# Semiblind Multiuser MIMO Channel Estimation Using Capon and MUSIC Techniques

Shahram Shahbazpanahi, *Member, IEEE*, Alex B. Gershman, *Fellow, IEEE*, and Georgios B. Giannakis, *Fellow, IEEE*

***Abstract*—Channel state information (CSI) is required at the receiver side of multiple-input multiple-output (MIMO) systems for decoding space–time codes. In this paper, new semiblind techniques are developed to jointly estimate the receiver CSI of multiple multiantenna transmitters that use orthogonal space–time block codes (OSTBCs). The proposed techniques are based on the extension of the concepts of the popular Capon and MUSIC spectral estimators to the problem of multiuser channel estimation. Our methods blindly estimate the subspace that contains the user channel matrices, and then use only a few training symbol blocks to extract the user CSI from this subspace. It is demonstrated that, in addition to improving the bandwidth efficiency, the proposed techniques offer better channel estimation accuracy as compared with the standard (nonblind) least-squares (LS) channel estimation method. Moreover, it is shown that using the proposed Capon- and MUSIC-based semiblind channel estimators along with the coherent maximum-likelihood (ML) or the minimum-variance (MV) multiuser MIMO receivers results in a substantially better performance as compared with the case when the standard LS channel estimates are used.**

***Index Terms*—Capon and MUSIC techniques, multiuser multiple-input multiple-output (MIMO) receivers, semiblind MIMO channel estimation.**

## I. Introduction

SPACE-TIME coding is a powerful means to exploit spatial diversity and to combat fading in multiple-input multiple-output (MIMO) wireless communication systems [1]–[3]. Among various space–time codes developed to the date, orthogonal space–time block codes (OSTBCs) [2], [4], [5] are of particular interest because they achieve full diversity while offering very simple maximum-likelihood (ML) decoding in the single-user case.

Manuscript received December 25, 2004; revised September 7, 2005. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Daniel Fuhrman. This work was supported is part by the Wolfgang Paul Award Program of the Alexander von Humboldt Foundation (Germany) and German Ministry of Education and Research; Premier's Research Excellence Award Program of the Ministry of Energy, Science, and Technology (MEST) of Ontario; Natural Sciences and Engineering Research Council (NSERC) of Canada; Communication and Information Technology Ontario (CITO); and National Science Foundation (NSF) Grant No. CCF-0515032. The results of this paper were presented it part at the IEEE Sensor Array and Multichannel Signal Processing Workshop, Sitges, Spain, June 2004.

S. Shahbazpanahi is with the Faculty of Engineering and Applied Science, University of Ontario Institute of Technology, Oshawa, ON L1H 7K4, Canada (e-mail: shahram.shahbazpanahi@uoit.ca).

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

G. B. Giannakis is with the Department of Electrical and Computer Engineering, University of Minnesota, Minneapolis, MN 55455 USA (e-mail: georgios@ece.umn.edu,).

Digital Object Identifier 10.1109/TSP.2006.877650

Point-to-point MIMO communications based on space–time block codes (STBCs) have been extensively studied in the literature. However, there are only a few papers where the problem of joint decoding and multiple-access interference (MAI) suppression in STBC-based multiuser communication systems has been addressed [6]–[9].

In all these papers, it is assumed that the channel state information (CSI) of the users of interest is available at the receiver. This implies that training should be used to estimate the user CSI. However, the use of training may substantially reduce the bandwidth efficiency and, therefore, blind and/or semiblind channel estimation techniques are of great interest [10].

Recently, numerous blind and semiblind decoding techniques have been proposed for point-to-point OSTBC-based MIMO communications (see [11]–[21] and references therein). These techniques include differential decoders [11]–[13], unitary space–time modulation-based space–time codes [14], least-squares (LS) blind and semiblind decoders for generalized STBCs (GSTBCs) [15], Khatri–Rao space–time codes [18], space–time decoding schemes based on subspace-based blind channel estimation [17], [21], and other approaches [16], [19], [20].

Unfortunately, most of these techniques are not directly applicable to multiuser MIMO scenarios because they are (explicitly or implicitly) based on the point-to-point MIMO assumption and treat MAI as a white noise. As a result, their performance may degrade severely in the multiuser case.

A promising semiblind parameter estimation method to detect the symbols of the desired user while rejecting MAI has been proposed in [22]. Note, however, that the MIMO case is not considered in this paper. That is, the authors of [22] assume that the receiver has multiple antennas, while each transmitter has a single antenna.

In this paper, we propose an approach to simultaneously estimate the CSI of multiple multiantenna transmitters that use OSTBCs to communicate with a single multiantenna receiver (base station). Our approach is based on the extension of the concepts of the popular Capon [23] and MUSIC [24] techniques commonly used in array processing and spectral analysis[1] to the problem of channel estimation in multiuser MIMO communications. As will be shown in the sequel, such extension is possible due to a close similarity between the array processing and multiuser MIMO channel estimation problems. It will be shown that, using such a similarity, the subspace that contains the user channel matrices can be estimated in a fully blind way. However, to extract the user CSI from this subspace, a few

[1]There are also several approaches that apply the ideas of Capon and MUSIC techniques to wireless communications, see, e.g., [25] and [26].

training blocks have to be used with the amount of training that is much smaller than required by the standard (nonblind) LS-based channel estimation method.

Our simulation results demonstrate that, in addition to improving the bandwidth efficiency, the proposed techniques also enjoy substantially lower channel estimation errors as compared to the LS channel estimation method. All these improvements are achieved at the expense of a moderate increase in decoding delay and computational complexity with respect to the LS estimator, as well as at the expense of the requirement of small channel variations during the interval where the sample covariance matrix is computed. It is shown that using the proposed semiblind channel estimators along with the ML or minimum-variance (MV) linear multiuser receivers results in a substantially better performance in terms of the symbol error rate (SER) as compared with the case when the standard LS channel estimator is used.

The remainder of this paper is organized as follows. In Section II, the multiuser MIMO model is developed. The ML and MV multiuser receivers are briefly reviewed in Section III. Section IV presents the standard LS and the proposed Capon- and MUSIC-based semiblind multiuser MIMO channel estimators. Simulations are given in Section V, and conclusions are drawn in Section VI.

## II. Multiuser MIMO Data Model

Let us consider the uplink multiuser MIMO case and assume that $P$ synchronous multiantenna transmitters[2] communicate with a single multiantenna receiver. All the transmitters are assumed to have the same number of antennas, $N$, while the receiver has $M$ antennas. We also assume that all the transmitters use the same OSTBC to encode the information symbols. Using these assumptions and considering the flat block-fading channel case, the received signal can be written as [9]

$$\mathbf{Y} = \sum_{p=1}^{P} \mathbf{X}(\mathbf{s}_p)\mathbf{H}_p + \mathbf{V} \tag{1}$$

where $\mathbf{Y}$ is the $T \times M$ matrix of the received signals, $\mathbf{s}_p$ is the $K \times 1$ vector of the information symbols of the $p$th user, $\mathbf{X}(\mathbf{s}_p)$ is the $T \times N$ matrix of the transmitted signals of this user, $\mathbf{H}_p$ is the $N \times M$ matrix of channel coefficients between the $p$th transmitter and the receiver, $\mathbf{V}$ is the $T \times M$ noise matrix, and $T$ denotes the block length. We assume that the entries of matrices $\{\mathbf{H}_p\}_{p=1}^{P}$ are independent random variables. This assumption implies that in the vector space of all $N \times M$ complex matrices, the channel matrices $\{\mathbf{H}_p\}_{p=1}^{P}$ are linearly independent, almost surely.[3]

It should be emphasized that, in contrast of code-division multiple-access (CDMA)-based approaches (see [6], [17], [27], and references therein), the considered underlying transmission scheme does not make any use of the user spreading codes, but exploits the fact that all the users have different channel vectors because of multiple antennas employed at the receiver and each of the transmitters.

The matrix $\mathbf{X}(\mathbf{s}_p)$ is assumed to correspond to some OSTBC [4]. The matrix $\mathbf{X}(\mathbf{s})$ is called an OSTBC if all elements of $\mathbf{X}(\mathbf{s})$ are linear functions of the $K$ complex variables $s_1, \ldots, s_K$ and their complex conjugates, and if for any arbitrary $\mathbf{s}$, it satisfies $\mathbf{X}^H(\mathbf{s})\mathbf{X}(\mathbf{s}) = \|\mathbf{s}\|^2\mathbf{I}_N$ where $\mathbf{I}_N$ is the $N \times N$ identity matrix, while $(\cdot)^H$ and $\|\cdot\|$ denote the Hermitian transpose and the Euclidean norm, respectively.

The matrix $\mathbf{X}(\mathbf{s})$ can be written as

$$\mathbf{X}(\mathbf{s}) = \sum_{k=1}^{K} (\mathbf{C}_k \mathrm{Re}\{s_k\} + \mathbf{D}_k \mathrm{Im}\{s_k\}) \tag{2}$$

where the matrices $\mathbf{C}_k$ and $\mathbf{D}_k$ are defined as

$$\mathbf{C}_k \triangleq \mathbf{X}(\mathbf{e}_k), \qquad \mathbf{D}_k \triangleq \mathbf{X}(j\mathbf{e}_k)$$

respectively, $j = \sqrt{-1}$, $\mathrm{Re}(\cdot)$ and $\mathrm{Im}(\cdot)$ denote the real and the imaginary parts, respectively, and $\mathbf{e}_k$ is the $K \times 1$ vector having one in its $k$th position and zeros elsewhere. Using (2), one can rewrite (1) in a vectorized form as [9], [28]

$$\underline{\mathbf{Y}} = \sum_{p=1}^{P} \mathbf{A}(\mathbf{h}_p)\, \underline{\mathbf{s}_p} + \underline{\mathbf{V}} \tag{3}$$

where

$$\mathbf{h}_p \triangleq \underline{\mathbf{H}_P}$$

is the equivalent channel vector of the $p$th user, and the "underline" operator for any matrix $\mathbf{P}$ is defined as

$$\underline{\mathbf{P}} \triangleq \begin{bmatrix} \mathrm{vec}\{\mathrm{Re}(\mathbf{P})\} \\ \mathrm{vec}\{\mathrm{Im}(\mathbf{P})\} \end{bmatrix}. \tag{4}$$

In (4), $\mathrm{vec}\{\cdot\}$ denotes the vectorization operator stacking all columns of a matrix on top of each other. The $2MT \times 2K$ real matrix $\mathbf{A}(\mathbf{h}_p)$ in (3) is given by [9], [28]

$$\begin{aligned} \mathbf{A}(\mathbf{h}_p) &\triangleq [\underline{\mathbf{C}_1\mathbf{H}_p} \ \ldots \ \underline{\mathbf{C}_K\mathbf{H}_p} \ \ \underline{\mathbf{D}_1\mathbf{H}_p} \ \ldots \ \underline{\mathbf{D}_K\mathbf{H}_p}] \\ &\triangleq [\mathbf{a}_1(\mathbf{h}_p) \ \mathbf{a}_2(\mathbf{h}_p) \ \ldots \ \mathbf{a}_{2K}(\mathbf{h}_p)]. \end{aligned}$$

This matrix has an important property that its columns have the same norms and are orthogonal to each other [28]

$$\mathbf{A}^T(\mathbf{h}_p)\mathbf{A}(\mathbf{h}_p) = \|\mathbf{h}_p\|^2\mathbf{I}_{2K} \tag{5}$$

where $(\cdot)^T$ stands for the transpose. As $\mathbf{A}(\mathbf{h}_p)$ is linear in $\mathbf{h}_p$, there exist $2K$ real matrices $\{\mathbf{\Phi}_k\}_{k=1}^{2K}$ with dimensions $2MT \times 2MN$ such that

$$\mathbf{a}_k(\mathbf{h}_p) = \mathbf{\Phi}_k \mathbf{h}_p \qquad \text{for} \quad k = 1, \ldots, 2K. \tag{6}$$

[2]The synchronous assumption is required for formulation of our semiblind techniques and, therefore, is adopted throughout the manuscript.

[3]This means that, with probability one, there does not exist any nonzero vector $\mathbf{c} = [c_1, \ldots, c_P]^T$ such that $\sum_{p=1}^{P} c_p \mathbf{H}_p = \mathbf{0}$.

Note that the matrices $\{\mathbf{\Phi}_k\}_{k=1}^{2K}$ are OSTBC-specific and known. Using (6), we have

$$\mathrm{vec}\{\mathbf{A}(\mathbf{h}_p)\} = \mathbf{\Phi}\mathbf{h}_p \tag{7}$$

where the $4KMT \times 2MN$ matrix $\mathbf{\Phi}$ is defined as

$$\mathbf{\Phi} \triangleq [\mathbf{\Phi}_1^T \;\; \mathbf{\Phi}_2^T \;\; \cdots \;\; \mathbf{\Phi}_{2K}^T]^T.$$

## III. Multiuser Receivers

In this section, we briefly review two known multiuser MIMO receivers that will be used in Section V.

Given the channel vectors $\{\mathbf{h}_p\}_{p=1}^P$, one can use the ML approach to detect the transmitted information symbols. This approach amounts to solving the following optimization problem:

$$\min_{\bar{\mathbf{s}}\in\mathcal{S}} \|\underline{\mathbf{Y}} - \mathbf{\Xi}\,\bar{\mathbf{s}}\| \tag{8}$$

where

$$\bar{\mathbf{s}} \triangleq \left[\underline{\mathbf{s}_1}^T \;\; \underline{\mathbf{s}_2}^T \;\; \ldots \;\; \underline{\mathbf{s}_P}^T\right]^T$$
$$\mathbf{\Xi} \triangleq [\mathbf{A}(\mathbf{h}_1) \;\; \mathbf{A}(\mathbf{h}_2) \;\; \ldots \;\; \mathbf{A}(\mathbf{h}_P)]$$

and $\mathcal{S}$ is the set of all possible values of $\bar{\mathbf{s}}$. To reduce the computational complexity of the ML multiuser receiver in (8), the sphere decoding technique can be used [29].

Unfortunately, the computational cost of sphere decoding algorithms may be prohibitively high when the dimension of $\bar{\mathbf{s}}$ is large (i.e., when $KP \gg 1$) and/or when the signal-to-noise ratio (SNR) is low (typically less than 5 dB). Alternatively, a much simpler MV linear receiver of [9] can be used. Linear receivers obtain the transmitted symbols of the $p$th user as [9]

$$\underline{\hat{\mathbf{s}}_p} = \mathbf{W}^T(\mathbf{h}_p)\underline{\mathbf{Y}}$$

where $\mathbf{W}(\mathbf{h}_p)$ is the $2MT \times 2K$ matrix of receiver coefficients. To detect the transmitted symbols, the estimate

$$\hat{\mathbf{s}}_p = [\mathbf{I}_k \; j\mathbf{I}_k]\underline{\hat{\mathbf{s}}_p}$$

has to be computed, and then the $k$th transmitted symbol should be detected as a point in the corresponding constellation which is the closest one to the $k$th element of $\hat{\mathbf{s}}_p$.

For the $p$th transmitter, the MV linear receiver of [9] can be written as

$$\mathbf{W}(\mathbf{h}_p) = \hat{\mathbf{R}}_{\mathrm{dl}}^{-1}\mathbf{A}(\mathbf{h}_p)\{\mathbf{A}^T(\mathbf{h}_p)\hat{\mathbf{R}}_{\mathrm{dl}}^{-1}\mathbf{A}(\mathbf{h}_p)\}^{-1} \tag{9}$$

where

$$\mathbf{R} \triangleq \mathrm{E}\{\underline{\mathbf{Y}}\;\underline{\mathbf{Y}}^T\}$$

is the covariance matrix of the vectorized received data, $\hat{\mathbf{R}}$ is its sample estimate, $\hat{\mathbf{R}}_{\mathrm{dl}} = \hat{\mathbf{R}} + \gamma\mathbf{I}$ is the diagonally loaded sample covariance matrix, and $\gamma$ is the loading factor. Diagonal loading is used in (9) to improve the receiver robustness.

Another class of robust MV MIMO receivers based on the idea of worst-case optimization-based beamforming [30] has been proposed in [31].

Note that the multiuser receivers in (8) and (9) are called "clairvoyant" if they assume that the true channel vectors $\{\mathbf{h}_p\}_{p=1}^P$ are available. Unfortunately, such an assumption can hardly be met in practical cases because of a limited number of training symbols and finite channel coherence time. Therefore, in practice, the channel vectors $\{\mathbf{h}_p\}_{p=1}^P$ have to be estimated first and then these estimates can be used in (8) or (9) to decode the information symbols.

## IV. Multiuser Channel Estimation

In this section, we first formulate the training-based LS channel estimation technique and then develop our semiblind Capon- and MUSIC-based channel estimators.

### A. Training-Based LS Channel Estimator

To recover the transmitted user symbols, the channel vectors $\{\mathbf{h}_p\}_{p=1}^P$ have to be estimated. One straightforward approach to estimate these vectors is to employ training and to use the LS method. Let us rewrite (3) as

$$\underline{\mathbf{Y}} = \sum_{p=1}^{P}\tilde{\mathbf{A}}(\underline{\mathbf{s}_p})\mathbf{h}_p + \underline{\mathbf{V}} \tag{10}$$

where $\tilde{\mathbf{A}}(\underline{\mathbf{s}_p})$ is a $2MT \times 2MN$ matrix whose $k$th column is given by

$$[\tilde{\mathbf{A}}(\underline{\mathbf{s}_p})]_k = \mathbf{A}(\mathbf{e}_k)\underline{\mathbf{s}_p}\,.$$

As before, $\mathbf{e}_k$ is the vector having one in its $k$th position and zeros elsewhere, but now its dimension is $2MN \times 1$.

Assuming that each user transmits $J_{\mathrm{t}}$ training blocks and using the data model (10), we can write

$$\mathbf{y}(n) \triangleq \underline{\mathbf{Y}(n)} = \sum_{p=1}^{P}\tilde{\mathbf{A}}(\underline{\mathbf{s}_p(n)})\mathbf{h}_p + \underline{\mathbf{V}(n)}, \quad n = 1,\ldots,J_{\mathrm{t}} \tag{11}$$

where $\mathbf{s}_p(n)$ is the $n$th known vector of training symbols transmitted by the $p$th user, and $\mathbf{Y}(n)$ and $\mathbf{V}(n)$ are, respectively, the received signal matrix and the noise matrix for the $n$th training block.

Using the notations

$$\mathbf{g} \triangleq [\mathbf{h}_1^T \; \mathbf{h}_2^T \; \ldots \; \mathbf{h}_p^T]^T$$
$$\boldsymbol{\mathcal{A}}(n) \triangleq [\tilde{\mathbf{A}}(\underline{\mathbf{s}_1(n)}) \;\; \tilde{\mathbf{A}}(\underline{\mathbf{s}_2(n)}) \;\; \cdots \;\; \tilde{\mathbf{A}}(\underline{\mathbf{s}_P(n)})]$$

(11) can be rewritten as

$$\mathbf{y}(n) = \boldsymbol{\mathcal{A}}(n)\mathbf{g} + \underline{\mathbf{V}(n)}, \quad n = 1, \dots, J_{\rm t}\,. \tag{12}$$

Stacking the vectors $\mathbf{y}(n)$ $(n = 1, \dots, J_{\rm t})$ in a longer $2MTJ_{\rm t} \times 1$vector

$$\mathbf{z} \triangleq [\mathbf{y}^T(1)\ \ \mathbf{y}^T(2)\ \ \cdots\ \ \mathbf{y}^T(J_{\rm t})]^T \tag{13}$$

and defining the $2MTJ_{\rm t} \times 2PMN$ matrix

$$\mathbb{A} \triangleq [\boldsymbol{\mathcal{A}}^T(1)\ \ \boldsymbol{\mathcal{A}}^T(2)\ \ \cdots\ \ \boldsymbol{\mathcal{A}}^T(J_{\rm t})]^T$$

and the $2MTJ_{\rm t} \times 1$ vector

$$\mathbf{n} \triangleq [\underline{\mathbf{V}^T(1)}\ \ \underline{\mathbf{V}^T(2)}\ \ \cdots\ \ \underline{\mathbf{V}^T(J_{\rm t})}]^T$$

we can rewrite (12) in a more compact form as

$$\mathbf{z} = \mathbb{A}\mathbf{g} + \mathbf{n}\,. \tag{14}$$

Based on (14), the LS estimate of the vector $\mathbf{g}$ can be expressed as

$$\hat{\mathbf{g}} = (\mathbb{A}^H\mathbb{A})^{-1}\mathbb{A}^H\mathbf{z}\,. \tag{15}$$

According to (15), one has to ensure that the matrix $\mathbb{A}$ is full column rank. Since the dimension of $\mathbb{A}$ is $2MTJ_{\rm t} \times 2PMN$, the full column rank condition of $\mathbb{A}$ implies that the choice of $J_{\rm t}$ should satisfy the following condition:

$$J_{\rm t} \geq \frac{PN}{T}\,. \tag{16}$$

Note that the LS estimate uses only the received data vectors corresponding to the training symbols. In fact, if the noise vectors $\underline{\mathbf{V}(n)}$ $(n = 1, \dots, J_{\rm t})$ are independent and identically distributed (i.i.d.) Gaussian, the LS estimate in (15) is ML optimal, i.e., it maximizes the likelihood function. In the next two subsections, we propose two novel channel estimation techniques that additionally use information-bearing (nontraining) received data to obtain improved channel estimates as compared to the LS estimate.

### *B. Capon-Based Technique*

Let us now develop a Capon-based technique to blindly estimate the subspace that contains the channel vectors $\{\mathbf{h}_p\}_{p=1}^P$. Having such estimate of this "channel subspace," we will be able to reduce the number of parameters required to be estimated to enable recovery of the channel vectors $\{\mathbf{h}_p\}_{p=1}^P$. As a result of such parsimonious channel characterization, the number of training blocks will be reduced as compared with the training-based LS approach, and, correspondingly, the bandwidth efficiency will be improved.

Applied to our problem, the Capon linear receiver can be interpreted as a sort of spatio–temporal filter that passes the signal of a hypothetical user with the channel vector $\mathbf{h}$ without any distortion while maximally rejecting the signals of the other users. More specifically, to linearly estimate the $k$th entry of some $2K \times 1$ real vector $\underline{\mathbf{s}}$ (which belongs to a hypothetical transmitter with the channel vector $\mathbf{h}$), one can obtain the coefficient vector $\mathbf{w}_k$ of the corresponding linear Capon receiver by solving the following optimization problem [6], [9]:

$$\min_{\mathbf{w}_k}\ \mathbf{w}_k^T\mathbf{R}\mathbf{w}_k \quad \text{subject to} \quad \mathbf{w}_k^T\mathbf{a}_k(\mathbf{h}) = 1\,. \tag{17}$$

The solution to (17) is well known to be

$$\mathbf{w}_k(\mathbf{h}) = \frac{1}{\mathbf{a}_k^T(\mathbf{h})\mathbf{R}^{-1}\mathbf{a}_k(\mathbf{h})}\,\mathbf{R}^{-1}\mathbf{a}_k(\mathbf{h})\,. \tag{18}$$

Note that, according to (18), a separate weight vector has to be obtained for each $k = 1, \dots, 2K$ (i.e., $2K$ linear receivers have to be used in parallel to estimate all entries of $\underline{\mathbf{s}}$).

For any channel vector $\mathbf{h}$ and the $k$th entry of $\underline{\mathbf{s}}$, we can define the Capon "spectrum" through the output of the corresponding Capon receiver as

$$P_{\rm C}^k(\mathbf{h}) \triangleq \mathbf{w}_k^T(\mathbf{h})\mathbf{R}\mathbf{w}_k(\mathbf{h}) = \frac{1}{\mathbf{a}_k^T(\mathbf{h})\mathbf{R}^{-1}\mathbf{a}_k(\mathbf{h})}\,. \tag{19}$$

As we defined the Capon "spectrum" in (19) as the output power of the $k$th linear receiver, it is expected to have a maximum for $\mathbf{h} = \mathbf{h}_p/\|\mathbf{h}_p\|$. Therefore, our goal will be to find the values of $\mathbf{h}$ which maximize the Capon function in (19). However, $\mathbf{a}_k(\mathbf{h})$ is linear in $\mathbf{h}$ and, therefore, the value of (19) can increase arbitrarily when $\mathbf{h} \to 0$. To avoid such a trivial solution, we constrain the norm of $\mathbf{h}$ as $\|\mathbf{h}\| = 1$.

Although any of the Capon functions $\{P_{\rm C}^k(\mathbf{h})\}_{k=1}^{2K}$ can be used to estimate the channel vectors, we propose to combine all of them in the following fashion:

$$\begin{aligned} Q_{\rm C}(\mathbf{h}) &\triangleq \sum_{k=1}^{2K} \frac{1}{P_{\rm C}^k(\mathbf{h})} \\ &= \sum_{k=1}^{2K} \mathbf{a}_k^T(\mathbf{h})\mathbf{R}^{-1}\mathbf{a}_k(\mathbf{h}) \\ &= \mathbf{h}^T\left(\sum_{k=1}^{2K} \boldsymbol{\Phi}_k^T\mathbf{R}^{-1}\boldsymbol{\Phi}_k\right)\mathbf{h} \end{aligned} \tag{20}$$

where the last line of (20) is obtained from (6), and $Q_{\rm C}(\mathbf{h})$ can be viewed as a null-spectrum.

The definition of $Q_{\rm C}(\mathbf{h})$ in (20) allows us to find the channel estimates in closed form. To show this, we note that if all Capon spectra $\{P_{\rm C}^k(\mathbf{h})\}_{k=1}^{2K}$ have a maximum for $\mathbf{h} = \mathbf{h}_p/\|\mathbf{h}_p\|$, then $Q_{\rm C}(\mathbf{h})$ will have a minimum for $\mathbf{h} = \mathbf{h}_p/\|\mathbf{h}_p\|$. Therefore, we propose to estimate the *normalized* user channel vectors as the

$P$ values of $\mathbf{h}$ which minimize $Q_{\rm C}(\mathbf{h})$. Hence, exploiting the fact that the channel vectors $\{\mathbf{h}_p\}_{p=1}^{P}$ are linearly independent, the true channel vectors $\{\mathbf{h}_p\}_{p=1}^{P}$ are expected to belong to the subspace spanned by the *minor* eigenvectors of the matrix

$$\boldsymbol{\Psi} \triangleq \sum_{k=1}^{2K} \boldsymbol{\Phi}_k^T \mathbf{R}^{-1} \boldsymbol{\Phi}_k$$

or, more specifically, the eigenvectors corresponding to the $P$ smallest eigenvalues of this matrix.

We emphasize that these $P$ smallest eigenvalues can have a multiplicity. To show this, note that the $p$th smallest eigenvalue of $\boldsymbol{\Psi}$ is equal to the value of the $p$th minimum of $Q_{\rm C}(\mathbf{h})$. Rewrite $Q_{\rm C}(\mathbf{h})$ as

$$\begin{aligned} Q_{\rm C}(\mathbf{h}) &= \mathbf{h}^T \boldsymbol{\Phi}^T (\mathbf{I}_{2K} \otimes \mathbf{R}^{-1}) \boldsymbol{\Phi} \mathbf{h} \\ &= \mathrm{tr}\{\mathbf{A}^T(\mathbf{h}) \mathbf{R}^{-1} \mathbf{A}(\mathbf{h})\} \end{aligned} \tag{21}$$

where $\mathrm{tr}\{\cdot\}$ and $\otimes$ denote the trace operator and the Kronecker product, respectively. From (21), it follows that if for any normalized vector $\mathbf{h}$ ($\|\mathbf{h}\| = 1$) there exists a normalized vector $\tilde{\mathbf{h}} \neq \pm\mathbf{h}$ ($\|\tilde{\mathbf{h}}\| = 1$) and a $2K \times 2K$ matrix $\mathbf{Q}$ such that

$$\mathbf{A}(\tilde{\mathbf{h}}) = \mathbf{A}(\mathbf{h})\mathbf{Q} \tag{22}$$

then $\tilde{\mathbf{h}}$ will satisfy

$$Q_{\rm C}(\tilde{\mathbf{h}}) = Q_{\rm C}(\mathbf{h}) .$$

To prove this fact, note that according to (5), $\mathbf{A}(\tilde{\mathbf{h}})^T \mathbf{A}(\tilde{\mathbf{h}}) = \mathbf{A}(\mathbf{h})^T \mathbf{A}(\mathbf{h}) = \mathbf{I}_{2K}$. Therefore, $\mathbf{Q}$ is unitary ($\mathbf{Q}^T\mathbf{Q} = \mathbf{Q}\mathbf{Q}^T = \mathbf{I}_{2K}$). Using the latter fact and (21), we have

$$\begin{aligned} Q_{\rm C}(\tilde{\mathbf{h}}) &= \mathrm{tr}\{\mathbf{A}^T(\tilde{\mathbf{h}}) \mathbf{R}^{-1} \mathbf{A}(\tilde{\mathbf{h}})\} \\ &= \mathrm{tr}\{\mathbf{Q}^T \mathbf{A}^T(\mathbf{h}) \mathbf{R}^{-1} \mathbf{A}(\mathbf{h}) \mathbf{Q}\} \\ &= \mathrm{tr}\{\mathbf{Q}\mathbf{Q}^T \mathbf{A}^T(\mathbf{h}) \mathbf{R}^{-1} \mathbf{A}(\mathbf{h})\} \\ &= \mathrm{tr}\{\mathbf{A}^T(\mathbf{h}) \mathbf{R}^{-1} \mathbf{A}(\mathbf{h})\} \\ &= Q_{\rm C}(\mathbf{h}) . \end{aligned}$$

Therefore, if (22) holds true, then the values of the null-spectrum $Q_{\rm C}(\cdot)$ for the vectors $\mathbf{h}$ and $\tilde{\mathbf{h}}$ are the same. This implies that if $Q_{\rm C}(\cdot)$ has a minimum for $\mathbf{h} = \mathbf{h}_p/\|\mathbf{h}_p\|$ ($p = 1, 2, \ldots, P$), and if there exist a normalized vector $\tilde{\mathbf{h}}$ and a matrix $\mathbf{Q}$ that satisfy

$$\mathbf{A}(\tilde{\mathbf{h}}) = \mathbf{A}\left(\frac{\mathbf{h}_p}{\|\mathbf{h}_p\|}\right)\mathbf{Q} \tag{23}$$

then the minimum values $Q_{\rm C}(\tilde{\mathbf{h}})$ and $Q_{\rm C}(\mathbf{h}_p/\|\mathbf{h}_p\|)$ are the same. In other words, all the values of $\tilde{\mathbf{h}}$ which satisfy (23) will result in the same minimum of $Q_{\rm C}(\cdot)$ as $\mathbf{h} = \mathbf{h}_p/\|\mathbf{h}_p\|$ does. Let us define the set

$$\mathcal{B}_p \triangleq \left\{\tilde{\mathbf{h}} \mid \mathbf{A}(\tilde{\mathbf{h}}) = \mathbf{A}\left(\frac{\mathbf{h}_p}{\|\mathbf{h}_p\|}\right)\mathbf{Q}\right\} .$$

It has been shown in [21] that this set is a subspace whose dimension depends on the underlying OSTBC and on the number of receive antennas but does not depend on the value of $\mathbf{h}_p$. Denoting the dimension of this subspace by $L$, from the results of [21] it follows that for most OSTBCs, $L = 1$ holds true, but there are a few OSTBCs (including the Alamouti's code) for which $L > 1$. For more details, see the table in [21] and [32], which summarizes the values of $L$ for different OSTBCs.

From the discussion above, it follows that if $Q_{\rm C}(\cdot)$ has a minimum for $\mathbf{h} = \mathbf{h}_p/\|\mathbf{h}_p\|$, then it will have the same minimum value for any member of $\mathcal{B}_p$. This, in turn, implies that each of the $P$ smallest eigenvalues of $\boldsymbol{\Psi}$ has the multiplicity order of $L$.

Denoting the eigenvectors corresponding to $P$ smallest eigenvalues of $\boldsymbol{\Psi}$ as $\mathbf{u}_k$ ($k = 1, \ldots, LP$) and the subspace spanned by these eigenvectors as $\mathcal{U}$, we have

$$\mathbf{h}_p = \sum_{k=1}^{LP} \alpha_{pk} \mathbf{u}_k . \tag{24}$$

Note that to find the vectors $\{\mathbf{u}_k\}_{k=1}^{LP}$, only the knowledge of the data covariance matrix $\mathbf{R}$ and the OSTBC is required. This matrix can be estimated without any training data. However, to determine the real coefficients $\alpha_{pk}$ ($p = 1, \ldots, P$; $k = 1, \ldots, LP$), training data should be used. In what follows, we show how these coefficients can be obtained based on the LS approach.

Let each transmitter send a total number of $J$ blocks to the receiver and let out of these $J$ blocks, the first $J_{\rm t}$ training blocks be known to the receiver while the remaining $J - J_{\rm t}$ blocks be used to convey the information symbols. We can use all $J$ blocks to obtain the sample estimate of $\mathbf{R}$ as

$$\hat{\mathbf{R}} = \frac{1}{J} \sum_{n=1}^{J} \mathbf{y}(n) \mathbf{y}^T(n) .$$

Replacing $\mathbf{R}$ with $\hat{\mathbf{R}}$, we can estimate the vectors $\{\mathbf{u}_k\}_{k=1}^{LP}$ as the minor eigenvectors $\{\hat{\mathbf{u}}_k\}_{k=1}^{LP}$ of the matrix

$$\hat{\boldsymbol{\Psi}} \triangleq \sum_{k=1}^{2K} \boldsymbol{\Phi}_k^T \hat{\mathbf{R}}^{-1} \boldsymbol{\Phi}_k .$$

Then, using (24), we have

$$\mathbf{A}(\hat{\mathbf{h}}_p) = \sum_{k=1}^{LP} \alpha_{pk} \mathbf{A}(\hat{\mathbf{u}}_k) \tag{25}$$

which follows from the linearity of $\mathbf{A}(\cdot)$ in its argument. Replacing $\mathbf{A}(\mathbf{h}_p)$ in (3) with $\mathbf{A}(\hat{\mathbf{h}}_p)$ of (25), we obtain

$$\begin{aligned} \mathbf{y}(n) &= \sum_{k=1}^{LP} \mathbf{B}_k(n) \boldsymbol{\alpha}_k + \underline{\mathbf{V}(n)} \\ &= \mathbf{B}(n)\boldsymbol{\alpha} + \underline{\mathbf{V}(n)}, \quad n = 1, \ldots J_{\rm t} \end{aligned} \tag{26}$$

where

$$\begin{aligned}\mathbf{B}_k(n) &\triangleq [\mathbf{A}(\hat{\mathbf{u}}_k)\underline{\mathbf{s}_1(n)}\ \mathbf{A}(\hat{\mathbf{u}}_k)\underline{\mathbf{s}_2(n)}\ \ldots\ \mathbf{A}(\hat{\mathbf{u}}_k)\underline{\mathbf{s}_P(n)}]\\ \mathbf{B}(n) &\triangleq [\mathbf{B}_1(n)\ \mathbf{B}_2(n)\ \cdots\ \mathbf{B}_{LP}(n)]\\ \boldsymbol{\alpha}_k &\triangleq [\alpha_{1,k}\ \cdots\ \alpha_{LP,k}]^T\\ \boldsymbol{\alpha} &\triangleq [\boldsymbol{\alpha}_1^T\ \boldsymbol{\alpha}_2^T\ \cdots\ \boldsymbol{\alpha}_P^T]^T.\end{aligned}$$

Stacking the vectors $\mathbf{y}(n)$ ($n = 1, \ldots, J_{\mathrm{t}}$) in the longer vector $\mathbf{z}$ defined in (13) and using the notation

$$\mathbb{B} \triangleq [\mathbf{B}^T(1)\ \mathbf{B}^T(2)\ \cdots\ \mathbf{B}^T(J_{\mathrm{t}})]^T$$

we obtain from (26) the following relationship:

$$\mathbf{z} = \mathbb{B}\boldsymbol{\alpha} + \mathbf{n}$$

and, therefore, the LS estimate of $\boldsymbol{\alpha}$ can be written as

$$\hat{\boldsymbol{\alpha}} = (\mathbb{B}^H\mathbb{B})^{-1}\mathbb{B}^H\mathbf{z}. \tag{27}$$

Note that if the noise is i.i.d. Gaussian, then $\hat{\boldsymbol{\alpha}}$ is the ML estimate of $\boldsymbol{\alpha}$. Once the estimate of $\boldsymbol{\alpha}$ is found, the channel estimate for the $p$th user can be obtained from (24).

According to (27), the $2MTJ_{\mathrm{t}} \times LP^2$ matrix $\mathbb{B}$ has to be full column rank. This implies that the choice of $J_{\mathrm{t}}$ should satisfy the following condition:

$$J_{\mathrm{t}} \geq \frac{LP^2}{2MT}. \tag{28}$$

Comparing (28) with (16), we obtain that the Capon-based technique requires less training blocks than the standard LS channel estimator provided that

$$P < \frac{2MN}{L}. \tag{29}$$

It should be emphasized here that multiuser receivers that employ the obtained channel estimates imply stronger restrictions on $P$ than (29). For example, from the structure of the ML receiver (8), it is clear that if $P \leq MT/K$ is not satisfied, then the resulting estimates of the information symbols are not unique. The same condition should be satisfied for the MV receiver (9) [9]. Taking into account that for the majority of OSTBCs $1/2 \leq K/T \leq 1$, we obtain that $P$ cannot exceed $2M$ anyway—otherwise, the ML and MV receivers are not applicable. This (along with the fact that if $M > 1$ then $L = 1$ for most of OSTBCs [21]) implies that condition (29) is satisfied and, therefore, the proposed Capon-based estimator has better bandwidth efficiency than the standard LS estimator.

### C. MUSIC-Based Technique

To develop our MUSIC-based technique for the considered multiuser channel estimation problem, the signal and the noise subspaces have to be defined first. Let us define the signal subspace as that spanned by the columns of the matrices $\{\mathbf{A}(\mathbf{h}_p)\}_{p=1}^P$ and the noise subspace as the orthogonal complement to the signal subspace. Note that the dimension of the signal subspace is (almost sure) $2KP$, and, therefore, the dimension of the noise subspace is $2MT - 2KP$. To ensure that the noise subspace is nondegenerate, we require the number of the transmitters $P$ to be smaller than $\lfloor MT/K \rfloor$, where $\lfloor r \rfloor$ denotes the largest integer smaller or equal to $r$.

It is well known [24] that the signal subspace is spanned by the principal eigenvectors of the data covariance matrix $\mathbf{R}$ and the remaining eigenvectors of this matrix span the noise subspace. Based on this fact, we have

$$\mathbf{A}^T(\mathbf{h}_p)\mathbf{E} = \mathbf{0}, \qquad p = 1, \ldots, P \tag{30}$$

where the $2MT \times (2MT - 2KP)$ matrix $\mathbf{E}$ is built from the $2(MT - KP)$ minor eigenvectors of $\mathbf{R}$.

According to (30), let us define the generalized MUSIC "spectrum" as

$$P_{\mathrm{MUSIC}}(\mathbf{h}) = \frac{1}{\|\mathbf{A}^T(\mathbf{h})\mathbf{E}\|_F^2} \tag{31}$$

where $\|\cdot\|_F$ denotes the Frobenius norm. To avoid the trivial zero solution, we assume as before that $\|\mathbf{h}\| = 1$. Doing so, the user normalized channel vector estimates are given as the values of $\mathbf{h}$ for which $P_{\mathrm{MUSIC}}(\mathbf{h})$ has its $P$ most prominent peaks.

Using (7), the MUSIC function of (31) can be simplified as

$$\begin{aligned}P_{\mathrm{MUSIC}}(\mathbf{h}) &\triangleq \frac{1}{\mathrm{tr}\{\mathbf{A}^T(\mathbf{h})\mathbf{E}\mathbf{E}^T\mathbf{A}(\mathbf{h})\}}\\ &= \frac{1}{\mathrm{vec}\{\mathbf{A}(\mathbf{h})\}^T(\mathbf{I}_{2K} \otimes \mathbf{E}\mathbf{E}^T)\mathrm{vec}\{\mathbf{A}(\mathbf{h})\}}\\ &= \frac{1}{\mathbf{h}^T\boldsymbol{\Phi}^T(\mathbf{I}_{2K} \otimes \mathbf{E}\mathbf{E}^T)\boldsymbol{\Phi}\mathbf{h}}.\end{aligned} \tag{32}$$

From (32), it can be seen that the channel vectors belong to the subspace spanned by the $LP$ minor eigenvectors of the matrix

$$\boldsymbol{\Gamma} \triangleq \boldsymbol{\Phi}^T(\mathbf{I}_{2K} \otimes \mathbf{E}\mathbf{E}^T)\boldsymbol{\Phi}.$$

Denoting these eigenvectors as $\mathbf{v}_k$ ($k = 1, \ldots, LP$) and indicating the subspace spanned by these eigenvectors as $\mathcal{V}$, we have that the estimates of $\mathbf{h}_p$ ($p = 1, \ldots, P$) can be found as

$$\hat{\mathbf{h}}_p = \sum_{k=1}^{LP} \beta_{pk}\hat{\mathbf{v}}_k \tag{33}$$

where $\{\hat{\mathbf{v}}_k\}_{k=1}^{LP}$ are the sample estimates of $\{\mathbf{v}_k\}_{k=1}^{LP}$, i.e., the $LP$ minor eigenvectors of the matrix

$$\hat{\boldsymbol{\Gamma}} = \boldsymbol{\Phi}^T(\mathbf{I}_{2K} \otimes \hat{\mathbf{E}}\hat{\mathbf{E}}^T)\boldsymbol{\Phi}$$

and $\hat{\mathbf{E}}$ is the sample estimate of $\mathbf{E}$ obtained through the eigendecomposition of $\hat{\mathbf{R}}$.

Using the LS-based procedure similar to that developed in the previous subsection for the Capon-based estimator, and introducing

$$\begin{aligned}\boldsymbol{\beta}_k &\triangleq [\beta_{1,k}\ \beta_{2,k}\ \cdots\ \beta_{LP,k}]^T\\ \boldsymbol{\beta} &\triangleq [\boldsymbol{\beta}_1^T\ \boldsymbol{\beta}_2^T\ \cdots\ \boldsymbol{\beta}_P^T]^T\end{aligned}$$

we have that the vector $\boldsymbol{\beta}$ can be estimated as

$$\hat{\boldsymbol{\beta}} = (\mathbb{F}^H\mathbb{F})^{-1}\mathbb{F}^H\mathbf{z} \tag{34}$$

where

$$\mathbb{F} \triangleq [\mathbf{F}^T(1)\ \mathbf{F}^T(2)\ \cdots\ \mathbf{F}^T(J_{\rm t})]^T$$
$$\mathbf{F}(n) \triangleq [\mathbf{F}_1(n)\ \mathbf{F}_2(n)\ \cdots\ \mathbf{F}_{LP}(n)]$$
$$\mathbf{F}_k(n) \triangleq [\mathbf{A}(\hat{\mathbf{v}}_k)\underline{\mathbf{s}_1(n)}\ \mathbf{A}(\hat{\mathbf{v}}_k)\underline{\mathbf{s}_2(n)}\ \cdots\ \mathbf{A}(\hat{\mathbf{v}}_k)\underline{\mathbf{s}_P(n)}].$$

Once $\boldsymbol{\beta}$ is estimated by means of (34), the channel vector estimate for the $p$th user can be found from (33).

According to (34), the $2MTJ_{\rm t} \times LP^2$ matrix $\mathbb{F}$ has to be full column rank, and, therefore, the condition similar to (28) should be satisfied for $J_{\rm t}$. This means that, similar to the Capon-based technique, the number of training blocks required for the MUSIC-based channel estimator is less than that required for the standard LS technique. Hence, both the Capon- and MUSIC-based estimators have better bandwidth efficiency than the LS estimator.

Improving the bandwidth efficiency using the proposed Capon- and MUSIC-based channel estimators can be explained by the fact that both of them are able to exploit the information-bearing (non-training) received data to estimate the subspace which contains the channel vectors. This reduces the number of remaining parameters to be estimated using the training data, and makes the channel parametrization more parsimonious than that used in the standard LS estimator. Indeed, after such blind estimate of the subspace of channel vectors is found, the proposed techniques need to obtain only the coefficient vectors $\boldsymbol{\alpha}$ or $\boldsymbol{\beta}$ from the training data to recover the user channels. Note, however, that the standard LS estimator needs to obtain all the channel vectors $\{\mathbf{h}_p\}_{p=1}^P$ from the training data. Therefore, our methods can be expected to outperform the standard LS technique.

It is well known that high-resolution techniques (including Capon and MUSIC estimators) may be very sensitive to model perturbations [33]–[35]. For example, in array processing a severe performance degradation of these techniques can be caused by various types of model errors in the array manifold (such as calibration errors, distorted array shape, propagation mismatches, etc.). It is worth noting that in the considered application to MIMO multiuser channel estimation, there are no model perturbations of such kind because for any value of $\mathbf{h}$, the matrix $\mathbf{A}(\mathbf{h})$ is fully determined by the underlying OSTBC which is known at the receiver. Therefore, in the considered case the performance of the proposed Capon- and MUSIC-based estimators is not limited by model errors.

The performance of the proposed Capon- and MUSIC-based estimators at low SNR values may be limited by subspace estimation errors (for example, subspace swaps that may affect the subspace estimates used in these techniques). Note, however, that this type of performance degradation is common for all subspace-based methods.

## V. Simulations

Throughout the simulations, the SNR for the $p$th user is defined as $\sigma_p^2/\sigma^2$, where $\sigma_p^2$ is the variance of each complex entry of $\mathbf{H}_p$ and $\sigma^2$ is the noise power. In each Monte-Carlo run, the entries of the channel matrices $\{\mathbf{H}_p\}_{p=1}^P$ are generated as complex zero-mean i.i.d. Gaussian random variables. In both examples, the 3/4 rate ($K = 3$, $T = 4$) amicable design-based OSTBC of [3] is used by all the transmitters to encode the information symbols. According to [21], if $M > 1$, then $L = 1$ for this code (i.e., there is no multiplicity in the eigenvalues of $\boldsymbol{\Psi}$). The sample covariance matrix is computed at the receiver using $J = 300$ blocks, and the channel is assumed to be fixed during these $J$ blocks (that is, only scenarios with slow fading are tested where the channel does not vary substantially for many data blocks).

In the first example, we consider $P = 2$ transmitters with $N = 4$ transmit antennas each, and a receiver with $M = 4$ antennas. The SNR of one of the users is assumed to be 2.5 dB smaller than that of the other user. The performance of the proposed Capon- and MUSIC-based semiblind channel estimators is compared to that of the standard LS estimator. Figs. 1 and 2

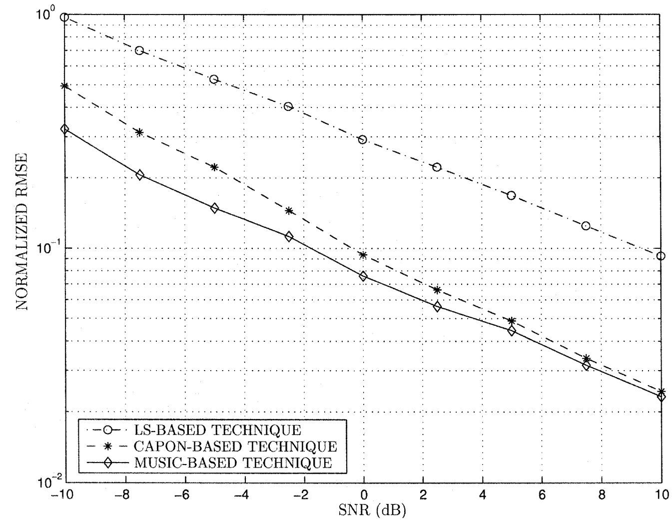



Fig. 1. Normalized RMSEs of channel estimates of the weaker transmitter versus SNR; first example.

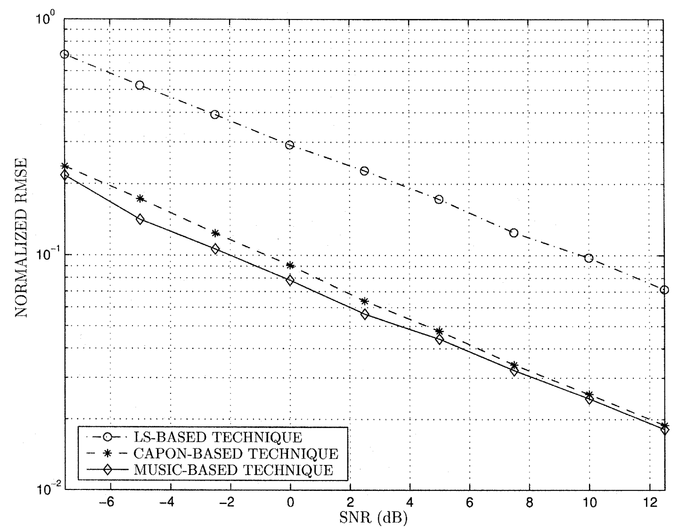



Fig. 2. Normalized RMSEs of channel estimates of the stronger transmitter versus SNR; first example.

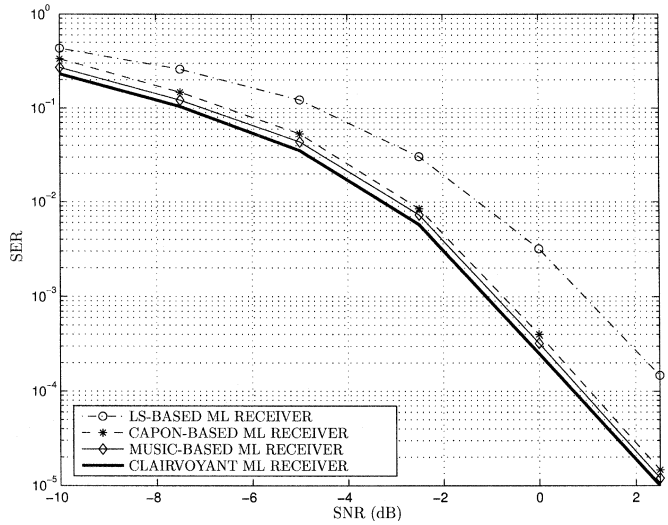



Fig. 3. SERs of the sphere decoding based ML receiver for the weaker transmitter versus SNR; first example.

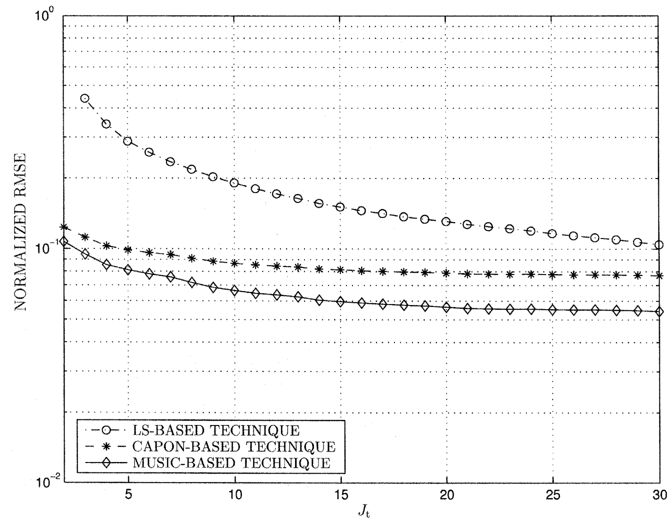



Fig. 5. Normalized RMSEs of channel estimates of the weaker transmitter versus number of training blocks; first example.

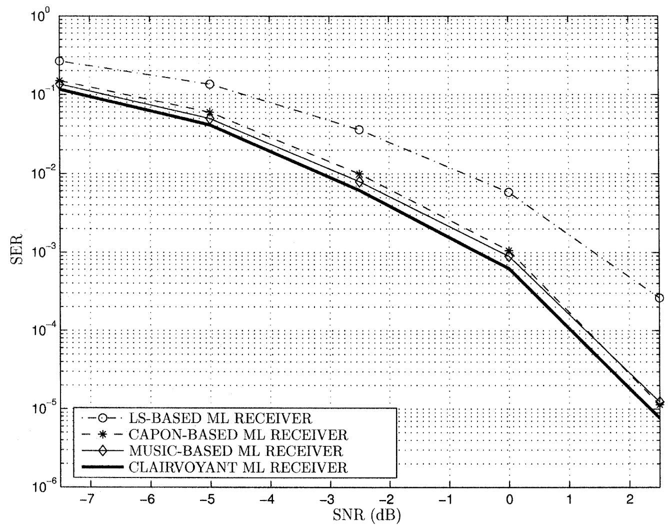



Fig. 4. SERs of the sphere decoding based ML receiver for the stronger transmitter versus SNR; first example.

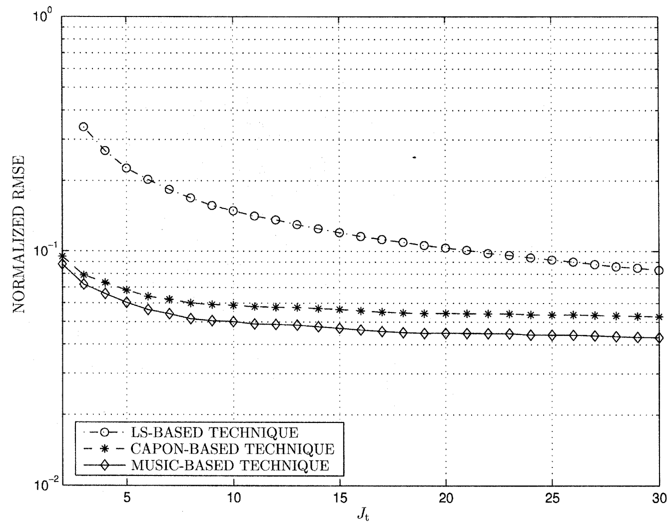



Fig. 6. Normalized RMSEs of channel estimates of the stronger transmitter versus number of training blocks; first example.

show the normalized root-mean-square errors (RMSEs) of the obtained channel estimates for the weaker and stronger transmitter, respectively, versus the SNR of the corresponding transmitter. In both of these figures, only $J_{\rm t} = 5$ training blocks are used. It can be seen from these figures that both the Capon- and the MUSIC-based techniques greatly outperform the standard LS-based method. Note that in both of these figures the MUSIC-based estimator performs better than the Capon-based technique.

Figs. 3 and 4 show the symbol error rates (SERs) of the ML multiuser receiver (8) that uses the Capon- MUSIC-, and LS-based channel estimates for the weaker and stronger transmitters, respectively, versus the SNR of the corresponding transmitter. In these figures, sphere decoding has been used to reduce the computational cost of the ML receiver. In addition, the SERs of the clairvoyant ML multiuser receiver are also shown in Figs. 3 and 4. Note that the clairvoyant ML receiver does not correspond to any practical situation and is used for comparison purposes only. As can be seen from Figs. 3 and 4, the performance of the MUSIC-based ML receiver is very close to that of the clairvoyant ML receiver, while the Capon-based ML receiver has slightly worse performance. It can also be observed from these two figures that the proposed Capon- and MUSIC-based channel estimation methods provide about 2 dB of improvement in the ML decoder performance as compared with the case when the LS-based channel estimates are used.

To illustrate bandwidth efficiency improvements achieved by the proposed techniques with respect to the LS estimator, the normalized RMSEs of the channel estimates of the weaker and stronger transmitter, respectively, are displayed in Figs. 5 and 6 versus the number of training blocks $J_{\rm t}$. In these figures, the SNRs of the weaker and stronger transmitters are 0 and 2.5 dB, respectively. As can be seen from these figures, the Capon- and MUSIC-based channel estimation techniques perform substantially better than the LS-based method. These performance improvements are especially pronounced when the number of training blocks is small. The corresponding SER curves of the ML receiver (8) are shown in Figs. 7 and 8 for the weaker and

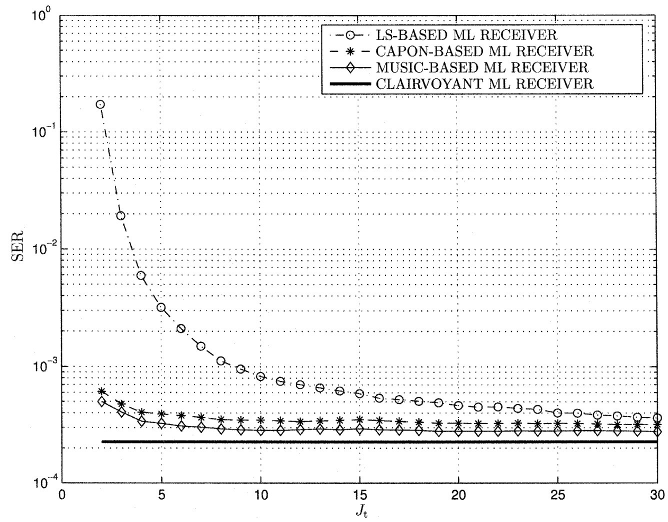



Fig. 7. SERs of the sphere decoding based ML receiver for the weaker transmitter versus number of training blocks; first example.

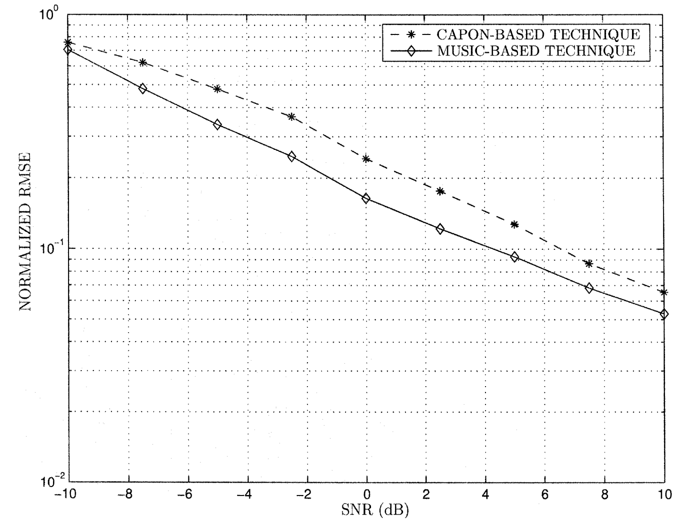



Fig. 9. Normalized RMSEs of channel estimates of the weaker transmitter versus SNR; second example.

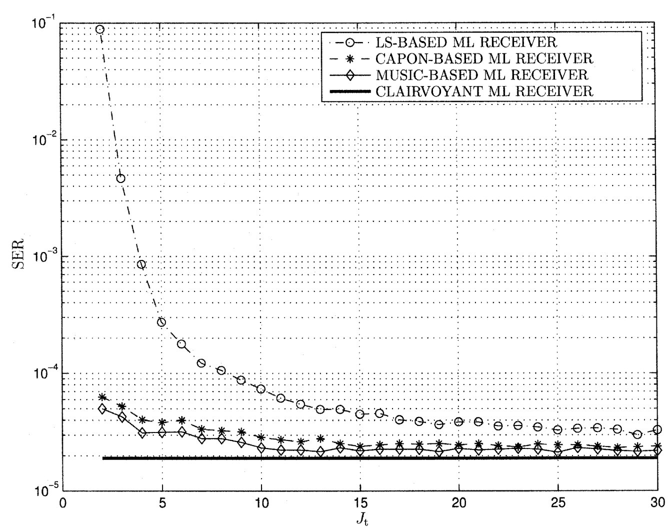



Fig. 8. SERs of the sphere decoding-based ML receiver for the stronger transmitter versus number of training blocks; first example.

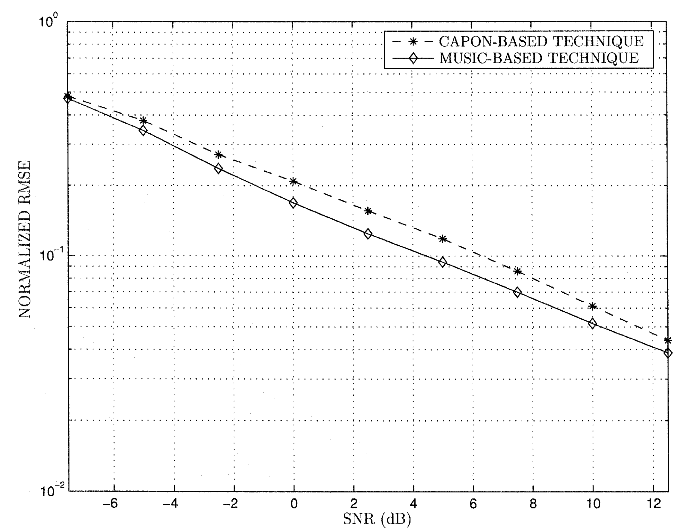



Fig. 10. Normalized RMSEs of channel estimates of one of the stronger transmitters versus SNR; second example.

stronger transmitter, respectively. Similar to Figs. 5 and 6, the SNRs of the weaker and stronger transmitters in Figs. 7 and 8 are 0 and 2.5 dB, respectively. The latter two figures show that the Capon- and MUSIC-based ML receivers offer substantially lower SERs than the LS-based ML receiver. These performance improvements are especially pronounced when the number of training blocks is low ($J_t = 2$ to $J_t = 10$).

In the second example, we consider $P = 8$ transmitters with $N = 4$ transmit antennas each, and a receiver with $M = 8$ antennas. $J_t = 5$ training blocks are used. Note that, according to (16), this number of training blocks is not sufficient for the standard LS-based technique, and hence, this technique cannot be used in this example. Moreover, with $P = 8$ transmitters the computational cost of the ML receiver is prohibitively high and, therefore, the MV receiver (9) is used to decode the information symbols. It is assumed that the SNR of one of the users is 2.5 dB smaller than of the other users (which have the same SNR). Figs. 9 and 10 display the normalized RMSEs of the channel estimates of the weaker and one of the stronger transmitters, respectively, versus the SNR of the corresponding transmitter. Figs. 11 and 12 show the corresponding SERs of the MV receiver (9) based on different channel estimates for the weaker and one of the stronger transmitters, respectively, versus the SNR of the corresponding transmitter. In addition to the performance of the MV receivers that use Capon-, MUSIC-, and LS-based channel estimates, the performance of the clairvoyant MV receiver is also shown in these two figures.

As can be seen from Figs. 9–12, the MUSIC-based technique has better performance as compared to the Capon-based approach, both in terms of RMSE and SER. It is noteworthy that for high values of SNR, the SER of the MUSIC-based MV receiver is very close to that of the clairvoyant MV receiver.

It is important to note that performance improvements achieved by the Capon- and MUSIC-based methods come at the price of a larger decoding delay and a moderate increase in computational complexity as compared to the LS

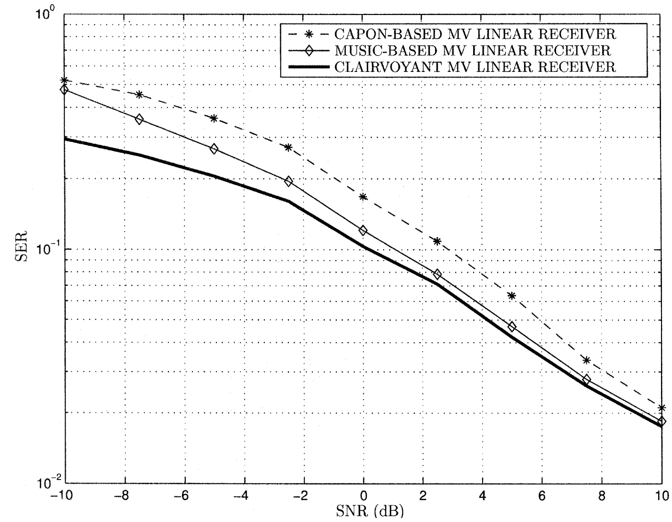



Fig. 11. SERs of the MV receiver for the weaker transmitter versus SNR; second example.

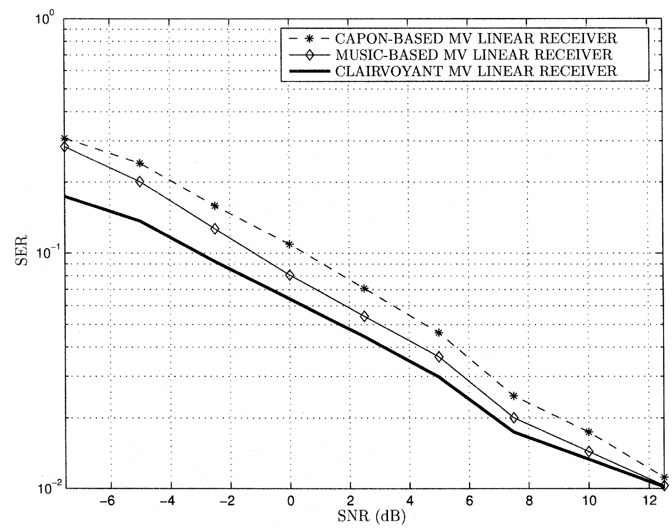



Fig. 12. SERs of the MV receiver for one of the stronger transmitters versus SNR; second example.

method. Indeed, both our techniques require enough number of data blocks to form a reliable sample estimate of the data covariance matrix, which may impose a decoding delay. In addition, our techniques involve the eigendecomposition of $4KMT \times 4KMT$ matrices, and, therefore, their computational complexity is higher than that of the LS method.

## VI. Conclusion

In this paper, two novel semiblind techniques for multiuser MIMO channel estimation have been proposed. These techniques are applicable when OSTBCs are used for data transmission. The proposed techniques are based on the extension of the concepts of the popular Capon and MUSIC methods to the problem of multiuser MIMO channel estimation. They exploit the inherent structure of OSTBCs to blindly estimate the subspace that contains the user channel matrices, and then employ only a few training blocks to extract the user channels from this subspace. Compared with the standard nonblind LS-based channel estimator, our techniques require less training blocks and, therefore, offer improvements of the bandwidth efficiency. Moreover, due to a parsimonious channel representation used in the proposed techniques, they substantially improve the channel estimation accuracy with respect to the nonblind LS channel estimation method, which does not enjoy such a parsimony.

Simulation results have shown that using the proposed Capon- and MUSIC-based semiblind channel estimators in the coherent ML and MV multiuser MIMO receivers results in a substantially improved performance as compared with the case when the standard LS channel estimates are used in these receivers.

## References

[1] V. Tarokh, N. Seshadri, and A. R. Calderbank, "Space–time codes for high data rates wireless communication: Performance criterion and code construction," *IEEE Trans. Inf. Theory*, vol. 44, no. 2, pp. 744–765, Mar. 1998.

[2] S. M. Alamouti, "A simple transmit diversity technique for wireless communications," *IEEE J. Sel. Areas Commun.*, vol. 16, no. 8, pp. 1451–1458, Oct. 1998.

[3] E. G. Larsson and P. Stoica, *Space-Time Block Coding for Wireless Communications*. Cambridge, U.K.: Cambridge Univ. Press, 2003.

[4] V. Tarokh, H. Jafarkhani, and A. R. Calderbank, "Space–time block codes from orthogonal designs," *IEEE Trans. Inf. Theory*, vol. 45, no. 5, pp. 1456–1467, Jul. 1999.

[5] G. Ganesan and P. Stoica, "Space–time block codes: A maximum SNR approach," *IEEE Trans. Inf. Theory*, vol. 47, no. 4, pp. 1650–1656, May 2001.

[6] H. Li, X. Lu, and G. B. Giannakis, "Capon multi-user receiver for CDMA systems with space–time coding," *IEEE Trans. Signal Process.*, vol. 50, no. 5, pp. 1193–1204, May 2002.

[7] D. Reynolds, X. Wang, and H. V. Poor, "Blind adaptive space–time multi-user detection with multiple transmitter and receiver antennas," *IEEE Trans. Signal Process.*, vol. 50, no. 6, pp. 1261–1276, Jun. 2002.

[8] A. F. Naguib, N. Seshadri, and A. R. Calderbank, "Applications of space–time block codes and interference suppression for high capacity and high data rate wireless systems," in *Proc. 32nd Asilomar Conf. Signals, Systems, Computers*, Pacific Grove, CA, Nov. 1998, pp. 1803–1810.

[9] S. Shahbazpanahi, M. Beheshti, A. B. Gershman, M. Gharavi-Alkhansari, and K. M. Wong, "Minimum variance linear receivers for multiaccess MIMO wireless systems with space–time block coding," *IEEE Trans. Signal Process.*, vol. 52, no. 12, pp. 3306–3313, Dec. 2004.

[10] D. Gesbert, M. Shafi, D. Shiu, P. J. Smith, and A. Naguib, "From theory to practice: An overview of MIMO space–time coded wireless systems," *IEEE J. Sel. Areas Commun.*, vol. 21, no. 3, pp. 281–302, Apr. 2003.

[11] B. L. Hughes, "Differential space–time modulation," *IEEE Trans. Inf. Theory*, vol. 46, pp. 2567–2578, Nov. 2000.

[12] H. Jafarkhani and V. Tarokh, "Multiple transmit antenna differential detection from generalized orthogonal design," *IEEE Trans. Inf. Theory*, vol. 47, no. 6, pp. 2626–2631, Sep. 2001.

[13] G. Ganesan and P. Stoica, "Differential modulation using space–time block codes," *IEEE Signal Process. Lett.*, vol. 9, no. 2, pp. 57–60, Feb. 2002.

[14] B. M. Hochwald and T. L. Marzetta, "Unitary space–time modulation for multiple-antenna communications in Rayleigh flat fading," *IEEE Trans. Inf. Theory*, vol. 46, no. 2, pp. 543–564, Mar. 2000.

[15] A. L. Swindlehurst and G. Leus, "Blind and semi-blind equalization for generalized space–time block codes," *IEEE Trans. Signal Process.*, vol. 50, no. 10, pp. 2489–2498, Oct. 2002.

[16] J. Choi, "Equalization and semi-blind channel estimation for space–time block coded signals over a frequency-selective fading channel," *IEEE Trans. Signal Process.*, vol. 52, no. 3, pp. 774–785, Mar. 2004.

[17] S. Zhou, B. Muquet, and G. B. Giannakis, "Subspace-based (semi-) blind channel estimation for block precoded space–time OFDM," *IEEE Trans. Signal Process.*, vol. 50, no. 5, pp. 1215–1228, May 2002.

[18] N. D. Sidiropoulos and R. S. Budampati, "Khatri–Rao space–time codes," *IEEE Trans. Signal Process.*, vol. 50, no. 10, pp. 2396–2407, Oct. 2002.

[19] P. Stoica and G. Ganesan, "Space–time block codes: Trained, blind, and semi-blind detection," *Digit. Signal Process.*, vol. 13, no. 1, pp. 93–105, Jan. 2003.

[20] E. Larsson, P. Stoica, and J. Li, "Space–time block codes: ML detection for unknown channels and unstructured interference," *IEEE Trans. Signal Process.*, vol. 51, no. 2, pp. 362–372, Feb. 2003.

[21] S. Shahbazpanahi, A. B. Gershman, and J. H. Manton, "Closed-form blind MIMO channel estimation for orthogonal space–time block codes," *IEEE Trans. Signal Process.*, vol. 53, no. 12, pp. 4506–4517, Dec. 2005.

[22] G. Klang and B. Ottersten, "Structured semi-blind interference rejection in dispersive multichannel systems," *IEEE Trans. Signal Process.*, vol. 50, no. 8, pp. 2027–2036, Aug. 2002.

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

[24] R. O. Schmidt, "Multiple emitter location and signal parameter estimation," in *Proc. RADC Spectral Estim. Workshop*, Rome, NY, 1979, pp. 234–258.

[25] A. Dogandzic and A. Nehorai, "Space–time fading channel estimation and symbol detection in unknown spatially correlated noise," *IEEE Trans. Signal Process.*, vol. 50, no. 3, pp. 457–474, Mar. 2002.

[26] T. Ostman, S. Parkvall, and B. Ottersten, "An improved MUSIC algorithm for estimation of time delays in asynchronous DS-CDMA systems," *IEEE Trans. Commun.*, vol. 47, no. 11, pp. 1628–1631, Nov. 1999.

[27] Z. Liu and G. B. Giannakis, "Space–time block coded multiple access through frequency-selective fading channels," *IEEE Trans. Commun.*, vol. 49, no. 6, pp. 1033–1044, Jun. 2001.

[28] M. Gharavi-Alkhansari and A. B. Gershman, "Constellation space invariance of orthogonal space–time block codes," *IEEE Trans. Inf. Theory*, vol. 51, no. 1, pp. 331–334, Jan. 2005.

[29] E. Agrell, T. Eriksson, A. Vardy, and K. Zeger, "Closest point search in lattices," *IEEE Trans. Inf. Theory*, vol. 48, no. 8, pp. 2201–2214, Aug. 2002.

[30] S. A. Vorobyov, A. B. Gershman, and Z.-Q. Luo, "Robust adaptive beamforming using worst-case performance optimization: A solution to the signal mismatch problem," *IEEE Trans. Signal Process.*, vol. 51, no. 2, pp. 313–324, Feb. 2003.

[31] Y. Rong, S. Shahbazpanahi, and A. B. Gershman, "Robust linear receivers for space–time block coded multi-access MIMO systems with imperfect channel state information," *IEEE Trans. Signal Process.*, vol. 53, no. 8, pp. 3081–3090, Aug. 2005.

[32] S. Shahbazpanahi, A. B. Gershman, and J. Manton, "A linear precoding approach to resolve ambiguity of blind decoding of orthogonal space–time block codes in slowly fading channels," in *Proc. IEEE Workshop Signal Processing Advances Wireless Communications*, Lisbon, Portugal, Jul. 2004, pp. 228–232.

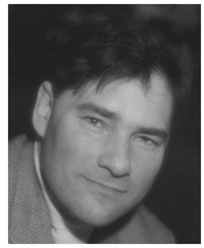
[33] B. Friedlander, "A sensitivity analysis of the MUSIC algorithm," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 38, no. 10, pp. 1740–1751, Oct. 1990.

[34] A. L. Swindlehurst and T. Kailath, "A performance analysis of subspace-based methods in the presence of model errors, Part I: The MUSIC algorithm," *IEEE Trans. Signal Process.*, vol. 40, no. 7, pp. 1758–1774, Jul. 1992.

[35] C. Vaidyanathan and K. M. Buckley, "Performance analysis of the MVDR spatial spectrum estimator," *IEEE Trans. Signal Process.*, vol. 43, no. 6, pp. 1427–1437, Jun. 1995.

**Shahram Shahbazpanahi** (M'02) was born in Sanandaj, Kurdistan, Iran. He received the B.Sc., M.Sc., and Ph.D. degrees in electrical engineering from Sharif University of Technology, Tehran, Iran, in 1992, 1994, and 2001, respectively.

From September 1994 to September 1996, he was a faculty member with the Department of Electrical Engineering, Razi University, Kermanshah, Iran. From July 2001 to March 2003, he was a Postdoctoral Fellow at the Department of Electrical and Computer Engineering, McMaster University, Hamilton, ON, Canada. From April 2003 to September 2004, he was a Visiting Researcher with the Department of Communication Systems, University of Duisburg-Essen, Duisburg, Germany. From September 2004 to April 2005, he was a Lecturer and Adjunct Professor with the Department of Electrical and Computer Engineering, McMaster University. Since July 2005, he has joined the Faculty of Engineering and Applied Science, University of Ontario Institute of Technology, Oshawa, ON, Canada, where he holds an Assistant Professor position. His research interests include statistical and array signal processing, space–time adaptive processing, detection and estimation, smart antennas, spread-spectrum techniques, MIMO communications, DSP programming, and hardware/real-time software design for telecommunication systems.



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

From 1984 to 1999, he held several full-time and visiting research appointments in Russia, Switzerland, and Germany. In 1999, he joined the Department of Electrical and Computer Engineering, McMaster University, Hamilton, ON, Canada, where he became a Professor in 2002. From 2003 to 2005, he held a visiting professorship at the University of Duisburg-Essen, Duisburg, Germany. Since April 2005, he has been a Professor of communication systems with the Darmstadt University of Technology, Darmstadt, Germany. His research interests are in the area of signal processing and communications with the primary emphasis on array processing, beamforming, MIMO and multiuser communications, and estimation and detection theory. He has co-edited two books and (co)authored several book chapters and more than 100 journal and 150 conference papers on these subjects.

Dr. Gershman is the recipient of several awards, including the 2004 IEEE Signal Processing Society Best Paper Award; the 2002 Young Explorers Prize from the Canadian Institute for Advanced Research (CIAR); the 2001 Wolfgang Paul Award from the Alexander von Humboldt Foundation, Germany; and the 2000 Premier's Research Excellence Award, Ontario, Canada. He is currently Editor-in-Chief for the IEEE SIGNAL PROCESSING LETTERS and is on Editorial Boards of the *EURASIP Journal on Wireless Communications and Networking* and the *EURASIP Signal Processing*. He was Associate Editor of the IEEE TRANSACTIONS ON SIGNAL PROCESSING. He is a member of both the Sensor Array and Multichannel (SAM) and the Signal Processing Theory and Methods (SPTM) Technical Committees (TCs) of the IEEE Signal Processing Society. He was Technical Co-Chair of the IEEE International Symposium on Signal Processing and Information Technology (ISSPIT), Darmstadt, Germany, December 2003; General Co-Chair of the First IEEE Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP), Puerto Vallarta, Mexico, December 2005; and Technical Co-Chair of the Fourth IEEE Sensor Array and Multichannel Signal Processing Workshop, Waltham, MA, June 2006.

**Georgios B. Giannakis** (F'97) received the Diploma degree in electrical engineering from the National Technical University of Athens, Greece, in 1981 and the M.Sc. degree in electrical engineering, the M.Sc. degree in mathematics, and the Ph.D. degree in electrical engineering all from the University of Southern California (USC), Los Angeles, in 1983, 1986, and 1986, respectively.

After lecturing for one year at USC, he joined the University of Virginia, Charlottesville, in 1987, where he became a Professor of electrical engineering in 1997. Since 1999 he has been a Professor with the Department of Electrical and Computer Engineering at the University of Minnesota, Minneapolis, where he now holds an ADC Chair in Wireless Telecommunications. His general interests span the areas of communications and signal processing, estimation and detection theory, time-series analysis, and system identification—subjects on which he has published more than 220 journal papers, 380 conference papers, and two edited books. His current research focuses on transmitter and receiver diversity techniques for single-user and multiuser fading communication channels, complex-field and space–time coding, multicarrier, ultra-wideband wireless communication systems, cross-layer designs, and sensor networks.

Dr. Giannakis is the (co)recipient of six paper awards from the IEEE Signal Processing (SP) and Communications Societies (1992, 1998, 2000, 2001, 2003, and 2004). He also received Technical Achievement Awards from the Signal Processing Society in 2000 and from EURASIP in 2005. He served as Editor-in-Chief for the IEEE SIGNAL PROCESSING LETTERS, as Associate Editor for the IEEE TRANSACTIONS ON SIGNAL PROCESSING and the IEEE SIGNAL PROCESSING LETTERS, as Secretary of the Signal Processing Conference Board, as Member of the Signal Processing Publications Board, as Member and Vice-Chair of the Statistical Signal and Array Processing (SSAP) Technical Committee (TC), as Chair of the Signal Processing for Communications (SPCOM) TC, and as a Member of the IEEE Fellows Election Committee. He has also served as a Member of the IEEE Signal Processing Society's Board of Governors, the Editorial Board for the PROCEEDINGS OF THE IEEE, and the steering committee of the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS.