Robust Adaptive Beamforming Based on Steering Vector Estimation With as Little as Possible Prior Information

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

Abstract—A general notion of robustness for robust adaptive beamforming (RAB) problem and a unified principle for minimum variance distortionless response (MVDR) RAB techniques design are formulated. This principle is to use standard MVDR beamformer in tandem with an estimate of the desired signal steering vector found based on some imprecise prior information. Differences between various MVDR RAB techniques occur only because of the differences in the assumed prior information and the corresponding signal steering vector estimation techniques. A new MVDR RAB technique, which uses as little as possible and easy to obtain imprecise prior information, is developed. The objective for estimating the steering vector is the maximization of the beamformer output power, while the constraints are the normalization condition and the requirement that the estimate does not converge to any of the interference steering vectors and their linear combinations. The prior information used is only the imprecise knowledge of the antenna array geometry and angular sector in which the actual steering vector lies. Mathematically, the proposed MVDR RAB is expressed as the well known non-convex quadratically constrained quadratic programming problem with two constraints, which can be efficiently and exactly solved. Some new results for the corresponding optimization problem such as a new algebraic way of finding the rank-one solution from the general-rank solution of the relaxed problem is guaranteed to be rank-one are derived. Our simulation results demonstrate the superiority of the proposed method over other previously developed RAB techniques.

Index Terms—Output power maximization, robust adaptive beamforming (RAB), quadratically constrained quadratic programming (QCQP), semi-definite programming (SDP) relaxation, steering vector estimation.

I. INTRODUCTION

Robust adaptive beamforming (RAB) has been an intensive research topic over several decades due to, on one hand, its importance in wireless communications, radar, sonar, microphone array speech processing, radioastronomy, medical imaging, and other fields; and on the other hand, because of the challenges related to the practical applications manifesting themselves in the robustness requirements. The presence of the desired signal component in the training data, small sample size, and imprecise knowledge of the desired signal steering vector are the main causes of performance degradation in adaptive beamforming. The traditional design approaches to adaptive beamforming [1]–[4] do not provide sufficient robustness and are not applicable in such situations. Thus, various RAB techniques have been developed [5]. Some examples of popular conventional RAB approaches are the diagonal loading technique [6], [7], the projection beamforming techniques [4], [8], and the eigenspace-based beamforming technique [9]. The disadvantages of these approaches such as the ad hoc nature of the former one and high probability of subspace swap at low signal-to-noise ratios (SNRs) for the latter one [10] are well known.

Among more recent RAB techniques based on minimum variance distortionless response (MVDR) principle are i) the worst-case-based adaptive beamforming technique proposed in [11] and [12] and further developed in [13]–[15]; ii) the doubly constrained robust Capon beamforming method [16], [17] (it is based on the same idea of the worst-case performance optimization as i); iii) the probabilistically constrained RAB technique [18]; iv) the RAB technique based on steering vector estimation [19]; and others. Although the relationships between different RAB MVDR techniques have been established,1 the general notion of robustness and unified principle for all RAB MVDR techniques have been missing.

In this paper, we rethink the notion of robustness and present a unified principle to MVDR RAB design,2 that is, to use standard MVDR beamformer in tandem with steering vector estimation based on some prior information and data covariance matrix estimation. This unified principle motivates us to develop a new technique which uses as little as possible, imprecise, and easy to obtain prior information about the desired

1The worst-case-based design is related to the diagonal loading principle [11], [13], while the probabilistically constrained design can be approximated by the worst-cases-based design [18]. Moreover, the worst-cases-based and doubly-constrained RAB techniques can be derived via steering vector estimation approach [13], [16].

2Some preliminary results have been presented in [20].
signal/source, the antenna array, and the propagation media. We develop such a new RAB technique in which the steering vector is estimated through the beamformer output power maximization under the requirement that the estimate does not converge to any of the interference steering vectors and their linear combinations. The only prior information used is the imprecise knowledge of the angular sector of the desired signal and antenna array geometry, while the knowledge of the presumed steering vector is not needed. Such MVDR RAB technique can be mathematically formulated as a non-convex (due to an additional steering vector normalization condition) quadratically constrained quadratic programming (QCQP) problem, which is known in the literature for decades [21]–[26] and which can be solved using the semi-definite programming (SDP) relaxation technique. Moreover, our specific optimization problem allows for an exact solution using, for example, the duality theory [17], [27] or the iterative rank reduction technique [28]. In the optimization context, we develop several new results when answering the questions of i) how to obtain a rank-one solution from a general-rank solution of the relaxed problem algebraically and ii) when it is guaranteed that the solution of the relaxed problem is rank-one. The latter question, for example, is important because it had been observed that the probability of obtaining a rank-one solution for the class of problems similar to the one considered in this paper is close to one, while the theoretical upper bound suggests a significantly smaller probability [24], [26]. Our result proves the correctness of the experimental observations about the high probability of a rank-one solution for the relaxed problem.

This paper is organized as follows. A general notion of robustness and a unified principle for MVDR RAB design are given in Section II and the existing MVDR RAB techniques are summarized and shown to satisfy the general principle under the requirement that the estimate does not converge to any of the interference steering vectors and their linear combinations. In Section III, we formulate a new MVDR RAB technique which uses as little as possible, imprecise, and easy to obtain prior information. An analysis of the performance of the problem as well as some new optimization related results are shown in Section IV. Simulation results comparing the performance of the proposed method to the existing methods are shown in Section V. Finally, Section VI presents our conclusions. This paper is reproducible research [29], and the software needed to generate the simulation results can be obtained from IEEE Xplore together with the paper.

II. PRELIMINARIES, ROBUSTNESS, AND THE UNIFIED PRINCIPLE TO MVDR RAB DESIGN

Consider a linear antenna array with \( M \) omni-directional antenna elements. The narrowband signal received by the antenna array at the time instant \( k \) can be written as

\[
x(k) = s(k) + i(k) + n(k)
\]  

where \( s(k) \), \( i(k) \), and \( n(k) \) denote the \( M \times 1 \) vectors of the desired signal, interference, and noise, respectively. The desired signal, interference, and noise components of the received signal (1) are assumed to be statistically independent to each other. The desired signal can be written as \( s(k) = s(k)a \), where \( s(k) \) is the signal waveform and \( a \) is the associated steering vector.

The beamformer output at the time instant \( k \) can be written as

\[
y(k) = w^H x(k)
\]  

where \( w \) is the \( M \times 1 \) complex weight (beamforming) vector of the antenna array and \( (\cdot)^H \) stands for the Hermitian transpose.

Assuming that the steering vector \( a \) is known precisely, the optimal weight vector can be obtained by maximizing the beamformer output signal-to-noise-plus-interference ratio (SINR) [1]

\[
\text{SINR} = \frac{\sigma_s^2 |w^H a|^2}{w^H R_{i+n} w}
\]  

where \( \sigma_s^2 \) is the desired signal power, \( R_{i+n} \triangleq E\{i(k) + n(k))(i(k)^H + n(k)^H)\} \) is the \( M \times M \) interference-plus-noise covariance matrix, and \( E\{\cdot\} \) stands for the statistical expectation. Since \( R_{i+n} \) is unknown in practice, it is substituted in (3) by the data sample covariance matrix

\[
\hat{R} \triangleq \frac{1}{K} \sum_{i=1}^{K} x(i)x^H(i)
\]  

where \( K \) is the number of training data samples which also include the desired signal component. Note that better estimates than (4) can be used [30], but this straightforward possibility is not discussed in this paper.

The problem of maximizing (3), where the sample estimate (4) is used instead of \( R_{i+n} \), is known as the MVDR sample matrix inversion (SMI) beamforming and it is mathematically equivalent to the following convex optimization problem:

\[
\min_w w^H \hat{R} w \quad \text{subject to} \quad w^H a = 1.
\]  

The solution of (5) can be easily found as \( w_{\text{MVDR SMI}} = \alpha \hat{R}^{-1} a \), where \( \alpha = 1/\hat{a}^H \hat{R}^{-1} a \) [1].

The MVDR-SMI beamformer is known to be not robust to any imperfect knowledge of the desired signal steering vector. Different RAB techniques have been developed which use different specific notions of robustness. However, the general meaning of robustness for any RAB technique is the ability to compute the beamforming vector so that the SINR is maximized despite possibly imperfect and little knowledge of prior information. Specifically, the main cause of performance degradation of MVDR-SMI beamformer is the situation when the desired signal steering vector is mixed with any of the interference steering vectors. Thus, if with little and imperfect prior information, an adaptive beamforming technique is able to estimate the desired signal steering vector so that it does not converge to any of the interferences and their linear combinations, such technique is called robust. Using this notion of robustness, the unified principle to MVDR RAB design can be formulated as follows. Use the standard MVDR-SMI beamformer (5) in tandem with steering vector estimation based on some prior information. Difference between various MVDR RAB techniques can be then shown to boil down to the differences in the assumed prior information, the specific notions of robustness used, and the corresponding steering vector estimation techniques used.

In the case of steering vector mismatch \( \delta \), the solution of (5) can be written as a function of unknown \( \delta \), that is, \( w(\delta) = \alpha \hat{R}^{-1} a = \alpha \hat{R}^{-1}(p + \delta) \), where \( a \) is the actual steering vector.
and \( p \) is the presumed one. Thus, the beamformer output power can also be written as a function of \( \delta \) as

\[
P(\delta) = \frac{1}{(p + \delta)^H R^{-1}(p + \delta)}
\]  

(6)

Then, the best estimate of \( \delta \), denoted as \( \hat{\delta} \) (the estimate of \( \alpha \) is \( \hat{\alpha} = p + \hat{\delta} \)), is the one which maximizes (6) under the constraint that \( \hat{\alpha} \) does not converge to any interference steering vectors and their linear combinations. In [19], for example, such convergence is avoided by means of the constraint \( P^{-1}(p + \delta) - p^{-1} \hat{\alpha} = 0 \), where \( P^{-1} \triangleq I - LL^H \), \( L \triangleq [l_1, l_2, \ldots, l_l] \), \( l_i, i = 1, \ldots, L \) are the \( L \) dominant eigenvectors of the matrix \( C \triangleq \Sigma_{\alpha} \Sigma_{\theta} \Sigma_{\alpha}^H \) which is the steering vector associated with direction \( \theta \) that has the structure defined by the antenna array geometry, \( \Theta \) is the angular sector in which the desired signal is located, and \( I \) is the identity matrix.

The steering vector estimation in [19] is based on splitting the mismatch vector \( \delta \) into the orthogonal component to the presumed steering vector \( p \) and the parallel one, i.e., \( \delta = \delta_\perp + \delta_\parallel \). The estimates \( \hat{\delta} \) and \( \hat{\alpha} \) can be found iteratively by solving the following convex optimization problem for \( \delta_\perp \):

\[
\min_{\delta_\perp} \langle p + \delta_\perp \rangle^H \hat{R}^{-1} \langle p + \delta_\perp \rangle \quad \text{subject to} \quad P^{-1}(p + \delta_\parallel) = 0, \quad P^H \delta_\parallel = 0
\]

(7)

\[
\|p + \delta_\parallel\|^2 \leq M
\]

(8)

\[
\langle p + \delta_\parallel \rangle^H C \langle p + \delta_\parallel \rangle \leq P^H Cp
\]

(9)

where \( C \triangleq \int_{\Omega} d(\theta) d\Omega(\theta) d\theta \) and \( \Omega \) is the complement of the sector \( \Theta \). The constraint (10) limits the noise power collected in \( \Theta \), while the orthogonality between \( \delta_\perp \) and \( p \) is imposed by \( p^H \delta_\perp = 0 \).

The prior information used in this approach is the presumed steering vector, the angular sector \( \Theta \), and the antenna array geometry knowledge. The general notion of robustness mentioned above is adopted for this approach. This technique can be further simplified for more structured uncertainties, for example, when it is known that the array is partially calibrated [31]. However, the amount of prior information about the uncertainty type and structure then increases. The disadvantages of this method are the high computational complexity and the fact that \( I \) has to be known precisely, which in practice is not known.

Other well-known MVDR RAB techniques are summarized in Table I where the corresponding notions of robustness and prior information used by the techniques are listed. In this table, the multi-rank beamformer matrix of the eigenvalue beamforming method of [32] is computed as

\[
W = R^{-1} \Psi H R^{-1} \Psi^{-1} Q, \quad \text{where} \quad Q \text{ is a data dependent left-orthogonal matrix, i.e.,} \quad Q H Q = I, \quad \Psi \text{ is the linear subspace in which the desired signal is located, and} \quad \text{Tr}(\cdot) \text{ denotes the trace of a matrix. For resolving a signal with a rank-one covariance matrix and an unknown but fixed angle of arrival, the columns of} \quad Q \text{ should be selected as the dominant eigenvectors of the error covariance matrix, i.e.,} \quad \Phi_e = (\Psi H R^{-1} \Psi)^{-1}.
\]

It is assumed that the signal lies in a known subspace, but the angle of arrival is unknown and unfixed (for example, it randomly changes from snapshot to snapshot), it is the subdominant eigenvectors of the error covariance matrix that should be used as the columns of the matrix \( Q \).

It can be seen from Table I that the main problem for any MVDR RAB technique is to estimate the steering vector, while avoiding its convergence to any of the interferences and their linear combinations. It is achieved in different techniques by exploiting different prior information and solving different optimization problems. The complexity of the corresponding steering vector estimation problems can vary from the complexity of eigenvalue decomposition to the complexity of solving QCQP programming problem. All known MVDR RAB techniques require the knowledge of the presumed steering vector that, in turn, implies that the antenna array geometry, propagation media, and desired source characteristics such as the presumed angle of arrival are known. Therefore, it is of great importance to develop an MVDR RAB technique that requires as little as possible and easy to obtain prior information.

III. NEW BEAMFORMING PROBLEM FORMULATION

For estimating the actual steering vector \( \alpha \), we first observe that the maximization of the output power (6) is equivalent to the minimization of the denominator of (6). One obvious constraint that must be imposed on the estimate \( \hat{\alpha} \) is the norm constraint \( \|\hat{\alpha}\|^2 = M \). To avoid the convergence of the estimate \( \hat{\alpha} \) to any of the interference steering vectors and their linear combinations, we introduce a new constraint in what follows.

In order to establish such a new constraint, we assume that the desired source is located in the known angular sector of \( \Theta = [\theta_{\min}, \theta_{\max}] \) which can be obtained, for example, using low resolution direction finding methods. This angular sector is assumed to be distinguishable from general locations of the interfering signals. Let \( C = UC \) denote the eigenvalue decomposition of the matrix \( C \), where \( U = \text{unitary and diagonal matrices, respectively. Column} \ i \text{ of the unitary matrix} \ U \text{ denoted as} \ u_i \text{ and the} \ i \text{th diagonal element of the diagonal matrix} \ \Lambda \text{ denoted as} \ \lambda_i \text{ are, respectively, the} \ i \text{th eigenvector and the} \ i \text{th eigenvalue of} \ C \). It is assumed without loss of generality that the eigenvalues \( \lambda_i, i = 1, \ldots, M \) are ordered in the descending order, i.e., \( \lambda_i \geq \lambda_{i+1}, i = 1, \ldots, M - 1 \). By splitting matrix \( U \) to the \( M \times K \) matrix \( U_1 \) and the \( M \times M - K \) matrix \( U_2 \) as \( U = [U_1 \ U_2] \), the matrix \( C \) can be decomposed as \( C = U_1 \Lambda_1 U_1^H + U_2 \Lambda_2 U_2^H \), where the \( K \times K \) diagonal matrix \( \Lambda_1 \) contains the \( K \) dominant eigenvalues, while the other \( M - K \times M - K \) diagonal matrix \( \Lambda_2 \) contains the \( M - K \) subdominant eigenvalues. Since the matrix \( C \) is computed by the integration over the complement of the desired sector, it can be concluded that for the properly chosen \( K \), the steering vector in the desired sector and its complement can be approximately expressed as linear combinations of the columns of \( U_2 \) and \( U_1 \), respectively, that is,

\[
d(\theta) \approx U_2 v_2, \quad \theta \in \Theta
\]

(11)

\[
d(\theta) \approx U_1 v_1, \quad \theta \in \Theta
\]

(12)

where \( v_1 \) and \( v_2 \) are some coefficient vectors.

Example I: As an illustrative example, let us consider the squared norm of the vectors that are obtained by projecting the vector \( \hat{d}(\theta) \) onto the subspaces spanned by \( U_1 \) and \( U_2 \). Fig. 1 depicts such squared norms, i.e.,

\[
||U_1 \hat{d}(\theta)||^2 = \|U_1 U_1^H d(\theta)\|^2 \quad \text{and} \quad ||U_2 \hat{d}(\theta)||^2 = \|U_2 U_2^H d(\theta)\|^2
\]

versus \( \theta \). The example set up is the following. The angular sector is \( \Theta = [-35^\circ, -15^\circ] \), \( K = 5 \), and \( M = 20 \). It can be observed
from the figure that the approximations (11) and (12) are accurate and the squared norms of the projections of the vector \( \mathbf{d}(\theta) \) on \( \mathbf{U}_2 \) and \( \mathbf{U}_1 \), i.e., inside and outside of the desired sector, respectively, are almost equal to \( M \).

Using (11) and (12), we know that \( ||\mathbf{v}_1||^2 = M \) and \( ||\mathbf{v}_2||^2 = M \), and the following equations are in order:

\[
\mathbf{d}^H(\theta)\mathbf{C}\mathbf{d}(\theta) \cong (\mathbf{U}_1 \mathbf{v}_1)^H \mathbf{C}(\mathbf{U}_1 \mathbf{v}_1) = \mathbf{v}_1^H \mathbf{A}_1 \mathbf{v}_1, \quad \theta \in \Theta
\]

\[
\mathbf{d}^H(\theta)\mathbf{C}\mathbf{d}(\theta) \cong (\mathbf{U}_2 \mathbf{v}_2)^H \mathbf{C}(\mathbf{U}_2 \mathbf{v}_2) = \mathbf{v}_2^H \mathbf{A}_2 \mathbf{v}_2, \quad \theta \in \Theta.
\]

Note that since \( \mathbf{A}_1 \) is the diagonal matrix of \( K \) dominant eigenvalues of \( \mathbf{C} \) and \( \mathbf{A}_2 \) contains the remaining eigenvalues, the quadratic form \( \mathbf{d}^H(\theta)\tilde{\mathbf{C}}\mathbf{d}(\theta) \) takes larger values outside of the desired sector. Based on this observation, the estimate \( \mathbf{\hat{a}} \) can be forced not to converge to any vector with direction located within the complement of \( \Theta \) including the interference steering vectors and their linear combinations by means of the following new constraint:

\[
\mathbf{\hat{a}}^H \mathbf{C}\mathbf{\hat{a}} \leq \Delta_0 \quad (15)
\]

where \( \Delta_0 \) is a uniquely selected value for a given angular sector \( \Theta \), that is,

\[
\Delta_0 = \max_{\theta \in \Theta} \mathbf{d}^H(\theta)\tilde{\mathbf{C}}\mathbf{d}(\theta).
\]
Using the definition of $\Delta_0$ (16) together with (14), we can find that

$$\Delta_0 = \max_{\theta \in \Theta} \mathbf{v}(\theta)^H \mathbf{A}_2 \mathbf{v}(\theta) \leq M \lambda_{K+1}. \quad (17)$$

It is interesting that as compared to [19], where the number of $K$ is required, the proposed constraint (15) does not depend on $K$. In order to explain how the constraint (15) avoids the convergence of the steering vector estimate to any linear combination of the steering vectors of the interferences, let $\mathbf{A} \triangleq [\mathbf{a}_1, \ldots, \mathbf{a}_J]$ denote a set of $J$ plane wave steering vectors corresponding to the interferences that are located outside of the desired sector. Using (11) and (12), these steering vectors can be approximated as

$$\mathbf{a}_l = \mathbf{U}_1 \mathbf{r}_l, \quad j = 1, \ldots, J$$

where $\mathbf{r}_l, \quad j = 1, \ldots, J$ are some $K \times 1$ coefficient vectors. Then, every linear combination of these steering vectors which has norm squared equal to $M$ can be expressed as

$$\mathbf{f} = \mathbf{A} \mathbf{z} = \mathbf{U}_2 \mathbf{z} \quad (18)$$

where $\mathbf{z}$ is a $J \times 1$ coefficient vector, $\mathbf{z} = [\mathbf{r}_1, \ldots, \mathbf{r}_J] \cdot \mathbf{v}$, $[\mathbf{r}_1, \ldots, \mathbf{r}_J]$ is $K \times J$ matrix, and $\|\mathbf{z}\|^2 = ||\mathbf{a}||^2 \triangleq M$. Then, the following is true:

$$\mathbf{f}^H \mathbf{C} \mathbf{f} = (\mathbf{U}_1 \mathbf{z})^H \mathbf{C} \mathbf{U}_1 \mathbf{z} = \mathbf{z}^H \mathbf{A}_2 \mathbf{z} \geq M \lambda_K \geq M \lambda_{K+1} \geq \Delta_0 \quad (19)$$

where the last inequality follows from (17). In fact, (19) implies that $\mathbf{f}$ obtained as a linear combination of all interference steering vectors does not satisfy the constraint (15). Thus, (15) prevents the convergence of the estimate $\hat{\mathbf{a}}$ to any of the interference steering vectors or their linear combination.

It is worth stressing that no restrictions/assumptions on the structure of the interferences are needed. Moreover, the interferences do not need to have the same structure as the desired signal. The constraint (15) does not use any of such information. Indeed, interferences may have a rank-one or multi-rank covariance matrix. Specifically, an interference source with a rank-one covariance matrix corresponds to either a plane wave source or a locally coherent scattered source [33]. An interference source with a multi-rank covariance matrix corresponds, for example, to a locally incoherent scattered source [32].

Steering vector of an interference with a locally coherent scattered source can be expressed as [33]

$$\hat{\mathbf{a}}_i = \mathbf{a}_i + \sum_{l=1}^{T} e^{j \psi_l} \mathbf{b}(\theta_l) \quad (20)$$

where $\mathbf{a}_i$ corresponds to the direct path of the interference and $\mathbf{b}(\theta_l), \quad l = 1, \ldots, T$ correspond to the coherently scattered paths. Here $\mathbf{b}(\theta_l)$ is a plane wave impinging on the array from the direction $\theta_l$ which is fixed for different snapshots. The parameters $\psi_l \in [0, 2\pi], \quad l = 1, \ldots, T$ denote the phase shift of different paths that are also fixed for different snapshots. Thus, the spatial signature of a locally coherent scattered source (20) is fixed over different snapshots and is a linear combination of plane wave steering vectors. A normalized linear combination of plane wave steering vectors that all lie outside of the desired sector does not satisfy the quadratic constraint (19). Thus, the interferences with rank-one covariance matrix will be avoided by means of the new constraint (15).

Steering vector of a locally incoherent scattered source is time-varying and can be modeled as [32]

$$\hat{\mathbf{a}}_i(k) = s_0(k) \mathbf{a}_i + \sum_{l=1}^{T} s_l(k) \mathbf{b}(\theta_l) \quad (21)$$

where $s_l(k), \quad l = 0, \ldots, T$ are independently identically distributed (i.i.d.) zero-mean random variables with variances $\nu_l, \quad l = 0, \ldots, T$. However, the random variables $s_l(k)$ change from snapshot to snapshot. The correlation matrix of a locally incoherent scattered source (21) can be written as

$$\mathbf{R}_i = \nu_0 \mathbf{a}_i \mathbf{a}_i^H + \sum_{l=1}^{T} \nu_l \mathbf{b}(\theta_l) \mathbf{b}(\theta_l)^H. \quad (22)$$

Based on (22), it can be concluded that a locally incoherent scattered source is equivalent to $T+1$ independent plane wave sources that impinge on the array from $T+1$ different directions. Thus, as long as all such plane wave sources lie outside of the desired sector $\Theta$, they do not satisfy the quadratic constraint (15) and the interferences with multi-rank covariance matrix will be avoided as well by means of the constraint (15).

To further illustrate how the constraint (15) works, let us consider the following example.

Example 2: Consider uniform linear array (ULA) of 10 omni-directional antenna elements spaced half wavelength apart from each other. Let the range of the desired signal angular locations be $\Theta = [0^\circ, 10^\circ]$. Fig. 2 depicts the values of the quadratic term $d^H(\theta) \mathbf{C} d(\theta)$ for different angles. The rectangular bar in the figure marks the directions within the presumed angular sector $\Theta$. It can be observed from this figure that the term $d^H(\theta) \mathbf{C} d(\theta)$ has the smallest values within the angular sector $\Theta$ and increases outside of the sector. Therefore, if $\Delta_0$ is selected to be equal to the maximum value of the term $d^H(\theta) \mathbf{C} d(\theta)$ within the presumed angular sector $\Theta$, the constraint (15) guarantees that the estimate of the desired signal steering vector does not converge to any of the interference steering vectors and their linear combinations. It is worth noting that $d^H(\theta) \mathbf{C} d(\theta) = \Delta_0$ must occur at one of the edges of $\Theta$. However, the value of the quadratic
Fig. 2. Values of the term $d^H(\theta)\hat{C}d(\theta)$ in the constraint (15) for different angles.

Fig. 3. Comparison between the quadratic term with and without array perturbations.

term at the other edge of $\Theta$ may be smaller than $\Delta_0$. Therefore, we define another sector $\Theta_n \supset \Theta$ at which the equality $d^H(\Theta_n)C(\theta) = \Delta_0$ holds at both edges, i.e., the sector $\Theta_n$ is the actual sector at which the constraint (15) must be satisfied.

In order to compute the matrix $C$, the presumed knowledge of the antenna array geometry is required. Due to the imperfect array calibration, the precise knowledge of the antenna array geometry may be unavailable. If array perturbation is present, the curve for the quadratic form $d^H(\Theta_n)\hat{C}d(\theta)$ versus $\theta$ can deviate from the one drawn under the assumption of no array perturbation. This situation is demonstrated in Fig. 3, which depicts the quadratic term $d^H(\Theta_n)\hat{C}d(\theta)$ versus $\theta$ in the presence and in the absence of array perturbations. Although the angular sector $\Theta_n$ computed for a given $\Delta_0$ using (16) under the assumption of no array perturbation may change if the array is perturbed, the constraint (15) still remains precise as long as $\Theta_n$ contains the desired signal and does not contain any interfering sources. Therefore, an inaccurate information about the antenna array geometry is sufficiently good. In fact, in one of our simulation examples we show that even if the perturbations of the antenna array geometry are the highest possible, the constraint (15) remains precise.

Taking into account the normalization constraint and the constraint (15), the problem of estimating the desired signal steering vector based on the knowledge of the sector $\Theta$ can be formulated as the following optimization problem:

$$\min_{\hat{a}} \quad d^H(\Theta)\hat{C}d(\theta)$$
subject to $$\|\hat{a}\|^2 = M$$
$$\hat{a}^H\hat{C}\hat{a} \leq \Delta_0.$$  \hfill (23)
 \hfill (24)
 \hfill (25)

Unlike the constraint (10) used in the problem (7)–(10) in order to avoid the noise power magnification, the constraint (25) is enforcing $\hat{a}$ not to converge to any steering vector associated with any of the interferences and their linear combinations. Compared to the other MVDR RAB methods, which require the knowledge of the presumed steering vector and, thus, the knowledge of the presumed antenna array geometry, propagation media, and source characteristics, only imprecise knowledge of the antenna array geometry and approximate knowledge of the angular sector $\Theta$ are needed for the proposed method.

Due to the non-convex equality constraint (24), the QCQP problem (23)–(25) is non-convex and NP-hard in general. However, such problems are well known for decades and many results for them are available - the most recent ones are in [17], [27], [28] where it has been shown that the strong duality holds for the problems of type (23)–(25) and, thus, the solution based on semi-definite programming (SDP) relaxation is the exact one. In the following section we develop some new results regarding this problem while looking for a new algebraic way of finding the rank-one solution from the general-rank solution of the relaxed problem. We also obtain the condition under which the solution of the relaxed problem is guaranteed to be rank-one.

It is interesting that the steering vector estimation problem in [19] can be also expressed as a QCQP problem that makes it possible to find a much simpler solution than the sequential quadratic programming of [19] and draw some insightful connections to the newly proposed problem (23)–(25). Let us first find the set of vectors satisfying the constraint $P^\dagger\hat{a} = 0$. Note that $P^\dagger\hat{a} = 0$ implies that $\hat{a} = LL^H\hat{a}$ and, therefore, we can write that

$$\hat{a} = Lb$$  \hfill (26)

where $b$ is an $L \times 1$ complex valued vector. Using (26), the optimization problem for estimating the steering vector in [19] can be equivalently rewritten in terms of $b$ as

$$\min \quad b^H L^H\hat{R}^{-1}Lb$$
subject to $$\|b\|^2 = M$$
$$b^H L^H\hat{C}Lb \leq p^H\hat{C}p$$  \hfill (27)
 \hfill (28)
 \hfill (29)

which is a QCQP problem. Thus, as compared to (27)–(29), where the constraint $P^\dagger\hat{a} = 0$ enforces the estimated steering vector to be a linear combination of $L$ dominant eigenvectors $\{\lambda_1, \lambda_2, \ldots, \lambda_L\}$, the steering vector in (23)–(25) is not restricted by such requirement, while the convergence to any of the interference steering vectors and their linear combinations is avoided by means of the constraint (15). As a result, the problem (23)–(25) has more degrees of freedom. Thus, it is
expected that the new RAB method will outperform the one of [19].

As it has been explained in [19], the solution of the problem (7)–(10) leads to a better performance for the corresponding RAB compared to the other RAB techniques and, particularly, the worst-case-based and probabilistically constrained techniques. This performance improvement is the result of forming the beam toward a single corrected steering vector yielding maximum output power, while the worst-case-based method maximizes the output power for all steering vectors in its corresponding uncertainty set. Thus, despite a significantly more relaxed assumptions on the prior information, the performance of the new MVDR RAB technique based on (23)–(25) is expected to be superior to that of the other RAB techniques.

IV. STEERING VECTOR ESTIMATION VIA SDP RELAXATION

The first step is to make sure that the problem (23)–(25) is feasible. Fortunately, it can be easily verified that (23)–(25) is feasible if and only if \( \Delta_0/M \) is greater than or equal to the smallest eigenvalue of the matrix \( \mathbf{C} \). Indeed, if the smallest eigenvalue of \( \mathbf{C} \) is larger than \( \Delta_0/M \), then the constraint (25) can not be satisfied for any estimate \( \hat{\mathbf{a}} \). The selection of \( \Delta_0 \) according to (16) satisfies the feasibility condition that guarantees the feasibility of (23)–(25).

A. SDP Relaxation

If the problem (23)–(25) is feasible, the equalities \( \hat{\mathbf{a}}^H \mathbf{R}^{-1} \hat{\mathbf{a}} - \mathbf{Tr}(\mathbf{R}^{-1} \hat{\mathbf{a}} \hat{\mathbf{a}}^H) \) and \( \hat{\mathbf{a}}^H \hat{\mathbf{C}} \hat{\mathbf{a}} - \mathbf{Tr}(\hat{\mathbf{C}} \hat{\mathbf{a}} \hat{\mathbf{a}}^H) \) can be used to rewrite it as

\[
\begin{align*}
\min_{\hat{\mathbf{a}}} & \quad \mathbf{Tr}(\mathbf{R}^{-1} \hat{\mathbf{a}} \hat{\mathbf{a}}^H) \\
\text{subject to} & \quad \mathbf{Tr}(\hat{\mathbf{a}} \hat{\mathbf{a}}^H) = M \\
& \quad \mathbf{Tr}(\hat{\mathbf{C}} \hat{\mathbf{a}} \hat{\mathbf{a}}^H) \leq \Delta_0.
\end{align*}
\]

Introducing the following positive semi-definite matrix variable \( \mathbf{A} \triangleq \hat{\mathbf{a}} \hat{\mathbf{a}}^H \), \( \mathbf{A} \succeq 0 \), the problem (30)–(32) can be recast as

\[
\begin{align*}
\min_{\mathbf{A}} & \quad \mathbf{Tr}(\mathbf{R}^{-1} \mathbf{A}) \\
\text{subject to} & \quad \mathbf{Tr}(\mathbf{A}) = M \\
& \quad \mathbf{Tr}(\hat{\mathbf{C}} \mathbf{A}) \leq \Delta_0 \\
& \quad \text{rank}(\mathbf{A}) = 1.
\end{align*}
\]

where \( \text{rank}(\cdot) \) stands for the rank of a matrix. The only non-convex constraint in (33)–(36) is the rank-one constraint (36) while all other constraints and the objective are linear in \( \mathbf{A} \). Using the SDP relaxation technique [24], the relaxed problem can be obtained by dropping the non-convex rank-one constraint (36) and requiring that \( \mathbf{A} \succeq 0 \). Thus, the problem (33)–(36) is replaced by the following relaxed convex problem:

\[
\begin{align*}
\min_{\mathbf{A}} & \quad \mathbf{Tr}(\mathbf{R}^{-1} \mathbf{A}) \\
\text{subject to} & \quad \mathbf{Tr}(\mathbf{A}) = M \\
& \quad \mathbf{Tr}(\hat{\mathbf{C}} \mathbf{A}) \leq \Delta_0 \\
& \quad \mathbf{A} \succeq 0.
\end{align*}
\]

B. Rank of the Optimal Solution

If \( \Delta_0 \) is selected differently from (16), the original problem may be infeasible, while the relaxed one is feasible. The following lemma, which is a by-product result for this paper, establishes an exact equivalence between feasibilities of the original and relaxed problems considered.

Lemma 1: The problem (37)–(40) is feasible if and only if the problem (23)–(25) is feasible.

Proof: See the Appendix.

It is worth noting that if the relaxed problem (37)–(40) has an optimal solution, Lemma 1 follows straightforwardly from the results of [17] and [27]. However, if the relaxed problem does not have an optimal solution, the connection between the feasibility of the relaxed problem (37)–(40) and that of the original problem (23)–(25) is more subtle. In this respect, the novelty of Lemma 1 is in its generality.

If the optimal solution \( \mathbf{A} \) of the relaxed problem (37)–(40) is a rank-one matrix, then the principal eigenvector of \( \mathbf{A} \) scaled by the square root of the largest eigenvalue is the exact solution of the problem (23)–(25). However, even if \( \mathbf{A} \) is not rank-one, it has been shown in [17] and [27] that the rank-one solution for the problems of type (23)–(25) can be found using the duality theory based on the fact that the strong duality holds for such problems. Moreover, the rank-one solution can be found using the well known rank reduction technique [28]. A new algebraic way of extracting the rank-one optimal solution of the problem (23)–(25) from the non-rank-one optimal solution of the problem (37)–(40) under the condition that (37)–(40) is feasible is summarized by means of the following new constructive theorem.

Theorem 1: Let \( \mathbf{A}^{*} \) be the rank \( r \) optimal minimizer of the relaxed problem (37)–(40), i.e., \( \mathbf{A}^{*} = \mathbf{Y} \mathbf{Y}^H \) where \( \mathbf{Y} \) is an \( M \times r \) full rank matrix. If \( r = 1 \), the optimal solution of the original problem simply equals \( \mathbf{Y} \). Otherwise, it equals \( \mathbf{Y} \mathbf{v} \) where \( \mathbf{v} \) is an \( r \times 1 \) vector such that \( \| \mathbf{Y} \mathbf{v} \| = \sqrt{M} \) and \( \mathbf{v}^H \mathbf{Y}^H \mathbf{C} \mathbf{Y} \mathbf{v} = \mathbf{Tr}(\mathbf{Y}^H \mathbf{C} \mathbf{Y}) \). One possible solution for the vector \( \mathbf{v} \) is proportional to the sum of the eigenvectors of the following \( r \times r \) matrix

\[
\mathbf{D} = \frac{1}{M} \mathbf{Y}^H \mathbf{Y} - \frac{\mathbf{Y}^H \mathbf{C} \mathbf{Y}}{\mathbf{Tr}(\mathbf{Y}^H \mathbf{C} \mathbf{Y})}.
\]

Proof: See the Appendix.

One more important question is under which condition the solution of the relaxed problem (37)–(40) is always rank-one. The importance of this question also follows from the fact that it has been observed that the probability of obtaining a rank-one solution for the class of considered problems is close to 1, while the theoretical upper bound suggests a significantly smaller probability [24]. Our next result precisely explains and approves the correctness of the experimental observation about the high probability of the rank-one solution for the relaxed problem (37)–(40).

It is worth noting that any phase rotation of \( \hat{\mathbf{a}} \) does not change the SINR at the output of the corresponding RAB. Therefore, we say that the optimal solution \( \hat{\mathbf{a}} \) is unique when the value of the output SINR or output power (6) is the same for any \( \hat{\mathbf{a}}' = \hat{\mathbf{a}} e^{j\Phi} \). Then, the following lemma holds.

Lemma 2: Under the condition that the solution of the original problem (23)–(25) is unique in the sense mentioned above, the solution of the relaxed problem (37)–(40) always has rank one.

Proof: See the Appendix.
Under the condition of Lemma 2, the solution of (37)–(40) is rank-one and the solution of (23)–(25) can be found as a scaled version of the dominant eigenvector of the solution of (37)–(40). If the uniqueness condition of Lemma 2 is not satisfied for (23)–(25), we resort to the constructive result of Theorem 1 for finding the rank-one solution of (23)–(25) algebraically. An example of a situation when the condition of Lemma 2 is not satisfied is given next. In general, such situations are rare that, in fact, has been also observed by means of simulations in other works.

**Example 3:** Let us consider a ULA with 10 omni-directional antenna elements. The presumed direction of arrival of the desired user is assumed to be $\theta_p = 3^\circ$ with no interfering sources and the range of the desired signal angular locations is equal to $\Theta = [\theta_p - 12^\circ, \theta_p + 12^\circ]$. The actual steering vector of the desired user is perturbed due to the incoherent local scattering effect and it can be expressed as $\hat{a}(k) = v_0(k)p + v_1(k)b$, where $p = d(3^\circ)$ is the steering vector of the direct path, $b$ is the steering vector of the scattered path, and $v_0(k)$ and $v_1(k)$ are i.i.d. zero mean complex Gaussian random variables with unit variance which change from snapshot to snapshot. If $b$ is orthogonal to $p$, that is the case when $b$ is selected as $d(-8.4916^\circ)$, both $p$ and $b$ are the eigenvectors of the matrix $R^{-1}$ which correspond to the smallest eigenvalue. Since, these vectors satisfy the constraints (24)–(25) and correspond to the minimum eigenvalue, both of them are optimal solutions of (23)–(25). Thus, the solution of (23)–(25) is not unique.

V. SIMULATION RESULTS

Throughout the simulations, a ULA of 10 omni-directional antenna elements with the inter-element spacing of half wavelength is considered unless otherwise is specified. Additive noise in antenna elements is modeled as spatially and temporally independent complex Gaussian noise with zero mean and unit variance. Two interfering sources are assumed to impinge on the antenna array from the directions 30° and 50°, while the presumed direction towards the desired signal is assumed to be $\theta_p = 3^\circ$ unless otherwise is specified. In all simulation examples, the interference-to-noise ratio (INR) equals 30 dB and the desired signal is always present in the training data. For obtaining each point in the curves, 100 independent runs are used.

The proposed beamformer is compared with the following four methods in terms of the output SINR: i) the eigenspace-based beamformer of [9]; ii) the worst-case-based RAB of [11]; iii) the beamformer of [19]; and iv) the diagonally loaded SMI (LSMI) beamformer [6]. Moreover, in the last simulation example, a comparison is made with the multi-rank eigenvalue beamformer of [32]. For the proposed beamformer and the beamformer of [19], the angular sector of interest $\Theta$ is assumed to be $\Theta = [\theta_p - 5^\circ, \theta_p + 5^\circ]$. The CVX Matlab toolbox is used for solving the optimization problem (37)–(40). The value $\delta = 0.1$ and 8 dominant eigenvectors of the matrix $C$ are used in the beamformer of [19] and the value $\varepsilon = 0.3M$ is used for the worst-case-based beamformer as it has been recommended in [11]. The dimension of the signal-plus-interference subspace is assumed to be always estimated correctly for the eigenspace-based beamformer of [9]. Diagonal loading factor of the SMI beamformer is selected as twice the noise power as recommended by Cox et al. in [6].

![Fig. 4. Simulation Example 1: Output SINR versus training sample size $K$ for fixed SNR = 20 dB and INR = 30 dB.](image)

![Fig. 5. Simulation Example 1: Output SINR versus SNR for training data size of $K = 30$ and INR = 30 dB.](image)

**Simulation Example 1:** *Exactly Known Signal Steering Vector:* In this example, we consider the case when the actual steering vector is known exactly. Even in this case, the presence of the desired signal in the training data can substantially reduce the convergence rates of adaptive beamforming algorithms as compared to the signal-free training data case [8]. In Fig. 4, the mean output SINRs for the four methods tested are illustrated versus the number of training snapshots for the fixed single-sensor SNR = 20 dB. Fig. 5 displays the mean output SINR of the same methods versus the SNR for fixed training data size of $K = 30$. It can be seen from these figures that the proposed beamforming technique outperforms the other techniques. Only in the situation when the number of snapshots is larger than 70, the technique of [19] results in slightly better performance. It is because if the desired signal steering vector is known precisely, the only source of error is the finite sample size used, but the difference between the sample data covariance matrix and the theoretical data covariance matrix due to finite sample size can be equivalently transferred to the error in the steering vector [8]. Therefore, as the number of samples increases and the variance of the steering vector mismatch decreases, the estimator with more restrictive constraints, that is the one of [19], may indeed result in a better performance.
Simulation Example 2: Desired Signal Steering Vector Mismatch due to Wavefront Distortion: In the second example, we consider the situation when the signal steering vector is distorted by wave propagation effects in an inhomogeneous medium. Specifically, independent-increment phase distortions are accumulated by the components of the presumed steering vector. It is assumed that the phase increments remain fixed in each simulation run and are independently chosen from a Gaussian random generator with zero mean and standard deviation 0.04.

The output SINR curves for the proposed, SMI, and worst-case-based methods are shown versus the SNR for fixed training data size $K = 30$ in Fig. 6. It can be seen that the proposed beamforming technique outperforms the worst-case-based one at low and moderate SNRs. However, when SNR is much larger than INR so that signal-to-interference ratio goes to infinity, the proposed and the worst-case-based MVDR RAB techniques perform almost equivalently. Indeed, in the latter case, it is guaranteed that the estimate of the desired signal steering vector does not converge to an interference steering vector and, thus, the solution of the problem coincides with the problem obtained after dropping the constraint.

Simulation Example 3: Effect of the Error in the Knowledge of the Antenna Array Geometry: In this example, we first aim at checking how the presence of antenna array perturbations affects the sector. Specifically, we want to characterize quantitatively the dependence of the sector on the level of antenna array perturbations, which grows from zero to its maximum value, for a given $\Delta_\theta$. The presumed angular sector is assumed to be $[0^\circ, 10^\circ]$ for $\theta_a = 5^\circ$. Let the antenna array perturbations be caused by errors in the antenna element positions which are drawn uniformly from $[-\alpha, \alpha]$, where $\alpha$ is the level of perturbations measured in wavelength. Table II illustrates the average width of $\Theta_a$ for different values of $\alpha$. It can be seen from this table that the deviation of $\Theta_a$ due to array perturbations compared to the case of no perturbations is 0% for small perturbations and near 0% even for the highest levels of perturbations. Here the term ‘deviation’ stands for the percent of non-overlap between angular sectors corresponding to the cases of no perturbations and near 0% even for the highest levels of perturbations. Based on this observation, one can conclude that the matrix $C$ required for implementing the constraint (15) can be computed using the presumed array geometry which is not required to be precise and can be, in fact, very approximate. It is worth noting that we also observed throughout extensive simulations that if the sector $\Theta$ gets away from the broadside (it is near $-90^\circ$ or $90^\circ$), then the length of the associated sector $\Theta_a$ increases. Such increase can be noticeable especially when the number of antenna elements in the antenna array is small. To avoid the situation when $\Theta_a$ may contain interference sources because of its bigger size, the presteering filter-type technique [34] can be used, for example, to ensure that the center of the sector $\Theta$ for the signal at the output of the presteering filter is around the broadside.

We also study the effect of the error in the knowledge of antenna array geometry used for the computation of the matrix $C$ on the performance of the proposed MVDR RAB technique. The difference between the presumed and actual positions of each antenna element is modeled as a uniform random variable distributed in the interval $[-0.05, 0.05]$ measured in wavelength. In addition to the antenna element displacements, the signal steering vector is distorted as in our simulation Example 2. Figs. 7 and 8 depict the output SINR performance of the RAB techniques tested versus the number of training snapshots for fixed single-sensor SNR = 20 dB and versus the SNR for fixed training data size $K = 30$, respectively. As it can be observed from the figures, the proposed method has a

![Fig. 6. Simulation Example 2: Output SINR versus SNR for training data size of $K = 30$ and INR = 30 dB.](image)

![Fig. 7. Simulation Example 3: Output SINR versus training sample size $K$ for fixed SNR = 20 dB and INR = 30 dB for the case of perturbations in antenna array geometry.](image)

<table>
<thead>
<tr>
<th>Level of perturbations $\alpha$</th>
<th>0</th>
<th>0.05</th>
<th>0.1</th>
<th>0.15</th>
<th>0.2</th>
<th>0.25</th>
</tr>
</thead>
<tbody>
<tr>
<td>Width of $\Theta_a$</td>
<td>10.80°</td>
<td>10.80°</td>
<td>10.72°</td>
<td>10.57°</td>
<td>10.53°</td>
<td>10.42°</td>
</tr>
<tr>
<td>Deviation(%)</td>
<td>0%</td>
<td>0%</td>
<td>0.7%</td>
<td>2.06%</td>
<td>2.48%</td>
<td>3.48%</td>
</tr>
</tbody>
</table>

TABLE II

WIDTH OF THE FEASIBLE SET $\Theta_a$ VERSUS THE LEVEL OF PERTURBATIONS $\alpha$
better performance even if there is an error in the knowledge of the antenna array geometry.

Simulation Example 4: Desired Signal Steering Vector Mismatch due to Coherent Local Scattering [33]: In this example, the desired signal steering vector is distorted by local scattering effects so that the actual steering vector is formed by five signal paths as $\mathbf{a} = \mathbf{p} + \sum_{i=1}^{4} e^{j\psi_i} \mathbf{b} (\theta_i)$, where $\mathbf{p}$ corresponds to the direct path and $\mathbf{b} (\theta_i)$, $i = 1, 2, 3, 4$ correspond to the coherently scattered paths. The $i$th path $\mathbf{b} (\theta_i)$ is modeled as a plane wave impinging on the array from the direction $\theta_i$. The angles $\theta_i$, $i = 1, 2, 3, 4$ are independently drawn in each simulation run from a uniform random generator with mean $3^\circ$ and standard deviation $1^\circ$. The parameters $\psi_i$, $i = 1, 2, 3, 4$ represent path phases that are independently and uniformly drawn from the interval $[0, 2\pi]$ in each simulation run. Note that $\theta_i$ and $\psi_i$, $i = 1, 2, 3, 4$ change from run to run but do not change from snapshot to snapshot.

Fig. 9 displays the output SINR performance of all four methods tested versus the number of training snapshots $K$ for fixed single-sensor SNR $= -20$ dB. Note that the SNR in this example is defined by taking into account all signal paths. The output SINR performance of the same methods versus SNR for the fixed training data size $K = 301$ is displayed in Fig. 10.

Similar to the previous example, the proposed beamformer significantly outperforms other beamformers due to its ability to estimate the desired signal steering vector with higher accuracy than other methods. As compared to the eigenspace-based method, the proposed technique does not suffer from the subspace swap phenomenon at low SNRs since it does not use eigenvalue decomposition of the sample covariance matrix.

Simulation Example 5: Comparison With Eigenvalue Beamforming-Based Methods of [32]: In this example, we consider the case where the desired and interference signals have the same structure and are modeled as signals with a rank-one covariance matrix from a $p$-dimensional subspace. Specifically, the model introduced in [32] for the desired and interference signals is adopted. The corresponding steering vector of the desired and interference signals are all modeled as $\mathbf{s} = \Psi \mathbf{b}_0 \delta$, where $\Psi$ is an $M \times p$ ($p < M$) matrix whose columns are orthogonal ($\Psi^H \Psi = \mathbf{I}_{p \times p}$) and $\mathbf{b}_0$ is an unknown but fixed vector from one snapshot to another. The matrix $\Psi$ (different for each signal) is obtained by choosing $p - 3$ dominant eigenvectors of the matrix $\int_{\phi = \Delta \phi} d(\theta) \mathbf{d}^H(\theta) d\theta$ as the columns of $\Psi$, where $\phi$ denotes the presumed location of the source and $\Delta \phi$ equals to $5^\circ$ for all the signals. In order to satisfy the assumption of the proposed RAB which requires the desired user to lie inside the desired sector, the number of the antenna elements is taken to be equal to $M = 30$. Note that if $M = 10$ the matrix $\int_{\phi = \Delta \phi} d(\theta) \mathbf{d}^H(\theta) d\theta$ does not have three dominant eigenvectors.

To obtain the signal subspace which corresponds to the desired signal, we find the maximum of the bearing pattern response defined as $P_\theta (\delta) = \text{Tr}(\mathbf{Q}^H (\Psi^H (\delta) \mathbf{R}^{-1} \Psi(\theta))^{-1} \mathbf{Q})$ in the desired sector where $\Psi(\theta)$ is the orthogonal matrix whose columns are equal to the three dominant eigenvectors of the $M \times p$ matrix $\int_{\phi = \Delta \phi} d(\theta) \mathbf{d}^H(\theta) d\theta$ [32]. Fig. 11 shows the sample bearing pattern response. As it is expected, the maximum occurs exactly around $3^\circ$.

It is noteworthy to mention that the proposed RAB, the RAB of [19], and the eigenvalue beamformer of [32] all use the knowledge of approximate antenna array geometry. In order to evaluate how the error in the knowledge of antenna array geometry affects these methods, two different cases are considered. In the first case, knowledge of antenna array geometry is
Fig. 11. Simulation Example 5: Bearing beampattern corresponding to the eigenvalue beamforming method of [32].

Fig. 12. Simulation Example 5: Output SINR versus SNR for training data size of $K = 30$ and INR = 30 dB.

Fig. 13. Simulation Example 5: Output SINR versus SNR for training data size of $K = 30$ and INR = 30 dB.

accurate while in the second case, antenna array perturbations are considered. Similar to our simulation Example 3, antenna array perturbations are modeled as errors in the antenna element positions which are drawn uniformly from the interval $[-0.05, 0.05]$ measured in wavelength. For the multi-rank beamformer of [32], we use $\text{Tr}(W^H R_s W)/\text{Tr}(W^H R_{i+n} W)$ as the output SINR, where $R_s$ denotes the correlation matrix of the desired signal. Figs. 12 and 13 illustrate the performance of the aforementioned methods for the cases when the knowledge of antenna array geometry is accurate and approximate, respectively. The best performance by the eigenvalue beamformer of [32] is obtained when $Q$ only contains the most dominant eigenvector of the error covariance matrix. As it can be seen from the figures, the proposed RAB method outperforms all other RAB methods.

VI. CONCLUSION

The MVDR RAB techniques have been considered from the viewpoint of a single unified principle, that is, to use standard MVDR beamforming in tandem with an estimate of the signal steering vector found based on some prior information. It has been demonstrated that differences between various MVDR RAB techniques occur only because of the differences in the assumed prior information and the corresponding signal steering vector estimation techniques. The latter fact has motivated us to develop a new MVDR RAB technique that uses as little as possible, imprecise, and easy to obtain prior information. The new MVDR RAB technique, which assumes only an imprecise knowledge of antenna array geometry and angular sector in which the actual steering vector lies, has been developed. It is mathematically expressed as the well known non-convex QCQPI problem with one convex quadratic inequality constraint and one non-convex quadratic equality constraint. A number of methods for finding efficiently the exact optimal solution for such problem is known. In addition to the existing methods, we have developed a new algebraic method of finding the rank-one solution from the general-rank solution of the relaxed problem. The condition under which the solution of the relaxed problem is guaranteed to be rank-one has also been derived. Our simulation results demonstrate the superior performance for the proposed MVDR RAB technique over the existing state of the art RAB methods.

APPENDIX

Proof of Lemma 1

Let $\tilde{a}$ be a feasible point for the original problem (23)–(25). It is straightforward to see that $A \triangleq \tilde{a}\tilde{a}^H$ is also a feasible point for the relaxed problem (37)–(40). Thus, the necessity statement of the lemma follows trivially.

In order to prove sufficiency, let $A = \sum_{i=1}^M \lambda_i b_i b_i^H$ be a feasible point for the relaxed problem (37)–(40), where $\lambda_i$ and $b_i$, $i = 1, \ldots, M$ are, respectively, the eigenvalues and eigenvectors of $A$. Also let $l$ denote the index of $b_l$, $i = 1, \ldots, M$ for which the quadratic form $b_l^H \tilde{c} b_l$ has the smallest value and $\tilde{a} = \sqrt{M} b_l$. Then, the following holds $\tilde{a}^H \tilde{a} = M b_l^H \tilde{c} b_l = M$ and the constraint (24) is satisfied. Moreover, we can write that

$$\tilde{a}^H \tilde{c} b_l = M b_l^H \tilde{c} b_l = \left( \sum_{i=1}^M \lambda_i \right) b_l^H \tilde{c} b_l = (42)$$
where \( \sum_{i=1}^{M} \lambda_i = \text{Tr}(A) = M \). Using the following inequality:

\[
\left( \sum_{i=1}^{M} \lambda_i \right) b_i^H \tilde{C} b_i \leq \sum_{i=1}^{M} \lambda_i b_i^H \tilde{C} b_i \tag{43}
\]

and (42), we obtain that

\[
\hat{a}^H \tilde{C} \hat{a} \leq \sum_{i=1}^{M} \lambda_i b_i^H \tilde{C} b_i \tag{44}
\]

The right hand side of (44) can be further rewritten as

\[
\sum_{i=1}^{M} \lambda_i b_i^H \tilde{C} b_i = \sum_{i=1}^{M} \lambda_i \text{Tr} \left( \tilde{C} b_i b_i^H \right) \tag{45}
\]

Using the property of the trace that a sum of traces is equal to the trace of a sum, we obtain that

\[
\sum_{i=1}^{M} \lambda_i \text{Tr} \left( \tilde{C} b_i b_i^H \right) = \text{Tr} \left( \tilde{C} \sum_{i=1}^{M} \lambda_i b_i b_i^H \right) \tag{46}
\]

Since \( \sum_{i=1}^{M} \lambda_i b_i b_i^H = A \), we have

\[
\sum_{i=1}^{M} \lambda_i b_i b_i^H \tilde{C} b_i = \text{Tr} \left( \tilde{C} \sum_{i=1}^{M} \lambda_i b_i b_i^H \right) - \text{Tr}(\tilde{C} A) \tag{47}
\]

Substituting (47) in the right hand side of (44), we finally obtain that

\[
\hat{a}^H \tilde{C} \hat{a} \leq \text{Tr}(\tilde{C} A) \leq \Delta_0 \tag{48}
\]

Therefore, the constraint (25) is also satisfied and, thus, \( \hat{a} = \sqrt{M} b_1 \) is a feasible point for (23)–(25) that completes the proof. \( \square \)

Proof of Theorem 1

Let \( \hat{A}^* \) be the optimal minimizer of the relaxed problem (37)–(40), and its rank be \( r \). Consider the following decomposition of \( \hat{A}^* \):

\[
\hat{A}^* = Y Y^H \tag{49}
\]

where \( Y \) is an \( M \times r \) complex valued full rank matrix. It is trivial to see that if the rank of the optimal minimizer \( \hat{A}^* \) of the relaxed problem (37)–(40) is one, then \( Y \) is also the optimal minimizer of the original problem (23)–(25). Thus, it is assumed in the following that \( r > 1 \).

We start by considering the following auxiliary optimization problem

\[
\min_{G} \text{Tr}(\tilde{R}^{-1} Y G Y^H) \tag{50}
\]

subject to

\[
\text{Tr}(Y G Y^H) = M \tag{51}
\]

\[
\text{Tr}(C Y G Y^H) = \text{Tr}(\tilde{C} A^*) \tag{52}
\]

\[
G \succeq 0 \tag{53}
\]

where \( G \) is an \( r \times r \) Hermitian matrix. The matrix \( A \) in (37)–(40) can be expressed as a function of the matrix \( G \) in (50)–(53) as \( A(G) = Y Y^H \). Moreover, it can be easily shown that if \( G \) is a positive semi-definite matrix, then \( A(G) \) is also a positive semi-definite matrix. In addition, if \( G \) is a feasible solution of (50)–(53), \( A(G) \) is also a feasible solution of (37)–(40). The latter is true because \( A(G) \) is a positive semi-definite matrix and it satisfies the constraints \( \text{Tr}(A(G)) = M \) and \( \text{Tr}(\tilde{C} A(G)) = \text{Tr}(\tilde{C} A^*) \leq \Delta_c \). This implies that the minimum value of the problem (50)–(53) is greater than or equal to the minimum value of the problem (37)–(40).

It is then easy to verify that \( G^* = I_{r \times r} \) is a feasible point of the auxiliary optimization problem (50)–(53). Moreover, the fact that \( \text{Tr}(\tilde{R}^{-1} Y G^* Y^H) = \text{Tr}(\tilde{R}^{-1} A^*) \) denotes the minimum value of the relaxed problem (37)–(40) together with the fact that the minimum value of the auxiliary problem (50)–(53) is greater than or equal to \( \beta \), implies that \( G^* = I_{r \times r} \) is the optimal solution of the auxiliary problem (50)–(53).

In what follows, we show that every feasible solution of the problem (50)–(53), denoted as \( G' \), is an optimal minimizer of the same problem and it is also an optimal minimizer of (50)–(53). Thus, \( A(G') = Y G' Y^H \) is an optimal minimizer of (37)–(40).

Let us start by considering the following problem dual to (50)–(53)

\[
\max_{\nu_1, \nu_2, Z} \nu_1 M + \nu_2 \text{Tr}(\tilde{C} A^*) \tag{54}
\]

subject to

\[
Y^H \tilde{R}^{-1} Y - \nu_1 Y^H Y - \nu_2 Y^H \tilde{C} Y \succeq Z \tag{55}
\]

\[
Z \succeq 0 \tag{56}
\]

where \( \nu_1 \) and \( \nu_2 \) are the Lagrange multipliers associated with the constraints (51) and (52), respectively, and \( Z \) is an \( r \times r \) Hermitian matrix of Lagrange multipliers associated with the constraint (53). The problem (50)–(53) is convex, and it satisfies the Slater’s condition because, as it was mentioned, the positive definite matrix \( G - I_{r \times r} \) is a strictly feasible point for (50)–(53). Thus, the strong duality between (50)–(53) and (54)–(56) holds.

Let \( \nu_1^*, \nu_2^* \) and \( Z^* \) be one possible optimal solution of the dual problem (54)–(56). Since strong duality holds, we can state that \( \nu_1^* M + \nu_2^* \text{Tr}(\tilde{C} A^*) = \beta \) and \( I_{r \times r} \) is a possible optimal solution of the primal problem (50)–(53). Moreover, the complementary slackness condition implies that

\[
\text{Tr}(G^* Z^*) = \text{Tr}(Z^*) = 0. \tag{57}
\]

Since \( Z^* \) has to be a positive semi-definite matrix, the condition (57) implies that \( Z^* = 0 \). Then it follows from (55) that

\[
Y^H \tilde{R}^{-1} Y - \nu_1^* Y^H Y - \nu_2^* Y^H \tilde{C} Y \succeq 0. \tag{58}
\]

The fact that \( \beta - \nu_1^* M - \nu_2^* \text{Tr}(\tilde{C} A^*) = 0 \) implies that \( \text{Tr}(Y^H \tilde{R}^{-1} Y - \nu_1^* Y^H Y - \nu_2^* Y^H \tilde{C} Y) = 0 \). As a result, it can be easily verified that the constraint (58) is active, i.e., it is satisfied as equality for optimal \( \nu_1^* \) and \( \nu_2^* \). Therefore, we can write that

\[
Y^H \tilde{R}^{-1} Y = \nu_1^* Y^H Y + \nu_2^* Y^H \tilde{C} Y. \tag{59}
\]

Let \( G' \) be another feasible solution of (50)–(53) different from \( I_{r \times r} \). Then the following conditions must hold

\[
\text{Tr}(Y^H Y G') = M \tag{60}
\]

\[
\text{Tr}(Y^H C Y G') = \text{Tr}(\tilde{C} A^*) \tag{61}
\]

\[
G' \succeq 0. \tag{62}
\]
Multiplying both sides of the (59) by \( G' \), we obtain
\[
Y^H \tilde{R}^{-1} YG' = v_1^* Y^H YG' + v_2^* Y^H \tilde{C} YG'.
\] (63)

Moreover, taking the trace of the right hand and left hand sides of (63), we have
\[
\text{Tr}(Y^H \tilde{R}^{-1} YG') = v_1^* \text{Tr}(Y^H YG') + v_2^* \text{Tr}(Y^H \tilde{C} YG') - v_1^* M + v_2^* \text{Tr}(\tilde{C} A^*) - \beta.
\] (64)

This implies that \( G' \) is also a possible optimal solution of (50)–(53). Therefore, every feasible solution of (50)–(53) is also a possible optimal solution.

Finally, we show that there exists a feasible solution of (50)–(53) whose rank is one. As it has been proved above, such feasible solution is also a possible optimal solution. Let \( H \triangleq vv^H \), and we are interested in finding such \( v \) that
\[
\text{Tr}(Y^H YH) = v^H Y^H Y v = M \tag{65}
\]
\[
\text{Tr}(Y^H \tilde{C} YH) = v^H Y^H \tilde{C} Y v = \text{Tr}(\tilde{C} A^*). \tag{66}
\]

Equivalently, the conditions (65) and (66) can be rewritten as
\[
\frac{1}{M} v^H Y^H Y v = 1, \tag{67}
\]
\[
v^H \frac{Y^H \tilde{C} Y}{\text{Tr}(\tilde{C} A^*)} v = 1. \tag{68}
\]

Moreover, equating the left hand side of (67) to the left hand side of (68), we obtain that
\[
\frac{1}{M} v^H Y^H Y v = v^H \frac{Y^H \tilde{C} Y}{\text{Tr}(\tilde{C} A^*)} v. \tag{69}
\]

Finding the difference between the left- and right-hand sides of (69), we also obtain that
\[
v^H \left( \frac{1}{M} Y^H Y - \frac{Y^H \tilde{C} Y}{\text{Tr}(\tilde{C} A^*)} \right) v = v^H D v = 0. \tag{70}
\]

Considering the fact that \( \text{Tr}(Y^H Y) = M \) and \( \text{Tr}(Y^H \tilde{C} Y) = \text{Tr}(\tilde{C} A^*) \), we find that \( \text{Tr}(Y^H Y/M - Y^H \tilde{C} Y/\text{Tr}(\tilde{C} A^*)) = \text{Tr}(D) = 0 \). Therefore, the vector \( v \) can be chosen proportional to the sum of the eigenvectors of the matrix \( D \) and it can be scaled so that
\[
v^H Y^H Y v = \text{Tr}(\tilde{C} A^*) \text{ and, thus, (65) and (66) are satisfied.}
\]

So far we have found a rank one solution for the auxiliary optimization problem (51)–(53), that is, \( G = vv^H \). Since \( G = vv^H \) is a possible optimal solution of the auxiliary problem (51)–(53), then \( YGY^H = (Yv)(Yv)^H \) is a possible optimal solution of the relaxed problem (37)–(40). Moreover, since the solution \( (Yv)(Yv)^H \) is rank-one, \( Yv \) is a possible optimal solution of the original optimization problem (23)–(25). This completes the proof.

**Proof of Lemma 2**

Let \( A^* \) be one possible optimal solution of the problem (37)–(40) whose rank \( r \) is greater than one. Using the rank-one decomposition of Hermitian matrices [35], the matrix \( A^* \) can be written as
\[
A^* = \sum_{j=1}^{r} z_j z_j^H \tag{71}
\]

where
\[
z_j^H z_j = \frac{1}{r} \text{Tr}(A^*) = \frac{M}{r}, \quad j = 1, \ldots, r \tag{72}
\]
\[
z_j^H \tilde{C} z_j = \frac{1}{r} \text{Tr}(\tilde{C} A^*), \quad j = 1, \ldots, r. \tag{73}
\]

Let us show that the terms \( z_j^H \tilde{R}^{-1} z_j, \ j = 1, \ldots, r \) are equal to each other for all \( j = 1, \ldots, r \). We prove it by contradiction assuming first that there exist such \( z_m \) and \( z_n \), \( m \neq n \) that
\[
z_m^H \tilde{R}^{-1} z_m < z_n^H \tilde{R}^{-1} z_n. \]

Let the matrix \( A_0^* \) be constructed as
\[
A_0^* = A^* - z_m z_m^H + z_n z_n^H. \]

It is easy to see that \( \text{Tr}(A^*) = \text{Tr}(A_0^*) \) and \( \text{Tr}(\tilde{C} A^*) = \text{Tr}(\tilde{C} A_0^*) \), which means that \( A_0^* \) is also a feasible solution of the problem (37)–(40). However, based on our assumption that \( z_m^H \tilde{R}^{-1} z_m < z_n^H \tilde{R}^{-1} z_n \), it can be concluded that \( \text{Tr}(R^{-1} A_0^*) < \text{Tr}(R^{-1} A^*) \), which is obviously a contradiction. Thus, all terms \( z_j^H \tilde{R}^{-1} z_j, \ j = 1, \ldots, r \) must take the same value. Using this fact together with the (72) and (73), we can conclude that \( rz_j^H z_j \) for any \( j = 1, \ldots, r \) is a possible optimal solution of the relaxed problem (37)–(40) which has rank one. Thus, the optimal solution of the original problem (23)–(25) is \( \sqrt{r} z_j^* \) for any \( j = 1, \ldots, r \). Since the vectors \( z_j^*, \ j = 1, \ldots, r \) in (71) are linearly independent and each of them represents a possible optimal solution of (23)–(25), we conclude that there are many possible optimal solution of (23)–(25). However, it contradicts the assumption that the optimal solution of (23)–(25) is unique (regardless possible phase rotation as explained earlier in the paper). Thus, the only optimal solution \( A^* \) of the relaxed problem (37)–(40) has to be rank-one. This completes the proof. □

**References**


Arash Khabbazibasmenj (S’08) received the B.Sc. and M.Sc. degrees in electrical engineering from Amirkabir University of Technology, Tehran, Iran, and the University of Tehran, Tehran, Iran, in 2006 and 2009, respectively. He is currently working toward the Ph.D. degree in electrical engineering at the University of Alberta, Edmonton, AB, Canada.

His research interests include signal processing and optimization methods in radar, communications and related fields.

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. Since 2006, he has been with the Department of Electrical and Computer Engineering, University of Alberta, Edmonton, AB, Canada where he became an Associate Professor in 2010 and Full Professor in 2012. Since his graduation, he also occupied various research and faculty positions at the Kharkiv National University of Radio Electronics, Ukraine; the Institute of Physical and Chemical Research (RIKEN), Japan; McMaster University, Hamilton, ON, Canada; Duisburg-Essen University and the Darmstadt University of Technology, Germany; and the Joint Research Institute between the Heriot-Watt and Edinburgh Universities, U.K. He also held visiting positions at the Technion, Haifa, Israel, in 2005 and the Ilmenau University of Technology, Ilmenau, Germany, in 2011. 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), and other awards. He was 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 is a member of the Array and Multi-Channel Signal Processing and Signal Processing for Communications and Networking Technical Committees of IEEE Signal Processing Society. He has served as a Track Chair in Asilomar 2011 and as a Technical Co-Chair in the IEEE CAMSAP 2011.