On the Constrained Stochastic Gradient Algorithm: Model, Performance, and Improved Version

Javier Ernesto Kolodziej, Orlando José Tobias, *Member, IEEE*, Rui Seara, *Senior Member, IEEE*, and Dennis R. Morgan, *Life Senior Member, IEEE*

***Abstract*—This paper discusses the constrained stochastic gradient (CSG) algorithm used for controlling antenna arrays, aiming to maximize the signal-to-interference-plus-noise ratio (SINR) in mobile communications. Firstly, analytical expressions for the first moment of the weight vector and the SINR characteristic of the standard CSG algorithm are derived for two interferer signals, considering small step-size conditions and assuming Gaussian signal, interference, and noise. From these model expressions, the CSG algorithm performance is assessed, which predicts undesired behavior (termed here unbalanced behavior, pertaining to an unbalance between maximizing signal power and minimizing interference power) when one or more interference angles-of-arrival are close to the signal angle-of-arrival and the angle-of-arrival spreads of the involved signals are small. Finally, by using the model expressions, an improved CSG (ICSG) algorithm is proposed to compensate the unbalanced behavior of the standard CSG algorithm. The accuracy of the proposed model and the effectiveness of the modified algorithm are assessed through numerical simulations.**

***Index Terms*—Adaptive antenna arrays, algorithm modeling, constrained stochastic gradient algorithm, mobile communication.**

I. Introduction

OVER the last few years, wireless communication networks have become more and more popular. Thus, the increasing number of users is causing serious problems in densely urbanized areas since the frequency spectrum is approaching the limit of its capacity [1]. The limited spectrum problem can be overcome by using sectorized cells and increasing frequency reuse, along with a reduction of cell size [2]. For instance, with a layout of seven cells per cluster [see Fig. 1(a)], having three sectors of 120° per cell (delimited by dashed lines in Fig. 1), the number of carrier frequencies available in each sector is 1/21 of the total number of available frequencies. On the other hand, in a layout of three cells per cluster [see Fig. 1(b)], the number of carrier frequencies available in each sector is 1/9 of the total, thus increasing the system capacity (in terms of carrier frequencies per sector) by a factor of approximately 7/3. However, distance d between co-channel cells then decreases by a factor of $\sqrt{21/9}$ (see Fig. 1), thus increasing the level of co-channel interfering signals in the system [3]. A strategy to reduce such interference is the use of antenna arrays at the base stations, so that the array radiation pattern is directed to the position of the mobile terminals inside the proper cell and canceling the co-channel interference to other cells [1].

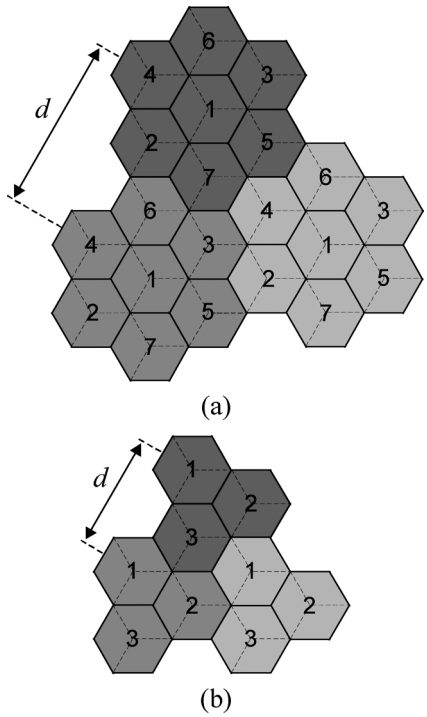

Fig. 1. Cell plan using (a) 7-cell and (b) 3-cell clusters with three 120° sectors.

Ideally, each cell must have knowledge of the position of all mobile terminals in its area, thus determining a global optimum solution. Unfortunately, such an approach is not practical. Currently, all information about the mobile terminals is obtained through the uplink channel. With this information available, an approach in which the neighboring cells reduce the transmitted signal power in the direction of the co-channel mobile terminals can be used. In this way, the interference level is globally decreased. By considering that the signal propagation of the uplink and downlink channels is similar, it can be assumed that the covariance matrices of the channels are similar [1]. For instance, in [4]–[6] such an assumption is used for controlling the adaptive array by using estimates of the desired and interference

Manuscript received March 06, 2008; revised October 22, 2008. First published December 02, 2008; current version published March 11, 2009. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Jonathon Chambers. This work was supported in part by the Brazilian National Council for Scientific and Technological Development (CNPq) and the Committee for Postgraduate Courses in Higher Education (CAPES).

J. E. Kolodziej, O. J. Tobias, and R. Seara are with the LINSE-Circuits and Signal Processing Laboratory, Department of Electrical Engineering, Federal University of Santa Catarina, SC 88040-900, Brazil (e-mail: javier@linse.ufsc.br; orlando@linse.ufsc.br; seara@linse.ufsc.br).

D. R. Morgan is with the Bell Laboratories, Alcatel-Lucent, Murray Hill, NJ 07974-0636 USA (e-mail: drrm@bell-labs.com).

Digital Object Identifier 10.1109/TSP.2008.2010375

downlink covariance matrices obtained from uplink measurements. Such a procedure is used in [1] to define a local objective function, thereby obtaining an approximate global optimum solution in a cooperative network, without the need of having communication between cells. From that objective function, the constrained stochastic gradient (CSG) algorithm is derived. The CSG algorithm is interesting because of its low computational complexity and very good convergence properties in comparison with other algorithms [1].

The performance of an adaptive array can be assessed from the signal-to-interference-plus-noise ratio (SINR) function [1]. To obtain adequate knowledge of such a function, it is necessary to consider a wide range of operating conditions. Specifically, several angle-of-arrival (AoA) combinations for the in-cell and co-channel interfering uplink signals must be considered. In this way, the performance assessment of the algorithm requires a large number of simulations to determine the mean behavior for each combination. In this context, the possibility of having an algorithm model becomes important since much simulation time can be saved. Such a model allows an exhaustive study of the algorithm performance in different working conditions, also providing useful insights into the algorithm behavior. Thus, this research work aims to develop model expressions for the mean weight behavior of the adaptive filter as well as for the SINR characteristic. The simplifying assumption considered to derive the analytical model is a small step-size condition. The presented expressions only consider the case of two interferers, although the approach could be generalized to three or more interferers.

Through an analysis of the obtained model expressions, the existence of unbalanced behavior in the standard CSG algorithm is demonstrated in this paper. Such behavior occurs under the following condition: one or more interference AoAs are close to the signal AoA, and the AoA spreads of the involved signals are small. In this situation, the CSG algorithm stresses the co-channel interference minimization more than the maximization of the power radiated to the in-cell mobile terminal. The proposed mean weight behavior expression allows us to address two central points:

i) to explain the unbalanced behavior of the standard CSG algorithm;
ii) to suggest a modification to the standard CSG algorithm to compensate the unbalanced behavior.

The modified algorithm is termed improved CSG (ICSG) algorithm. Numerical simulations are used to assess the accuracy of the proposed model and the effectiveness of the modified algorithm.

This paper is organized as follows. Section II presents the standard CSG algorithm characteristics and its updating equations. In Section III, model expressions describing the mean weight and SINR behaviors are derived. Section IV discusses the unbalanced behavior of the standard CSG algorithm through the model expressions, and proposes a modification to the CSG algorithm. Simulation results are presented in Section V. Finally, Section VI presents concluding remarks.

II. Standard CSG Algorithm

In adaptive antenna array systems, the adaptive algorithm aims to maximize the irradiated power to the mobile terminal (MT) inside the cell P_{IC} and to minimize that power to the co-channel MT in other cells P_{CC}. Such powers are expressed, respectively, as

$$P_{\mathrm{IC}} = \mathbf{w}^{\mathrm{H}}\mathbf{R}_{\mathrm{IC}}\mathbf{w} \tag{1}$$

and

$$P_{\mathrm{CC}} = \mathbf{w}^{\mathrm{H}}\mathbf{R}_{\mathrm{CC}}\mathbf{w} \tag{2}$$

where $\mathbf{w} = [w_1\ w_2\ \cdots\ w_M]^{\mathrm{T}}$ denotes the complex weight vector, and $\mathbf{R}_{\mathrm{IC}}$ and $\mathbf{R}_{\mathrm{CC}}$ are, respectively, the in-cell and co-channel interference downlink covariance matrices. Then, the adaptive algorithm must maximize the following expression over $\mathbf{w}$ [1]:

$$\xi = \frac{\mathbf{w}^{\mathrm{H}}\mathbf{R}_{\mathrm{IC}}\mathbf{w}}{\mathbf{w}^{\mathrm{H}}(\mathbf{R}_{\mathrm{CC}}+\mathbf{I})\mathbf{w}} \tag{3}$$

where $\mathbf{I}$ is the identity matrix, introduced to account for normalized additive noise. Note that the complex weight vector $\mathbf{w}$ in (3) is the independent variable of that deterministic cost function. In [1], the CSG algorithm is derived by means of an intuitive interpretation of (3), by considering $\mathbf{w}(n)$ instead of $\mathbf{w}$, where $\mathbf{w}(n)$ is now a time-varying variable arising from the adaptive process discussed in the following. Thereby, to maximize (3), the steepest descent algorithm is used for the co-channel interference power along with the steepest ascent algorithm for the in-cell signal power. Since two optimization processes are simultaneously running, such a strategy may result in slow algorithm convergence. A heuristic solution to circumvent the convergence problem is to use a two-step updating scheme. Firstly, the numerator of (3) is adapted while the denominator is fixed; next, the denominator is adapted, keeping the numerator fixed. Mathematically, such a scheme can be implemented by incorporating a projection matrix into the updating expression. As such, the CSG algorithm uses both the in-cell uplink signal $s(n)$ and the co-channel uplink signal $u(n)$ to estimate the downlink covariance matrices $\mathbf{R}_{\mathrm{IC}}$ and $\mathbf{R}_{\mathrm{CC}}$, respectively. For the case of two interferers, the updating equations corresponding to the aforementioned two-step optimization processes are given by [1]

$$\mathbf{v}_1(n) = \mathbf{w}(n) + \mu_{\mathrm{s}}\left[\mathbf{I} - \frac{\hat{\mathbf{R}}_{\mathrm{u}_1}(n)}{\|\mathbf{u}_1(n)\|^2} - \frac{\hat{\mathbf{R}}_{\mathrm{u}_2}(n)}{\|\mathbf{u}_2(n)\|^2}\right]\hat{\mathbf{R}}_{\mathrm{s}}(n)\mathbf{w}(n) \tag{4}$$

and

$$\mathbf{v}_2(n) = \mathbf{v}_1(n) - \mu_{\mathrm{u}}\left[\mathbf{I} - \frac{\hat{\mathbf{R}}_{\mathrm{s}}(n)}{\|\mathbf{s}(n)\|^2}\right]\left[\hat{\mathbf{R}}_{\mathrm{u}_1}(n) + \hat{\mathbf{R}}_{\mathrm{u}_2}(n)\right]\mathbf{w}(n) \tag{5}$$

with

$$\mathbf{w}(n+1) = \frac{\mathbf{v}_2(n)}{\|\mathbf{v}_2(n)\|} \tag{6}$$

where $\hat{\mathbf{R}}_{\mathrm{s}}(n) = \mathbf{s}(n)\mathbf{s}^{\mathrm{H}}(n)$ and $\hat{\mathbf{R}}_{\mathrm{u}_k}(n) = \mathbf{u}_k(n)\mathbf{u}_k^{\mathrm{H}}(n)$, for $k = 1, 2$, are the instantaneous estimates of $\mathbf{R}_{\mathrm{IC}}$ and $\mathbf{R}_{\mathrm{CC}}$, respectively. Vectors $\mathbf{s}(n)$ and $\mathbf{u}_k(n)$ are, respectively, the in-cell uplink and the kth co-channel interference vectors. Parameters μ_{s} and μ_{u} are the corresponding algorithm step sizes. The weight vector is normalized at each iteration, maintaining the term $\mathbf{w}^{\mathrm{H}}(n)\mathbf{I}\mathbf{w}(n)$ constant in the denominator of (3).

III. Stochastic Model of the Standard CSG Algorithm

This section develops a stochastic model for the weight vector behavior as well as an expression for the SINR, which depends on that vector. For such, we assume that the elements of the input vectors $\mathbf{s}(n)$ and $\mathbf{u}_k(n)$, for $k = 1, 2$, are samples of a stochastic process. As a consequence, the same assumption also applies to the weight vector.

A. Mean Weight Behavior

The first step to obtain a model describing the mean weight behavior of the adaptive vector is to determine the expected value of both sides of (6). Thus,

$$\mathrm{E}\left[\mathbf{w}(n+1)\right] = \mathrm{E}\left[\frac{\mathbf{v}_2(n)}{\|\mathbf{v}_2(n)\|}\right]. \tag{7}$$

Obtaining such an expected value is not a trivial task. However, by considering that the evolution of $\mathbf{v}_2(n)$ is smooth (small dispersion condition, i.e., when the standard deviation is much smaller than the mean), the following approximation can be used [7]:

$$\mathrm{E}\left[\mathbf{w}(n+1)\right] \approx \frac{\mathrm{E}\left[\mathbf{v}_2(n)\right]}{\|\mathrm{E}\left[\mathbf{v}_2(n)\right]\|}. \tag{8}$$

From the modeling point of view, a small dispersion in $\mathbf{v}_2(n)$ is ensured under a small-step-size condition, whereby (8) becomes valid. The weight vector and the received signal vectors are assumed independent, which is justifiable from the small-step-size condition [8]. Then, by taking the expected value of both sides of (4) and (5), one obtains

$$\mathrm{E}\left[\mathbf{v}_1(n)\right] = \mathrm{E}\left[\mathbf{w}(n)\right] + \mu_{\mathrm{s}}\left(\mathbf{I} - \mathbf{R}_{\mathrm{u}_1}^{\mathrm{N}} - \mathbf{R}_{\mathrm{u}_2}^{\mathrm{N}}\right)\mathbf{R}_{\mathrm{s}}\mathrm{E}\left[\mathbf{w}(n)\right] \tag{9}$$

and

$$\mathrm{E}\left[\mathbf{v}_2(n)\right] = \mathrm{E}\left[\mathbf{v}_1(n)\right] - \mu_{\mathrm{u}}\left(\mathbf{I} - \mathbf{R}_{\mathrm{s}}^{\mathrm{N}}\right)\left(\mathbf{R}_{\mathrm{u}_1} + \mathbf{R}_{\mathrm{u}_2}\right)\mathrm{E}\left[\mathbf{w}(n)\right] \tag{10}$$

where

$$\mathbf{R}_{\mathrm{s}}^{\mathrm{N}} = \mathrm{E}\left[\frac{\mathbf{s}(n)\mathbf{s}^{\mathrm{H}}(n)}{\mathbf{s}^{\mathrm{H}}(n)\mathbf{s}(n)}\right] = \mathrm{E}\left[\frac{\hat{\mathbf{R}}_{\mathrm{s}}(n)}{\mathbf{s}^{\mathrm{H}}(n)\mathbf{s}(n)}\right] \tag{11}$$

and

$$\mathbf{R}_{\mathrm{u}_k}^{\mathrm{N}} = \mathrm{E}\left[\frac{\mathbf{u}_k(n)\mathbf{u}_k^{\mathrm{H}}(n)}{\mathbf{u}_k^{\mathrm{H}}(n)\mathbf{u}_k(n)}\right] = \mathrm{E}\left[\frac{\hat{\mathbf{R}}_{\mathrm{u}_k}(n)}{\mathbf{u}_k^{\mathrm{H}}(n)\mathbf{u}_k(n)}\right] \tag{12}$$

are the normalized sample covariance matrices, and $\mathbf{R}_{\mathrm{s}} = \mathrm{E}[\hat{\mathbf{R}}_{\mathrm{s}}(n)]$ and $\mathbf{R}_{\mathrm{u}_k} = \mathrm{E}[\hat{\mathbf{R}}_{\mathrm{u}_k}(n)]$. The calculation of matrices (11) and (12) is shown in Appendix I. By substituting (9) into (10) and the resulting expression into (8), the expression for mean weight behavior is obtained as

$$\mathrm{E}\left[\mathbf{w}(n+1)\right] = \frac{\mathbf{A}\mathrm{E}\left[\mathbf{w}(n)\right]}{\sqrt{\mathrm{E}\left[\mathbf{w}^{\mathrm{H}}(n)\right]\mathbf{A}^{\mathrm{H}}\mathbf{A}\mathrm{E}\left[\mathbf{w}(n)\right]}} \tag{13}$$

where

$$\mathbf{A} = \mathbf{I} + \mu_{\mathrm{s}}\left(\mathbf{I} - \mathbf{R}_{\mathrm{u}_1}^{\mathrm{N}} - \mathbf{R}_{\mathrm{u}_2}^{\mathrm{N}}\right)\mathbf{R}_{\mathrm{s}} - \mu_{\mathrm{u}}\left(\mathbf{I} - \mathbf{R}_{\mathrm{s}}^{\mathrm{N}}\right)\left(\mathbf{R}_{\mathrm{u}_1} + \mathbf{R}_{\mathrm{u}_2}\right). \tag{14}$$

Expression (13) can also be interpreted as the well-known power method, used to obtain the eigenvector corresponding to the dominant eigenvalue of a diagonizable matrix. Therefore, the steady-state value for the mean-weight vector is the eigenvector associated with the dominant eigenvalue of $\mathbf{A}$ [9].

At this point, it is important to observe that the analysis of the weight error vector, defined as the difference between the current weight vector and optimum weight, is not useful, since in array systems the optimum weight vector is neither intrinsic to the system nor particularly relevant, unlike the usual adaptive filter problem performing a system identification of an explicit "true" impulse response. For more details, such background is well covered in [1].

B. SINR Expression

The basic objective of the adaptive array is to maximize the downlink SINR. Therefore, the SINR behavior is an important metric governing the array performance. By definition, the SINR is given by

$$\gamma = \frac{\mathrm{E}[P_{\mathrm{IC}}]}{\mathrm{E}[P_{\mathrm{CC}}] + 1} \tag{15}$$

where a normalized noise power of value 1 is assumed in the denominator. Since the weight vector and the downlink signals are independent, the evolution of the SINR can be written as

$$\gamma(n) = \frac{\mathrm{E}\left[\mathbf{w}^{\mathrm{H}}(n)\mathbf{R}_{\mathrm{IC}}\mathbf{w}(n)\right]}{\mathrm{E}\left[\mathbf{w}^{\mathrm{H}}(n)\mathbf{R}_{\mathrm{CC}}\mathbf{w}(n)\right] + 1}. \tag{16}$$

Alternatively, (16) can also be expressed as

$$\gamma(n) = \frac{\mathrm{tr}\left[\mathbf{R}_{\mathrm{IC}}\mathbf{K}(n)\right]}{\mathrm{tr}\left[\mathbf{R}_{\mathrm{CC}}\mathbf{K}(n)\right] + 1} \tag{17}$$

where $\mathbf{K}(n) = \mathrm{E}[\mathbf{w}(n)\mathbf{w}^{\mathrm{H}}(n)]$ is the second moment of the adaptive weight vector.

Note from (17) that the second moment must be known. Due to the small dispersion of the CSG algorithm adaptive weight evolution, the following approximation is considered for the sake of mathematical simplicity [10]:

$$\mathbf{K}(n) = \mathrm{E}\left[\mathbf{w}(n)\mathbf{w}^{\mathrm{H}}(n)\right] \approx \mathrm{E}\left[\mathbf{w}(n)\right]\mathrm{E}\left[\mathbf{w}^{\mathrm{H}}(n)\right]. \tag{18}$$

Thus, (16) can now be computed from the mean weight expression using

$$\gamma(n) \approx \frac{\mathrm{E}\left[\mathbf{w}^{\mathrm{H}}(n)\right]\mathbf{R}_{\mathrm{IC}}\mathrm{E}\left[\mathbf{w}(n)\right]}{\mathrm{E}\left[\mathbf{w}^{\mathrm{H}}(n)\right]\mathbf{R}_{\mathrm{CC}}\mathrm{E}\left[\mathbf{w}(n)\right] + 1}. \tag{19}$$

IV. Insights From the Model Expressions

A. Unbalanced Behavior of the Standard CSG Algorithm

The algorithm controlling the adaptive antenna array must adjust the weights to fulfill two goals taking into account the irradiated power: the power must be large for the intended in-cell MT (P_{IC}) and small for the co-channel MT in neighboring cells (P_{CC}). A crucial point in this process is to adequately balance these two objectives. From the derived model (19), the standard CSG algorithm behavior can be assessed for different working conditions. Undesired behavior is demonstrated here for the combination of two conditions: one or more co-channel interference and in-cell signals arrive from almost the same direction, and the respective AoA spread is small. In that case, the standard CSG algorithm leads to minimizing P_{CC} more than maximizing P_{IC}. With the aid of the proposed model, it is possible to understand and propose a solution for this undesired behavior, which is discussed in the following. Although the SINR (19) is the ultimate arbiter, here we shall focus our discussion on the mean weight update (9), which is responsible for maximizing P_{IC}.

First, let us consider the case in which the in-cell signal $s(n)$ arrives from the same direction as the interference signal $u_1(n)$. In this case, the covariance matrices $\mathbf{R}_{\mathrm{s}}$ and $\mathbf{R}_{\mathrm{u}_1}$ are numerically similar, which also occurs with the normalized sample covariance matrices $\mathbf{R}_{\mathrm{s}}^{\mathrm{N}}$ and $\mathbf{R}_{\mathrm{u}_1}^{\mathrm{N}}$. Thereby, replacing $\mathbf{R}_{\mathrm{u}_1}$ by $\mathbf{R}_{\mathrm{s}}$ and $\mathbf{R}_{\mathrm{u}_1}^{\mathrm{N}}$ by $\mathbf{R}_{\mathrm{s}}^{\mathrm{N}}$ in (9), we get after some algebra

$$\mathrm{E}\left[\mathbf{v}_1(n)\right] \approx \mathrm{E}\left[\mathbf{w}(n)\right] + \mu_{\mathrm{s}}\left\{\mathbf{R}_{\mathrm{s}}\mathrm{E}\left[\mathbf{w}(n)\right] - \mathbf{R}_{\mathrm{s}}^{\mathrm{N}}\mathbf{R}_{\mathrm{s}}\mathrm{E}\left[\mathbf{w}(n)\right] - \mathbf{R}_{\mathrm{u}_2}^{\mathrm{N}}\mathbf{R}_{\mathrm{s}}\mathrm{E}\left[\mathbf{w}(n)\right]\right\}. \quad (20)$$

A generic matrix $\mathbf{R}$ (for subscript s or u_k) and the corresponding normalized $\mathbf{R}^{\mathrm{N}}$ have the same eigenvectors (see Appendix I), so for each matrix product having the form $\mathbf{R}^{\mathrm{N}}\mathbf{R}$, we can write

$$\mathbf{R}^{\mathrm{N}}\mathbf{R} = \mathbf{Q}\mathbf{\Lambda}^{\mathbf{N}}\mathbf{Q}^{\mathrm{H}}\mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^{\mathrm{H}} = \mathbf{Q}\mathbf{\Lambda}^{\mathbf{N}}\mathbf{\Lambda}\mathbf{Q}^{\mathrm{H}} \quad (21)$$

where $\mathbf{Q}$ is the eigenvector matrix of $\mathbf{R}$, and $\mathbf{\Lambda}$ and $\mathbf{\Lambda}^{\mathrm{N}}$ are diagonal matrices containing the eigenvalues λ_i and λ_i^{N} ($i = 1, 2, \ldots, M$) of $\mathbf{R}$ and $\mathbf{R}^{\mathrm{N}}$, respectively. When the AoA spread (defined as Δ in [1]) is small (for instance, $\Delta = 3^{\circ}$), the covariance matrices obtained from the signal model used in [1] have a large eigenvalue dispersion, resulting in a dominant eigenvalue λ_1, i.e., $\lambda_1 \gg \lambda_2 > \cdots > \lambda_M$. In this situation, matrix $\mathbf{R}^{\mathrm{N}}$ has a dominant eigenvalue with a value slightly less than 1 [see Appendix II]. As a result, one can make the following approximations: $\mathbf{\Lambda}^{\mathrm{N}}\mathbf{\Lambda} \cong \alpha\mathbf{\Lambda}$ and $\mathbf{R}^{\mathrm{N}}\mathbf{R} \cong \alpha\mathbf{R}$, for $0 < \alpha < 1$, leading to

$$\mathbf{R}_{\mathrm{s}}^{\mathrm{N}}\mathbf{R}_{\mathrm{s}}\mathrm{E}\left[\mathbf{w}(n)\right] \approx \alpha\mathbf{R}_{\mathrm{s}}\mathrm{E}\left[\mathbf{w}(n)\right]. \quad (22)$$

Then, substituting (22) into (20), we obtain

$$\mathrm{E}\left[\mathbf{v}_1(n)\right] \approx \mathrm{E}\left[\mathbf{w}(n)\right] + \mu_{\mathrm{s}}\left\{(1-\alpha)\mathbf{R}_{\mathrm{s}}\mathrm{E}\left[\mathbf{w}(n)\right] - \mathbf{R}_{\mathrm{u}_2}^{\mathrm{N}}\mathbf{R}_{\mathrm{s}}\mathrm{E}\left[\mathbf{w}(n)\right]\right\}. \quad (23)$$

Now, by analyzing (23) we observe the following:

i) Term $\mathbf{R}_{\mathrm{s}}\mathrm{E}[\mathbf{w}(n)]$ aims to maximize the power radiated to the in-cell MT. This is inferred from the fact that the gradient of the numerator of (19) is $\mathbf{R}_{\mathrm{IC}}\mathrm{E}[\mathbf{w}(n)]$ and $\mathbf{R}_{\mathrm{s}}$ is equivalent to $\mathbf{R}_{\mathrm{IC}}$.

ii) Term $\mathbf{R}_{\mathrm{u}_2}^{\mathrm{N}}\mathbf{R}_{\mathrm{s}}\mathrm{E}[\mathbf{w}(n)]$ avoids radiating power to the co-channel MT. This characteristic is intrinsic to the algorithm nature.

As a result, the term $\mathbf{R}_{\mathrm{u}_2}^{\mathrm{N}}\mathbf{R}_{\mathrm{s}}\mathrm{E}[\mathbf{w}(n)]$ in (23) is dominant as compared with $(1-\alpha)\mathbf{R}_{\mathrm{s}}\mathrm{E}[\mathbf{w}(n)]$, since under unbalanced conditions, α becomes approximately equal to 1. Therefore, through model expressions it is verified that, under certain conditions, the standard CSG algorithm gives more importance to minimizing the co-channel interference along with considerably reducing the power radiated to the in-cell MT. A similar effect can be observed when $\mathbf{R}_{\mathrm{u}_2} \cong \mathbf{R}_{\mathrm{s}}$.

Now, let us consider the case when two interferers have AoAs close to the in-cell AoA. For this case, the update vector in (9), defined as $\mathbf{g}$, is given by

$$\mathbf{g} = \left(\mathbf{I} - \mathbf{R}_{\mathrm{u}_1}^{\mathrm{N}} - \mathbf{R}_{\mathrm{u}_2}^{\mathrm{N}}\right)\mathbf{R}_{\mathrm{s}}\mathrm{E}\left[\mathbf{w}(n)\right]. \quad (24)$$

Then, the projection of $\mathbf{g}$ in the direction maximizing P_{IC} results in

$$\begin{aligned}
&\mathrm{E}\left[\mathbf{w}^{\mathrm{H}}(n)\right]\mathbf{R}_{\mathrm{s}}^{\mathrm{H}}\mathbf{g} \\
&\quad = \mathrm{E}\left[\mathbf{w}^{\mathrm{H}}(n)\right]\mathbf{R}_{\mathrm{s}}^{\mathrm{H}}\left(\mathbf{I} - \mathbf{R}_{\mathrm{u}_1}^{\mathrm{N}} - \mathbf{R}_{\mathrm{u}_2}^{\mathrm{N}}\right)\mathbf{R}_{\mathrm{s}}\mathrm{E}\left[\mathbf{w}(n)\right] \\
&\quad = \mathrm{E}\left[\mathbf{w}^{\mathrm{H}}(n)\right]\mathbf{R}_{\mathrm{s}}^{\mathrm{H}}\mathbf{R}_{\mathrm{s}}\mathrm{E}\left[\mathbf{w}(n)\right] \\
&\qquad - \mathrm{E}\left[\mathbf{w}^{\mathrm{H}}(n)\right]\mathbf{R}_{\mathrm{s}}^{\mathrm{H}}\mathbf{R}_{\mathrm{u}_1}^{\mathrm{N}}\mathbf{R}_{\mathrm{s}}\mathrm{E}\left[\mathbf{w}(n)\right] \\
&\qquad - \mathrm{E}\left[\mathbf{w}^{\mathrm{H}}(n)\right]\mathbf{R}_{\mathrm{s}}^{\mathrm{H}}\mathbf{R}_{\mathrm{u}_2}^{\mathrm{N}}\mathbf{R}_{\mathrm{s}}\mathrm{E}\left[\mathbf{w}(n)\right]. \quad (25)
\end{aligned}$$

In (25), $\mathbf{R}_{\mathrm{s}}^{\mathrm{H}}\mathbf{R}_{\mathrm{s}}$, $\mathbf{R}_{\mathrm{s}}^{\mathrm{H}}\mathbf{R}_{\mathrm{u}_1}^{\mathrm{N}}\mathbf{R}_{\mathrm{s}}$, and $\mathbf{R}_{\mathrm{s}}^{\mathrm{H}}\mathbf{R}_{\mathrm{u}_2}^{\mathrm{N}}\mathbf{R}_{\mathrm{s}}$ are semi-positive-definite matrices, so that all terms of (25) are greater than or equal to zero. Now, if the statistics of the interference signals are similar to the desired signal (similar AoA), the sum of the second and third r.h.s. terms in (25) may be larger than the first one. In this way, the update direction is opposite to the direction that maximizes the power radiated to the in-cell MT. In addition, unbalanced behavior may also occur when the effect of both interferers is combined, even for AoA interferers somewhat separated (see Section V-B, Example 3).

B. Improved CSG Algorithm

In this section, a modification to the standard CSG algorithm is proposed. The modified version, here termed the improved CSG (ICSG) algorithm, circumvents the unbalanced behavior previously discussed. Thus, considering the critical case $\mathbf{R}_{\mathrm{u}_1} \cong \mathbf{R}_{\mathrm{s}}$, a possible compensation for the unbalanced behavior is to reduce the effect of $\mathbf{R}_{\mathrm{u}_2}^{\mathrm{N}}\mathbf{R}_{\mathrm{s}}\mathrm{E}[\mathbf{w}(n)]$ in the same way as $\mathbf{R}_{\mathrm{s}}\mathrm{E}[\mathbf{w}(n)]$ is diminished by $\mathbf{R}_{\mathrm{u}_1}^{\mathrm{N}}\mathbf{R}_{\mathrm{s}}\mathrm{E}[\mathbf{w}(n)]$. A straightforward way to do this is to include the term $\mathbf{R}_{\mathrm{u}_2}^{\mathbf{N}}\mathbf{R}_{\mathrm{u}_1}^{\mathbf{N}}\mathbf{R}_{\mathrm{s}}\mathrm{E}[\mathbf{w}(n)]$ into the updating expression (9), resulting in

$$\mathrm{E}\left[\mathbf{v}_1(n)\right] = \mathrm{E}\left[\mathbf{w}(n)\right] + \mu_{\mathrm{s}}\left\{\mathbf{R}_{\mathrm{s}}\mathrm{E}\left[\mathbf{w}(n)\right] - \mathbf{R}_{\mathrm{u}_1}^{\mathrm{N}}\mathbf{R}_{\mathrm{s}}\mathrm{E}\left[\mathbf{w}(n)\right] - \mathbf{R}_{\mathrm{u}_2}^{\mathrm{N}}\mathbf{R}_{\mathrm{s}}\mathrm{E}\left[\mathbf{w}(n)\right] + \mathbf{R}_{\mathrm{u}_2}^{\mathrm{N}}\mathbf{R}_{\mathrm{u}_1}^{\mathrm{N}}\mathbf{R}_{\mathrm{s}}\mathrm{E}\left[\mathbf{w}(n)\right]\right\}. \quad (26)$$

TABLE I
CSG AND ICSG ALGORITHM COMPUTATIONAL BURDEN COMPARISON

Algorithm	Number of real operations per iterations			
	Multiplications	Additions	Divisions	Square root
CSG	$56M+8$	$38M-8$	$8M$	1
ICSG	$72M+18$	$54M-8$	$12M$	1

TABLE II
PARAMETER VALUES USED FOR THE NUMERICAL SIMULATIONS

Case	AoA spread, Δ	M	$\mu_s=\mu_u$
1	30°	4	0.01
2	30°	8	0.007
3	3°	4	0.05
4	3°	8	0.01

Vector $\mathbf{R}_{u_2}^N\mathbf{R}_{u_1}^N\mathbf{R}_s\mathrm{E}[\mathbf{w}(n)]$ compensates for the mentioned unbalanced behavior, as follows. Replacing $\mathbf{R}_{u_1}$ by $\mathbf{R}_s$ in (26), we obtain

$$\mathrm{E}[\mathbf{v}_1(n)] \approx \mathrm{E}[\mathbf{w}(n)] + \mu_s\left\{\mathbf{R}_s\mathrm{E}[\mathbf{w}(n)] - \mathbf{R}_s^N\mathbf{R}_s\mathrm{E}[\mathbf{w}(n)] - \mathbf{R}_{u_2}^N\mathbf{R}_s\mathrm{E}[\mathbf{w}(n)] + \mathbf{R}_{u_2}^N\mathbf{R}_s^N\mathbf{R}_s\mathrm{E}[\mathbf{w}(n)]\right\}. \quad (27)$$

In the same way as before, approximations $\mathbf{R}_s^N\mathbf{R}_s \approx \alpha\mathbf{R}_s$ and $\mathbf{R}_{u_2}^N\mathbf{R}_s^N\mathbf{R}_s \approx \alpha\mathbf{R}_{u_2}^N\mathbf{R}_s$ can be verified. Note that $\mathbf{R}_s^N\mathbf{R}_s\mathrm{E}[\mathbf{w}(n)]$ still reduces the effect of $\mathbf{R}_s\mathrm{E}[\mathbf{w}(n)]$; however, the term $\mathbf{R}_{u_2}^N\mathbf{R}_s\mathrm{E}[\mathbf{w}(n)]$ is also reduced accordingly to $\mathbf{R}_{u_2}^N\mathbf{R}_s^N\mathbf{R}_s\mathrm{E}[\mathbf{w}(n)]$. Hence, vector $\mathbf{R}_{u_2}^N\mathbf{R}_{u_1}^N\mathbf{R}_s\mathrm{E}[\mathbf{w}(n)]$ circumvents the unbalanced behavior of (9) when $\mathbf{R}_{u_1}\cong\mathbf{R}_s$.

Similarly as before, considering the case $\mathbf{R}_{u_2}\cong\mathbf{R}_s$, matrix $\mathbf{R}_{u_1}^N\mathbf{R}_{u_2}^N$ must also be included in the updating expression (9), resulting in

$$\mathrm{E}[\mathbf{v}_1(n)] = \mathrm{E}[\mathbf{w}(n)] + \mu_s\left(\mathbf{I} - \mathbf{R}_{u_1}^N - \mathbf{R}_{u_2}^N + \mathbf{R}_{u_1}^N\mathbf{R}_{u_2}^N + \mathbf{R}_{u_2}^N\mathbf{R}_{u_1}^N\right)\mathbf{R}_s\mathrm{E}[\mathbf{w}(n)]. \quad (28)$$

Now, let us consider the case when two interferers have AoAs close to the in cell AoA. Similarly as before, (28) is analyzed by projecting the update vector, defined now as

$$\mathbf{g}' = \left(\mathbf{I} - \mathbf{R}_{u_1}^N - \mathbf{R}_{u_2}^N + \mathbf{R}_{u_1}^N\mathbf{R}_{u_2}^N + \mathbf{R}_{u_2}^N\mathbf{R}_{u_1}^N\right)\mathbf{R}_s\mathrm{E}[\mathbf{w}(n)] \quad (29)$$

in the direction that maximizes P_{IC}. Thereby, we obtain

$$\begin{aligned}&\mathrm{E}\left[\mathbf{w}^H(n)\right]\mathbf{R}_s^H\mathbf{g}'\\ &= \mathrm{E}\left[\mathbf{w}^H(n)\right]\mathbf{R}_s^H\left(\mathbf{I} - \mathbf{R}_{u_1}^N - \mathbf{R}_{u_2}^N + \mathbf{R}_{u_1}^N\mathbf{R}_{u_2}^N + \mathbf{R}_{u_2}^N\mathbf{R}_{u_1}^N\right)\mathbf{R}_s\mathrm{E}[\mathbf{w}(n)]\\ &= \mathrm{E}\left[\mathbf{w}^H(n)\right]\mathbf{R}_s^H\mathbf{R}_s\mathrm{E}[\mathbf{w}(n)]\\ &\quad - \mathrm{E}\left[\mathbf{w}^H(n)\right]\mathbf{R}_s^H\mathbf{R}_{u_1}^N\mathbf{R}_s\mathrm{E}[\mathbf{w}(n)]\\ &\quad - \mathrm{E}\left[\mathbf{w}^H(n)\right]\mathbf{R}_s^H\mathbf{R}_{u_2}^N\mathbf{R}_s\mathrm{E}[\mathbf{w}(n)]\\ &\quad + \mathrm{E}\left[\mathbf{w}^H(n)\right]\mathbf{R}_s^H\mathbf{R}_{u_1}^N\mathbf{R}_{u_2}^N\mathbf{R}_s\mathrm{E}[\mathbf{w}(n)]\\ &\quad + \mathrm{E}\left[\mathbf{w}^H(n)\right]\mathbf{R}_s^H\mathbf{R}_{u_2}^N\mathbf{R}_{u_1}^N\mathbf{R}_s\mathrm{E}[\mathbf{w}(n)]. \end{aligned} \quad (30)$$

In the considered condition, from (30) we verify that the incorporated terms act in favor of maximizing P_{IC}. However, in uncritical conditions, the impact of those terms is insignificant, since matrices $\mathbf{R}_s^H\mathbf{R}_{u_2}^N\mathbf{R}_{u_1}^N\mathbf{R}_s$ and $\mathbf{R}_s^H\mathbf{R}_{u_1}^N\mathbf{R}_{u_2}^N\mathbf{R}_s$ are negligible as compared to $\mathbf{R}_s^H\mathbf{R}_s$, $\mathbf{R}_s^H\mathbf{R}_{u_1}^N\mathbf{R}_s$, and $\mathbf{R}_s^H\mathbf{R}_{u_2}^N\mathbf{R}_s$, leading (28) to have similar behavior to (9).

Since better performance is verified considering model expressions (10) and (28), we can now formulate the ICSG algorithm including instantaneous estimates of matrices $\mathbf{R}_{u_1}^N\mathbf{R}_{u_2}^N$ and $\mathbf{R}_{u_2}^N\mathbf{R}_{u_1}^N$ into the standard CSG algorithm expressions. Therefore, (4) is modified to

$$\mathbf{v}_1(n) = \mathbf{w}(n) + \mu_s\left[\mathbf{I} - \frac{\hat{\mathbf{R}}_{u_1}(n)}{\|\mathbf{u}_1(n)\|^2} - \frac{\hat{\mathbf{R}}_{u_2}(n)}{\|\mathbf{u}_2(n)\|^2} + \frac{\hat{\mathbf{R}}_{u_1}(n)}{\|\mathbf{u}_1(n)\|^2}\cdot\frac{\hat{\mathbf{R}}_{u_2}(n)}{\|\mathbf{u}_2(n)\|^2} + \frac{\hat{\mathbf{R}}_{u_2}(n)}{\|\mathbf{u}_2(n)\|^2}\cdot\frac{\hat{\mathbf{R}}_{u_1}(n)}{\|\mathbf{u}_1(n)\|^2}\right]\hat{\mathbf{R}}_s(n)\mathbf{w}(n) \quad (31)$$

while (5) and (6) remain unchanged.

The proposed algorithm also outperforms the standard CSG algorithm for $\theta_{u_1}\cong\theta_s$ and $\theta_{u_2}\cong\theta_s$. The addition of the compensated terms in the improved algorithm version leads to a slight increase in the computational burden, as detailed in Table I.

V. SIMULATION RESULTS

In this section, Monte Carlo (MC) simulations are compared with the stochastic model predictions for both the standard CSG and ICSG algorithms for the two-interferer case. In the examples shown, exact downlink covariance matrices are used and determined as in [1], considering twelve independent fading paths ($I=12$). For the downlink signals, an in-cell signal-to-noise ratio $\rho_{IC}=36$ dB and a co-channel signal-to-noise ratio $\rho_{CC}=15$ dB are considered [1]. For the MC simulations, 100 independent realizations are used. Comments regarding the simulation results are given at the end of this section.

A. Normal Behavior

Here, a comparison between the characteristics of the standard CSG and ICSG algorithms under normal conditions is presented. For each algorithm, curves resulting from MC simulations and stochastic model expressions are shown. For the uplink channel, the in-cell signal angle-of-arrival is $\theta_s=60^\circ$ and the interference angles-of-arrival are $\theta_{u_1}=-16^\circ$ and $\theta_{u_2}=36^\circ$.

1) SINR: To assess the model expressions with respect to the behavior of the SINR $\gamma(n)$, (17) is used with the matrix $\mathbf{K}(n)$ determined from both MC simulations and the proposed analytical model. Four situations are considered (see Table II).

In Table II, the use of $\mu_s=\mu_u$ is reasonable for equal signal and interference power. Moreover, from numerical simulations, no great advantage for using unequal step-size values has been observed. In Fig. 2, the resulting SINR curves for each case are shown. Note that the considered conditions do not lead to unbalanced behavior of the standard CSG algorithm. In addition, very good accuracy for the model prediction, obtained by (19), is verified.

2) Mean Weight Behavior: To evaluate the mean weight behavior of the adaptive weight vector, Cases 1, 2, 3, and 4 are considered. In Fig. 3, the mean weight real part evolution for all cases are shown, presenting the results obtained from the MC simulations and the prediction models. Similar results have been obtained for the imaginary part.

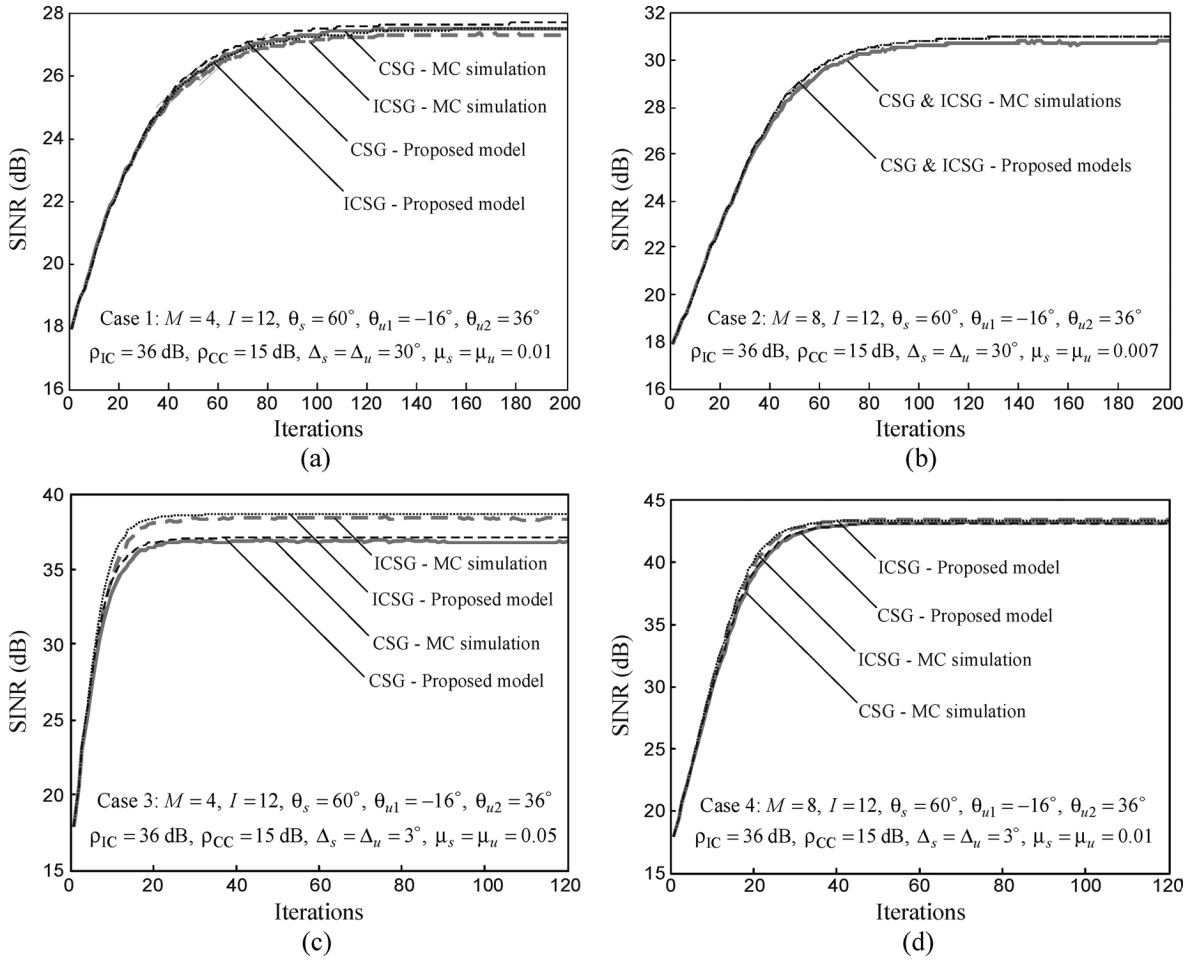

Fig. 2. Behavior of SINR $\gamma(n)$ under normal conditions. Standard CSG algorithm: (gray ragged lines) MC simulations, (dark dashed lines) proposed model. ICSG algorithm: (gray dashed lines) MC simulations, (dark dotted lines) proposed model. (a) Case 1. (b) Case 2. (c) Case 3. (d) Case 4.

B. Unbalanced Behavior

In this section, a comparison between the behavior of the SINR curves for both the standard CSG and ICSG algorithms under unbalanced conditions is presented. For each algorithm, the curves are obtained from MC simulations and the proposed stochastic models. Three examples are shown.

1) Example 1: In this example, the case $\theta_{u_1} \cong \theta_s$ is considered. The algorithm parameters are $M = 4$, $I = 12$, $\theta_s = \theta_{u_1} = -12^\circ$, $\theta_{u_2} = 36^\circ$, $\Delta_s = \Delta_u = 3^\circ$, and $\mu_s = \mu_u = 0.05$. Fig. 4 shows the SINR characteristic for the CSG and ICSG algorithms. In Fig. 5, the radiation pattern of the antenna array is plotted, considering that the coefficients have reached the mean steady-state value as determined by (13).

2) Example 2: In this example, the case $\theta_{u_2} \cong \theta_s$ is assessed. The algorithm parameters are $M = 8, I = 12, \theta_s = \theta_{u_2} = 36^\circ$, $\theta_{u_1} = -36^\circ$, $\Delta_s = \Delta_u = 3^\circ$, and $\mu_s = \mu_u = 0.01$. Fig. 6 shows the SINR curves for the CSG and ICSG algorithms and Fig. 7 presents the associated radiation patterns considering the steady-state weight vectors.

3) Example 3: In this example, the case where both θ_{u_1} and θ_{u_2} are within the nominal beamwidth centered at θ_s is considered. The algorithm parameters are the same as in Example 1, except that now $\theta_s = 0^\circ$, $\theta_{u_1} = -12^\circ$, $\theta_{u_2} = 12^\circ$. Fig. 8 shows the SINR characteristics for the CSG and ICSG algorithms and Fig. 9, the corresponding radiation diagrams in the steady-state conditions.

C. Verification of Key Assumptions

In this section, numerical simulations demonstrating the validity of assumptions (8) and (18) are presented. Fig. 10 illustrates assumption (8) considering two cases: Fig. 10(a) shows each of the vector components for a balanced case (see Section V-A, Case 1) and Fig. 10(b), for an unbalanced case (see Section V-B, Example 3). From the obtained plots, we verify the reasonability of the used approximation, given the very good match between the curves. In Fig. 11, assumption (18) is likewise verified through the comparison of (16) and (19), by using the same cases as in Fig. 10.

D. Discussion

From the presented figures, very good agreement between simulation and model is evident. Fig. 2 illustrates the behavior of the ICSG algorithm under normal working conditions, which is similar to the standard CSG. Especially, for some balanced conditions, for instance Case 3, a slightly larger steady-state SINR value is noticed for the ICSG algorithm as compared with the standard CSG. This can be explained by the fact that the ICSG algorithm favors maximizing P_{IC} instead of minimizing P_{CC}.

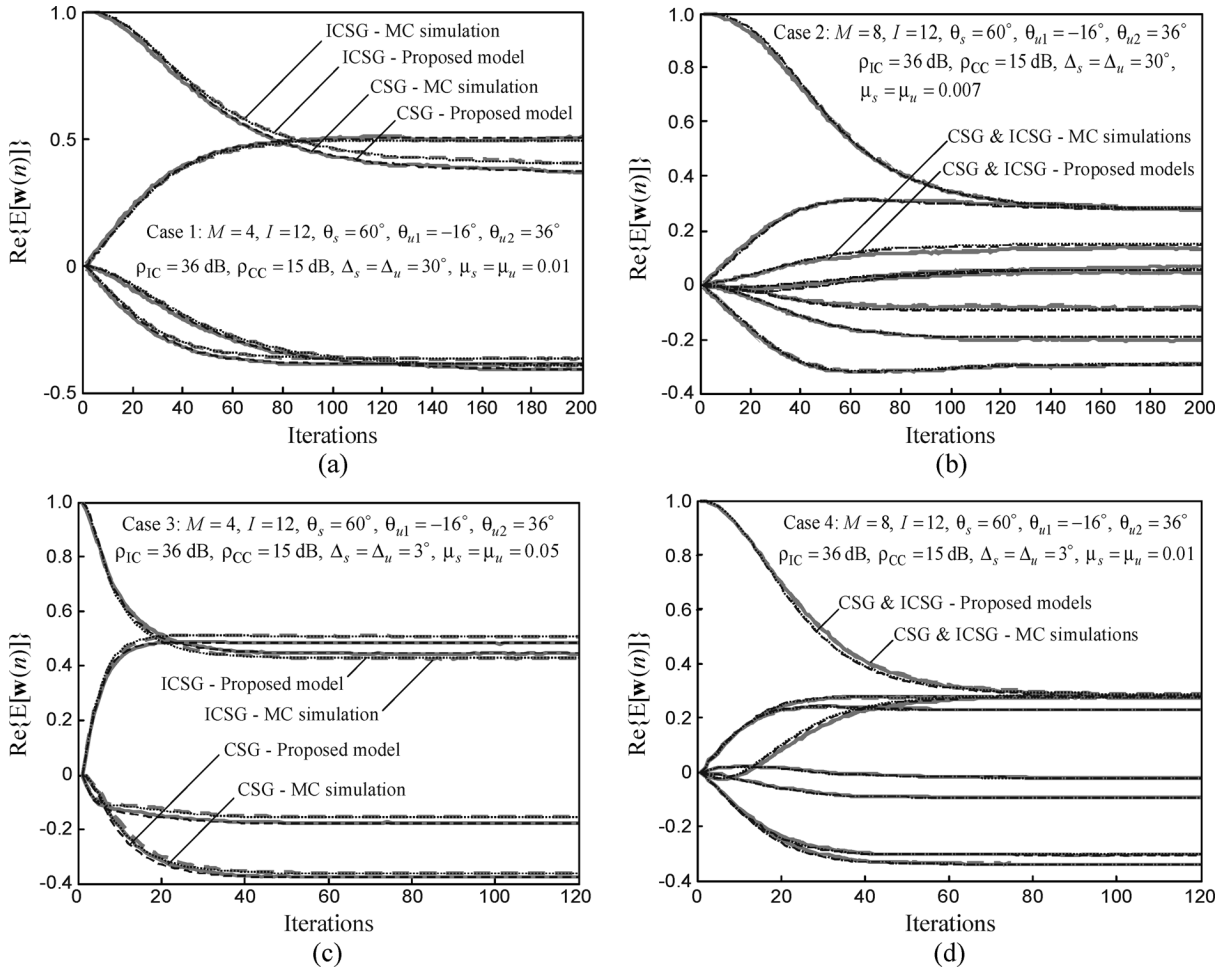

Fig. 3. Real part of mean weight behavior under normal conditions. Standard CSG algorithm: (gray ragged lines) MC simulations, (dark dashed lines) proposed model. ICSG algorithm: (gray dashed lines) MC simulations, (dark dotted lines) proposed model. (a) Case 1. (b) Case 2. (c) Case 3. (d) Case 4.

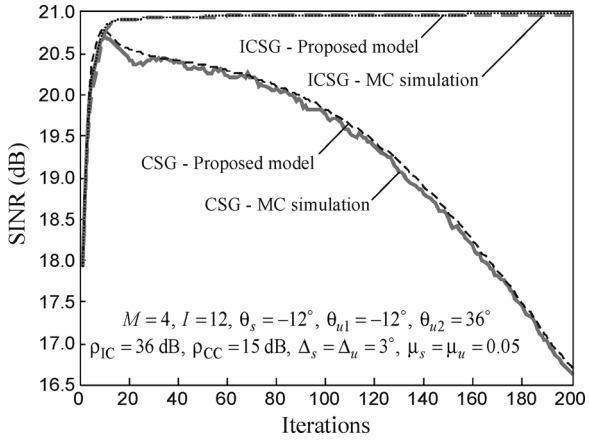

Fig. 4. Unbalanced behavior, Example 1: SINR curve. Standard CSG algorithm: (gray ragged lines) MC simulations, (dark dashed lines) proposed model. ICSG algorithm: (gray dashed lines) MC simulations, (dark dotted lines) proposed model.

In Figs. 4, 6, and 8 for unbalanced conditions, a peak in the SINR curves of the standard CSG algorithm is verified. Such a characteristic is due to the term $\mathrm{E}[P_{\mathrm{IC}}]$ in (15), which reduces more slowly than $\mathrm{E}[P_{\mathrm{CC}}]$: while $\mathrm{E}[P_{\mathrm{CC}}] \geq 1$, the SINR is growing, whereas when $\mathrm{E}[P_{\mathrm{CC}}] < 1$, the denominator of (15) becomes almost constant as the SINR decreases. The difference between simulation and theory observed in Figs. 6 and 8 is due to the approximations used in the model derivation (see Section V-C).

From the radiation patterns shown in Figs. 5, 7, and 9, one can verify that during the unbalanced behavior of the standard CSG algorithm, the MT does not receive enough power from the base station, indicating that service has been interrupted. On the other hand, for the antennas controlled by the ICSG algorithm, sufficient power is always provided to the MT.

VI. Concluding Remarks

In this work, analytic models for the first moment of the weight vector and the SINR characteristic were presented. The models are obtained assuming a slow adaptation condition. A very good match between MC simulations and the prediction model is obtained for both transient and steady-state behavior. By using the proposed model, the behavior of the adaptive array can be efficiently assessed over a wide range of working conditions. Thereby, an unbalanced behavior of the standard CSG algorithm can be verified. In addition, a study of this unbalanced behavior leads to an improved version of the CSG algorithm, which compensates for such undesired behavior.

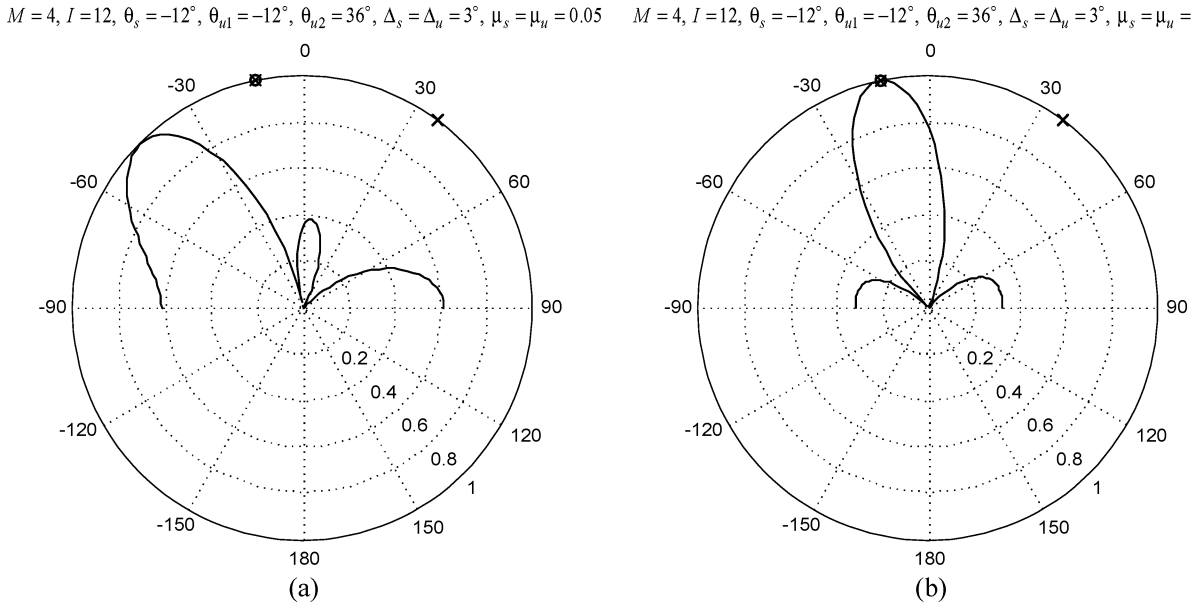

Fig. 5. Example 1: radiation patterns for $\theta_{\rm s} = -12^\circ$ ($\circ$), $\theta_{\rm u_1} = -12^\circ$ ($\times$), and $\theta_{\rm u_2} = 36^\circ$ ($\times$). (a) Standard CSG algorithm. (b) ICSG algorithm.

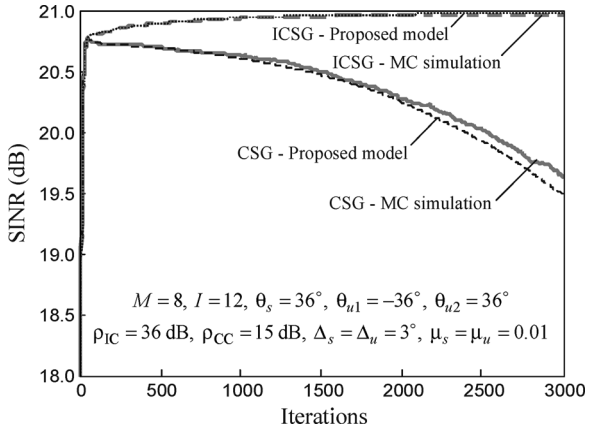

Fig. 6. Unbalanced behavior, Example 2: SINR curves. Standard CSG algorithm: (gray ragged lines) MC simulations, (dark dashed lines) proposed model. ICSG algorithm: (gray dashed lines) MC simulations, (dark dotted lines) proposed model.

Appendix I
Determination of the Normalized Sample Covariance Matrix

In this Appendix, an approach for determining a normalized sample covariance matrix $\mathbf{R}^{\rm N}$, defined as

$$\mathbf{R}^{\rm N} = \mathrm{E}\left[\frac{\mathbf{x}(n)\mathbf{x}^{\rm H}(n)}{\mathbf{x}^{\rm H}(n)\mathbf{x}(n)}\right] \tag{32}$$

is presented. Here, as in [1], a complex Gaussian signal vector $\mathbf{x}(n)$, having a covariance matrix $\mathbf{R}$, is assumed.

To determine (32), we have used an approach similar to the one presented in [11], applied now to the complex Gaussian case [12]. Thereby, an auxiliary matrix function $\mathbf{F}(\omega)$ is defined with elements given by

$$f_{i,j}(\omega) = \frac{1}{\pi^M \det(\mathbf{R})} \underbrace{\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty}}_{M \text{ fold}} \frac{x_i x_j^*}{\mathbf{x}^{\rm H}\mathbf{x}} \mathrm{e}^{-\mathbf{x}^{\rm H}\mathbf{L}^{-1}(\omega)\mathbf{x}} d\mathbf{x} \tag{33}$$

with

$$\mathbf{L}(\omega) = (\mathbf{R}^{-1} + \omega\mathbf{I})^{-1}. \tag{34}$$

Note that for $\omega = 0$, (33) is by definition the desired expectation, i.e.,

$$\mathrm{E}\left[\frac{\mathbf{x}(n)\mathbf{x}^{\rm H}(n)}{\mathbf{x}^{\rm H}(n)\mathbf{x}(n)}\right] = \mathbf{F}(0). \tag{35}$$

Applying partial differentiation in (33) with respect to ω, the term $\mathbf{x}^{\rm H}\mathbf{x}$ in the denominator of its integrand is eliminated, resulting in the expression

$$\frac{\partial f_{i,j}(\omega)}{\partial \omega} = \frac{-\det[\mathbf{L}(\omega)]}{\det(\mathbf{R})} A(\omega) \tag{36}$$

where

$$A(\omega) = \frac{1}{\pi^M \det[\mathbf{L}(\omega)]} \underbrace{\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty}}_{M \text{ fold}} x_i x_j^* \mathrm{e}^{-\mathbf{x}^{\rm H}\mathbf{L}^{-1}(\omega)\mathbf{x}} d\mathbf{x} \tag{37}$$

corresponds to the cross-correlation between x_i and x_j when x_i and x_j are jointly Gaussian complex random variables with covariance matrix $\mathbf{L}(\omega)$. Therefore, the factor $A(\omega) = [\mathbf{L}(\omega)]_{i,j}$, and integrating (36) with respect to ω, considering that

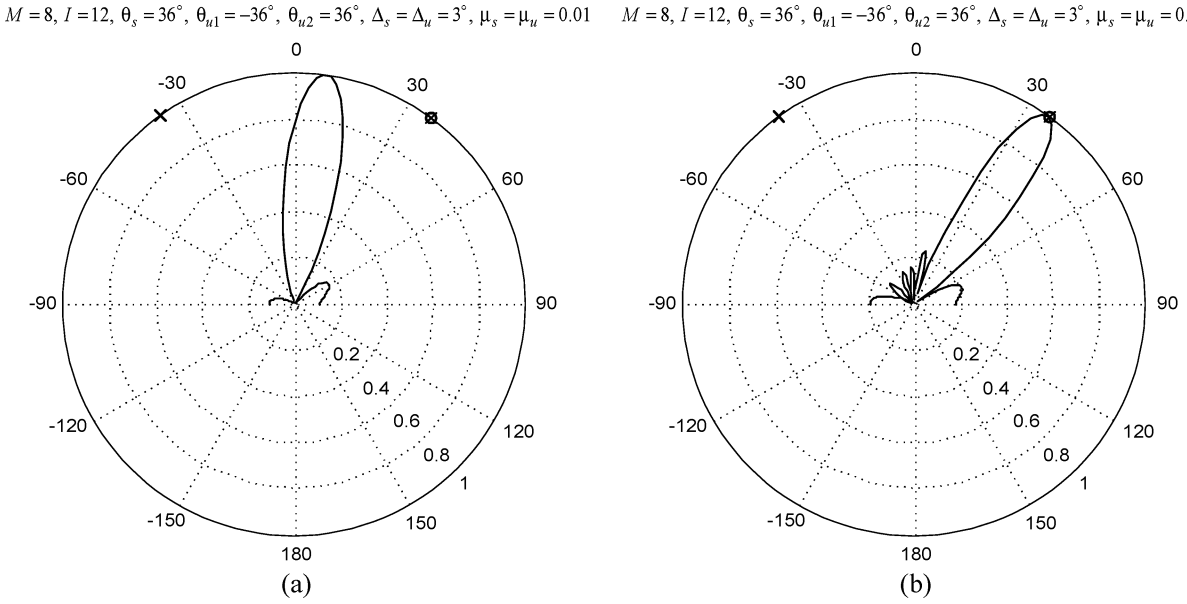

Fig. 7. Example 2: radiation patterns for $\theta_{\rm s} = 36^\circ$ ($\circ$), $\theta_{\rm u_1} = -36^\circ$ ($\times$), and $\theta_{\rm u_2} = 36^\circ$ ($\times$). (a) Standard CSG algorithm. (b) ICSG algorithm.

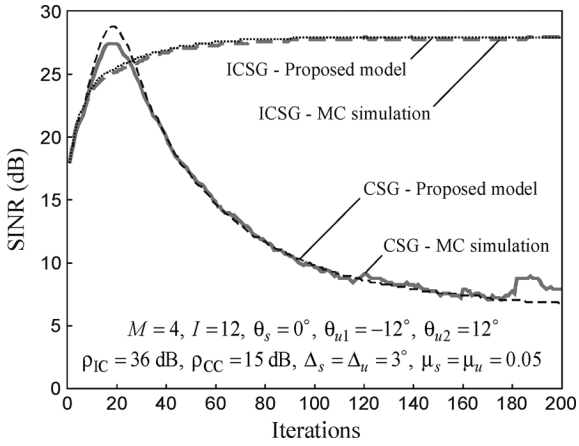

Fig. 8. Unbalanced behavior, Example 3: SINR curves. CSG algorithm: (gray ragged lines) MC simulations, (dark dashed lines) proposed model. ICSG algorithm: (gray dashed lines) MC simulations, (dark dotted lines) proposed model.

$\lim_{\omega\to\infty} f_{i,j}(\omega) = 0$, obtains a simpler expression for the elements of (35), in the same way as [11]. Thereby, (32) can be written as

$$\left\{ \mathrm{E}\left[\frac{\mathbf{x}(n)\mathbf{x}^{\mathrm{H}}(n)}{\mathbf{x}^{\mathrm{H}}(n)\mathbf{x}(n)} \right] \right\}_{i,j} = \int_0^\infty \frac{\{\mathbf{R}[\mathbf{I}+\omega\mathbf{R}]^{-1}\}_{i,j}}{\det(\mathbf{I}+\omega\mathbf{R})} d\omega. \quad (38)$$

By expressing $\mathbf{R} = \mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^{\mathrm{H}}$, where $\mathbf{Q}$ is the eigenvector matrix of $\mathbf{R}$ and $\mathbf{\Lambda}$ is a diagonal matrix containing the eigenvalues λ_i of $\mathbf{R}$ for $i = 1, 2, \ldots, M$, we can write

$$\mathrm{E}\left[\frac{\mathbf{x}(n)\mathbf{x}^{\mathrm{H}}(n)}{\mathbf{x}^{\mathrm{H}}(n)\mathbf{x}(n)} \right] = \mathbf{Q}\mathbf{H}\mathbf{Q}^{\mathrm{H}} \quad (39)$$

where $\mathbf{H}$ is a diagonal matrix with elements

$$h_k = \int_0^\infty \frac{\lambda_k}{(1+\omega\lambda_k)\det(\mathbf{I}+\omega\Lambda)} d\omega. \quad (40)$$

Now, (40) can be calculated by using the partial fraction expansion technique. For instance, by considering distinct eigenvalues, the elements of matrix $\mathbf{H}$ are given by

$$h_k = A_{1,k}\lambda_k + A_{2,k}\ln(\lambda_k) + \sum_{\substack{i=1 \\ i\neq k}}^{M} B_{i,k}\ln(\lambda_i) \quad (41)$$

where

$$A_{1,k} = \frac{\lambda_k^{M-2}}{\prod_{\substack{i=1 \\ i\neq k}}^{M}(\lambda_k - \lambda_i)} \quad (42)$$

and

$$B_{i,k} = \frac{\lambda_k \lambda_i^{M-1}}{(\lambda_i - \lambda_k)\prod_{\substack{j=1 \\ j\neq i}}^{M}(\lambda_i - \lambda_j)} \quad (43)$$

with

$$A_{2,k} = -A_{1,k}\lambda_k \sum_{\substack{i=1 \\ i\neq k}}^{M} \frac{\lambda_i}{\lambda_k - \lambda_i}. \quad (44)$$

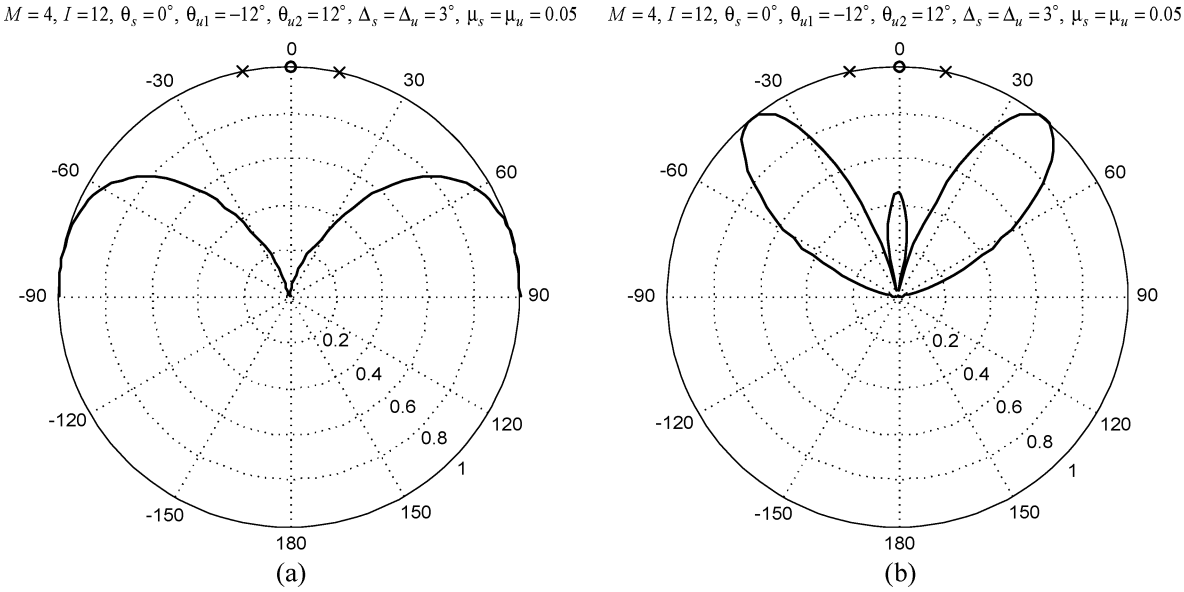

Fig. 9. Example 3: radiation patterns for $\theta_{\mathrm{s}} = 0^\circ$ ($\circ$), $\theta_{\mathrm{u}_1} = -12^\circ$ ($\times$), and $\theta_{\mathrm{u}_2} = 12^\circ$ ($\times$). (a) Standard CSG algorithm. (b) ICSG algorithm.

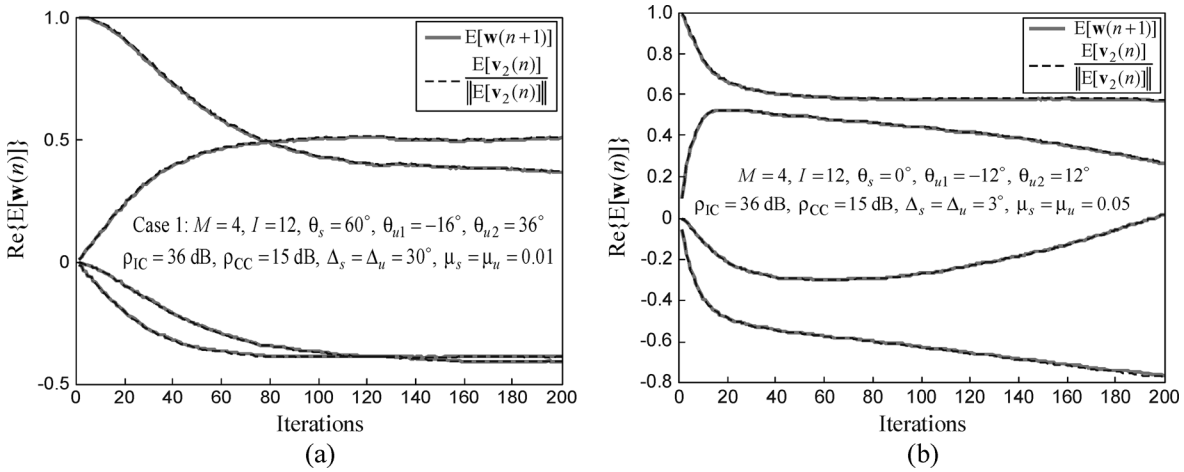

Fig. 10. Verification of assumption (8). (a) Balanced case. (b) Unbalanced case.

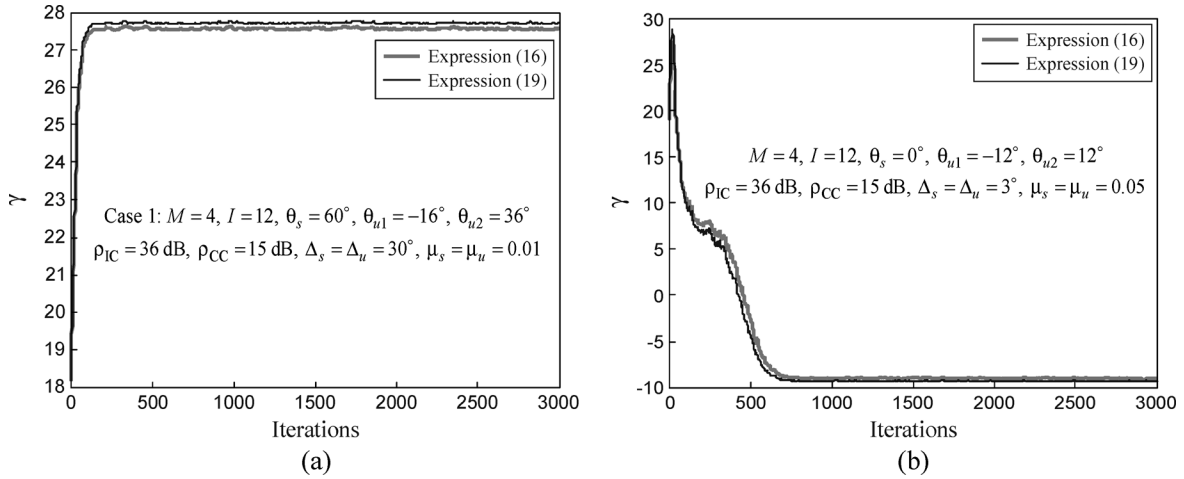

Fig. 11. Verification of assumption (18). (a) Balanced case. (b) Unbalanced case.

APPENDIX II
DOMINANT EIGENVALUE VERIFICATION

In this appendix, we verify the analysis assumption which uses the dominant eigenvalue approximation. To this end, considering that the covariance matrices, obtained from the signal model used in [1], have a large eigenvalue dispersion for small angle spread $\Delta \rightarrow 0$, well separated eigenvalues result, i.e., $\lambda_1 > \lambda_2 > \cdots > \lambda_M$. In this case, the following relations are valid:

$$\lim_{\Delta\rightarrow 0} \frac{\lambda_i}{\lambda_j} = 0, \quad j < i \tag{45}$$

and

$$\lim_{\Delta\rightarrow 0} \frac{\lambda_i}{\lambda_j} = \infty, \quad i < j. \tag{46}$$

From (41), we have for $k = 1$

$$h_1 = A_{1,1}\lambda_1 + A_{2,1}\ln(\lambda_1) + \sum_{i=2}^{M} B_{i,1}\ln(\lambda_i). \tag{47}$$

Now, each term of (47) is analyzed using (45) and (46) as follows:

i) Considering the coefficient $A_{1,1}$, we get

$$\lim_{\Delta\rightarrow 0} A_{1,1} = \lim_{\Delta\rightarrow 0} \frac{\lambda_1^{M-2}}{\prod_{l\neq 1}(\lambda_1 - \lambda_l)} = \lim_{\Delta\rightarrow 0} \frac{\lambda_1^{-1}}{\prod_{l\neq 1}\left(1 - \frac{\lambda_l}{\lambda_1}\right)} = \lambda_1^{-1} \tag{48}$$

which results in $\lim_{\Delta\rightarrow 0} A_{1,1}\lambda_1 = 1$.

ii) Now, for the coefficient $A_{2,1}$, we have

$$\lim_{\Delta\rightarrow 0} A_{2,1} = -\lim_{\Delta\rightarrow 0} \lambda_1 A_{1,1} \sum_{j\neq 1} \frac{\lambda_j}{\lambda_j - \lambda_1} = \lim_{\Delta\rightarrow 0} \sum_{j\neq 1} \frac{1}{1 - \frac{\lambda_1}{\lambda_j}} = 0 \tag{49}$$

thereby, $\lim_{\Delta\rightarrow 0} A_{2,1}\ln(\lambda_1) = 0$.

iii) Finally, for the coefficient $B_{i,1}$, we obtain

$$\begin{aligned} \lim_{\Delta\rightarrow 0} B_{i,1} &= \lim_{\Delta\rightarrow 0} \frac{\lambda_1\lambda_i^{M-1}}{(\lambda_i - \lambda_1)\prod_{l\neq i}(\lambda_i - \lambda_l)} \\ &= \lim_{\Delta\rightarrow 0} \frac{1}{\left(\frac{\lambda_i}{\lambda_1} - 1\right)\prod_{l\neq i}\left(1 - \frac{\lambda_l}{\lambda_i}\right)} = 0, \quad i \neq 1 \end{aligned} \tag{50}$$

resulting in $\lim_{\Delta\rightarrow 0} \sum_{i=2}^{M} B_{i,1}\ln(\lambda_i) = 0$.

Now, by using (i), (ii), and (iii) in (47), we obtain $\lim_{\Delta\rightarrow 0} h_1 = 1$.

Then, taking h_k for $k \neq 1$, using again (45) and (46), similarly as before, the following results are obtained:

$$\begin{aligned} \lim_{\Delta\rightarrow 0} A_{1,k}\lambda_k &= \lim_{\Delta\rightarrow 0} \frac{\lambda_k^{M-1}}{\prod_{l\neq k}(\lambda_k - \lambda_l)} \\ &= \lim_{\Delta\rightarrow 0} \frac{1}{\prod_{l\neq k}\left(1 - \frac{\lambda_l}{\lambda_k}\right)} = 0 \end{aligned} \tag{51}$$

$$\begin{aligned} \lim_{\Delta\rightarrow 0} A_{2,k} &= -\lim_{\Delta\rightarrow 0} \lambda_k A_{1,k} \sum_{j\neq k} \frac{\lambda_j}{\lambda_j - \lambda_k} \\ &= -\lim_{\Delta\rightarrow 0} \lambda_k A_{1,k} \sum_{j\neq k} \frac{1}{1 - \frac{\lambda_k}{\lambda_j}} = 0 \end{aligned} \tag{52}$$

and

$$\begin{aligned} \lim_{\Delta\rightarrow 0} B_{j,k} &= \lim_{\Delta\rightarrow 0} \frac{\lambda_k\lambda_j^{M-1}}{(\lambda_j - \lambda_k)\prod_{l\neq j}(\lambda_j - \lambda_l)} \\ &= \lim_{\Delta\rightarrow 0} \frac{1}{\left(\frac{\lambda_j}{\lambda_k} - 1\right)\prod_{l\neq j}\left(1 - \frac{\lambda_l}{\lambda_j}\right)} = 0, \quad j \neq k \end{aligned} \tag{53}$$

which leads to $\lim_{\Delta\rightarrow 0} h_k = 0$ for $k \neq 1$.

ACKNOWLEDGMENT

The authors would like to thank the Associate Editor and the anonymous reviewers for their valuable and constructive comments and suggestions, from which the revision of this paper has benefited significantly.

REFERENCES

[1] D. R. Morgan, "Downlink adaptive array algorithms for cellular mobile communications," *IEEE Trans. Commun.*, vol. 51, no. 3, pp. 476–488, Mar. 2003.

[2] J. Litva and T. K.-Y. Lo, *Digital Beamforming in Wireless Communications*. Norwood, MA: Artech House, 1996.

[3] L. C. Godara, "Cellular systems," in *Handbook in Antennas in Wireless Communications*, L. C. Godara, Ed. Boca Raton, FL: CRC Press, 2001.

[4] J. M. Goldberg and J. R. Fonollosa, "Downlink beamforming for spatially distributed sources in cellular mobile communications," *Signal Process.*, vol. 65, no. 2, pp. 181–197, Mar. 1998.

[5] B. M. Hochwald and T. L. Marzetta, "Adapting a downlink array from uplink measurements," *IEEE Trans. Signal Process.*, vol. 49, no. 3, pp. 642–653, Mar. 2001.

[6] Y.-C. Liang and F. P. S. Chin, "Downlink channel covariance matrix (DCCM) estimation and its applications in wireless DS-CDMA systems," *IEEE J. Sel. Areas Commun.*, vol. 19, no. 2, pp. 222–232, Feb. 2001.

[7] A. Papoulis and S. U. Pillai, *Probability, Random Variables and Stochastic Processes*. New York: McGraw-Hill, 2002.

[8] B. Farhang-Boroujeny, *Adaptive Filters: Theory and Applications*. New York: Wiley, 1998.

[9] J. Demmel, *Applied Numerical Linear Algebra*. Philadelphia, PA: SIAM, 1997.

[10] N. J. Bershad, P. Celka, and J.-M. Vesin, "Stochastic analysis of gradient adaptive identification of nonlinear systems with memory for Gaussian data and noisy input and output measurements," *IEEE Trans. Signal Process.*, vol. 47, no. 3, pp. 675–689, Mar. 1999.

[11] M. Rupp, "The behavior of LMS and NLMS algorithms in the presence of spherically invariant processes," *IEEE Trans. Signal Process.*, vol. 41, no. 3, pp. 1149–1160, Mar. 1993.

[12] C. W. Therrien, *Discrete Random Signals and Statistical Signal Processing*. Englewood Cliffs, NJ: Prentice-Hall, 1992.

Javier Ernesto Kolodziej was born in Posadas, Mns., Argentina. In 2002, he received the B.S. degree from the National University of Misiones, Argentina, and the M.Sc. degree, in 2006, from the Federal University of Santa Catarina, Florianopolis, Brazil, where he is currently working towards the Ph.D. degree.

He joined the Department of Electronic Engineering at the National University of Misiones in 2001. His current research interests include adaptive signal processing theory and its applications.

Orlando José Tobias (S'94–M'00) was born in Mar del Plata, Argentina. He received the B.S. degree in electrical engineering from the National University of La Plata, Argentina, in 1988 and the M.Sc. and Ph.D. degrees in electrical engineering from the Federal University of Santa Catarina, Florianópolis, Brazil, in 1995 and 1999, respectively.

He is now with the LINSE-Circuits and Signal Processing Laboratory at the Federal University of Santa Catarina, Brazil. He also joined the Department of Electrical and Communications Engineering, University of Blumenau (FURB), Blumenau, Brazil, in 2003. He is currently teaching communication systems. His present research interests include statistical analysis of adaptive algorithms, active noise and vibrations control, and image processing.

Rui Seara (M'93–SM'04) was born in Florianópolis, SC, Brazil. He received the B.E. and M.Sc. degrees in electrical engineering from the Federal University of Santa Catarina, Brazil, in 1975 and 1980, respectively, and the Doctoral degree in electrical engineering from the Paris-Sud University, Paris, France, in 1984.

In 1976, he joined the Electrical Engineering Department at the Federal University of Santa Catarina, Brazil, where he is currently a Professor of Electrical Engineering and Director of LINSE-Circuits and Signal Processing Laboratory. His research interests include digital and analog filtering, adaptive signal processing algorithms, image and speech processing, and digital communications.

Dennis R. Morgan (S'63–S'68–M'69–SM'92–LS'08) was born in Cincinnati, OH, on February 19, 1942. He received the B.S. degree from the University of Cincinnati, OH, in 1965 and the M.S. and Ph.D. degrees from Syracuse University, Syracuse, NY, in 1968 and 1970, respectively, all in electrical engineering.

From 1965 to 1984, he was with the General Electric Company, Electronics Laboratory, Syracuse, NY, specializing in the analysis and design of signal processing systems used in radar, sonar, and communications. He is now a Distinguished Member of Technical Staff with Bell Laboratories, Alcatel-Lucent (formerly Lucent Technologies, formerly AT&T), where he has been employed since 1984: from 1984 to 1990 he was with the Special Systems Analysis Department, Whippany NJ, where he was involved in the analysis and development of advanced signal processing techniques associated with communications, array processing, detection and estimation, and active noise control; from 1990 to 2002, he was with the Acoustics Research Department, Murray Hill, NJ, where he was engaged in research on adaptive signal processing techniques applied to electroacoustic systems, including adaptive microphones, echo cancellation, talker direction finders, and blind source separation; since 2002, he has been with the Wireless Research Laboratory, Wireless and Broadband Access Research Center, and Radio Access Research Domain, Murray Hill, NJ, where he is involved in research on adaptive signal processing applied to RF and optical communication systems. He has authored numerous journal publications and is coauthor of *Active Noise Control Systems: Algorithms and DSP Implementations* (New York: Wiley, 1996) and *Advances in Network and Acoustic Echo Cancellation* (New York: Springer-Verlag, 2001).

Dr. Morgan served as Associate Editor for the IEEE TRANSACTIONS ON SPEECH AND AUDIO PROCESSING from 1995 to 2000, and Associate Editor for the IEEE TRANSACTIONS ON SIGNAL PROCESSING from 2001 to 2004 and from 2008 to the present. Since 2004 he has been a member of the Signal Processing Theory & Methods Technical Committee of the IEEE Signal Processing Society.