Performance Tradeoffs Among Adaptive Beamforming Criteria

Yue Rong, *Member, IEEE*, Yonina C. Eldar, *Senior Member, IEEE*, and Alex B. Gershman, *Fellow, IEEE*

***Abstract*—In this paper, we study the tradeoffs among three main criteria for adaptive beamformer design: maximal signal-to-interference-plus-noise ratio (MSINR), minimal mean-squared error (MMSE), and minimal least-squares error (MLSE). When the power and steering vector of the signal-of-interest (SOI) are exactly known, there are beamformers that can simultaneously meet the MMSE and MSINR criteria. However, this is no longer true when the exact knowledge of the steering vector is unavailable. To account for steering vector errors, a meaningful approach, which we adopt in this paper, is to model the actual steering vector as random. In this setting, we show that the MMSE and MSINR criteria cannot be simultaneously attained. Therefore, using convex analysis tools, we study the achievable region in the MSE-SINR plane and propose an adaptive beamformer that can attain the frontier of operating points on the boundary of this region, providing an optimal performance tradeoff between SINR and MSE. In contrast, we show that even in the presence of steering-vector uncertainties, the MLSE and MSINR criteria are simultaneously achievable, and develop an adaptive beamformer which is optimal under both these criteria.**

***Index Terms*—Adaptive beamforming, least-squares beamformer, MSE-SINR tradeoff, random steering vector errors.**

I. Introduction

ADAPTIVE beamforming has found numerous applications in radar, sonar, wireless communications, biomedical engineering, and other fields [1]–[4]. In many scenarios such as in radar and sonar, the main criterion for beamformer design is to maximize the output signal-to-interference-plus-noise ratio (SINR). The resulting beamformer is referred to as the maximal SINR (MSINR) method [1]. However, maximizing the SINR does not necessarily guarantee an acceptably good estimate of the signal [5]. For example, in an estimation context, where our goal is to design a beamformer to estimate the signal-of-interest (SOI) waveform, it is more important to minimize the signal estimation error rather than maximize the SINR. In such cases, minimal mean-squared error (MMSE) beamforming is of great interest [4], [5]. In contrast to the MSINR approach, the computation of the MMSE beamformer requires exact knowledge of the signal power. If the SOI steering vector is exactly known, then the MMSE beamformer also maximizes the output SINR [5].

In practical scenarios, due to different imperfections such as signal pointing errors, array calibration errors, and environmental nonstationarities, there may be a mismatch between the actual and presumed steering vectors [7]–[13]. Therefore, many authors have addressed the problem of robust beamformer design based on different steering vector mismatch models [8]–[13]. For example, in [8]–[11], worst-case robust beamformer designs have been proposed where the mismatch effect is taken into account by treating the actual SOI steering vector as a deterministic vector located inside a multidimensional sphere. An alternative approach, which we adopt in this paper, is to treat the SOI steering vector as random with some known distribution [12]–[14]. In what follows, we assume that the actual steering vector is Gaussian with known mean and covariance matrix. This provides a general framework that can accommodate many practical cases of interest [13]. As we will show below, in this case the MMSE beamformer does not maximize the output SINR, so that the MMSE and MSINR criteria cannot be simultaneously achieved.

To characterize the tradeoffs between the MSE and SINR criteria, we consider the MSE-SINR plane, and use the fact that any beamformer is associated with a certain point in this plane. Clearly, good beamformers should yield small MSE and large SINR achieving a proper MSE-SINR tradeoff. Here, we propose a new class of beamformers that provide optimal MSE-SINR tradeoff by minimizing the MSE subject to a certain given SINR constraint. We refer to the resulting methods as OPTO (OPtimal TradeOff) beamformers. Each of these beamformers is Pareto optimal [15], [16] in the sense that there is no other approach which yields both smaller MSE and larger SINR. In terms of the MSE-SINR plane, the points corresponding to the OPTO class of methods form an optimal tradeoff curve which is the frontier that divides the MSE-SINR plane into two parts that contain achievable and non-achievable points, respectively.

The knowledge about the optimal tradeoff between the SINR and MSE can be important in practical applications where the objective is to optimize a combination of these two criteria. Such a beamformer can be easily obtained using the optimal tradeoff curve.

The design of OPTO beamformers requires the SOI power to be known. However, in practical scenarios the power may not be known precisely. One approach in this case can be to use some estimate of the signal power instead of its true value. The per-

Manuscript received February 14, 2007; revised September 5, 2007; accepted September 5, 2007. The work of Y. C. Eldar was supported by the Israel Science Foundation under Grant 536-04. The work of A. B. Gershman was supported by the German Research Foundation (DFG) under Grant GE 1881/1-1. This paper was presented in part at the 2006 IEEE Sensor Array and Multichannel Signal Processing Workshop, Waltham, MA, June 2006. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Timothy N. Davidson.

Y. Rong is with the Department of Electrical Engineering, University of California, Riverside, CA 92521 USA (e-mail: yue.rong@ieee.org).

Y. C. Eldar is with the Department of Electrical Engineering, Technion—Israel Institute of Technology, Haifa 32000, Israel (e-mail: yonina@ee.technion.ac.il).

A. B. Gershman is with the Communication Systems Group, Institute of Telecommunications, Darmstadt University of Technology, D-64283 Darmstadt, Germany (e-mail: gershman@ieee.org).

Digital Object Identifier 10.1109/JSTSP.2007.909379

1932-4553/$25.00 © 2007 IEEE

formance of the resulting MMSE beamformer will then depend on the accuracy of the power estimate. Alternatively, to circumvent the need for power estimation, we may use the minimal least-squares error (MLSE) beamformer [5], [6]. The MLSE strategy is based on minimizing the least-squares error (LSE) of the array observations, and does not require any knowledge of the signal power. Interestingly, we show that even when the steering vector is random, the MLSE and MSINR criteria can be simultaneously achieved. We then develop a new beamformer that is optimal under both of these criteria.

The rest of our paper is organized as follows. The system model and some necessary background on adaptive beamforming are presented in Section II. Section III studies the performance tradeoffs among different beamforming approaches. Simulations are presented in Section IV and conclusions are drawn in Section V.

II. Background

Consider an M-sensor antenna array whose $M \times 1$ observation vector at time t can be modelled as

$$\mathbf{y}(t) = s(t)\mathbf{a} + \mathbf{i}(t) + \mathbf{v}(t) \tag{1}$$

where $\mathbf{a}$ is the SOI steering vector, $s(t)$ is the SOI waveform, and $\mathbf{i}(t)$ and $\mathbf{v}(t)$ are the interference and noise vectors, respectively. In the sequel, we will omit the dependence of t for notational simplicity.

In many practical applications, it is difficult to obtain information about the actual steering vector, for example because of pointing errors, imperfect array calibration, and environmental nonstationarities. The problem of robust beamformer design when only partial information about the actual steering vector is available has been studied extensively [7]–[13]. In this paper, we consider the scenario where the SOI steering vector is complex Gaussian distributed [12], [13] with mean $\mathbf{m}$ and covariance matrix $\mathbf{C}$, that is, $\mathbf{a} \sim \mathcal{CN}(\mathbf{m}, \mathbf{C})$. The mean $\mathbf{m}$ corresponds to the perturbation-free steering vector, while the covariance matrix $\mathbf{C}$ captures potential uncertainties in this vector. We assume that $\mathbf{m}$ and $\mathbf{C}$ are both known and that $\mathbf{m} \neq 0$.

The output of a narrow-band beamformer is given by

$$\hat{s} = \mathbf{w}^H \mathbf{y} \tag{2}$$

where $\hat{s}$ is the estimated signal waveform, $\mathbf{w}$ is the $M \times 1$ weight vector, and $(\cdot)^H$ stands for the Hermitian transpose. Several different strategies have been proposed to design the beamforming weights $\mathbf{w}$.

The goal of the MSINR beamformer is to maximize the output SINR. In the case of a random steering vector, the output SINR is given by

$$\mathrm{SINR} = \mathrm{E}_{\mathbf{a}}\left\{\frac{\sigma_s^2 |\mathbf{w}^H \mathbf{a}|^2}{\mathbf{w}^H \mathbf{R}_{i+n} \mathbf{w}}\right\} = \frac{\sigma_s^2 \mathbf{w}^H \mathbf{R}_a \mathbf{w}}{\mathbf{w}^H \mathbf{R}_{i+n} \mathbf{w}} \tag{3}$$

where σ_s^2 denotes the signal power, $\mathbf{R}_{i+n} \triangleq \mathrm{E}\{(\mathbf{i}+\mathbf{v})(\mathbf{i}+\mathbf{v})^H\}$ is the interference-plus-noise covariance matrix, and $\mathbf{R}_a \triangleq \mathrm{E}_{\mathbf{a}}\{\mathbf{a}\mathbf{a}^H\} = \mathbf{C} + \mathbf{m}\mathbf{m}^H$ is the correlation matrix of the steering vector. We assume that $\mathbf{R}_{i+n} \succ 0$ so that $\mathbf{R}_{i+n}$ is invertible. Throughout the paper, $\mathrm{E}_{\mathbf{a}}\{\cdot\}$ denotes the statistical expectation with respect to the random vector $\mathbf{a}$, whereas $\mathrm{E}\{\cdot\}$ stands for the statistical expectation with respect to the random noise, interference, and signal waveform. Maximizing (3) with respect to $\mathbf{w}$, the weight vector of the MSINR approach can be written as [1]

$$\mathbf{w}_{\mathrm{MSINR}} = \alpha \mathcal{P}\left\{\mathbf{R}_{i+n}^{-1} \mathbf{R}_a\right\} \tag{4}$$

where α is an arbitrary nonzero scalar and $\mathcal{P}\{\cdot\}$ stands for the principal eigenvector of a matrix. A common approach to choose α is to use the minimum variance distortionless response (MVDR) formulation of the MSINR problem

$$\min_{\mathbf{w}} \mathbf{w}^H \mathbf{R}_{i+n} \mathbf{w} \quad \text{s.t.} \quad \mathbf{w}^H \mathbf{R}_a \mathbf{w} = 1. \tag{5}$$

According to (5), the scalar α in (4) is given by

$$\alpha = \left(\mathcal{P}\left\{\mathbf{R}_{i+n}^{-1} \mathbf{R}_a\right\}^H \mathbf{R}_a \mathcal{P}\left\{\mathbf{R}_{i+n}^{-1} \mathbf{R}_a\right\}\right)^{-1/2}. \tag{6}$$

The maximal output SINR of the MSINR beamformer can be expressed as

$$\mathrm{SINR_o} = \sigma_s^2 \lambda_{\max}\left\{\mathbf{R}_{i+n}^{-1} \mathbf{R}_a\right\} \tag{7}$$

where $\lambda_{\max}\{\cdot\}$ stands for the maximal eigenvalue of a matrix.

As mentioned in the introduction, maximizing the output SINR does not necessarily lead to a good estimate of the signal waveform. When it is more important to minimize the waveform estimation error than maximize the SINR, the MMSE beamformer can be used instead of the MSINR technique. The MSE of the signal waveform estimation for one realization of the steering vector $\mathbf{a}$ is given by

$$\begin{aligned}\mathrm{E}\left\{|\hat{s} - s|^2\right\} &= \mathrm{E}\left\{\left|\mathbf{w}^H(\mathbf{a}s + \mathbf{i} + \mathbf{n}) - s\right|^2\right\} \\ &= \sigma_s^2 |1 - \mathbf{w}^H \mathbf{a}|^2 + \mathbf{w}^H \mathbf{R}_{i+n} \mathbf{w}.\end{aligned} \tag{8}$$

Taking the expectation of (8) with respect to $\mathbf{a}$ yields

$$\begin{aligned}\mathrm{MSE} &= \mathrm{E}_{\mathbf{a}}\left\{\sigma_s^2 |1 - \mathbf{w}^H \mathbf{a}|^2 + \mathbf{w}^H \mathbf{R}_{i+n} \mathbf{w}\right\} \\ &= \mathbf{w}^H \left(\mathbf{R}_{i+n} + \sigma_s^2 \mathbf{R}_a\right) \mathbf{w} - 2\mathrm{Re}\left\{\sigma_s^2 \mathbf{w}^H \mathbf{m}\right\} + \sigma_s^2\end{aligned} \tag{9}$$

where $\mathrm{Re}\{\cdot\}$ denotes the real part. The weight vector of the MMSE beamformer can be obtained by minimizing (9), resulting in [5]

$$\mathbf{w}_{\mathrm{MMSE}} = \sigma_s^2 \left(\mathbf{R}_{i+n} + \sigma_s^2 \mathbf{R}_a\right)^{-1} \mathbf{m}. \tag{10}$$

Substituting (10) back into (9), we obtain that the minimal MSE can be expressed as

$$\mathrm{MSE_o} = \sigma_s^2 - \sigma_s^4 \mathbf{m}^H \left(\mathbf{R}_{i+n} + \sigma_s^2 \mathbf{R}_a\right)^{-1} \mathbf{m}. \tag{11}$$

From (10), it is evident that the exact knowledge of the signal power is required in the MMSE beamformer. However, in practice this knowledge may be difficult to obtain. In such cases, the

MLSE approach can be applied instead of the MMSE technique. The weighted LSE for a single array snapshot can be written as

$$(\mathbf{y}-\mathbf{a}\hat{s})^H\mathbf{R}_{i+n}^{-1}(\mathbf{y}-\mathbf{a}\hat{s}). \tag{12}$$

The expectation of (12) with respect to $\mathbf{a}$ is given by

$$\begin{aligned}\mathrm{LSE} &= \mathrm{E}_{\mathbf{a}}\left\{(\mathbf{y}-\mathbf{a}\hat{s})^H\mathbf{R}_{i+n}^{-1}(\mathbf{y}-\mathbf{a}\hat{s})\right\}\\ &= (\mathbf{y}-\mathbf{m}\hat{s})^H\mathbf{R}_{i+n}^{-1}(\mathbf{y}-\mathbf{m}\hat{s})+\hat{s}^2\mathrm{Tr}\left\{\mathbf{R}_{i+n}^{-1}\mathbf{C}\right\}. \end{aligned}\tag{13}$$

Differentiating (13) with respect to $\hat{s}$ and equating the result to zero, we obtain

$$\hat{s} = \frac{1}{\mathbf{m}^H\mathbf{R}_{i+n}^{-1}\mathbf{m}+\mathrm{Tr}\left\{\mathbf{R}_{i+n}^{-1}\mathbf{C}\right\}}\mathbf{m}^H\mathbf{R}_{i+n}^{-1}\mathbf{y}. \tag{14}$$

Thus, the weight vector of the MLSE beamformer can be expressed in the following general form:

$$\mathbf{w}_{\mathrm{MLSE}} = \frac{1}{\mathbf{m}^H\mathbf{R}_{i+n}^{-1}\mathbf{m}+\mathrm{Tr}\left\{\mathbf{R}_{i+n}^{-1}\mathbf{C}\right\}}\mathbf{R}_{i+n}^{-1}\mathbf{m}+\mathbf{P}_{\mathbf{y}}^{\perp}\mathbf{u} \tag{15}$$

where $\mathbf{u}$ is an arbitrary $M\times 1$ vector, and

$$\mathbf{P}_{\mathbf{y}}^{\perp} = \mathbf{I}-\mathbf{y}\mathbf{y}^H/\mathbf{y}^H\mathbf{y}$$

is the orthogonal projector onto the subspace orthogonal to the snapshot vector $\mathbf{y}$. In Section III, we will show that by a judicious choice of $\mathbf{u}$, the MLSE beamformer can also maximize the output SINR. Choosing $\mathbf{u}=0$ results in the standard MLSE beamformer [5], [6]

$$\mathbf{w}_{\mathrm{MLSE}} = \frac{1}{\mathbf{m}^H\mathbf{R}_{i+n}^{-1}\mathbf{m}+\mathrm{Tr}\left\{\mathbf{R}_{i+n}^{-1}\mathbf{C}\right\}}\mathbf{R}_{i+n}^{-1}\mathbf{m}. \tag{16}$$

Clearly, the MLSE beamformer does not require any knowledge of the signal power. Substituting (16) back into (13), the minimal LSE can be written as

$$\mathrm{LSE_o} = \mathbf{y}^H\mathbf{R}_{i+n}^{-1}\mathbf{y}-\frac{\left|\mathbf{y}^H\mathbf{R}_{i+n}^{-1}\mathbf{m}\right|^2}{\mathbf{m}^H\mathbf{R}_{i+n}^{-1}\mathbf{m}+\mathrm{Tr}\left\{\mathbf{R}_{i+n}^{-1}\mathbf{C}\right\}}. \tag{17}$$

Note that $\mathrm{LSE_o}$ does not depend on the choice of $\mathbf{u}$.

If the knowledge of $\mathbf{R}_{i+n}$ is unavailable, then the sample covariance matrix

$$\hat{\mathbf{R}} = \frac{1}{J}\sum_{l=1}^{J}\mathbf{y}(l)\mathbf{y}^H(l) \tag{18}$$

can be used as an estimate of $\mathbf{R}_{i+n}$, where J is the number of training snapshots. It is well known that, if the sample covariance matrix $\hat{\mathbf{R}}$ is used instead of $\mathbf{R}_{i+n}$ in (5) and if the signal components are present in the training snapshots, then the performance of the MVDR beamformer can significantly degrade because of signal self-nulling [7]. Different approaches have been suggested to combat this performance degradation phenomenon. The diagonal loading (DL) approach [17], [18] is one of the most traditional techniques of that type. The diagonally loaded sample covariance matrix can be written as

$$\hat{\mathbf{R}}_{\mathrm{dl}} = \hat{\mathbf{R}}+\xi\mathbf{I} \tag{19}$$

where ξ denotes the DL factor and $\mathbf{I}$ is the identity matrix.

In practical cases where the knowledge of $\mathbf{R}_{i+n}$ is unavailable, the matrix (19) is commonly used in (4), (10) and (15) instead of $\mathbf{R}_{i+n}$.

From (4), (10), and (16), it follows that when the exact SOI steering vector is known (i.e., $\mathbf{C}=0$), the weight vectors of the MSINR, MMSE, and MLSE beamforming approaches can be simplified to

$$\mathbf{w}_{\mathrm{MSINR}} = \tilde{\alpha}\mathbf{R}_{i+n}^{-1}\mathbf{m} \tag{20}$$

$$\mathbf{w}_{\mathrm{MMSE}} = \frac{\sigma_s^2}{1+\sigma_s^2\mathbf{m}^H\mathbf{R}_{i+n}^{-1}\mathbf{m}}\mathbf{R}_{i+n}^{-1}\mathbf{m} \tag{21}$$

$$\mathbf{w}_{\mathrm{MLSE}} = \frac{1}{\mathbf{m}^H\mathbf{R}_{i+n}^{-1}\mathbf{m}}\mathbf{R}_{i+n}^{-1}\mathbf{m} \tag{22}$$

where $\tilde{\alpha}$ is an arbitrary nonzero scalar. In the case of the MVDR beamformer, $\tilde{\alpha}=1/(\mathbf{m}^H\mathbf{R}_{i+n}^{-1}\mathbf{m})$. Evidently, in this case the weight vectors differ only in their scaling factors. Therefore, when the SOI steering vector and power are exactly known, the MMSE beamformer should be chosen, since it simultaneously minimizes the MSE and maximizes the SINR. If the signal power is unknown, then the MLSE beamformer can be used because it optimizes both the MLSE and MSINR criteria. Moreover, comparing (21) with (22), we see that the MLSE beamformer asymptotically minimizes the MSE (that is, for a high signal power).

Unfortunately, these conclusions do not hold in the case of random steering vector, i.e., when $\mathbf{C}\neq 0$. This case will be considered in the next section.

III. PERFORMANCE TRADEOFFS

A. MMSE Versus MSINR

We begin by discussing the tradeoff between the MSE and SINR criteria. Comparing (4) with (10), we see that in general the MMSE beamformer is not proportional to the SINR beamformer, and, therefore, will not in general maximize the output SINR. As any method is associated with some point in the MSE-SINR plane, a good approach should yield large SINR and small MSE. However, these two objectives are typically contradictive to each other, that is, not all points on the plane are achievable.

To characterize the fundamental tradeoff between SINR and MSE, we consider the boundary of the achievable points, namely the points $(\gamma, \min_{\mathbf{w}\in\mathcal{W}}\mathrm{MSE})$ where $\mathcal{W}$ is the set of beamformers for which $1/\mathrm{SINR}\leq\gamma$. This boundary defines a curve in the MSE-SINR plane which divides this plane into two parts. All points on one part of the plane are achievable, while the points on the other part cannot be achieved. Fig. 1 illustrates an example of such a boundary in the MSE-SINR plane. Here, in order to bring the values of MSE and SINR to the same order, they are normalized with respect to $\mathrm{MSE_o}$ and $\mathrm{SINR_o}$, respectively. For some given γ, the diamond mark is associated with the OPTO beamformer, defined below in (23), while the achievable part of the MSE-SINR plane for this value of γ is marked by the shaded area. Two important points are also shown in Fig. 1. The square mark on the curve corresponds to the beamformer that achieves the MMSE criterion with the

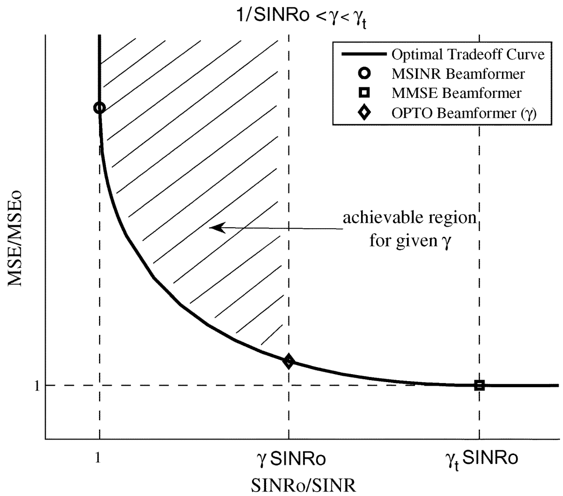


Fig. 1. Qualitative example of the optimal tradeoff curve in the MSE-SINR plane.

largest attainable SINR, while the circle mark represents the beamformer which has the smallest MSE among all MSINR beamformers. The meaning of the parameter γ_t will be clarified below.

The points on the boundary of the MSE-SINR tradeoff curve can be found using the OPTO beamformer which minimizes the MSE subject to some given SINR constraint:

$$\begin{aligned}\min_{\mathbf{w}} \quad & \text{MSE} \\ \text{s.t.} \quad & \frac{1}{\text{SINR}} \leq \gamma. \end{aligned} \tag{23}$$

Here, SINR and MSE are defined by (3) and (9), respectively, and γ is some given positive scalar. The OPTO beamformer is Pareto optimal [15], [16], that is, there is no other approach which yields both smaller MSE and larger SINR values.

In order to have a feasible optimization problem, γ must satisfy $\gamma \geq 1/\text{SINR}_\text{o}$, where SINR_o is given by (7). Substituting (3) and (9) into (23) and assuming that the signal power is known, our problem becomes

$$\begin{aligned}\min_{\mathbf{w}} \quad & \mathbf{w}^H\left(\mathbf{R}_{i+n} + \sigma_s^2\mathbf{R}_a\right)\mathbf{w} - 2\text{Re}\left\{\sigma_s^2\mathbf{w}^H\mathbf{m}\right\} + \sigma_s^2 \\ \text{s.t.} \quad & \mathbf{w}^H\left(\mathbf{R}_{i+n} - \gamma\sigma_s^2\mathbf{R}_a\right)\mathbf{w} \leq 0. \end{aligned} \tag{24}$$

We now develop the solution to (24) for different regions of γ.

Case 1: When $\gamma = 1/\text{SINR}_\text{o}$, problem (24) is equivalent to the design of a beamformer that minimizes the MSE under the maximal (optimal) SINR constraint. The weight vector of such a beamformer can be represented as $\mathbf{w} = c\mathcal{P}\{\mathbf{R}_{i+n}^{-1}\mathbf{R}_a\}$, where c is a nonzero scalar that should be chosen to yield the minimal MSE. Thus, the optimal c can be computed as

$$\begin{aligned}c_\text{opt} &= \arg\min_c \left\{\mathbf{w}^H\left(\mathbf{R}_{i+n} + \sigma_s^2\mathbf{R}_a\right)\mathbf{w} \right.\\ &\qquad \left. -2\text{Re}\left\{\sigma_s^2\mathbf{w}^H\mathbf{m}\right\} + \sigma_s^2\right\} \\ &= \arg\min_c \left\{t_1|c|^2 - 2\text{Re}\{t_2 c\} + \sigma_s^2\right\} \end{aligned} \tag{25}$$

where

$$\mathbf{w} = c\mathcal{P}\left\{\mathbf{R}_{i+n}^{-1}\mathbf{R}_a\right\} \tag{26}$$

$$t_1 \triangleq \mathcal{P}\left\{\mathbf{R}_{i+n}^{-1}\mathbf{R}_a\right\}^H\left(\mathbf{R}_{i+n} + \sigma_s^2\mathbf{R}_a\right)\mathcal{P}\left\{\mathbf{R}_{i+n}^{-1}\mathbf{R}_a\right\} \tag{27}$$

$$t_2 \triangleq \sigma_s^2\mathbf{m}^H\mathcal{P}\left\{\mathbf{R}_{i+n}^{-1}\mathbf{R}_a\right\}. \tag{28}$$

Since $t_1 > 0$, we have $c_\text{opt} = t_2^*/t_1$, where $(\cdot)^*$ stands for the complex conjugate and therefore

$$\mathbf{w} = \frac{t_2^*}{t_1}\mathcal{P}\left\{\mathbf{R}_{i+n}^{-1}\mathbf{R}_a\right\}. \tag{29}$$

Inserting (29) into (9), the MSE can be written as

$$\text{MSE}_1 = \sigma_s^2 - \frac{\sigma_s^4\left|\mathcal{P}\left\{\mathbf{R}_{i+n}^{-1}\mathbf{R}_a\right\}^H\mathbf{m}\right|^2}{\mathcal{P}\left\{\mathbf{R}_{i+n}^{-1}\mathbf{R}_a\right\}^H\left(\mathbf{R}_{i+n} + \sigma_s^2\mathbf{R}_a\right)\mathcal{P}\left\{\mathbf{R}_{i+n}^{-1}\mathbf{R}_a\right\}}. \tag{30}$$

If $\mathbf{C} = 0$, then $\mathbf{R}_a = \mathbf{m}\mathbf{m}^H$ and (30) can be simplified to

$$\text{MSE}_1 = \frac{\sigma_s^2}{1 + \sigma_s^2\mathbf{m}^H\mathbf{R}_{i+n}^{-1}\mathbf{m}} \tag{31}$$

which is equivalent to the MSE of the MMSE beamformer. This is an expected result because if $\mathbf{C} = 0$, then (29) is equivalent to the weight vector (21) of the MMSE beamformer.

If $\mathbf{C} \neq 0$, then using (30), we have that

$$\begin{aligned}\text{MSE}_1 &\geq \min_{\mathbf{v}}\left\{\sigma_s^2 - \frac{\sigma_s^4\left|\mathbf{v}^H\mathbf{m}\right|^2}{\mathbf{v}^H\left(\mathbf{R}_{i+n} + \sigma_s^2\mathbf{R}_a\right)\mathbf{v}}\right\} \\ &= \sigma_s^2 - \mathbf{m}^H\left(\mathbf{R}_{i+n} + \sigma_s^2\mathbf{R}_a\right)^{-1}\mathbf{m} \\ &= \text{MSE}_\text{o} \end{aligned} \tag{32}$$

where we have taken into account that the function $\sigma_s^2 - \sigma_s^4|\mathbf{v}^H\mathbf{m}|^2/\mathbf{v}^H(\mathbf{R}_{i+n} + \sigma_s^2\mathbf{R}_a)\mathbf{v}$ is minimized by $\mathbf{v} \propto (\mathbf{R}_{i+n} + \sigma_s^2\mathbf{R}_a)^{-1}\mathbf{m}$. Equality in (32) can be achieved if and only if

$$\mathcal{P}\left\{\mathbf{R}_{i+n}^{-1}\mathbf{R}_a\right\} \propto \left(\mathbf{R}_{i+n} + \sigma_s^2\mathbf{R}_a\right)^{-1}\mathbf{m} \tag{33}$$

which is true only in rather specific cases. In these situations the SINR and MSE can be jointly optimized. One such case is when $\mathbf{R}_a$ is proportional to $\mathbf{R}_{i+n}$, i.e.

$$\mathbf{R}_a = \rho\mathbf{R}_{i+n} \tag{34}$$

where ρ is a positive scalar. Under this condition we have $\mathbf{R}_{i+n}^{-1}\mathbf{R}_a = \rho\mathbf{I}$, and from (4) it is clear that $\mathbf{w}_\text{MSINR}$ can be any $M \times 1$ nonzero vector. Thus, both MSE and SINR can be simultaneously optimized using the MMSE beamformer. Another example is when $\mathbf{C} = k\mathbf{m}\mathbf{m}^H$ for some $k \geq 0$.

However, in the general case it follows from (32) that the beamformer (29) has a *larger* MSE than the MMSE technique. This indicates that typically the optimal MSE and SINR values cannot be achieved simultaneously. □

Case 2: We next consider the interval $\gamma \geq \gamma_t$ where $1/\gamma_t$ is the SINR achieved when using the MMSE beamformer of (10). The parameter γ_t is given by

$$\gamma_t \triangleq \frac{\mathbf{m}^H \left(\mathbf{R}_{i+n} + \sigma_s^2 \mathbf{R}_a\right)^{-1} \mathbf{R}_{i+n} \left(\mathbf{R}_{i+n} + \sigma_s^2 \mathbf{R}_a\right)^{-1} \mathbf{m}}{\sigma_s^2 \mathbf{m}^H \left(\mathbf{R}_{i+n} + \sigma_s^2 \mathbf{R}_a\right)^{-1} \mathbf{R}_a \left(\mathbf{R}_{i+n} + \sigma_s^2 \mathbf{R}_a\right)^{-1} \mathbf{m}}. \quad (35)$$

Clearly, in this regime the MSE can be minimized since the constraint in (24) will be satisfied for the MMSE beamformer. Thus, the optimal beamformer is $\mathbf{w}_{\mathrm{MMSE}}$. □

Case 3: The most interesting case is when $1/\mathrm{SINR_o} < \gamma < \gamma_t$. In this regime, the matrix $\mathbf{R}_{i+n} - \gamma\sigma_s^2\mathbf{R}_a$ in the constraint (24) is not positive semi-definite, as incorporated in the following proposition.

Proposition 1: Let $1/\mathrm{SINR_o} < \gamma < \gamma_t$. Then the matrix $\mathbf{Z} = \mathbf{R}_{i+n} - \gamma\sigma_s^2\mathbf{R}_a$ is not positive semi-definite.

Proof: To prove the proposition recall from (7) that $\mathrm{SINR_o} = \sigma_s^2 \lambda_{\max}\{\mathbf{R}_{i+n}^{-1}\mathbf{R}_a\}$. Therefore

$$\sigma_s^2 \mathbf{R}_{i+n}^{-1/2} \mathbf{R}_a \mathbf{R}_{i+n}^{-1/2} \preceq \mathrm{SINR_o}\mathbf{I}$$

or

$$\frac{\sigma_s^2}{\mathrm{SINR_o}} \mathbf{R}_a \preceq \mathbf{R}_{i+n}$$

and, furthermore, $\mathbf{R}_{i+n} - (\sigma_s^2/\mathrm{SINR_o})\mathbf{R}_a$ has at least one zero eigenvalue. Since $\gamma > 1/\mathrm{SINR_o}$ it follows immediately that $\mathbf{Z}$ has at least one negative eigenvalue, completing the proof.

From Proposition 1 it follows that (24) is nonconvex. We can also conclude that there always exists some $\mathbf{w}$ such that $\mathbf{w}^H(\mathbf{R}_{i+n} - \gamma\sigma_s^2\mathbf{R}_a)\mathbf{w} < 0$, so that (24) is strictly feasible. Therefore, this problem belongs to the class of quadratically constrained quadratic optimization problems. It is well known that for this class of problems, even when the constraint is nonconvex, strong duality holds [19]–[21]. The solution to (24) can thus be obtained via the dual program and the necessary and sufficient optimality conditions that follow from strong duality.

To obtain the dual problem, we first write the Lagrangian function as

$$\begin{aligned}\mathcal{L}(\mathbf{w}, \mu) &= \mathbf{w}^H \left(\mathbf{R}_{i+n} + \sigma_s^2 \mathbf{R}_a\right) \mathbf{w} - 2\mathrm{Re}\left(\sigma_s^2 \mathbf{w}^H \mathbf{m}\right) \\ &\quad + \sigma_s^2 + \mu \mathbf{w}^H \left(\mathbf{R}_{i+n} - \gamma\sigma_s^2 \mathbf{R}_a\right) \mathbf{w} \\ &= \mathbf{w}^H \mathbf{T}(\mu) \mathbf{w} - 2\mathrm{Re}\left(\sigma_s^2 \mathbf{w}^H \mathbf{m}\right) + \sigma_s^2 \end{aligned} \quad (36)$$

where $\mu \geq 0$ is the Lagrange multiplier, and

$$\mathbf{T}(\mu) \triangleq \mathbf{R}_{i+n} + \sigma_s^2 \mathbf{R}_a + \mu \left(\mathbf{R}_{i+n} - \gamma\sigma_s^2 \mathbf{R}_a\right). \quad (37)$$

To ensure that the Lagrangian is bounded from below, we must have that $\mathbf{T}(\mu) \succeq 0$. The Lagrangian dual function associated with (24) can then be obtained by minimizing (36), resulting in

$$g(\mu) = \sigma_s^2 - \sigma_s^4 \mathbf{m}^H \mathbf{T}^\dagger(\mu) \mathbf{m} \quad (38)$$

where $(\cdot)^\dagger$ denotes the pseudo-inverse that is used here because $\mathbf{T}(\mu)$ can be rank-deficient. Finally, the dual problem is obtained by maximizing (38), which can be equivalently written as [19], [20]

$$\begin{aligned} \max_{\delta,\mu} \quad & \delta \\ \text{s.t.} \quad & \begin{bmatrix} \mathbf{T}(\mu) & \sigma_s^2 \mathbf{m} \\ \sigma_s^2 \mathbf{m}^H & \sigma_s^2 - \delta \end{bmatrix} \succeq 0 \\ & \mu \geq 0. \end{aligned} \quad (39)$$

This is a convex semi-definite programming (SDP) problem that can be solved using modern convex optimization tools [22].

Using the optimality conditions for this class of problems [20], it follows that if μ is a solution of (39), then $\mathbf{w}$ is optimal for (24) if and only if

$$\mathbf{T}(\mu)\mathbf{w} = \sigma_s^2 \mathbf{m} \quad (40)$$

$$\mu \mathbf{w}^H \left(\mathbf{R}_{i+n} - \gamma\sigma_s^2 \mathbf{R}_a\right) \mathbf{w} = 0 \quad (41)$$

$$\mathbf{w}^H \left(\mathbf{R}_{i+n} - \gamma\sigma_s^2 \mathbf{R}_a\right) \mathbf{w} \leq 0. \quad (42)$$

Once we solve the dual program, we can explicitly compute $\mathbf{T}(\mu)$. If $\mathbf{T}(\mu) \succ 0$, then it follows from (40) that

$$\begin{aligned}\mathbf{w} &= \sigma_s^2 \mathbf{T}^{-1}(\mu) \mathbf{m} \\ &= \sigma_s^2 \left(\mathbf{R}_{i+n} + \sigma_s^2 \mathbf{R}_a + \mu \left(\mathbf{R}_{i+n} - \gamma\sigma_s^2 \mathbf{R}_a\right)\right)^{-1} \mathbf{m}. \end{aligned} \quad (43)$$

Comparing the weight vectors (10) and (43), it can be seen that their structure is similar. In particular, (43) has an additional term $\mu(\mathbf{R}_{i+n} - \gamma\sigma_s^2\mathbf{R}_a)$ with respect to (10). This extra term depends on the variables μ and γ and reflects the effect of taking into account both the SINR and MSE criteria in our beamformer design.

An interesting special case of (43) results when $\mathbf{C} = d\mathbf{I}$, where $d > 0$. In this case, (43) can be rewritten as

$$\mathbf{w} = \frac{1}{(1+\mu)/\sigma_s^2 + (1-\mu\gamma)\mathbf{m}^H(\mathbf{R}_{i+n} + \varepsilon\mathbf{I})^{-1}\mathbf{m}} \cdot (\mathbf{R}_{i+n} + \varepsilon\mathbf{I})^{-1}\mathbf{m} \quad (44)$$

where $\varepsilon \triangleq \sigma_s^2 d(1-\mu\gamma)/(1+\mu)$. The term $\varepsilon\mathbf{I}$ in (44) can be interpreted as an adaptive diagonal loading term added to the covariance matrix $\mathbf{R}_{i+n}$.

Computing the optimal $\mathbf{w}$ when $\mathbf{T}(\mu)$ is not positive definite is more involved. In this case, (40) can be written as

$$\mathbf{w} = \sigma_s^2 \mathbf{T}^\dagger(\mu)\mathbf{m} + \mathbf{B}\mathbf{w}_2 \quad (45)$$

where the columns of $\mathbf{B}$ form a basis for the null-space of $\mathbf{T}(\mu)$, and $\mathbf{w}_2$ is a $(M - \mathrm{rank}\{\mathbf{T}(\mu)\}) \times 1$ vector. Note that in this case, we must have $\mu \neq 0$, since, otherwise, if $\mu = 0$, then $\mathbf{T} = \mathbf{R}_{i+n} + \sigma_s^2\mathbf{R}_a$ is a positive definite matrix. It follows that (45) is an optimal solution of (24) if and only if (42) is satisfied with equality which leads to a quadratic equation. The problem can be easily solved using the method described in [20] and [21]. □

Quantitative examples of the optimal tradeoff curves are shown in Fig. 2. In this example, a uniform linear array (ULA) is assumed with $M = 5$ sensors spaced half a wavelength apart. A random SOI steering vector with mean value $\mathbf{m} = [1, e^{j\pi\sin\theta_s}, \cdots, e^{j(M-1)\pi\sin\theta_s}]^T$ is assumed where θ_s is the nominal SOI direction-of-arrival (DOA) that is equal to 30° with respect to the array broadside direction. An exponential decay model is used for $\mathbf{C}$, i.e., the (l, k)-th element of this

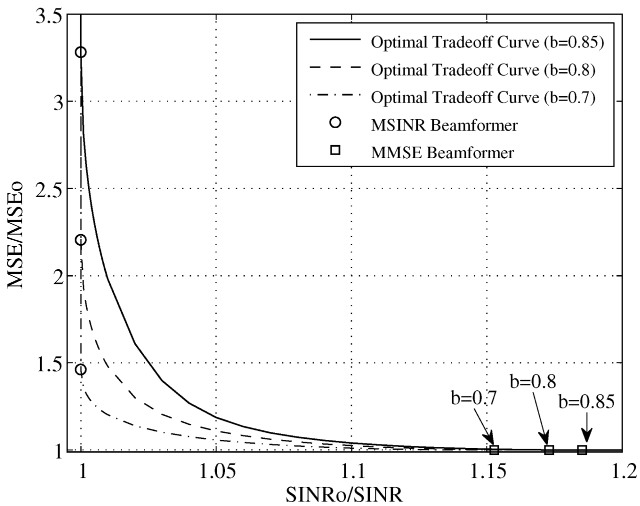


Fig. 2. Quantitative example of the optimal tradeoff curves in the MSE-SINR plane.

matrix is given by $[\mathbf{C}]_{l,k} = b^{|l-k|}; l,k = 1,\ldots,M$. Here, $0 \leq b \leq 1$ is a parameter that characterizes the level of the steering vector uncertainty and is chosen equal to 0.7, 0.8, and 0.85. The interference is modelled as $\mathbf{i}(t) = i(t)\mathbf{a}_i$ where $i(t)$ is a zero mean complex Gaussian interference waveform, and $\mathbf{a}_i = [1, e^{j\pi\sin\theta_i}, \cdots, e^{j(M-1)\pi\sin\theta_i}]^T$ is a non-random interference steering vector. The interference DOA is $\theta_i = -30^\circ$ with respect to the array broadside direction. The signal and interference powers σ_s^2 and σ_i^2 are set to be 10 and 20 dB above the noise level, respectively.

Fig. 2 displays the optimal tradeoff curves which represent the frontiers between the achievable and non-achievable regions in the MSE-SINR plane for different values of b. For each curve, the region above the curve is attainable, while the region below the curve is non-achievable. The square mark on each curve corresponds to the MMSE beamformer for which $\mathrm{MSE_o}$ is achieved and $\mathrm{SINR} = \gamma_t$. We see that when $\gamma \geq \gamma_t$, $\mathrm{MSE_o}$ is always achieved. The circle mark represents the beamformer (29), which has the smallest MSE among all MSINR beamformers. Evidently, achieving the MSE-SINR tradeoff is a more challenging task for larger values of b.

The optimal tradeoff curves appear to be an important tool to analyze the MSE and SINR tradeoffs because a beamformer associated with any point on this curve can be obtained from (43).

Note that in our development so far, we have assumed that σ_s^2 is exactly known. In practice, σ_s^2 can be estimated from the received data, for example, using conventional beamforming techniques. However, such estimates may be inaccurate in certain cases, and this may affect the beamformer design. Therefore, it is interesting to obtain an alternative performance measure that does not require any knowledge of the signal power. Such an approach is discussed in the next section.

B. MLSE Versus MSINR

To circumvent the need for estimating the SOI power, we consider the MLSE beamformer [5], [6]. In this subsection, we show that the optimal LSE and SINR, namely $\mathrm{LSE_o}$ and $\mathrm{SINR_o}$, can be jointly achieved by the MLSE method (15) with a judiciously chosen $\mathbf{u}$. We refer to the resulting approach as the *optimal* MLSE beamformer.

For simplicity, define

$$a_1 \triangleq \mathbf{m}^H\mathbf{R}_{i+n}^{-1}\mathbf{m} + \mathrm{Tr}\left\{\mathbf{R}_{i+n}^{-1}\mathbf{C}\right\} \tag{46}$$

$$a_2 \triangleq \mathbf{y}^H\mathbf{R}_{i+n}^{-1}\mathbf{m}. \tag{47}$$

Since we have freedom in choosing $\mathbf{u}$ in (15), we suggest to select it such that the resulting beamformer maximizes the output SINR. Comparing (4) and (15), we find that such an optimal choice of $\mathbf{u}$ should satisfy

$$\beta\mathcal{P}\left\{\mathbf{R}_{i+n}^{-1}\mathbf{R}_a\right\} = \frac{1}{a_1}\mathbf{R}_{i+n}^{-1}\mathbf{m} + \mathbf{P}_{\mathbf{y}}^{\perp}\mathbf{u} \tag{48}$$

where β is a scalar. Orthogonally projecting both sides of (48) onto the subspace spanned by $\mathbf{y}$ and onto its orthogonal compliment, respectively, we have

$$\beta\mathbf{P}_{\mathbf{y}}\mathcal{P}\left\{\mathbf{R}_{i+n}^{-1}\mathbf{R}_a\right\} = \frac{1}{a_1}\mathbf{P}_{\mathbf{y}}\mathbf{R}_{i+n}^{-1}\mathbf{m} \tag{49}$$

$$\beta\mathbf{P}_{\mathbf{y}}^{\perp}\mathcal{P}\left\{\mathbf{R}_{i+n}^{-1}\mathbf{R}_a\right\} = \frac{1}{a_1}\mathbf{P}_{\mathbf{y}}^{\perp}\mathbf{R}_{i+n}^{-1}\mathbf{m} + \mathbf{P}_{\mathbf{y}}^{\perp}\mathbf{u} \tag{50}$$

where $\mathbf{P}_{\mathbf{y}} = \mathbf{y}\mathbf{y}^H/\mathbf{y}^H\mathbf{y}$. From (49)

$$\beta = \frac{\mathbf{y}^H\mathbf{R}_{i+n}^{-1}\mathbf{m}}{a_1\mathbf{y}^H\mathcal{P}\left\{\mathbf{R}_{i+n}^{-1}\mathbf{R}_a\right\}}. \tag{51}$$

Inserting (51) into (50), we have

$$\mathbf{P}_{\mathbf{y}}^{\perp}\mathbf{u} = \frac{1}{a_1}\left[\frac{\mathbf{y}^H\mathbf{R}_{i+n}^{-1}\mathbf{m}}{\mathbf{y}^H\mathcal{P}\left\{\mathbf{R}_{i+n}^{-1}\mathbf{R}_a\right\}}\mathcal{P}\left\{\mathbf{R}_{i+n}^{-1}\mathbf{R}_a\right\} - \mathbf{R}_{i+n}^{-1}\mathbf{m}\right]. \tag{52}$$

Finally, substituting (51) and (52) back into (15), we obtain

$$\mathbf{w}_{\mathrm{MLSE,o}} = \frac{\mathbf{y}^H\mathbf{R}_{i+n}^{-1}\mathbf{m}}{\left(\mathbf{m}^H\mathbf{R}_{i+n}^{-1}\mathbf{m} + \mathrm{Tr}\left\{\mathbf{R}_{i+n}^{-1}\mathbf{C}\right\}\right)\mathbf{y}^H\mathcal{P}\left\{\mathbf{R}_{i+n}^{-1}\mathbf{R}_a\right\}} \cdot\mathcal{P}\left\{\mathbf{R}_{i+n}^{-1}\mathbf{R}_a\right\} \tag{53}$$

where the subscript "o" implies that (53) is the weight vector of the optimal MLSE beamformer.

Comparing (53) and (4), we see that in the random signal case the optimal MLSE beamformer is a particular choice of the MSINR beamformer and, therefore, in addition to minimizing LSE this approach also maximizes SINR. In fact, (53) re-scales the weight vector of the MSINR beamformer by a properly chosen time-varying factor that corresponds to the minimal LSE. Note that the scalar coefficient β in (51) depends on the array observation $\mathbf{y}$ and, therefore, this coefficient should be updated every new snapshot. However, as the computation of β requires only two vector multiplications, such an update can be easily computed. Interestingly, if $\mathbf{C} = 0$, then $\mathcal{P}\{\mathbf{R}_{i+n}^{-1}\mathbf{R}_a\} \propto \mathbf{R}_{i+n}^{-1}\mathbf{m}$ and (53) becomes equivalent to the conventional MVDR method.

Fig. 3 displays the LSE-SINR tradeoff curve illustrating improvements in SINR achieved by the optimal MLSE beamformer with respect to the standard one (corresponding to $\mathbf{u} = 0$). In this figure, we assume a covariance matrix $\mathbf{C}$ with $[\mathbf{C}]_{l,k} = 3.5 \times 0.9^{|l-k|}; l,k = 1,\ldots,M$. All other parameters are the same as in Fig. 2. It can be seen from Fig. 3 that the

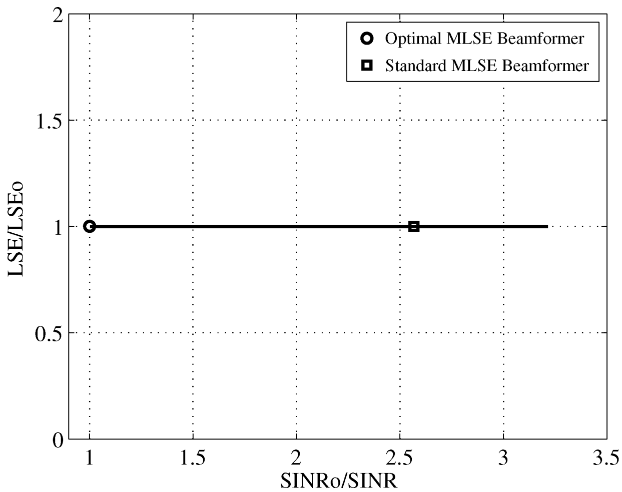


Fig. 3. Performance gain of the optimal MLSE beamformer over the standard MLSE beamformer.

optimal values of SINR and LSE are jointly achievable and quite a dramatic SINR improvement can be obtained when using the optimal MLSE beamformer with appropriate choice of $\mathbf{u}$.

C. Applications of Sensitivity Theory

As shown in Section III-A, in the general case, the MMSE and MSINR criteria cannot be simultaneously attained. The performance tradeoff between these two criteria can be qualitatively characterized by means of the *sensitivity theory* of constrained optimization [19].

Denote the objective and the constraint functions of (24) as

$$q(\mathbf{w}) \triangleq \mathbf{w}^H\left(\mathbf{R}_{i+n}+\sigma_s^2\mathbf{R}_a\right)\mathbf{w}-2\mathrm{Re}\left\{\sigma_s^2\mathbf{w}^H\mathbf{m}\right\}+\sigma_s^2 \quad (54)$$

$$h(\mathbf{w}) \triangleq \frac{\mathbf{w}^H\mathbf{R}_{i+n}\mathbf{w}}{\sigma_s^2\mathbf{w}^H\mathbf{R}_a\mathbf{w}}-\gamma \quad (55)$$

respectively. Using these simplified notations, (24) can be compactly written as

$$\inf\left\{q(\mathbf{w})|h(\mathbf{w})\leq 0\right\}. \quad (56)$$

Let us introduce a new variable ν as a perturbation to the constraint function of (24), that is, $h(\mathbf{w}) \leq \nu$. Correspondingly, we denote the optimal value of the perturbed optimization problem as $p_{\mathrm{o}}(\nu)$, that is

$$p_{\mathrm{o}}(\nu)=\inf\left\{q(\mathbf{w})|h(\mathbf{w})\leq \nu\right\}. \quad (57)$$

Clearly, the optimal value of the unperturbed problem (24) is $p_{\mathrm{o}}(0)$. Denoting the optimal solution for μ in the dual problem (39) as μ_{o}, from the properties of the Lagrange dual function in (38), we have

$$g(\mu_{\mathrm{o}}) \leq q(\mathbf{w})+\mu_{\mathrm{o}}h(\mathbf{w}). \quad (58)$$

As the strong duality property holds for the optimization problem (24) and its dual problem (that is, the optimal duality gap is zero), we have $p_{\mathrm{o}}(0)=g(\mu_{\mathrm{o}}) \leq q(\mathbf{w})+\mu_{\mathrm{o}}h(\mathbf{w}) \leq q(\mathbf{w})+\mu_{\mathrm{o}}\nu$, which implies that $q(\mathbf{w}) \geq p_{\mathrm{o}}(0)-\mu_{\mathrm{o}}\nu$. Hence

$$p_{\mathrm{o}}(\nu) \geq p_{\mathrm{o}}(0)-\mu_{\mathrm{o}}\nu. \quad (59)$$

From (59), we can see that if the value of the Lagrange multiplier μ_{o} is large, then the optimal value $p_{\mathrm{o}}(\nu)$ is more sensitive to negative perturbations ν of the constraint function. Thus, the performance tradeoff represents a more challenging problem in this particular region of the MSE-SINR plane than in other regions where μ_{o} is small.

We now discuss three particular examples to support the latter conclusion.

Example 1: $\mathbf{R}_a \propto \mathbf{R}_{i+n}$ so that $\mathbf{R}_{i+n}^{-1}\mathbf{R}_a=\rho\mathbf{I}$. The dual problem can be written as

$$\max_{\delta,\mu}\delta \quad \text{s.t.} \quad \delta \leq \sigma_s^2-\frac{\sigma_s^4\mathbf{m}^H\mathbf{R}_{i+n}^{-1}\mathbf{m}}{1+\sigma_s^2\rho+\mu\left(1-\gamma\rho\sigma_s^2\right)}, \quad \mu\geq 0. \quad (60)$$

From (60), it can be seen that, if $\gamma > 1/\mathrm{SINR}_{\mathrm{o}}=1/(\rho\sigma_s^2)$, then $\mu_{\mathrm{o}}=0$. This agrees with our analysis in Section III-A, where it has been shown that $\mathrm{SINR}_{\mathrm{o}}$ and $\mathrm{MSE}_{\mathrm{o}}$ can be jointly achieved in this particular case.

Example 2: $\mathbf{C}=0$. From the dual problem (39)

$$\begin{aligned}\delta &\leq \sigma_s^2-\sigma_s^4\mathbf{m}^H \\ &\quad\cdot\left[\mathbf{R}_{i+n}+\sigma_s^2\mathbf{m}\mathbf{m}^H+\mu\left(\mathbf{R}_{i+n}-\gamma\sigma_s^2\mathbf{m}\mathbf{m}^H\right)\right]^{-1}\mathbf{m} \\ &=\sigma_s^2-\frac{\sigma_s^4\mathbf{m}^H\mathbf{R}_{i+n}^{-1}\mathbf{m}}{1+\sigma_s^2\mathbf{m}^H\mathbf{R}_{i+n}^{-1}\mathbf{m}+\mu\left(1-\sigma_s^2\gamma\mathbf{m}^H\mathbf{R}_{i+n}^{-1}\mathbf{m}\right)}. \quad (61)\end{aligned}$$

Thus, if $\gamma > 1/\mathrm{SINR}_{\mathrm{o}}=1/(\sigma_s^2\mathbf{m}^H\mathbf{R}_{i+n}^{-1}\mathbf{m})$, then $\mu_{\mathrm{o}}=0$. This is in agreement with our result that $\mathrm{SINR}_{\mathrm{o}}$ and $\mathrm{MSE}_{\mathrm{o}}$ can be jointly attained in this setting.

Example 3: $\gamma \geq \gamma_t$. In this region of γ, we have $\mu_{\mathrm{o}}=0$ and $\mathrm{MSE}_{\mathrm{o}}$ is achievable.

In the general case, the analysis of the dual problem (39) reveals that, if $\gamma \to 1/\mathrm{SINR}_{\mathrm{o}}$, then maximizing δ leads to $\mu_{\mathrm{o}} \to \infty$. Thus, the performance tradeoff is more difficult to achieve in this particular region. This observation will be confirmed by numerical simulations in the next section.

IV. Simulations

Since the MLSE and MSINR criteria are simultaneously achievable, we only need to study the tradeoffs between the MMSE and MSINR criteria. In our simulations, a ULA with $M=5$ antenna elements spaced half a wavelength apart is assumed. Throughout the examples, we use a random SOI steering vector with mean $\mathbf{m}=[1,e^{j\pi\sin\theta_s},\cdots,e^{j(M-1)\pi\sin\theta_s}]^T$ and $\theta_s=30^\circ$. The interference is modelled as $\mathbf{i}(t)=i(t)\mathbf{a}_i$ where $i(t)$ is zero mean complex Gaussian and $\mathbf{a}_i$ is a non-random (plane-wave) interference steering vector with $\theta_i=-30^\circ$.

In each example, the results are averaged over 10^4 independent simulation runs. In each simulation run, the actual SOI steering vector $\mathbf{a}$ is generated as a complex circular Gaussian distributed random vector with mean $\mathbf{m}$ and covariance matrix $\mathbf{C}$ with $[\mathbf{C}]_{l,k}=b^{|l-k|}; l,k=1,\ldots,M$. The MSE of the signal

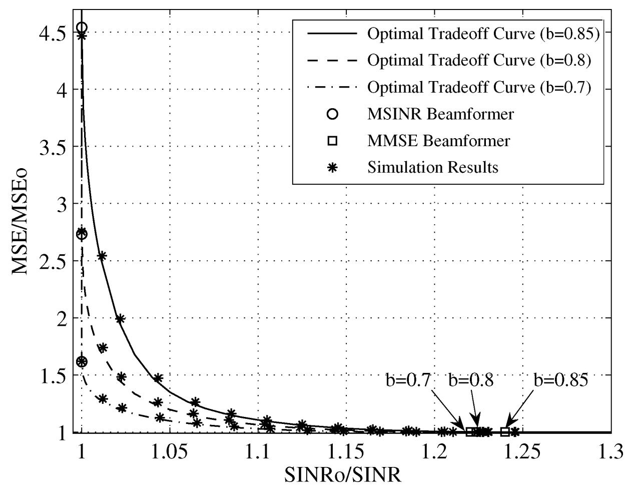


Fig. 4. MSE-SINR tradeoff for $\mathrm{INR} = 15$ dB and $\mathrm{SNR} = 20$ dB with exactly known $\mathbf{R}_{i+n}$.

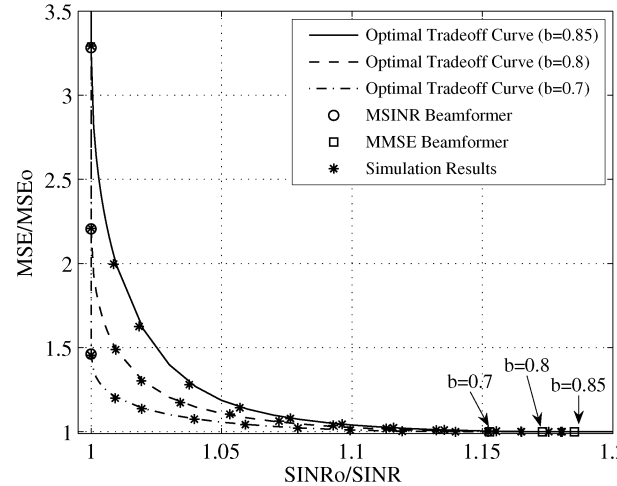


Fig. 5. MSE-SINR tradeoff for $\mathrm{INR} = 20$ dB and $\mathrm{SNR} = 10$ dB with exactly known $\mathbf{R}_{i+n}$.

waveform estimation and the output SINR in each run are averaged over $L = 10^4$ data realizations and computed as

$$\mathrm{MSE} = \frac{1}{L}\sum_{t=1}^{L}\left|\mathbf{w}^H\mathbf{y}(t) - s(t)\right|^2 \tag{62}$$

$$\mathrm{SINR} = \frac{\sum_{t=1}^{L}\left|\mathbf{w}^H\mathbf{a}s(t)\right|^2}{\sum_{t=1}^{L}\left|\mathbf{w}^H\left(\mathbf{i}(t)+\mathbf{v}(t)\right)\right|^2} \tag{63}$$

respectively. Throughout all our examples, we assume that σ_s^2 is exactly known.

In the first example, $\mathbf{R}_{i+n}$ is assumed to be available without the signal component, i.e., it is computed as

$$\mathbf{R}_{i+n} = \sigma_i^2\mathbf{a}_i\mathbf{a}_i^H + \sigma_v^2\mathbf{I} \tag{64}$$

where σ_i^2 and σ_v^2 denote the interference and noise powers, respectively. We simulate a scenario where the interference-to-noise ratio (INR) is 15 dB and the signal-to-noise ratio (SNR) is 20 dB in a single antenna element, respectively. Fig. 4 compares the theoretical and simulated MSE-SINR tradeoff curves for $b = 0.7$, $b = 0.8$, and $b = 0.85$. The theoretical MSEs and SINRs are calculated using (9) and (3), while the simulated points are obtained using the proposed beamformer (24) with different values of γ.

In the second example, we consider a scenario where $\mathrm{INR} \gg \mathrm{SNR}$ and take $\mathrm{INR} = 20$ dB and $\mathrm{SNR} = 10$ dB in a single antenna element. For these parameters, Fig. 5 displays similar curves, as shown in Fig. 4.

From Figs. 4 and 5, it can be seen that the numerical and theoretical results coincide, and that the curves in the MSE-SINR plane describe the optimal tradeoff between the MMSE and MSINR criteria. It is also evident from both figures, that achieving the MSE-SINR tradeoff is a more challenging task for larger values of b. Comparing Fig. 4 with Fig. 5, we observe that the shape of the tradeoff curves also depends on the particular scenario parameters, for example, the relationship between SNR and INR.

In the third example, we assume that the exact knowledge of $\mathbf{R}_{i+n}$ is unavailable. Instead we use the diagonally loaded sample covariance matrix $\hat{\mathbf{R}}_{\mathrm{dl}}$ with $J = 1000$ and $\xi = 10\sigma_v^2$.

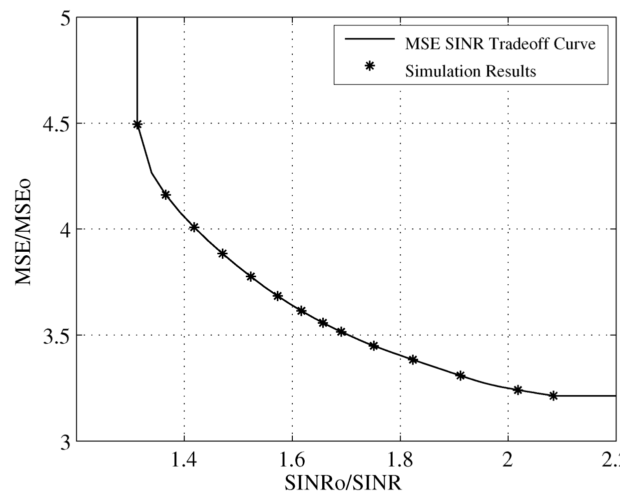


Fig. 6. MSE-SINR tradeoff for $\mathrm{INR} = 20$ dB, $\mathrm{SNR} = 0$ dB and $b = 0.7$. The diagonally loaded sample covariance matrix $\hat{\mathbf{R}}_{\mathrm{dl}}$ is used in lieu of the true covariance matrix $\mathbf{R}_{i+n}$.

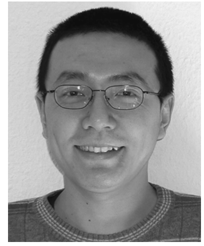
The remaining parameters are chosen as $\mathrm{INR} = 20$ dB and $\mathrm{SNR} = 0$ dB in a single antenna element, and $b = 0.7$. The MSE-SINR tradeoff curve and the beamformer simulation points are shown in Fig. 6. Despite the fact that the sample covariance matrix has been used instead of the true one, an excellent coincidence of the theoretical curves and simulated points can be seen from this figure.

V. Conclusions

Three main criteria for adaptive beamformer design have been discussed assuming a random steering vector. It has been shown that there is a fundamental tradeoff between the MSE and SINR criteria. To characterize this tradeoff, the achievable region in the MSE-SINR plane and a frontier of operating points on the boundary of this region have been studied. A new class of convex optimization theory based adaptive beamformers has been proposed that can attain the aforementioned frontier of points and, therefore, provide an optimal performance tradeoff between SINR and MSE. It has also been shown that the MLSE and MSINR criteria are simultaneously achievable and a new adaptive beamformer has been developed that is optimal under both these criteria.

REFERENCES

[1] H. Van Trees, *Optimum Array Processing*. New York: Wiley, 2002.
[2] R. Mozingo and T. Miller, *Introduction to Adaptive Arrays*. New York: Wiley, 1980.
[3] B. D. Van Veen and K. M. Buckley, "Beamforming: A versatile approach to spatial filtering," *IEEE Acoust., Speech, Signal Process. Mag.*, vol. 5, no. 2, pp. 4–24, Apr. 1988.
[4] H. Krim and M. Viberg, "Two decades of array signal processing research," *IEEE Signal Process. Mag.*, vol. 13, no. 4, pp. 67–94, Jul. 1996.
[5] Y. C. Eldar and A. Nehorai, "Mean-squared error beamforming for signal estimation: A competitive approach," in *Robust Adaptive Beamforming*, J. Li and P. Stoica, Eds. New York: Wiley, 2005, pp. 259–298.
[6] Y. C. Eldar, A. Nehorai, and P. S. L. Rosa, "An expected least-squares beamforming approach to signal estimation with steering vector uncertainties," *IEEE Signal Process. Lett.*, vol. 13, no. 5, pp. 288–291, May 2006.
[7] A. B. Gershman, "Robust adaptive beamforming," *Int. J. Electron. Commun.*, vol. 53, pp. 305–314, Dec. 1999.
[8] 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, no. 2, pp. 313–324, Feb. 2003.
[9] 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, no. 9, pp. 2257–2269, Sep. 2003.
[10] J. Li, P. Stoica, and Z. Wang, "On robust Capon beamforming and diagonal loading," *IEEE Trans. Signal Process.*, vol. 51, no. 7, pp. 1702–1715, Jul. 2003.
[11] R. G. Lorenz and S. P. Boyd, "Robust minimum variance beamforming," *IEEE Trans. Signal Process.*, vol. 53, no. 5, pp. 1684–1696, May 2005.
[12] O. Besson and F. Vincent, "Performance analysis of beamformers using generalized loading of the covariance matrix in the presence of random steering vector errors," *IEEE Trans. Signal Process.*, vol. 53, no. 2, pp. 452–459, Feb. 2005.
[13] O. Besson, A. A. Monakov, and C. Chalus, "Signal waveform estimation in the presence of uncertainties about the steering vector," *IEEE Trans. Signal Process.*, vol. 52, no. 9, pp. 2432–2440, Sep. 2004.
[14] K. L. Bell, Y. Ephraim, and H. L. Van Trees, "A Bayesian approach to robust adaptive beamforming," *IEEE Trans. Signal Process.*, vol. 48, no. 2, pp. 386–398, Feb. 2000.
[15] D. G. Luenberger, *Microeconomic Theory*. New York: McGraw-Hill, 1995.
[16] M. Ehrgott, *Multicriteria Optimization*, 2nd ed. Berlin, Germany: Springer-Verlag, 2004.
[17] Y. I. Abramovich, "Controlled method for adaptive optimization of filters using the criterion of maximum SNR," *Radio Eng. and Electron. Physics*, vol. 26, pp. 87–95, Mar. 1981.
[18] B. D. Carlson, "Covariance matrix estimation errors and diagonal loading in adaptive arrays," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 24, no. 4, pp. 397–401, Jul. 1988.
[19] S. P. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge University Press, 2004.
[20] A. Beck and Y. C. Eldar, "Strong duality in noncovex quadratic optimization with two quadratic constraints," *SIAM J. Optimiz.*, vol. 17, no. 3, pp. 844–860, 2006.
[21] A. Beck and Y. C. Eldar, "Doubly constrained robust Capon beamformer with ellipsoidal uncertainty sets," *IEEE Trans. Signal Process.*, vol. 55, no. 2, pp. 753–758, Feb. 2007.
[22] J. F. Sturm, "Using SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones," *Optimiz. Methods and Software*, vol. 11–12, pp. 625–653, 1999.



Yue Rong (S'03–M'06) was born in Jiangsu, China, in 1976. He received the B.E. degree from Shanghai Jiao Tong University, China, the M.Sc. degree from the University of Duisburg-Essen, Duisburg, Germany, and the Ph.D. degree (*summa cum laude*) from Darmstadt University of Technology, Darmstadt, Germany, all in electrical engineering, in 1999, 2002, and 2005, respectively.

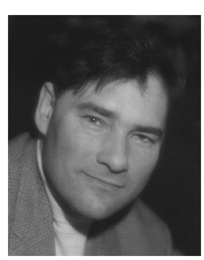
From April 2001 to October 2001, he was a Research Assistant at the Fraunhofer Institute of Microelectronic Circuits and Systems, Duisburg, Germany. From October 2001 to March 2002, he was with Nokia Ltd., Bochum, Germany. From November 2002 to March 2005, he was a Research Associate at the Department of Communication Systems, University of Duisburg-Essen. From April 2005 to January 2006, he was with the Institute of Telecommunications, Darmstadt University of Technology, as a Research Associate. Since February 2006, he has been a Postdoctoral Researcher with the Department of Electrical Engineering, University of California, Riverside. His research interests include signal processing for communications, wireless communications, wireless networks, applications of linear algebra and optimization methods, and statistical and array signal processing.

Dr. Rong received the 2001-2002 Graduate Sponsoring Asia scholarship of DAAD/ABB (Germany) and the 2004 Chinese Government Award for Outstanding Self-Financed Students Abroad (China).

Yonina C. Eldar (SM'07) received the B.Sc. degree in physics in 1995 and the B.Sc. degree in electrical engineering in 1996 both from Tel-Aviv University (TAU), Tel-Aviv, Israel, and the Ph.D. degree in electrical engineering and computer science in 2001 from the Massachusetts Institute of Technology (MIT), Cambridge.

From January 2002 to July 2002, she was a Postdoctoral Fellow at the Digital Signal Processing Group at MIT. She is currently an Associate Professor in the Department of Electrical Engineering at Technion—Israel Institute of Technology, Haifa. She is also a Research Affiliate with the Research Laboratory of Electronics at MIT.

Dr. Eldar was in the program for outstanding students at TAU from 1992 to 1996. In 1998, she held the Rosenblith Fellowship for study in Electrical Engineering at MIT, and in 2000, she held an IBM Research Fellowship. From 2002–2005 she was a Horev Fellow of the Leaders in Science and Technology program at Technion and an Alon Fellow. In 2004, she was awarded the Wolf Foundation Krill Prize for Excellence in Scientific Research, in 2005, the Andre and Bella Meyer Lectureship, and in 2007 the Henry Taub Prize for Excellence in Research. She is a member of the IEEE Signal Processing Theory and Methods technical committee, an Associate Editor for the IEEE TRANSACTIONS ON SIGNAL PROCESSING and the EURASIP *Journal of Signal Processing*, and on the Editorial Board of *Foundations and Trends in Signal Processing*.



Alex B. Gershman (M'97–SM'98–F'06) received his Diploma and Ph.D. degrees in Radiophysics from the Nizhny Novgorod State University, Russia, in 1984 and 1990, respectively.

From 1984 to 1999, he held several full-time and visiting research appointments in Russia, Switzerland, and Germany. In 1999, he joined the Department of Electrical and Computer Engineering, McMaster University, Hamilton, Ontario, Canada, where he became a Professor in 2002. From April 2005, he has been with the Darmstadt University of Technology, Darmstadt, Germany, as a Professor of Communication Systems. His research interests are in the area of signal processing and communications with the primary emphasis on array processing; beamforming; multi-antenna, multiuser and multi-carrier communications; and estimation and detection theory.

Dr. Gershman is the recipient of several awards, including the 2004 IEEE Signal Processing Society (SPS) Best Paper Award, the 2004 IEEE Aerospace and Electronic Systems Society Barry Carlton Award, and the 2001 Wolfgang Paul Award from the Alexander von Humboldt Foundation, Germany. He is currently Editor-in-Chief for the IEEE SIGNAL PROCESSING LETTERS and Chair of the Sensor Array and Multichannel (SAM) Technical Committee of the IEEE SPS. He was Associate Editor for the IEEE TRANSACTIONS ON SIGNAL PROCESSING (1999-2005). He is serving as General Co-Chair of the 5th IEEE SAM Workshop, to be held in Darmstadt, Germany, in July 2008.