MINIMUM MEAN-SQUARED ERROR COVARIANCE SHAPING

Yonina C. Eldar

Department of Electrical Engineering
Technion–Israel Institute of Technology
Haifa, 32000, Israel

ABSTRACT

This paper develops and explores applications of a linear shaping transformation that minimizes the mean squared error (MSE) between the original and shaped data, *i.e.,* that results in an output vector with the desired covariance that is as close as possible to the input, in an MSE sense. Three applications of minimum MSE shaping are considered, specifically matched filter detection, multiuser detection and linear least-squares parameter estimation.

1. INTRODUCTION

Covariance shaping arises in a variety of contexts in which it is useful to shape the covariance of a data vector either prior to subsequent processing, or to control the spectral shape after processing.

It is well known that the linear transformation that shapes the covariance of a data vector is not unique. While in some applications certain conditions might be imposed on the transformation such as causality or symmetry, there have been no general assertions of optimality for various choices of a shaping transformation.

Shaping the covariance of a data vector introduces distortion to the values of the data relative to the unshaped data. In certain applications, it is desirable to shape the data while minimizing this distortion, *i.e.,* to choose the shaping transformation in an optimal sense. In [1, 2], an optimal whitening transformation was developed and used to improve the detection performance in a multi-signature system. The whitening transformation, referred to as the minimum mean squared error (MMSE) whitening transformation, is designed to minimize the mean squared error (MSE) between the original and whitened data.

In this paper we extend the concept of MMSE whitening to include other forms of covariance shaping. Specifically, we develop and explore applications of a linear shaping transformation that minimizes a possibly weighted MSE between the original and shaped data, *i.e.,* that results in an output with the desired covariance that is as close as possible to the input in MSE. We refer to such a transformation as an MMSE shaping transformation.

In Section 2 we derive the linear MMSE shaping transformation. For simplicity, we restrict our attention to the case in which the original and output covariance matrices as well as the weighting matrix are all positive definite; the more general case of noninvertible matrices is treated in [3]. Section 3 discusses several applications of linear MMSE shaping. The first is a summary of the use of MMSE whitening in the context of MF detection. This application is developed and explored in considerably more detail in [1, 3]. The second application described is to multiuser detection in multiple access channels. The third application is to estimation of a set of unknown deterministic parameters in a linear model.

2. OPTIMAL COVARIANCE SHAPING

Let $\mathbf{a} \in \mathbb{C}^m$ denote a zero-mean[1] random vector with positive definite covariance matrix $\mathbf{C}_a$. We wish to shape the covariance of $\mathbf{a}$ using a shaping transformation $\mathbf{T}$ to obtain the random vector $\mathbf{b} = \mathbf{Ta}$, where the covariance matrix of $\mathbf{b}$ is given by $\mathbf{C}_b = c^2\mathbf{R}$ for some $c > 0$ and positive definite covariance matrix $\mathbf{R}$. Thus we seek a transformation $\mathbf{T}$ such that

$$\mathbf{C}_b = \mathbf{T}\mathbf{C}_a\mathbf{T}^* = c^2\mathbf{R}, \tag{1}$$

for some $c > 0$. We refer to any $\mathbf{T}$ satisfying (1) as a covariance shaping transformation.

Given a covariance matrix $\mathbf{C}_a$, there are many ways to choose a transformation $\mathbf{T}$ satisfying (1). Although there are an unlimited number of covariance shaping transformations, no general assertion of optimality is known for the output $\mathbf{b} = \mathbf{Ta}$ of these different transformations. In particular, the vector $\mathbf{b}$ may not be "close" to the input vector $\mathbf{a}$. If $\mathbf{b}$ undergoes some noninvertible processing, or is used as an estimator of some unknown parameters represented by the data $\mathbf{a}$, then we may wish to choose the shaping transformation in a way that $\mathbf{b}$ is close to $\mathbf{a}$ in some sense. For example, if $\mathbf{b}$ is the input to a detector, then we may wish to shape the covariance of $\mathbf{a}$ prior to detection, but at the same time minimize the distortion to $\mathbf{a}$ by choosing $\mathbf{T}$ so that $\mathbf{b}$ is close to $\mathbf{a}$. Applications of this type are explored in Section 3.

We therefore propose a shaping transformation that is optimal in the sense that it results in a random vector $\mathbf{b}$ that is as close as possible to $\mathbf{a}$ in MSE. Specifically, among all possible covariance shaping transformations we seek the one that minimizes the MSE

$$\begin{aligned}\varepsilon_{\mathrm{MSE}} &= E\left((\mathbf{a}-\mathbf{b})^*(\mathbf{a}-\mathbf{b})\right)\\ &= \mathrm{Tr}\left(E((\mathbf{a}-\mathbf{b})(\mathbf{a}-\mathbf{b})^*)\right)\\ &= \mathrm{Tr}(\mathbf{C}_a) + \mathrm{Tr}(\mathbf{T}\mathbf{C}_a\mathbf{T}^*) - \mathrm{Tr}(\mathbf{C}_a(\mathbf{T}^*+\mathbf{T})),\end{aligned} \tag{2}$$

subject to (1). We may wish to constraint the constant c in (1), or may choose c to minimize the MSE.

Our approach to determining the shaping transformation that minimizes (2) is to show that this problem can be interpreted as an MMSE whitening problem, and then rely on results obtained in that context [1, 3].

Let $\tilde{\mathbf{a}} = \mathbf{R}^{1/2}\mathbf{a}$, $\widetilde{\mathbf{T}} = \mathbf{R}^{-1/2}\mathbf{T}\mathbf{R}^{-1/2}$, and

$$\tilde{\varepsilon}_{\mathrm{MSE}} = E\left((\tilde{\mathbf{a}} - \widetilde{\mathbf{T}}\tilde{\mathbf{a}})^*(\tilde{\mathbf{a}} - \widetilde{\mathbf{T}}\tilde{\mathbf{a}})\right). \tag{3}$$

We may then express $\varepsilon_{\mathrm{MSE}}$ as

$$\varepsilon_{\mathrm{MSE}} = \tilde{\varepsilon}_{\mathrm{MSE}} + \mathrm{Tr}\left((\mathbf{I}-\mathbf{R})(\mathbf{C}_a - c^2\mathbf{I})\right), \tag{4}$$

[1] If the mean $E(\mathbf{a})$ is not zero, then we can always define $\mathbf{a}' = \mathbf{a} - E(\mathbf{a})$ so that the results hold for $\mathbf{a}'$.

where we used the fact that $\mathbf{T}$ satisfies (1). Furthermore, with $\mathbf{C}_{\tilde{a}} = \mathbf{R}^{1/2}\mathbf{C}_a\mathbf{R}^{1/2}$ denoting the covariance of $\tilde{\mathbf{a}}$, we have that

$$\widetilde{\mathbf{T}}\mathbf{C}_{\tilde{a}}\widetilde{\mathbf{T}}^* = \mathbf{R}^{-1/2}\mathbf{T}\mathbf{C}_a\mathbf{T}^*\mathbf{R}^{-1/2} = c^2\mathbf{I}. \tag{5}$$

Therefore, for fixed c, finding $\mathbf{T}$ to minimize $\varepsilon_{\mathrm{MSE}}$ subject to (1) is equivalent to finding the transformation $\widetilde{\mathbf{T}} = \mathbf{R}^{-1/2}\mathbf{T}\mathbf{R}^{-1/2}$ to minimize $\tilde{\varepsilon}_{\mathrm{MSE}}$ of (3) subject to (5). Since from (5) the covariance of $\widetilde{\mathbf{T}}\tilde{\mathbf{a}}$ is proportional to $\mathbf{I}$, $\widetilde{\mathbf{T}}$ is a whitening transformation. From all possible whitening transformations we seek the one that minimizes the MSE (3). This problem is therefore equivalent to the MMSE whitening problem considered in [1, 3], the solution of which is incorporated in the following theorem:

Theorem 1 (MMSE whitening [3]) *Let $\mathbf{z} \in \mathbb{C}^m$ be a zero-mean random vector with positive definite covariance $\mathbf{C}_z$. Among all possible whitening transformations, let $\widehat{\mathbf{T}}$ denote the whitening transformation that minimizes the MSE,*

$$\varepsilon_{\mathrm{MSE}} = E\left((\mathbf{z}-\mathbf{h})^*(\mathbf{z}-\mathbf{h})\right),$$

between the input $\mathbf{z}$ and the output $\mathbf{h} = \mathbf{T}\mathbf{z}$ with covariance $\mathbf{C}_h = c^2\mathbf{I}$. Then

$$\widehat{\mathbf{T}} = c\mathbf{C}_z^{-1/2}.$$

From Theorem 1 the optimal value of $\widetilde{\mathbf{T}}$, denoted $\widehat{\widetilde{\mathbf{T}}}$, is

$$\widehat{\widetilde{\mathbf{T}}} = c\mathbf{C}_{\tilde{a}}^{-1/2} = c(\mathbf{R}^{1/2}\mathbf{C}_a\mathbf{R}^{1/2})^{-1/2}. \tag{6}$$

The optimal shaping transformation, denoted $\widehat{\mathbf{T}}$, is then equal to

$$\widehat{\mathbf{T}} = c\mathbf{R}^{1/2}(\mathbf{R}^{1/2}\mathbf{C}_a\mathbf{R}^{1/2})^{-1/2}\mathbf{R}^{1/2} = c(\mathbf{R}\mathbf{C}_a)^{-1/2}\mathbf{R}, \tag{7}$$

where we used the fact that for any invertible matrices $\mathbf{A}$ and $\mathbf{B}$, $\mathbf{A}\mathbf{B}^{-1/2}\mathbf{A}^{-1} = \left(\mathbf{A}\mathbf{B}\mathbf{A}^{-1}\right)^{-1/2}$ [3].

We may further wish to choose c such that (2) is minimized. Substituting $\mathbf{b} = \widehat{\mathbf{T}}\mathbf{a}$ into (2), c is chosen to minimize

$$\varepsilon_{\mathrm{MSE}} = \mathrm{Tr}(\mathbf{C}_a) + c^2\mathrm{Tr}(\mathbf{R}) - 2c\mathrm{Tr}\left((\mathbf{R}\mathbf{C}_a)^{1/2}\right). \tag{8}$$

Differentiating with respect to c and equating to 0, the optimal value of c, denoted $\hat{c}$, is given by

$$\hat{c} = \frac{\mathrm{Tr}\left((\mathbf{R}\mathbf{C}_a)^{1/2}\right)}{\mathrm{Tr}(\mathbf{R})}. \tag{9}$$

We may also consider a weighted MMSE covariance shaping problem in which we seek a transformation $\mathbf{T}$ such that $\mathbf{b} = \mathbf{T}\mathbf{a}$ has covariance $\mathbf{C}_b = c^2\mathbf{R}$ for some $c > 0$ and positive definite covariance matrix $\mathbf{R}$, and such that

$$\varepsilon_{\mathrm{MSE}}^w = E\left((\mathbf{a}-\mathbf{b})^*\mathbf{W}(\mathbf{a}-\mathbf{b})\right), \tag{10}$$

is minimized, where $\mathbf{W}$ is some positive definite Hermitian weighting matrix.

To determine the weighted MMSE covariance shaping transformation we note that

$$E\left((\mathbf{a}-\mathbf{b})^*\mathbf{W}(\mathbf{a}-\mathbf{b})\right) = E\left((\bar{\mathbf{a}}-\bar{\mathbf{b}})^*(\bar{\mathbf{a}}-\bar{\mathbf{b}})\right), \tag{11}$$

where $\bar{\mathbf{a}} = \mathbf{W}^{1/2}\mathbf{a}$ and $\bar{\mathbf{b}} = \mathbf{W}^{1/2}\mathbf{b}$. Thus we may first seek the transformation $\widehat{\overline{\mathbf{T}}}$ that minimizes the MSE between the random vector $\bar{\mathbf{a}}$ with covariance $\mathbf{C}_{\bar{a}} = \mathbf{W}^{1/2}\mathbf{C}_a\mathbf{W}^{1/2}$, and the random vector $\bar{\mathbf{b}} = \overline{\mathbf{T}}\bar{\mathbf{a}}$ with covariance $\mathbf{C}_{\bar{b}} = \overline{\mathbf{T}}\mathbf{C}_{\bar{a}}\overline{\mathbf{T}} = c^2\overline{\mathbf{R}}$, where $\overline{\mathbf{R}} = \mathbf{W}^{1/2}\mathbf{R}\mathbf{W}^{1/2}$. From (7) and (9) we have,

$$\begin{aligned}\widehat{\overline{\mathbf{T}}} &= \beta(\overline{\mathbf{R}}\mathbf{C}_{\bar{a}})^{-1/2}\overline{\mathbf{R}} \\ &= \beta\mathbf{W}^{1/2}(\mathbf{R}\mathbf{W}\mathbf{C}_a\mathbf{W})^{-1/2}\mathbf{R}\mathbf{W}^{1/2},\end{aligned} \tag{12}$$

where $\beta = c$ if c is fixed and $\beta = \hat{c}$ if c is chosen to minimize the MSE with

$$\hat{c} = \frac{\mathrm{Tr}\left((\mathbf{R}\mathbf{W}\mathbf{C}_a\mathbf{W})^{1/2}\right)}{\mathrm{Tr}(\mathbf{R}\mathbf{W})}. \tag{13}$$

The weighted MMSE covariance shaping transformation is then

$$\widehat{\mathbf{T}} = \mathbf{W}^{-1/2}\widehat{\overline{\mathbf{T}}}\mathbf{W}^{1/2} = \beta(\mathbf{R}\mathbf{W}\mathbf{C}_a\mathbf{W})^{-1/2}\mathbf{R}\mathbf{W}. \tag{14}$$

In the applications in Section 3 we use a weighted MMSE covariance shaping transformation with $\mathbf{W} = \mathbf{C}_a^{-1}$, in which case $\widehat{\mathbf{T}} = \beta(\mathbf{R}{\mathbf{C}_a}^{-1})^{1/2}$. With this choice of weighting, if the elements of $\mathbf{a}$ are uncorrelated, then the weight given to each of the errors $a_i - b_i$ is inversely proportional to the variance of a_i, thereby emphasizing the contributions of the elements of $\mathbf{a}$ that tend to be more reliable. Note that this choice of weighting matrix is also reminiscent of the Gauss-Markov weighting in least-squares estimation [4].

We summarize our results in the following theorem:

Theorem 2 (MMSE shaping) *Let $\mathbf{a} \in \mathbb{C}^m$ be a zero-mean random vector with positive definite covariance $\mathbf{C}_a$. Among all possible shaping transformations, let $\widehat{\mathbf{T}}$ denote the shaping transformation that minimizes the weighted MSE given by (10) between the input $\mathbf{a}$ and the output $\mathbf{b} = \mathbf{T}\mathbf{a}$ with covariance $\mathbf{C}_b = c^2\mathbf{R}$ for some positive definite covariance $\mathbf{R}$. Then*

$$\widehat{\mathbf{T}} = \beta(\mathbf{R}\mathbf{W}\mathbf{C}_a\mathbf{W})^{-1/2}\mathbf{R}\mathbf{W},$$

where

1. *if c is specified then $\beta = c$;*
2. *if c is chosen to minimize (10) then $\beta = \hat{c}$ given by (13).*

3. APPLICATIONS

In this section we consider applications of MMSE shaping to matched filter (MF) detection, multiuser detection, and least-squares (LS) parameter estimation. In all the applications, we assume that a weighted MMSE transformation is used with $\mathbf{W} = \mathbf{C}_a^{-1}$.

3.1. Matched Filter Detection

A generic and well studied problem is that of detecting in the presence of additive noise, which one from a set of known signals has been received. When the additive noise is white and Gaussian, the receiver which maximizes the probability of correct detection P_D is a MF demodulator comprised of a bank of correlators with correlating signals equal to the transmitted set, followed by a detector which chooses as the detected signal the one for which the output of the correlator is maximum [5]. If the noise is not Gaussian, then the MF detector does not necessarily maximize P_D. The problem then is to design a simple linear receiver that does not depend on the noise distribution, and that leads to improved performance over MF detection for non-Gaussian noise.

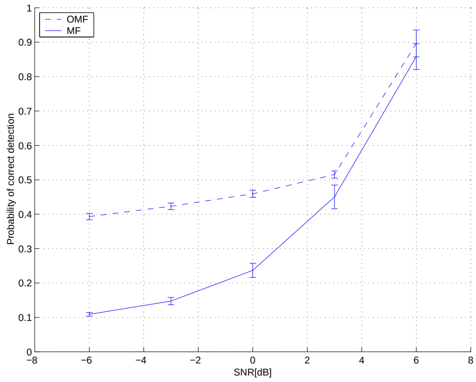

Fig. 1. Mean P_D using the OMF and MF detectors in Beta-distributed noise, as a function of SNR. The vertical lines indicate the standard deviation of the corresponding P_D.

To try and improve the performance over the MF receiver, we propose optimally shaping the output vector of the MF prior to detection using an MMSE shaping transformation, and basing the detection on the shaped vector. In the special case in which the shaping transformation is chosen as a whitening transformation, the resulting receiver is referred to as the orthogonal MF (OMF) receiver [1]. Simulations presented in [1, 3] show that when the noise is non-Gaussian this approach can in fact lead to improved performance over conventional MF detection in many cases.

In Fig. 1 we plot the mean and standard deviation of P_D for the OMF and MF detectors as a function of SNR for transmitted constellations of 13 signals, in Beta-distributed noise with parameters $a = b = 0.1$. The signals have dimension 13 and the samples of the signals are mutually independent zero-mean Gaussian random variables with variance $1/\sqrt{13}$, scaled to have norm 1. The results in the figure were obtained by generating 500 realizations of signals. For each signal realization, we determined P_D by recording the number of successful detections over 500 noise realizations.

Preliminary simulations demonstrate that in a variety of cases choosing the output covariance to be non-diagonal can further improve the performance of the modified receiver over the MF receiver. An interesting direction for future research is to design an optimality criterion for choosing the desired output covariance based on knowledge of the transmitted signals.

3.2. Multiuser Detection

In a manner similar to the MF detection problem, MMSE shaping can also be used to improve the performance in a multiuser communication system. In this context, the background noise is inherently non-Gaussian since it is comprised of additive white noise and multiple access interference (MAI) from other users.

The received signal $\mathbf{y}$ in a CDMA system is modelled as $\mathbf{y} = \mathbf{SAx} + \mathbf{w}$, where $\mathbf{S}$ is the matrix of columns $\mathbf{s}_i$ with $\mathbf{s}_i$ being the signature vector of the ith user, $\mathbf{A}$ is the diagonal matrix with diagonal elements $A_i > 0$ with A_i being the received amplitude of the ith user's signal, $\mathbf{x}$ is the data vector with components $x_i \in \{1, -1\}$ with x_i being the ith user's transmitted symbol, and $\mathbf{w}$ is a noise vector whose elements are independent $\mathcal{CN}(0, \sigma^2)$. We assume for simplicity that the vectors $\mathbf{s}_i$ are linearly independent. Based on the observed signal $\mathbf{y}$, we design a receiver to detect the information transmitted by each user. We restrict our attention to linear receivers that do not require knowledge of A_i or σ^2.

A linear multiuser receiver of this form is the decorrelator [6], which cross-correlates $\mathbf{y}$ with each of the columns $\mathbf{v}_i$ of $\mathbf{V} = \mathbf{S}(\mathbf{S}^*\mathbf{S})^{-1}$, to yield the outputs $a_i = \mathbf{v}_i^*\mathbf{y}$. The ith users' bit is then detected as $\hat{x}_i = \text{sgn}\,(a_i)$. The decorrelator optimally rejects the MAI but does not compensate for the white noise. Indeed, the covariance $\mathbf{C}_a$ of the noise component in the vector output $\mathbf{a}$ of the decorrelator is given by

$$\mathbf{C}_a = \sigma^2\mathbf{V}^*\mathbf{V} = \sigma^2(\mathbf{S}^*\mathbf{S})^{-1}, \tag{15}$$

so that the decorrelator tends to enhance the noise on the channel. To improve the performance of the decorrelator receiver, we propose a modified receiver, which we refer to as the covariance shaping multiuser (CSMU) receiver, that consists of optimally shaping the noise component in the output of the decorrelator prior to detection, using a weighted MMSE shaping transformation with scaling σ^2.

From (15) and Theorem 2 it follows that the CSMU receiver cross-correlates the received vector $\mathbf{y}$ with each of the columns $\mathbf{q}_i$ of $\mathbf{Q} = \mathbf{SR}(\mathbf{S}^*\mathbf{SR})^{-1/2}$. The ith users' bit is then detected as $\hat{x}_i = \text{sgn}\,(\mathbf{q}_i^*\mathbf{y})$. The choice of shaping $\mathbf{R}$ can be tailored to the specific set of signatures.

To demonstrate the performance advantage in using the CSMU receiver, we consider the case in which the signature vectors are chosen as PN sequences with norm 1 and equal inner products $-1/N$, and the shaping $\mathbf{R}$ is chosen as a circulant matrix with $[\mathbf{R}]_{ii} = 1$ and $[\mathbf{R}]_{ij} = \rho$ for $i \neq j$.

In Fig. 2 we plot the theoretical probability of bit error of the CSMU receiver in the case of 5 users with $\rho = 0.2$, where the first user, the desired user, has 4 interferers such that $A_i/A_1 = 0.5$ for $i = 2, 3, 4, 5$. The corresponding curves for the decorrelator, single-user MF and linear MMSE receivers [6] are plotted for comparison. We see that the CSMU receiver performs better than the decorrelator and the MF and performs similarly to the linear MMSE receiver which is the optimal linear receiver that assumes knowledge of the channel parameters.

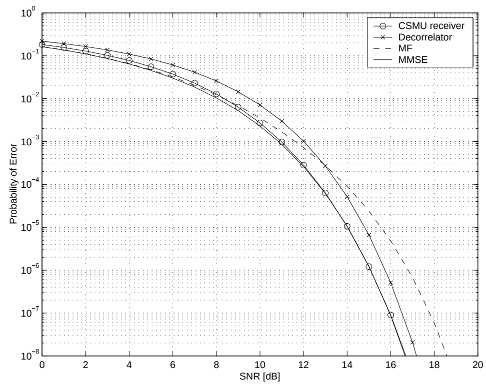

Fig. 2. Probability of bit error with 5 users, $\rho = 0.2$, and $A_i/A_1 = 0.5$ for $i = 2, 3, 4, 5$, as a function of SNR.

In Fig. 3 we plot the probability of bit error of the CSMU receiver in the case of 10 users with $\rho = 0.35$, and with accurate power control so that $A_i = 1$ for all i. Here again, the CSMU receiver performs better than the decorrelator and the MF and performs similarly to the linear MMSE receiver.

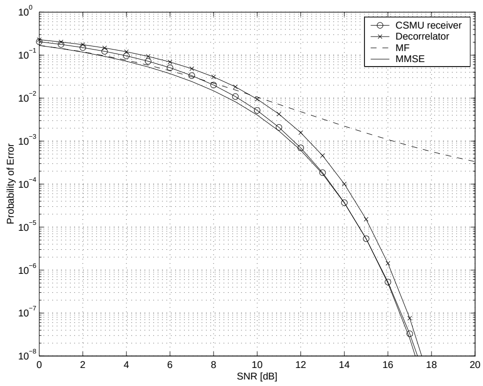

Fig. 3. Probability of bit error with 10 users, $\rho = 0.35$, and accurate power control, as a function of SNR.

3.3. Least-Squares Estimation

As a third application of MMSE shaping, we consider estimating the unknown deterministic parameters $\mathbf{x}$ in the linear model

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{w}, \tag{16}$$

where $\mathbf{H}$ is a known $n \times m$ matrix with full column rank, and $\mathbf{w}$ is a zero mean random vector with positive definite covariance $\mathbf{C}_w$.

Many signal processing estimation problems can be represented by the linear model (16), and consequently this problem has been studied extensively in the literature. A common approach to estimating the parameters $\mathbf{x}$ is to restrict the estimator $\hat{\mathbf{x}}$ to be linear in the data, and then find the linear estimator that is unbiased and has minimum variance [4]. The resulting estimator, denoted $\hat{\mathbf{x}}_{\mathrm{LS}}$, is also the LS estimate of $\mathbf{x}$ from the data $\mathbf{y}$, *i.e.*, it minimizes the LS error $(\mathbf{y} - \hat{\mathbf{y}})^* \mathbf{C}_w^{-1} (\mathbf{y} - \hat{\mathbf{y}})$, and is given by

$$\hat{\mathbf{x}}_{\mathrm{LS}} = (\mathbf{H}^* \mathbf{C}_w^{-1} \mathbf{H})^{-1} \mathbf{H}^* \mathbf{C}_w^{-1} \mathbf{y} = \mathbf{x} + \tilde{\mathbf{w}}, \tag{17}$$

where $\tilde{\mathbf{w}} = (\mathbf{H}^* \mathbf{C}_w^{-1} \mathbf{H})^{-1} \mathbf{H}^* \mathbf{C}_w^{-1} \mathbf{w}$ is the estimation error.

Since the covariance of $\tilde{\mathbf{w}}$ is given by $(\mathbf{H}^* \mathbf{C}_w^{-1} \mathbf{H})^{-1}$, a difficulty often encountered in this estimation problem is that the error in the estimation can have a large variance and a covariance structure with a very high dynamic range. To control the dynamic range and spectral shape of the covariance of the estimation error, we propose a modification of the LS estimator based on the concept of MMSE shaping, in which we optimally shape the covariance of the estimation error. From Theorem 2 the resulting estimator, referred to as the covariance shaping LS (CSLS) estimator and denoted $\hat{\mathbf{x}}_{\mathrm{CSLS}}$, is given by

$$\begin{aligned} \hat{\mathbf{x}}_{\mathrm{CSLS}} &= \hat{c} (\mathbf{R} \mathbf{H}^* \mathbf{C}_w^{-1} \mathbf{H})^{1/2} \hat{\mathbf{x}}_{\mathrm{LS}} \\ &= \hat{c} (\mathbf{R} \mathbf{H}^* \mathbf{C}_w^{-1} \mathbf{H})^{-1/2} \mathbf{R} \mathbf{H}^* \mathbf{C}_w^{-1} \mathbf{y}, \end{aligned} \tag{18}$$

where $\hat{c} = \mathrm{Tr}((\mathbf{R} \mathbf{H}^* \mathbf{C}_w^{-1} \mathbf{H})^{1/2}) / \mathrm{Tr}(\mathbf{R} \mathbf{H}^* \mathbf{C}_w^{-1} \mathbf{H})$.

It can be shown that the CSLS estimator can alternatively be derived as the estimator that minimizes the weighted total error variance in the observations subject to a constraint on the covariance of the estimation error [7]. It is also shown in [7] that the CSLS estimator achieves the Cramer-Rao lower bound for biased estimators. Furthermore, analysis of the MSE of both the CSLS estimator and the LS estimator demonstrates that the covariance of the estimation error can be chosen such that there is a threshold SNR, below which the CSLS estimator yields a lower MSE than the LS estimator, for all values of the unknown parameters $\mathbf{x}$.

Simulations presented in [3] strongly suggest that the CSLS estimator can significantly decrease the MSE of the estimation error over the LS estimator over a wide range of SNR values. In Fig. 4 we illustrate the performance advantage with one simulation from [3]. In this figure we plot the MSE in estimating a set of AR parameters in an ARMA model contaminated by white noise, using the CSLS estimator with $\mathbf{R} = \mathbf{I}$ and the LS estimator, from 20 noisy observations of the channel, averaged over 2000 noise realizations, as a function of SNR.

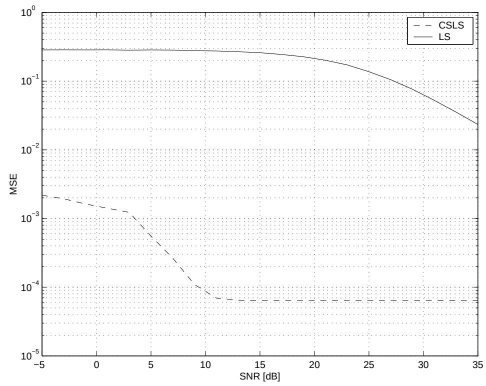

Fig. 4. Mean-squared error in estimating a set of AR parameters using the LS estimator and the CSLS estimator.

4. REFERENCES

[1] Y. C. Eldar and A. V. Oppenheim, “Orthogonal matched filter detection,” *Proc. Int. Conf. Acoust., Speech, Signal Processing (ICASSP-2001), (Salt Lake, UT)*, 2001.

[2] Y. C. Eldar and A. V. Oppenheim, “MMSE whitening and subspace whitening,” *IEEE Trans. Inform. Theory*, to appear.

[3] Y. C. Eldar, *Quantum Signal Processing*, Ph.D. thesis, Massachusetts Institute of Technology, Dec. 2001, Available: `http://allegro.mit.edu/dspg/publications/TechRep/index.html`.

[4] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*, Upper Saddle River, NJ: Prentice Hall, Inc., 1993.

[5] J. G. Proakis, *Digital Communications*, McGraw-Hill, Inc., third edition, 1995.

[6] S. Verdu, *Multiuser Detection*, Cambridge, UK: Cambridge Univ. Press, 1998.

[7] Y. C. Eldar and A. V. Oppenheim, “Covariance shaping least-squares estimation,” *IEEE Trans. Signal Processing*, March 2003.