SEMI-BLIND MULTI-USER MIMO CHANNEL ESTIMATION BASED ON CAPON AND MUSIC TECHNIQUES

Shahram Shahbazpanahi[*] *Alex B. Gershman*[*] *Georgios B. Giannakis*[†]

[*]Dept. Electrical and Computer Engineering, McMaster University, Hamilton, ON, Canada
[†]Dept. of Electrical and Computer Engineering, University of Minnesota, Minneapolis, MN, US

ABSTRACT

We consider the problem of simultaneous estimation of the channel state information (CSI) of several transmitters that use orthogonal space-time block codes to communicate with a single receiver. Based on the generalizations of the Capon and MUSIC techniques, we propose two novel algorithms to estimate multi-user MIMO channels. These algorithms estimate the subspace spanned by the user channels blindly and use only a few training blocks to extract the users CSI from this subspace.

1. INTRODUCTION

Point-to-point communications based on orthogonal space-time block codes (OSTBC) have been extensively studied in the literature [1], [2]. However, there are only a few papers where the problem of joint decoding and multi-access interference (MAI) suppression in OSTBC-based multiuser communication systems has been addressed [3]-[5]. 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 reduces the bandwidth efficiency and, therefore, blind and/or semi-blind channel estimation techniques are of great interest.

In this paper, we propose two novel methods to simultaneously estimate the CSI of several transmitters that use OSTBCs to communicate with a single receiver. Our approach is based on the generalizations of the Capon and MUSIC techniques to the problem of channel estimation in multiuser MIMO communications. It will be shown that using such generalization, the subspace spanned by the user channels can be estimated in a fully blind way. However, to extract the user CSI from this subspace, a few training blocks have to be used with the amount of training that is much smaller than that required by the standard non-blind LS-based channel estimation method. Our simulation results show that, in addition to an improved bandwidth efficiency, the proposed techniques also have substantially lower channel estimation errors as compared to the standard LS channel estimation method. Moreover, it is shown that using the proposed semi-blind channel estimators in conjunction with the minimum variance (MV) linear multiuser receiver of [5] improves its performance compared to the case when it uses the standard non-blind LS channel estimator.

2. BACKGROUND

We assume that P synchronous multi-antenna transmitters communicate with a single multi-antenna receiver. All transmitters are assumed to use the same OSTBC to encode the information symbols and to have the same number of antennas, N. The receiver is also assumed to have M antennas. The flat block-fading channel is considered. Based on these assumptions, the received signal is given by

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

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 pth user, $\mathbf{X}(\mathbf{s}_p)$ is the $T \times N$ matrix of its transmitted signals, $\mathbf{H}_p$ is the $N \times M$ matrix of channel coefficients between the pth 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[1].

The matrix $\mathbf{X}(\mathbf{s}_p)$ is assumed to correspond to OSTBCs [2]. The $T \times N$ 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, $(\cdot)^H$ stands for the Hermitian transpose, and $\|\cdot\|$ denotes the Euclidean norm. The matrix $\mathbf{X}(\mathbf{s})$ can then be written as

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

Here, the matrices $\mathbf{C}_k$ and $\mathbf{D}_k$ are defined, respectively, as $\mathbf{C}_k \triangleq \mathbf{X}(\mathbf{e}_k)$ and $\mathbf{D}_k \triangleq \mathbf{X}(j\mathbf{e}_k)$, where $j = \sqrt{-1}$ and $\mathbf{e}_k$ is the $K \times 1$ vector having one in the kth position and zeros elsewhere. Using (1), one can rewrite (2) as [5]

$$\underline{\mathbf{Y}} = \sum_{p=1}^{P} \mathbf{A}(\mathbf{h}_p)\, \underline{\mathbf{s}_p} + \underline{\mathbf{V}} \tag{2}$$

where the "underline" operator for any matrix $\mathbf{P}$ is defined as $\underline{\mathbf{P}} \triangleq \left[\mathrm{vec}^T\{\mathrm{Re}(\mathbf{P})\} \;\; \mathrm{vec}^T\{\mathrm{Im}(\mathbf{P})\}\right]^T$ where $\mathrm{vec}\{\cdot\}$ is the vectorization operator stacking all columns of a matrix on top of each other, $\mathrm{Re}(\cdot)$ and $\mathrm{Im}(\cdot)$ denote, respectively, the real and the imaginary parts, and $(\cdot)^T$ stands for the transpose. In (2), the channel vector of the pth user is defined as $\mathbf{h}_p \triangleq \underline{\mathbf{H}_P}$ and the $2MT \times 2K$ real matrix $\mathbf{A}(\mathbf{h}_p)$ is given by

$$\begin{aligned} \mathbf{A}(\mathbf{h}_p) &\triangleq \left[\underline{\mathbf{C}_1\mathbf{H}_p} \cdots \underline{\mathbf{C}_K\mathbf{H}_p} \;\; \underline{\mathbf{D}_1\mathbf{H}_p} \cdots \underline{\mathbf{D}_K\mathbf{H}_p}\right] \\ &\triangleq \left[\mathbf{a}_1(\mathbf{h}_p)\; \mathbf{a}_2(\mathbf{h}_p)\; \cdots\; \mathbf{a}_{2K}(\mathbf{h}_p)\right] . \end{aligned} \tag{3}$$

This matrix has an important property that its columns have the same norms and are orthogonal to each other [7]:

$$\mathbf{A}^T(\mathbf{h}_p)\mathbf{A}(\mathbf{h}_p) = \|\mathbf{h}_p\|^2 \mathbf{I}_{2K} . \tag{4}$$

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

$$\mathbf{a}_k(\mathbf{h}_p) = \mathbf{\Phi}_k \mathbf{h}_p \quad \text{for} \quad k = 1, \ldots, 2K . \tag{5}$$

[1]This means that, with probability one, there does not exist any non-zero 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 (5), one can write

$$\mathrm{vec}\{\mathbf{A}(\mathbf{h}_p)\} = \mathbf{\Phi}\mathbf{h}_p \tag{6}$$

where the $4KMT \times 2MN$ matrix $\mathbf{\Phi} \triangleq [\mathbf{\Phi}_1^T \cdots \mathbf{\Phi}_{2K}^T]^T$.

Although the multiuser MIMO channel estimation techniques considered below can be used in conjunction with any multiuser receiver (including the computationally expensive ML receiver), we will use them in conjunction with the MV linear receiver of [5] because it is computationally much simpler than the ML receiver.

Any linear receiver for detecting the transmitted symbols of the pth user can represented by the weight matrix $\mathbf{W}(\mathbf{h}_p)$ that is applied to the received vectorized data $\underline{\mathbf{Y}}$ to estimate $\underline{\mathbf{s}_p}$ as $\underline{\hat{\mathbf{s}}_p} = \mathbf{W}^T(\mathbf{h}_p)\underline{\mathbf{Y}}$. 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 kth transmitted symbol should be detected as a point in the corresponding constellation which is the closest one to the kth element of $\hat{\mathbf{s}}_p$.

For the pth transmitter, the MV linear receiver of [5] is given by

$$\mathbf{W}(\mathbf{h}_p) = \mathbf{R}_d^{-1}\mathbf{A}(\mathbf{h}_p)(\mathbf{A}^T(\mathbf{h}_p)\mathbf{R}_d^{-1}\mathbf{A}(\mathbf{h}_p))^{-1} \tag{7}$$

where $\mathbf{R} \triangleq \mathrm{E}\{\underline{\mathbf{Y}}\underline{\mathbf{Y}}^T\}$ is the covariance matrix of the vectorized data, $\hat{\mathbf{R}}$ is the sample estimate of $\mathbf{R}$, $\mathbf{R}_d \triangleq \hat{\mathbf{R}} + \gamma\mathbf{I}$ is the diagonally loaded sample covariance matrix, and γ is the loading factor. It has been shown in [5] that for OSTBCs, the matrix $\mathbf{A}^T(\mathbf{h}_p)\mathbf{R}_d^{-1}\mathbf{A}(\mathbf{h}_p)$ in (7) is invertible regardless of the channel vector value.

Note that the "clairvoyant" linear receiver in (7) assumes that the true channel vector $\mathbf{h}_p$ is available. In practice, such an assumption can hardly be met. Hence, the linear receiver in (7) does not correspond to any practical scenario and can only be used for comparison purposes. In practice, one has to estimate $\mathbf{h}_p$ and to use this estimate in (7).

3. MULTIUSER CHANNEL ESTIMATION

3.1. Standard LS-Based Technique

To recover the user transmitted symbols, one needs to estimate the user channel vectors $\{\mathbf{h}_p\}_{p=1}^P$. A straightforward approach to estimate these vectors is to employ training and to use the LS approach to obtain $\{\mathbf{h}_p\}_{p=1}^P$. To show this, let us rewrite (2) as

$$\underline{\mathbf{Y}} = \sum_{p=1}^{P} \tilde{\mathbf{A}}(\underline{\mathbf{s}_p})\mathbf{h}_p + \underline{\mathbf{V}} \tag{8}$$

where $\tilde{\mathbf{A}}(\underline{\mathbf{s}_p})$ is a $2MT \times 2MN$ matrix whose kth 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 kth position and zeros elsewhere but now its dimension is $2MN \times 1$. Assuming that each user transmits J_t training blocks and using the data model in (8), we can write

$$\mathbf{y}(n) \triangleq \underline{\mathbf{Y}(n)} = \sum_{p=1}^{P} \tilde{\mathbf{A}}(\underline{\mathbf{s}_p(n)})\mathbf{h}_p + \underline{\mathbf{V}(n)}, \quad n = 1, \ldots, J_t \tag{9}$$

where $\mathbf{s}_p(n)$ is the nth known vector transmitted by the pth user, and $\mathbf{Y}(n)$ and $\mathbf{V}(n)$ are, respectively, the received signal matrix and the noise matrix.

Using the notations $\mathbf{g} \triangleq [\mathbf{h}_1^T \;\; \mathbf{h}_2^T \;\; \ldots \;\; \mathbf{h}_p^T]^T$ and $\boldsymbol{\mathcal{A}}(n) \triangleq [\tilde{\mathbf{A}}(\underline{\mathbf{s}_1(n)}) \;\; \tilde{\mathbf{A}}(\underline{\mathbf{s}_2(n)}) \;\; \cdots \;\; \tilde{\mathbf{A}}(\underline{\mathbf{s}_P(n)})]$, we can rewrite (9) as

$$\mathbf{y}(n) = \boldsymbol{\mathcal{A}}(n)\mathbf{g} + \underline{\mathbf{V}(n)}, \quad n = 1, \ldots, J_t\,. \tag{10}$$

Using (10), the LS estimate of the vector $\mathbf{g}$ is given by

$$\hat{\mathbf{g}} = (\mathbb{A}^H\mathbb{A})^{-1}\mathbb{A}^H\mathbf{z} \tag{11}$$

where $\mathbf{z} \triangleq [\mathbf{y}^T(1) \cdots \mathbf{y}^T(J_t)]^T$ and $\mathbb{A} \triangleq [\boldsymbol{\mathcal{A}}(1)^T \cdots \boldsymbol{\mathcal{A}}(J_t)^T]^T$.

It is noteworthy that, according to (11), one has to ensure that the matrix $\mathbb{A}$ is full column rank. Since the dimension of $\mathbb{A}$ is $2MTJ_t \times 2PMN$, the full column rank condition of $\mathbb{A}$ implies that $J_t \geq PN/T$ has to be chosen.

3.2. Capon-Based Technique

In application to our problem, the Capon linear receiver can be interpreted as a sort of spatio-temporal filter which passes the signal of a hypothetical user with the channel vector $\mathbf{h}$ without any distortion while maximally rejecting the signals of other users. More specifically, to linearly estimate the kth entry of some $2K \times 1$ real vector $\underline{\mathbf{s}}$ (which belongs to a hypothetical transmitter with a 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 [3], [5]:

$$\min_{\mathbf{w}_k} \mathbf{w}_k^T\mathbf{R}\mathbf{w}_k \quad \text{subject to} \quad \mathbf{w}_k^T\mathbf{a}_k(\mathbf{h}) = 1\,. \tag{12}$$

The solution to (12) is well known to be

$$\mathbf{w}_k(\mathbf{h}) = (\mathbf{a}_k^T(\mathbf{h})\mathbf{R}^{-1}\mathbf{a}_k(\mathbf{h}))^{-1}\,\mathbf{R}^{-1}\mathbf{a}_k(\mathbf{h})\,. \tag{13}$$

Note that, according to (13), a separate weight vector has to be obtained for each $k = 1, \ldots, 2K$ (i.e., $2K$ linear receivers have to be designed to estimate all entries of $\underline{\mathbf{s}}$).

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

$$P_{\mathrm{C}}^k(\mathbf{h}) \triangleq \mathbf{w}_k^T(\mathbf{h})\mathbf{R}\mathbf{w}_k(\mathbf{h}) = (\mathbf{a}_k^T(\mathbf{h})\mathbf{R}^{-1}\mathbf{a}_k(\mathbf{h}))^{-1}\,. \tag{14}$$

Since we defined the Capon spectrum in (14) as the output power of the kth linear receiver, it is expected to have maxima for $\mathbf{h} = \mathbf{h}_p/\|\mathbf{h}_p\|$ for all $p = 1, \ldots P$. Therefore, our goal will be to find the values of $\mathbf{h}$ which maximize the Capon function in (14). However, as $\mathbf{a}_k(\mathbf{h})$ is linear in $\mathbf{h}$, the value of Capon spectrum can increase arbitrarily when $\mathbf{h} \to 0$. To avoid such a trivial solution, we assume that $\|\mathbf{h}\| = 1$. Although any of the Capon spectra $\{P_{\mathrm{C}}^k(\mathbf{h})\}_{k=1}^{2K}$ can be used to estimate the channel vectors, we combine all of them as

$$Q_C(\mathbf{h}) \triangleq \sum_{k=1}^{2K}(P_C^k(\mathbf{h}))^{-1} = \mathbf{h}^T\Big(\sum_{k=1}^{2K}\mathbf{\Phi}_k^T\mathbf{R}^{-1}\mathbf{\Phi}_k\Big)\mathbf{h} \tag{15}$$

where $Q_C(\mathbf{h})$ can be viewed as null-spectrum.

The definition of $Q_C(\mathbf{h})$ in (15) allows us to find the channel estimates in a closed form. To show this, we note that if all Capon spectra $\{P_{\mathrm{C}}^k(\mathbf{h})\}_{k=1}^{2K}$ have a maximum for $\mathbf{h} = \mathbf{h}_p/\|\mathbf{h}_p\|$, then $Q_C(\mathbf{h})$ will have a minimum for $\mathbf{h} = \mathbf{h}_p/\|\mathbf{h}_p\|$. Hence, we propose to estimate the *normalized* user channel vectors as the P values of $\mathbf{h}$ which minimize $Q_C(\mathbf{h})$. Therefore, 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 P minor eigenvectors of the matrix $\mathbf{\Psi} \triangleq \sum_{k=1}^{2K}\mathbf{\Phi}_k^T\mathbf{R}^{-1}\mathbf{\Phi}_k$. Denoting the P minor eigenvectors of this matrix as $\mathbf{u}_k$ ($k = 1, \ldots, P$), we have

$$\mathbf{h}_p = \sum_{k=1}^{P}\alpha_{pk}\mathbf{u}_k\,. \tag{16}$$

Note that to find the vectors $\{\mathbf{u}_k\}_{k=1}^P$, only the knowledge of the data covariance matrix $\mathbf{R}$ is required. This matrix can be estimated without any training data. However, to determine the real coefficients α_{pk}, a few training data blocks have to be used. In the sequel, we show how these coefficients can be obtained based on the LS approach.

Suppose that each transmitter sends a total number of J blocks to the receiver and assume that, out of these J blocks, the first J_t training blocks are known at the receiver while the remaining $J - J_t$ blocks are used to convey the information symbols. One can then 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 $\{\mathbf{u}_k\}_{k=1}^P$ as the P minor eigenvectors $\{\hat{\mathbf{u}}_k\}_{k=1}^P$ of the matrix $\hat{\mathbf{\Psi}} \triangleq \sum_{k=1}^{2K}\mathbf{\Phi}_k^T\hat{\mathbf{R}}^{-1}\mathbf{\Phi}_k$. Then, using (16), we have

$$\mathbf{A}(\hat{\mathbf{h}}_p) = \sum_{p=1}^{P}\alpha_{pk}\mathbf{A}(\hat{\mathbf{u}}_k) \tag{17}$$

which follows from the linearity of $\mathbf{A}(\cdot)$ in its argument. Replacing $\mathbf{A}(\mathbf{h}_p)$ in (9) with $\mathbf{A}(\hat{\mathbf{h}}_p)$ of (17), we can write (9) as

$$\mathbf{y}(n) = \sum_{k=1}^{P}\mathbf{B}_k(n)\boldsymbol{\alpha}_k + \underline{\mathbf{V}(n)} = \mathbf{B}(n)\boldsymbol{\alpha} + \underline{\mathbf{V}(n)} \tag{18}$$

where $\mathbf{B}_k(n) \triangleq [\mathbf{A}(\hat{\mathbf{u}}_k)\underline{\mathbf{s}_1(n)}\ \mathbf{A}(\hat{\mathbf{u}}_k)\underline{\mathbf{s}_2(n)}\ \cdots\ \mathbf{A}(\hat{\mathbf{u}}_k)\underline{\mathbf{s}_P(n)}]$, $\mathbf{B}(n) \triangleq [\mathbf{B}_1(n)\ \cdots\ \mathbf{B}_P(n)]$, $\mathbb{B} \triangleq [\mathbf{B}^T(1)\ \cdots\ \mathbf{B}^T(J_t)]^T$, $\boldsymbol{\alpha}_k q \triangleq [\alpha_{1,k}\ \cdots\ \alpha_{P,k}]^T$, and $\boldsymbol{\alpha} \triangleq [\boldsymbol{\alpha}_1^T\ \cdots\ \boldsymbol{\alpha}_P^T]^T$. Using these notations and employing (18), the LS estimate of $\boldsymbol{\alpha}$ can be written as

$$\hat{\boldsymbol{\alpha}} = (\mathbb{B}^H\mathbb{B})^{-1}\mathbb{B}^H\mathbf{z}\,. \tag{19}$$

Once the estimate of $\boldsymbol{\alpha}$ is found, one can obtain the channel estimate of the pth user by means of (16).

To guarantee the uniqueness of the LS solution for $\boldsymbol{\alpha}$, the $2MTJ_t \times P^2$ matrix $\mathbb{B}$ has to be full column rank, i.e. $J_t \geq P^2/(2MT)$ has to be chosen.

3.3. MUSIC-Based Technique

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 that orthogonal to the signal subspace. Note that the dimension of the signal subspace is, at most, $2KP$ and the dimension of the noise subspace is, at least, $2MT - 2KP$. To ensure that the noise subspace is non-degenerate, 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.

Denoting the dimension of the signal subspace as d ($d \leq 2KP$), it can be easily proven that the signal subspace is spanned by the d principal eigenvectors of the data covariance matrix $\mathbf{R}$ and the rest $2MT - d$ eigenvectors of this matrix span the noise subspace. Based on this fact, we can write

$$\mathbf{A}^T(\mathbf{h}_p)\mathbf{E} = \mathbf{0} \quad \text{for} \quad p = 1, \ldots, P \tag{20}$$

where $\mathbf{E}$ is a is a $2MT \times (2MT - 2KP)$ matrix whose columns are the eigenvectors of the data covariance matrix $\mathbf{R}$ corresponding to the $2MT - 2KP$ smallest eigenvalues.

In practice, $\mathbf{E}$ can be estimated through the eigendecomposition of the sample covariance matrix $\hat{\mathbf{R}}$. Having such an estimate $\hat{\mathbf{E}}$ of $\mathbf{E}$, let us define the generalized MUSIC "spectrum" as

$$P_{\text{MUSIC}}(\mathbf{h}) = 1/\|\mathbf{A}^T(\mathbf{h})\hat{\mathbf{E}}\|_F^2 \tag{21}$$

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

Note that the MUSIC spectrum defined in (21) can be further simplified as

$$P_{\text{MUSIC}}(\mathbf{h}) \triangleq \frac{1}{\text{tr}\{\mathbf{A}^T(\mathbf{h})\hat{\mathbf{E}}\hat{\mathbf{E}}^T\mathbf{A}(\mathbf{h})\}} = \frac{1}{\mathbf{h}^T\mathbf{\Phi}^T(\mathbf{I}_{2K}\otimes\hat{\mathbf{E}}\hat{\mathbf{E}}^T)\mathbf{\Phi}\mathbf{h}}$$

where $\text{tr}\{\cdot\}$ denotes the trace operator, $\otimes$ stands for the Kronecker product, and (6) has been used to obtain the last line of (22). From (22), we see that the channel vector estimates belong to the subspace spanned by the P eigenvectors of the matrix $\mathbf{\Gamma} \triangleq \mathbf{\Phi}^T(\mathbf{I}_{2K}\otimes\hat{\mathbf{E}}\hat{\mathbf{E}}^T)\mathbf{\Phi}$ that correspond to the P smallest eigenvalues. Denoting the P minor eigenvectors of the matrix $\mathbf{\Gamma}$ as $\mathbf{v}_k$ ($k = 1, \ldots, P$), one can write

$$\hat{\mathbf{h}}_p = \sum_{k=1}^{P}\beta_{pk}\mathbf{v}_k\,. \tag{22}$$

Using the training procedure similar to that developed in the previous subsection for the Capon-based channel estimator, the coefficients β_{pk} ($k, p = 1, \ldots, P$) can be estimated as

$$\hat{\boldsymbol{\beta}} = (\mathbb{F}^H\mathbb{F})^{-1}\mathbb{F}^H\mathbf{z} \tag{23}$$

where $\boldsymbol{\beta} \triangleq [\boldsymbol{\beta}_1^T\ \boldsymbol{\beta}_2^T\ \cdots\ \boldsymbol{\beta}_P^T]^T$, $\boldsymbol{\beta}_k \triangleq [\beta_{1,k}\ \beta_{2,k}\ \cdots\ \beta_{P,k}]^T$, $\mathbb{F} \triangleq [\mathbf{F}^T(1)\ \cdots\ \mathbf{F}^T(J_t)]^T$, $\mathbf{F}(n) \triangleq [\mathbf{F}_1(n)\ \cdots\ \mathbf{F}_P(n)]$, and $\mathbf{F}_k(n) \triangleq [\mathbf{A}(\mathbf{v}_k)\underline{\mathbf{s}_1(n)}\ \cdots\ \mathbf{A}(\mathbf{v}_k)\underline{\mathbf{s}_P(n)}]$.

Once $\boldsymbol{\beta}$ is estimated by means of (23), the channel vector estimate of the pth user can be found from (22) by using the proper elements of $\hat{\boldsymbol{\beta}}$ in this equation.

In order to have a unique LS estimate for the vector $\boldsymbol{\beta}$, the $2MTJ_t \times P^2$ matrix $\mathbb{F}$ has to be full column rank, and hence, $J_t \geq P^2/(2MT)$ has to be chosen. Therefore, the minimal value of J_t for the MUSIC-based estimator is the same as for the Capon-based estimator.

It should be noted that the proposed Capon and MUSIC-based techniques can exploit all received data blocks to estimate the basis for the subspace spanned by the channel vectors $\{\mathbf{h}_p\}_{p=1}^P$. This is a parsimonious approach, because after obtaining such a basis, one needs to determine only the coefficient vectors $\boldsymbol{\alpha}$ or $\boldsymbol{\beta}$ rather than all the channel vectors $\{\mathbf{h}_p\}_{p=1}^P$. Such a reduction in the number of parameters allows us to reduce the number of training blocks compared to the case where the channel vectors $\{\mathbf{h}_p\}_{p=1}^P$ are estimated directly. Indeed, for the Capon and MUSIC-based techniques, the minimum value for J_t is $P^2/(2MT)$ while for the standard LS approach the minimal value of J_t is PN/T. Hence, if $P < 2MN$, then the minimum number of training blocks in our Capon and MUSIC-based channel estimators is $2NM/P$ times smaller than when the standard LS channel estimator is used. That is why we call our techniques *semi-blind*.

Note that in application to array processing, due to uncertainties in the array manifold (such as calibration errors, distorted array shape, propagation mismatches, etc.), the MUSIC technique may break down. However, in application to the MIMO multiuser channel estimation, the MUSIC method does not suffer from such uncertainties because for any value of $\tilde{\mathbf{h}}$, the matrix $\mathbf{A}(\tilde{\mathbf{h}})$ is exactly determined by the underlying OSTBC.

4. SIMULATIONS

Throughout our simulations, the signal-to-noise ratio (SNR) for the pth user is defined as σ_p^2/σ^2 where σ_p^2 is the variance of each complex entry of $\mathbf{H}_p$ and σ^2 is the noise power. In each 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, i.e., Rayleigh fading is considered. The 3/4 rate ($K = 3$, $T = 4$) amicable design-based OSTBC of [7] is used by the transmitters as the underlying OSTBC to encode the information symbols. All the results are averaged over 100 simulation runs. In all our examples, only $J_t = 5$ training blocks are used and the sample covariance matrix is computed using $J = 300$ blocks. The diagonal loading factor $\gamma = 10\sigma^2$ is assumed.

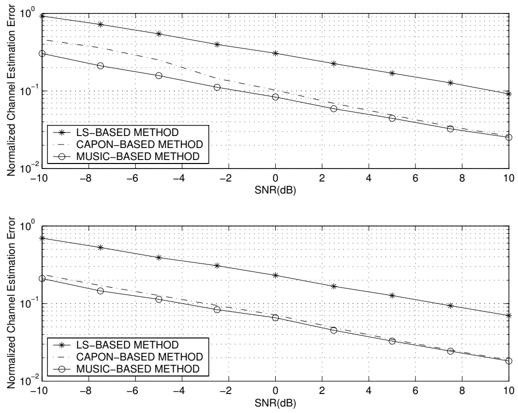

Fig. 1. Normalized channel estimation error versus the SNR for the weaker (top) and stronger (bottom) transmitter. First example.

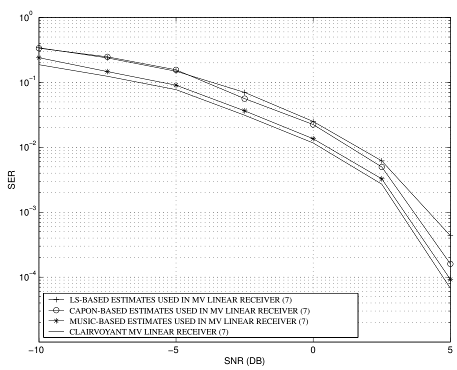

Fig. 2. SER versus SNR for the weaker transmitter. First example.

In the first example, we consider $P = 2$ transmitters with $N = 4$ and a receiver with $M = 4$ antennas. We assume that the SNR of one of the users is 2.5 dB smaller than that of the other user. In this example, we compare the performance of three different multiuser MIMO channel estimation techniques: the standard LS approach, the Capon-based technique, and the MUSIC-based method. Figure 1 shows the normalized channel estimation error for different techniques and for both transmitters. As can be seen from this figure, the Capon and MUSIC-based techniques greatly outperform the standard LS-based method. It is also noteworthy that the MUSIC-based approach performs slightly better than the Capon-based algorithm. Figure 2 shows the symbol error rates (SERs) versus the SNR for the MV linear receiver (7) when it uses channel estimates obtained by different methods for both transmitters. In this figure, the SERs for the clairvoyant MV linear receiver are also shown. Note that the latter receiver does not correspond to any practical application and is used for comparison purposes only. As can be seen from Figure 2, the performance of the MUSIC-based MV linear receiver is very close to that of the clairvoyant MV linear receiver, while the Capon-based MV receiver has slightly worse performance. The LS-based MV receiver has the worst performance among the methods tested.

In the second example, we consider $P = 8$ transmitters with $N = 4$ and a receiver with $M = 8$ antennas. Note that in this example, $J_t = 5$ training blocks are not sufficient for the standard LS-based technique which cannot be used in this case.

Figure 3 shows the SERs versus the SNR for the weaker transmitter. In this figure, we compare the clairvoyant MV linear receiver with the MV receivers in (7) that use the Capon and MUSIC-based channel estimates. As can be seen from this figure, the

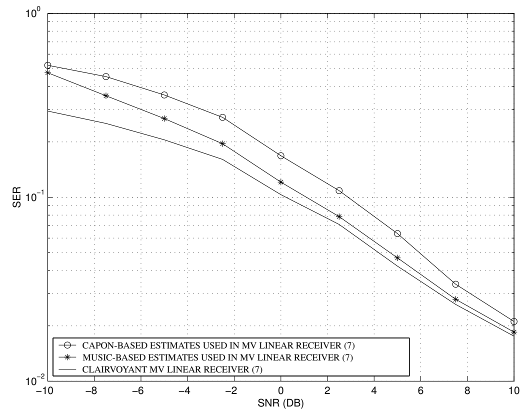

Fig. 3. The SER versus the SNR for the weaker transmitter. Second example.

MUSIC-based multiuser MIMO channel estimation method performs better than the Capon-based technique. For high values of SNR, the MUSIC-based technique has its SER quite close to that of the clairvoyant MV linear receiver.

5. CONCLUSIONS

In this paper, we proposed two novel semi-blind techniques for multiuser MIMO channel estimation that are applicable to the case when OSTBCs are used for data transmission. Our approach is based on the extension of the concepts of the Capon and MUSIC methods to the problem of multiuser MIMO channel estimation. Compared to the standard LS-based channel estimation approach, the proposed techniques provide an improved bandwidth efficiency and also improve the channel estimation accuracy at the expense of only moderate increase in complexity.

6. 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. Inform. Theory,* vol. 44, pp. 744-765, March 1998.

[2] V. Tarokh, H. Jafarkhani, and A. R. Calderbank, "Space-time block codes from orthogonal designs," *IEEE Trans. Inform. Theory,* vol. 45, pp. 1456-1467, July 1999.

[3] H. Li, X. Lu, and G. B. Giannakis, "Capon multiuser receiver for CDMA systems with space-time coding," *IEEE Trans. on Signal Processing,* vol. 50, pp. 1193-1204, May 2002.

[4] D. Reynolds, X. Wang, and H. V. Poor, "Blind adaptive space-time multiuser detection with multiple transmitter and receiver antennas," *IEEE Trans. Signal Processing,* vol. 50, pp. 1261-1276, June 2002.

[5] S. Shahbazpanahi, M. Beheshti, A. B. Gershman, M. Gharavi-Alkhansari, and K. M. Wong, "Minimum variance linear receivers for multi-access MIMO wireless systems with space-time block coding," to appear in *IEEE Trans. Signal Processing.*

[6] P. Stoica and R. Moses, *Introduction to Spectral Analysis*, Prentice-Hall, Upper Saddle River, NJ, 1997.

[7] E. G. Larsson and P. Stoica, *Space-Time Block Coding for Wireless Communications,* Cambridge University Press, Cambridge, UK, 2003.