Robust Adaptive Beamforming for General-Rank Signal Model With Positive Semi-Definite Constraint via POTDC

Arash Khabbazibasmenj, *Student Member, IEEE*, and Sergiy A. Vorobyov, *Senior Member, IEEE*

***Abstract*—The robust adaptive beamforming (RAB) problem for general-rank signal model with an additional positive semi-definite constraint is considered. Using the principle of the worst-case performance optimization, such RAB problem leads to a difference-of-convex functions (DC) optimization problem. The existing approaches for solving the resulted non-convex DC problem are based on approximations and find only suboptimal solutions. Here, we aim at finding the globally optimal solution for the non-convex DC problem and clarify the conditions under which the solution is guaranteed to be globally optimal. Particularly, we rewrite the problem as the minimization of a one-dimensional optimal value function (OVF). Then, the OVF is replaced with another equivalent one, for which the corresponding optimization problem is convex. The new one-dimensional OVF is minimized iteratively via polynomial time DC (POTDC) algorithm. We show that the POTDC converges to a point that satisfies Karush-Kuhn-Tucker (KKT) optimality conditions, and such point is the global optimum under certain conditions. Towards this conclusion, we prove that the proposed algorithm finds the globally optimal solution if the presumed norm of the mismatch matrix that corresponds to the desired signal covariance matrix is sufficiently small. The new RAB method shows superior performance compared to the other state-of-the-art general-rank RAB methods.**

***Index Terms*—Difference-of-convex functions (DC) programming, non-convex programming, semi-definite programming relaxation, robust adaptive beamforming, general-rank signal model, polynomial time DC (POTDC).**

I. Introduction

IT is well known that when the desired signal is present in the training data, the performance of adaptive beamforming methods degrades dramatically in the presence of even a very slight mismatch in the knowledge of the desired signal covariance matrix. The mismatch between the presumed and actual source covariance matrices occurs because of, for example, displacement of antenna elements, time varying environment, imperfections of propagation medium, etc. The main goal of any robust adaptive beamforming (RAB) technique is to provide robustness against any such mismatches.

Most of the RAB methods have been developed for the case of point source signals when the rank of the desired signal covariance matrix is equal to one [1]–[11]. Among the principles used for such RAB methods design are i) the worst-case performance optimization [2]–[6]; ii) probabilistic based performance optimization [8]; and iii) estimation of the actual steering vector of the desired signal [9]–[11]. In many practical applications such as, for example, the incoherently scattered signal source or source with fluctuating (randomly distorted) wavefront, the rank of the source covariance matrix is higher than one. Although the RAB methods of [1]–[11] provide excellent robustness against any mismatch of the underlying point source assumption, they are not perfectly suited to the case when the rank of the desired signal covariance matrix is higher than one.

The RAB for the general-rank signal model based on the explicit modeling of the error mismatches has been developed in [12] based on the worst-case performance optimization principle. Although the RAB of [12] has a simple closed form solution, it is overly conservative because the worst-case correlation matrix of the desired signal may be indefinite or even negative definite [13]–[15]. Thus, less conservative approaches have been developed in [13]–[15] by considering an additional positive semi-definite (PSD) constraint to the worst-case signal covariance matrix. The major shortcoming of the RAB methods of [13]–[15] is that they find only a suboptimal solution and there may be a significant gap to the global optimal solution. For example, the RAB of [13] finds a suboptimal solution in an iterative way, but there is no guarantee that such iterative method converges [15]. A closed-form approximate suboptimal solution is proposed in [14], however, this solution may be quite far from the globally optimal one as well. All these shortcomings motivate us to look for new efficient ways to solve the aforementioned non-convex problem globally optimally.[1]

We propose a new method that is based on recasting the original non-convex difference-of-convex functions (DC) programming problem as the minimization of a one dimensional optimal value function (OVF). Although the corresponding optimization problem of the newly introduced OVF is non-convex, it can be replaced with another equivalent problem. Such optimization problem is convex and can be solved efficiently. The new

Manuscript received December 11, 2012; revised May 18, 2013 and August 25, 2013; accepted August 29, 2013. Date of publication September 10, 2013; date of current version November 05, 2013. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Antonio De Maio. This work was supported in part by the Natural Science and Engineering Research Council (NSERC) of Canada. Parts of this paper were presented at CAMSAP San Juan, Puerto Rico, 2011.

A. Khabbazibasmenj is with the Department of Electrical and Computer Engineering, University of Alberta, Edmonton, AB T6G 2V4 Canada (e-mail: khabbazi@ualberta.ca).

S. A. Vorobyov is with the Department of Signal Processing and Acoustics, Aalto University, FI-00076 AALTO, Finland, on leave from the Department of Electrical and Computer Engineering, University of Alberta, Edmonton, AB T6G 2V4, Canada (e-mail: svor@ieee.org).

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.2013.2281301

[1]Some preliminary results have been presented in [16].

one-dimensional OVF is then minimized by the means of the polynomial time DC (POTDC) algorithm (see also [17], [18]). We prove that the point found by the POTDC algorithm for the RAB for general-rank signal model with positive semi-definite constraint is a Karush-Kuhn-Tucker (KKT) optimal point. Moreover, we prove a number of results that lead us to the equivalence between the claim of global optimality for the problem considered and the convexity or strict quasi-convexity of the newly obtained one-dimensional OVF. The global optimality of the proposed POTDC method is then proved under some conditions. As an additional check, we also develop a tight lower-bound for such OVF that is used in the simulations to further confirming global optimality.

The rest of the paper is organized as follows. System model and preliminaries are given in Section II, while the problem is formulated in Section III. The new proposed method is developed in Section IV followed by our simulation results in Section V. Finally, Section VI presents our conclusions. This paper is reproducible research, and the software needed to generate the simulation results can be obtained from the IEEE Xplore together with the paper.

II. System Model and Preliminaries

The narrowband signal received by a linear antenna array with M omni-directional antenna elements at the time instant k can be expressed as

$$\mathbf{x}(k) = \mathbf{s}(k) + \mathbf{i}(k) + \mathbf{n}(k) \tag{1}$$

where $\mathbf{s}(k), \mathbf{i}(k)$, and $\mathbf{n}(k)$ are the statistically independent $M \times 1$ vectors of the desired signal, interferences, and noise, respectively. The beamformer output at the time instant k is given as

$$y(k) = \mathbf{w}^H \mathbf{x}(k) \tag{2}$$

where $\mathbf{w}$ is the $M \times 1$ complex beamforming vector of the antenna array and $(\cdot)^H$ stands for the Hermitian transpose. The beamforming problem is formulated as finding the beamforming vector $\mathbf{w}$ which maximizes the beamformer output signal-to-interference-plus-noise ratio (SINR) given as

$$\mathrm{SINR} = \frac{\mathbf{w}^H \mathbf{R}_{\mathrm{s}} \mathbf{w}}{\mathbf{w}^H \mathbf{R}_{\mathrm{i+n}} \mathbf{w}} \tag{3}$$

where $\mathbf{R}_{\mathrm{s}} \triangleq E\{\mathbf{s}(k)\mathbf{s}^H(k)\}$ and $\mathbf{R}_{\mathrm{i+n}} \triangleq E\{(\mathbf{i}(k) + \mathbf{n}(k))(\mathbf{i}(k) + \mathbf{n}(k))^H\}$ are the desired signal and interference-plus-noise covariance matrices, respectively, and $E\{\cdot\}$ stands for the statistical expectation.

Depending on the nature of the desired signal source, its corresponding covariance matrix can be of an arbitrary rank, i.e., $1 \leq \mathrm{rank}\{\mathbf{R}_{\mathrm{s}}\} \leq M$, where rank $\{\cdot\}$ denotes the rank operator. Indeed, in many practical applications, for example, in the scenarios with incoherently scattered signal sources or signals with randomly fluctuating wavefronts, the rank of the desired signal covariance matrix $\mathbf{R}_{\mathrm{s}}$ is greater than one [12]. The only particular case in which, the rank of $\mathbf{R}_{\mathrm{s}}$ is equal to one is the case of the point source.

The interference-plus-noise covariance matrix $\mathbf{R}_{\mathrm{i+n}}$ is typically unavailable in practice and it is substituted by the data sample covariance matrix

$$\hat{\mathbf{R}} = \frac{1}{K} \sum_{k=1}^{K} \mathbf{x}(k)\mathbf{x}^H(k) \tag{4}$$

where K is number of the training data samples. The problem of maximizing the SINR (3) (here we always use sample matrix estimate $\hat{\mathbf{R}}$ instead of $\mathbf{R}_{\mathrm{i+n}}$) is known as minimum variance distortionless response (MVDR) beamforming and can be mathematically formulated as

$$\min_{\mathbf{w}} \mathbf{w}^H \hat{\mathbf{R}} \mathbf{w} \quad \text{s.t.} \quad \mathbf{w}^H \mathbf{R}_{\mathrm{s}} \mathbf{w} = 1. \tag{5}$$

The solution to the MVDR beamforming problem (5) can be found as [1]

$$\mathbf{w}_{\mathrm{SMI-MVDR}} = \boldsymbol{\mathcal{P}}\{\hat{\mathbf{R}}^{-1}\mathbf{R}_{\mathrm{s}}\} \tag{6}$$

which is known as the sample matrix inversion (SMI) MVDR beamformer for general-rank signal model. Here $\boldsymbol{\mathcal{P}}\{\cdot\}$ stands for the principal eigenvector operator.

In practice, the actual desired signal covariance matrix $\mathbf{R}_{\mathrm{s}}$ is usually unknown and only its presumed value is available. The actual source correlation matrix can be modeled as $\mathbf{R}_{\mathrm{s}} = \tilde{\mathbf{R}}_{\mathrm{s}} + \boldsymbol{\Delta}_1$, where $\boldsymbol{\Delta}_1$ and $\tilde{\mathbf{R}}_{\mathrm{s}}$ denote an unknown mismatch and the presumed correlation matrices, respectively. It is well known that the MVDR beamformer is very sensitive to such mismatches [12]. RABs also address the situation when the sample estimate of the data covariance matrix (4) is inaccurate (for example, because of small sample size) and $\mathbf{R} = \hat{\mathbf{R}} + \boldsymbol{\Delta}_2$, where $\boldsymbol{\Delta}_2$ is an unknown mismatch matrix to the data sample covariance matrix. In order to provide robustness against the norm-bounded mismatches $\|\boldsymbol{\Delta}_1\| \leq \epsilon$ and $\|\boldsymbol{\Delta}_2\| \leq \gamma$ (here $\|\cdot\|$ denotes the Frobenius norm of a matrix), the RAB of [12] uses the worst-case performance optimization principle of [2] and finds the solution as

$$\mathbf{w} = \boldsymbol{\mathcal{P}}\{(\hat{\mathbf{R}} + \gamma\mathbf{I})^{-1}(\tilde{\mathbf{R}}_{\mathrm{s}} - \epsilon\mathbf{I})\}. \tag{7}$$

Although the RAB of (7) has a simple closed-form expression, it is overly conservative because the constraint that the matrix $\tilde{\mathbf{R}}_{\mathrm{s}} + \boldsymbol{\Delta}_1$ has to be PSD is not considered [13]. For example, the worst-case desired signal covariance matrix $\tilde{\mathbf{R}}_{\mathrm{s}} - \epsilon\mathbf{I}$ in (7) can be indefinite or negative definite if $\tilde{\mathbf{R}}_{\mathrm{s}}$ is rank deficient. Indeed, in the case of incoherently scattered source, $\tilde{\mathbf{R}}_{\mathrm{s}}$ has the following form $\tilde{\mathbf{R}}_{\mathrm{s}} = \sigma_{\mathrm{s}}^2 \int_{-\pi/2}^{\pi/2} \zeta(\theta)\mathbf{a}(\theta)\mathbf{a}^H(\theta)d\theta$, where $\zeta(\theta)$ denotes the normalized angular power density, σ_{s}^2 is the desired signal power, and $\mathbf{a}(\theta)$ is the steering vector towards direction θ. For a uniform angular power density on the angular bandwidth Φ, the approximate numerical rank of $\tilde{\mathbf{R}}_{\mathrm{s}}$ is equal to $(\Phi/\pi) \cdot M$ [19]. This leads to a rank deficient matrix $\tilde{\mathbf{R}}_{\mathrm{s}}$ if the angular power density does not cover all the directions. Therefore, the worst-case covariance matrix $\tilde{\mathbf{R}}_{\mathrm{s}} - \epsilon\mathbf{I}$ is indefinite or negative definite. Note that the worst-case data sample covariance matrix $\hat{\mathbf{R}} + \gamma\mathbf{I}$ is always positive definite.

III. Problem Formulation

Decomposing $\mathbf{R}_s$ as $\mathbf{R}_s = \mathbf{Q}^H\mathbf{Q}$, the RAB problem for a norm-bounded mismatch $\|\boldsymbol{\Delta}\| \leq \eta$ to the matrix $\mathbf{Q}$ is given as [13]

$$\begin{aligned}\min_{\mathbf{w}} \quad & \max_{\|\boldsymbol{\Delta}_2\|\leq\gamma} \mathbf{w}^H(\hat{\mathbf{R}}+\boldsymbol{\Delta}_2)\mathbf{w}\\ \text{s.t.} \quad & \min_{\|\boldsymbol{\Delta}\|\leq\eta} \mathbf{w}^H(\mathbf{Q}+\boldsymbol{\Delta})^H(\mathbf{Q}+\boldsymbol{\Delta})\mathbf{w} \geq 1.\end{aligned} \tag{8}$$

For every $\boldsymbol{\Delta}$ in the optimization problem (8) whose norm is less than or equal to η, the expression $\mathbf{w}^H(\mathbf{Q}+\boldsymbol{\Delta})^H(\mathbf{Q}+\boldsymbol{\Delta})\mathbf{w} \geq 1$ represents a non-convex quadratic constraint with respect to $\mathbf{w}$. Because there exists infinite number of mismatches $\boldsymbol{\Delta}$, there also exists infinite number of such non-convex quadratic constraints. By finding the minimum possible value of the quadratic term $\mathbf{w}^H(\mathbf{Q}+\boldsymbol{\Delta})^H(\mathbf{Q}+\boldsymbol{\Delta})\mathbf{w}$ with respect to $\boldsymbol{\Delta}$ for a fixed $\mathbf{w}$, the infinite number of such constraints can be replaced with a single constraint. Hence, we consider the following optimization problem

$$\begin{aligned}\min_{\boldsymbol{\Delta}} \quad & \mathbf{w}^H(\mathbf{Q}+\boldsymbol{\Delta})^H(\mathbf{Q}+\boldsymbol{\Delta})\mathbf{w}\\ \text{s.t.} \quad & \|\boldsymbol{\Delta}\|^2 \leq \eta^2.\end{aligned} \tag{9}$$

This problem is convex and its optimal value can be expressed as a function of $\mathbf{w}$ as given by the following lemma.

Lemma 1: The optimal value of the optimization problem (9) as a function of $\mathbf{w}$ is equal to

$$\begin{aligned}&\min_{\|\boldsymbol{\Delta}\|^2\leq\eta^2} \mathbf{w}^H(\mathbf{Q}+\boldsymbol{\Delta})^H(\mathbf{Q}+\boldsymbol{\Delta})\mathbf{w}\\ &\quad = \begin{cases}(\|\mathbf{Q}\mathbf{w}\|-\eta\|\mathbf{w}\|)^2, & \|\mathbf{Q}\mathbf{w}\|\geq\eta\|\mathbf{w}\|\\ 0, & \text{otherwise.}\end{cases}\end{aligned} \tag{10}$$

Proof: See Appendix I-A. ■

It follows from (10) that the desired signal can be totally removed from the beamformer output if $\|\mathbf{Q}\mathbf{w}\| < \eta\|\mathbf{w}\|$. Based on Lemma 1, the constraint in (8) can be equivalently replaced by the constraint

$$\|\mathbf{Q}\mathbf{w}\| - \eta\|\mathbf{w}\| \geq 1. \tag{11}$$

Moreover, the maximum of the quadratic term $\mathbf{w}^H(\hat{\mathbf{R}}+\boldsymbol{\Delta}_2)\mathbf{w}$ in the objective function of the problem in (8) with respect to $\boldsymbol{\Delta}_2$, $\|\boldsymbol{\Delta}_2\|\leq\gamma$ can be easily derived as $\mathbf{w}^H(\hat{\mathbf{R}}+\gamma\mathbf{I})\mathbf{w}$. Therefore, the RAB problem (8) can be equivalently written in a simpler form as

$$\begin{aligned}&\min_{\mathbf{w}} \mathbf{w}^H(\hat{\mathbf{R}}+\gamma\mathbf{I})\mathbf{w}\\ &\text{s.t.}\ \|\mathbf{Q}\mathbf{w}\| - \eta\|\mathbf{w}\| \geq 1.\end{aligned} \tag{12}$$

Due to the non-convex DC constraint, the problem (12) is a non-convex DC programming problem [17], [18]. DC optimization problems are believed to be NP-hard in general [20], [21]. There is a number of methods that can be applied to address DC programming problems of type (12). Among these methods are the generalized poly block algorithm, the extended general power iterative (GPI) algorithm [22], DC iteration-based method [23], etc. However, the existing methods do not guarantee to find the globally optimal solution of a DC programming problem in polynomial time.

Recently, the problem (12) has also been suboptimally solved using an iterative semi-definite relaxation (SDR)-based algorithm in [13] which also does not result in the globally optimal solution and for which the convergence even to a KKT optimal point is not guaranteed. A closed-form suboptimal solution for the aforementioned non-convex DC problem has been also derived in [14]. Despite its computational simplicity, the performance of the method of [14] may be far from the global optimum and even the KKT optimal point. Another iterative algorithm has been proposed in [15], but it modifies the problem (12) and solves the modified problem instead, which again gives no guarantees of finding the globally optimal solution of the original problem (12).

IV. New Proposed Method

A. Main Idea and OVF

Here, we aim at solving the problem (12) globally optimally in polynomial time. For this goal, we design a POTDC-type algorithm (see also [17], [18]) that can be used for solving a class of DC programming problems in polynomial time. By introducing the auxiliary optimization variable $\alpha \geq 1$ and setting $\|\mathbf{Q}\mathbf{w}\| = \sqrt{\alpha}$, the problem (12) can be equivalently rewritten as

$$\begin{aligned}\min_{\mathbf{w},\alpha} \quad & \mathbf{w}^H(\hat{\mathbf{R}}+\gamma\mathbf{I})\mathbf{w}\\ \text{s.t.} \quad & \mathbf{w}^H\mathbf{Q}^H\mathbf{Q}\mathbf{w} = \alpha\\ & \mathbf{w}^H\mathbf{w} \leq \frac{(\sqrt{\alpha}-1)^2}{\eta^2}, \quad \alpha \geq 1.\end{aligned} \tag{13}$$

Note that α is restricted to be greater than or equal to one because $\|\mathbf{Q}\mathbf{w}\|$ is greater than or equal to one due to the constraint of the problem (12). For future needs, we find the set of all α's for which the problem (13) is feasible. Let us define the following set for a fixed value of $\alpha \geq 1$

$$S(\alpha) \triangleq \{\mathbf{w} \mid \mathbf{w}^H\mathbf{w} \leq (\sqrt{\alpha}-1)^2/\eta^2\}. \tag{14}$$

It is trivial that for every $\mathbf{w} \in S(\alpha)$, the quadratic term $\mathbf{w}^H\mathbf{Q}^H\mathbf{Q}\mathbf{w}$ is non-negative as $\mathbf{Q}^H\mathbf{Q}$ is a positive semi-definite matrix. Using the minimax theorem [24], it can be easily verified that the maximum value of the quadratic term $\mathbf{w}^H\mathbf{Q}^H\mathbf{Q}\mathbf{w}$ over $\mathbf{w} \in S(\alpha)$ is equal to $\left((\sqrt{\alpha}-1)^2/\eta^2\right)\lambda_{\max}\{\mathbf{Q}^H\mathbf{Q}\}$ and this value is achieved by

$$\mathbf{w}_\alpha = \frac{\sqrt{\alpha}-1}{\eta}\boldsymbol{\mathcal{P}}\{\mathbf{Q}^H\mathbf{Q}\} \in S(\alpha). \tag{15}$$

Here $\lambda_{\max}\{\cdot\}$ stands for the largest eigenvalue operator. Due to the fact that for any $0 \leq \beta \leq 1$, the scaled vector $\beta\mathbf{w}_\alpha$ lies inside the set $S(\alpha)$, the quadratic term $\mathbf{w}^H\mathbf{Q}^H\mathbf{Q}\mathbf{w}$ can take values only in the interval $[0, \left((\sqrt{\alpha}-1)^2/\eta^2\right)\lambda_{\max}\{\mathbf{Q}^H\mathbf{Q}\}]$ over $\mathbf{w} \in S(\alpha)$.

Considering the later fact and also the optimization problem (13), it can be concluded that α is feasible if and only if $\alpha \in [0, \left((\sqrt{\alpha}-1)^2/\eta^2\right)\cdot\lambda_{\max}\{\mathbf{Q}^H\mathbf{Q}\}]$ which implies that

$$\frac{(\sqrt{\alpha}-1)^2}{\eta^2}\lambda_{\max}\{\mathbf{Q}^H\mathbf{Q}\} \geq \alpha \tag{16}$$

or, equivalently, that

$$\frac{(\sqrt{\alpha}-1)^2}{\alpha} \geq \frac{\eta^2}{\lambda_{\max}\{\mathbf{Q}^H\mathbf{Q}\}}. \tag{17}$$

The function $(\sqrt{\alpha}-1)^2/\alpha$ is strictly increasing and it is also less than or equal to one for $\alpha \geq 1$. Therefore, it can be immediately found that the problem (13) is infeasible for any $\alpha \geq 1$ if $\lambda_{\max}\{\mathbf{Q}^H\mathbf{Q}\} \leq \eta^2$. Thus, hereafter, it is assumed that $\lambda_{\max}\{\mathbf{Q}^H\mathbf{Q}\} > \eta^2$. Moreover, using (17) and the fact that the function $(\sqrt{\alpha}-1)^2/\alpha$ is strictly increasing, it can be found that the feasible set of the problem (13) corresponds to

$$\alpha \geq \frac{1}{\left(1-\frac{\eta}{\sqrt{\lambda_{\max}\{\mathbf{Q}^H\mathbf{Q}\}}}\right)^2} \geq 1. \tag{18}$$

As we will see in the following sections, for developing the POTDC algorithm for the problem (13), an upper-bound for the optimal value of α in (13) is needed. Such upper-bound is obtained in terms of the following lemma.

Lemma 2: The optimal value of the optimization variable α in the problem (13) is upper-bounded by $\lambda_{\max}\left\{(\hat{\mathbf{R}}+\gamma\mathbf{I})^{-1}\mathbf{Q}^H\mathbf{Q}\right\}\mathbf{w}_0^H(\hat{\mathbf{R}}+\gamma\mathbf{I})\mathbf{w}_0$, where $\mathbf{w}_0$ is any arbitrary feasible point of the problem (13).

Proof: See Appendix I-B. ■

Using Lemma 2, the problem (13) can be equivalently stated as

$$\begin{aligned} \min_{\theta_1\leq\alpha\leq\theta_2} \quad & \overbrace{\min_{\mathbf{w}} \mathbf{w}^H(\hat{\mathbf{R}}+\gamma\mathbf{I})\mathbf{w}}^{\text{InnerProblem}} \\ \text{s.t.} \quad & \mathbf{w}^H\mathbf{Q}^H\mathbf{Q}\mathbf{w} = \alpha \\ & \mathbf{w}^H\mathbf{w} \leq \frac{(\sqrt{\alpha}-1)^2}{\eta^2} \end{aligned} \tag{19}$$

where

$$\theta_1 = \frac{1}{\left(1-\frac{\eta}{\sqrt{\lambda_{\max}\{\mathbf{Q}^H\mathbf{Q}\}}}\right)^2} \tag{20}$$

and

$$\theta_2 = \lambda_{\max}\left\{(\hat{\mathbf{R}}+\gamma\mathbf{I})^{-1}\mathbf{Q}^H\mathbf{Q}\right\}\mathbf{w}_0^H(\hat{\mathbf{R}}+\gamma\mathbf{I})\mathbf{w}_0. \tag{21}$$

For a fixed value of α, the inner optimization problem in (19) is non-convex with respect to $\mathbf{w}$. Based on the inner optimization problem in (19) when α is fixed, we define the following OVF

$$h(\alpha) \triangleq \left\{\min_{\mathbf{w}} \mathbf{w}^H(\hat{\mathbf{R}}+\gamma\mathbf{I})\mathbf{w} \mid \mathbf{w}^H\mathbf{Q}^H\mathbf{Q}\mathbf{w}=\alpha,\right.$$
$$\left.\mathbf{w}^H\mathbf{w} \leq \frac{(\sqrt{\alpha}-1)^2}{\eta^2}\right\}, \quad \theta_1 \leq \alpha \leq \theta_2. \tag{22}$$

Using the OVF (22), the problem (19) can be equivalently expressed as

$$\min_{\alpha} \quad h(\alpha) \quad \text{s.t.} \quad \theta_1 \leq \alpha \leq \theta_2. \tag{23}$$

The corresponding optimization problem of $h(\alpha)$ for a fixed value of α is non-convex. In what follows, we aim at replacing $h(\alpha)$ with an equivalent OVF whose corresponding optimization problem is convex.

Introducing the matrix $\mathbf{W} \triangleq \mathbf{w}\mathbf{w}^H$ and using the fact that for any arbitrary matrix $\mathbf{A}$, $\mathbf{w}^H\mathbf{A}\mathbf{w} = \text{tr}\{\mathbf{A}\mathbf{w}\mathbf{w}^H\}$ (here $\text{tr}\{\cdot\}$ stands for the trace of a matrix), the OVF (22) can be equivalently recast as

$$h(\alpha) = \left\{\min_{\mathbf{W}} \text{tr}\{(\hat{\mathbf{R}}+\gamma\mathbf{I})\mathbf{W}\} \mid \text{tr}\{\mathbf{Q}^H\mathbf{Q}\mathbf{W}\}=\alpha,\right.$$
$$\left.\text{tr}\{\mathbf{W}\} \leq \frac{(\sqrt{\alpha}-1)^2}{\eta^2},\ \mathbf{W}\succeq\mathbf{0},\ \text{rank}\{\mathbf{W}\}=1\right\},$$
$$\theta_1 \leq \alpha \leq \theta_2. \tag{24}$$

By dropping the rank-one constraint in the corresponding optimization problem of $h(\alpha)$ for a fixed value of α, $(\theta_1 \leq \alpha \leq \theta_2)$, a new OVF denoted as $k(\alpha)$ can be defined as

$$k(\alpha) \triangleq \left\{\min_{\mathbf{W}} \text{tr}\{(\hat{\mathbf{R}}+\gamma\mathbf{I})\mathbf{W}\} \mid \text{tr}\{\mathbf{Q}^H\mathbf{Q}\mathbf{W}\}=\alpha,\right.$$
$$\left.\text{tr}\{\mathbf{W}\} \leq \frac{(\sqrt{\alpha}-1)^2}{\eta^2},\ \mathbf{W}\succeq\mathbf{0}\right\},$$
$$\theta_1 \leq \alpha \leq \theta_2. \tag{25}$$

For brevity, we will refer to the optimization problems that correspond to the OVFs $h(\alpha)$ and $k(\alpha)$ when α is fixed, as the optimization problems of $h(\alpha)$ and $k(\alpha)$, respectively. Note also that compared to the optimization problem of $h(\alpha)$, the optimization problem of $k(\alpha)$ is convex. Moreover, it is easy to check that the optimization problem of $h(\alpha)$ is a hidden convex problem, i.e., the duality gap between this problem and its dual is zero [25]–[28]. Since both of the optimization problems of $h(\alpha)$ and $k(\alpha)$ have the same dual problem, it can be immediately concluded that the OVFs $h(\alpha)$ and $k(\alpha)$ are equivalent, i.e., $h(\alpha)=k(\alpha)$ for any $\alpha\in[\theta_1,\theta_2]$. Furthermore, based on the optimal solution of the optimization problem of $k(\alpha)$ when α is fixed, the optimal solution of the optimization problem of $h(\alpha)$ can be constructed [25]–[28]. Based on the later fact, the original problem (23) can be expressed as

$$\min_{\alpha} \quad k(\alpha) \quad \text{s.t.} \quad \theta_1 \leq \alpha \leq \theta_2. \tag{26}$$

It is noteworthy to mention that based on the optimal solution of (26) denoted as α_{opt}, we can easily obtain the optimal solution of the original problem (23) or, equivalently, the optimal solution of the problem (19). Specifically, since the OVFs $h(\alpha)$ and $k(\alpha)$ are equivalent, α_{opt} is also the optimal solution of the problem (23) and, thus, also the problem (19). Moreover, the optimization problem of $k(\alpha_{\text{opt}})$ is convex and can be easily solved. In addition, using the results in [25]–[28] and based on the optimal solution of the optimization problem of $k(\alpha_{\text{opt}})$, the optimal solution of the optimization problem of $h(\alpha_{\text{opt}})$ can be constructed. Thus, we concentrate on the problem (26).

Since for every fixed value of α, the corresponding optimization problem of $k(\alpha)$ is a convex semi-definite programming (SDP) problem, one possible approach for solving (26) is based on exhaustive search over α. In other words, α can be found by using an exhaustive search over a fine grid on the interval of $[\theta_1, \theta_2]$. Although this search method is inefficient, it can be used as a benchmark.

Using the definition of the OVF $k(\alpha)$, the problem (26) can be equivalently expressed as

$$\begin{aligned} \min_{\mathbf{W},\alpha} \quad & \mathrm{tr}\left\{(\hat{\mathbf{R}}+\gamma\mathbf{I})\mathbf{W}\right\} \\ \text{s.t.} \quad & \mathrm{tr}\{\mathbf{Q}^H\mathbf{Q}\mathbf{W}\} = \alpha \\ & \eta^2\mathrm{tr}\{\mathbf{W}\} \leq (\sqrt{\alpha}-1)^2 \\ & \mathbf{W} \succeq 0, \ \theta_1 \leq \alpha \leq \theta_2. \end{aligned} \tag{27}$$

Note that replacing $h(\alpha)$ by $k(\alpha)$ results in a much simpler problem. Indeed, compared to the original problem (19), in which the first constraint is non-convex, the corresponding first constraint of (27) is convex. All the constraints and the objective function of the problem (27) are convex except for the constraint $\mathrm{tr}\{\mathbf{W}\} \leq (\sqrt{\alpha}-1)^2/\eta^2$ which is non-convex only in a single variable α. It makes the problem (27) non-convex overall. This single non-convex constraint can be rewritten as $\eta^2\mathrm{tr}\{\mathbf{W}\} - (\alpha+1) + 2\sqrt{\alpha} \leq 0$ where all the terms are linear with respect to $\mathbf{W}$ and α except for the concave term of $\sqrt{\alpha}$. The latter constraint can be handled iteratively by building a POTDC-type algorithm (see also [17], [18]) based on the iterative linear approximation of the non-convex term $\sqrt{\alpha}$ around suitably selected points. It is interesting to mention that this iterative linear approximation can be also interpreted in terms of the DC-iteration approach over the single non-convex term $\sqrt{\alpha}$. The fact that iterations are needed only over a single variable helps to reduce dramatically the number of iterations as compared to the traditional DC-iteration approach and allows for simple algorithm shown below.

B. Iterative POTDC Algorithm

Let us consider the optimization problem (27) and replace the term $\sqrt{\alpha}$ by its linear approximation around α_c, i.e., $\sqrt{\alpha} \approx \sqrt{\alpha_c} + (\alpha - \alpha_c)/(2\sqrt{\alpha_c})$. It leads to the following SDP problem

$$\begin{aligned} \min_{\mathbf{W},\alpha} \quad & \mathrm{tr}\left\{(\hat{\mathbf{R}}+\gamma\mathbf{I})\mathbf{W}\right\} \\ \text{s.t.} \quad & \mathrm{tr}\{\mathbf{Q}^H\mathbf{Q}\mathbf{W}\} = \alpha \\ & \eta^2\mathrm{tr}\{\mathbf{W}\} + (\sqrt{\alpha_c}-1) + \alpha\left(\frac{1}{\sqrt{\alpha_c}} - 1\right) \leq 0 \\ & \mathbf{W} \succeq 0, \ \ \theta_1 \leq \alpha \leq \theta_2. \end{aligned} \tag{28}$$

To demonstrate the POTDC algorithm graphically and also to see how the linearization points are selected in different iterations, let us define the following OVF based on the optimization problem (28)

$$\begin{aligned} l(\alpha,\alpha_c) \triangleq \Big\{ & \min_{\mathbf{W}} \mathrm{tr}\{(\hat{\mathbf{R}}+\gamma\mathbf{I})\mathbf{W}\} \mid \mathrm{tr}\{\mathbf{Q}^H\mathbf{Q}\mathbf{W}\} = \alpha, \\ & \eta^2\mathrm{tr}\{\mathbf{W}\} + (\sqrt{\alpha_c}-1) + \alpha\left(\frac{1}{\sqrt{\alpha_c}} - 1\right) \leq 0, \\ & \mathbf{W} \succeq \mathbf{0}\Big\}, \ \theta_1 \leq \alpha \leq \theta_2 \end{aligned} \tag{29}$$

where α_c in $l(\alpha,\alpha_c)$ denotes the linearization point. The OVF $l(\alpha,\alpha_c)$ can be also obtained through $k(\alpha)$ in (25) by replacing the term $\sqrt{\alpha}$ in $\eta^2\mathrm{tr}\{\mathbf{W}\} - (\alpha+1) + 2\sqrt{\alpha} \leq 0$ with its linear approximation around α_c. Since $\sqrt{\alpha}$ and its linear approximation have the same values at α_c, $l(\alpha,\alpha_c)$ and $k(\alpha)$ take the same values at this point. The following lemma establishes the relationship between the OVFs $k(\alpha)$ and $l(\alpha,\alpha_c)$.

Lemma 3: The OVF $l(\alpha,\alpha_c)$ is a convex upper-bound of $k(\alpha)$ for any arbitrary $\alpha_c \in [\theta_1,\theta_2]$, i.e., $l(\alpha,\alpha_c) \geq k(\alpha), \forall \alpha \in [\theta_1,\theta_2]$ and $l(\alpha,\alpha_c)$ is convex with respect to α. Furthermore, the OVFs $k(\alpha)$ and $l(\alpha,\alpha_c)$ are directionally differentiable at the point $\alpha = \alpha_c$ and the values of these OVFs as well as their right and left derivatives are equal at $\alpha = \alpha_c$. In other words, under the condition that $k(\alpha)$ is differentiable at α_c, $l(\alpha,\alpha_c)$ is tangent to $k(\alpha)$ at the point $\alpha = \alpha_c$.

Proof: See Appendix I-C. ■

In what follows, we explain intuitively how the proposed POTDC method works. For the sake of clarity, it is assumed in this explanation only that the OVF $k(\alpha)$ is differentiable over the interval (θ_1,θ_2), however, similar interpretation can be made generally even for non-differentiable $k(\alpha)$. Moreover, as we will see later, the differentiability of the OVF $k(\alpha)$ is not needed for establishing the optimality results for the POTDC method.

Let us consider an arbitrary point, denoted as $\alpha_0 \in (\theta_1,\theta_2)$, as an initial linearization point, i.e., $\alpha_c = \alpha_0$. Based on Lemma 3, $l(\alpha,\alpha_0)$ is a convex function with respect to α which is the tangent to $k(\alpha)$ at the linearization point $\alpha = \alpha_0$, and it is also an upper-bound to $k(\alpha)$. Let α_1 denote the global minimizer of $l(\alpha,\alpha_0)$ that can be easily obtained due to the convexity of $l(\alpha,\alpha_0)$ with polynomial time complexity.

Since $l(\alpha,\alpha_0)$ is the tangent to $k(\alpha)$ at $\alpha = \alpha_0$ and it is also an upper-bound for $k(\alpha)$, it can be concluded that α_1 is a descent point for $k(\alpha)$, i.e., $k(\alpha_1) \leq k(\alpha_0)$ as it is shown in Fig. 1. Specifically, the fact that $l(\alpha,\alpha_0)$ is the tangent to $k(\alpha)$ at $\alpha = \alpha_0$ and α_1 is the global minimizer of $l(\alpha,\alpha_0)$ implies that

$$l(\alpha_1,\alpha_0) \leq l(\alpha_0,\alpha_0) = k(\alpha_0). \tag{30}$$

Furthermore, since $l(\alpha,\alpha_0)$ is an upper-bound for $k(\alpha)$, $k(\alpha_1) \leq l(\alpha_1,\alpha_0)$. Due to the later fact and also the (30), it is concluded that $k(\alpha_1) \leq k(\alpha_0)$.

Choosing α_1 as the linearization point in the second iteration, and finding the global minimizer of $l(\alpha,\alpha_1)$ over the interval $[\theta_1,\theta_2]$ denoted as α_2, another descent point can be obtained, i.e., $k(\alpha_2) \leq k(\alpha_1)$. This process is continued until convergence.

The iterative descent method can be described as shown in Algorithm 1. The following lemma about the convergence of Algorithm 1 and the optimality of the solution obtained by this algorithm is in order. Note that this lemma makes no assumptions about the differentiability of the OVF $k(\alpha)$.

Lemma 4: The following statements regarding Algorithm 1 are true:

i) The optimal value of the optimization problem in Algorithm 1 is non-increasing over iterations, i.e.,

$$\mathrm{tr}\left\{(\hat{\mathbf{R}}+\gamma\mathbf{I})\mathbf{W}_{\mathrm{opt,i+1}}\right\} \leq \mathrm{tr}\left\{(\hat{\mathbf{R}}+\gamma\mathbf{I})\mathbf{W}_{\mathrm{opt,i}}\right\}, \\ i \geq 1.$$

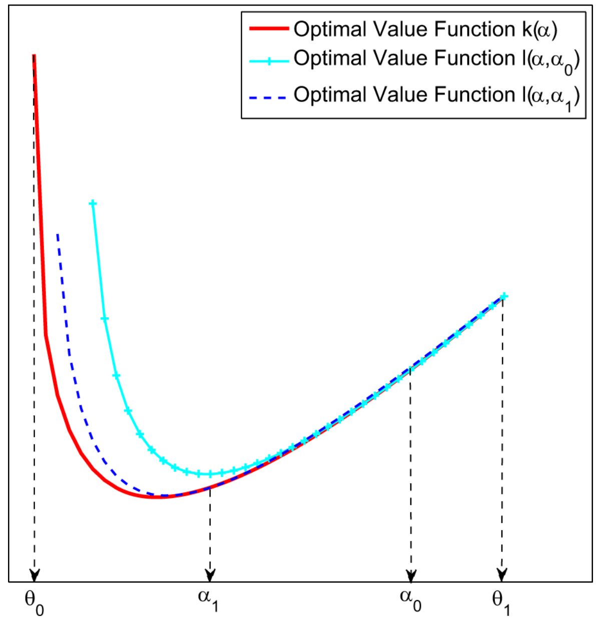

Fig. 1. Iterative method for minimizing the OVF $k(\alpha)$. The convex OVF $l(\alpha,\alpha_0)$ is the upper-bound to $k(\alpha)$ which is tangent to it at $\alpha=\alpha_0$, and its minimum is denoted as α_1. The point α_1 is used to establish another convex upper-bound function denoted as $l(\alpha,\alpha_1)$ and this process continues.

ii) The sequence of the optimal values in Algorithm 1 converges. Note that the termination condition is not considered for this statement.

iii) If Algorithm 1 converges (without considering termination condition), such a limiting point is regular and it satisfies the KKT optimality conditions.

Proof: See Appendix I-D. ■

Algorithm 1: The iterative POTDC algorithm

Require: An arbitrary $\alpha_c=\alpha_0\in[\theta_1,\theta_2]$, the termination threshold ζ, set i equal to 1.

repeat

Solve the following optimization problem using α_c to obtain $\mathbf{W}_{\text{opt}}$ and α_{opt}

$$\begin{aligned}\min_{\mathbf{W},\alpha}\quad & \text{tr}\left\{(\hat{\mathbf{R}}+\gamma\mathbf{I})\mathbf{W}\right\}\\ \text{s.t.}\quad & \text{tr}\{\mathbf{Q}^H\mathbf{Q}\mathbf{W}\}=\alpha\\ & \eta^2\text{tr}\{\mathbf{W}\}+(\sqrt{\alpha_c}-1)+\alpha\left(\frac{1}{\sqrt{\alpha_c}}-1\right)\leq 0\\ & \mathbf{W}\succeq 0,\ \theta_1\leq\alpha\leq\theta_2\end{aligned}$$

and set

$$\mathbf{W}_{\text{opt,i}}\leftarrow\mathbf{W}_{\text{opt}},\ \alpha_{\text{opt,i}}\leftarrow\alpha_{\text{opt}}$$
$$\alpha_c\leftarrow\alpha_{\text{opt}},\ i\leftarrow i+1$$

until

$\text{tr}\left\{(\hat{\mathbf{R}}+\gamma\mathbf{I})\mathbf{W}_{\text{opt,i-1}}\right\}-\text{tr}\left\{(\hat{\mathbf{R}}+\gamma\mathbf{I})\mathbf{W}_{\text{opt,i}}\right\}\leq\zeta$ for $i\geq 2$.

Note that the termination condition in Algorithm 1, i.e., $\text{tr}\left\{(\hat{\mathbf{R}}+\gamma\mathbf{I})\mathbf{W}_{\text{opt,i-1}}\right\}-\text{tr}\left\{(\hat{\mathbf{R}}+\gamma\mathbf{I})\mathbf{W}_{\text{opt,i}}\right\}\leq\zeta, i\geq 2$ is used for stopping the algorithm when the value achieved is deemed to be close enough to the optimal solution. The fact that the sequence of optimal values generated by Algorithm 1 is non-increasing and convergent has been used for choosing the termination condition. Despite its simplicity, this termination condition may stop the iterative algorithm prematurely. In order to avoid this situation, one can define the termination condition based on the approximate satisfaction of the KKT optimality conditions.

The point obtained by Algorithm 1 is guaranteed to be the global optimum of the problem considered if the OVF $k(\alpha)$ is a convex function of α. It is also worth noting that even a more relaxed property of the OVF $k(\alpha)$ is sufficient to guarantee global optimality. Specifically, if $k(\alpha)$ defined in (25) is a strictly quasi-convex function of $\alpha\in[\theta_1,\theta_2]$, then it is still guaranteed that we find the global optimum of the optimization problem (12) [29].

The worst-case computational complexity of a general standard SDP problem can be expressed as $\mathcal{O}(n_c^2\, n_v^{2.5}+n_c\, n_v^{3.5})$, where n_c and n_v denote, respectively, the number of constraints and the number of variables of the standard SDP problem [30] and $\mathcal{O}(\cdot)$ stands for the big-O (the highest order of complexity). The total number of variables in the SDP problem in Algorithm 1, which includes the real and imaginary parts of $\mathbf{W}$ and the real variable α, is equal to M^2+1. The computational complexity of Algorithm 1 is equal to that of the SDP optimization problem in Algorithm 1, that is, thus $\mathcal{O}(M^7)$, times the number of iterations (see also Simulation Example 1 in the next section).

The RAB algorithm of [13] is iterative as well and its computational complexity is equal to $\mathcal{O}(M^7)$ times the number of iterations. The complexity of the RABs of [12] and [14] is $\mathcal{O}(M^3)$. The comparison of the overall complexity of the proposed POTDC algorithm with that of the DC iteration-based method is also performed in Simulation Example 4 in the next section. Although the computational complexity of the new proposed method may be slightly higher than that of some other RABs, it finds the globally optimal solution as it is shown in Section IV-C. Moreover, it results in a superior performance as it is shown in Section V. Thus, next we show that under certain conditions the proposed POTDC method is guaranteed to find the globally optimal solution of a reformulated optimization problem that corresponds to the general-rank RAB problem.

C. Global Optimality

For studying the conditions under which the proposed POTDC method is guaranteed to find the global optimum of (12), we consider a reformulation of (12). Specifically, since this problem is feasible, it can be equivalently expressed as

$$\begin{aligned}\max_{\mathbf{w}}\quad & \frac{\|\mathbf{Q}\mathbf{w}\|-\eta\|\mathbf{w}\|}{\sqrt{\mathbf{w}^H(\hat{\mathbf{R}}+\gamma\mathbf{I})\mathbf{w}}}\\ \text{s.t.}\quad & \|\mathbf{Q}\mathbf{w}\|-\eta\|\mathbf{w}\|>0.\end{aligned}\tag{31}$$

Note that the constraint $\|\mathbf{Q}\mathbf{w}\|-\eta\|\mathbf{w}\|>0$ can be dropped as maximizing the objective function in (31) implies that this constraint is satisfied at the optimal point. By dropping this constraint, the problem (31) can be further expressed as the following homogenous problem

$$\max_{\mathbf{w}} \quad \frac{\|\mathbf{Q}\mathbf{w}\| - \eta\|\mathbf{w}\|}{\sqrt{\mathbf{w}^H(\hat{\mathbf{R}} + \gamma\mathbf{I})\mathbf{w}}}. \tag{32}$$

Since (32) is homogenous, without loss of generality, the term $\mathbf{w}^H(\hat{\mathbf{R}} + \gamma\mathbf{I})\mathbf{w}$ can be fixed to be equal to one. By doing so and introducing the auxiliary variable β, the problem (32) can be equivalently rewritten as

$$\begin{aligned} \max_{\mathbf{w},\beta} \quad & \sqrt{\mathbf{w}^H\mathbf{Q}^H\mathbf{Q}\mathbf{w}} - \eta\sqrt{\beta} \\ \text{s.t.} \quad & \mathbf{w}^H(\hat{\mathbf{R}} + \gamma\mathbf{I})\mathbf{w} = 1,\ \mathbf{w}^H\mathbf{w} = \beta \end{aligned} \tag{33}$$

where β takes values in a closed interval as it is shown next. Specifically, the problem (33) is feasible if and only if $\beta \in [\gamma_1, \gamma_2]$ where $\gamma_1 \triangleq \lambda_{\min}\{(\hat{\mathbf{R}} + \gamma\mathbf{I})^{-1}\}$, $\gamma_2 \triangleq \lambda_{\max}\{(\hat{\mathbf{R}} + \gamma\mathbf{I})^{-1}\}$, and $\lambda_{\min}\{\cdot\}$ stands for the smallest eigenvalue operator. In similar steps as in Section IV-A, i.e., by introducing $\mathbf{W} \triangleq \mathbf{w}\mathbf{w}^H$ and relaxing rank-one constraint, (33) can be equivalently recast as

$$\begin{aligned} \max_{\mathbf{W},\beta} \quad & \sqrt{\text{tr}\{\mathbf{Q}^H\mathbf{Q}\mathbf{W}\}} - \eta\sqrt{\beta} \\ \text{s.t.} \quad & \text{tr}\{(\hat{\mathbf{R}} + \gamma\mathbf{I})\mathbf{W}\} = 1, \text{tr}\{\mathbf{W}\} = \beta, \mathbf{W} \succeq \mathbf{0} \end{aligned} \tag{34}$$

where the optimal solution of (33) can be extracted precisely from the optimal solution of the problem (34). Thus, hereafter we focus on the problem (34). This problem is a DC optimization problem which can be addressed using the POTDC algorithm. Specifically, the proposed POTDC method can be applied to (34) by successively linearizing the term $\sqrt{\beta}$ around suitably selected points. Moreover, all the related results hold true in this case. In order to find the conditions which guarantee the global optimality of the POTDC method, let us introduce the following OVF

$$\begin{aligned} m(\beta) \triangleq \Big\{ \max_{\mathbf{W}} \text{tr}\{\mathbf{Q}^H\mathbf{Q}\mathbf{W}\} \mid \text{tr}\{(\hat{\mathbf{R}} + \gamma\mathbf{I})\mathbf{W}\} = 1, \\ \text{tr}\{\mathbf{W}\} = \beta, \mathbf{W} \succeq \mathbf{0} \Big\}, \gamma_1 \le \beta \le \gamma_2. \end{aligned} \tag{35}$$

Similar to the convexity proof for the OVF $l(\alpha, \alpha_c)$ (see Lemma 4), it can be easily verified that the OVF $m(\beta)$ is concave. Based on the definition of the OVF $m(\beta)$, the problem (34) can be further simplified as

$$\max_{\beta} \quad \sqrt{m(\beta)} - \eta\sqrt{\beta} \quad \text{s.t.} \quad \gamma_1 \le \beta \le \gamma_2. \tag{36}$$

Note that since $m(\beta)$ is a concave function, $\sqrt{m(\beta)}$ is also a concave function and as a result, the objective function of the problem (36) is the difference of two concave functions. The following theorem shows when the problem (34), or equivalently, (36) is guaranteed to be solvable globally optimally by our proposed method.

Theorem 1: For any arbitrary $\hat{\mathbf{R}}$ and $\mathbf{R}_{\text{s}} = \mathbf{Q}^H\mathbf{Q}$ whose corresponding OVF $m(\beta)$ is strictly concave and continuously differentiable, and provided that η is sufficiently small, the proposed POTDC method finds the globally optimal solution of the problem (34), or equivalently, (36).

Proof: See Appendix I-E. The explicit condition for η to be sufficiently small is specified in the proof and it is not repeated in the theorem formulation because of the space limitations. ■

Note that the result of Theorem 1 also holds when OVF $m(\beta)$ is non-differentiable. In this case, the proof follows similar steps, but it is slightly more technical and therefore omitted because of the space limitations.

Additionally, note that Theorem 1 does not imply that if η is not sufficiently small or the OVF $m(\beta)$ is not continuously differentiable, the POTDC method does not find the globally optimal solution. In other words, the condition of the theorem is sufficient but not necessary. Indeed, according to our numerical results, the globally optimal solution is always achieved by the proposed method.

D. Lower-Bounds on the Optimal Value of Problem (27)

We also aim at developing a tight lower-bound for the optimal value of the optimization problem (27). Such lower-bound is also used for assessing the performance of the proposed iterative algorithm.

As it was mentioned earlier, although the objective function of the optimization problem (27) is convex, its feasible set is non-convex due to the second constraint of (27). A lower-bound for the optimal value of (27) can be achieved by replacing the second constraint of (27) by its corresponding convex-hull. However, such lower-bound may not be tight. In order to obtain a tight lower-bound, we can divide the sector $[\theta_1, \theta_2]$ into N subsectors and solve the optimization problem (27) over each subsector in which the second constraint of (27) has been replaced with the corresponding convex hull. The minimum of the optimal values of such optimization problem over the subsectors is the lower-bound for the problem (27). It is obvious that by increasing N, the lower-bound becomes tighter.

V. Simulation Results

Let us consider a uniform linear array (ULA) of 10 omni-directional antenna elements with the inter-element spacing of half wavelength. Additive noise in antenna elements is modeled as spatially and temporally independent complex Gaussian noise with zero mean and unit variance. Throughout all simulation examples, it is assumed that in addition to the desired source, an interference source with the interference-to-noise ratio (INR) of 30 dB impinges on the antenna array. For obtaining each point in the simulation examples, 100 independent runs are used unless otherwise is specified and the sample data covariance matrix is estimated using $K = 50$ snapshots.

The new proposed method is compared in terms of the output SINR to the general-rank RAB methods of [12]–[14] and to the rank-one worst-case RAB of [2]. Moreover, the proposed method and the aforementioned general-rank RAB methods are also compared in terms of the achieved values for the objective function of the problem (12). The diagonal loading parameters of $\gamma = 10$ and $\eta = 0.5\sqrt{\text{tr}\{\mathbf{R}_{\text{s}}\}}$ are chosen for the proposed RAB and the RAB methods of [13] and [14], and the parameters of $\gamma = 10$ and $\epsilon = 8\sigma_{\text{s}}^2$ are chosen for the RAB of [12]. The initial point α_0 (see Algorithm 1) in the first iteration of the proposed method equals to $(\theta_1 + \theta_2)/2$ unless otherwise is

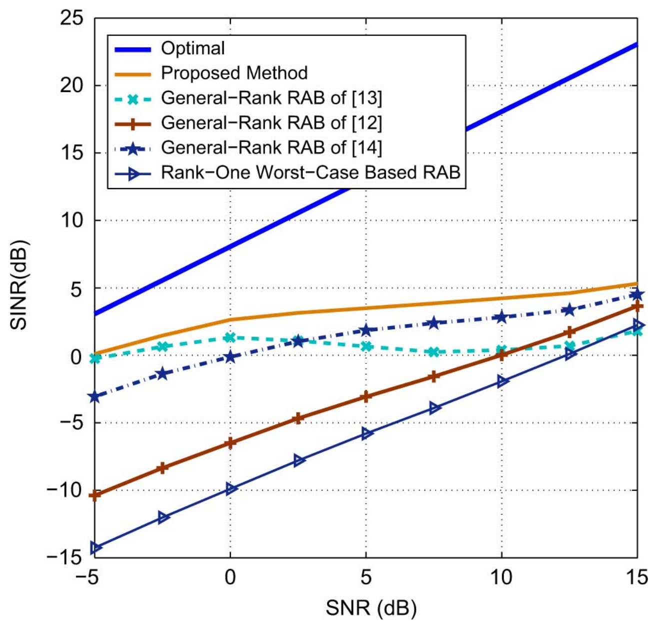

Fig. 2. Example 1: Output SINR versus SNR;INR=30 dB and $K = 50$.

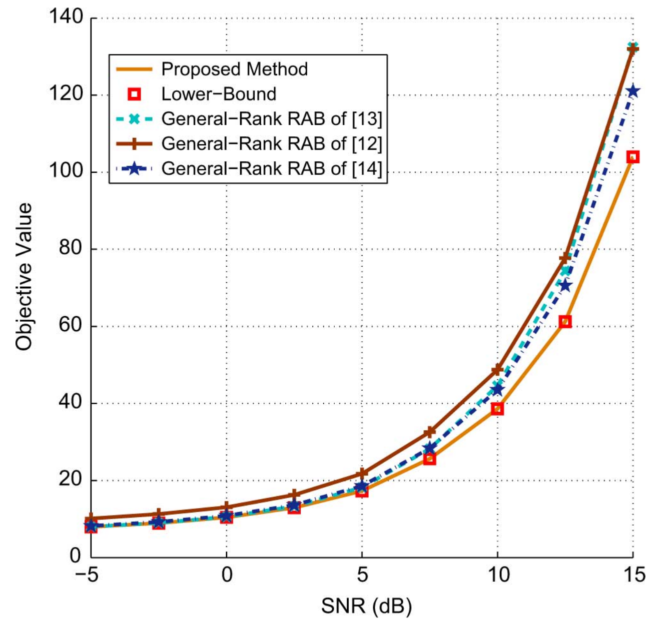

Fig. 3. Example 1: Objective function value of the problem (12) versus SNR; INR=30 dB, and $K = 50$.

specified. The termination threshold ζ for the proposed method is chosen to be equal to 10^{-6}. For obtaining a lower-bound on the optimal value of the optimization problem (27), the interval $[\theta_1, \theta_2]$ is divided in to 50 subsectors.

A. Simulation Example 1

In this example, the desired and interference sources are locally incoherently scattered with Gaussian and uniform angular power densities with central angles of 30° and 10°, respectively. The angular spreads of the desired and the interfering sources are assumed to be 4° and 10°, respectively. The presumed knowledge of the desired source is different from the actual one and is characterized by an incoherently scattered source with Gaussian angular power density whose central angle and angular spread are 34° and 6°, respectively. Note that, the presumed knowledge about the shape of the angular power density of the desired source is correct while the presumed central angle and angular spread deviate from the actual one.

In Figs. 2 and 3, the output SINR and the objective function values of the problem (12), respectively, are plotted versus SNR. It can be observed from the figures that the proposed new method based on the POTDC algorithm has superior performance over the other RABs. Moreover, Fig. 3 confirms that the new proposed method achieves the global minimum of the optimization problem (12) since the corresponding objective value coincides with the lower-bound on the objective function of the problem (12). Fig. 4 shows the convergence of the iterative POTDC method in terms of the average of the optimal value found by the algorithm over iterations for SNR=15 dB. It can be observed that the proposed algorithm converges to the global optimum in about 4 iterations.

B. Simulation Example 2

In the second example, we study how the rank of the actual correlation matrix of the desired source $\mathbf{R}_s$ affects the performance of the proposed general-rank RAB and other methods

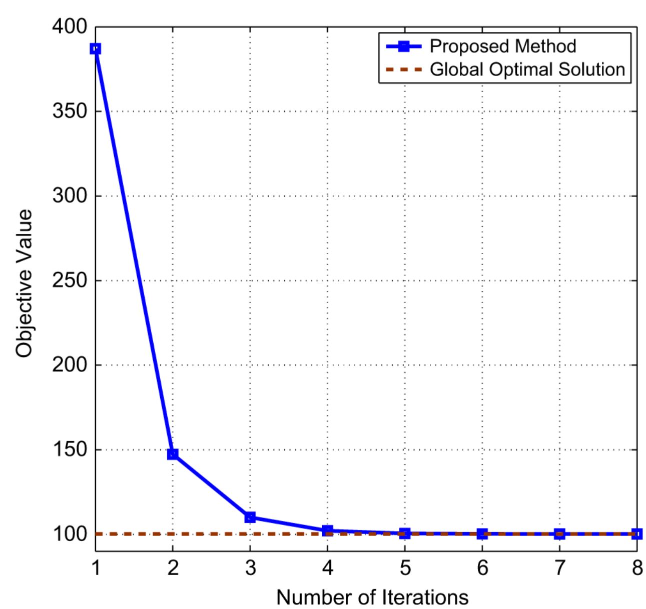

Fig. 4. Example 1: Objective function value of the problem (12) versus the number of iterations;SNR=15 dB, INR=30 dB, and $K = 50$.

tested. The same simulation set up as in the previous example is considered. The only difference is that the actual angular spread of the desired source varies and so does the actual rank of the desired source covariance matrix. The angular spread of the desired user is chosen to be 1°, 2°, 5°, 9°, and 14°. Figs. 5 and 6 show, respectively, the output SINR and the objective function values of the problem (12) versus the rank of the actual correlation matrix of the desired source for different methods when SNR=10 dB. It can be seen from the figures that the proposed method outperforms the other methods in all rank in terms of the objective value of the optimization problem (12) and it

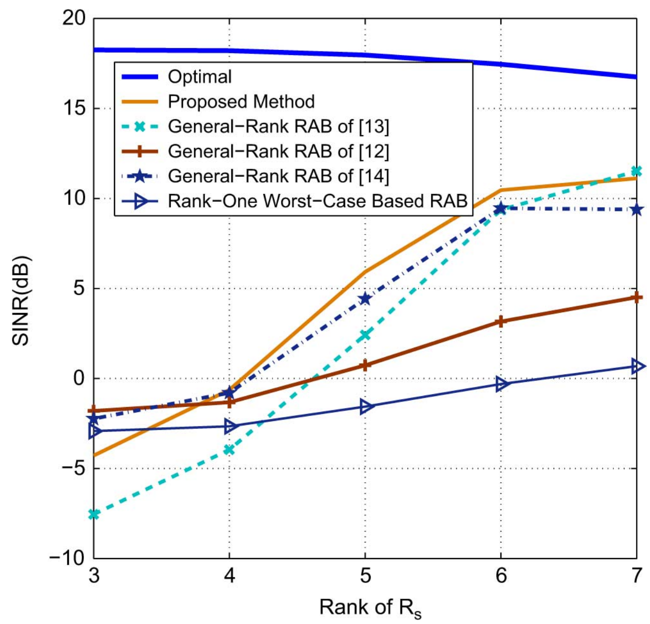

Fig. 5. Example 2: Output SINR versus the actual rank of $\mathbf{R}_s$; SNR=10 dB, INR=30 dB, and $K = 50$.

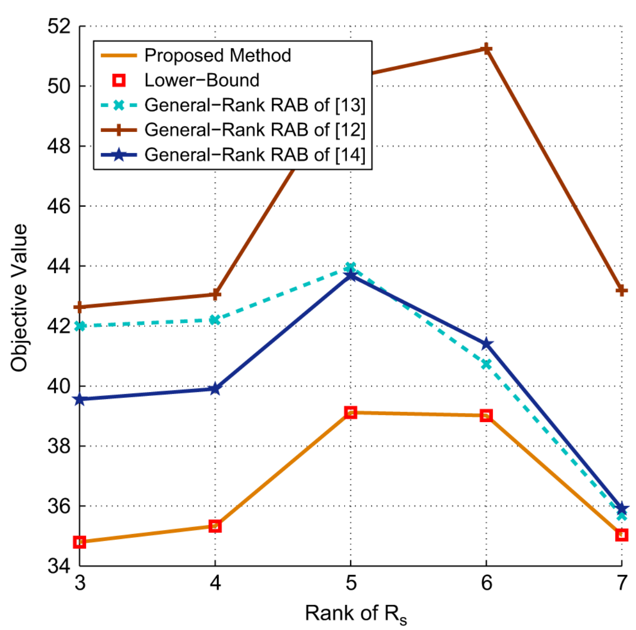

Fig. 6. Example 2: Objective function value of the problem (12) versus the actual rank of $\mathbf{R}_s$; SNR=10 dB, INR=30 dB, and $K = 50$.

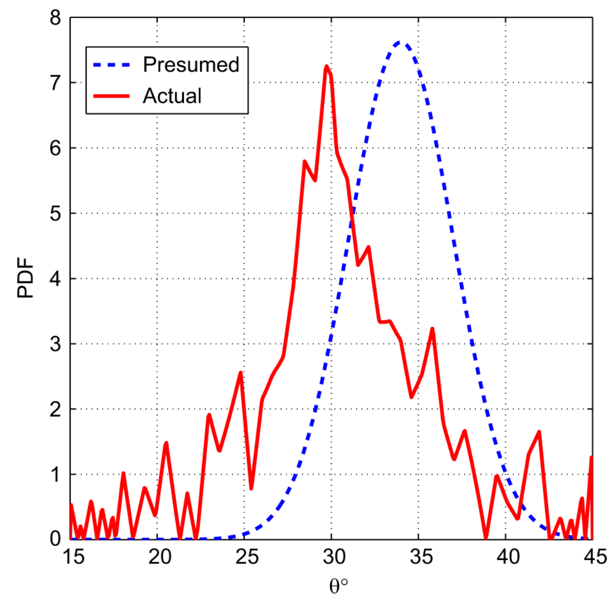

Fig. 7. Example 3: Actual and presumed angular power densities of general-rank source.

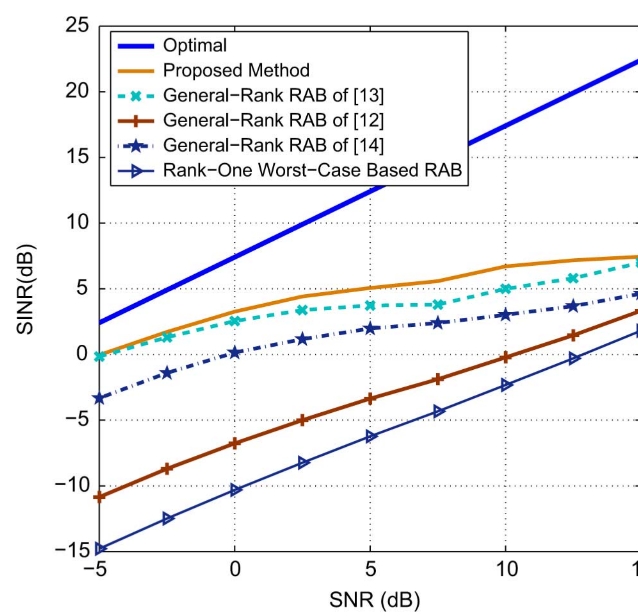

Fig. 8. Example 3: Output SINR versus SNR; INR=30 dB and $K = 50$.

achieves the globally optimal solution as it coincides with the lower-bound.

C. Simulation Example 3

In this example, we also consider the locally incoherently scattered desired and interference sources. However, compared to the previous examples, there is a substantial error in the knowledge of the desired source angular power density.

The interference source is modeled as in the previous examples, while the angular power density of the desired source is assumed to be a truncated Laplacian function distorted by severe fluctuations. The central angle and the scale parameter of the Laplacian distribution is assumed to be 30° and 0.1, respectively, and it is assumed to be equal to zero outside of the interval $[15^\circ, 45^\circ]$ as it has been shown in Fig. 7. The presumed knowledge of the desired source is different from the actual one and is characterized by an incoherently scattered source with Gaussian angular power density whose central angle and angular spread are 34° and 6°, respectively.

Fig. 8 depicts the corresponding output SINR of the problem (12) obtained by the beamforming methods tested versus SNR. It can be concluded from the figure that the proposed method has superior performance over the other methods.

TABLE I
AVERAGE NUMBER OF THE ITERATIONS

Array size	8	10	12	14	16	18	20
POTDC	3.260	3.265	3.000	3.140	3.080	3.285	3.235
DC iteration	4.930	5.765	5.950	6.750	7.510	8.240	8.835

TABLE II
AVERAGE CPU TIME

Array size	8	10	12	14	16	18	20
POTDC	0.674	0.740	0.726	0.812	0.845	0.972	1.027
DC iteration	1.243	1.480	1.583	1.875	2.118	2.399	2.598

D. Simulation Example 4

Finally, we compare the efficiency of the proposed POTDC method to that of the DC iteration-based method that can be written for the problem (12) as

$$\begin{aligned} &\min_{\mathbf{w}} \mathbf{w}^H(\hat{\mathbf{R}} + \gamma\mathbf{I})\mathbf{w} \\ &\text{s.t. } f(\mathbf{w}^{(k)}) + \langle \nabla f(\mathbf{w}^{(k)}), \mathbf{w} - \mathbf{w}^{(k)} \rangle - \eta\|\mathbf{w}\| \geq 1 \end{aligned} \quad (37)$$

where $\langle \cdot, \cdot \rangle$ denotes the inner product, $\nabla(\cdot)$ stands for the gradient operator, and the function $f(\mathbf{w}) \triangleq \|\mathbf{Q}\mathbf{w}\|$ is replaced with the first two terms of the Taylor expansion of $f(\mathbf{w})$ around $\mathbf{w}^{(k)}$. First, $\mathbf{w}^{(0)}$ is initialized, and in the next iterations, $\mathbf{w}^{(k)}$ is selected as the optimal $\mathbf{w}$ obtained from solving (37) in the previous iteration. Thus, the iterations are performed over the whole vector of variables of the problem.

The simulation set up is the same as in Simulation Example 1 except that different number of antennas is used. For a fair comparison, the initial point α_0 in the proposed method and $\mathbf{w}^{(0)}$ in (37) are chosen randomly. Particularly, the initialization point for the proposed POTDC method is chosen uniformly over the interval $[\theta_1, \theta_2]$ while the imaginary and real parts of the initial vector $\mathbf{w}^{(0)}$ in (37) are chosen independently as zero mean, unit variance, Gaussian random variables. If the so-generated $\mathbf{w}^{(0)}$ is not feasible, another initialization point is generated and this process continues until a feasible point is resulted. Note that the time which is consumed during the generation of a feasible point is negligible and it has not been considered in the average CPU time comparison. Table I shows the average number of iterations till convergence for the aforementioned methods versus the size of the antenna array. The termination threshold is set to 10^{-6}, $\text{SNR} = -5$ dB, and each number in the table is obtained by averaging the results over 200 runs. It can be seen from the table that the number of iterations for the proposed method is essentially fixed while it increases for the DC-iteration method as the size of the array, and thus the size of the problem (12), increases. The latter phenomenon can be justified by considering the DC iteration-type interpretation of the proposed method over the one dimensional OVF of $k(\alpha)$. The dimension of $k(\alpha)$ is independent of the size of the array (thus, the size of the optimization problem), while the size of the search space over iterations for the DC iteration-based method (37), that is, $2M$, increases as M increases. The average (over 200 runs) CPU time for the aforementioned methods is also shown in Table II. Both methods have been implemented in Matlab using CVX software and run on the same PC with Intel(R) Core(TM)2 CPU 2.66 GHz.

Table II confirms that the proposed method is more efficient than the DC iteration-based one in terms of the time required for convergence. It is worth noting also that although the number of variables in the matrix $\mathbf{W}$ of the optimization problem (28) is in general M^2 (since $\mathbf{W}$ has to be a Hermitian matrix) after the rank-one constraint is relaxed, the probability that the optimal $\mathbf{W}$ has rank one is very high as shown in [11], [31]–[33]. Thus, in almost all cases, for different data sets, the actual dimension of the problem (28) is $2M + 1$. As a result, the average complexity of solving (28) is significantly smaller than the worst-case complexity, which is also guaranteed to be polynomial.

VI. CONCLUSION

We have considered the RAB problem for general-rank signal model with additional positive semi-definite constraint. Such RAB problem corresponds to a non-convex DC optimization problem. We have studied this non-convex DC problem and designed the POTDC-type algorithm for solving it. It has been proved that the point found by the POTDC algorithm for the RAB for general-rank signal model with positive semi-definite constraint is a KKT optimal point. Moreover, the problem considered can be solved globally optimally under certain conditions. Specifically, we have proved that if the presumed norm of the mismatch that corresponds to the covariance matrix of the desired source is sufficiently small, then the proposed POTDC method finds the globally optimal solution of the corresponding optimization problem. The resulted RAB method shows superior performance compared to the other existing methods in terms of the output SINR and the resulted objective value. It also has complexity that is guaranteed to be polynomial. None of the existing methods used for DC programming problems guarantee that the global optimum can be found in polynomial time, even under some conditions. Thus, the fundamental development of this work is the claim of global optimality and the fact that this claim boils down to convexity of the OVF (25). It implies that certain relatively simple DC programming problems, which have been believed to be NP-hard, are actually not NP-hard under certain conditions.

APPENDIX

A. Proof of Lemma 1

The optimization problem (9) can be equivalently expressed as

$$\begin{aligned} \min_{\mathbf{\Delta}} \quad & \|\mathbf{Q}\mathbf{w} + \mathbf{\Delta}\mathbf{w}\| \\ \text{s.t.} \quad & \|\mathbf{\Delta}\| \leq \eta. \end{aligned} \quad (38)$$

First, note that based on the Cauchy-Schwarz inequality it can be found that $\|\mathbf{\Delta}\mathbf{w}\| \leq \|\mathbf{\Delta}\| \cdot \|\mathbf{w}\| \leq \eta\|\mathbf{w}\|$. The latter implies that under the condition that $\|\mathbf{\Delta}\| \leq \eta$, the norm of the vector $\mathbf{\Delta}\mathbf{w}$ is always less than or equal to $\eta\|\mathbf{w}\|$. Depending on whether the norm of $\mathbf{Q}\mathbf{w}$ is greater than or smaller than $\eta\|\mathbf{w}\|$, two different cases are possible. First, let us consider the case that $\|\mathbf{Q}\mathbf{w}\| \leq \eta\|\mathbf{w}\|$. Then, by choosing $\mathbf{\Delta}$ as $\mathbf{\Delta}_0 = -\mathbf{Q}\mathbf{w}\mathbf{w}^H/\|\mathbf{w}\|^2$ it is guaranteed that $\|\mathbf{\Delta}_0\| \leq \eta$ and the matrix product $\mathbf{\Delta}_0\mathbf{w}$ becomes equal to $-\mathbf{Q}\mathbf{w}$. The former can be verified simply as follows

$$\begin{aligned}\|\mathbf{\Delta}_0\|^2 &= \frac{1}{\|\mathbf{w}\|^4}\|\mathbf{Q}\mathbf{w}\mathbf{w}^H\|^2 \\ &= \frac{1}{\|\mathbf{w}\|^4}\mathrm{tr}\{\mathbf{Q}\mathbf{w}\mathbf{w}^H\mathbf{w}\mathbf{w}^H\mathbf{Q}^H\} \\ &= \frac{1}{\|\mathbf{w}\|^4}\mathrm{tr}\{\mathbf{Q}\mathbf{w}\mathbf{w}^H\mathbf{Q}^H\}\mathbf{w}^H\mathbf{w} \\ &= \frac{\|\mathbf{Q}\mathbf{w}\|^2}{\|\mathbf{w}\|^2} \leq \eta^2 \end{aligned} \tag{39}$$

where the last inequality is due to the assumption that $\|\mathbf{Q}\mathbf{w}\| \leq \eta\|\mathbf{w}\|$. By such $\mathbf{\Delta}_0$, the objective value of the problem (9) becomes equal to its smallest non-negative value, i.e., zero.

Next we consider the case when $\|\mathbf{Q}\mathbf{w}\| > \eta\|\mathbf{w}\|$. Then, choosing $\mathbf{\Delta}_1 = -\eta\mathbf{Q}\mathbf{w}\mathbf{w}^H/(\|\mathbf{Q}\mathbf{w}\|\cdot\|\mathbf{w}\|)$ results in the vector $\mathbf{\Delta}_1\mathbf{w}$ to be parallel to the vector $\mathbf{Q}\mathbf{w}$. Since $\|\mathbf{\Delta}_1\mathbf{w}\| = \eta\|\mathbf{w}\|$ (it can be verified by following similar steps as in (39)) and, as it was discussed earlier, the norm of the vector $\mathbf{\Delta}\mathbf{w}$ is always less than or equal to $\eta\|\mathbf{w}\|$, it can be concluded that $\mathbf{\Delta}_1\mathbf{w}$ is parallel to $\mathbf{Q}\mathbf{w}$ and it has the largest possible magnitude. In what follows, we show that the optimal solution in this case is equal to $\mathbf{\Delta}_1$. The following train of inequalities is in order

$$\|\mathbf{Q}\mathbf{w} + \mathbf{\Delta}\mathbf{w}\| \geq \|\mathbf{Q}\mathbf{w}\| - \|\mathbf{\Delta}\mathbf{w}\| \tag{40}$$
$$\geq \|\mathbf{Q}\mathbf{w}\| - \eta\|\mathbf{w}\| \tag{41}$$

where the first inequality is due to the triangular inequality and the second one is due to the Cauchy-Schwarz inequality. Since $\mathbf{\Delta}_1\mathbf{w}$ is parallel to $\mathbf{Q}\mathbf{w}$ and it has the largest possible magnitude, the inequalities (40) and (41) are both active when $\mathbf{\Delta} = \mathbf{\Delta}_1$, i.e., the equality holds, and therefore, $\mathbf{\Delta}_1$ is the optimal solution. This completes the proof. ■

B. Proof of Lemma 2

First, we verify whether the optimal solution of the optimization problem (13), or equivalently, the following problem

$$\begin{aligned}\min_{\mathbf{w}} \quad & \mathbf{w}^H(\hat{\mathbf{R}} + \gamma\mathbf{I})\mathbf{w} \\ \text{s.t.} \quad & \|\mathbf{Q}\mathbf{w}\| - \eta\|\mathbf{w}\| \geq 1 \end{aligned} \tag{42}$$

is achievable or not.

Let $\mathbf{w}_0$ denote any arbitrary feasible point of the problem (42). It is easy to see that if $\mathbf{w}^H\mathbf{w} \geq \mathbf{w}_0^H(\hat{\mathbf{R}} + \gamma\mathbf{I})\mathbf{w}_0/\lambda_{\min}\{\hat{\mathbf{R}} + \gamma\mathbf{I}\}$, then $\mathbf{w}^H(\hat{\mathbf{R}} + \gamma\mathbf{I})\mathbf{w}$ is greater than or equal to $\mathbf{w}_0^H(\hat{\mathbf{R}} + \gamma\mathbf{I})\mathbf{w}_0$. The latter implies that if the optimal solution is achievable, it lies inside the sphere of $\mathbf{w}^H\mathbf{w} \leq \mathbf{w}_0^H(\hat{\mathbf{R}} + \gamma\mathbf{I})\mathbf{w}_0/\lambda_{\min}\{\hat{\mathbf{R}} + \gamma\mathbf{I}\}$. Based on this fact, the optimization problem (42) can be recast as

$$\begin{aligned}\min_{\mathbf{w}} \quad & \mathbf{w}^H(\hat{\mathbf{R}} + \gamma\mathbf{I})\mathbf{w} \\ \text{s.t.} \quad & \|\mathbf{Q}\mathbf{w}\| - \eta\|\mathbf{w}\| \geq 1, \\ & \mathbf{w}^H\mathbf{w} \leq \frac{\mathbf{w}_0^H(\hat{\mathbf{R}} + \gamma\mathbf{I})\mathbf{w}_0}{\lambda_{\min}\{\hat{\mathbf{R}} + \gamma\mathbf{I}\}}. \end{aligned} \tag{43}$$

The feasible set of the new constraint in (43) is bounded and closed. Moreover, it can be easily shown that the feasible set of the constraint $\|\mathbf{Q}\mathbf{w}\| - \eta\|\mathbf{w}\| \geq 1$ is also closed. Specifically, due to the fact that first constraint of the problem (43) is a sublevel set of the following continuous function $q(\mathbf{w}) \triangleq \eta\|\mathbf{w}\| - \|\mathbf{Q}\mathbf{w}\|$, its feasible set is closed [34]. Since both of the feasible sets of the constraints are closed and one of them is bounded, the feasible set of the problem (43), which is the intersection of these two sets, is also closed and bounded. The latter implies that the feasible set of the problem (43) is compact. Therefore, also based on the fact that the objective function of (43) is continues, the optimal solution of (43), or equivalently (13), is achievable.

Let $(\mathbf{w}_{\mathrm{opt}}, \alpha_{\mathrm{opt}})$ denote the optimal solution of the problem (13) and let us define the following auxiliary optimization problem

$$\begin{aligned}\min_{\mathbf{w}} \quad & \mathbf{w}^H(\hat{\mathbf{R}} + \gamma\mathbf{I})\mathbf{w} \\ \text{s.t.} \quad & \mathbf{w}^H\mathbf{Q}^H\mathbf{Q}\mathbf{w} = \alpha_{\mathrm{opt}} \\ & \mathbf{w}^H\mathbf{w} \leq \frac{(\sqrt{\alpha_{\mathrm{opt}}} - 1)^2}{\eta^2}. \end{aligned} \tag{44}$$

It can be seen that if $\mathbf{w}$ is a feasible point of (44), then the pair $(\mathbf{w}, \alpha_{\mathrm{opt}})$ is also a feasible point of (13), which implies that the optimal value of (44) is greater than or equal to that of (13). However, since $\mathbf{w}_{\mathrm{opt}}$ is a feasible point of (44) and the value of the objective function at this feasible point is equal to the optimal value of (13), i.e., it is equivalent to $\mathbf{w}_{\mathrm{opt}}^H(\hat{\mathbf{R}}+\gamma\mathbf{I})\mathbf{w}_{\mathrm{opt}}$, it can be concluded that both of the optimization problems (13) and (44) have the same optimal value.

Let us define another auxiliary optimization problem based on (44) as

$$\begin{aligned} g \triangleq \min_{\mathbf{w}} \; & \mathbf{w}^H(\hat{\mathbf{R}} + \gamma\mathbf{I})\mathbf{w} \\ \text{s.t.} \quad & \mathbf{w}^H\mathbf{Q}^H\mathbf{Q}\mathbf{w} = \alpha_{\mathrm{opt}} \end{aligned} \tag{45}$$

which is obtained from (44) by dropping the last constraint of (44). The feasible set of (44) is a subset of the feasible set of (45). Thus, the optimal value g of (45) is smaller than or equal to the optimal value of (44), and thus also, the optimal value of (13). Using the maximin theorem [24], it is easy to verify that $g = \alpha_{\mathrm{opt}}/\lambda_{\max}\left\{(\hat{\mathbf{R}} + \gamma\mathbf{I})^{-1}\mathbf{Q}^H\mathbf{Q}\right\}$. Since g is smaller than or equal to the optimal value of (13), it is upper-bounded by $\mathbf{w}_0^H(\hat{\mathbf{R}} + \gamma\mathbf{I})\mathbf{w}_0$, where $\mathbf{w}_0$ is an arbitrary feasible point of (13). The latter implies that $\alpha_{\mathrm{opt}} \leq \lambda_{\max}\left\{(\hat{\mathbf{R}} + \gamma\mathbf{I})^{-1}\mathbf{Q}^H\mathbf{Q}\right\}\mathbf{w}_0^H(\hat{\mathbf{R}} + \gamma\mathbf{I})\mathbf{w}_0$. This completes the proof. ■

C. Proof of Lemma 3

First, we prove that $l(\alpha, \alpha_{\mathrm{c}})$ is a convex function with respect to α. For this goal, let $\mathbf{W}_{\alpha_1}$ and $\mathbf{W}_{\alpha_2}$ denote the optimal solutions of the optimization problems of $l(\alpha_1, \alpha_{\mathrm{c}})$ and $l(\alpha_2, \alpha_{\mathrm{c}})$, respectively, i.e., $l(\alpha_1, \alpha_{\mathrm{c}}) = \mathrm{tr}\left\{(\hat{\mathbf{R}} + \gamma\mathbf{I})\mathbf{W}_{\alpha_1}\right\}$ and $l(\alpha_2, \alpha_{\mathrm{c}}) = \mathrm{tr}\left\{(\hat{\mathbf{R}} + \gamma\mathbf{I})\mathbf{W}_{\alpha_2}\right\}$, where α_1 and α_2 are any two arbitrary points in the interval $[\theta_1, \theta_2]$. It is trivial to verify that $\theta\mathbf{W}_{\alpha_1} + (1-\theta)\mathbf{W}_{\alpha_2}$ is a feasible point of the corresponding optimization problem of $l(\theta\alpha_1 + (1-\theta)\alpha_2, \alpha_c)$ (see the definition (29)). Therefore,

$$\begin{aligned} l(\theta\alpha_1 + &(1-\theta)\alpha_2, \alpha_c) \\ &\leq \mathrm{tr}\big\{(\hat{\mathbf{R}} + \gamma\mathbf{I})(\theta\mathbf{W}_{\alpha_1} + (1-\theta)\mathbf{W}_{\alpha_2})\big\} \\ &= \theta\mathrm{tr}\big\{(\hat{\mathbf{R}} + \gamma\mathbf{I})\mathbf{W}_{\alpha_1}\big\} \\ &\quad + (1-\theta)\mathrm{tr}\big\{(\hat{\mathbf{R}} + \gamma\mathbf{I})\mathbf{W}_{\alpha_2}\big\} \\ &= \theta l(\alpha_1, \alpha_c) + (1-\theta) l(\alpha_2, \alpha_c) \end{aligned} \tag{46}$$

which proves that $l(\alpha, \alpha_c)$ is a convex function with respect to α.

In order to show that $l(\alpha, \alpha_c)$ is greater than or equal to $k(\alpha)$, it suffices to show that the feasible set of the optimization problem of $l(\alpha, \alpha_c)$ is a subset of the feasible set of the optimization problem of $k(\alpha)$. Let $\mathbf{W}_1$ denote a feasible point of the optimization problem of $l(\alpha, \alpha_c)$, it is easy to verify that $\mathbf{W}_1$ is also a feasible point of the optimization problem of $k(\alpha)$ if the inequality $\sqrt{\alpha} \leq \sqrt{\alpha_c} + (\alpha - \alpha_c)/(2\sqrt{\alpha_c})$ holds. This inequality can be rearranged as

$$(\sqrt{\alpha} - \sqrt{\alpha_c})^2 \geq 0 \tag{47}$$

and it is valid for any arbitrary α. Therefore, $\mathbf{W}_1$ is also a feasible point of the optimization problem of $k(\alpha)$ which implies that $l(\alpha, \alpha_c) \geq k(\alpha)$.

In order to show that the right and left derivatives are equal, we use the result of [35] which gives expressions for the directional derivatives of a parametric SDP. Specifically, the directional derivatives for the following OVF

$$\psi(\mathbf{u}) \triangleq \{\min_{\mathbf{y}} f(\mathbf{y}, \mathbf{u}) \mid \mathbf{G}(\mathbf{y}, \mathbf{u}) \preceq \mathbf{0}_{n\times n}\} \tag{48}$$

are derived in [35], where $f(\mathbf{y}, \mathbf{u})$ and $\mathbf{G}(\mathbf{y}, \mathbf{u})$ are a scalar and an $n \times n$ matrix, respectively, $\mathbf{y}$ is the $m \times 1$ real valued vector of optimization variables, and $\mathbf{u}$ is the $k \times 1$ real valued vector of optimization parameters. Let $\mathbf{u}_c$ be an arbitrary fixed point. If the optimization problem of $\psi(\mathbf{u}_c)$ poses certain properties, then according to [35, Theorem 10] it is directionally differentiable at $\mathbf{u}_c$. These properties are (i) the functions $f(\mathbf{y}, \mathbf{u})$ and $\mathbf{G}(\mathbf{y}, \mathbf{u})$ are continuously differentiable, (ii) the optimization problem of $\psi(\mathbf{u}_c)$ is convex, (iii) the set of optimal solutions of the optimization problem of $\psi(\mathbf{u}_c)$ denoted as $\mathcal{M}$ is nonempty and bounded, (iv) the Slater condition for the optimization problem of $\psi(\mathbf{u}_c)$ holds true, and (v) the *inf-compactness* condition is satisfied. Here inf-compactness condition refers to the condition of the existence of $\alpha > \psi(\mathbf{u}_c)$ and a compact set $S \subset \mathcal{R}^m$ such that $\{\mathbf{y} | f(\mathbf{y}, \mathbf{u}) \leq \alpha, \mathbf{G}(\mathbf{y}, \mathbf{u}) \preceq \mathbf{0}\} \subset S$ for all $\mathbf{u}$ in a neighborhood of $\mathbf{u}_c$ where $\mathcal{R}^m$ denotes the m-dimensional Euclidean space. If for all $\mathbf{u}$ the optimization problem of $\psi(\mathbf{u})$ is convex and the set of optimal solutions of $\psi(\mathbf{u})$ is non-empty and bounded, then the inf-compactness conditions holds automatically.

The directional derivative of $\psi(\mathbf{u})$ at $\mathbf{u}_c$ in a direction $\mathbf{d} \in \mathcal{R}^k$ is given by

$$\psi'(\mathbf{u}_c, \mathbf{d}) = \min_{\mathbf{y}\in\mathcal{M}} \max_{\mathbf{\Omega}\in\mathcal{Z}} \mathbf{d}^T \nabla_{\mathbf{u}} L(\mathbf{y}, \mathbf{\Omega}, \mathbf{u}_c) \tag{49}$$

where $\mathcal{Z}$ is the set of optimal solutions that corresponds to the dual problem of the optimization problem of $\psi(\mathbf{u}_c)$, and $L(\mathbf{y}, \mathbf{\Omega}, \mathbf{u})$ denotes the Lagrangian defined as

$$L(\mathbf{y}, \mathbf{\Omega}, \mathbf{u}) \triangleq f(\mathbf{y}, \mathbf{u}) + \mathrm{tr}\left(\mathbf{\Omega}\mathbf{G}(\mathbf{y}, \mathbf{u})\right) \tag{50}$$

where $\mathbf{\Omega}$ denotes the Lagrange multiplier matrix.

Let us look again to the definitions of the OVFs $k(\alpha)$ and $l(\alpha, \alpha_c)$(25) and (29), respectively, and define the following block diagonal matrix

$$\begin{aligned} &\mathbf{G}_1(\mathbf{W}, \alpha) \\ &\triangleq \begin{pmatrix} -\mathbf{W} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \eta^2\mathrm{tr}\{\mathbf{W}\} - (\sqrt{\alpha}-1)^2 & 0 & 0 \\ \mathbf{0} & 0 & \mathrm{tr}\{\mathbf{Q}^H\mathbf{Q}\mathbf{W}\} - \alpha & 0 \\ \mathbf{0} & 0 & 0 & \alpha - \mathrm{tr}\{\mathbf{Q}^H\mathbf{Q}\mathbf{W}\} \end{pmatrix} \end{aligned} \tag{51}$$

as well as another block diagonal matrix denoted as $\mathbf{G}_2(\mathbf{W}, \alpha)$ which has exactly same structure as the matrix $\mathbf{G}_1(\mathbf{W}, \alpha)$ with only difference that the element $\eta^2\mathrm{tr}\{\mathbf{W}\} - (\sqrt{\alpha}-1)^2$ in $\mathbf{G}_1(\mathbf{W}, \alpha)$ is replaced by $\eta^2\mathrm{tr}\{\mathbf{W}\} + (\sqrt{\alpha_c} - 1) + \alpha\left(1/\sqrt{\alpha_c} - 1\right)$ in $\mathbf{G}_2(\mathbf{W}, \alpha)$. Then the OVFs $k(\alpha)$ and $l(\alpha, \alpha_c)$ can be equivalently recast as

$$\begin{aligned} k(\alpha) = &\left\{\min_{\mathbf{W}} \mathrm{tr}\left\{(\hat{\mathbf{R}} + \gamma\mathbf{I})\mathbf{W}\right\} \;|\mathbf{G}_1(\mathbf{W}, \alpha) \preceq \mathbf{0}\right\}, \\ &\theta_1 \leq \alpha \leq \theta_2 \end{aligned} \tag{52}$$

and

$$\begin{aligned} l(\alpha, \alpha_c) = &\left\{\min_{\mathbf{W}} \mathrm{tr}\left\{(\hat{\mathbf{R}} + \gamma\mathbf{I})\mathbf{W}\right\} \;|\mathbf{G}_2(\mathbf{W}, \alpha) \preceq \mathbf{0}\right\}, \\ &\theta_1 \leq \alpha \leq \theta_2. \end{aligned} \tag{53}$$

It is straightforward to see that the functions $\mathrm{tr}\left\{(\hat{\mathbf{R}} + \gamma\mathbf{I})\mathbf{W}\right\}$, $\mathbf{G}_1(\mathbf{W}, \alpha)$, and $\mathbf{G}_2(\mathbf{W}, \alpha)$ are continuously differentiable. Furthermore, it is easy to verify that both optimization problems of $k(\alpha_c)$ and $l(\alpha_c, \alpha_c)$ can be expressed as

$$\begin{aligned} \min_{\mathbf{W}} \quad & \mathrm{tr}\{(\hat{\mathbf{R}} + \gamma\mathbf{I})\mathbf{W}\} \\ \text{s.t.} \quad & \mathrm{tr}\{\mathbf{Q}^H\mathbf{Q}\mathbf{W}\} = \alpha_c \\ & \mathrm{tr}\{\mathbf{W}\} \leq \frac{(\sqrt{\alpha_c}-1)^2}{\eta^2} \\ & \mathbf{W} \succeq \mathbf{0}. \end{aligned} \tag{54}$$

The problem (54) is convex and its solution set is non-empty and bounded. Indeed, let $\mathbf{W}_1$ and $\mathbf{W}_2$ denote two optimal solutions of the problem above. The Euclidean distance between $\mathbf{W}_1$ and $\mathbf{W}_2$ can be expressed as

$$\begin{aligned} \|\mathbf{W}_1 - \mathbf{W}_2\| &= \sqrt{\mathrm{tr}\{\mathbf{W}_1^2\} + \mathrm{tr}\{\mathbf{W}_2^2\} - 2\mathrm{tr}\{\mathbf{W}_1\mathbf{W}_2\}} \\ &\leq \sqrt{2}\frac{(\sqrt{\alpha_c}-1)^2}{\eta^2} \end{aligned} \tag{55}$$

where the last line is due to the fact that the matrix product $\mathbf{W}_1\mathbf{W}_2$ is positive semi-definite and, therefore, $\mathrm{tr}\{\mathbf{W}_1\mathbf{W}_2\} \geq 0$, and also the fact that for any arbitrary positive semi-definite matrix $\mathrm{tr}\{\mathbf{A}^2\} \leq (\mathrm{tr}\{\mathbf{A}\})^2$. From (55), it can be seen that the distance between any two arbitrary optimal solutions of (54) is finite and, therefore, the solution set is bounded. Moreover, it is easy to verify that the dual problem for the optimization problem (54) can be expressed as

$$\begin{aligned}\max_{\tau,\psi} \quad & \tau\alpha_{\mathrm{c}} - \psi(\sqrt{\alpha_{\mathrm{c}}}-1)^2/\eta^2 \\ \text{s.t.} \quad & (\hat{\mathbf{R}}+\gamma\mathbf{I}) - \tau\mathbf{Q}^H\mathbf{Q} + \psi\mathbf{I} \succeq \mathbf{0} \\ & \psi \geq 0 \end{aligned} \tag{56}$$

where τ and ψ are the Lagrange multipliers. The optimization problem (54) is a convex SDP problem which satisfies the Slater's conditions as the point $(\tau = 0,\ \psi = 1)$ is a strictly feasible point for its dual problem (56). Thus, the optimization problem (54) satisfies the strong duality. It can also be shown that the inf-compactness condition is satisfied by verifying that the optimization problems of $k(\alpha)$ and $l(\alpha, \alpha_{\mathrm{c}})$ are convex and their corresponding solution sets are bounded for any α. Therefore, both of the OVFs $k(\alpha)$ and $l(\alpha, \alpha_{\mathrm{c}})$ are directionally differentiable at α_{c}.

Using the result of [35, Theorem 10], the directional derivatives of $k(\alpha)$ and $l(\alpha, \alpha_{\mathrm{c}})$ can be respectively computed as

$$k'(\alpha, d) = \min_{\mathbf{W}\in\mathcal{M}} \max_{\mathbf{\Omega}\in\mathcal{Z}} d\left(\mathrm{tr}\{\mathbf{\Omega}\frac{d}{d\alpha}\mathbf{G}_1(\mathbf{W},\alpha)\big|_{\alpha=\alpha_{\mathrm{c}}}\}\right) \tag{57}$$

and

$$l'(\alpha, \alpha_{\mathrm{c}}, d) = \min_{\mathbf{W}\in\mathcal{M}} \max_{\mathbf{\Omega}\in\mathcal{Z}} d\left(\mathrm{tr}\{\mathbf{\Omega}\frac{d}{d\alpha}\mathbf{G}_2(\mathbf{W},\alpha)\big|_{\alpha=\alpha_{\mathrm{c}}}\}\right) \tag{58}$$

where $\mathcal{M}$ and $\mathcal{Z}$ denote the optimal solution sets of the optimization problem (54) and its dual problem, respectively. Using the definitions of $\mathbf{G}_1(\mathbf{W},\alpha)$ and $\mathbf{G}_2(\mathbf{W},\alpha)$, it can be seen that the terms $d\mathbf{G}_1(\mathbf{W},\alpha)/d\alpha$ and $d\mathbf{G}_2(\mathbf{W},\alpha)/d\alpha$ are equal at $\alpha = \alpha_{\mathrm{c}}$ and, therefore, the directional derivatives are equivalent. The latter implies that the left and right derivatives of $k(\alpha)$ and $l(\alpha, \alpha_{\mathrm{c}})$ are equal at $\alpha = \alpha_{\mathrm{c}}$. ■

D. Proof of Lemma 4

i) The optimization problem in Algorithm 1 at iteration $i,\ i \geq 2$ is obtained by linearizing $\sqrt{\alpha}$ around $\alpha_{\mathrm{opt,i-1}}$. Since $\mathbf{W}_{\mathrm{opt,i-1}}$ and $\alpha_{\mathrm{opt,i-1}}$ are feasible for the optimization problem at iteration i, it can be straightforwardly concluded that the optimal value of the objective at iteration i is less than or equal to the optimal value at the previous iteration, i.e., $\mathrm{tr}\left\{(\hat{\mathbf{R}}+\gamma\mathbf{I})\mathbf{W}_{\mathrm{opt,i}}\right\} \leq \mathrm{tr}\left\{(\hat{\mathbf{R}}+\gamma\mathbf{I})\mathbf{W}_{\mathrm{opt,i-1}}\right\}$.

ii) Since the sequence of the optimal values, i.e., $\mathrm{tr}\left\{(\hat{\mathbf{R}}+\gamma\mathbf{I})\mathbf{W}_{\mathrm{opt,i}}\right\}, i \geq 1$ is non-increasing and bounded from below (every optimal value is non-negative), the sequence of the optimal values converges.

iii) The proof follows straightforwardly from [36, Proposition 3.2]. Moreover, every feasible point of the problem (27) is a regular point. Specifically, if $\mathbf{W}_0$ and α_0 denote a feasible point of the problem (27), the gradients of the equality and inequality constraints of this problem at $\mathbf{W}_0$ and α_0 can be expressed, respectively, as

$$\mathbf{g}_1 = \left(\mathrm{Re}\{\mathrm{vec}\{\mathbf{R}_{\mathrm{s}}\}^T\}, -\mathrm{Im}\{\mathrm{vec}\{\mathbf{R}_{\mathrm{s}}\}^T\}, -1\right)^T \tag{59}$$

$$\mathbf{g}_2 = \left(\eta^2\mathrm{vec}\{\mathbf{I}\}^T, \mathbf{0}_{M^2\times1}, -1+\frac{1}{\sqrt{\alpha_0}}\right)^T \tag{60}$$

$$\mathbf{g}_3 = \left(\mathbf{0}_{M^2\times1}, \mathbf{0}_{M^2\times1}, 1\right)^T \tag{61}$$

$$\mathbf{g}_4 = \left(\mathbf{0}_{M^2\times1}, \mathbf{0}_{M^2\times1}, -1\right)^T \tag{62}$$

where $\mathrm{vec}\{\cdot\}$ denotes the vectorization operator, $\mathrm{Re}\{\cdot\}$ and $\mathrm{Im}\{\cdot\}$ denote, respectively, the real and imaginary parts of a complex number. Note that only one of the constraints $\alpha \leq \theta_2$ or $\theta_1 \leq \alpha$ can be active. The gradients $\mathbf{g}_1$, $\mathbf{g}_2$, and $\mathbf{g}_3$ (or $\mathbf{g}_4$) are linearly independent unless $\mathbf{R}_{\mathrm{s}}$ is proportional to the identity matrix $\mathbf{I}$, i.e., $\mathbf{R}_{\mathrm{s}} = c^2\mathbf{I}$ where c is some coefficient. Therefore, assuming first that $\mathbf{R}_{\mathrm{s}}$ is not proportional to the identity matrix, the linear independence constraint qualification (LICQ) holds at every feasible point. In the case when $\mathbf{R}_{\mathrm{s}}$ is proportional to the identity matrix, the problem (12) can be expressed as

$$\begin{aligned}\min_{\mathbf{w}} \quad & \mathbf{w}^H(\hat{\mathbf{R}}+\gamma\mathbf{I})\mathbf{w} \\ \text{s.t.} \quad & c\sqrt{\mathbf{w}^H\mathbf{w}} - \eta\sqrt{\mathbf{w}^H\mathbf{w}} \geq 1. \end{aligned} \tag{63}$$

Then, the optimal solution of the problem (63) can be trivially obtained as $\mathbf{w}_{\mathrm{opt}} = \boldsymbol{\mathcal{P}}\{(\hat{\mathbf{R}}+\gamma\mathbf{I})^{-1}\}/(c-\eta)$ provided that $c > \eta$. For the case when $c \leq \eta$, the problem (63) is not feasible.

Since every point obtained in any iteration of Algorithm 1 is a feasible point of the problem (27), the sequence generated by Algorithm 1 is a sequence of all regular points. Thus, this sequence of regular points converges to a regular point and such point satisfies the KKT optimality conditions. ■

E. Proof of Theorem 1

The OVF $m(\beta)$ is continuously differentiable under the condition that the optimal solution of its corresponding optimization problem is unique [37]. Although the following proof is established based on the continuous differentiability assumption of the OVF $m(\beta)$, it can be generalized to the case when the OVF $m(\beta)$ is not continuously differentiable.

Let $\beta_{\mathrm{m}}, \gamma_1 \leq \beta_{\mathrm{m}} \leq \gamma_2$ denote the optimal maximizer of the concave OVF $m(\beta)$. Note that since the OVF $m(\beta)$ is assumed to be strictly concave, its global maximizer, i.e., β_{m}, is unique. Moreover, the strict concavity of the OVF $m(\beta)$ together the fact that $m(\beta) \geq 0$ for any $\beta \in [\gamma_1, \gamma_2]$ imply that $m(\beta) > 0$ for $\beta \in (\gamma_1, \gamma_2)$. If $\beta_{\mathrm{m}} = \gamma_1$, then it is obvious that the function $\sqrt{m(\beta)} - \eta\sqrt{\beta}$ is strictly decreasing over the interval $[\gamma_1, \gamma_2]$ for any arbitrary $\eta > 0$ and, therefore, the POTDC method finds the globally optimal solution, i.e., γ_1. Next, we assume that $\beta_{\mathrm{m}} > \gamma_1$ which implies that $m'(\beta) > 0,\ \beta \in (\gamma_1, \beta_{\mathrm{m}})$ due to the strict concavity of $m(\beta)$. The optimal solution of the problem (36) denoted as β_{opt} is less than or equal to β_{m}. In order to show it, let us assume that $\beta_{\mathrm{opt}} > \beta_{\mathrm{m}}$. Using the fact that β_{opt} is the optimal solution of (36), it can be trivially concluded that

$$\sqrt{m(\beta_{\mathrm{opt}})} - \eta\sqrt{\beta_{\mathrm{opt}}} \geq \sqrt{m(\beta_{\mathrm{m}})} - \eta\sqrt{\beta_{\mathrm{m}}}. \tag{64}$$

Since it was assumed that $\beta_{\rm opt} > \beta_{\rm m}$, the inequality (64) can be rewritten as

$$m(\beta_{\rm opt}) \geq m(\beta_{\rm m}) + \eta^2(\sqrt{\beta_{\rm opt}} - \sqrt{\beta_{\rm m}})^2 + 2\eta\sqrt{m(\beta_{\rm m})}(\sqrt{\beta_{\rm opt}} - \sqrt{\beta_{\rm m}}). \quad (65)$$

From (65), it can be concluded that $m(\beta_{\rm opt}) > m(\beta_{\rm m})$ assuming $\beta_{\rm opt} > \beta_{\rm m}$. The latter contradicts the fact that $\beta_{\rm m}$ is the global maximizer of the OVF $m(\beta)$ and, thus, $\beta_{\rm opt} \leq \beta_{\rm m}$.

Using the fact that $\beta_{\rm opt} \leq \beta_{\rm m}$, the problem (36) can be further rewritten as

$$\max_{\beta} \sqrt{m(\beta)} - \eta\sqrt{\beta} \quad \text{s.t.} \quad \gamma_1 \leq \beta \leq \beta_{\rm m}. \quad (66)$$

Let us now assume that $\beta_{\rm m} < \gamma_2$ while the following arguments can be straightforwardly resulted also when $\beta_{\rm m} = \gamma_2$. It is obvious that the function $\sqrt{m(\beta)} - \eta\sqrt{\beta}$ is strictly decreasing over the interval $(\beta_{\rm m}, \gamma_2]$ and, therefore, the POTDC method will not be stuck over this interval for any arbitrary $\eta > 0$.

Since $\beta_{\rm m} \in (\gamma_1, \gamma_2)$ is the global maximizer of the OVF $m(\beta)$, we equivalently say that $m'(\beta_{\rm m}) = 0$. In what follows, we show that there exists η_0 such that for any $\eta \leq \eta_0$ the function $\sqrt{m(\beta)} - \eta\sqrt{\beta}$ is strictly quasi-concave over the interval $[\gamma_1, \beta_{\rm m}]$. For this goal, we define the following new function

$$p(\beta) \triangleq 2\sqrt{\beta}\frac{d(\sqrt{m(\beta)})}{d\beta} = \sqrt{\beta}\frac{m'(\beta)}{\sqrt{m(\beta)}}, \quad \beta \in (\gamma_1, \gamma_2). \quad (67)$$

Note that since $m(\beta)$ is continuously differentiable, $m'(\beta)$ is continuous and, therefore, $p(\beta)$ is a continuous function. Based on the definition of $\beta_{\rm m}$ and due to the strict concavity of $m(\beta)$, it can be concluded that the multiplicative term $m'(\beta)/\sqrt{m(\beta)}$ in the definition of $p(\beta)$ is a strictly decreasing function on the interval $(\gamma_1, \beta_{\rm m}]$ which approaches zero as β approaches $\beta_{\rm m}$. Moreover, the term $m'(\beta)/\sqrt{m(\beta)}$ is equal to zero at the point $\beta = \beta_{\rm m}$. The latter implies that the function $p(\beta)$ is non-zero over the interval $(\gamma_1, \beta_{\rm m})$, while $p(\beta) = 0$ at the point $\beta = \beta_{\rm m}$.

Using the properties that the term $m'(\beta)/\sqrt{m(\beta)}$ is strictly decreasing over the interval $(\gamma_1, \beta_{\rm m}]$ and the OVF $m(\beta)$ is strictly concave together with the fact that the the term $m'(\beta)/\sqrt{m(\beta)}$ approaches zero when $\beta \rightarrow \beta_{\rm m}$, it is straightforward to show that there exists ξ, $\gamma_1 < \xi < \beta_{\rm m}$ such that the function $p(\beta)$ is strictly decreasing over the interval $[\xi, \beta_{\rm m}]$. Moreover, using the strictly decreasing property of the term $m'(\beta)/\sqrt{m(\beta)}$ over the interval $(\gamma_1, \beta_{\rm m}]$, it can be easily concluded that $p(\beta) > p(\xi) \cdot (\sqrt{\gamma_1}/\sqrt{\xi})$ for any $\beta \in (\gamma_1, \xi]$. Since the lower-bound of $p(\beta)$ over the interval $(\gamma_1, \xi]$, i.e., $p(\xi)(\sqrt{\gamma_1}/\sqrt{\xi})$, is less than $p(\xi)$ and also $p(\beta)$ is a continuous function, there exists $\phi \in (\xi, \beta_{\rm m})$ such that $p(\phi) = p(\xi) \cdot (\sqrt{\gamma_1}/\sqrt{\xi})$. The latter implies that

$$p(\beta) > p(\phi), \qquad \forall\beta \in (\gamma_1, \phi). \quad (68)$$

Based on the fact that $p(\beta)$ is strictly decreasing over the interval $[\xi, \beta_{\rm m}]$ and, therefore, over the interval $[\phi, \beta_{\rm m}]$ and also based on (68), it can be concluded that for any $\nu \in [\phi, \beta_{\rm m})$, $p(\beta) > p(\nu)$ if $\gamma_1 < \beta < \nu$, while $p(\beta) < p(\nu)$ if $\nu < \beta < \beta_{\rm m}$. In other words,

$$\frac{d(\sqrt{m(\beta)})}{d\beta} - p(\nu)\frac{1}{2\sqrt{\beta}} > 0, \quad \gamma_1 < \beta < \nu \quad (69)$$

and

$$\frac{d(\sqrt{m(\beta)})}{d\beta} - p(\nu)\frac{1}{2\sqrt{\beta}} < 0, \quad \nu < \beta < \beta_{\rm m} \quad (70)$$

which implies that by selecting η equal to $p(\nu)$ where $\nu \in [\phi, \beta_{\rm m})$, or equivalently, if $\eta \leq \eta_0 \triangleq p(\phi)$, then the function $\sqrt{m(\beta)} - \eta\sqrt{\beta}$ is strictly quasi-concave, and therefore, the POTDC method finds the globally optimal solution. This completes the proof. ■

REFERENCES

[1] A. B. Gershman, "Robust adaptive beamforming in sensor arrays," *J. Electron. Commun.*, vol. 53, no. 6, pp. 305–314, Dec. 1999.

[2] S. A. Vorobyov, A. B. Gershman, and Z.-Q. Luo, "Robust adaptive beamforming using worst-case performance optimization: A solution to the signal mismatch problem," *IEEE Trans. Signal Process.*, vol. 51, pp. 313–324, Feb. 2003.

[3] J. Li, P. Stoica, and Z. Wang, "On robust capon beamforming and diagonal loading," *IEEE Trans. Signal Process.*, vol. 51, pp. 1702–1715, Jul. 2003.

[4] S. J. Kim, A. Magnani, A. Mutapcic, S. Boyd, and Z. Q. Luo, "Robust beamforming via worst-case SINR maximization," *IEEE Trans. Signal Process.*, vol. 56, no. 6, pp. 1539–1547, Apr. 2008.

[5] S. A. Vorobyov, A. B. Gershman, Z.-Q. Luo, and N. Ma, "Adaptive beamforming with joint robustness against mismatched signal steering vector and interference nonstationarity," *IEEE Signal Process. Lett.*, vol. 11, no. 2, pp. 108–111, Feb. 2004.

[6] J. Li, P. Stoica, and Z. Wang, "Doubly constrained robust capon beamformer," *IEEE Trans. Signal Process.*, vol. 52, pp. 2407–2423, Sep. 2004.

[7] R. G. Lorenz and S. P. Boyd, "Robust minimum variance beamforming," *IEEE Trans. Signal Process.*, vol. 53, pp. 1684–1696, May 2005.

[8] S. A. Vorobyov, H. Chen, and A. B. Gershman, "On the relationship between robust minimum variance beamformers with probabilistic and worst-case distortionless response constraints," *IEEE Trans. Signal Process.*, vol. 56, pp. 5719–5724, Nov. 2008.

[9] A. Hassanien, S. A. Vorobyov, and K. M. Wong, "Robust adaptive beamforming using sequential programming: An iterative solution to the mismatch problem," *IEEE Signal Process. Lett.*, vol. 15, pp. 733–736, 2008.

[10] A. Khabbazibasmenj, S. A. Vorobyov, and A. Hassanien, "Robust adaptive beamforming via estimating steering vector based on semidefinite relaxation," in *Proc. 44th Ann. Asilomar Conf. Signals, Syst., Comput.*, Pacific Grove, CA, Nov. 2010, pp. 1102–1106.

[11] A. Khabbazibasmenj, S. A. Vorobyov, and A. Hassanien, "Robust adaptive beamforming based on steering vector estimation with as little as possible prior information," *IEEE Trans. Signal Process.*, vol. 60, pp. 2974–2987, Jun. 2012.

[12] S. Shahbazpanahi, A. B. Gershman, Z.-Q. Luo, and K. M. Wong, "Robust adaptive beamforming for general rank signal models," *IEEE Trans. Signal Process.*, vol. 51, pp. 2257–2269, Sep. 2003.

[13] H. H. Chen and A. B. Gershman, "Robust adaptive beamforming for general-rank signal models with positive semi-definite constraints," in *Proc. IEEE ICASSP*, Las Vegas, NV, Apr. 2008, pp. 2341–2344.

[14] C. W. Xing, S. D. Ma, and Y. C. Wu, "On low complexity robust beamforming with positive semidefinite constraints," *IEEE Trans. Signal Process.*, vol. 57, pp. 4942–4945, Dec. 2009.

[15] H. H. Chen and A. B. Gershman, "Worst-case based robust adaptive beamforming for general-rank signal models using positive semi-definite covariance constraint," in *Proc. IEEE ICASSP*, Prague, Czech Republic, May 2011, pp. 2628–2631.

[16] A. Khabbazibasmenj and S. A. Vorobyov, "A computationally efficient robust adaptive beamforming for general-rank signal models with positive semi-definiteness constraint," in *Proc. IEEE CAMSAP*, San Juan, Puerto Rico, Dec. 2011, pp. 185–188.

[17] A. Khabbazibasmenj, S. A. Vorobyov, F. Roemer, and M. Haardt, "Polynomial-time DC (POTDC) for sum-rate maximization in two-way AF MIMO relaying," in *Proc. IEEE ICASSP*, Kyoto, Japan, Mar. 2012, pp. 2889–2892.
[18] A. Khabbazibasmenj, F. Roemer, S. A. Vorobyov, and M. Haardt, "Sum-rate maximization in two-way AF MIMO relaying: Polynomial time solutions to a class of DC programming problems," *IEEE Trans. Signal Process.*, vol. 60, no. 10, pp. 5478–5493, Oct. 2012.
[19] A. Pezeshki, B. D. Van Veen, L. L. Scharf, H. Cox, and M. Lundberg, "Eigenvalue beamforming using a multi-rank MVDR beamformer and subspace selection," *IEEE Trans. Signal Process.*, vol. 56, pp. 1954–1967, May 2008.
[20] R. Horst, P. M. Pardalos, and N. V. Thoai, *Introduction to Global Optimization*. Dordrecht, The Netherlands: Kluwer Academic, 1995.
[21] R. Horst and H. Tuy, *Global Optimization: Deterministic Approaches*. New York, NY, USA: Springer, 1996.
[22] J. Zhang, F. Roemer, M. Haardt, A. Khabbazibasmenj, and S. A. Vorobyov, "Sum rate maximization for multi-pair two-way relaying with single-antenna amplify and forward relays," in *Proc. ICASSP*, Kyoto, Japan, Mar. 2012, pp. 2477–2480.
[23] A. L. Yuille and A. Rangarajan, "The concave-convex procedure," *Neural Comput.*, vol. 15, pp. 915–936, 2003.
[24] S. Haykin, *Adaptive Filter Theory*, 3rd ed. Englewood Cliffs, NJ, USA: Prentice-Hall, 1995.
[25] A. Beck and Y. C. Eldar, "Strong duality in nonconvex quadratic optimization with two quadratic constraints," *SIAM J. Optimiz.*, vol. 17, no. 3, p. 844860, 2006.
[26] Y. Huang and D. P. Palomar, "Rank-constrained separable semidefinite programming with applications to optimal beamforming," *IEEE Trans. Signal Process.*, vol. 58, pp. 664–678, Feb. 2010.
[27] Y. Ye and S. Zhang, "New results on quadratic minimization," *SIAM J. Optimiz.*, vol. 14, pp. 245–267, 2003.
[28] W. Ai, Y. Huang, and S. Zhang, "New results on Hermitian matrix rank-one decomposition," *Math. Program.*, pp. 253–383, Aug. 2009.
[29] M. S. Bazaraa, C. M. Shetty, and H. D. Sherali, *Nonlinear Programming: Theory and Algorithms*, 2nd ed. New York, NY, USA: Wiley, 1993.
[30] Y. Nesterov and A. Nemirovsky, *Interior Point Polynomial Algorithms in Convex Programming*. Philadelphia, PA, USA: SIAM, 1994.
[31] Z.-Q. Luo, W.-K. Ma, A. M.-C. So, Y. Ye, and S. Zhang, "Semidefinite relaxation of quadratic optimization problems," *IEEE Signal Process. Mag.*, vol. 27, no. 3, pp. 20–34, May 2010.
[32] K. T. Phan, S. A. Vorobyov, N. D. Sidiropoulos, and C. Tellambura, "Spectrum sharing in wireless networks via QoS-aware secondary multicast beamforming," *IEEE Trans. Signal Process.*, vol. 57, pp. 2323–2335, Jun. 2009.
[33] Z.-Q. Luo, N. D. Sidiropoulos, P. Tseng, and S. Zhang, "Approximation bounds for quadratic optimization with homogeneous quadratic constraints," *SIAM J. Optimiz.*, vol. 18, no. 1, pp. 1–28, Feb. 2007.
[34] C. Udriste, "Convex functions and optimization methods on Riemannian manifolds," in *Series: Mathematics and Its Applications*. New York, NY, USA: Springer, 1994, vol. 297.
[35] A. Shapiro, "First and second order analysis of nonlinear semidefinite programs," *Math. Programm. Ser.B*, vol. 77, pp. 301–320, 1997.
[36] A. Beck, A. Ben-Tal, and L. Tetruashvili, "A sequential parametric convex approximation method with applications to nonconvex Truss topology design problems," *J. Global Optimiz.*, vol. 47, no. 1, pp. 29–51, 2010.
[37] D. Goldfarb and K. Scheinberg, "On parametric semidefinite programming," *Appl. Numer. Math.*, vol. 29, no. 3, pp. 361–377, 1999.

Arash Khabbazibasmenj (S'08) received the B.Sc. and M.Sc. degrees in electrical engineering (communications) from Amirkabir University of Technology, Tehran, Iran, and the University of Tehran, Tehran, Iran, in 2006 and 2009, respectively, and the Ph.D. degree in electrical engineering (signal processing) from the University of Alberta, Edmonton, Alberta, Canada, in 2013.

He is currently working as a Postdoctoral Fellow in the Department of Electrical and Computer Engineering, University of Alberta. During spring and summer 2011, he was also a visiting student at Ilmenau University of Technology, Germany. His research interests include signal processing and optimization methods in radar, communications, and related fields.

Dr. Khabbazibasmenj is a recipient of the Alberta Innovates Graduate Award in ICT.

Sergiy A. Vorobyov (M'02–SM'05) received the M.Sc. and Ph.D. degrees in systems and control from Kharkiv National University of Radio Electronics, Ukraine, in 1994 and 1997, respectively.

He is a Professor with the Department of Signal Processing and Acoustics, Aalto University, Finland, and is currently on leave from the Department of Electrical and Computer Engineering, University of Alberta, Edmonton, Canada. He has been with the University of Alberta as an Assistant Professor from 2006 to 2010, Associate Professor from 2010 to 2012, and Full Professor since 2012. Since his graduation, he also held various research and faculty positions at Kharkiv National University of Radio Electronics, Ukraine; the Institute of Physical and Chemical Research (RIKEN), Japan; McMaster University, Canada; Duisburg-Essen University and Darmstadt University of Technology, Germany; and the Joint Research Institute between Heriot-Watt University and Edinburgh University, U.K. He has also held short-term visiting positions at Technion, Haifa, Israel and Ilmenau University of Technology, Ilmenau, Germany. His research interests include statistical and array signal processing, applications of linear algebra, optimization, and game theory methods in signal processing and communications, estimation, detection and sampling theories, and cognitive systems.

Dr. Vorobyov is a recipient of the 2004 IEEE Signal Processing Society Best Paper Award, the 2007 Alberta Ingenuity New Faculty Award, the 2011 Carl Zeiss Award (Germany), the 2012 NSERC Discovery Accelerator Award, and other awards. He served as an Associate Editor for the IEEE TRANSACTIONS ON SIGNAL PROCESSING from 2006 to 2010 and for the IEEE SIGNAL PROCESSING LETTERS from 2007 to 2009. He was a member of the Sensor Array and Multi-Channel Signal Processing Committee of the IEEE Signal Processing Society from 2007 to 2012. He is a member of the Signal Processing for Communications and Networking Committee since 2010. He has served as the Track Chair for Asilomar 2011, Pacific Grove, CA, the Technical Co-Chair for IEEE CAMSAP 2011, Puerto Rico, and the Tutorial Chair for ISWCS 2013, Ilmenau, Germany.