingly solve for them until convergence.[2] In this approach, the solution for $\mathbf{g}$ is

$$\begin{aligned}\widehat{\mathbf{g}} &= \arg\min_{\mathbf{g}} \Big\| (\overline{\mathbf{W}}_c \otimes \mathbf{W}_c)\mathrm{vec}(\widehat{\mathbf{R}} - \boldsymbol{\Sigma}_n) + \\ &\quad -(\overline{\mathbf{W}}_c \otimes \mathbf{W}_c)\left((\mathbf{R}_0\mathbf{G}^H) \circ \mathbf{I}\right)\mathbf{g} \Big\|_F^2 \\ &= \left(\overline{\mathbf{R}_0^H\mathbf{G}^H\mathbf{R}^{-1}\mathbf{G}\mathbf{R}_0} \odot \mathbf{R}^{-1}\right)^{-1} \\ &\quad \times \left(\overline{\mathbf{R}^{-1}\mathbf{G}\mathbf{R}_0} \circ \mathbf{R}^{-1}\right)^H \mathrm{vec}(\widehat{\mathbf{R}} - \boldsymbol{\Sigma}_n).\end{aligned} \tag{29}$$

Since the result for $\overline{\mathbf{g}}$ is simply the complex conjugate of this relation, it is sufficient to apply (29) repeatedly until convergence. Although ALS ensures that the value of the cost function decreases in each iteration, it does not guarantee convergence to the global minimum, especially if the initial estimate is poor. The number of iterations required for convergence also depends strongly on the vicinity of the initial estimate to the true value. We are thus interested in obtaining a good initial estimate for $\mathbf{g}$.

Fuhrmann [20] has proposed to use a suboptimal closed form solution to initialize the Newton iterations used to solve the ML cost function under the assumption that $\boldsymbol{\Sigma}_n$ is known (or actually no noise is present). A similar problem (with unknown $\boldsymbol{\Sigma}_n$ and $\mathbf{R}_0$ a rank 1 matrix, i.e., a single calibrator source) was also studied by us in [6], where a "column ratio method" (COLR) was proposed. Below, we generalize and improve that technique.[3] All these techniques are suboptimal in the case of multiple calibrator sources, but they can be used to provide the starting point for iterative refinement.

The closed form solution by [20] is derived as follows. If the structure in $\mathbf{v} = \overline{\mathbf{g}} \otimes \mathbf{g}$ is neglected and $\boldsymbol{\Sigma}_n$ is known or negligible, we can solve for $\mathbf{v}$ in the least squares sense by

$$\widehat{\mathbf{v}} = \mathrm{diag}\left(\mathrm{vec}(\mathbf{R}_0)\right)^{-1}\mathrm{vec}(\widehat{\mathbf{R}} - \boldsymbol{\Sigma}_n) \tag{30}$$

or equivalently

$$\widehat{\mathbf{g}\mathbf{g}^H} = (\widehat{\mathbf{R}} - \boldsymbol{\Sigma}_n) \oslash \mathbf{R}_0. \tag{31}$$

Note that the weights cancel since we solve for one parameter for each entry of $\widehat{\mathbf{R}}$. Subsequently, use the structure of $\mathbf{v}$ and assume that the estimate of $\mathbf{g}\mathbf{g}^H$ obtained in the first step has rank 1. An eigenvalue decomposition is used to extract $\widehat{\mathbf{g}}$ as the dominant eigenvector from $\widehat{\mathbf{v}}$. It is clear however that the noise on specific elements of $\widehat{\mathbf{v}}$ may be increased considerably if some entries of $\mathbf{R}_0$ have a small value. Also, the method is not applicable if $\boldsymbol{\Sigma}_n$ is unknown.

In [7] another approach is therefore suggested based on the observation that $g_i\overline{g}_k = R_{ik}/R_{0,ik}$ holds for all off-diagonal elements of $\mathbf{R}$, i.e., for $i \neq k$. This implies that

$$\frac{g_i}{g_j} = \frac{g_i\overline{g}_k}{g_j\overline{g}_k} = \frac{R_{ik}/R_{0,ik}}{R_{jk}/R_{0,jk}} = \frac{R_{ik}R_{0,jk}}{R_{0,ik}R_{jk}} \quad (k \neq i, j). \tag{32}$$

This relation is similar to the closure amplitude relation in astronomy [5], [24], which states that in an observation on a *single* point source or, more generally, in the case of a rank 1 model, the

[2]We thank one of the reviewers for pointing this out.

[3]An initial presentation of this was in [7].

Authorized licensed use limited to: Technische Universiteit Delft. Downloaded on August 20, 2009 at 05:56 from IEEE Xplore. Restrictions apply.

amplitude ratios are related as indicated by the second equality sign in (32).

Since the index k can be chosen freely as long as $k \neq i, j$, we can introduce $\mathbf{c}_{1,ij}$ being the column vector containing the values $R_{ik}R_{0,jk}$ and $\mathbf{c}_{2,ij}$ being the column vector containing the values $R_{0,ik}R_{jk}$ for all possible values of $k \neq i, j$. We can now write (32) in its more general form

$$\frac{g_i}{g_j}\mathbf{c}_{2,ij} = \mathbf{c}_{1,ij}, \tag{33}$$

which has the well-known solution

$$\frac{g_i}{g_j} = \mathbf{c}_{2,ij}^{\dagger}\mathbf{c}_{1,ij}. \tag{34}$$

All possible gain ratios can be collected in a matrix $\mathbf{M}$ with entries $M_{ij} = \mathbf{c}_{2,ij}^{\dagger}\mathbf{c}_{1,ij}$. Since the model for $\mathbf{M}$ is $M_{ij} = g_i/g_j$, the matrix $\mathbf{M}$ is expected to be of rank 1, and $\mathbf{g}$ can be extracted from this matrix using an eigenvalue decomposition: let $\mathbf{u}_1$ be the eigenvector corresponding to the largest eigenvalue of $\mathbf{M}$, then $\mathbf{g} = \alpha\mathbf{u}_1$, where the scaling α needs to be determined separately since the quotient g_i/g_j is insensitive to modification by a constant scaling factor applied to all gains. This scaling can be found by minimizing

$$\widehat{\alpha} = \arg\min_{\alpha} \|\alpha\mathbf{G}\mathbf{R}_0\mathbf{G}^H\overline{\alpha} - \widehat{\mathbf{R}}\|_F^2 \tag{35}$$

for all off-diagonal elements. By introducing the vectors $\mathbf{r}_0 = \mathrm{vec}_{-}(\mathbf{G}\mathbf{R}_0\mathbf{G}^H)$ and $\mathbf{r} = \mathrm{vec}_{-}(\widehat{\mathbf{R}})$, where $\mathrm{vec}_{-}(\cdot)$ operates like the $\mathrm{vec}(\cdot)$ operator but leaves out the elements on the main diagonal of its argument, this cost function can rewritten as

$$\widehat{\alpha} = \arg\min_{\alpha} \left\||\alpha|^2\mathbf{r}_0 - \mathbf{r}\right\|_F^2 \tag{36}$$

which can be solved in the least squares sense by $\widehat{\alpha} = \sqrt{\mathbf{r}_0^{\dagger}\mathbf{r}}$.

Equation (34) extends the column ratio method proposed in [6] to all elements of the matrix. The column ratio method was introduced to reconstruct the main diagonal of a rank 1 source model, which allows to estimate this rank 1 model using an eigenvalue decomposition without distortion of the results due to unknown sensor noise powers. In our case this allows us to neglect the unknown receiver noise powers. The CRB analysis in [25] shows that it does not matter whether one simultaneously estimates the receiver noise powers and the omnidirectional complex gains exploiting all data, i.e., including the autocorrelations, or ignore the autocorrelations and solve for the direction independent complex gains only.

The Monte Carlo simulations presented in [7] suggest that the method outlined above provides a statistically efficient estimate of $\widehat{\mathbf{g}}$; the pseudoinverse ensures that the noise on the entries of $\mathbf{M}$ does not increase dramatically due to small values in either $\widehat{\mathbf{R}}$ or $\mathbf{R}_0$. It also provides a modification of the initial estimate proposed by Fuhrmann [20] which finds a near optimal solution to the least squares gain estimation problem without further optimization using, e.g., the Newton algorithm. This will generally save computational effort. In the simulations in this paper, we will use (34) to demonstrate that this method gives statistically efficient results without requiring additional iterations using (29). Moreover, if three or more iterations are required to ensure convergence using (29), it also requires less computational effort, as discussed in Section IV-G.

Although the pseudoinverse in (34) ensures robustness against small entries of either $\widehat{\mathbf{R}}$ or $\mathbf{R}_0$, the fact that the method relies on gain ratios may lead to poor performance if there are small entries in $\mathbf{g}$, e.g., due to failing array elements. This risk can be mitigated by rewriting (33) to

$$\mathbf{c}_{2,ij}g_i = \mathbf{c}_{1,ij}g_j. \tag{37}$$

By defining

$$\mathbf{Y}_i = \left[\mathbf{c}_{2,i1}^T, \mathbf{c}_{2,i2}^T, \cdots, \mathbf{c}_{2,ip}^T\right]^T$$

$$\mathbf{Y}_i = \begin{bmatrix} \mathbf{c}_{1,i1} & & & \\ & \mathbf{c}_{1,i2} & & \\ & & \ddots & \\ & & & \mathbf{c}_{1,ip} \end{bmatrix}$$

$$\mathbf{C}_1 = \left[\mathbf{Y}_1^T, \mathbf{Y}_2^T, \cdots, \mathbf{Y}_p^T\right]^T$$

$$\mathbf{C}_2 = \begin{bmatrix} \mathbf{Y}_1 & & & \\ & \mathbf{Y}_2 & & \\ & & \ddots & \\ & & & \mathbf{Y}_p \end{bmatrix}$$

we can enumerate i and j and collect all relations in a single matrix equation

$$\mathbf{C}_2\mathbf{g} = \mathbf{C}_1\mathbf{g}. \tag{38}$$

This suggests that $\mathbf{g}$ can be found by searching for the null space of the $p(p-1)^2 \times p$ matrix $\mathbf{C}_2 - \mathbf{C}_1$. By substituting the details of $\mathbf{c}_{1,ij}$ and $\mathbf{c}_{2,ij}$ in $\mathbf{C}_2 - \mathbf{C}_1$, it is easily seen that $\mathbf{g}$ lies indeed in the null space and that this null space is one dimensional. However, this is only true for noise free data. In the practical case of noisy data, there will not be a null space and $\mathbf{g}$ lies in the noise subspace. Furthermore, finding $\mathbf{g}$ will involve a singular value decomposition on a very large matrix which may be computationally prohibitive.

An alternative approach follows from recognizing that $\mathbf{g}$ is obtained from the principal right eigenvector of

$$\begin{aligned}\mathbf{C}_2^{\dagger}\mathbf{C}_1 &= \left[\frac{\sum_k \overline{R}_{0,ik}\overline{R}_{jk}R_{0,jk}R_{ik}}{\sum_j\sum_k \overline{R}_{0,ik}\overline{R}_{jk}R_{0,ik}R_{jk}}\right] \\ &= \left[\frac{g_i\overline{g}_j\sum_k |g_k|^2|R_{0,ik}|^2|R_{0,jk}|^2}{\sum_j |g_j|^2\sum_k |g_k|^2|R_{0,ik}|^2|R_{0,jk}|^2}\right]\end{aligned} \tag{39}$$

which is easily found by substitution of the definition of $\mathbf{c}_{1,ij}$ and $\mathbf{c}_{2,ij}$ and use of the Moore–Penrose left inverse. Equation (39) shows that $\mathbf{g}$ is the principal right eigenvector of $\mathbf{C}_2^{\dagger}\mathbf{C}_1$, which has eigenvalue 1 in the noise free case. Simulations with completely randomized $\mathbf{R}_0$ and $\mathbf{g}$ suggest that the other eigenvalues are considerably lower than 1. Since it is easily demonstrated that the noise on actual data will lower the main eigenvalue, this is an important result; the contrast between the largest and the second-largest eigenvalue determines the susceptibility of this method to noise on the data.

Note that these methods are still insensitive to a constant scaling factor applied to all gains. This ambiguity is solved by (36). Also note that all methods presented above enumerate over

Authorized licensed use limited to: Technische Universiteit Delft. Downloaded on August 20, 2009 at 05:56 from IEEE Xplore. Restrictions apply.

$i, j \neq k$ to avoid the diagonal entries of the array covariance matrix which are affected by the system noise. By imposing more stringent restrictions on the choice of i and j, this method can also handle cases in which some elements in the array experience correlator noise, i.e. cases in which $\mathbf{\Sigma}_n$ is not diagonal but still has sufficient zero entries to allow the above methods to extract the required information for every combination of g_i and g_j.

C. Source Power Estimation

Based on the unweighted cost function $\kappa(\boldsymbol{\theta})$, $\widehat{\boldsymbol{\sigma}}$ is found by

$$\begin{aligned}\widehat{\boldsymbol{\sigma}} &= \arg\min_{\boldsymbol{\sigma}} \left\| \widehat{\mathbf{R}} - \mathbf{G}\mathbf{A}\mathbf{\Sigma}\mathbf{A}^H\mathbf{G}^H - \mathbf{\Sigma}_n \right\|_F^2 \\ &= \arg\min_{\boldsymbol{\sigma}} \left\| \mathrm{vec}\left((\widehat{\mathbf{R}} - \mathbf{\Sigma}_n) - (\mathbf{G}\mathbf{A})\mathbf{\Sigma}(\mathbf{G}\mathbf{A})^H \right) \right\|_F^2 \\ &= \arg\min_{\boldsymbol{\sigma}} \left\| \mathrm{vec}(\widehat{\mathbf{R}} - \mathbf{\Sigma}_n) - (\overline{\mathbf{G}\mathbf{A}} \circ \mathbf{G}\mathbf{A})\boldsymbol{\sigma} \right\|_F^2 \\ &= (\overline{\mathbf{G}\mathbf{A}} \circ \mathbf{G}\mathbf{A})^\dagger \mathrm{vec}(\widehat{\mathbf{R}} - \mathbf{\Sigma}_n). \end{aligned} \tag{40}$$

If weighting is applied, $\widehat{\boldsymbol{\sigma}}$ follows from

$$\begin{aligned}\widehat{\boldsymbol{\sigma}} &= \arg\min_{\boldsymbol{\sigma}} \Big\| \mathbf{W}_c(\widehat{\mathbf{R}} - \mathbf{\Sigma}_n)\mathbf{W}_c \\ &\quad - \mathbf{W}_c\mathbf{G}\mathbf{A}\mathbf{\Sigma}\mathbf{A}^H\mathbf{G}^H\mathbf{W}_c \Big\|_F^2 \\ &= \arg\min_{\boldsymbol{\sigma}} \Big\| (\overline{\mathbf{W}}_c \otimes \mathbf{W}_c)\mathrm{vec}(\widehat{\mathbf{R}} - \mathbf{\Sigma}_n) \\ &\quad - (\overline{\mathbf{W}_c\mathbf{G}\mathbf{A}}) \circ (\mathbf{W}_c\mathbf{G}\mathbf{A})\boldsymbol{\sigma}_n \Big\|_F^2 \\ &= \left((\overline{\mathbf{W}_c\mathbf{G}\mathbf{A}}) \circ (\mathbf{W}_c\mathbf{G}\mathbf{A}) \right)^\dagger \\ &\quad \times (\overline{\mathbf{W}}_c \otimes \mathbf{W}_c)\mathrm{vec}(\widehat{\mathbf{R}} - \mathbf{\Sigma}_n). \end{aligned} \tag{41}$$

Using standard relations for Khatri–Rao and Kronecker products and substituting $\mathbf{W}_c = \mathbf{R}^{-(1/2)}$ we obtain

$$\begin{aligned}\widehat{\boldsymbol{\sigma}} &= \left((\overline{\mathbf{A}^H\mathbf{G}^H\mathbf{R}^{-1}\mathbf{G}\mathbf{A}}) \odot (\mathbf{A}^H\mathbf{G}^H\mathbf{R}^{-1}\mathbf{G}\mathbf{A}) \right)^{-1} \\ &\quad \times \mathrm{vecdiag}\left(\mathbf{A}^H\mathbf{G}^H\mathbf{R}^{-1}(\widehat{\mathbf{R}} - \mathbf{\Sigma})\mathbf{R}^{-1}\mathbf{G}\mathbf{A} \right). \end{aligned} \tag{42}$$

This result confirms the observation by Ottersten, Stoica and Roy [23] that although the derivation involves a square root of the array covariance matrix, the final result only depends on the array covariance matrix and its inverse.

D. Estimating Receiver Noise Powers

If $\mathbf{\Sigma}_n = \mathrm{diag}(\boldsymbol{\sigma}_n)$ and no weighting is applied to the cost function, then $\widehat{\boldsymbol{\sigma}}_n$ is found by solving

$$\widehat{\boldsymbol{\sigma}}_n = \arg\min_{\boldsymbol{\sigma}_n} \left\| \widehat{\mathbf{R}} - \mathbf{G}\mathbf{A}\mathbf{\Sigma}\mathbf{A}^H\mathbf{G}^H - \mathbf{\Sigma}_n \right\|_F^2. \tag{43}$$

This estimation problem is the same as the sensor noise estimation problem treated in [6], so we just state the result, which is found to be

$$\widehat{\boldsymbol{\sigma}}_n = \mathrm{diag}(\widehat{\mathbf{R}} - \mathbf{G}\mathbf{A}\mathbf{\Sigma}\mathbf{A}^H\mathbf{G}^H). \tag{44}$$

If $\mathbf{\Sigma}_n = \sigma_n^2\mathbf{I}$, a similar derivation gives

$$\widehat{\sigma}_n^2 = \frac{1}{p}\mathrm{tr}(\widehat{\mathbf{R}} - \mathbf{G}\mathbf{A}\mathbf{\Sigma}\mathbf{A}^H\mathbf{G}^H). \tag{45}$$

This result is just the average of the sensor noise estimates obtained when they are estimated individually.

If the weighted cost function $\kappa_{\mathbf{W}}(\boldsymbol{\theta})$ is used, we can follow a derivation similar to the one that led us to the estimate for $\widehat{\boldsymbol{\sigma}}$ in the previous section. We first note that

$$\begin{aligned}\widehat{\boldsymbol{\sigma}}_n &= \arg\min_{\boldsymbol{\sigma}_n} \left\| \mathbf{W}_c(\widehat{\mathbf{R}} - \mathbf{G}\mathbf{R}_0\mathbf{G}^H)\mathbf{W}_c - \mathbf{W}_c\mathbf{\Sigma}_n\mathbf{W}_c \right\|_F^2 \\ &= \arg\min_{\boldsymbol{\sigma}_n} \Big\| (\overline{\mathbf{W}}_c \otimes \mathbf{W}_c)\mathrm{vec}(\widehat{\mathbf{R}} - \mathbf{G}\mathbf{R}_0\mathbf{G}^H) \\ &\qquad - (\overline{\mathbf{W}}_c \circ \mathbf{W}_c)\boldsymbol{\sigma}_n \Big\|_F^2 \\ &= (\overline{\mathbf{W}}_c \circ \mathbf{W}_c)^\dagger(\overline{\mathbf{W}}_c \otimes \mathbf{W}_c)\mathrm{vec}(\widehat{\mathbf{R}} - \mathbf{G}\mathbf{R}_0\mathbf{G}^H). \end{aligned} \tag{46}$$

Applying the standard Khatri–Rao and Kronecker product relations and inserting $\mathbf{W}_c = \mathbf{R}^{-(1/2)}$ we get

$$\begin{aligned}\widehat{\boldsymbol{\sigma}}_n &= (\overline{\mathbf{R}}^{-1} \odot \mathbf{R}^{-1})^{-1} \\ &\quad \times \mathrm{vecdiag}\left(\mathbf{R}^{-1}(\widehat{\mathbf{R}} - \mathbf{G}\mathbf{R}_0\mathbf{G}^H)\mathbf{R}^{-1} \right). \end{aligned} \tag{47}$$

A similar derivation for $\mathbf{\Sigma}_n = \sigma_n^2\mathbf{I}$ gives

$$\widehat{\sigma}_n^2 = \|\mathbf{R}^{-1}\|_F^{-2}\mathrm{tr}\left(\mathbf{R}^{-1}(\widehat{\mathbf{R}} - \mathbf{G}\mathbf{R}_0\mathbf{G}^H)\mathbf{R}^{-1} \right). \tag{48}$$

The true value of the covariance matrix $\mathbf{R}$ is not known in practical situations. Therefore the measured covariance matrix $\widehat{\mathbf{R}}$ is generally taken as estimate of $\mathbf{R}$. It can be shown that this conventional approach leads to a noticeable bias in the estimate of $\widehat{\boldsymbol{\sigma}}_n$ for a finite number of samples N. For simplicity of the argument we will prove this statement for the case where $\mathbf{R} = \sigma_n^2\mathbf{I}$, such that $\mathbf{G}\mathbf{R}_0\mathbf{G}^H = 0$. Inserting this in (48) gives

$$\begin{aligned}\widehat{\sigma}_n^2 &= \|\widehat{\mathbf{R}}^{-1}\|_F^{-2}\mathrm{tr}(\widehat{\mathbf{R}}^{-1}\widehat{\mathbf{R}}\widehat{\mathbf{R}}^{-1}) \\ &= \|\widehat{\mathbf{R}}^{-1}\|_F^{-2}\mathrm{tr}(\widehat{\mathbf{R}}^{-1}). \end{aligned} \tag{49}$$

Since $\mathbf{R}^{-1} = \sigma_n^{-2}\mathbf{I}$, we can write $\widehat{\mathbf{R}}^{-1} = \sigma_n^{-2}(\mathbf{I} + \mathbf{E})$ where $\sigma_n^{-2}\mathbf{E} = \widehat{\mathbf{R}}^{-1} - \mathbf{R}^{-1}$ represents a specific realization of the noise on the elements of $\widehat{\mathbf{R}}^{-1}$. After substitution in (49) we obtain

$$\begin{aligned}\widehat{\sigma}_n^2 &= \left\| \sigma_n^{-2}(\mathbf{I} + \mathbf{E}) \right\|_F^{-2} \mathrm{tr}\left(\sigma_n^{-2}(\mathbf{I} + \mathbf{E}) \right) \\ &= \sigma_n^2 \mathrm{tr}\left((\mathbf{I} + \mathbf{E})(\mathbf{I} + \mathbf{E})^H \right)^{-1} \mathrm{tr}(\mathbf{I} + \mathbf{E}) \\ &= \sigma_n^2 \left(\mathrm{tr}(\mathbf{I}) + \mathrm{tr}(\mathbf{E}) + \mathrm{tr}(\mathbf{E}^H) + \mathrm{tr}(\mathbf{E}\mathbf{E}^H) \right)^{-1} \\ &\quad \times \left(\mathrm{tr}(\mathbf{I}) + \mathrm{tr}(\mathbf{E}) \right). \end{aligned} \tag{50}$$

Since $\mathbf{E}$ represents the noise on the data, the expected value of $\mathrm{tr}(\mathbf{E})$ and $\mathrm{tr}(\mathbf{E}^H)$ is zero. This implies that

$$\begin{aligned}\widehat{\sigma}_n^2 &\approx \sigma_n^2 \left(\mathrm{tr}(\mathbf{I}) + \mathrm{tr}(\mathbf{E}\mathbf{E}^H) \right)^{-1} \mathrm{tr}(\mathbf{I}) \\ &= \sigma_n^2 \frac{\|\mathbf{I}\|_F^2}{\|\mathbf{I}\|_F^2 + \|\mathbf{E}\|_F^2}. \end{aligned} \tag{51}$$

Authorized licensed use limited to: Technische Universiteit Delft. Downloaded on August 20, 2009 at 05:56 from IEEE Xplore. Restrictions apply.

This shows that the presence of noise systematically lowers the value of the estimate of σ_n^2. Since $\widehat{\mathbf{R}}^{-1}$ asymptotically in N converges to $\mathbf{R}^{-1}$, $\mathbf{E}$ converges to zero. Therefore this result also shows that the σ_n^2 estimate asymptotically converges to the true value if the number of samples N approaches infinity.

However, this bias can be avoided by using the best available knowledge of the estimated parameters, i.e., in (48) use

$$\mathbf{R} = \widehat{\mathbf{\Gamma}}\mathbf{A}(\widehat{\mathcal{L}})\widehat{\mathbf{\Sigma}}\mathbf{A}^H(\widehat{\mathcal{L}})\widehat{\mathbf{\Gamma}}^H + \widehat{\mathbf{\Sigma}}_n. \tag{52}$$

In the first iteration of an alternating least squares algorithm, initial estimates of the parameters are used. These values are replaced by increasingly accurate estimated values in consecutive iterations.

E. DOA Estimation

The problem of estimating $\mathcal{L}$ is that of estimating the direction of arrival of signals impinging on a sensor array, and has been studied extensively. MUSIC [26] and weighted subspace fitting (WSF) [27], [28] are well-known statistically efficient DOA estimation methods applicable to arbitrary sensor arrays. In either case the eigenvalue decomposition of $\mathbf{R}$ is interpreted in terms of a noise subspace and a signal subspace, i.e.

$$\mathbf{R} = \sum_{i=1}^{p} \lambda_i \mathbf{e}_i \mathbf{e}_i^H = \mathbf{E}_s \mathbf{\Lambda}_s \mathbf{E}_s^H + \mathbf{E}_n \mathbf{\Lambda}_n \mathbf{E}_n^H \tag{53}$$

where $\lambda_1 > \cdots > \lambda_q > \lambda_{q+1} \geq \cdots \geq \lambda_p$. The noise eigenvalues of the whitened true covariance matrix are all equal and the number of sources can be derived from the distribution of the eigenvalues of $\mathbf{R}$; we will assume that it is known. In practical situations in which only an estimate of the true covariance matrix is available, methods like the exponential fitting test [29], [30] or information theoretic criteria [31] must be used to determine the number of signals.

It is quite straightforward to adapt the WSF method to our needs. The data model assumed in [28] can be described as

$$\mathbf{R} = \widetilde{\mathbf{A}}(\mathcal{L})\mathbf{\Sigma}\widetilde{\mathbf{A}}^H(\mathcal{L}) + \sigma_n^2\mathbf{I} \tag{54}$$

where we have used $\widetilde{\mathbf{A}}$ to avoid confusion with $\mathbf{A}$ introduced in this paper. To map our data model on the data model described by (54) we should whiten our array covariance matrix using

$$\mathbf{R}_w = \mathbf{\Sigma}_n^{-\frac{1}{2}}\mathbf{R}\mathbf{\Sigma}_n^{-\frac{1}{2}} = \mathbf{\Sigma}_n^{-\frac{1}{2}}\mathbf{G}\mathbf{A}\mathbf{\Sigma}\mathbf{A}^H\mathbf{G}^H\mathbf{\Sigma}_n^{-\frac{1}{2}} + \mathbf{I} \tag{55}$$

and then set $\widetilde{\mathbf{A}} = \mathbf{\Sigma}_n^{-1/2}\mathbf{G}\mathbf{A}$. With these substitutions it is straightforward to implement the procedures described in [28].

F. Alternating Least Squares (ALS) and Weighted Alternating Least Squares (WALS)

The ingredients of the previous four subsections can be combined to formulate a WALS solution to the stated optimization problems. We start by introducing an algorithm handling the first two scenarios identified in Section II. DOA estimation is then added in a straightforward way.

To estimate $\widehat{\mathbf{g}}$, $\widehat{\boldsymbol{\sigma}}$ and $\widehat{\boldsymbol{\sigma}}_n$ we propose the following (W)ALS algorithm.

1) *Initialization* Set the iteration counter $i = 1$ and initialize $\widehat{\boldsymbol{\sigma}}^{[0]}$ based on knowledge of $\boldsymbol{\sigma}_s$ and directional response of the sensors. For WALS, define the weight $\mathbf{W}_c = \widehat{\mathbf{R}}^{-(1/2)}$. Initialize $\widehat{\mathcal{L}}^{[0]}$ based on knowledge of the nominal position of the calibrator sources.

2) *Estimate* $\widehat{\mathbf{g}}^{[i]}$ by an eigenvalue decomposition of the matrix $\mathbf{M}$ with its entries obtained by averaging (32) [resp. (34)] over all $k \neq i, j$ or by an eigenvalue decomposition of the matrix $\mathbf{C}_2^\dagger\mathbf{C}_1$ as given by (39) using $\mathbf{A}$ and $\widehat{\boldsymbol{\sigma}}^{[i-1]}$ as prior knowledge. Note that neither gain calibration approach does require knowledge of the sensor noise powers $\boldsymbol{\sigma}_n$ and that the latter approach is advisable if some elements may have a very low gain.

3) *Estimate* $\widehat{\boldsymbol{\sigma}}_n^{[i]}$ using either (44) or (45) [resp. (47) or (48)] applying available knowledge of $\widehat{\mathbf{g}}^{[i]}$, $\widehat{\boldsymbol{\sigma}}^{[i-1]}$ and $\mathbf{A}$.

4) *Estimate* $\widehat{\boldsymbol{\sigma}}^{[i]}$ using (40) [resp. (42)] and knowledge of $\widehat{\mathbf{g}}^{[i]}$, $\widehat{\boldsymbol{\sigma}}_n^{[i]}$ and $\mathbf{A}$.

5) If the DOAs are inaccurately known (scenarios 3 and 4), *estimate* $\widehat{\mathcal{L}}^{[i]}$ using WSF as described in Section IV-E using knowledge of $\widehat{\mathbf{g}}^{[i]}$ and $\widehat{\boldsymbol{\sigma}}_n^{[i]}$ and initial estimate $\widehat{\mathcal{L}}^{[i-1]}$.

6) *Check for convergence or stop criterion* If $|\boldsymbol{\theta}^{[i-1]\dagger}\boldsymbol{\theta}^{[i]} - 1| < \delta$ or $i > i_{max}$, stop, otherwise increase i by 1 and continue with step 2.

The proposed criterion for convergence is based on a measure of the average relative error in all parameters, and will work even if the parameter values differ by orders of magnitude.

The extension with step 5 is referred to as the Extended ALS (xALS) resp. xWALS algorithm.

An algorithm that alternatingly optimizes for distinct groups of parameters can be proven to converge if the value of the cost function decreases in each iteration. In [27], [28] it is demonstrated that WSF minimizes the least squares cost function w.r.t. the parameterization of $\mathbf{A}$, thus providing a partial solution to the least squares problem considered here. In step 2, $\widehat{\mathbf{g}}$ is estimated using a method that only provides a near optimal solution to the least squares cost function. Its solution could, however, not be discerned from the true solution in Monte Carlo simulations, as demonstrated later in this paper. Optionally, step 2 could be augmented with one or two iterations of (29) to assure minimization of the least squares cost function. Alternatively, one could use the proposed estimate in the first iteration and use (29) in consecutive iterations in which a proper initial estimate is available from the previous iteration. The other parameters are estimated using well known standard solutions for least squares estimation problems. Therefore the value of the cost function is reduced in each step, thus ensuring convergence.

G. Computational Complexity

Table I summarizes the numerical complexity of different stages of the ALS and WALS algorithms per iteration expressed in the number of complex multiplications. N_{iter} is the number

Authorized licensed use limited to: Technische Universiteit Delft. Downloaded on August 20, 2009 at 05:56 from IEEE Xplore. Restrictions apply.

TABLE I
COMPLEXITY OF ALS AND WALS ITERATIONS

ALS	$\widehat{\mathbf{g}}$	$4\frac{1}{2}p^3 + pq^2 + o(\cdots)$
	$\widehat{\boldsymbol{\sigma}}$	$2p^2q + 12pq^2 + 9q^3 + o(\cdots)$
	$\widehat{\boldsymbol{\sigma}}_n$	0
	$\widehat{\sigma}_n^2$	0
	total	$4\frac{1}{2}p^3 + 2p^2q + 13pq^2 + 9q^3 + o(\cdots)$
WALS	$\widehat{\mathbf{g}}$	$5p^3 + pq^2 + o(\cdots)$
	$\widehat{\mathbf{g}}$ (Eq. (29))	$N_{iter}\left(2\frac{2}{3}p^3 + 4p^2q + 2q^2p + o(\cdots)\right)$
	$\widehat{\boldsymbol{\sigma}}$	$\frac{2}{3}p^3 + 2p^2q + pq^2 + \frac{2}{3}q^3 + o(\cdots)$
	$\widehat{\boldsymbol{\sigma}}_n$	$\frac{5}{3}p^3 + o(\cdots)$
	$\widehat{\sigma}_n^2$	$p^3 + o(\cdots)$
	total	$7\frac{1}{3}p^3 + 2p^2q + 2pq^2 + \frac{2}{3}q^3 + o(\cdots)$ $(\widehat{\boldsymbol{\sigma}}_n)$
		$6\frac{2}{3}p^3 + 2p^2q + 2pq^2 + \frac{2}{3}q^3 + o(\cdots)$ $(\widehat{\sigma}_n^2)$

of iterations required for convergence when using (29). In this table, the notation $o(\cdots)$ is used to denote all the neglected lower order terms. The complexity of the omnidirectional gain estimates is dominated by the eigenvalue decomposition on $\mathbf{M}$, which takes approximately $4p^3$ multiplications assuming use of the divide and conquer method [32]. This step dominates the overall complexity as well, so it may be worthwhile to estimate the number of iterations of the power method [33] required to obtain sufficient accuracy, especially if p is large.

The pseudoinverse in (40) can be implemented by a singular value decomposition which is computationally more efficient than direct computation of the Moore-Penrose inverse [33]. Some intermediate results, such as $\mathbf{G}\mathbf{A}\boldsymbol{\Sigma}\mathbf{A}^H\mathbf{G}^H$, are required more than once. In our calculation we assume that these terms are computed only once and stored for future use. Under this assumption the number of complex multiplications required to compute the noise power reduces to zero. The number of array elements p will generally be considerably larger than the number of source signals q. Therefore the q^3-, pq^2- and p^2q-terms may be considered negligible compared to the p^3-term in most cases. These results suggest that the better statistical performance of the WALS algorithm compared to the ALS algorithm as demonstrated by the simulation results presented in the next section comes with only a minor increase in complexity.

Both algorithm can be augmented with source position estimates using WSF. In [28] it is demonstrated that WSF can be implemented with $O(pq^2)$ complex operations per Gauss-Newton-type iteration of the proposed modified variable projection algorithm once the result from the eigenvalue decomposition is available. The cost of WSF is therefore dominated by the eigenvalue decomposition on $\widehat{\mathbf{R}}$ requiring about $4p^3$ complex multiplications. This makes the computational cost of source location estimation comparable to the cost of estimating the direction independent gains.

V. SIMULATION RESULTS

The methods proposed in the previous section were tested using Monte Carlo simulations and compared with the CRB. For these simulations a five-armed array was defined, each arm being an eight-element one wavelength spaced ULA. The first element of each arm formed an equally spaced circular array with half wavelength spacing between the elements. The source model used in the simulations is presented in Table II. This source model was created using random number generators to demonstrate that the proposed approach indeed works for arbitrary source models.

TABLE II
SOURCE POWERS AND SOURCE LOCATIONS USED IN THE SIMULATIONS

#	l	m	σ_q^2
1	0.24651	-0.71637	1.00000
2	-0.34346	0.76883	0.88051
3	-0.13125	-0.31463	0.79079
4	-0.29941	-0.52339	0.74654
5	0.39290	0.58902	0.69781

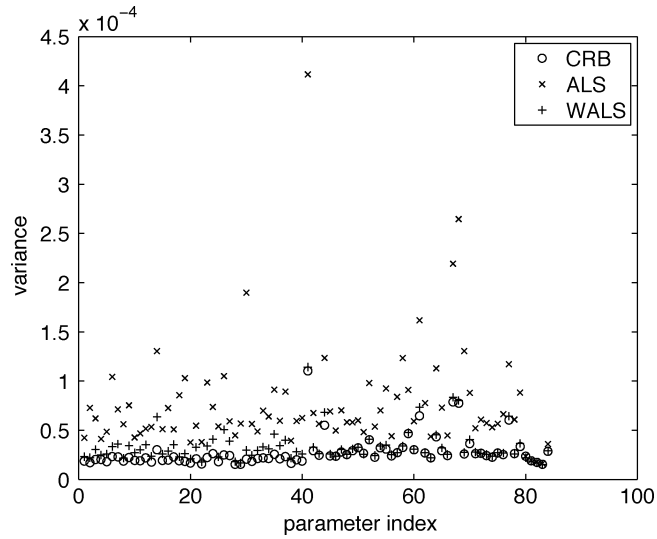

Fig. 2. Variance of the parameter vector $\boldsymbol{\theta} = [\boldsymbol{\gamma}^T, \phi_2, \cdots, \phi_{40}, \sigma_2^2, \cdots, \sigma_5^2, \sigma_n^2]^T$ estimates obtained in Monte Carlo simulations for the weighted and unweighted versions of the alternating least squares algorithm and compared these results with the CRB.

In the first simulation the direction independent gains, the source powers and the receiver noise powers were estimated assuming an array of equal elements, i.e., the parameter vector was defined as $\boldsymbol{\theta} = [\boldsymbol{\gamma}^T, \phi_2, \cdots, \phi_{40}, \sigma_2^2, \cdots \sigma_5^2, \sigma_n^2]^T$ corresponding with the first scenario mentioned in Section II. Data were generated assuming $\sigma_n^2 = 10$ and $N = 10^5$. This value for σ_n^2 implies an instantaneous SNR on the strongest source of only 0.1 per array element, a typical situation for radio astronomers. Fig. 2 shows the variance of the estimated parameters based on 1000 runs and compares these values with the CRB. As expected the WALS method appears to be asymptotically statistically efficient while the ALS approach does not attain the CRB and shows clear outliers.

Fig. 3 shows the variance found on a number of representative parameters as function of the number of samples and compares this with the corresponding CRBs. This plot confirms the conjecture raised in the previous paragraph that the WALS method is asymptotically statistically efficient. Fig. 4 shows the bias found in the Monte Carlo simulations and compares this with the statistical error of 1σ for a single realization based on the

Authorized licensed use limited to: Technische Universiteit Delft. Downloaded on August 20, 2009 at 05:56 from IEEE Xplore. Restrictions apply.

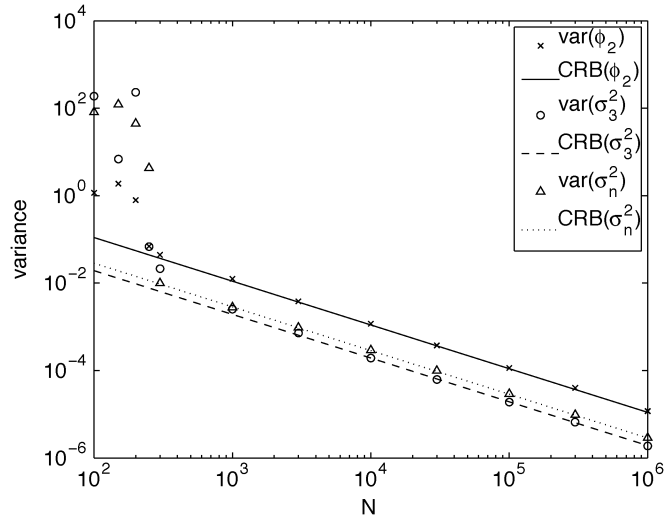

Fig. 3. Variances on ϕ_2. σ_3^2 and σ_n^2 as function of the number of samples N, compared to the corresponding CRBs.

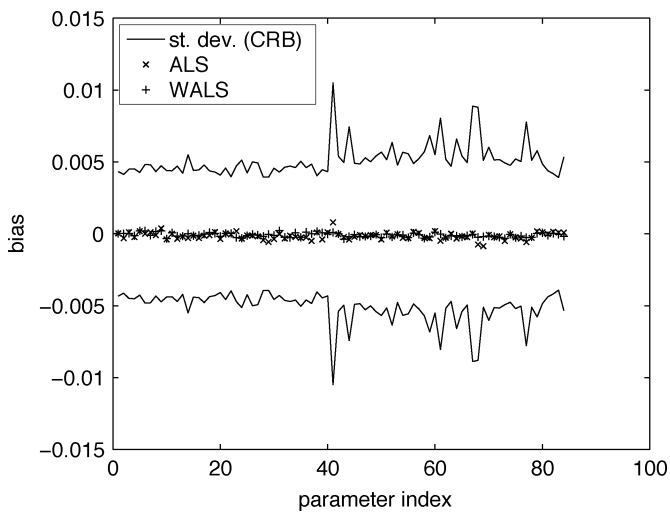

Fig. 4. Bias on the estimated parameters ($\boldsymbol{\theta} = [\boldsymbol{\gamma}^T, \phi_2, \cdots, \phi_{40}, \sigma_2^2, \cdots, \sigma_q^2, \sigma_n^2]^T$) for the weighted and unweighted version of the alternating least squares algorithm in the Monte Carlo simulations. The bias is compared with the standard deviation on the estimates derived from the CRB.

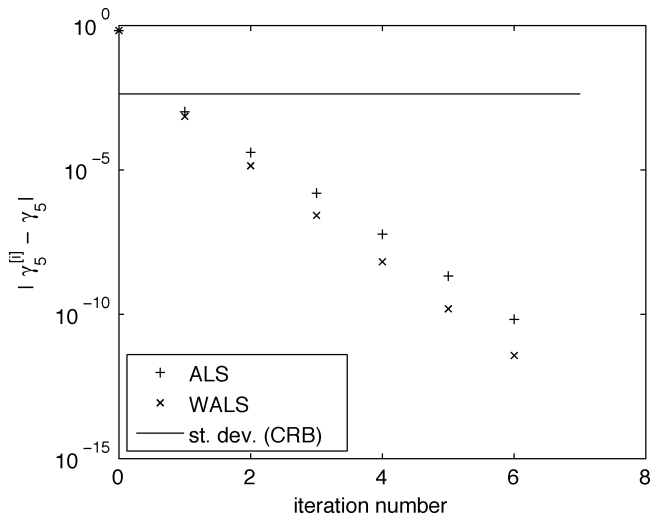

Fig. 5. The difference between the value of $\widehat{\gamma}_5$ at the end of each iteration and its final value is plotted versus the iteration number for ALS and WALS obtained in one of the Monte Carlo simulation runs. The behavior shown in this plot is representative for the results obtained for other parameters and from other runs. The standard deviation based on the CRB is also shown to demonstrate that one but preferably two iterations are sufficient to obtain a sufficiently accurate result.

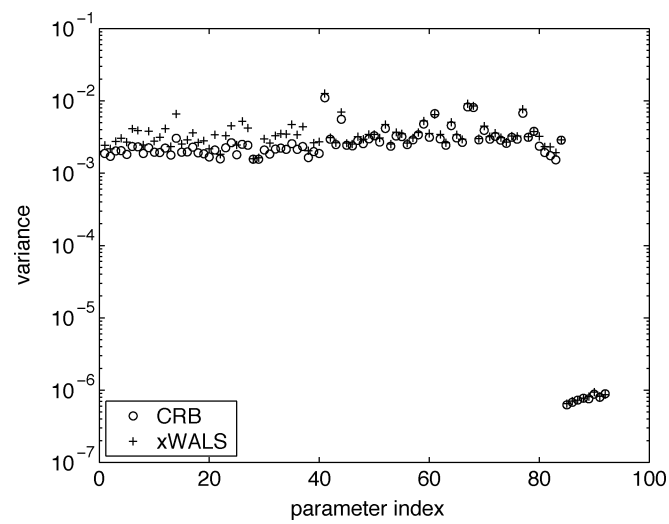

Fig. 6. Variance on the parameter vector $\boldsymbol{\theta} = [\boldsymbol{\gamma}^T, \phi_2, \cdots, \phi_{40}, \sigma_2^2, \cdots, \sigma_5^2, \sigma_n^2, l_2, \cdots, l_5, m_2 \cdots, m_5]^T$ estimates obtained in Monte Carlo simulations for the xWALS method, compared with the CRB.

CRB. This result indicates that both methods are unbiased for all parameters.

Fig. 5 shows the difference of the parameter value at the end of each iteration and its final value obtained from one run of the Monte Carlo simulation for a representative case, i.e. similar results were found for other parameters and other runs. The standard deviation based on the CRB is plotted as reference to facilitate the interpretation of the vertical scale. The results indicate that only one or two iterations are needed to reach the CRB in this scenario. With regard to other scenarios, this result suggests exponential convergence, i.e. each iteration adds about one significant digit to the parameter estimate, and it indicates that the WALS method converges more rapidly than the ALS method. In these simulation the stop criterion was $|\boldsymbol{\theta}^{[i-1]\dagger}\boldsymbol{\theta}^{[i]} - 1| < 10^{-10}$.

In the second simulation, the direction independent gains, the apparent source powers and source locations and the receiver noise powers were estimated assuming an array of ideal elements. This corresponds to the third scenario described in Section II. The parameter vector to be estimated is, thus, $\boldsymbol{\theta} = [\boldsymbol{\gamma}^T, \phi_2, \cdots, \phi_{40}, \sigma_2^2, \cdots, \sigma_5^2, \sigma_n^2, l_2, \cdots, l_5, m_2, \cdots, m_5]^T$. Data were generated assuming $\sigma_n^2 = 10$ and $N = 1000$. This again implies an instantaneous SNR of only 0.1 per array element but in this case the SNR per receiver is only 3.2 for the strongest source after integration. Fig. 6 shows the variance on the estimated parameters based on 1000 runs and compares these results with the CRB. In these simulations we only used the xWALS algorithm, since the previous simulations demonstrated that the ALS method gives inferior results compared to the WALS approach. These results show that the variance on the parameter estimates are close to the CRB.

Authorized licensed use limited to: Technische Universiteit Delft. Downloaded on August 20, 2009 at 05:56 from IEEE Xplore. Restrictions apply.

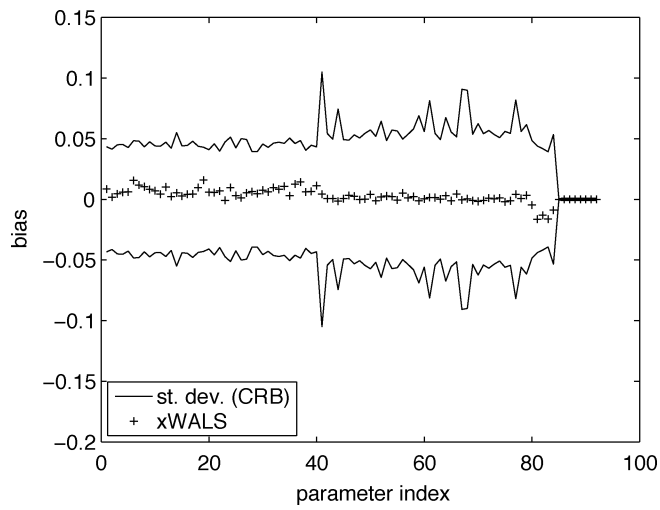

Fig. 7. Bias on the estimated parameters ($\boldsymbol{\theta} = [\boldsymbol{\gamma}^T, \phi_2, \cdots, \phi_{40}, \sigma_2^2, \cdots, \sigma_5^2, \sigma_n^2, l_2, \cdots, l_5, m_2 \cdots, m_5]^T$) for the xWALS algorithm obtained in Monte Carlo simulations. The results are compared with the standard deviation on the estimated derived from the CRB.

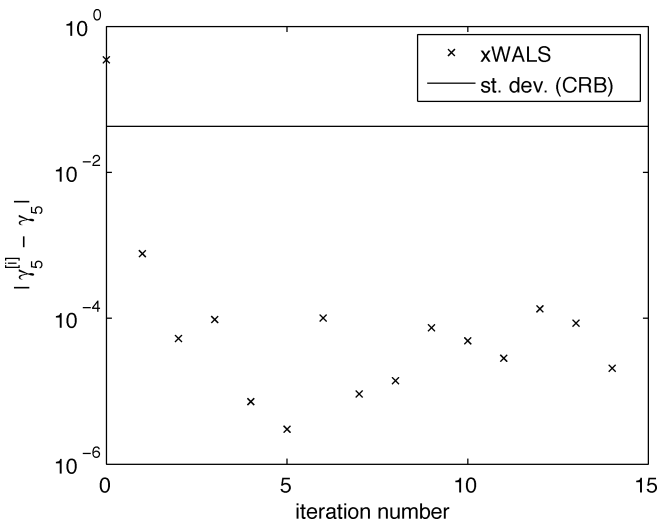

Fig. 8. The difference between the value of $\widehat{\gamma}_5$ at the end of each iteration and its final value is plotted versus the iteration number for one of the Monte Carlo simulation runs. The behavior shown in this plot is representative for the results obtained for other parameters and from other runs. The standard deviation based on the CRB is also shown to demonstrate that one but preferably two iterations are sufficient to obtain a sufficiently accurate result.

Fig. 7 shows the bias on the estimated parameters found in these Monte Carlo simulations. These results indicate that the estimates are unbiased.

The convergence of a representative parameter is shown in Fig. 8. Again one but preferably two iterations are sufficient to get accurate results. However, in this scenario the stop criterion $|\boldsymbol{\theta}^{[i-1]\dagger}\boldsymbol{\theta}^{[i]} - 1| < 10^{-10}$ was not reached and the algorithm stops because the maximum number of iterations, which was set to 15, was reached. This indicates that the algorithm tries to interpret the noise as real signal. In the first iteration first the omnidirectional gains, apparent source powers and noise power are estimated. These values are then used in the WSF to find the source locations. The source locations are then used to update the sky model, leading to an update of $\widehat{\mathbf{g}}$, $\widehat{\boldsymbol{\sigma}}$, and $\widehat{\boldsymbol{\sigma}}_n$ and the cycle is complete. In the first simulation, the SNR after integration defined as $\sqrt{N}\sigma_i^2/\sigma_n^2$ was a factor 10 better than in this simulation in which the SNR per array element per source after integration is only 2.2 to 3.2.

VI. Summary and Conclusion

In this paper, we have developed a weighted alternating least squares (WALS) algorithm to solve for direction independent complex gains, apparent source powers and receiver noise powers simultaneously in a snapshot observation by a sensor array. Our solution for finding the direction independent gains extends and improves the available methods in the literature in the case of multiple calibrator sources and provides robustness to small entries in the array covariance matrix and small gain values (due to, e.g., failing elements). Although we have assumed independent sensor noise powers, our gain estimation method is easily applied to cases in which some pairs of elements experience correlated noise. We also found that unbiased estimation of the receiver noise powers requires weighting with the best available array covariance matrix model instead of weighting with the measured array covariance matrix, which is common practice in signal processing.

The statistical performance was compared with an unweighted alternating least squares (ALS) algorithm and the CRB and found to provide a statistically efficient estimate thereby outperforming the ALS method. The computational complexity of ALS and WALS scale with $4(1/2)p^3$ and about $7p^3$, respectively, assuming that the number of calibrator sources is much smaller than the number of array elements, i.e., the WALS method comes with only a minor increase in computational burden. Simulations indicate that only one or two iterations are sufficient to reach the CRB and that every iteration adds about one significant digit, the WALS method converging slightly faster than the ALS method.

We extended these methods with source location estimation using weighted subspace fitting. Simulations indicate that this extension to the WALS algorithm provides a statistically efficient simultaneous estimate of omnidirectional gains, apparent source powers, source locations and receiver noise powers. Again, one to two iterations proved to be sufficient to reach the CRB. However, the convergence of the parameter is blocked at some level lower than the CRB by the noise in the measurement, at which point the algorithm keeps jumping from one local optimum to another.

Acknowledgment

This work has been carried out in close collaboration with the LOFAR-team. This helped to develop our ideas up to the point at which our method could be implemented to calibrate the LOFAR stations. Especially the interaction with A.-J. Boonstra and J. Bregman at ASTRON on LOFAR station calibration is gratefully acknowledged. The authors would also like to thank the reviewers whose comments helped to improve the original manuscript.

References

[1] J. D. Bregman, "LOFAR approaching the critical design review," in *Proc. XXVIIIth General Assembly of the Int. Union of Radio Science (URSI GA)*, New Delhi, India, Oct. 23–29, 2005.

Authorized licensed use limited to: Technische Universiteit Delft. Downloaded on August 20, 2009 at 05:56 from IEEE Xplore. Restrictions apply.

[2] C. J. Lonsdale, "The Murchison widefield array," in *Proc. XXIXth General Assembly of the Int. Union of Radio Sci. (URSI GA)*, Chicago, IL, Oct. 7–16, 2008.

[3] J. B. Peterson, U. L. Pen, and X. P. Wu, "The primeval structure telescope: Goals and status," in *Proc. XXVIIIth General Assembly of the Int. Union of Radio Sci. (URSI GA)*, New Delhi, India, Oct. 23–29, 2005.

[4] C. Lonsdale, Calibration Approaches MIT Haystack, Tech. Rep. LFD memo 015, Dec. 8, 2004.

[5] T. Cornwell and E. B. Fomalont, "Self-calibration," in *Synthesis Imaging in Radio Astronomy*, R. A. Perley, F. R. Schwab, and A. H. Bridle, Eds. San Francisco, CA: BookCrafters, 1994, vol. 6, Astronom. Soc. Pacific Conf. Ser.

[6] A. J. Boonstra and A. J. van der Veen, "Gain calibration methods for radio telescope arrays," *IEEE Trans. Signal Process.*, vol. 51, no. 1, pp. 25–38, Jan. 2003.

[7] S. J. Wijnholds and A. J. Boonstra, "A multisource calibration method for phased array radio telescopes," in *Proc. 4th IEEE Workshop on Sens. Array and Multi-Channel Process. (SAM)*, Waltham (MA), Jul. 12–14, 2006.

[8] S. van der Tol, B. D. Jeffs, and A. J. van der Veen, "Self calibration for the LOFAR radio astronomical array," *IEEE Trans. Signal Process.*, vol. 55, no. 9, pp. 4497–4510, Sep. 2007.

[9] B. P. Flanagan and K. L. Bell, "Array self-calibration with large sensor position errors," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Proces. (ICASSP)*, 1999.

[10] D. Astély, A. L. Swindlehurst, and B. Ottersten, "Spatial signature estimation for uniform linear arrays with unknown receiver gains and phases," *IEEE Trans. Signal Process.*, vol. 47, no. 8, pp. 2128–2138, Aug. 1999.

[11] M. Pesavento, A. B. Gershman, and K. M. Wong, "Direction finding in partly calibrated sensor arrays composed of multiple subarrays," *IEEE Trans. Signal Process.*, vol. 50, no. 9, pp. 2103–2115, Sep. 2002.

[12] C. M. S. See and A. B. Gershman, "Direction of arrival estimation in partly calibrated subarray-based sensor arrays," *IEEE Trans. Signal Process.*, vol. 52, no. 2, pp. 329–338, Feb. 2004.

[13] A. J. Weiss and B. Friedlander, "Array shape calibration using sources in unknown locations - a maximum likelihood approach," *IEEE Trans. Acoustics, Speech and Signal Processing*, vol. 37, no. 12, Dec. 1989.

[14] C. M. S. See, "Method for array calibration in high-resolution sensor array processing," *IEE Proc. Radar, Sonar and Navig.*, vol. 142, no. 3, Jun. 1995.

[15] J. Pierre and M. Kaveh, "Experimental performance of calibration and direction finding algorithms," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, May 1991, vol. 2, pp. 1365–1368.

[16] B. C. Ng and C. M. S. See, "Sensor array calibration using a maximum likelihood approach," *IEEE Trans. Antennas Propag.*, vol. 44, no. 6, Jun. 1996.

[17] H. Li, P. Stoica, and J. Li, "Computationally efficient maximum likelihood estimation of structured covariance matrices," *IEEE Trans. Signal Process.*, vol. 47, no. 5, pp. 1314–1323, May 1999.

[18] A. J. Weiss and B. Friedlander, "'Almost Blind' signal estimation using second-order moments," *IEE Proc. Radar, Sonar and Navig.*, vol. 142, no. 5, Oct. 1995.

[19] M. Zatman, "How narrow is narrowband," *IEE Proc. Radar, Sonar and Navig.*, vol. 145, no. 2, pp. 85–91, Apr. 1998.

[20] D. R. Fuhrmann, "Estimation of sensor gain and phase," *IEEE Trans. Signal Process.*, vol. 42, no. 1, pp. 77–87, Jan. 1994.

[21] S. J. Wijnholds and A. J. van der Veen, "Effects of parametric constraints on the CRLB in gain and phase estimation problems," *IEEE Signal Process. Lett.*, vol. 13, no. 10, pp. 620–623, Oct. 2006.

[22] S. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Englewood Cliffs, NJ: Prentice-Hall, 1993, vol. 1.

[23] B. Ottersten, P. Stoica, and R. Roy, "Covariance matching estimation techniques for array signal processing applications," *Digital Signal Process., A Rev. J.*, vol. 8, pp. 185–210, Jul. 1998.

[24] T. J. Pearson and A. C. S. Readhead, "Image formation by self- calibration in radio astronomy," *Ann. Rev. Astron. Astrophys.*, vol. 22, pp. 97–130, 1984.

[25] S. van der Tol and S. J. Wijnholds, "CRB analysis of the impact of unknown receiver noise on phased array calibration," in *Proc. 4th IEEE Workshop on Sens. Array and Multi-channel Process. (SAM)*, Waltham, MA, Jul. 12–14, 2006.

[26] R. O. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE Trams. Antennas Propag.*, vol. AP-34, no. 3, Mar. 1986.

[27] M. Viberg and B. Ottersten, "Sensor array processing based on subspace fitting," *IEEE Trans. Signal Process.*, vol. 39, no. 5, pp. 1110–1121, May 1991.

[28] M. Viberg, B. Ottersten, and T. Kailath, "Detection and estimation in sensor arrays using weighted subspace fitting," *IEEE Trans. Signal Process.*, vol. 39, no. 11, pp. 2436–2448, Nov. 1991.

[29] A. Quinlan, J.-P. Barbot, and P. Larzabal, "Automatic determination of the number of targets present when using the time reversal operator," *J. Acoust. Soc. Amer.*, vol. 119, no. 4, pp. 2220–2225, Apr. 2006.

[30] A. Quinlan, J. P. Barbot, P. Larzabal, and M. Haardt, "Model order selection for short data: An exponential fitting test (EFT)," *EURASIP J. Adv. Signal Process.*, 2007.

[31] M. Wax and T. Kailath, "Detection of signals by information theoretic criteria," *IEEE Trans. Acoust., Speech Signal Process.*, vol. ASSP-33, no. 2, pp. 387–392, Apr. 1985.

[32] Divide-and-Conquer Eigenvalue Algorithm [Online]. Available: http://en.wikipedia.org/wiki/Divide-and-conquer eigenvalue algorithm [Online]. Available:

[33] T. K. Moon and W. C. Stirling, *Mathematical Methods and Algorithms for Signal Processing*. Englewood Cliffs, NJ: Prentice-Hall, 2000.

Stefan J. Wijnholds (S'06) was born in The Netherlands in 1978. He received the M.Sc. degrees in astronomy and applied physics (both *cum laude*) from the University of Groningen, Groningen, The Netherlands, in 2003.

After his graduation he joined the R&D Department, ASTRON, the Netherlands Institute for Radio Astronomy, Dwingeloo, where he works with the system design and integration group on the development of the next generation of radio telescopes. Since 2006, he has also been with the Delft University of Technology, Delft, The Netherlands, where he is pursuing the Ph.D. degree. His research interests lie in the area of array signal processing, specifically calibration and imaging.

Alle-Jan van der Veen (F'05) was born in The Netherlands in 1966. He received the Ph.D. degree (*cum laude*) from Delft University of Technology (TU Delft), Delft, The Netherlands, in 1993.

Throughout 1994, he was a Postdoctoral scholar at Stanford University, Stanford, CA. At present, he is a Full Professor in Signal Processing with TU Delft. His research interests are in the general area of system theory applied to signal processing, and in particular, algebraic methods for array signal processing, with applications to wireless communications and radio astronomy.

Dr. van der Veen is the recipient of a 1994 and 1997 IEEE Signal Processing Society (SPS) Young Author paper award, and was an Associate Editor for the IEEE TRANSACTIONS ON SIGNAL PROCESSING (1998–2001), Chairman of IEEE SPS Signal Processing for Communications Technical Committee (2002–2004), Editor-in-Chief of the IEEE SIGNAL PROCESSING LETTERS (2002–2005), and Editor-in-Chief of the IEEE TRANSACTIONS ON SIGNAL PROCESSING. He is currently Member-At-Large of the Board of Governors of IEEE SPS.

Authorized licensed use limited to: Technische Universiteit Delft. Downloaded on August 20, 2009 at 05:56 from IEEE Xplore. Restrictions apply.