1 Mean-Squared Error Beamforming for Signal Estimation: A Competitive Approach

YONINA C. ELDAR and ARYE NEHORAI

1.1 INTRODUCTION

Beamforming is a classical method of processing temporal sensor array measurements for signal estimation, interference cancellation, source direction, and spectrum estimation. It has ubiquitously been applied in areas such as radar, sonar, wireless communications, speech processing, and medical imaging (see, for example, [1, 2, 3, 4, 5, 6, 7, 8] and the references therein.)

Conventional approaches for designing data dependent 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 [9, 10, 11, 12], whereas others are developed to also overcome uncertainty in the steering vector, for example [13, 14, 15, 16, 17, 18, 19] (see also Chapters 2 and 3 in this book and the references therein.) Despite the fact that the SINR has been used as a measure of beamforming performance and as a design criterion in many beamforming approaches, we note that maximizing SINR may not guarantee a good estimate of

the signal. In an *estimation* context, where our goal is to design a beamformer in order to obtain an estimate of the signal amplitude that is close to its true value, it would make more sense to choose the weights to minimize an objective that is related to the estimation error, *i.e.*, the difference between the true signal amplitude and its estimate, rather than the SINR. Furthermore, in comparing performance of different beamforming methods, it may be more informative to consider the estimation error as a measure of performance.

In this chapter we derive five beamformers for estimating a signal in the presence of interference and noise using the mean-squared error (MSE) as the performance criterion assuming either known steering vectors or random steering vectors with known second-order statistics. Computing the MSE shows, however, that it depends explicitly on the unknown signal magnitude in the deterministic case, or the unknown signal second order moment in the stochastic case [8], 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 rely on a recent minimax estimation framework that has been developed for solving robust estimation problems [20, 21]. This framework considers a general linear estimation problem, and suggests two classes of linear estimators that optimize an MSE-based criterion. In the first approach, developed in [20], a linear estimator is developed to minimize the worst-case MSE over an ellipsoidal region of uncertainty on the parameter set. In the second approach, developed in [21], a linear estimator is designed whose performance is as close as possible to that of the optimal linear estimator for the case of known model parameters. 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. Note that as we explain further in Section 1.2.2, even when the signal magnitude is known, we cannot achieve a zero MSE with a *linear* estimator.

Using the general minimax MSE framework [20, 21] we develop two classes of robust beamformers: Minimax MSE beamformers and minimax regret beamformers. The minimax MSE beamformers minimize the worst-case MSE over all signals whose

magnitude (or variance, in the zero-mean stochastic signals case) is bounded by a constant. The minimax regret beamformers minimize the worst-case regret over all bounded signals, where this approach considers both an upper and a lower bound on the signal magnitude. We note that in practice, if bounds on the signal magnitude are not known, then they can be estimated from the data, as we demonstrate in the numerical examples.

We first consider the case in which the steering vector is known completely and develop a minimax MSE and minimax regret beamformer. In this case we show that the minimax beamformers are scaled versions of the classical SINR-based beamformers. We then consider the case in which the steering vector is not completely known or fully calibrated, for example due to errors in sensor positions, gains or phases, coherent and incoherent local scatters, receiver fluctuations due to temperature changes, quantization effects, etc. (see [22, 14] and the references therein). To model the uncertainties in the steering vector, we assume that it is a random vector with known mean and covariance. Under this model, we develop three possible beamformers: minimax MSE, minimax regret, and, following the ideas in [23], a least-squares beamformer. While the minimax beamformers require bounds on the signal magnitude, the least-squares beamformer does not require such bounds. As we show, in the case of a random steering vector, the beamformers resulting from the minimax approaches are fundamentally different than those resulting from the SINR approach, and are not just scaled versions of each other as in the known steering vector case.

To illustrate the advantages of our methods we present several numerical examples comparing the proposed methods with conventional SINR-based methods, and several recently proposed robust methods. For a known steering vector, the minimax MSE and minimax regret beamformers are shown to consistently have the best performance, particularly for negative SNR values. For random steering vectors, the minimax and least-squares beamformers are shown to have the best performance for low SNR values (-10 to 5 dB). As we show, the least-squares approach, which does not require bounds on the signal magnitude, often performs better than the recently proposed robust methods [14, 15, 16] for dealing with steering vector uncertainty. In this case,

the improvement in performance resulting from the proposed methods is often quite substantial.

The chapter is organized as follows. In Section 1.2 we present the problem formulation and review existing methods. In Section 1.3 we develop the minimax MSE and minimax regret beamformers for the case in which the steering vector is known. The case of a random steering vector is considered in Section 1.4. In Sections 1.5 and 1.6 we discuss practical considerations and present numerical examples illustrating the advantages of the proposed beamformers over several existing standard and robust beamformers, for a wide range of SNR values. The chapter is summarized in Section 1.7.

1.2 BACKGROUND AND 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. The matrix $\mathbf{I}$ denotes the identity matrix of the appropriate dimension, $(\cdot)^*$ denotes the Hermitian conjugate of the corresponding matrix, and $\hat{(\cdot)}$ denotes an estimated variable. The eigenvector of a matrix $\mathbf{A}$ associated with the largest eigenvalue is denoted by $\mathcal{P}\{\mathbf{A}\}$.

1.2.1 Background

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

$$\mathbf{y}(t) = s(t)\mathbf{a} + \mathbf{i}(t) + \mathbf{e}(t), \quad 1 \le t \le N, \tag{1.1}$$

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, $s(t)$ is the signal amplitude, $\mathbf{a}$ is the signal steering vector, $\mathbf{i}(t)$ is the interference, $\mathbf{e}(t)$ is a Gaussian noise vector and N is the number of snapshots [4, 6, 8]. In the above model we implicitly made the common assumption of a narrow-band signal.

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.

For concreteness, in our development below we will treat $s(t)$ as a deterministic signal. However, as we show analytically in Section 1.2.2 and through simulations in Section 1.6, the optimality properties of the algorithms we develop are valid also in the case of stochastic signals.

In some applications, such as in the case of a fully calibrated array, the steering vector can be assumed to be known exactly. In this case, we treat $\mathbf{a}$ as a known deterministic vector. However, in practice the array response may have some uncertainties or perturbations in the steering vectors. These perturbations may be due to errors in sensor positions, gains or phases, mutual couplings between sensors, receiver fluctuations due to temperature changes, quantization effects, and coherent and incoherent local scatters [22, 14]. To account for these uncertainties, several authors have tried modeling some of their effects [24, 25]. However these perturbations often take place simultaneously, which significantly complicates the model. Instead, the uncertainty in $\mathbf{a}$ can be taken into account by treating it as a deterministic vector that lies in an ellipsoid centered at a nominal steering vector [14, 15]. An alternative approach has been to treat the steering vector as a random vector assuming knowledge of its distribution [13] or the second-order statistics [26, 27, 28, 29, 30, 31, 32, 16]. In the latter case the mean value of $\mathbf{a}$ corresponds to the nominal steering vector, and the covariance matrix captures its perturbations. In this chapter we consider the cases where $\mathbf{a}$ is known (Section 1.3) or random with known second-order statistics (Section 1.4).

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

$$\hat{s}(t) = \mathbf{w}^*(t)\mathbf{y}(t), \quad 1 \leq t \leq N. \tag{1.2}$$

To illustrate our approach, in this section we focus primarily on the case in which the steering vector $\mathbf{a}$ is assumed to be known exactly. As we show in Section 1.4, the essential ideas we outline for this case can also be applied to the case in which $\mathbf{a}$ is random.

Traditionally, the beamformer weights $\mathbf{w}(t) = \mathbf{w}$ (where we omitted the index t for brevity) are chosen to maximize the SINR, which in the case of a known steering

vector is given by

$$\mathrm{SINR} \propto \frac{|\mathbf{w}^*\mathbf{a}|^2}{\mathbf{w}^*\mathbf{R}\mathbf{w}}, \tag{1.3}$$

where

$$\mathbf{R} = E\left\{(\mathbf{i}+\mathbf{n})(\mathbf{i}+\mathbf{n})^*\right\} \tag{1.4}$$

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

$$\mathbf{w}_{\mathrm{MVDR}} = \frac{1}{\mathbf{a}^*\mathbf{R}^{-1}\mathbf{a}}\mathbf{R}^{-1}\mathbf{a}. \tag{1.5}$$

The solution (1.5) is commonly referred to as the minimum variance distortionless response (MVDR) beamformer, since it can also be obtained as the solution to

$$\min_{\mathbf{w}} \mathbf{w}^*\mathbf{R}\mathbf{w} \quad \text{subject to} \quad \mathbf{w}^*\mathbf{a} = 1. \tag{1.6}$$

In practice, the interference-plus-noise covariance matrix $\mathbf{R}$ is often not available. In such cases, the exact covariance $\mathbf{R}$ in (1.5) is replaced by an estimated covariance. Various methods exist for estimating the covariance $\mathbf{R}$. The simplest approach is to choose the estimate as the sample covariance

$$\widehat{\mathbf{R}}_{sm} = \frac{1}{N}\sum_{t=1}^{N}\mathbf{y}(t)\mathbf{y}(t)^*. \tag{1.7}$$

The resulting beamformer is referred to as the sample matrix inversion (SMI) beamformer or the Capon beamformer [33, 34]. (For simplicity, we assume here that the sample covariance matrix is invertible.) If the signal is present in the training data, then it is well known that the performance of the MVDR beamformer with $\mathbf{R}$ replaced by $\widehat{\mathbf{R}}_{sm}$ of (1.7) degrades considerably [11].

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

$$\widehat{\mathbf{R}}_{dl} = \widehat{\mathbf{R}}_{sm} + \xi\mathbf{I} = \frac{1}{N}\sum_{t=1}^{N}\mathbf{y}(t)\mathbf{y}(t)^* + \xi\mathbf{I}, \tag{1.8}$$

where ξ is the diagonal loading factor. The resulting beamformer is referred to as the loaded SMI or the loaded Capon beamformer [9, 10]. Various methods have been proposed for choosing the diagonal loading factor ξ; see *e.g.,* [10]. 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 $\mathbf{R}$ is the eigenspace approach [11], in which the inverse of the covariance matrix is estimated as

$$\left(\widehat{\mathbf{R}}_{eig}\right)^{-1} = \left(\widehat{\mathbf{R}}_{sm}\right)^{-1} \mathbf{P}_s, \tag{1.9}$$

where $\mathbf{P}_s$ is the orthogonal projection onto the sample signal+interference subspace, *i.e.*, the subspace corresponding to the $D+1$ largest eigenvalues of $\widehat{\mathbf{R}}_{sm}$, were D is the known rank of the interference subspace.

Note that the Capon, loaded Capon and eigenspace beamformers, can all be viewed as MVDR beamformers with a particular estimate of $\mathbf{R}$.

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 [14, 15, 16]. 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-plus-noise covariance, namely the finite sample effects, and the fact that the signal is typically present in the training data. Each of the above robust methods is designed to maximize a measure of SINR on the uncertainty set. Specifically, in [14], the authors suggest minimizing $\mathbf{w}^*\widehat{\mathbf{R}}_{sm}\mathbf{w}$ 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

$$\mathbf{w} = \frac{\lambda}{\lambda \mathbf{a}^* \left(\widehat{\mathbf{R}}_{sm} + \lambda\epsilon^2\mathbf{I}\right)^{-1}\mathbf{a} - 1} \left(\widehat{\mathbf{R}}_{sm} + \lambda\epsilon^2\mathbf{I}\right)^{-1}\mathbf{a}, \tag{1.10}$$

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 [15] the authors consider a similar approach in which they first estimate the steering vector by minimizing $\tilde{\mathbf{a}}^*\widehat{\mathbf{R}}_{sm}^{-1}\tilde{\mathbf{a}}$ with respect to $\tilde{\mathbf{a}}$ subject to $\|\tilde{\mathbf{a}}-\mathbf{a}\|^2 = \epsilon$, and then use $\sqrt{M}\tilde{\mathbf{a}}/\|\tilde{\mathbf{a}}\|$ in the MVDR

beamformer, which results in the beamformer

$$\mathbf{w} = \frac{\alpha}{\sqrt{M}} \cdot \frac{\left(\lambda\widehat{\mathbf{R}}_{sm} + \mathbf{I}\right)^{-1}\mathbf{a}}{\mathbf{a}^*\left(\lambda\widehat{\mathbf{R}}_{sm} + \mathbf{I}\right)^{-1}\widehat{\mathbf{R}}_{sm}\left(\lambda\widehat{\mathbf{R}}_{sm} + \mathbf{I}\right)^{-1}\mathbf{a}}. \tag{1.11}$$

Here λ is chosen such that $\left\|\left(\mathbf{I} + \lambda\widehat{\mathbf{R}}_{sm}\right)^{-1}\mathbf{a}\right\|^2 = \epsilon$, and $\alpha = \left\|\left(\widehat{\mathbf{R}}_{sm}^{-1} + \lambda\mathbf{I}\right)^{-1}\mathbf{a}\right\|$. Finally, in [16] 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 $\mathbf{w}^*\widehat{\mathbf{R}}_{dl}\mathbf{w}$ subject to $|\mathbf{w}^*\mathbf{a}|^2 \geq 1 - \mathbf{w}^*\Delta\mathbf{w}$ for all $\|\Delta\| \leq \epsilon$, and is given by

$$\mathbf{w} = \alpha\mathcal{P}\left\{\widehat{\mathbf{R}}_{dl}^{-1}\left(\mathbf{a}\mathbf{a}^* - \epsilon\mathbf{I}\right)\right\}, \tag{1.12}$$

where α is chosen such that $\mathbf{w}^*(\mathbf{a}\mathbf{a}^* - \epsilon\mathbf{I})\mathbf{w} = 1$.

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 design 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 $\hat{s}(t) - s(t)$.

1.2.2 MSE Beamforming

If $\hat{s} = \mathbf{w}^*\mathbf{y}$, then, assuming that s is deterministic, the MSE between s and $\hat{s}$ is given by

$$E\left\{|\hat{s} - s|^2\right\} = V(\hat{s}) + |B(\hat{s})|^2 = \mathbf{w}^*\mathbf{R}\mathbf{w} + |s|^2|1 - \mathbf{w}^*\mathbf{a}|^2, \tag{1.13}$$

where $V(\hat{s}) = E\left\{|\hat{s} - E\left\{\hat{s}\right\}|^2\right\}$ is the variance of the estimate $\hat{s}$ and $B(\hat{s}) = E\left\{\hat{s}\right\} - s$ is the bias. In the case in which s is a zero-mean random variable with variance σ_s^2, the MSE is given by

$$E\left\{|\hat{s} - s|^2\right\} = \mathbf{w}^*\mathbf{R}\mathbf{w} + \sigma_s^2|1 - \mathbf{w}^*\mathbf{a}|^2. \tag{1.14}$$

Comparing (1.13) and (1.14) we see that the expressions for the MSE have the same form in the deterministic and stochastic cases, where $|s|^2$ in the deterministic case is replaced by σ_s^2 in the stochastic case. For concreteness, in the discussion in the rest of the chapter we assume the deterministic model. However, all the results hold true for the stochastic model where we replace $|s|^2$ everywhere with σ_s^2. In particular, in the development of the minimax MSE and regret beamformers in the stochastic case, the bounds on $|s|^2$ are replaced by bounds on the signal variance σ_s^2.

The minimum MSE (MMSE) beamformer minimizing the MSE when $|s|$ is known is obtained by differentiating (1.13) with respect to $\mathbf{w}$ and equating to 0, which results in

$$\mathbf{w}(s) = |s|^2(\mathbf{R} + |s|^2\mathbf{a}\mathbf{a}^*)^{-1}\mathbf{a}. \tag{1.15}$$

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

$$\mathbf{w}(s) = \frac{|s|^2}{1 + |s|^2\mathbf{a}^*\mathbf{R}^{-1}\mathbf{a}}\mathbf{R}^{-1}\mathbf{a}. \tag{1.16}$$

The MMSE beamformer can alternatively be expressed as

$$\mathbf{w}(s) = \frac{|s|^2\mathbf{a}^*\mathbf{R}^{-1}\mathbf{a}}{1 + |s|^2\mathbf{a}^*\mathbf{R}^{-1}\mathbf{a}} \cdot \frac{\mathbf{R}^{-1}\mathbf{a}}{\mathbf{a}^*\mathbf{R}^{-1}\mathbf{a}} = \beta(s)\mathbf{w}_{\mathrm{MVDR}}, \tag{1.17}$$

where

$$\beta(s) = \frac{|s|^2\mathbf{a}^*\mathbf{R}^{-1}\mathbf{a}}{1 + |s|^2\mathbf{a}^*\mathbf{R}^{-1}\mathbf{a}}. \tag{1.18}$$

The scaling $\beta(s)$ satisfies $0 \leq \beta(s) \leq 1$, and is monotonically increasing in $|s|^2$. Therefore, for any $|s|^2$, $\|\mathbf{w}(s)\| \leq \|\mathbf{w}_{\mathrm{MVDR}}\|$. Substituting $\mathbf{w}(s)$ back into (1.13), the smallest possible MSE, which we denote by $\mathrm{MSE}_{\mathrm{OPT}}$, is given by

$$\mathrm{MSE}_{\mathrm{OPT}} = \frac{|s|^2}{1 + |s|^2\mathbf{a}^*\mathbf{R}^{-1}\mathbf{a}}. \tag{1.19}$$

Using (1.13) we can also compute the MSE of the MVDR beamformer (1.5), which maximizes the SINR, assuming that $\mathbf{R}$ is known. Substituting $\mathbf{w} = (1/\mathbf{a}^*\mathbf{R}^{-1}\mathbf{a})\mathbf{R}^{-1}\mathbf{a}$ into (1.13), the MSE is

$$\mathrm{MSE}_{\mathrm{MVDR}} = \frac{1}{\mathbf{a}^*\mathbf{R}^{-1}\mathbf{a}}. \tag{1.20}$$

Comparing (1.19) with (1.20),

$$\mathrm{MSE}_{\mathrm{OPT}} = \frac{|s|^2}{1+|s|^2\mathbf{a}^*\mathbf{R}^{-1}\mathbf{a}} = \frac{1}{1/|s|^2+\mathbf{a}^*\mathbf{R}^{-1}\mathbf{a}} \leq \frac{1}{\mathbf{a}^*\mathbf{R}^{-1}\mathbf{a}} = \mathrm{MSE}_{\mathrm{MVDR}}. \tag{1.21}$$

Thus, as we expect, the MMSE beamformer always results in a smaller MSE than the MVDR beamformer. Therefore, in an estimation context where our goal is to estimate the signal amplitude, the MMSE beamformer will lead to better average performance.

From (1.17) we see that the MMSE beamformer is just a shrinkage of the MVDR beamformer (as we will show below, this is no longer true in the case of a random steering vector). Therefore, the two beamformers will result in the same SINR, so that the MMSE beamformer also maximizes the SINR. However, as (1.21) shows, this shrinkage factor impacts the MSE so that the MMSE beamformer has better MSE performance. To illustrate the advantage of the MMSE beamformer, we consider a numerical example. The scenario consists of a uniform linear array (ULA) of $M = 20$ omnidirectional sensors spaced half a wavelength apart. We choose s as a complex sinewave with varying amplitude to obtain the desired SNR in each sensor; its plane-wave has a DOA of 30^o relative to the array normal. The noise $\mathbf{e}$ is a zero-mean, Gaussian, complex random vector, temporally and spatially white, with a power of 0 dB in each sensor. The interference is given by $\mathbf{i} = \mathbf{a}_i i$ where i is a zero-mean, Gaussian, complex process temporally white with interference-plus-noise ratio (INR) of 20 dB and $\mathbf{a}_i$ is the interference steering vector with DOA = -30^o. Assuming knowledge of $|s|^2$, $\mathbf{R}$ and $\mathbf{a}$, we evaluate the square-root of the normalized MSE (NMSE) over a time-window of $N = 100$ samples, where each result is obtained by averaging 200 Monte Carlo simulations.

Figure 1.1 illustrates the NMSE of the MMSE and MVDR beamformers when estimating the complex sinewave, as a function of SNR. It can be seen that the MMSE beamformer outperforms the MVDR beamformer in all the illustrated SNR range. As the SNR increases, the MVDR beamformer converges to the MMSE beamformer, since $\beta(s)$ in (1.17) converges to 1. The absolute value of the original signal and its estimates obtained from the MMSE and MVDR beamformers are illustrated in

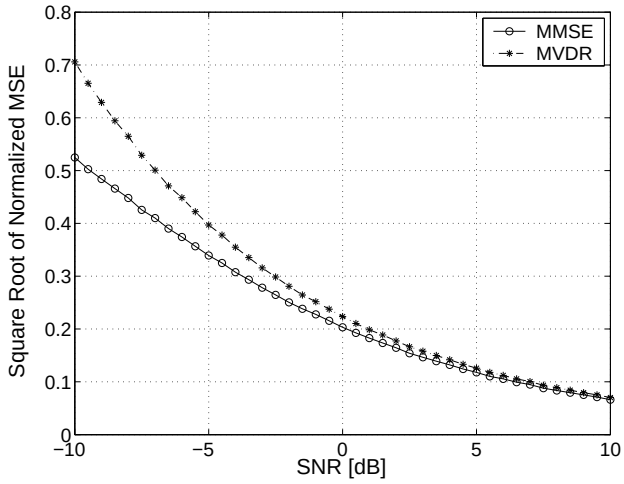

Fig. 1.1 Square-root of the normalized MSE as a function of SNR using the MMSE and MVDR beamformers when estimating a complex sinewave with DOA 30^o, in the presence of an interference with INR = 20 dB and DOA = -30^o. We assume that $|s|^2$, the noise-plus-interference covariance matrix $\mathbf{R}$ and the steering vector $\mathbf{a}$ are known.

Fig. 1.2, for an SNR of -10 dB. Clearly, the MMSE beamformer leads to a better estimate than the MVDR beamformer.

The difference between the MSE and SINR based approaches is more pronounced in the case of a random steering vector. Suppose that $\mathbf{a}$ is a random vector with mean $\mathbf{m}$ and covariance matrix $\mathbf{C}$. In Section 1.4 we consider this case in detail, and show that the beamformer minimizing the MSE is

$$\mathbf{w}(s) = \frac{|s|^2}{1 + |s|^2 \mathbf{m}^* \left(\mathbf{R} + |s|^2 \mathbf{C}\right)^{-1} \mathbf{m}} \left(\mathbf{R} + |s|^2 \mathbf{C}\right)^{-1} \mathbf{m}. \tag{1.22}$$

On the other hand, the beamformer maximizing the SINR is given by

$$\mathbf{w} = \alpha \mathcal{P}\left\{\mathbf{R}^{-1}(\mathbf{C} + \mathbf{m}\mathbf{m}^*)\right\}, \tag{1.23}$$

where α is chosen such that $\mathbf{w}^*(\mathbf{C} + \mathbf{m}\mathbf{m}^*)\mathbf{w} = 1$. We refer to this beamformer as the principal eigenvector (PEIG) beamformer [35]. Comparing (1.22) and (1.23) we see that in this case the two beamformers are in general not scalar versions of

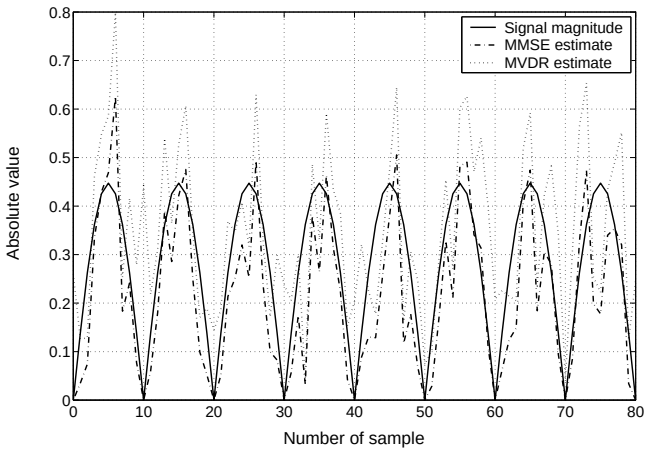

Fig. 1.2 Absolute values of the true complex sinewave, and its estimates obtained by the MMSE and MVDR beamformers for SNR = -10 dB, in the presence of an interference with INR = 20 dB and DOA = -30^o. We assume that $|s|^2$, the noise-plus-interference covariance matrix $\mathbf{R}$ and the steering vector $\mathbf{a}$ are known.

each other. As we now illustrate, the MMSE beamformer can result in a much better estimate of the signal amplitude and waveform than the SINR-based PEIG beamformer.

To illustrate the advantage of the MMSE beamformer in the case of a random steering vector, we consider an example in which the DOA of the signal is random. The scenario is similar to that of the previous example, where in place of a constant signal DOA, the DOA is now given by a Gaussian random variable with mean equal to 30^o and standard deviation equal to 1 (about $\pm 3^o$). The mean $\mathbf{m}$ and covariance matrix $\mathbf{C}$ of the steering vector are estimated from 2000 realizations of the steering vector. Assuming knowledge of $|s|^2$ and $\mathbf{R}$, we evaluate the NMSE over a time-window of $N = 100$ samples, where each result is obtained by averaging 200 Monte Carlo simulations.

Figure 1.3 illustrates the NMSE of the MMSE and PEIG beamformers as a function of SNR. It can be seen that the NMSE of the MMSE beamformer is substantially lower than that of the PEIG beamformer. The absolute value of the original signal

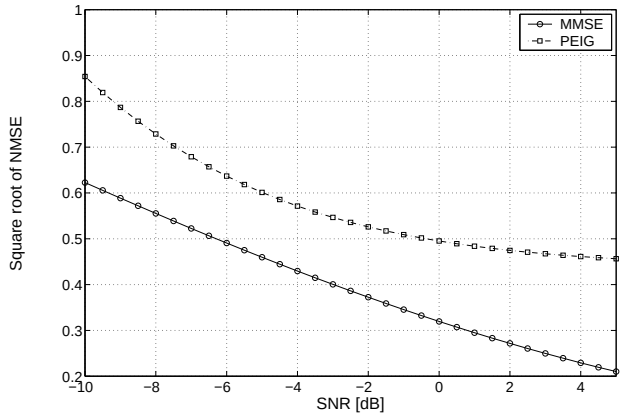

Fig. 1.3 Square-root of the normalized MSE as a function of SNR using the MMSE and PEIG beamformers when estimating a complex sinewave with random DOA, in the presence of an interference with INR = 20 dB and DOA = -30^o. We assume that $|s|^2$, the noise-plus-interference covariance matrix $\mathbf{R}$, and the mean $\mathbf{m}$ and covariance $\mathbf{C}$ of the steering vector $\mathbf{a}$ are known.

and its estimates obtained from the MMSE and PEIG beamformers are illustrated in Fig. 1.4 for an SNR of -10 dB. Clearly, the MMSE beamformer leads to a better signal estimate than the PEIG beamformer. It is also evident from this example that the signal waveform estimate obtained from both beamformers is different: the MMSE beamformer leads to a signal waveform that is much closer to the original waveform than the PEIG beamformer.

1.2.3 Robust MMSE Beamforming

Unfortunately, both in the case of known $\mathbf{a}$ and in the case of random $\mathbf{a}$ the MMSE beamformer depends explicitly on $|s|$ which is typically unknown. Therefore, in practice, we cannot implement the MMSE beamformer. The problem stems from the fact that the MSE depends explicitly on $|s|$. To illustrate the main ideas, in the remainder of this section we focus on the case of a deterministic steering vector.

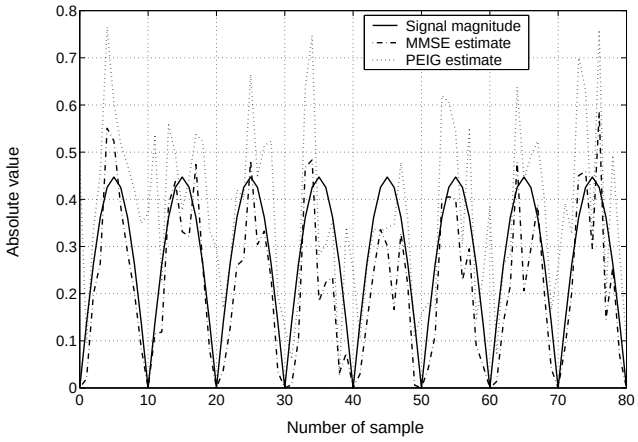

Fig. 1.4 Absolute values of the true complex sinewave with random DOA, and its estimates obtained by the MMSE and PEIG beamformers for SNR = -10 dB, in the presence of an interference with INR = 20 dB and DOA = -30^o. We assume that $|s|^2$, the noise-plus-interference covariance matrix $\mathbf{R}$, and the mean $\mathbf{m}$ and covariance $\mathbf{C}$ of the steering vector $\mathbf{a}$ are known.

One approach to obtain a beamformer that does not depend on $|s|$ in this case is to force the term depending on $|s|$, namely the bias, to 0, and then minimize the MSE, *i.e.*,

$$\min_{\mathbf{w}} \mathbf{w}^*\mathbf{R}\mathbf{w} \quad \text{subject to} \quad \mathbf{w}^*\mathbf{a} = 1, \tag{1.24}$$

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 may be far from s. Indeed, it is well known that unbiased estimators may often lead to large MSE values. The attractiveness of the SINR criterion is the fact that it is easy to solve, and leads to a beamformer that does not depend on $|s|$. Its drawback is that it does not necessarily lead to a small MSE, as can also be seen in Figs. 1.1 and 1.3.

We note, that as we discuss in Section 1.4.1, the property of the MVDR beamformer that it minimizes the MSE subject to a zero bias constrain no longer holds in the random steering vector case. In fact, when $\mathbf{a}$ is random with positive covariance matrix, there is in general no choice of linear beamformer for which the MSE is independent of $|s|$.

Instead of forcing the term depending on $|s|$ to zero, it would be desirable to design a robust beamformer whose MSE is reasonably small across all possible values of $|s|$. To this end we need to define the set of possible values of $|s|$. In some practical applications, we may know *a-priori* bounds on $|s|$, for example when the type of the source and the possible distances from the array are known, as can happen for instance in wireless communications and underwater source localization. We may have an upper bound of the form $|s| \leq U$, or we may have both an upper and (nonzero) lower bound, so that

$$L \leq |s| \leq U. \tag{1.25}$$

In our development of the minimax robust beamformers, we will assume that the bounds L and U are known. In practice, if no such bounds are known *a-priori*, then we can estimate them from the data, as we elaborate on further in Sections 1.5 and 1.6. This is similar in spirit to the MVDR-based beamformers: In developing the

MVDR beamformer it is assumed that the interference-plus-noise covariance matrix $\mathbf{R}$ is known; however, in practice, this matrix is estimated from the data.

Given an uncertainty set of the form (1.25), we may seek a beamformer that minimizes a worst-case MSE measure on this set. In the next section, we rely on ideas of [21] and [20], and propose two robust beamformers. We first assume that only an upper bound on $|s|$ is given, and develop a minimax MSE beamformer that minimizes the worst-case MSE over all $|s| \leq U$. As we show in (1.34), this beamformer is also minimax subject to (1.25). We then develop a minimax regret beamformer over the set defined by (1.25), that minimizes the worst-case difference between the MSE attainable with a beamformer that does not know $|s|$, and the optimal MSE of the MMSE beamformer that minimizes the MSE when $|s|$ is known. In Section 1.4 we develop minimax MSE and minimax regret beamformers for the case of a random steering vector.

1.3 MINIMAX MSE BEAMFORMING FOR KNOWN STEERING VECTOR

We now consider two MSE-based criteria for developing robust beamformers when the steering vector is known: In Section 1.3.1 we consider a minimax MSE approach and in Section 1.3.2 we consider a minimax regret approach. In our development, we assume that the covariance matrix $\mathbf{R}$ is known. In practice, as we discuss in Sections 1.5 and 1.6, the unknown $\mathbf{R}$ is replaced by an estimate $\widehat{\mathbf{R}}$.

1.3.1 Minimax MSE Beamforming

The first approach we consider for developing a robust beamformer it to minimize the worst-case MSE over all bounded values of $|s|$. Thus, we seek the beamformer that is the solution to

$$\min_{\mathbf{w}} \max_{|s|\leq U} E\left\{|\hat{s}-s|^2\right\} = \min_{\mathbf{w}} \max_{|s|\leq U} \left\{\mathbf{w}^*\mathbf{R}\mathbf{w} + |s|^2|1-\mathbf{w}^*\mathbf{a}|^2\right\}. \qquad (1.26)$$

This ensures that in the worst case, the MSE of the resulting beamformer is minimized.

Now,

$$\max_{|s|\leq U} \left\{ \mathbf{w}^* \mathbf{R} \mathbf{w} + |s|^2 |1 - \mathbf{w}^* \mathbf{a}|^2 \right\} = \mathbf{w}^* \mathbf{R} \mathbf{w} + U^2 |1 - \mathbf{w}^* \mathbf{a}|^2, \tag{1.27}$$

which is equal to the MSE of (1.13) with $|s|^2 = U$. It follows that the minimax MSE beamformer, denoted $\mathbf{w}_{\text{MXM}}$, is an MMSE beamformer of the form (1.16) matched to $|s| = U$:

$$\mathbf{w}_{\text{MXM}} = \frac{U^2}{1 + U^2 \mathbf{a}^* \mathbf{R}^{-1} \mathbf{a}} \mathbf{R}^{-1} \mathbf{a} = \beta_{\text{MXM}} \mathbf{w}_{\text{MVDR}}, \tag{1.28}$$

where

$$\beta_{\text{MXM}} = \frac{U^2 \mathbf{a}^* \mathbf{R}^{-1} \mathbf{a}}{1 + U^2 \mathbf{a}^* \mathbf{R}^{-1} \mathbf{a}}. \tag{1.29}$$

The resulting MSE is

$$\text{MSE}_{\text{MXM}} = \frac{U^4 \mathbf{a}^* \mathbf{R}^{-1} \mathbf{a} + |s|^2}{(1 + U^2 \mathbf{a}^* \mathbf{R}^{-1} \mathbf{a})^2}. \tag{1.30}$$

For any $|s|^2 \leq U^2$, we have that

$$\text{MSE}_{\text{MXM}} \leq \frac{U^2}{1 + U^2 \mathbf{a}^* \mathbf{R}^{-1} \mathbf{a}}. \tag{1.31}$$

For comparison, we have seen in (1.20) that the MSE of the MVDR beamformer is

$$\text{MSE}_{\text{MVDR}} = \frac{1}{\mathbf{a}^* \mathbf{R}^{-1} \mathbf{a}}, \tag{1.32}$$

from which we have immediately that for any choice of U,

$$\text{MSE}_{\text{MVDR}} > \text{MSE}_{\text{MXM}}. \tag{1.33}$$

The inequality (1.33) is valid as long as the true covariance $\mathbf{R}$ is known. When $\mathbf{R}$ is estimated from the data, (1.33) is no longer true in general. Nonetheless, as we will see in the simulations in Section 1.6, (1.33) typically holds even when both U and $\mathbf{R}$ are estimated from the data.

Finally, we point out that the minimax MSE beamformer $\mathbf{w}_{\text{MXM}}$ of (1.30) is also the solution in the case in which we have both an upper and lower bound on the norm

of s. This follows from the fact that

$$\max_{L\leq|s|\leq U} |s|^2|1-\mathbf{w}^*\mathbf{a}|^2 = \max_{|s|\leq U} |s|^2|1-\mathbf{w}^*\mathbf{a}|^2, \tag{1.34}$$

since the maximum is obtained at $|s| = U$. Thus, in contrast with the minimax regret beamformer developed in the next section, the minimax MSE beamformer does not take the lower bound (if available) into account and therefore may be overconservative when such a lower bound is known.

1.3.2 Minimax Regret Beamforming

We have seen that the minimax MSE beamformer is an MMSE beamformer matched to the worst possible choice of $|s|$, namely $|s| = U$. In some practical applications, particulary when a lower bound on $|s|$ is known, this approach may be overconservative. Although it optimizes the performance in the worst case, it may lead to deteriorated performance in other cases. To overcome this possible limitation, in this section we develop a minimax regret beamformer whose performance is as close as possible to that of the MMSE 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.

In [21], 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 positive definite matrix $\mathbf{T}$. It was shown that the linear minimax regret 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 below (see also [36]). 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 (1.25).

The minimax regret beamformer $\mathbf{w}_{\rm MXR}$ is designed to minimize the worst-case regret subject to the constraint $L \leq |s| \leq U$, where the regret, denoted $\mathcal{R}(s, \mathbf{w})$,

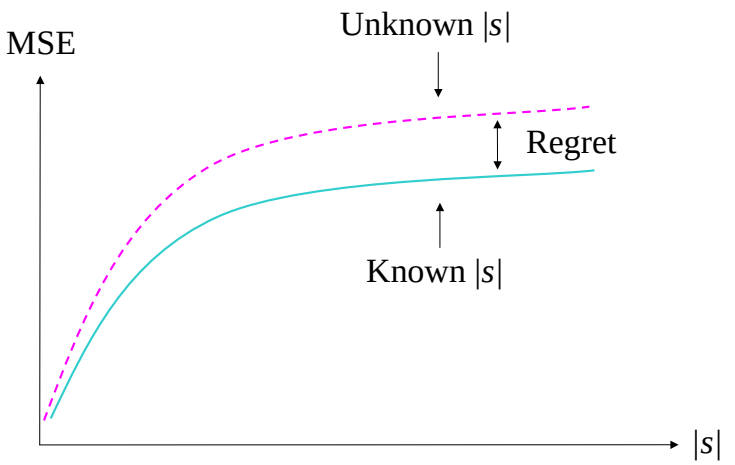

Fig. 1.5 The solid line represents the best attainable MSE as a function of $|s|$ when $|s|$ is known. The dashed line represents a desirable graph of MSE with small regret as a function of $|s|$ using some linear estimator that does not depend on $|s|$.

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, so that $\mathbf{w}$ can depend explicitly on s. We have seen in (1.19) that since we are restricting ourselves to linear beamformers, even in the case in which the beamformer can depend on s, the minimal attainable MSE is not generally equal to zero. The best possible MSE is illustrated schematically in Fig. 1.5. Instead of seeking an estimator to minimize the worst-case MSE, we therefore propose seeking an estimator to minimize the worst-case difference between its MSE and the best possible MSE, as illustrated in Fig. 1.5.

Using (1.19), we can express the regret as

$$\begin{aligned} \mathcal{R}(s,\mathbf{w}) &= E\left\{|\mathbf{w}^*\mathbf{y} - s|^2\right\} - \mathrm{MSE}_{\mathrm{OPT}} \\ &= \mathbf{w}^*\mathbf{R}\mathbf{w} + |s|^2|1 - \mathbf{w}^*\mathbf{a}|^2 - \frac{|s|^2}{1 + |s|^2\mathbf{a}^*\mathbf{R}^{-1}\mathbf{a}}. \end{aligned} \tag{1.35}$$

Thus $\mathbf{w}_{\mathrm{MXR}}$ is the solution to

$$\min_{\mathbf{w}} \max_{L\leq|s|\leq U} \mathcal{R}(s,\mathbf{w}) = \\ = \min_{\mathbf{w}} \left\{ \mathbf{w}^*\mathbf{R}\mathbf{w} + \max_{L\leq|s|\leq U} \left\{ |s|^2|1-\mathbf{w}^*\mathbf{a}|^2 - \frac{|s|^2}{1+|s|^2\mathbf{a}^*\mathbf{R}^{-1}\mathbf{a}} \right\} \right\}. \quad (1.36)$$

The minimax regret beamformer is given by the following theorem.

Theorem 1. *Let s denote an unknown signal amplitude 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 the problem*

$$\min_{\hat{s}=\mathbf{w}^*\mathbf{y}} \max_{L\leq|s|\leq U} \left\{ E\left\{|\hat{s}-s|^2\right\} - \min_{\hat{s}=\mathbf{w}^*(s)\mathbf{y}} E\left\{|\hat{s}-s|^2\right\} \right\} = \\ = \min_{\mathbf{w}} \max_{L\leq|s|\leq U} \left\{ \mathbf{w}^*\mathbf{R}\mathbf{w} + |s|^2(1-\mathbf{w}^*\mathbf{a})^2 - \frac{|s|^2}{1+|s|^2\mathbf{a}^*\mathbf{R}^{-1}\mathbf{a}} \right\}$$

is

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

Before proving Theorem 1, we first comment on some of the properties of the minimax regret beamformer, which, from Theorem 1, is given by

$$\mathbf{w}_{\mathrm{MXR}} = \left(1 - \frac{1}{\sqrt{(1+L^2\mathbf{a}^*\mathbf{R}^{-1}\mathbf{a})(1+U^2\mathbf{a}^*\mathbf{R}^{-1}\mathbf{a})}}\right) \frac{1}{\mathbf{a}^*\mathbf{R}^{-1}\mathbf{a}}\mathbf{R}^{-1}\mathbf{a}. \quad (1.37)$$

Comparing (1.37) with (1.5) we see that the minimax regret beamformer is a scaled version of the MVDR beamformer, i.e. $\mathbf{w}_{\mathrm{MXR}} = \beta_{\mathrm{MXR}}\mathbf{w}_{\mathrm{MVDR}}$ where

$$\beta_{\mathrm{MXR}} = 1 - \frac{1}{\sqrt{(1+L^2\mathbf{a}^*\mathbf{R}^{-1}\mathbf{a})(1+U^2\mathbf{a}^*\mathbf{R}^{-1}\mathbf{a})}}. \quad (1.38)$$

Clearly, the scaling β_{MXR} satisfies $0 \leq \beta_{\mathrm{MXR}} \leq 1$, and is monotonically increasing in L and U. In addition, when $U = L$,

$$\beta_{\mathrm{MXR}} = 1 - \frac{1}{1+U^2\mathbf{a}^*\mathbf{R}^{-1}\mathbf{a}^2} = \frac{U^2\mathbf{a}^*\mathbf{R}^{-1}\mathbf{a}}{1+U^2\mathbf{a}^*\mathbf{R}^{-1}\mathbf{a}}, \quad (1.39)$$

in which case the minimax regret beamformer is equal to the minimax MSE beamformer of (1.30). For $L < U$, $\|\mathbf{w}_{\mathrm{MXR}}\| < \|\mathbf{w}_{\mathrm{MXM}}\|$.

It is also interesting to note that the minimax regret beamformer can be viewed as an MMSE beamformer of the form (1.16), matched to

$$|s|^2 = \frac{1}{\mathbf{a}^*\mathbf{R}^{-1}\mathbf{a}} \left(\sqrt{(1+L^2\mathbf{a}^*\mathbf{R}^{-1}\mathbf{a})(1+U^2\mathbf{a}^*\mathbf{R}^{-1}\mathbf{a})} - 1 \right), \tag{1.40}$$

for arbitrary choices of L and U. This follows immediately from substituting $|s|^2$ of (1.40) into (1.16). Since the minimax regret estimator minimizes the MSE for the signal power given by (1.40), we may view this power as the "least-favorable" signal power in the regret sense.

The signal power (1.40) can be viewed as an estimate of the true, unknown signal power. To gain some insight into the estimate of (1.40) we note that

$$\begin{aligned} \left(\sqrt{(1+L^2\gamma)(1+U^2\gamma)} - 1 \right) \left(\sqrt{1+L^2\gamma} + \sqrt{1+U^2\gamma} \right) = \\ = \quad U^2\gamma\sqrt{1+L^2\gamma} + L^2\gamma\sqrt{1+U^2\gamma}, \end{aligned} \tag{1.41}$$

where for brevity, we denoted $\gamma = \mathbf{a}^*\mathbf{R}^{-1}\mathbf{a}$. Substituting (1.41) into (1.40), we have that

$$|s|^2 = \frac{U^2\sqrt{1+L^2\mathbf{a}^*\mathbf{R}^{-1}\mathbf{a}} + L^2\sqrt{1+U^2\mathbf{a}^*\mathbf{R}^{-1}\mathbf{a}}}{\sqrt{1+L^2\mathbf{a}^*\mathbf{R}^{-1}\mathbf{a}} + \sqrt{1+U^2\mathbf{a}^*\mathbf{R}^{-1}\mathbf{a}}}. \tag{1.42}$$

From (1.42) it follows that the unknown signal power is estimated as a weighted combination of the power bounds U^2 and L^2, where the weights depend explicitly on the uncertainty set and on $\mu(L)$ and $\mu(U)$, where

$$\mu(T) = T^2\mathbf{a}^*\mathbf{R}^{-1}\mathbf{a}, \tag{1.43}$$

can be viewed as the SNR in the observations, when the signal power is T^2.

If $\mu(L) \gg 1$, then from (1.40),

$$|s|^2 \approx \sqrt{U^2L^2}, \tag{1.44}$$

so that in this case the unknown signal power is estimated as the geometric mean of the power bounds. If, on the other hand, $\mu(L) \ll 1$, then from (1.42),

$$|s|^2 \approx \frac{1}{2}\left(L^2 + U^2\right). \tag{1.45}$$

Thus, in this case the unknown signal power is estimated as the algebraic mean of the power bounds.

It is interesting to note that while the minimax MSE estimator of (1.28) is matched to a signal power U^2, the minimax difference regret estimator of (1.37) is matched to a signal power $|s|^2 \leq (U^2 + L^2)/2$. This follows from (1.40) by using the inequality $\sqrt{ab} \leq (a+b)/2$.

We now prove Theorem 1. Although parts of the proof are similar to the proofs in [37], we repeat the arguments for completeness.

Proof. To develop a solution to (1.36), we first consider the inner maximization problem

$$\begin{aligned} f(\mathbf{w}) &= \max_{L \leq |s| \leq U} \left\{ |s|^2 |1 - \mathbf{w}^* \mathbf{a}|^2 - \frac{|s|^2}{1 + |s|^2 \gamma} \right\} \\ &= \max_{L^2 \leq x \leq U^2} \left\{ x |1 - \mathbf{w}^* \mathbf{a}|^2 - \frac{x}{1 + x\gamma} \right\}, \end{aligned} \tag{1.46}$$

where $x = |s|^2$. To derive an explicit expression for $f(\mathbf{w})$ we note that the function

$$h(x) = ax - \frac{bx}{c + dx} \tag{1.47}$$

with $b, c, d > 0$ is convex in $x \geq 0$. Indeed,

$$\frac{d^2 h}{dx^2} = 2 \frac{bd}{(c + dx)^3} > 0, \quad x \geq 0. \tag{1.48}$$

It follows that for fixed $\mathbf{w}$,

$$g(x) = x |1 - \mathbf{w}^* \mathbf{a}|^2 - \frac{x}{1 + x\gamma} \tag{1.49}$$

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,

$$\begin{aligned} f(\mathbf{w}) = & \\ &= \max_{L^2 \leq x \leq U^2} g(x) \\ &= \max \left(g(L^2), g(U^2) \right) \\ &= \max \left(L^2 |1 - \mathbf{w}^* \mathbf{a}|^2 - \frac{L^2}{1 + L^2 \gamma}, U^2 |1 - \mathbf{w}^* \mathbf{a}|^2 - \frac{U^2}{1 + U^2 \gamma} \right), \end{aligned} \tag{1.50}$$

and the problem (1.36) reduces to

$$\min_{\mathbf{w}} \left\{ \mathbf{w}^*\mathbf{R}\mathbf{w} + \max\left(L^2|1-\mathbf{w}^*\mathbf{a}|^2 - \frac{L^2}{1+L^2\gamma}, U^2|1-\mathbf{w}^*\mathbf{a}|^2 - \frac{U^2}{1+U^2\gamma} \right) \right\}. \tag{1.51}$$

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

$$\mathbf{w} = d(\mathbf{a}^*\mathbf{R}^{-1}\mathbf{a})^{-1}\mathbf{R}^{-1}\mathbf{a} = \frac{d}{\gamma}\mathbf{R}^{-1}\mathbf{a}, \tag{1.52}$$

for some (possibly complex) d. To this end, we first note that the objective in (1.51) 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

$$\mathbf{w} = \frac{\mathbf{a}^*\tilde{\mathbf{w}}}{\gamma}\mathbf{R}^{-1}\mathbf{a}. \tag{1.53}$$

Then

$$\mathbf{w}^*\mathbf{a} = \frac{\tilde{\mathbf{w}}^*\mathbf{a}}{\gamma}\mathbf{a}^*\mathbf{R}^{-1}\mathbf{a} = \tilde{\mathbf{w}}^*\mathbf{a}, \tag{1.54}$$

and

$$\mathbf{w}^*\mathbf{R}\mathbf{w} = \frac{|\mathbf{a}^*\tilde{\mathbf{w}}|^2}{\gamma^2}\mathbf{a}^*\mathbf{R}^{-1}\mathbf{a} = \frac{|\mathbf{a}^*\tilde{\mathbf{w}}|^2}{\gamma}. \tag{1.55}$$

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

$$|\mathbf{a}^*\mathbf{x}|^2 = |(\mathbf{R}^{-1/2}\mathbf{a})^*\mathbf{R}^{1/2}\mathbf{x}|^2 \leq \mathbf{a}^*\mathbf{R}^{-1}\mathbf{a}\mathbf{x}^*\mathbf{R}\mathbf{x} = \gamma\mathbf{x}^*\mathbf{R}\mathbf{x}. \tag{1.56}$$

Substituting (1.56) with $\mathbf{x} = \tilde{\mathbf{w}}$ into (1.55), we have that

$$\mathbf{w}^*\mathbf{R}\mathbf{w} \leq \frac{|\mathbf{a}^*\tilde{\mathbf{w}}|^2}{\gamma} \leq \tilde{\mathbf{w}}^*\mathbf{R}\tilde{\mathbf{w}}. \tag{1.57}$$

It follows from (1.54) and (1.57) that $\mathbf{w}$ is at least as good as $\tilde{\mathbf{w}}$ in the sense of minimizing (1.51). Therefore, the optimal value of $\mathbf{w}$ satisfies

$$\mathbf{w} = \frac{\mathbf{a}^*\mathbf{w}}{\gamma}\mathbf{R}^{-1}\mathbf{a}, \tag{1.58}$$

which implies that

$$\mathbf{w} = \frac{d}{\gamma}\mathbf{R}^{-1}\mathbf{a}, \tag{1.59}$$

for some d.

Combining (1.59) and (1.51), our problem reduces to

$$\min_d \left\{ \frac{|d|^2}{\gamma} + \max\left(L^2|1-d|^2 - \frac{L^2}{1+\gamma L^2}, U^2|1-d|^2 - \frac{U^2}{1+\gamma U^2} \right) \right\}. \quad (1.60)$$

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 (1.60) as

$$\min_{t,d} t \quad (1.61)$$

subject to

$$\begin{aligned} \frac{d^2}{\gamma} + L^2(1-d)^2 - \frac{L^2}{1+\gamma L^2} &\leq t; \\ \frac{d^2}{\gamma} + U^2(1-d)^2 - \frac{U^2}{1+\gamma U^2} &\leq t. \end{aligned} \quad (1.62)$$

The constraints (1.62) can be equivalently written as

$$\begin{aligned} f_L(d) &\triangleq \left(\frac{1}{\gamma} + L^2\right)\left(d - \frac{\gamma L^2}{1+\gamma L^2}\right)^2 \leq t; \\ f_U(d) &\triangleq \left(\frac{1}{\gamma} + U^2\right)\left(d - \frac{\gamma U^2}{1+\gamma U^2}\right)^2 \leq t. \end{aligned} \quad (1.63)$$

To develop a solution to (1.61) subject to (1.63), 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

$$\begin{aligned} d_L &= \frac{\gamma L^2}{1+\gamma L^2}; \\ d_U &= \frac{\gamma U^2}{1+\gamma U^2}. \end{aligned} \quad (1.64)$$

It therefore follows, that the optimal value of d, denoted d_0, satisfies

$$d_L \leq d_0 \leq d_U. \quad (1.65)$$

Indeed, let $t(d) = \max(f_L(d), f_U(d))$, and let $t_0 = t(d_0)$ be the optimal value of (1.61) subject to (1.63). 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

$$(1-d)^2 = \frac{1}{(1+\gamma L^2)(1+\gamma U^2)}, \tag{1.66}$$

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

$$d_\pm = 1 \pm \frac{1}{\sqrt{(1+\gamma L^2)(1+\gamma U^2)}}. \tag{1.67}$$

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

$$\frac{1}{1+\gamma U^2} \leq \frac{1}{\sqrt{(1+\gamma L^2)(1+\gamma U^2)}} \leq \frac{1}{1+\gamma L^2}, \tag{1.68}$$

we have that $d_- \in \mathcal{I}$.

In Fig. 1.6, we illustrate schematically the functions $f_L(d)$ and $f_U(d)$, where

$$d_e = d_- = 1 - \frac{1}{\sqrt{(1+\gamma L^2)(1+\gamma U^2)}} \tag{1.69}$$

is the unique intersection point of $f_L(d)$ and $f_U(d)$ in $\mathcal{I}$. For the specific choices of $f_L(d)$ and $f_U(d)$ drawn in the figure, it can be seen that the optimal value of d is $d_0 = d_e$. We now show that this conclusion holds true for any choice of the parameters. Indeed, if $L = U$, then $d_e = d_L = d_U$ so that from (1.65), $d_0 = d_e$. 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 $t_e = t(d_e)$ and noting that $t_e = f_L(d_e) = f_U(d_e)$, we conclude that for $d_e < d \leq d_L$, $f_U(d) > t_e$, and for $d_U \leq d < d_e$, $f_L(d) > t_e$ so that $t(d) > t_e$ for any $d \in \mathcal{I}$ such that $d \neq d_e$, and therefore $d_0 = d_e$. □

One advantage of the minimax regret beamformer is that it explicitly accounts for both the upper and the lower bounds on $|s|$, while the minimax MSE beamformer depends only on the upper bound. Therefore, in applications in which both bounds are available, the minimax regret beamformer can lead to better performance. As an example, in Fig. 1.7 we compare the square-root of the NMSE of the minimax regret

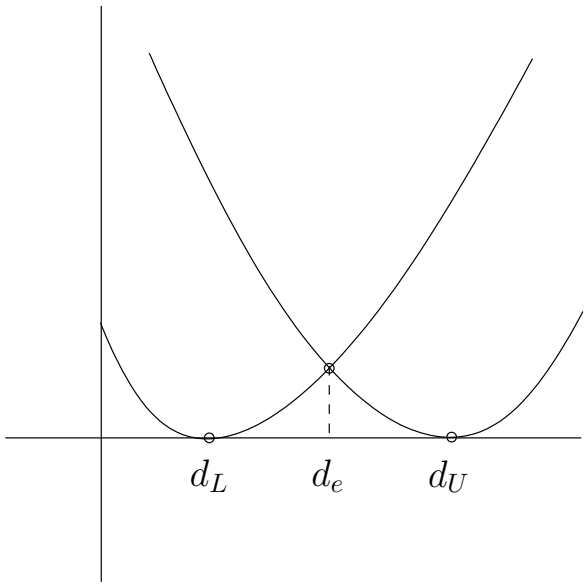

Fig. 1.6 Illustration of the functions $f_L(d)$ and $f_U(d)$ of (1.63).

(MXR) and minimax MSE (MXM) beamformers, where for comparison we also plot the NMSE of the MVDR beamformer, and the MMSE beamformer. Note that the MMSE beamformer cannot be implemented if $|s|$ is not known, however it serves as a bound on the NMSE. Each result was obtained by averaging 200 Monte Carlo simulations. The scenario we consider consists of a ULA of $M = 20$ omnidirectional sensors spaced half a wavelength apart. The signal of interest s is a complex random process, temporally white, whose amplitude has a uniform distribution between the values 3 and 6 and its plane-wave has a DOA of 30^o relative to the array normal. The noise $\mathbf{e}$ consists of a zero-mean complex Gaussian random vector, spatially and temporally white, with a varying power to obtain the desired SNR. The interference is given by $\mathbf{i} = \mathbf{a}_i i$ where i is a zero-mean complex Gaussian random process with INR = 20 dB and $\mathbf{a}_i$ its steering vector with DOA = -30^o. We assume that $\mathbf{R}$, $\mathbf{a}$ and the bounds $U = 6$ and $L = 3$ are known.

As we expect, the minimax regret beamformer outperforms the minimax MSE and MVDR beamformers in all the illustrated SNR range, and approaches the performance

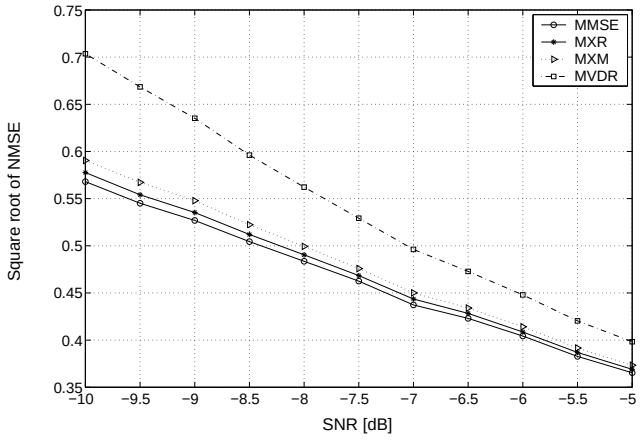

Fig. 1.7 Square-root of the normalized MSE as a function of SNR using the MMSE, MXR, MXM and MVDR beamformers in estimating a complex random process with amplitude uniformly distributed between 3 and 6 and DOA = 30^o, in the presence of an interference with INR = 20 dB and DOA = -30^o.

of the MMSE beamformer. In Section 1.6, Example 1, we show the advantages of the minimax MSE and minimax regret beamformers over several other beamformers for the case where $\mathbf{a}$ is known but $\mathbf{R}$, U and L are estimated from training data containing s.

1.4 RANDOM STEERING VECTOR

In Section 1.3, we explicitly assumed that the steering vector was deterministic and known. However, it is well known that the performance of adaptive beamformers degrades due to uncertainties or errors in the assumed array steering vectors [38, 28, 39, 40]. Several authors have considered this problem by modeling the steering vector as a random vector with known distribution [13] or known second-order statistics [26, 27, 28, 29, 30, 31, 32, 16]. In this section, we consider the case in which the steering vector is a random vector with mean $\mathbf{m}$ and covariance matrix $\mathbf{C}$. In this case, the mean $\mathbf{m}$ corresponds to a perfectly calibrated array, *i.e.*, the

perturbation-free steering vector, and $\mathbf{C}$ represents the perturbations to the steering vector.

In the case of a random steering vector, the SINR is given by [16, 35]

$$\mathrm{SINR} \propto \frac{\mathbf{w}^* \mathbf{R}_s \mathbf{w}}{\mathbf{w}^* \mathbf{R} \mathbf{w}}, \tag{1.70}$$

where

$$\mathbf{R}_s = E\left\{\mathbf{a}\mathbf{a}^*\right\} = \mathbf{C} + \mathbf{m}\mathbf{m}^* \tag{1.71}$$

is the signal correlation matrix, whose rank can be between 1 and M. In the case in which $\mathbf{a}$ is deterministic so that $\mathbf{a} = \mathbf{m}$, $\mathbf{R}_s = \mathbf{a}\mathbf{a}^*$ and the SINR of (1.70) reduces to the SINR of (1.3).

As in the case of deterministic $\mathbf{a}$, the most common approach to designing a beamformer is to maximize the SINR, which results in the principle eigenvector beamformer

$$\mathbf{w} = \alpha \mathcal{P}\left\{\mathbf{R}^{-1}\mathbf{R}_s\right\}, \tag{1.72}$$

where α is chosen such that $\mathbf{w}^*\mathbf{R}_s\mathbf{w} = 1$. The beamformer (1.72) can also be obtained as the solution to

$$\min_{\mathbf{w}} \mathbf{w}^*\mathbf{R}\mathbf{w} \quad \text{subject to} \quad \mathbf{w}^*\mathbf{R}_s\mathbf{w} = 1. \tag{1.73}$$

In practice, the interference-plus-noise matrix $\mathbf{R}$ is replaced by an estimate.

However, choosing $\mathbf{w}$ to maximize the SINR does not necessarily result in an estimated signal amplitude $\hat{s}$ that is close to s. Instead, it would be desirable to minimize the MSE. In the case of a random steering vector, the MSE between s and $\hat{s}$ is given by the expectation of the MSE of (1.13) with respect to $\mathbf{a}$, so that

$$\begin{aligned} E\left\{|\hat{s} - s|^2\right\} &= \mathbf{w}^*\mathbf{R}\mathbf{w} + |s|^2 E\left\{|1 - \mathbf{w}^*\mathbf{a}|^2\right\} \\ &= \mathbf{w}^*\mathbf{R}\mathbf{w} + |s|^2 E\left\{|1 - \mathbf{w}^*\mathbf{m} - \mathbf{w}^*(\mathbf{a} - \mathbf{m})|^2\right\} \\ &= \mathbf{w}^*\mathbf{R}\mathbf{w} + |s|^2 \left(|1 - \mathbf{w}^*\mathbf{m}|^2 + \mathbf{w}^*\mathbf{C}\mathbf{w}\right). \end{aligned} \tag{1.74}$$

The MMSE beamformer minimizing the MSE when $|s|$ is known is obtained by differentiating (1.74) with respect to $\mathbf{w}$ and equating to 0, which results in [23]

$$\begin{aligned}\mathbf{w}(s) &= |s|^2(\mathbf{R}+|s|^2\mathbf{C}+|s|^2\mathbf{m}\mathbf{m}^*)^{-1}\mathbf{m} \\ &= \frac{|s|^2}{1+|s|^2\mathbf{m}^*\left(\mathbf{R}+|s|^2\mathbf{C}\right)^{-1}\mathbf{m}}\left(\mathbf{R}+|s|^2\mathbf{C}\right)^{-1}\mathbf{m}. \qquad (1.75)\end{aligned}$$

Note, that if $\mathbf{C}=\mathbf{0}$, so that $\mathbf{a}=\mathbf{m}$ (with probability one), then (1.75) reduces to

$$\mathbf{w}(s)=\frac{|s|^2}{1+|s|^2\mathbf{m}^*\mathbf{R}^{-1}\mathbf{m}}\mathbf{R}^{-1}\mathbf{m}, \qquad (1.76)$$

which is equal to the MMSE beamformer of (1.16) with $\mathbf{a}=\mathbf{m}$. Comparing (1.75) with (1.72) we see that in general the MMSE beamformer and the SINR beamformer are not scaled versions of each other, as in the known steering vector case. Substituting $\mathbf{w}(s)$ back into (1.74), the smallest possible MSE, which we denote by $\mathrm{MSE}_{\mathrm{OPT}}$, is given by

$$\mathrm{MSE}_{\mathrm{OPT}}=|s|^2-|s|^4\mathbf{m}^*\left(\mathbf{R}+|s|^2\mathbf{C}\right)^{-1}\mathbf{m}. \qquad (1.77)$$

Since the optimal beamformer (1.76) depends explicitly on $|s|$, it cannot be implemented if $|s|$ is not known. Following the same framework as in the deterministic steering vector case, we consider two alternative approaches for designing a beamformer when $|s|$ is not known: a minimax MSE approach that minimizes the worst-case MSE over all $|s|\leq U$, and a minimax MSE regret that minimizes the worst-case regret, which in the case of a random steering vector is given by

$$\begin{aligned}\mathcal{R}(s,\mathbf{w}) &= \mathbf{w}^*\mathbf{R}\mathbf{w}+|s|^2\left(-\mathbf{w}^*\mathbf{m}-\mathbf{m}^*\mathbf{w}+\mathbf{w}^*(\mathbf{C}+\mathbf{m}\mathbf{m}^*)\mathbf{w}\right) \\ &\quad +|s|^4\mathbf{m}^*\left(\mathbf{R}+|s|^2\mathbf{C}\right)^{-1}\mathbf{m}. \qquad (1.78)\end{aligned}$$

The minimax MSE and minimax regret estimators for the linear estimation problem of estimating s in the model $\mathbf{y}=\mathbf{h}s+\mathbf{n}$, where $\mathbf{h}$ is a random vector with mean $\mathbf{m}$ and covariance $\mathbf{C}$, and $\mathbf{n}$ is a noise vector with covariance $\mathbf{R}$, has been considered in [23]. From the results in [23], the minimax MSE beamformer is

$$\mathbf{w}_{\mathrm{MXMR}}=\frac{U^2}{1+U^2\mathbf{m}^*\left(\mathbf{R}+U^2\mathbf{C}\right)^{-1}\mathbf{m}}\left(\mathbf{R}+U^2\mathbf{C}\right)^{-1}\mathbf{m}, \qquad (1.79)$$

which is just an MMSE estimator matched to $|s| = U$. The minimax regret beamformer for this problem is

$$\mathbf{w}_{\mathrm{MXRR}} = \frac{\gamma}{1 + \gamma \mathbf{m}^* \left(\mathbf{R} + \gamma \mathbf{C}\right)^{-1} \mathbf{m}} \left(\mathbf{R} + \gamma \mathbf{C}\right)^{-1} \mathbf{m}, \tag{1.80}$$

where $L^2 \leq \gamma \leq U^2$ is the unique root of $\mathcal{G}(\gamma)$ defined by

$$\mathcal{G}(\gamma) = \sum_{i=1}^{n} \frac{|y_i|^2 \left(U^2 - L^2\right)}{(1 + U^2\lambda_i)(1 + L^2\lambda_i)\lambda_i} \left(\frac{(1 + U^2\lambda_i)(1 + L^2\lambda_i)}{(1 + \gamma\lambda_i)^2} - 1 \right). \tag{1.81}$$

Here y_i is the ith component of $\mathbf{y} = \mathbf{V}^* \mathbf{R}^{-1/2} \mathbf{m}$, $\mathbf{V}$ is the unitary matrix in the eigendecomposition of $\mathbf{A} = \mathbf{R}^{-1/2}(\mathbf{C} + \mathbf{m}\mathbf{m}^*)\mathbf{R}^{-1/2}$, and λ_i is the ith eigenvalue of $\mathbf{A}$.

We see that in the case of a random steering vector, the minimax beamformers are in general no longer scaled versions of the SINR-based principle eigenvector beamformer (1.72), but rather point in different directions. Through several numerical examples (see Section 1.6, Example 2) we demonstrate the advantages of the minimax MSE and minimax regret beamformers over the principal eigenvector solution as well as over some alternative robust beamformers [14, 15, 16] for a wide range of SNR values.

1.4.1 Least-Squares Approach

In the far-field point source case, we have seen that the MVDR approach, which consists of maximizing the SINR, is equivalent to minimizing the MSE subject to the constraint that the beamformer is unbiased. The MVDR beamformer is also the least-squares beamformer, which minimizes the weighted least-squares error

$$\epsilon_{\mathrm{LS}} = (\mathbf{y} - \mathbf{a}\hat{s})\mathbf{R}^{-1}(\mathbf{y} - \mathbf{a}\hat{s}). \tag{1.82}$$

As we now show, these equivalences no longer hold in the case of a random steering vector. First, we note that in this case of a random steering vector, the variance also depends on the unknown power $|s|^2$. Therefore, in this case, the unbiased beamformer that minimizes the variance depends explicitly on s and therefore cannot be implemented. Furthermore, for a full-rank model in which $\mathbf{C}$ is positive definite,

there is no choice of beamformer that will result in an MSE that is independent of s (unless, of course, $s = 0$). This follows from the fact that the term depending on s in the MSE is

$$|s|^2 \left(|1 - \mathbf{w}^* \mathbf{m}|^2 + \mathbf{w}^* \mathbf{C} \mathbf{w} \right). \tag{1.83}$$

Since $\mathbf{w}^* \mathbf{C} \mathbf{w} > 0$ for any nonzero $\mathbf{w}$, this term cannot be equal to zero.

Following the ideas in [23], we now consider the least-squares beamformer for the case of a random steering vector. In this case, the least-squares error (1.82) depends on $\mathbf{a}$, which is random. Thus, instead of minimizing the error $\epsilon_{\rm LS}$ directly, we may consider minimizing the expectation of $\epsilon_{\rm LS}$ with respect to $\mathbf{a}$, which is given by

$$\begin{aligned} E\left\{\epsilon_{\rm LS}\right\} &= E\left\{(\mathbf{y} - \mathbf{m}\hat{s} - (\mathbf{a} - \mathbf{m})\hat{s})^* \mathbf{R}^{-1} (\mathbf{y} - \mathbf{m}\hat{s} - (\mathbf{a} - \mathbf{m})\hat{s})\right\} \\ &= (\mathbf{y} - \mathbf{m}\hat{s})^* \mathbf{R}^{-1} (\mathbf{y} - \mathbf{m}\hat{s}) + \hat{s}^2 E\left\{(\mathbf{a} - \mathbf{m})^* \mathbf{R}^{-1} (\mathbf{a} - \mathbf{m})\right\} \\ &= (\mathbf{y} - \mathbf{m}\hat{s})^* \mathbf{R}^{-1} (\mathbf{y} - \mathbf{m}\hat{s}) + \hat{s}^2 \mathrm{Tr}\left(\mathbf{R}^{-1}\mathbf{C}\right). \end{aligned} \tag{1.84}$$

Differentiating (1.84) with respect to $\hat{s}$ and equating to 0, we have that

$$\hat{s} = \frac{1}{\mathrm{Tr}\left(\mathbf{R}^{-1}\mathbf{C}\right) + \mathbf{m}^* \mathbf{R}^{-1} \mathbf{m}} \mathbf{m}^* \mathbf{R}^{-1} \mathbf{y}. \tag{1.85}$$

Thus, the least-squares beamformer is

$$\mathbf{w}_{\rm LS} = \frac{1}{\mathrm{Tr}\left(\mathbf{R}^{-1}\mathbf{C}\right) + \mathbf{m}^* \mathbf{R}^{-1} \mathbf{m}} \mathbf{R}^{-1} \mathbf{m}, \tag{1.86}$$

which, in general, is different than the MVDR beamformer of (1.72). It is interesting to note that the beamformer of (1.86) is a scaled version of the MVDR beamformer for known $\mathbf{a} = \mathbf{m}$.

The advantage of the least-squares approach is that it does not require bounds on the signal magnitude; in fact, it assumes the same knowledge as the principle eigenvector beamformer which maximizes the SINR. In Section 1.6 Example 2 we illustrate through numerical examples that for a wide range of SNR values the least-squares beamformer has smaller NMSE than the principal eigenvector beamformer (1.72) as well as the robust solutions [14, 15, 16]. These observations are true even when $\mathbf{m}$ and $\mathbf{C}$ are not known exactly, but are rather chosen in an ad-hoc manner. Therefore, in terms of NMSE, the least-squares approach appears to often be preferable over

standard and robust methods, while requiring the same prior knowledge. As we show, the NMSE performance can be improved further by using the minimax MSE and minimax regret methods; however, these methods require prior estimates of the signal magnitude bounds.

1.5 PRACTICAL CONSIDERATIONS

In our development of the minimax MSE and minimax regret beamformers, we assumed that there exists an upper bound U on the magnitude of the signal to be estimated, as well as a lower bound L for the minimax regret beamformer. For random steering vectors we assumed also the knowledge of their mean $\mathbf{m}$ and covariance $\mathbf{C}$.

In some applications, the bounds U and L may be known, for example based on *a-priori* knowledge on the type of the source and its possible range of distances from the array. 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 $s(t)$ as $\hat{s}(t) = \mathbf{w}_c^*\mathbf{y}(t)$ (the dependence on the time index t is presented for clarity). We may then use this estimate to obtain approximate values for U and L. In the simulations below, we use $U = (1+\beta)^2\|\mathbf{w}_c^*\mathbf{Y}\|$ and $L = (1-\beta)^2\|\mathbf{w}_c^*\mathbf{Y}\|$ for some β, where $\mathbf{Y}$ is the training data matrix of dimension $M \times N$ and $\|\cdot\|$ is the average norm over the training interval.

Since in most applications the true covariance $\mathbf{R}$ is not available we have to estimate it, e.g., using (1.7). However, as we discussed in Section 1.2, if $s(t)$ is present in the training data, then a diagonal loading (1.8) may perform better than (1.7). Therefore, in the simulations, the true $\mathbf{R}$ is replaced by (1.8).

In the case of a random steering vector, our beamformers rely on knowledge of the steering vector $\mathbf{m}$ and covariance $\mathbf{C}$. These parameters can either be estimated from observations of the steering vector, or, they can be approximated by choosing $\mathbf{m}$ to be equal to a nominal steering vector, and choosing $\mathbf{C} = \nu\mathbf{I}$, where ν reflects the uncertainty size around the nominal steering vector.

1.6 NUMERICAL EXAMPLES

To evaluate and compare the performance of our methods with other techniques, we conducted numerical examples using scenarios similar to [16]. Specifically, we consider a uniform linear array of $M = 20$ omnidirectional sensors spaced half a wavelength apart. In all the examples below, we choose $s(t)$ to be either a complex sinewave or a zero-mean complex Gaussian random process, temporally white, with varying amplitude or variance respectively, to obtain the desired SNR in each sensor. The signal $s(t)$ is continuously present throughout the training data. The noise $\mathbf{e}(t)$ is a zero-mean, Gaussian, complex random vector, temporally and spatially white, with a power of 0 dB in each sensor. The interference is given by $\mathbf{i}(t) = \mathbf{a}_i i(t)$ where $i(t)$ is a zero-mean, Gaussian, complex process temporally white and $\mathbf{a}_i$ is the interference steering vector. To illustrate the performance of the beamformers we use the square-root of the NMSE, which is obtained by averaging 200 Monte Carlo simulations. Unless otherwise stated, we use a signal with SNR=-5 dB, and an interference with DOA = -30^o, INR = 20 dB and $N = 50$ training snapshots. For brevity, in the remainder of this section we use the following notation for the different beamformers: MXM (minimax MSE), MXR (minimax regret), MXMR (minimax MSE for random $\mathbf{a}$), MXRR (minimax regret for random $\mathbf{a}$), LSR (least-squares for random $\mathbf{a}$); see also Table 1.1.

We consider two examples: In Example 1 we evaluate the performance of the MXM and MXR beamformers using the exact knowledge of the steering vector. In Example 2 we evaluate the MXMR, MXRR and LSR beamformers for a mismatch in the signal DOA. We focus on low SNR values (important e.g. in sonar) and compare the performance of the proposed methods against seven alternative methods: the Capon beamformer (CAPON) [33, 34], loading Capon beamformer (L-CAPON) [9, 10], eigenspace-based beamformer (EIG) [11, 12], and the robust beamformers of (1.10), (1.11) and (1.12) which we refer to, respectively, as ROB1, ROB2, and ROB3 [14, 15, 16]. In Example 2 we compare our methods against the principal eigenvector beamformer [35], which we refer to as PEIG, with $\mathbf{R}$ given by (1.8) and $\mathbf{R}_s$ given by exact knowledge or an ad-hoc estimate (see Example 2 for more details.)

Beamformer	Expression	Parameters	Ref.
L-CAPON	$\frac{1}{\mathbf{a}^*\widehat{\mathbf{R}}_{dl}^{-1}\mathbf{a}}\widehat{\mathbf{R}}_{dl}^{-1}\mathbf{a}$	ξ = 10	[10]
EIG	$\frac{1}{\mathbf{a}^*\widehat{\mathbf{R}}_{eig}^{-1}\mathbf{a}}\widehat{\mathbf{R}}_{eig}^{-1}\mathbf{a}$	D = 1	[11]
ROB1	$\frac{\lambda}{\lambda\mathbf{a}^*\left(\widehat{\mathbf{R}}_{sm}+\lambda\epsilon^2\mathbf{I}\right)^{-1}\mathbf{a}-1}\left(\widehat{\mathbf{R}}_{sm}+\lambda\epsilon^2\mathbf{I}\right)^{-1}\mathbf{a}$	ϵ = 3	[14]
ROB2	$\frac{\alpha_1\left(\lambda\widehat{\mathbf{R}}_{sm}+\mathbf{I}\right)^{-1}\mathbf{a}}{\sqrt{M}\mathbf{a}^*\left(\lambda\widehat{\mathbf{R}}_{sm}+\mathbf{I}\right)^{-1}\widehat{\mathbf{R}}_{sm}\left(\lambda\widehat{\mathbf{R}}_{sm}+\mathbf{I}\right)^{-1}\mathbf{a}}$	ϵ = 3.5	[15]
ROB3	$\alpha_2\mathcal{P}\left\{\widehat{\mathbf{R}}_{dl}^{-1}\left(\mathbf{a}\mathbf{a}^*-\epsilon\mathbf{I}\right)\right\}$	$\xi = 30$, ϵ = 9	[16]
PEIG	$\alpha_3\mathcal{P}\left\{\mathbf{R}^{-1}\mathbf{R}_s\right\}$		[35]
MXM	$\frac{U^2}{1+U^2\mathbf{a}^*\mathbf{R}^{-1}\mathbf{a}}\mathbf{R}^{-1}\mathbf{a}$	U	[41]
MXR	$\frac{\alpha_4}{\mathbf{a}^*\mathbf{R}^{-1}\mathbf{a}}\mathbf{R}^{-1}\mathbf{a}$	U, L	[36]
MXMR	$\frac{U^2}{1+U^2\mathbf{m}^*(\mathbf{R}+U^2\mathbf{C})^{-1}\mathbf{m}}\left(\mathbf{R}+U^2\mathbf{C}\right)^{-1}\mathbf{m}$	U, $\mathbf{m}$, $\mathbf{C}$	
MXRR	$\frac{\gamma}{1+\gamma\mathbf{m}^*(\mathbf{R}+\gamma\mathbf{C})^{-1}\mathbf{m}}\left(\mathbf{R}+\gamma\mathbf{C}\right)^{-1}\mathbf{m}$	U, L, $\mathbf{m}$, $\mathbf{C}$	
LSR	$\frac{1}{\mathrm{Tr}(\mathbf{R}^{-1}\mathbf{C})+\mathbf{m}^*\mathbf{R}^{-1}\mathbf{m}}\mathbf{R}^{-1}\mathbf{m}$	$\mathbf{m}$, $\mathbf{C}$	

Table 1.1 Beamformers used in the numerical examples.

The parameters of each of the compared methods were chosen as suggested in the literature. Namely, for L-CAPON (1.8) and PEIG (1.72) the diagonal loading was set as $\xi = 10\sigma_w^2$ [14, 15] with σ_w^2 being the variance of the noise in each sensor, assumed to be known ($\sigma_w^2 = 1$ in these examples); for the EIG beamformer (1.9) it was assumed that the low-rank condition and number of interferers are known. For the alternative robust methods, the parameters were chosen as follows: For ROB1 (1.10) the upper bound on the steering vector uncertainty was set as $\epsilon = 3$ [14], for ROB2 (1.11) $\epsilon = 3.5$, and for ROB3 (1.12), $\epsilon = 9$ and the diagonal loading was chosen as $\xi = 30$. Table 1.1 summarizes the beamformers implemented in the simulations, where $\alpha_1 = \left\|\left(\widehat{\mathbf{R}}_{sm}^{-1}+\lambda\mathbf{I}\right)^{-1}\mathbf{a}\right\|$, α_2 is chosen such that the corresponding beamformer satisfies $\mathbf{w}^*(\mathbf{a}\mathbf{a}^*-\epsilon\mathbf{I})\mathbf{w} = 1$, α_3 is chosen such that $\mathbf{w}^*\mathbf{R}_s\mathbf{w} = 1$, and $\alpha_4 = 1 - 1/\sqrt{(1+L^2\mathbf{a}^*\mathbf{R}^{-1}\mathbf{a})(1+U^2\mathbf{a}^*\mathbf{R}^{-1}\mathbf{a})}$.

Example 1 - Known steering vector: In this example we assume that the steering vector $\mathbf{a}$ is known. We first choose $s(t)$ as a complex sinewave with DOA of its

Beamformer	$\mathbf{R}$	$\mathbf{w}_c$	β
MXM	$\mathbf{R}_{dl}$, ξ = 10	L-Capon, ξ = 10	9
MXR	$\mathbf{R}_{dl}$, ξ = 10	L-Capon, ξ = 10	6

Table 1.2 Specification of the MXM and MXR beamformers in Example 1.

plane-wave equal to 30^o relative to the array normal. We implemented the MXR and MXM beamformers with the sample covariance matrix estimated using a loading factor $\xi = 10$ [14, 15], $\mathbf{w}_c$ given by the L-CAPON beamformer with $\xi = 10$ and β set as 6 and 9, for these beamformers, respectively. The values of β selected for MXM and MXR were those that gave the best performance over a wide range of negative SNR values. Table 1.2 summaries the parameters chosen for MXM and MXR in this example.

In Fig. 1.8 we plot the square-root of the NMSE as a function of the SNR using the MXR, MXM, EIG, L-CAPON, ROB2, and ROB3 beamformers. Since in all the scenarios considered in this example the NMSE of the CAPON and ROB1 beamformers was out of the illustrated scales, we do not plot the performance of these methods. It can be seen in Fig. 1.8 that the MXM beamformer has the best performance for SNR values between -10 to -3 dB, and the MXR beamformer has the best performance for SNR values between -2 to 2 dB. In Fig. 1.9 we plot the square-root of the NMSE as a function of the number of training data with SNR=-5 dB. The performance of the proposed methods as a function of the difference between the signal and interference DOAs is illustrated in Fig. 1.10. The NMSE of all the methods remains constant for DOA differences between 10^o and 90^o, but deteriorates for DOA differences close to 0^o. Despite this, MXR and MXM continue to outperform the other methods. The performance of ROB2 and ROB3 is out of the illustrated scale for DOA differences less than 3^o, because the uncertainty region of their steering vectors overlaps with that of the interference. Figure 1.11 illustrates the performance as a function of the signal-to-interference-ratio (SIR) for an SNR of -5 dB. As expected, the NMSE of all the beamformers decreases as the SIR increases,

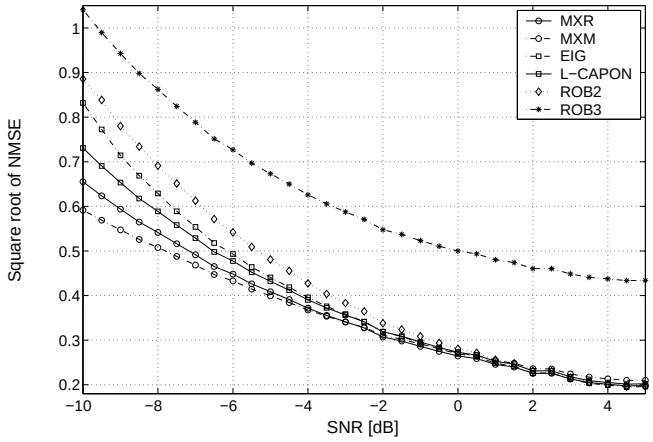

Fig. 1.8 Square-root of the normalized MSE as a function of SNR when estimating a complex sinewave with known $\mathbf{a}$ and with DOA=30^o, using the MXR, MXM, EIG, L-CAPON, ROB2, and ROB3 beamformers.

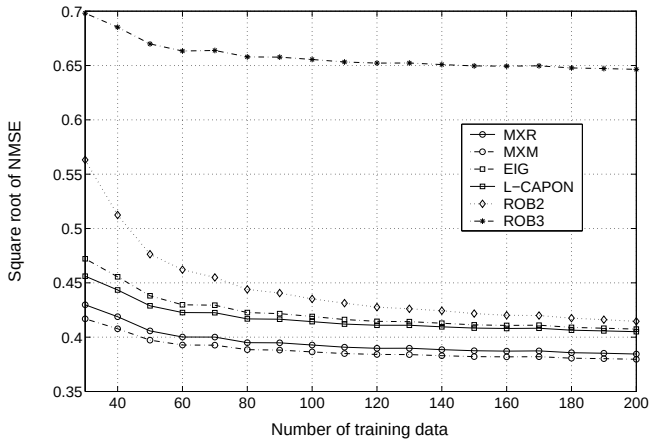

Fig. 1.9 Square-root of the normalized MSE as a function of the number of training snapshots when estimating a complex sinewave with known $\mathbf{a}$ and DOA=30^o, using the MXR, MXM, EIG, L-CAPON, ROB2, and ROB3 beamformers.

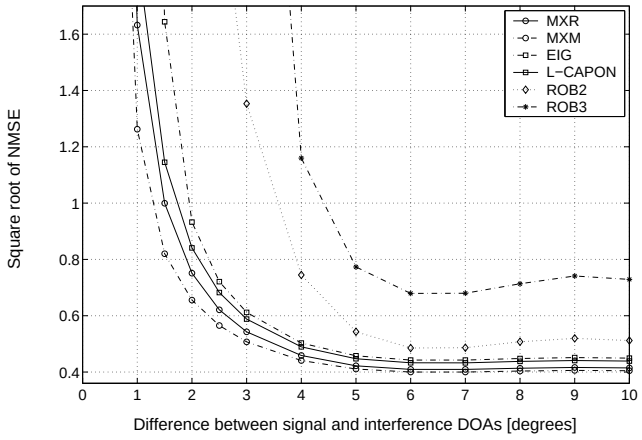

Fig. 1.10 Square-root of the normalized MSE as a function of the difference between the signal and interference DOAs when estimating a complex sinewave with known $\mathbf{a}$ and DOA=30^o, using the MXR, MXM, EIG, L-CAPON, ROB2, and ROB3 beamformers.

however the minimax methods still outperform the alternative methods for all SIR values shown.

We next repeat the simulations for $s(t)$ chosen to be a zero-mean complex Gaussian random signal, temporally white. The square-root of the NMSE as a function of the SNR, the number of training data, the difference between signal and interference DOAs, and the SIR is depicted in Figs. 1.12, 1.13, 1.14 and 1.15, respectively. It can be seen in Fig. 1.12 that the MXM has the best performance for SNR values between -10 to -4 dB and the MXR has the best performance for SNR values between -3 to 1.5 dB. The performance in Figs. 1.13, 1.14 and 1.15 is similar to the case of a deterministic sinewave.

Example 2 - Steering vector with signal DOA uncertainties: We illustrate the robustness of the algorithms for random steering vectors developed in Section 1.4 against a mismatch in the assumed signal DOA. Specifically, the DOA of the signal is given by a Gaussian random variable with mean equal to 30^o and standard deviation equal to 1 (about $\pm 3^o$). This DOA was independently drawn in each simulation run. To estimate the signal in this case, we implemented the MXRR, MXMR, LSR, and

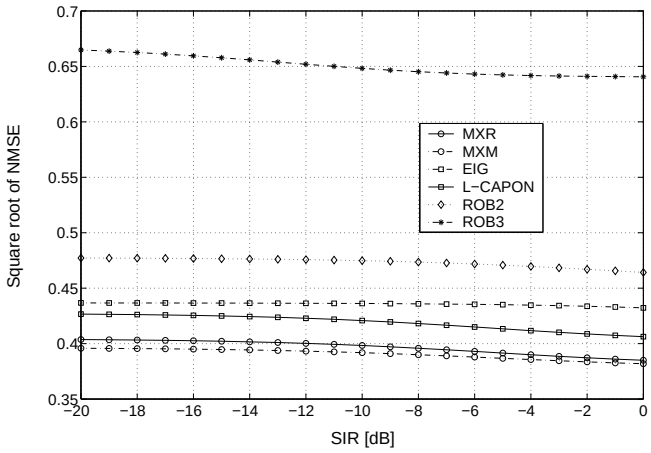

Fig. 1.11 Square-root of the normalized MSE as a function of SIR when estimating a complex sinewave with known $\mathbf{a}$ and DOA=30^o, using the MXR, MXM, EIG, L-CAPON, ROB2, and ROB3 beamformers.

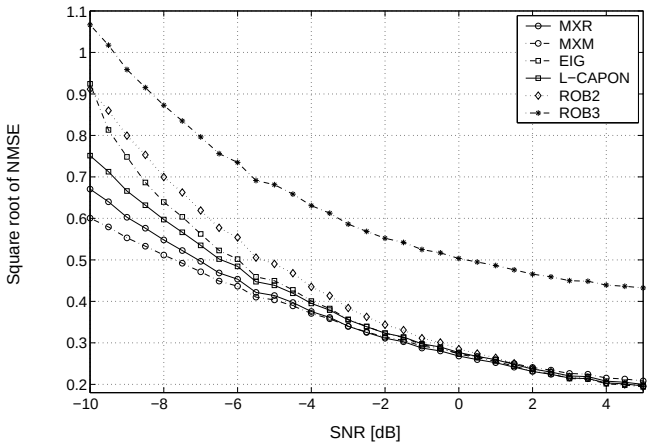

Fig. 1.12 Square-root of the normalized MSE as a function of SNR when estimating a zero-mean complex Gaussian random signal with known $\mathbf{a}$ and DOA=30^o, using the MXR, MXM, EIG, L-CAPON, ROB2, and ROB3 beamformers.

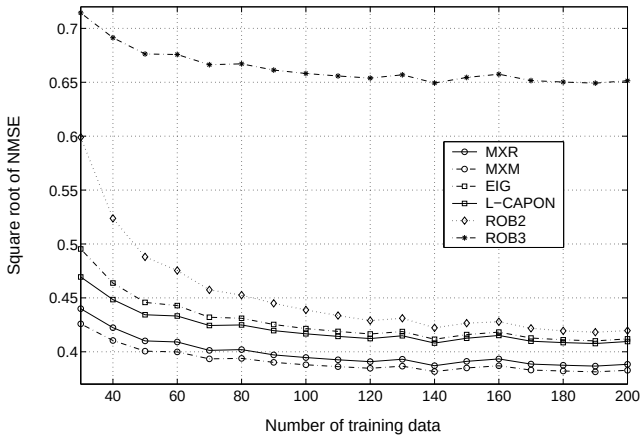

Fig. 1.13 Square-root of the normalized MSE as a function of the number of training snapshots when estimating a zero-mean complex Gaussian random signal with known $\mathbf{a}$ and DOA=30^o, using the MXR, MXM, EIG, L-CAPON, ROB2, and ROB3 beamformers.

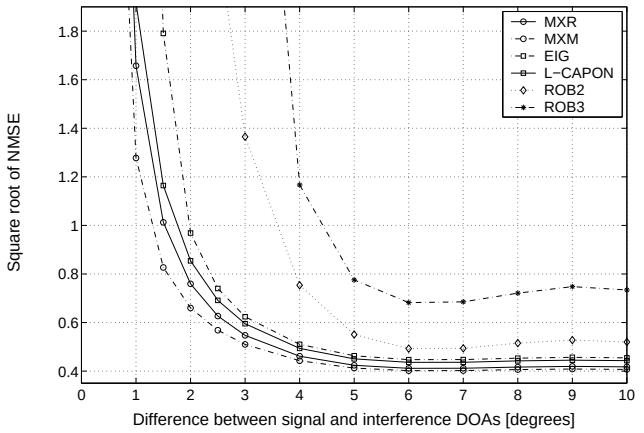

Fig. 1.14 Square-root of the normalized MSE as a function of the difference between the signal and interference DOAs when estimating a zero-mean complex Gaussian random signal with known $\mathbf{a}$ and DOA=30^o, using the MXR, MXM, EIG, L-CAPON, ROB2, and ROB3 beamformers.

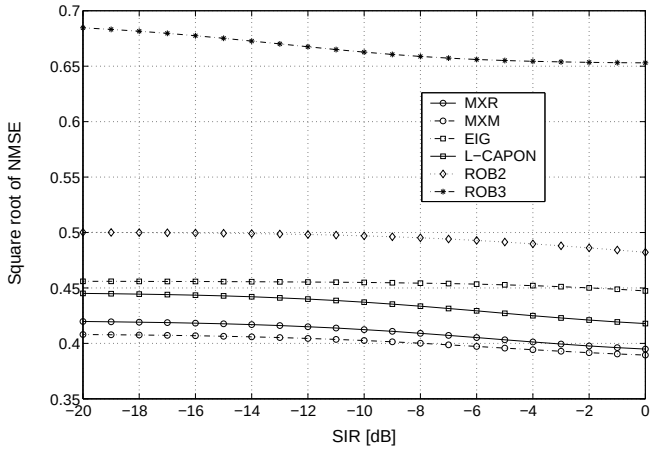

Fig. 1.15 Square-root of the normalized MSE as a function of SIR when estimating a zero-mean complex Gaussian random signal with known $\mathbf{a}$ and DOA=30^o, using the MXR, MXM, EIG, L-CAPON, ROB2, and ROB3 beamformers.

Beamformer	$\mathbf{R}$	$\mathbf{w}_c$	β
MXMR	$\mathbf{R}_{dl}$, ξ = 10	L-Capon, ξ = 10	9
MXRR	$\mathbf{R}_{dl}$, ξ = 10	L-Capon, ξ = 10	4
LSR	$\mathbf{R}_{dl}$, ξ = 10	—	—
PEIG	$\mathbf{R}_{dl}$, ξ = 10	—	—

Table 1.3 Specification of the beamformers used in Example 2.

PEIG beamformers with parameters given by Table 1.3, for two choices of $\mathbf{m}$ and $\mathbf{C}$: the true values (estimated from 2000 realizations of the steering vector), and ad-hoc values. The ad-hoc value of $\mathbf{m}$ was chosen as the steering vector with DOA = 30^o and for the covariance we chose $\mathbf{C} = \nu\mathbf{I}$ with $\nu = \epsilon/M$, where ϵ = 3.5 is the norm value of the steering vector error (size of the uncertainty) used by ROB2 [15] and $M = 20$ is the number of sensors. For the ROB1, ROB2 and ROB3 beamformers $\mathbf{a}$ is the steering vector for a signal DOA = 30^o.

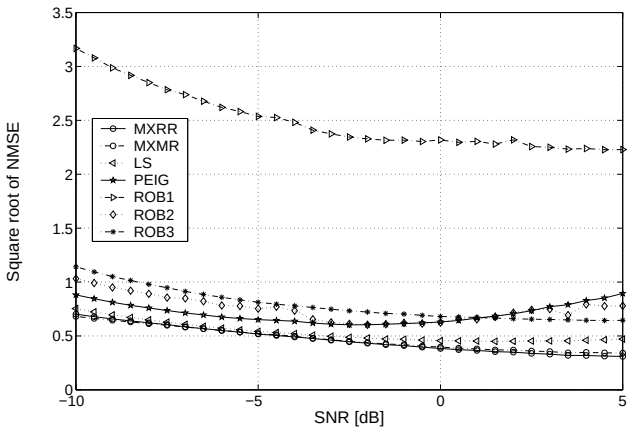

Fig. 1.16 Square-root of the normalized MSE as a function of SNR when estimating a complex sinewave with random DOA and known $\mathbf{m}$ and $\mathbf{C}$, using the MXRR, MXMR, LSR, PEIG, ROB1, ROB2, and ROB3 beamformers.

In Figs. 1.16 and 1.17 we depict the NMSE when estimating a complex sinewave as a function of SNR, using the MXRR, MXMR, LSR, PEIG, ROB1, ROB2, and ROB3 beamformers, with known and ad-hoc values of $\mathbf{m}$ and $\mathbf{C}$, respectively. As can be seen from the figures, the MXMR, MXRR and LSR methods perform better than all the other methods both in the case of known $\mathbf{m}$ and $\mathbf{C}$ and when ad-hoc values are chosen. Although the performance of the MXMR, MXRR and LSR methods deteriorates when the ad-hoc values are used in place of the true values, the difference in performance is minor. A surprising observation from the figures is that the standard PEIG method performs better, in terms of NMSE, than ROB1, ROB2 and ROB3. We also note that the LSR method, which does not use any additional parameters such as magnitude bounds, outperforms all previously proposed methods.

The performance of the MXRR, MXMR, LSR, PEIG, ROB1, ROB2, and ROB3 beamformers, assuming known $\mathbf{m}$ and $\mathbf{C}$, as a function of the number of training snapshots and the difference between the signal and interference DOAs is illustrated in Figs. 1.18 and 1.19, respectively. The NMSE of all methods remain almost constant for differences above 6^o. Below this range all the beamformers decrease their performance significantly, except for the PEIG and LSR beamformers. As in

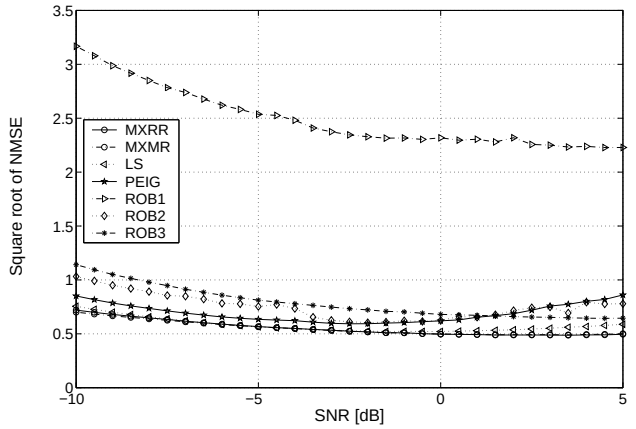

Fig. 1.17 Square-root of the normalized MSE as a function of SNR when estimating a complex sinewave with random DOA and ad-hoc values of $\mathbf{m}$ and $\mathbf{C}$, using the MXRR, MXMR, LSR, PEIG, ROB1, ROB2, and ROB3 beamformers.

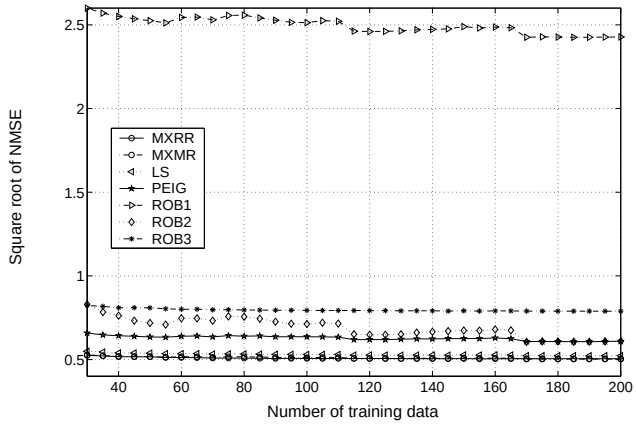

Fig. 1.18 Square-root of the normalized MSE as a function of the number of training snapshots when estimating a complex sinewave with random DOA and known $\mathbf{m}$ and $\mathbf{C}$ using the MXRR, MXMR, LSR, PEIG, ROB1, ROB2, and ROB3 beamformers.

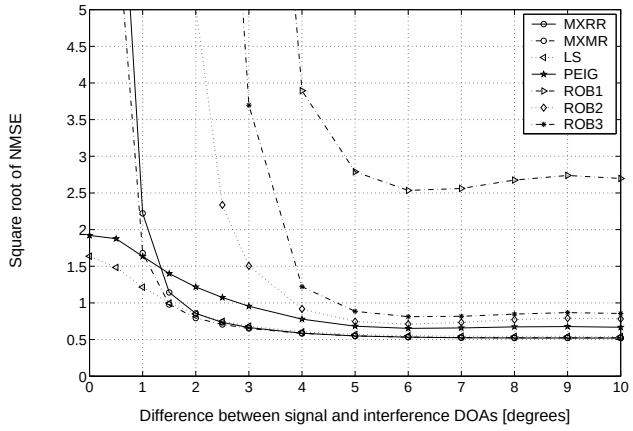

Fig. 1.19 Square-root of the normalized MSE as a function of the difference between the signal and interference DOAs when estimating a complex sinewave with random DOA and known $\mathbf{m}$ and $\mathbf{C}$ using the MXRR, MXMR, LSR, PEIG, ROB1, ROB2, and ROB3 beamformers.

the case of deterministic $\mathbf{a}$, the performance of all the methods improves slightly as a function of negative SIR, and is therefore not shown.

In Figs. 1.20 and 1.21 we plot the NMSE as a function of the SNR when estimating a complex Gaussian random signal, temporally white, with random DOA, as a function of SNR. As can be seen by comparing these figures with Figs. 1.16 and 1.17, the performance of all of the methods is similar to the case of a deterministic sinewave. The NMSE as a function of the number of training snapshots and the difference between the signal and interference DOAs is also similar to the deterministic sinewave case, and therefore not shown.

1.7 SUMMARY

We considered the problem of designing linear beamformers to estimate a source signal $s(t)$ from sensor array observations, where the goal is to obtain an estimate $\hat{s}(t)$ that is close to $s(t)$. Although standard beamforming approaches are aimed at maximizing the SINR, maximizing SINR does not necessarily guarantee a small

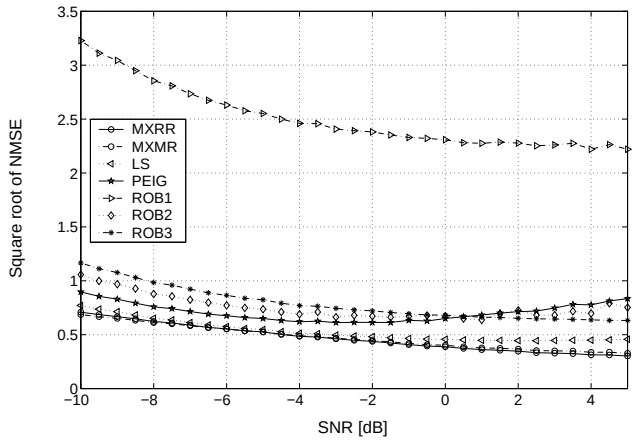

Fig. 1.20 Square-root of the normalized MSE as a function of SNR when estimating a zero-mean complex Gaussian random signal with random DOA and known $\mathbf{m}$ and $\mathbf{C}$, using the MXRR, MXMR, LSR, PEIG, ROB1, ROB2, and ROB3 beamformers.

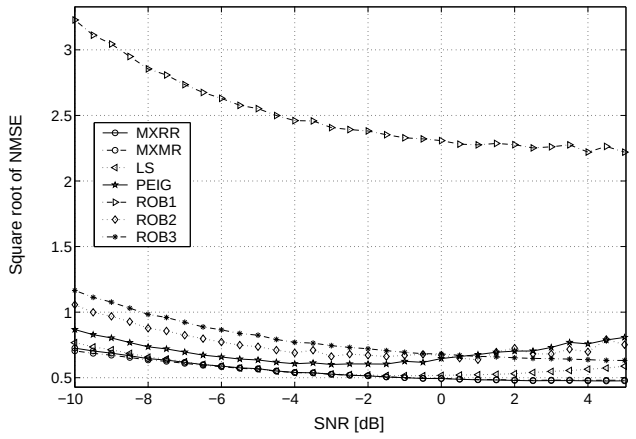

Fig. 1.21 Square-root of the normalized MSE as a function of SNR when estimating a zero-mean complex Gaussian random signal with random DOA and ad-hoc values of $\mathbf{m}$ and $\mathbf{C}$, using the MXRR, MXMR, LSR, PEIG, ROB1, ROB2, and ROB3 beamformers.

MSE, hence on average a signal estimate maximizing the SINR can be far from $s(t)$. To ensure that $\hat{s}(t)$ is close to $s(t)$, we proposed 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. Instead, we suggested beamforming methods that minimize a worst-case measure of MSE assuming known and random steering vectors with known second-order statistics. We first considered a minimax MSE beamformer that minimizes the worst-case MSE. We then considered a minimax regret beamformer that minimizes the worst-case difference between the MSE using a beamformer ignorant of $s(t)$ and the smallest possible MSE attainable with a beamformer that knows $s(t)$. As we showed, even if $s(t)$ is known, we cannot achieve a zero MSE with a linear estimator. In the case of a random steering vector we also proposed a least-squares beamformer that does not require bounds on the signal magnitude.

In the numerical examples, we clearly illustrated the advantages of our methods in terms of the MSE. For both known and random steering vectors, the minimax beamformers consistently have the best performance, particularly for negative SNR values. Quite surprisingly, it was observed that the least-squares beamformer, which does not require bounds on the signal magnitude, performs better than the recently proposed robust methods in the case of random DOAs. It was also observed that for a small difference between signal and interference directions of arrival, all our methods show better performance. The performance of our methods was similar when the signal was chosen as a deterministic sinewave or a zero-mean complex Gaussian random signal.

Acknowledgments

The authors are grateful to Mr. Patricio S. La Rosa for conducting the numerical examples and for carefully reviewing several earlier versions of this chapter. The work of A. Nehorai was supported by the Air Force Office of Scientific Research Grant F49620-02-1-0339, and the National Science Foundation Grants CCR-0105334 and CCR-0330342.

References

1. R. Mozingo and T. Miller, *Introduction to Adaptive Arrays*. Wiley and Sons, New York, 1980.

2. J. E. Hudson, *Adaptive Array Principles*. Peter Peregrinus Ltd, London, 1981.

3. S. Haykin, *Array Signal Processing*. Prentice-Hall, Englewood Cliffs, New Jersey, 1985.

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

5. S. Haykin and A. Steinhardt, *Adaptive Radar Detection and Estimation*. Wiley, New York, 1992.

6. 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.

7. M. Hawkes and A. Nehorai, "Acoustic vector-sensor beamforming and capon direction estimation," *IEEE Trans. on Signal Proc.*, vol. SP–46, pp. 2291–2304, Sept. 1998.

8. H. L. V. Trees, *Optimal Array Processing (Detection, Estimation, and Modulation Theory, Part IV)*. Wiley-Interscience, 2002.

9. 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.

10. 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.

11. 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.

12. N. L. Owsley, "Enhanced minimum variance beamforming," in *Underwater Acoustic Data Processing* (Y. T. Chan, ed.), pp. 285–292, Kluwer, 1989.

13. K. L. Bell, Y. Ephraim, and H. L. V. Trees, "A bayesian approach to robust adaptive beamforming," *IEEE Trans. Signal Processing*, vol. 48, pp. 386–398, Feb. 2000.

14. 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.

15. 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.

16. 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.

17. C. J. Lam and A. C. Singer, "Performance analysis of the bayesian beamforming," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, vol. II, pp. 197–200, 2004.

18. C. D. Richmond, "The capon-mvdr algorithm threshold region performance prediction and its probability of resolution," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, p. (submitted), 2004.

19. O. Besson and F. Vincent, "Performance analysis of the bayesian beamforming," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, vol. II, pp. 169–172, 2004.

20. Y. C. Eldar, A. Ben-Tal, and A. Nemirovski, "Robust mean-squared error estimation in the presence of model uncertainties," *IEEE Trans. Signal Processing*, to appear.

21. Y. C. Eldar, A. Ben-Tal, and A. Nemirovski, "Linear minimax regret estimation of deterministic parameters with bounded data uncertainties," *IEEE Trans. Signal Processing*, vol. 52, pp. 2177–2188, Aug. 2004.

22. J. Yang and A. Swindlehurst, "Signal copy with array calibration errors," *Signals, Systems and Computers, 1993. Conference Record of The Twenty-Seventh Asilomar Conference on , 1-3 Nov.*, vol. 2, pp. 405–413, 1993.

23. Y. C. Eldar, "Robust estimation in linear models with a random model matrix," submitted to *IEEE Trans. Signal Processing*.

24. L. Seymour, C. Cowan, and P. Grant, "Bearing estimation in the presence of sensors positioning errors," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, pp. 2264–2267, 1987.

25. B. Friendlander and A. Weiss, "Direction finding in the presence of mutual coupling," *IEEE Trans. Antennas and Propagation*, vol. AP-39, pp. 273–284, 1991.

26. A. B. Gershman, V. I. Turchin, and V. A. Zverev, "Experimental results of localization of moving underwater signal by adaptive beamforming," *IEEE Trans. Signal Processing*, vol. 43, pp. 2249–2257, Oct. 1995.

27. E. Y. Gorodetskaya, A. I. Malekhanov, A. G. Sazontov, and N. K. Vdovicheva, "Deep-water acoustic coherence at long ranges: Theoretical prediction and effects on large-array signal processing," *IEEE J. Ocean. Eng.*, vol. 24, pp. 156–171, Apr. 1999.

28. H. Cox, "Effects of random steering vector errors in the applebaum array," *J. Acoust. Soc. Amer.*, vol. 54, pp. 771–785, Sept. 1973.

29. D. R. Morgan and T. M. Smith, "Coherence effects on the detection performance of quadratic array processors with application to largearray matched-field beamforming," *J. Acoust. Soc. Amer.*, vol. 87, pp. 737–747, Feb. 1988.

30. A. Paulraj and T. Kailath, "Direction of arrival estimation by eigenstructure methods with imperfect spatial coherence ofwave fronts," *J. Acoust. Soc. Amer.*, vol. 83, pp. 1034–1040, Mar. 1988.

31. A. B. Gershman, C. F. Mecklenbräuker, and J. F. Böhme, "Matrix fitting approach to direction of arrival estimation with imperfect spatial coherence of wavefronts," *IEEE Trans. Signal Processing*, vol. 45, pp. 1894–1899, July 1997.

32. M. Wax and Y. Anu, "Performance analysis of the minimum variance beamformer in the presence of steering vector errors," *IEEE Trans. Signal Processing*, vol. 44, pp. 938–947, Apr. 1996.

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

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

35. A. B. Gershman, "Robust adaptive beamforming in sensor arrays," *Int. J. Electron. Commun.*, vol. 53, pp. 305–324, Dec. 1999.

36. Y. C. Eldar and A. Nehorai, "Competitive mean-squared error beamforming," in *Proc. 12th Annu. Workshop Adaptive Sensor Array Processing, Lincoln Laboratory, MIT, Lexington, MA*, Mar. 2004.

37. Y. C. Eldar, "Robust competitive estimation with signal and noise covariance uncertainties," submitted to *IEEE Trans. Inform. Theory*.

38. C. L. Zahm, "Effects of errors in the direction of incidence on the performance of an adaptive array," *Proc. IEEE*, vol. 60, pp. 1068–1069, Aug. 1972.

39. R. T. Compton, "Effects of random steering vector errors in the applebaum array," *IEEE Trans. Aerosp. Electron. Syst.*, vol. AES-18, pp. 392–400, Sept. 1982.

40. L. C. Godara, "Error analysis of the optimal antenna array processors," *IEEE Trans. Aerosp. Electron. Syst.*, vol. AES-22, pp. 395–409, July 1986.

41. Y. C. Eldar and A. Nehorai, "Uniformly robust mean-squared error beamforming," in *Proc. 3rd IEEE Sensor Array and Multichannel Signal Processing Workshop, Barcelona, Spain*, Jul. 2004.

Index