Competitive Mean-Squared Error Beamforming

Yonina C. Eldar
Department of Electrical Engineering
Technion, Israel Institute of Technology
email: yonina@ee.technion.ac.il

Arye Nehorai
Department of Electrical and Computer Engineering
The University of Illinois at Chicago
email: nehorai@ece.uic.edu

Abstract Beamforming methods are used extensively in a variety of different areas, where one of their main goals is to estimate the source signal amplitude $s(t)$ from the array observations $\mathbf{y}(t) = s(t)\mathbf{a} + \mathbf{i}(t) + \mathbf{e}(t)$, $t = 1,2,...$, where $\mathbf{a}$ is the steering vector, $\mathbf{i}(t)$ is the interference, and $\mathbf{e}(t)$ is a Gaussian noise vector [1, 2].

To estimate $s(t)$, we may use a beamformer with weights $\mathbf{w}$ so that $\hat{s}(t) = \mathbf{w}^*\mathbf{y}(t)$, where $s(t)$ is an estimate of $\hat{s}(t)$. To ensure that $\hat{s}(t)$ is close to $s(t)$ in some sense, we may design the beamformer weights to minimize the MSE. However, since the MSE of a linear beamformer depends in general on the unknown signal power σ_s^2, we cannot directly minimize the MSE. A possible approach is to minimize the MSE subject to a constraint on the beamformer that ensures that the MSE does not depend on σ_s^2, which, as we show below, results in the conventional Capon beamformer. However, this approach does not guarantee a small MSE, so that on average, the resulting estimate of $s(t)$ may be far from $s(t)$. Instead, it would be desirable to design a robust beamformer whose performance is reasonably good access all possible signal powers.

In our work, we propose a minimax regret beamformer whose MSE is uniformly as close as possible to that of the optimal beamformer that knows σ_s^2, for all possible values of σ_s^2. Thus, we ensure that over a wide range power signal of values, our beamformer will result in a relatively low MSE.

Specifically, the MSE corresponding to a beamformer $\mathbf{w}$ is

$$E(|\hat{s}(t) - s(t)|^2) = \mathbf{w}^*\mathbf{R}\mathbf{w} + \sigma_s^2|1 - \mathbf{w}^*\mathbf{a}|^2, \tag{1}$$

where $\mathbf{R}$ is the interference+noise covariance matrix, which is typically replaced by the sample covariance. Since σ_s^2 is not known, we cannot choose an estimator to minimize the MSE directly. One approach is to force the term depending on σ_s^2 to 0, and then minimize the MSE, *i.e.*,

$$\min_{\mathbf{w}} \mathbf{w}^*\mathbf{R}\mathbf{w} \text{ subject to } \mathbf{w}^*\mathbf{a} = 1, \tag{2}$$

which leads to the Capon beamformer. However, this does not guarantee a small MSE.

Based on the ideas of [3], we propose a competitive beamforming approach, in which the beamformer $\mathbf{w}$ is designed to minimize the worst-case *regret* over all values $\sigma_s^2 \le U^2$, where the regret is defined as the difference between the MSE using a beamformer ignorant of the signal power, and the smallest possible MSE attainable with a beamformer that knows the power. This ensures that the performance of the resulting minimax regret beamformer minimizing the MSE when the signal power is known. The upper bound U is estimated as $U(t) = |\mathbf{w}^*\mathbf{y}(t)|$ using one of the conventional beamformers $\mathbf{w}$. The resulting estimate of $s(t)$ is

$$\hat{s}(t) = \left(1 - \frac{1}{\sqrt{(1 + U^2(t)\ \mathbf{a}^*\mathbf{R}^{-1}\mathbf{a})}}\right) \frac{\mathbf{a}^*\mathbf{R}^{-1}\mathbf{y}(t)}{\mathbf{a}^*\mathbf{R}^{-1}\mathbf{a}}, \quad t = 1, 2,... \tag{3}$$

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
20 DEC 2004	**N/A**	**-**

4. TITLE AND SUBTITLE	5a. CONTRACT NUMBER
Competitive Mean-Squared Error Beamforming	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)	8. PERFORMING ORGANIZATION REPORT NUMBER
Dept. of Electrical Engineering TechnionIsrael Institute of Technology Haifa, Israel 32000; Dept. of Electrical and Computer Engineering The University of Illinois at Chicago Chicago, IL 60607, USA	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S)
	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release, distribution unlimited

13. SUPPLEMENTARY NOTES
See also, ADM001741 Proceedings of the Twelfth Annual Adaptive Sensor Array Processing Workshop, 16-18 March 2004 (ASAP-12, Volume 1)., The original document contains color images.

14. ABSTRACT

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			
unclassified	**unclassified**	**unclassified**	**UU**	**9**	

COMPETITIVE MEAN-SQUARED ERROR BEAMFORMING

Yonina C. Eldar

Dept. of Electrical Engineering
Technion—Israel Institute of Technology
Haifa, Israel 32000
yonina@ee.technion.ac.il

Arye Nehorai

Dept. of Electrical and Computer Engineering
The University of Illinois at Chicago
Chicago, IL 60607, USA
nehorai@ece.uic.edu

ABSTRACT

We consider the problem of designing a linear beamformer to estimate a source signal $s(t)$ from array observations. Conventional beamforming methods typically aim at maximizing the signal-to-interference-plus-noise ratio (SINR). However this does not guarantee a small mean-squared error (MSE), hence on average their resulting signal estimate $\hat{s}(t)$ can be far from $s(t)$. To ensure that $\hat{s}(t)$ is close to $s(t)$, we propose using the more appropriate design criterion of MSE. Since the MSE depends in general on $s(t)$ which is unknown, it cannot be minimized directly. Therefore we develop a competitive beamforming approach, in which the beamformer is designed to minimize the worst-case* regret *over all $s(t)$, where the regret is the difference between the MSE using a beamformer ignorant of $s(t)$ and the smallest possible MSE attainable with a beamformer that knows $s(t)$. Thus, we ensure that over a wide range of signal values, our beamformer will result in a relatively low MSE. We demonstrate through numerical examples that the proposed minimax regret beamformer (MMR) outperforms several existing standard and robust beamformers, for wide range of SNR values.

1. INTRODUCTION

Beamforming is a classical method of processing temporal sensor array measurements for signal estimation, interference cancellation, or source direction and spectrum estimation [1, 2, 3]. It has ubiquitously been applied in areas such as radar, sonar, wireless communications, speech processing, medical imaging, radioastronomy, etc.

Conventional approaches for designing beamformers typically attempt to maximize the signal-to-interference-plus-noise ratio (SINR). Maximizing the SINR requires knowledge of the interference-plus-noise covariance matrix and the array steering vector. Since this covariance is unknown, it is often replaced by the sample covariance of the measurements, resulting in deterioration of performance with higher signal-to-noise ratio (SNR) when the signal is present in the training data. Some beamforming techniques are designed to mitigate this effect [4, 5, 6, 7, 8], whereas others are developed to also overcome uncertainty in the steering vector [9, 10, 11]. However, maximizing SINR may not guarantee a good estimate of the signal. In an *estimation* context, where our goal is to deign a beamformer in order to obtain an estimate of the signal amplitude that is close to the true amplitude, it would make more sense to choose the weights to minimize an objective that is related to the estimation error, rather than the SINR.

In this paper we derive a beamforming method for estimating a signal in the presence of interference and noise using the mean-squared error (MSE) as the performance criterion. Computing the MSE shows, however, that it depends explicitly on the unknown signal in the deterministic case, or the unknown signal power in the stochastic case, hence cannot be minimized directly. Thus, we aim at designing a robust beamformer whose performance in terms of MSE is good across all possible values of the unknowns. To develop a beamformer with this property, we employ a new competitive estimation framework, which has been recently proposed for solving robust estimation problems [12]. This framework considers a general linear estimation problem, and suggests a linear estimator whose performance is as close as possible to that of the optimal linear estimator for the case in which the model parameters are completely known. Specifically, the estimator is designed to minimize the worst-case *regret*, which is the difference between the MSE of the estimator in the presence of uncertainties, and the smallest attainable MSE with a linear estimator that knows the exact model. Based on this framework, we propose a minimax regret (MMR) beamformer whose MSE is uniformly as close as possible to that of the optimal beamformer that knows the signal parameters, for all possible parameter values, as we detail below in the next section. Thus, we ensure that our proposed beamformer will result in a relatively low MSE over a wide range of signal values.

In Section 2 we present the problem formulation and review existing methods. In Section 3 we develop the proposed MMR beamformer. In Sections 4 and 5 we discuss practical considerations and present numerical examples illustrating the advantage of the proposed MMR beamformer over several existing standard and robust beamformers, for wide range of SNR values.

2. PROBLEM FORMULATION

We denote vectors in $\mathbb{C}^m$ by boldface lowercase letters and matrices in $\mathbb{C}^{n\times m}$ by boldface uppercase letters. $\mathbf{I}$ denotes the identity matrix of appropriate dimension, $(\cdot)^*$ denotes the Hermitian conjugate, and $\hat{(\cdot)}$ denotes an estimated vector or matrix.

Beamforming methods are used extensively in a variety of different areas, where one of their main goals is to estimate the source signal amplitude $s(t)$ from the array observations

$$\mathbf{y}(t) = s(t)\mathbf{a} + \mathbf{i}(t) + \mathbf{e}(t), \quad t = 1, \ldots, N \tag{1}$$

The work of A. Nehorai was supported by the Air Force Office of Scientific Research Grant F49620-02-1-0339, the National Science Foundation Grants CCR-0105334 and CCR-0330342, and the Office of Naval Research Grant N00014-01-1-0681.

where $\mathbf{y}(t) \in \mathbb{C}^M$ is the complex vector of array observations at time t with M being the number of array sensors, $\mathbf{a}$ is the signal steering vector which depends on the degree of arrival (DOA) of the wavefront plane associated with $s(t)$ with respect to a uniform linear array (ULA) of sensors, $s(t)$ is the signal amplitude, $\mathbf{i}(t)$ is the interference, $\mathbf{e}(t)$ is a Gaussian noise vector and N is the number of snapshots [1, 2].

The source signal amplitude $s(t)$ may be a deterministic unknown signal, such as a complex sinusoid, or a stochastic stationary process with unknown signal power. Since the statistics of $s(t)$ are unknown, in our development below we will treat $s(t)$ as a deterministic signal. Note, that as we demonstrate in Section 5, the algorithms we develop can be applied to both deterministic and stochastic signals.

Our goal is to estimate the signal amplitude $s(t)$ from the observations $\mathbf{y}(t)$ using a set of beamformer weights $\mathbf{w}(t)$, where the output of a narrowband beamformer is given by

$$\hat{s}(t) = \mathbf{w}^*(t)\mathbf{y}(t), \quad t = 1, \ldots, N. \tag{2}$$

Traditionally, the beamformer weights $\mathbf{w}(t) = \mathbf{w}$ are chosen to maximize the SINR

$$\text{SINR} \propto \frac{|\mathbf{w}^*\mathbf{a}|^2}{\mathbf{w}^*\mathbf{R}_{i+e}\mathbf{w}}, \tag{3}$$

where

$$\mathbf{R}_{i+e} = E\left\{(\mathbf{i}+\mathbf{n})(\mathbf{i}+\mathbf{n})^*\right\} \tag{4}$$

is the interference+noise covariance matrix. The weight vector maximizing the SINR is given by

$$\mathbf{w} = \frac{\mathbf{R}_{i+e}^{-1}\mathbf{a}}{\mathbf{a}^*\mathbf{R}_{i+e}\mathbf{a}}. \tag{5}$$

The solution (5) is also referred to as the minimum variance distortionless response (MVDR) beamformer, since it can be obtained as the solution to

$$\min_{\mathbf{w}} \mathbf{w}^*\mathbf{R}_{i+e}\mathbf{w} \quad \text{subject to} \quad \mathbf{w}^*\mathbf{a} = 1. \tag{6}$$

In practice, the interference+noise covariance matrix $\mathbf{R}_{i+e}$ is often not available. In such cases, the exact covariance $\mathbf{R}_{i+e}$ in (5) is replaced by an estimated covariance. Various methods exist for estimating the covariance $\mathbf{R}_{i+e}$. The simplest approach is to choose the estimate as the sample covariance

$$\widehat{\mathbf{R}} = \frac{1}{N}\sum_{t=1}^{N}\mathbf{y}(t)\mathbf{y}(t)^*. \tag{7}$$

The resulting beamformer is commonly referred to as the sample matrix inversion (SMI) beamformer or the Capon beamformer [13, 14]. If the signal is present in the training data, then it is well known that the performance of the MVDR beamformer with $\mathbf{R}_{i+n}$ replaced by $\widehat{\mathbf{R}}$ of (7) degrades considerably [6].

An alternative approach for estimating $\mathbf{R}_{i+n}$ is the diagonal loading approach, in which the estimate is chosen as

$$\widehat{\mathbf{R}}_{dl} = \widehat{\mathbf{R}} + \xi\mathbf{I} = \frac{1}{N}\sum_{t=1}^{N}\mathbf{y}(t)\mathbf{y}(t)^* + \xi\mathbf{I}, \tag{8}$$

where ξ is the diagonal loading factor. The resulting beamformer is referred to as the loaded SMI (LSMI) beamformer [4, 5], or the loaded Capon beamformer. Various methods have been proposed for choosing the diagonal loading factor ξ; see *e.g.*, [5]. A heuristic choice for ξ, which is common in applications, is $\xi \approx 10\sigma^2$, where σ^2 is the noise power in a single sensor.

Another popular approach to estimating $\mathbf{R}_{i+e}$ is the eigenspace approach [6, 7, 8], in which the covariance matrix is estimated as

$$\widehat{\mathbf{R}}_{eig} = \widehat{\mathbf{R}}P_s, \tag{9}$$

where P_s is the orthogonal projection onto the subspace corresponding to the $D+1$ largest eigenvalues of $\widehat{\mathbf{R}}$, were D is the known rank of the interference subspace.

The Capon, loaded Capon and eigenspace beamformers, can all be viewed as MVDR beamformers with a particular estimate of $\mathbf{R}_{i+e}$. In the sequel, we use $\mathbf{R}$ to denote the interference+noise covariance matrix, where $\mathbf{R}$ is equal to $\mathbf{R}_{i+e}$ when $\mathbf{R}_{i+e}$ is available, and is otherwise chosen as any of the estimates above.

The class of MVDR beamformers assumes explicitly that the steering vector $\mathbf{a}$ is known exactly. Recently, several robust beamformers have been proposed for the case in which the steering vector is not known precisely, but rather lies in some uncertainty set [9, 10, 11]. Although originally developed to deal with steering vector mismatch, the authors of the referenced papers suggest using these robust methods even in the case in which $\mathbf{a}$ is known, in order to deal with the mismatch in the interference+noise covariance. Each of the above robust methods is designed to maximize a measure of SINR on the uncertainty set. Specifically, in [9], the authors suggest minimizing $\mathbf{w}^*\widehat{\mathbf{R}}\mathbf{w}$ subject to the constraint that $|\mathbf{w}^*\mathbf{c}| \geq 1$ for all possible values of the steering vector $\mathbf{c}$, where $\|\mathbf{c}-\mathbf{a}\| \leq \epsilon$. The resulting beamformer is given by

$$\mathbf{w} = \frac{\lambda}{\lambda\mathbf{a}^*\left(\widehat{\mathbf{R}}+\lambda\epsilon^2\mathbf{I}\right)^{-1}\mathbf{a}-1}\left(\widehat{\mathbf{R}}+\lambda\epsilon^2\mathbf{I}\right)^{-1}\mathbf{a}, \tag{10}$$

where λ is chosen such that $|\mathbf{w}^*\mathbf{a}-1|^2 = \epsilon^2\mathbf{w}^*\mathbf{w}$. In practice, the solution can be found by using a second order cone program. In [10] the authors consider a similar approach in which they seek to minimize $\mathbf{w}^*\widehat{\mathbf{R}}\mathbf{w}$ subject to $\|\mathbf{w}-\mathbf{a}\| \leq \epsilon$, which results in the beamformer

$$\mathbf{w} = \frac{\left(\widehat{\mathbf{R}}+\frac{1}{\lambda}\mathbf{I}\right)^{-1}\mathbf{a}}{\mathbf{a}^*\left(\widehat{\mathbf{R}}+\frac{1}{\lambda}\mathbf{I}\right)^{-1}\widehat{\mathbf{R}}\left(\widehat{\mathbf{R}}+\frac{1}{\lambda}\mathbf{I}\right)\mathbf{a}}, \tag{11}$$

where λ is chosen such that $\left\|\left(\mathbf{I}+\lambda\widehat{\mathbf{R}}\right)^{-1}\mathbf{a}\right\|^2 = \epsilon$. Finally, in [11] the authors consider a general-rank signal model. Adapting their results to the rank-one steering vector case, their beamformer is the solution to minimizing $\mathbf{w}^*\widehat{\mathbf{R}}\mathbf{w}$ subject to $|\mathbf{w}^*\mathbf{a}|^2 \geq 1 - \mathbf{w}^*\Delta\mathbf{w}$ for all $\|\Delta\| \leq \epsilon$, and is given by

$$\mathbf{w} = P\left(\left(\widehat{\mathbf{R}}+\lambda\mathbf{I}\right)^{-1}\left(\mathbf{a}\mathbf{a}^*-\epsilon\mathbf{I}\right)\right), \tag{12}$$

where $P(\mathbf{A})$ is the eigenvector associated with the largest eigenvalue of $\mathbf{A}$, and λ is a diagonal loading factor.

The motivation behind the class of MVDR beamformers and the robust beamformers is to maximize the SINR. However, choosing $\mathbf{w}$ to maximize the SINR does not necessarily result in an estimated signal amplitude $\hat{s}(t)$ that is close to $s(t)$. In an *estimation* context, where our goal is to deign a beamformer in order to obtain

an estimate $\hat{s}(t)$ that is close to $s(t)$, it would make more sense to choose the weights $\mathbf{w}$ to minimize the MSE rather than to maximize the SINR, which is not directly related to the estimation error $\hat{s}(t) - s(t)$.

If $\hat{s} = \mathbf{w}^*\mathbf{y}$, where for brevity we omitted the index t, then the MSE between s and $\hat{s}$ is given by

$$\begin{aligned} E(|\hat{s} - s|^2) &= V(\hat{s}) + |B(\hat{s})|^2 \\ &= \mathbf{w}^*\mathbf{R}\mathbf{w} + |s|^2|1 - \mathbf{w}^*\mathbf{a}|^2, \end{aligned} \quad (13)$$

where $V(\hat{s}) = E(|\hat{s} - E(\hat{s})|^2)$ is the variance of the estimate $\hat{s}$ and $B(\hat{s}) = E(\hat{s}) - s$ is the bias. Since s is not known, we cannot choose a beamformer to minimize the MSE of (13). One approach is to force the term depending on $|s|$, namely the bias, to 0, and then minimize the MSE, *i.e.,*

$$\min_{\mathbf{w}} \mathbf{w}^*\mathbf{R}\mathbf{w} \text{ subject to } \mathbf{w}^*\mathbf{a} = 1, \quad (14)$$

which leads to the class of MVDR beamformers. Thus, in addition to maximizing the SINR, the MVDR beamformer minimizes the MSE subject to the constraint that the bias in the estimator $\hat{s}$ is equal to 0. However, this does not guarantee a small MSE, so that on average, the resulting estimate of $s(t)$ may be far from $s(t)$. Indeed, it is well known that unbiased estimators may often lead to large MSE values.

Instead of forcing the term depending on s to zero, it would be desirable to design a robust beamformer whose performance is reasonably good across all possible values of s. Based on the ideas of [12], we propose an MMR beamformer whose MSE is uniformly as close as possible to that of the optimal beamformer that knows s, for all possible values of s in a prespecified region of uncertainty. Thus, we ensure that over a wide range of values of s, our beamformer will result in a relatively low MSE. Specifically, we seek a beamformer that minimizes the worst-case difference regret, namely the worst-case difference between its MSE and the best possible MSE attainable using a linear beamformer when s is known, over a bounded set of values $|s|$. When s is known, the beamformer output has the form $\hat{s} = \mathbf{w}^*(s)\mathbf{y}$, where $\mathbf{w}$ can depend on s. As we show below, even in the case in which s is known, we cannot achieve a zero MSE, when restricting ourselves to linear beamformers.

In [12], a minimax difference regret estimator was derived for the problem of estimating an unknown vector $\mathbf{x}$ in a linear model $\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n}$, where $\mathbf{H}$ is a known linear transformation, and $\mathbf{n}$ is a noise vector with known covariance matrix. The estimator was designed to minimize the worst-case regret over all bounded vectors $\mathbf{x}$, namely vectors satisfying $\mathbf{x}^*\mathbf{T}\mathbf{x} \leq U^2$ for some $U > 0$ and some positive definite matrix $\mathbf{T}$. It was shown that the linear MMR estimator can be found as a solution to a convex optimization problem that can be solved very efficiently.

In our problem, the unknown parameter $\mathbf{x} = s$ is a scalar, so that an explicit solution can be derived, as we show in Section 3. Furthermore, in our development we consider both lower and upper bounds on $|s|$, so that we seek the beamformer that minimizes the worst-case regret over the uncertainty region $L \leq |s| \leq U$. The bounds L and U can either be determined based on prior information on the signal amplitude, or, in cases in which no such information is available, these bounds can be estimated from the data, as we discuss in Section 4. Thus, in practice, the only prior information needed in order to implement the MMR beamformer is knowledge of the steering vector.

3. THE MINIMAX REGRET BEAMFORMER

The MMR beamformer is designed to minimize the worst-case regret $\mathcal{R}(s, \mathbf{w})$, which is defined as the difference between the MSE using an estimator $\hat{s} = \mathbf{w}^*\mathbf{y}$ and the smallest possible MSE attainable with an estimator of the form $\hat{s} = \mathbf{w}^*(s)\mathbf{y}$ when s is known, which we denote by MSE^o.

To develop an explicit expression for MSE^o we first determine the estimator $\hat{s} = \mathbf{w}^*(s)\mathbf{y}$ that minimizes the MSE when s is known. To this end we differentiate MSE of (13) with respect to $\mathbf{w}$ and equate to 0, which results in

$$\mathbf{R}\mathbf{w}(s) + |s|^2(\mathbf{a}^*\mathbf{w}(s) - 1)\mathbf{a} = 0, \quad (15)$$

so that

$$\mathbf{w}(s) = |s|^2(\mathbf{R} + |s|^2\mathbf{a}\mathbf{a}^*)^{-1}\mathbf{a}. \quad (16)$$

Using the Matrix Inversion Lemma we can express $\mathbf{w}(s)$ as

$$\mathbf{w}(s) = \frac{|s|^2}{1 + |\mathbf{s}|^2\mathbf{a}^*\mathbf{R}\mathbf{a}}\mathbf{R}^{-1}\mathbf{a}. \quad (17)$$

Substituting $\mathbf{w}(s)$ back into (13), MSE^0 is given by

$$\mathrm{MSE}^o = \frac{|s|^2}{1 + |s|^2\mathbf{a}^*\mathbf{R}^{-1}\mathbf{a}} = \frac{|s|^2}{1 + |s|^2\alpha}, \quad (18)$$

where for brevity we denote $\alpha = \mathbf{a}^*\mathbf{R}^{-1}\mathbf{a}$.

Since s is unknown, we cannot implement the optimal beamformer (17). Instead we seek the beamformer $\hat{s} = \mathbf{w}^*\mathbf{y}$ that minimizes the worst-case regret $\mathcal{R}(s, \mathbf{w}) = E(|\mathbf{w}^*\mathbf{y} - s|^2) - \mathrm{MSE}^o$ subject to the constraint $L \leq |s| \leq U$. Thus we seek the beamformer $\mathbf{w}$ that is the solution to the problem

$$\begin{aligned} \min_{\mathbf{w}} \max_{L \leq |s| \leq U} \mathcal{R}(s, \mathbf{w}) = \min_{\mathbf{w}} \Bigg\{ & \mathbf{w}^*\mathbf{R}\mathbf{w} + \\ & + \max_{L \leq |s| \leq U} \left\{ |s|^2|1 - \mathbf{w}^*\mathbf{a}|^2 - \frac{|s|^2}{1 + |s|^2\alpha} \right\} \Bigg\}. \end{aligned} \quad (19)$$

To develop a solution to (19), we first consider the inner maximization problem

$$\begin{aligned} f(\mathbf{w}) &= \max_{L \leq |s| \leq U} \left\{ |s|^2|1 - \mathbf{w}^*\mathbf{a}|^2 - \frac{|s|^2}{1 + |s|^2\alpha} \right\} \\ &= \max_{L^2 \leq x \leq U^2} \left\{ x|1 - \mathbf{w}^*\mathbf{a}|^2 - \frac{x}{1 + x\alpha} \right\}, \end{aligned} \quad (20)$$

where $x = |s|^2$. Noting that the function $h(x) = ax - bx/(c + dx)$ with $b, c, d > 0$ is convex in $x \geq 0$, we have that for fixed $\mathbf{w}$,

$$g(x) = x|1 - \mathbf{w}^*\mathbf{a}|^2 - \frac{x}{1 + x\alpha} \quad (21)$$

is convex in $x \geq 0$, and consequently the maximum of $g(x)$ over a closed interval is obtained at one of the boundaries. Thus,

$$f(\mathbf{w}) = \max_{L^2 \leq x \leq U^2} g(x) = \max\left(g(L^2), g(U^2)\right), \quad (22)$$

and the problem (19) reduces to

$$\begin{aligned} \min_{\mathbf{w}} \Bigg\{ \mathbf{w}^*\mathbf{R}\mathbf{w} + \max \Bigg(& L^2|1 - \mathbf{w}^*\mathbf{a}|^2 - \frac{L^2}{1 + L^2\alpha}, \\ & U^2|1 - \mathbf{w}^*\mathbf{a}|^2 - \frac{U^2}{1 + U^2\alpha} \Bigg) \Bigg\}. \end{aligned} \quad (23)$$

We now show that the optimal value of $\mathbf{w}$ has the form

$$\mathbf{w} = d(\mathbf{a}^*\mathbf{R}^{-1}\mathbf{a})^{-1}\mathbf{R}^{-1}\mathbf{a} = \frac{d}{\alpha}\mathbf{R}^{-1}\mathbf{a}, \tag{24}$$

for some d. To this end, we first note that the objective in (23) depends on $\mathbf{w}$ only through $\mathbf{w}^*\mathbf{a}$ and $\mathbf{w}^*\mathbf{R}\mathbf{w}$. Now, suppose that we are given a beamformer $\tilde{\mathbf{w}}$, and let

$$\mathbf{w} = \frac{\mathbf{a}^*\tilde{\mathbf{w}}}{\alpha}\mathbf{R}^{-1}\mathbf{a}. \tag{25}$$

Then

$$\mathbf{w}^*\mathbf{a} = \frac{\tilde{\mathbf{w}}^*\mathbf{a}}{\alpha}\mathbf{a}^*\mathbf{R}^{-1}\mathbf{a} = \tilde{\mathbf{w}}^*\mathbf{a}, \tag{26}$$

and

$$\mathbf{w}^*\mathbf{R}\mathbf{w} = \frac{|\mathbf{a}^*\tilde{\mathbf{w}}|^2}{\alpha^2}\mathbf{a}^*\mathbf{R}^{-1}\mathbf{a} = \frac{|\mathbf{a}^*\tilde{\mathbf{w}}|^2}{\alpha}. \tag{27}$$

¿From the Cauchy-Schwarz inequality, for any vector $\mathbf{x}$,

$$|\mathbf{a}^*\mathbf{x}|^2 \leq \mathbf{a}^*\mathbf{R}^{-1}\mathbf{a}\mathbf{x}^*\mathbf{R}\mathbf{x} = \alpha\mathbf{x}^*\mathbf{R}\mathbf{x}. \tag{28}$$

Substituting (28) with $\mathbf{x} = \tilde{\mathbf{w}}$ into (27), we have that

$$\mathbf{w}^*\mathbf{R}\mathbf{w} \leq \frac{|\mathbf{a}^*\tilde{\mathbf{w}}|^2}{\alpha} \leq \tilde{\mathbf{w}}^*\mathbf{R}\tilde{\mathbf{w}}. \tag{29}$$

It follows from (26) and (29) that $\mathbf{w}$ is at least as good as $\tilde{\mathbf{w}}$ for minimizing (23). Therefore, the optimal value of $\mathbf{w}$ satisfies

$$\mathbf{w} = \frac{\mathbf{a}^*\mathbf{w}}{\alpha}\mathbf{R}^{-1}\mathbf{a}, \tag{30}$$

which implies that $\mathbf{w}$ has the form (24) for some d.

Combining (24) and (23), our problem reduces to

$$\min_d \left\{ \frac{|d|^2}{\alpha} + \max\left(L^2|1-d|^2 - \frac{L^2}{1+\alpha L^2}, \right.\right. \\ \left.\left. U^2|1-d|^2 - \frac{U^2}{1+\alpha U^2} \right) \right\}. \tag{31}$$

Since d is in general complex, we can write $d = |d|e^{j\phi}$ for some $0 \leq \phi \leq 2\pi$. Using the fact that $|1-d|^2 = 1 + |d|^2 - 2\cos(\phi)$, it is clear that at the optimal solution, $\phi = 0$. Therefore, without loss of generality, we assume in the sequel that $d \geq 0$. We can then express the problem of (31) as

$$\min_{t,d} t \tag{32}$$

subject to

$$\begin{aligned} \frac{d^2}{\alpha} + L^2(1-d)^2 - \frac{L^2}{1+\alpha L^2} &\leq t; \\ \frac{d^2}{\alpha} + U^2(1-d)^2 - \frac{U^2}{1+\alpha U^2} &\leq t. \end{aligned} \tag{33}$$

The constraints (33) can be equivalently written as

$$\begin{aligned} f_L(d) &\triangleq \left(\frac{1}{\alpha} + L^2\right)\left(d - \frac{\alpha L^2}{1+\alpha L^2}\right)^2 \leq t; \\ f_U(d) &\triangleq \left(\frac{1}{\alpha} + U^2\right)\left(d - \frac{\alpha U^2}{1+\alpha U^2}\right)^2 \leq t. \end{aligned} \tag{34}$$

To develop a solution to (32) subject to (34), we note that both $f_L(d)$ and $f_U(d)$ are quadratic functions in d, that obtain a minimum at d_L and d_U respectively, where

$$\begin{aligned} d_L &= \frac{\alpha L^2}{1+\alpha L^2}; \\ d_U &= \frac{\alpha U^2}{1+\alpha U^2}. \end{aligned} \tag{35}$$

Therefore, the optimal value of d, denoted d_0, satisfies

$$d_L \leq d_0 \leq d_U. \tag{36}$$

Indeed, let $t(d) = \max(f_L(d), f_U(d))$, and let $t_0 = t(d_0)$ be the optimal value of (32) subject to (34). Since both $f_L(d)$ and $f_U(d)$ are monotonically decreasing for $d < d_L$, $t(d) > t(d_L) \geq t_0$ for $d < d_L$ so that $d_0 \geq d_L$. Similarly, since both $f_L(d)$ and $f_U(d)$ are monotonically increasing for $d > d_U$, $t(d) > t(d_U) \geq t_0$ for $d > d_U$ so that $d \leq d_U$.

Since $f_L(d)$ and $f_U(d)$ are both quadratic, they intersect at most at two points. If $f_L(d) = f_U(d)$, then

$$(1-d)^2 = \frac{1}{(1+\alpha L^2)(1+\alpha U^2)}, \tag{37}$$

so that $f_L(d) = f_U(d)$ for $d = d_+$ and $d = d_-$, where

$$d_\pm = 1 \pm \frac{1}{\sqrt{(1+\alpha L^2)(1+\alpha U^2)}}. \tag{38}$$

Denoting by $\mathcal{I}$ the interval $\mathcal{I} = [d_L, d_U]$, since $d_+ > 1$, clearly $d_+ \notin \mathcal{I}$. Using the fact that

$$\frac{1}{1+\alpha U^2} \leq \frac{1}{\sqrt{(1+\alpha L^2)(1+\alpha U^2)}} \leq \frac{1}{1+\alpha L^2}, \tag{39}$$

we have that $d_- \in \mathcal{I}$. We now show that the optimal value of d is $d_0 = d_-$. If $L = U$, then $d_- = d_L = d_U$ so that from (36), $d_0 = d_-$. Next, assume that $L < U$. In this case, for $d \in \mathcal{I}$, $f_L(d)$ is monotonically increasing and $f_U(d)$ is monotonically decreasing. Denoting $t_- = t(d_-)$ and noting that $t_- = f_L(d_-) = f_U(d_-)$, we conclude that for $d_- < d \leq d_L$, $f_U(d) > t_-$, and for $d_U \leq d < d_-$, $f_L(d) > t_-$ so that $t(d) > t_-$ for any $d \in \mathcal{I}$ such that $d \neq d_-$, and therefore $d_0 = d_-$.

We summarize our results in the following theorem.

Theorem 1. *Let s denote an unknown parameter in the model $\mathbf{y} = s\mathbf{a} + \mathbf{n}$, where $\mathbf{a}$ is a known length-M vector, and $\mathbf{n}$ is a zero-mean random vector with covariance $\mathbf{R}$. Then the solution to*

$$\min_{\hat{s}=\mathbf{w}^*\mathbf{y}} \max_{L\leq|s|\leq U} \left\{ E(|\hat{s}-s|^2) - \min_{\hat{s}=\mathbf{w}^*(s)\mathbf{y}} E(|\hat{s}-s|^2) \right\}$$

is

$$\hat{s} = \left(1 - \frac{1}{\sqrt{(1+L^2\mathbf{a}^*\mathbf{R}^{-1}\mathbf{a})(1+U^2\mathbf{a}^*\mathbf{R}^{-1}\mathbf{a})}}\right)\frac{\mathbf{a}^*\mathbf{R}^{-1}\mathbf{y}}{\mathbf{a}^*\mathbf{R}^{-1}\mathbf{a}}.$$

4. PRACTICAL CONSIDERATIONS

In our development of the MMR beamformer, we assumed that there exists bounds L and U on the magnitude of the signal to be estimated. In some applications, such bounds may exist, for example when the type of the source and the possible distances from the

array are known. If no such bounds are available, then we may estimate them from the data using one of the conventional beamformers. Specifically, let $\mathbf{w}_c$ denote one of the conventional beamformers. Then, using this beamformer we can estimate s as $\hat{s} = \mathbf{w}_c^* \mathbf{y}$. We may then use this estimate to obtain approximate values for L and U. In the examples below, we choose $U = \beta|\mathbf{w}_c^* \mathbf{y}|$ for some $\beta > 0$, and $L = 0$. Assuming that $\mathbf{a}$ is known, with this choice of bounds the MMR beamformer becomes

$$\mathbf{w} = \left(1 - \frac{1}{\sqrt{(1+\beta^2|\mathbf{w}_c^*\mathbf{y}|^2 \mathbf{a}^* \mathbf{R}^{-1}\mathbf{a})}}\right) \frac{\mathbf{R}^{-1}\mathbf{a}}{\mathbf{a}^* \mathbf{R}^{-1}\mathbf{a}}. \quad (40)$$

Since in most applications the true covaraince is not available we have to estimate it, e.g. using (7). However, as we discussed in Section 2, if $s(t)$ is present in the training data, then a diagonal loading (8) or a projection approach (9) may perform better than (7). Thus in the examples below, the true covariance is replaced by (8) or (9).

5. NUMERICAL EXAMPLES

To evaluate and compare the performance of the proposed MMR method with other techniques, we conducted numerical examples using the same scenarios as in [11]. These consist of a uniform linear array of 20 omnidirectional sensors spaced half a wavelength apart. In all the examples below the interference and noise are zero-mean complex white Gaussian processes. The signal $s(t)$ is continuously present throughout the training data and the steering vector $\mathbf{a}$ is known. The plane-waves of $s(t)$ and interference $\mathbf{i}(t)$ have directions of arrival (DOAs) of 30^o and -30^o, respectively, relative to the array normal. The power of the noise $\mathbf{e}(t)$ is one and the interference-to-noise ratio (INR) in a single sensor is 20 dB. The merit function we use to evaluate the performance is the square root of the normalized mean-square error NMSE = $E(|\hat{s}(t) - s(t))|^2)/E(|s(t)|^2)$. Each result presented below was obtained as a sample mean from 100 Monte Carlo experiments.

The performance of the proposed method was compared against six methods: the Capon beamformer (CAPON) [13, 14], loading Capon beamformer (L-CAPON) [4, 5], eigenspace-based beamformer (EIG) [6, 7, 8] and robust beamformers (ROB1, ROB2 and ROB3) [9, 10, 11]. The parameters of each of the compared methods were chosen as suggested in the literature. For the L-CAPON (8) the diagonal loading was set as ξ = 10 σ_w^2 [9, 10] with σ_w^2 being the variance of the noise in each sensor, assumed to be known (σ_w^2 = 1 in this example); for the EIG beamformer (9) it was assumed that the low rank condition and number of interferers is known. For the alternative robust methods we have that, for ROB1 (10) the upper bound on the steering vector uncertainty was set as ϵ = 3 [9], for ROB2 (11), the upper bound on the steering vector uncertainty was set as ϵ = 3.5 [10] and the diagonal loads for ROB3 (12) [11] were chosen as λ = 30 and ϵ = 9. In order to show the best possible performance by each approach, the optimal CAPON and optimal MMR beamformer are also included. By optimal we refer to the case when the covariance matrix $\mathbf{R}$, and $|s(t)|^2$ for $t = 1, \ldots, N$ are known, in addition to the steering vector $\mathbf{a}$.

Example 1 - Deterministic signal. We chose $s(t)$ to be a complex sine wave with varying amplitude to obtain the desired SNR, and used 30 training snapshots. We implemented MMR1 with a sample covariance matrix estimated with a loading factor, ξ = 10 [9, 10], $\beta = 8$ and $\mathbf{w}_c$ given by the L-CAPON beamformer with ξ = 10; MMR2 was applied by using (9) with D = 1, $\beta = 8$ and $\mathbf{w}_c$ given by the EIG beamformer. It can be seen in Figure 1 that MMR1 has the best performance for SNR values between -6 to -4 dB, whereas MMR2 has the best performance between -3 and 4 dB. Figure 2 shows the performance as a function of the number of training data for a fixed SNR = -5 dB. It can be seen that MMR1 and MMR2 have improved performance with larger number of training snapshots.

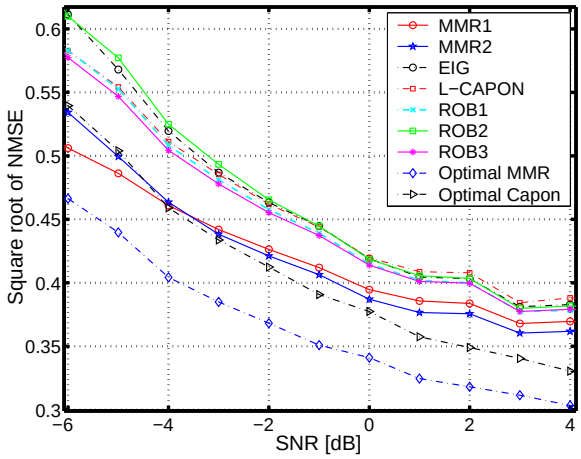

Figure 1: Performance comparison of the different methods in terms of the square root of the normalized mean-squared error for estimating a complex sinewave, as a function of signal-to-noise ratio for a training data of 30 snapshots.

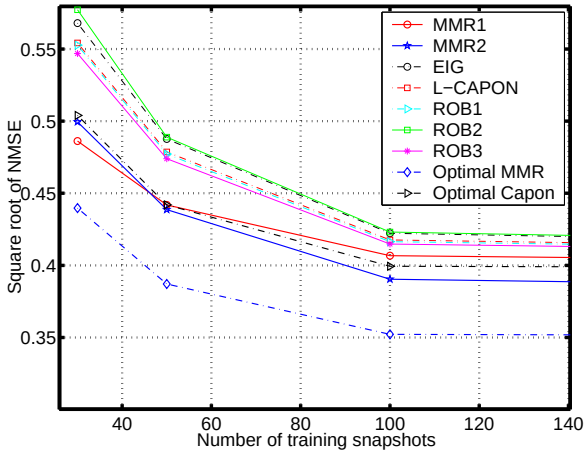

Figure 2: Performance comparison of the different methods in terms of the square root of the normalized mean-squared error for estimating a complex sinewave, as a function of number of training snapshots for SNR = -5 dB.

Example 2 - Stochastic signal. In this example we chose $s(t)$ to be a zero-mean complex Gaussian signal with varying variance to obtain the desired SNR, and used 30 training snapshots. We implemented MMR1 and MMR2 as described in the last example,

but with β = 12. From Figure 3, we can draw similar conclusions as in the last example about the performance of optimal and practical beamformers presented. It can be seen that MMR1 with ξ = 10 and β = 12, has the best performance for SNR values between -8 to -5 dB, whereas MMR2 with β = 12 has the best performance between -4 and 2 dB. Also, it can be seen in Figure 4 that MMR1 and MMR2 improve their performances with a larger number of training snapshots for SNR of -5 dB.

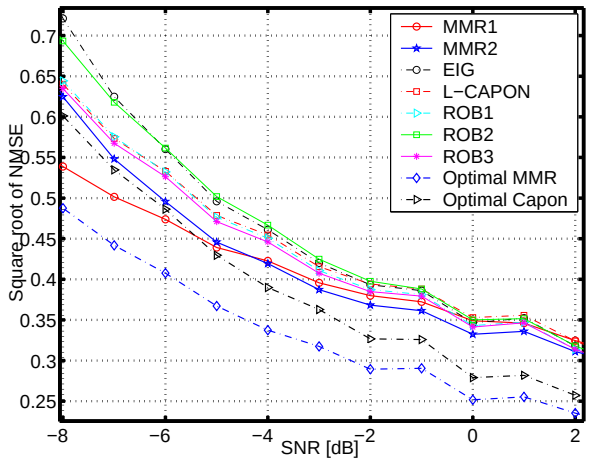

Figure 3: Performance comparison of the different methods in terms of the square root of the normalized mean-squared error for estimating a zero mean complex Gaussian signal, as a function of signal-to-noise ratio for a training data of 30 snapshots.

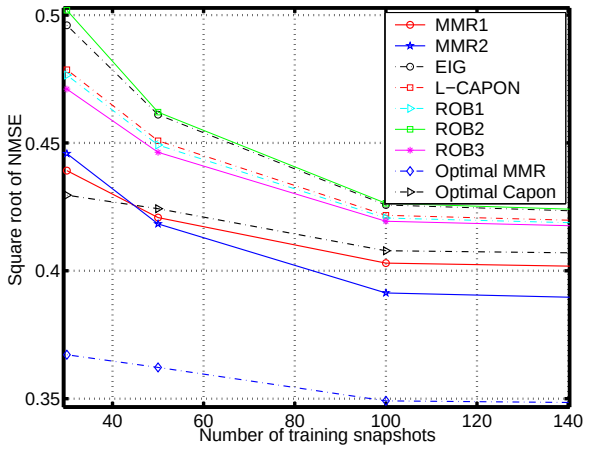

Figure 4: Performance comparison of the different methods in terms of the square root of the normalized mean-squared error for estimating a zero mean complex Gaussian signal, as a function of number of training snapshots for SNR = -5 dB.

Acknowledgment

The authors are grateful to Mr. Patricio S. La Rosa for conducting the numerical examples in this paper.

6. REFERENCES

[1] B. D. Van Veen and K. M. Buckley, "Beamforming: A versatile approach to spatial filtering," *IEEE Signal Proc. Magazine*, vol. 5, pp. 4–24, Apr. 1988.

[2] H. Krim and M. Viberg, "Two decades of array signal processing research: The parametric approach," *IEEE Signal Proc. Magazine*, vol. 13, pp. 67–94, July 1996.

[3] H. L. Van Trees, *Optimal Array Processing (Detection, Estimation, and Modulation Theory, Part IV)*, Wiley-Interscience, 2002.

[4] B. D. Carlson, "Covariance matrix estimation errors and diagonal loading in adaptive arrays," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 24, pp. 397–401, July 1988.

[5] H. Cox, R. M. Zeskind, and M. M. Owen, "Robust adaptive beamforming," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-35, pp. 1365–1376, Oct. 1987.

[6] D. D. Feldman and L. J. Griffiths, "A projection approach for robust adaptive beamforming," *IEEE Trans. Signal Processing*, vol. 42, pp. 867–876, Apr. 1994.

[7] N. L. Owsley, "Signal subspace based minimum variance spatial array processing," *Proceedings of the 19th Asilomar Conf. On Signals, Systems and Computers*, pp. 94–97, Nov. 1985.

[8] N. L. Owsley, "Enhanced minimum variance beamforming," in *Underwater Acoustic Data Processing*, Y. T. Chan, Ed. 1989, pp. 285–292, Kluwer.

[9] S. A. Vorobyov, A. B. Gershman, and Z.-Q. Luo, "Robust adaptive beamforming using worst case performance optimization," *IEEE Trans. Signal Proc.*, vol. 51, pp. 313–324, Feb. 2003.

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

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

[12] Y. C. Eldar, A. Ben-Tal, and A. Nemirovski, "Linear minimax regret estimation with bounded data uncertainties," Tech. Rep. CCIT 437, Technion—Israel Institute of Technology, Haifa, Israel, July 2003, *IEEE Trans. Signal Processing*, to appear.

[13] J. Capon, "High resolution frequency-wavenumber spectrum analysis," *Proc. IEEE*, vol. 57, pp. 1408–1418, Aug. 1969.

[14] R. T. Lacos, "Data adaptive spectral analysis methods," *Geophys.*, vol. 36, pp. 661–675, Aug. 1971.

Competitive Mean-Squared Error Beamforming

Yonina C. Eldar[(1)] *and Arye Nehorai*[(2)]

[(1)] Dept. of EE, Technion, Israel Institute of Technology
[(2)] Dept. of ECE, University of Illinois at Chicago

Competitive Mean-Squared Error Beamforming

- Conventional beamformers typically maximize the signal-to-interference-plus-noise ratio (SINR). However, this may not guarantee a good estimate of the signal.
- We derive beamformers for estimating a signal in the presence of interference and noise using the mean-squared error (MSE) as the performance criterion.
- Challenge: the MSE depends on the signal in the deterministic case, or its power in the stochastic case, both are unknown.
- To solve this problem, we minimize the worst-case $regret$, i.e. the difference between the true MSE and smallest attainable MSE with a linear estimator that knows the exact model. This results in a low MSE over a wide range of signal values.