CLOSED-FORM BLIND MIMO CHANNEL ESTIMATION FOR OSTBCS: RESOLVING AMBIGUITIES IN ROTATABLE CODES

Nima Sarmadi *Marius Pesavento*

Communication Systems Group
Technische Universität Darmstadt
64283 Darmstadt, Germany

ABSTRACT

In this paper, the problem of blind subspace-based channel estimation in multiple-input multiple-output (MIMO) systems under orthogonal space-time block coded (OSTBC) transmission is investigated. We introduce a virtual snapshot model in which the redundancies in the OSTBC are exploited to augment the received data. We show that the vector of true channel parameters is scaled version of the normalized principal eigenvector of the associated augmented data covariance matrix, which in the case of rotatable OSTBCs is not unique. We propose a simple weighting of the different virtual snapshots in the computation of a modified covariance matrix and derive general conditions that guarantee uniqueness of the channel estimates from the principal eigenvector of that matrix. Further, we prove that the blind estimation schemes of [8] and [9] can be viewed as a particular examples satisfying these uniqueness conditions. In previous works, the uniqueness of these schemes has only been concluded from simulation results but it has not been proven analytically before.

***Index Terms*—** blind MIMO channel estimation, OSTBC, covariance matching, uniqueness conditions, subspace model

1. INTRODUCTION

In MIMO wireless communication systems, space-time coding has been used to exploit the spatial diversity for improving reliability and transmission rate [1]. In particular, the popular class of orthogonal space-time block codes (OSTBCs) are known not only to maximize the diversity gain, but also to offer a simple decoding scheme provided that the channel state information (CSI) is perfectly known at the receiver [2]. In practical systems, the CSI is commonly acquired from known training symbols inserted in the transmission at the expense of a reduced bandwidth efficiency. If differential encoding schemes are applied then costly training symbols can be omitted, however, at the expense of a 3-dB performance penalty [3]. Several blind decoding and channel estimation techniques have been recently proposed in [4]-[9] that avoid the latter drawbacks.

Methods of [4]-[6] exhibit comparably high computational complexity and rely on a specific choice of the symbol constellation. Moreover, they need to be performed in a block-wise manner and therefore can not benefit from averaging over successive blocks of received data. The approximate maximum-likelihood (ML) approach of [7] is based on the iterative optimization of the ML function with respect to the unknown transmitted symbols and the desired channel parameters. However, it does not provide a closed-form solution for the channel estimate, it requires proper initialization, and its convergence can not be generally guaranteed. The approach of [8] that is solely based on second-order statistics of the received data offers low computational complexity and is independent of the symbol constellation. Nevertheless, it does not allow the unique estimation of the channel for the important class of rotatable OSTBCs [10], which includes the popular Alamouti code, without performing linear precoding. Although the proposed schemes of [9] are able to resolve these ambiguities, the uniqueness of the estimates has never been proven analytically.

Below, we generalize our recent work of [11] for the class of rotatable OSTBCs. First, we derive an augmented snapshot model based on the redundancies of the OSTBCs and show that the vector of channel parameters can be estimated as a principal eigenvector of the corresponding covariance matrix. Then, we exploit the specific properties of the rotatable OSTBCs to prove that the respective largest eigenvalue exhibits multiplicities, which generally lead to ambiguities in the estimation of the channel vector. To resolve such ambiguities, we propose a general weighing scheme of the augmented snapshots in the computation of a modified covariance matrix. In the case of rotatable OSTBCs, the latter matrix is proven to exhibit a unique principal eigenvalue from which the true channel vector can be recovered provided that specific conditions on the weighting coefficients are satisfied. Based on these results, we show that the linear precoding method of [8] and the correlation matching method of [9] can be viewed as special cases of the proposed method associated with a weighting scheme that satisfies the uniqueness conditions. The uniqueness of the latter methods has not been proven analytically before. It is noteworthy to mention that the weighting strategy proposed in this paper can be extended straightforwardly to MIMO-OFDM systems based on the coherent subcarrier processing approach of [11].

2. SIGNAL MODELS AND CODE PROPERTIES

The input-output relationship of a MIMO system with N transmit and M receive antennas can be expressed as [8]

$$\mathbf{Y}(p) = \mathbf{X}(p)\,\mathbf{H} + \mathbf{V}(p)\,, \tag{1}$$

where $\mathbf{Y}(p) \in \mathbb{C}^{T\times M}$ is the received data matrix at block index p, $\mathbf{X}(p) \in \mathbb{C}^{T\times N}$ is the code matrix of the transmitted symbols with the code length of T, $\mathbf{H} \in \mathbb{C}^{N\times M}$ is the MIMO flat fading channel matrix and $\mathbf{V}(p) \in \mathbb{C}^{T\times M}$ is the matrix of additive noise. The noise is assumed to be spatially and temporally white complex Gaussian with the variance σ^2. We consider the block fading scenario in which the coherence time of the channel is much larger than the code block length T. Consider the K complex information symbols corresponding to the p-th block of data prior to encoding are given by [8]

$$\mathbf{s}(p) \triangleq [s_1(p), s_2(p), \ldots, s_K(p)]^T,$$

This work was supported in parts by the European Research Council (ERC) Advanced Investigator Grants Program under Grant 227477-ROSE.

where $(\cdot)^T$ denotes the transpose. Also, consider that each code matrix $\mathbf{X}(p) \triangleq \mathbf{X}(\mathbf{s}(p))$ is an OSTBC-type matrix, hence, [2]

$$\mathbf{X}^H(p)\mathbf{X}(p) = \|\mathbf{s}(p)\|^2\mathbf{I}_N, \tag{2}$$

$$\mathbf{X}(p) = \sum_{k=1}^{K} \left(\mathbf{C}_k \mathrm{Re}\left(s_k(p)\right) + \mathbf{C}_{k+K} \mathrm{Im}\left(s_k(p)\right)\right), \tag{3}$$

where $(\cdot)^H$ is the conjugate transpose, $\mathbf{I}_L$ is the $L \times L$ identity matrix, $\|\cdot\|$ is the Frobenius norm of a matrix or the Euclidean norm of a vector, $\mathrm{Re}(\cdot)$ and $\mathrm{Im}(\cdot)$ represent the real and imaginary parts, and $\{\mathbf{C}_k\}_{k=1}^{2K}$ are the OSTBC basis matrices that are known at the receiver. Further, an OSTBC $\mathbf{X}(p)$ is called rotatable if there exists a code rotation matrix $\mathbf{Q} \in \mathbb{C}^{N\times N}$ such that for any $\mathbf{s}(p) \in \mathbb{C}^{K\times 1}$ [10]

$$\mathbf{X}(\mathbf{s}(p))\mathbf{Q} = \mathbf{X}(\breve{\mathbf{s}}(p)), \tag{4}$$

for some $\breve{\mathbf{s}}(p) \in \mathbb{C}^{K\times 1}$ with $\breve{\mathbf{s}}(p) \neq \pm\mathbf{s}(p)$. Otherwise, the code matrix $\mathbf{X}(p)$ is non-rotatable. Using (2) and (4), it can be readily verified that $\mathbf{Q}^H\mathbf{Q} = \mathbf{Q}\mathbf{Q}^H = \mathbf{I}_N$. In order to convert the complex-valued signal model into an equivalent real-valued one, let us define the following operator for any complex-valued matrix $\mathbf{B}$ [8]

$$\underline{\mathbf{B}} \triangleq \left[\ \mathrm{vec}\{\mathrm{Re}(\mathbf{B})\}^T,\ \mathrm{vec}\{\mathrm{Im}(\mathbf{B})\}^T\ \right]^T, \tag{5}$$

where $\mathrm{vec}\{\cdot\}$ is the vectorization operator. Making use of (3) and (5) we rewrite (1) as [8]

$$\mathbf{y}(p) = \mathbf{A}(\mathbf{h})\underline{\mathbf{s}}(p) + \mathbf{v}(p), \tag{6}$$

where $\mathbf{y}(p) \triangleq \underline{\mathbf{Y}(p)} \in \mathbb{R}^{2MT\times 1}$, $\mathbf{h} \triangleq \underline{\mathbf{H}} \in \mathbb{R}^{2MN\times 1}$, $\mathbf{v}(p) \triangleq \underline{\mathbf{V}(p)} \in \mathbb{R}^{2MT\times 1}$ and [11]

$$\mathbf{A}(\mathbf{h}) \triangleq [\mathbf{a}_1(\mathbf{h}), \ldots, \mathbf{a}_{2K}(\mathbf{h})] = [\underline{\mathbf{C}_1\mathbf{H}}, \ldots, \underline{\mathbf{C}_{2K}\mathbf{H}}]. \tag{7}$$

Using (2), it can be proved that for OSTBCs, regardless of the value of the channel vector $\mathbf{h}$, the following orthogonality property holds [8]:

$$\mathbf{A}^T(\mathbf{h})\mathbf{A}(\mathbf{h}) = \|\mathbf{h}\|^2\mathbf{I}_{2K}. \tag{8}$$

As $\mathbf{A}(\mathbf{h})$ is linear in $\mathbf{h}$, there exists a unique set of matrices $\{\mathbf{\Phi}_k\}_{k=1}^{2K} \in \mathbb{R}^{2MT\times 2MN}$ such that [11]

$$\mathbf{a}_k(\mathbf{h}) = \mathbf{\Phi}_k\mathbf{h}, \quad k = 1, \ldots, 2K, \tag{9}$$

where [11]

$$\mathbf{\Phi}_k \triangleq \begin{bmatrix} \mathrm{Re}(\mathbf{I}_M \otimes \mathbf{C}_k) & -\mathrm{Im}(\mathbf{I}_M \otimes \mathbf{C}_k) \\ \mathrm{Im}(\mathbf{I}_M \otimes \mathbf{C}_k) & \mathrm{Re}(\mathbf{I}_M \otimes \mathbf{C}_k) \end{bmatrix}, \tag{10}$$

with $\otimes$ denoting the Kronecker matrix product and [11]

$$\mathbf{\Phi}_k^T\mathbf{\Phi}_l \triangleq \begin{cases} \mathbf{I}_{2MN}, & \text{if } k = l \\ -\mathbf{\Phi}_l^T\mathbf{\Phi}_k, & \text{if } k \neq l \end{cases}. \tag{11}$$

Note that the matrices $\{\mathbf{\Phi}_k\}_{k=1}^{2K}$ are also OSTBC-specific and, hence, known to the receiver. Using (7) and (9), we have [8]

$$\mathrm{vec}\{\mathbf{A}(\mathbf{h})\} = \mathbf{\Phi}\mathbf{h}, \tag{12}$$

where the matrix $\mathbf{\Phi} \in \mathbb{R}^{4MTK\times 2MN}$ is defined as

$$\mathbf{\Phi} \triangleq \left[\mathbf{\Phi}_1^T, \ldots, \mathbf{\Phi}_{2K}^T\right]^T. \tag{13}$$

3. BLIND CHANNEL ESTIMATION

The specific structure of the OSTBCs reflected in (7) and (9) can be exploited to create a set of $2K$ *virtual snapshots* from which the weighted covariance matrix can be formed. This is of particular importance in fast fading channel scenarios in which the subspace estimates are known to severely degrade in performance. Using (11) along with the relations in (6), (7) and (9), the $2K$ virtual snapshots can be defined as [11]

$$\begin{aligned}\tilde{\mathbf{y}}(k,p) &\triangleq \mathbf{\Phi}_k^T\mathbf{y}(p) \\ &= \tilde{\mathbf{A}}_k(\mathbf{h})\underline{\mathbf{s}}(p) + \mathbf{\Phi}_k^T\mathbf{v}(p),\ k = 1, \ldots, 2K,\end{aligned} \tag{14}$$

where

$$\begin{aligned}\tilde{\mathbf{A}}_1(\mathbf{h}) &\triangleq \mathbf{\Phi}_1^T\mathbf{A}(\mathbf{h}) = \left[\mathbf{h}, \mathbf{\Phi}_1^T\mathbf{\Phi}_2\mathbf{h}, \cdots, \mathbf{\Phi}_1^T\mathbf{\Phi}_{2K}\mathbf{h}\right] \\ \tilde{\mathbf{A}}_2(\mathbf{h}) &\triangleq \mathbf{\Phi}_2^T\mathbf{A}(\mathbf{h}) = \left[\mathbf{\Phi}_2^T\mathbf{\Phi}_1\mathbf{h}, \mathbf{h}, \cdots, \mathbf{\Phi}_2^T\mathbf{\Phi}_{2K}\mathbf{h}\right] \\ &\vdots \qquad \vdots \\ \tilde{\mathbf{A}}_{2K}(\mathbf{h}) &\triangleq \mathbf{\Phi}_{2K}^T\mathbf{A}(\mathbf{h}) = \left[\mathbf{\Phi}_{2K}^T\mathbf{\Phi}_1\mathbf{h}, \mathbf{\Phi}_{2K}^T\mathbf{\Phi}_2\mathbf{h}, \cdots, \mathbf{h}\right],\end{aligned} \tag{15}$$

define the virtual signal matrices corresponding to the respective virtual snapshots. We note from (11) that for OSTBCs, the signal component $\mathbf{h}$ in (15) is orthogonal to the remaining signal components in the virtual signal matrices $\tilde{\mathbf{A}}_k(\mathbf{h})$, for $k = 1, \ldots 2K$, as $\mathbf{h}^T\mathbf{\Phi}_k^T\mathbf{\Phi}_l\mathbf{h} = -\mathbf{h}^T\mathbf{\Phi}_k^T\mathbf{\Phi}_l\mathbf{h} = 0$ for any $k \neq l$. A necessary and sufficient condition for OSTBC to be rotatable is that the code rotation matrix $\mathbf{Q}$ in (4) satisfies [10]

$$\mathbf{C}_k\mathbf{Q} = d_k\mathbf{C}_{n_k}, \qquad k = 1, \ldots, 2K, \tag{16}$$

where $d_k \in \{\pm 1\}$, and $n_k \in \{1, \ldots, 2K\}$ is an index with $n_k \neq k$ and $n_k \neq n_l$ for $k \neq l$. Hence,

$$n_k = \mathcal{M}(k), \qquad k = 1, \ldots, 2K, \tag{17}$$

where $\mathcal{M}$ represents the specific one-to-one mapping which corresponds to the specific $\mathbf{Q}$ in (16) such that the ordered index set $\{n_1, n_2, \ldots, n_{2K}\}$ describes the permutation corresponding to the ordered index set $\{1, 2, \ldots, 2K\}$. Let us define

$$\breve{\mathbf{Q}} \triangleq \begin{bmatrix} \mathrm{Re}(\mathbf{I}_M \otimes \mathbf{Q}) & -\mathrm{Im}(\mathbf{I}_M \otimes \mathbf{Q}) \\ \mathrm{Im}(\mathbf{I}_M \otimes \mathbf{Q}) & \mathrm{Re}(\mathbf{I}_M \otimes \mathbf{Q}) \end{bmatrix}, \tag{18}$$

such that

$$\breve{\mathbf{Q}}^T\breve{\mathbf{Q}} = \breve{\mathbf{Q}}\breve{\mathbf{Q}}^T = \mathbf{I}_{2MN}, \tag{19}$$

follows from the unitary property of the code rotation matrix $\mathbf{Q}$. Using (10) together with (16) and (18), we have

$$\mathbf{\Phi}_k\breve{\mathbf{Q}} = \begin{bmatrix} \mathrm{Re}\{(\mathbf{I}_M \otimes \mathbf{C}_k\mathbf{Q})\} & -\mathrm{Im}\{(\mathbf{I}_M \otimes \mathbf{C}_k\mathbf{Q})\} \\ \mathrm{Im}\{(\mathbf{I}_M \otimes \mathbf{C}_k\mathbf{Q})\} & \mathrm{Re}\{(\mathbf{I}_M \otimes \mathbf{C}_k\mathbf{Q})\} \end{bmatrix} = d_k\mathbf{\Phi}_{n_k}, \tag{20}$$

where $d_k \in \{\pm 1\}$, and $n_k \in \{1, \ldots, 2K\}$ is an index with $n_k \neq k$ and $n_k \neq n_l$ for $k \neq l$. The properties (11) and (20) imply that if code rotation matrix $\mathbf{Q}$ exists, then

$$\breve{\mathbf{Q}} = d_k\mathbf{\Phi}_k^T\mathbf{\Phi}_{n_k} \quad k = 1, \ldots, 2K,\ n_k \neq k,\ n_k \neq n_l \text{ for } k \neq l. \tag{21}$$

Hence, we conclude that $\breve{\mathbf{Q}}$ belongs to each of the following sets, i.e.,

$$\begin{aligned}\breve{\mathbf{Q}} &\in \mathcal{C}_1 \triangleq \{\pm\mathbf{\Phi}_1^T\mathbf{\Phi}_2, \pm\mathbf{\Phi}_1^T\mathbf{\Phi}_3, \ldots, \pm\mathbf{\Phi}_1^T\mathbf{\Phi}_{2K}\} \\ \breve{\mathbf{Q}} &\in \mathcal{C}_2 \triangleq \{\pm\mathbf{\Phi}_2^T\mathbf{\Phi}_1, \pm\mathbf{\Phi}_2^T\mathbf{\Phi}_3, \ldots, \pm\mathbf{\Phi}_2^T\mathbf{\Phi}_{2K}\} \\ &\vdots \\ \breve{\mathbf{Q}} &\in \mathcal{C}_{2K} \triangleq \{\pm\mathbf{\Phi}_{2K}^T\mathbf{\Phi}_1, \pm\mathbf{\Phi}_{2K}^T\mathbf{\Phi}_2, \ldots, \pm\mathbf{\Phi}_{2K}^T\mathbf{\Phi}_{2K-1}\}.\end{aligned} \tag{22}$$

From (15) and (22), we further conclude that the vector $\breve{\mathbf{h}} \triangleq \breve{\mathbf{Q}}^T\mathbf{h}$ which is orthogonal to $\mathbf{h}$, is a column of each $\tilde{\mathbf{A}}_k(\mathbf{h})$ for $k = 1,\ldots,2K$. In particular, we have

$$\breve{\mathbf{Q}} = \pm\mathbf{\Phi}_1^T\mathbf{\Phi}_{n_1} = \pm\mathbf{\Phi}_2^T\mathbf{\Phi}_{n_2} = \ldots = \pm\mathbf{\Phi}_{2K}^T\mathbf{\Phi}_{n_{2K}}, \quad (23)$$

with $n_1 \neq n_2 \neq \ldots \neq n_{2K}$. In other words, in each of the virtual signal matrices in (15) the vector $\breve{\mathbf{h}}$ appears at a different column position. Further, using (11) and (21), we obtain that $\breve{\mathbf{Q}}^T = -\breve{\mathbf{Q}}$ and, therefore, $\breve{\mathbf{h}}^T\mathbf{\Phi}_k^T\mathbf{\Phi}_l\mathbf{h} = \mathbf{h}^T\breve{\mathbf{Q}}\mathbf{\Phi}_k^T\mathbf{\Phi}_l\mathbf{h} = -\mathbf{h}^T\mathbf{\Phi}_{n_k}^T\mathbf{\Phi}_l\mathbf{h} = 0$ for $n_k \neq l$. Hence, we conclude that all remaining columns in $\tilde{\mathbf{A}}_k(\mathbf{h})$ are orthogonal to $\breve{\mathbf{h}}$. The following lemma summarizes the properties above:

Lemma 1: For rotatable OSTBCs, the vector $\breve{\mathbf{h}} = \breve{\mathbf{Q}}^T\mathbf{h}$ is identical to the n_k-th column (up to sign) of the k-th virtual signal matrices $\tilde{\mathbf{A}}_k(\mathbf{h})$, $k = 1,\ldots,2K$, in (15) for some $n_k \neq k$ with $n_k \neq n_l$ for $k \neq l$. Also, the remaining columns in $\tilde{\mathbf{A}}_k(\mathbf{h})$ are orthogonal to it.

Next, we propose to weight the different virtual snapshots in (14) by different scalars γ_k in order to enhance the desired signal component $\mathbf{h}$ in the virtual signal matrices of (15). Towards this aim, we define the weighted covariance matrix obtained from the $2K$ virtual snapshots in (14) as [11]

$$\begin{aligned}\mathcal{X}(\boldsymbol{\gamma}) &\triangleq \mathrm{E}\,\{\sum_{k=1}^{2K}\gamma_k\;\tilde{\mathbf{y}}(k,p)\tilde{\mathbf{y}}(k,p)^T\} \\ &= \sum_{k=1}^{2K}\gamma_k\left(\tilde{\mathbf{A}}_k(\mathbf{h})\mathrm{E}\{\underline{\mathbf{s}}\,\underline{\mathbf{s}}^T\}\tilde{\mathbf{A}}_k(\mathbf{h})^T\right) + \sum_{k=1}^{2K}\frac{\gamma_k\sigma^2}{2}\,\mathbf{I}_{2MN}, \quad (24)\end{aligned}$$

where $\mathrm{E}\{\cdot\}$ stands for the statistical expectation and the positive real weight coefficients are contained in the vector $\boldsymbol{\gamma} \triangleq [\gamma_1,\ldots,\gamma_{2K}]^T$. Let us assume that the symbol streams are mutually independent and independent of the sensor noise. Inserting (15) in (24), we have

$$\begin{aligned}\mathcal{X}(\boldsymbol{\gamma}) \;=\; &\underbrace{\sum_{k=1}^{2K}\gamma_k\,\mathrm{E}\{|s_k|^2\}\mathbf{h}\mathbf{h}^T}_{\text{desired component}} + \underbrace{\sum_{k=1}^{2K}\frac{\gamma_k\sigma^2}{2}\,\mathbf{I}_{2MN}}_{\text{noise contribution}} \\ &+ \underbrace{\sum_{k=1}^{2K}\sum_{l=1;l\neq k}^{2K}\gamma_k\,\mathrm{E}\{|s_l|^2\}\mathbf{\Phi}_k^T\mathbf{\Phi}_l\mathbf{h}\mathbf{h}^T\mathbf{\Phi}_l^T\mathbf{\Phi}_k}_{\text{orthogonal to the desired component}}. \quad (25)\end{aligned}$$

It can be readily verified from the skew-symmetry property in (11) and the covariance model in (25) that $\mathbf{h}$ is the eigenvector of $\mathcal{X}(\boldsymbol{\gamma})$ with the corresponding eigenvalue

$$\lambda_{\mathbf{h}} \triangleq \sum_{k=1}^{2K}\gamma_k\,\mathrm{E}\{|s_k|^2\}\|\mathbf{h}\|^2 + \sum_{k=1}^{2K}\gamma_k\frac{\sigma^2}{2}. \quad (26)$$

It has been proved in [11] that for uniform weighting in (25), i.e., for $\boldsymbol{\gamma} = \boldsymbol{\gamma}_{\mathrm{u}} \triangleq \gamma\,[1,\ldots,1]^T$ with arbitrary real $\gamma > 0$, the true channel vector $\mathbf{h}$ is the principal eigenvector of $\mathcal{X}(\boldsymbol{\gamma}_{\mathrm{u}})$ with the associated eigenvalue given as

$$\lambda_{\mathbf{h}} = \gamma\,\|\mathbf{h}\|^2\sum_{k=1}^{2K}\mathrm{E}\{|s_k|^2\} + \gamma\sum_{k=1}^{2K}\frac{\sigma^2}{2}. \quad (27)$$

If the principal eigenvalue is unique, then the channel vector can be estimated up to some real scalar ambiguity. Otherwise, there exist a set of linearly independent eigenvectors corresponding to the principal eigenvalue. This is, for example, the case in rotatable OSTBCs as stated by the following lemma:

Lemma 2: For rotatable OSTBCs and uniform weighting in (25), i.e. $\boldsymbol{\gamma} = \boldsymbol{\gamma}_{\mathrm{u}}$, the principal eigenvalue of $\mathcal{X}(\boldsymbol{\gamma})$ is not unique.

Using (20), it can be readily verified that multiplying $\mathcal{X}(\boldsymbol{\gamma}_{\mathrm{u}})$ in (24) from the left and the right by $\breve{\mathbf{Q}}^T$ and $\breve{\mathbf{Q}}$, respectively, only changes the sequence in which the summations are performed in (24) but does not change the individual components or the result of the summation. Hence, the following property holds

$$\breve{\mathbf{Q}}^T\mathcal{X}(\boldsymbol{\gamma}_{\mathrm{u}})\breve{\mathbf{Q}} = \mathcal{X}(\boldsymbol{\gamma}_{\mathrm{u}}). \quad (28)$$

Using (28), we have

$$\begin{aligned}\mathcal{X}(\boldsymbol{\gamma}_{\mathrm{u}})\;\mathbf{h} &= \lambda_{\max}\left(\mathcal{X}(\boldsymbol{\gamma}_{\mathrm{u}})\right)\mathbf{h} \\ &\Rightarrow \mathcal{X}(\boldsymbol{\gamma}_{\mathrm{u}})\breve{\mathbf{Q}}\breve{\mathbf{Q}}^T\mathbf{h} = \lambda_{\max}\left(\mathcal{X}(\boldsymbol{\gamma}_{\mathrm{u}})\right)\mathbf{h} \\ &\Rightarrow \breve{\mathbf{Q}}^T\mathcal{X}(\boldsymbol{\gamma}_{\mathrm{u}})\breve{\mathbf{Q}}\breve{\mathbf{Q}}^T\mathbf{h} = \lambda_{\max}\left(\mathcal{X}(\boldsymbol{\gamma}_{\mathrm{u}})\right)\breve{\mathbf{Q}}^T\mathbf{h} \\ &\Rightarrow \quad \mathcal{X}(\boldsymbol{\gamma}_{\mathrm{u}})\breve{\mathbf{h}} = \lambda_{\max}\left(\mathcal{X}(\boldsymbol{\gamma}_{\mathrm{u}})\right)\breve{\mathbf{h}}. \quad (29)\end{aligned}$$

As a result, we obtain that for the uniform weighting $\boldsymbol{\gamma} = \boldsymbol{\gamma}_{\mathrm{u}}$, the true channel vector $\mathbf{h}$ and its transformed version $\breve{\mathbf{h}}$ are both principal eigenvectors of $\mathcal{X}(\boldsymbol{\gamma})$, even in the case when different information symbols in $\underline{\mathbf{s}}(p)$ exhibit different powers.

The main idea in this paper is to exploit aforementioned properties of the rotatable OSTBCs and the degrees of freedom provided by the weighting coefficients, $\{\gamma_k\}_{k=1}^{2K}$ in (25) to resolve such a multiplicity shown in (29). The devised strategy is based on the assumption that the transmitted symbols have nonuniform powers such that $\mathrm{E}\{|s_k|^2\} \neq \mathrm{E}\{|s_{n_k}|^2\}$, at least for one $k \in \{1,\ldots,2K\}$ in the mapping (17). It is noteworthy to emphasize that the latter assumption includes two important examples as its special case. First, when all transmitted symbol powers are distinct and second, when there exists one symbol with unique transmitted power. Without loss of generality, we assume that the symbol powers are arranged in non-increasing order, i.e.,

$$\mathrm{E}\{|s_1|^2\} \geq \mathrm{E}\{|s_2|^2\} \geq \cdots \geq \mathrm{E}\{|s_{2K}|^2\};\;\; \mathrm{E}\{|s_k|^2\} \neq \mathrm{E}\{|s_{n_k}|^2\}, \quad (30)$$

for some $k \in \{1,\ldots,2K\}$. Further, consider a nonuniform weight vector $\boldsymbol{\gamma} = [\gamma_1,\ldots,\gamma_{2K}]^T$ chosen in correspondence to the symbol powers in (30), such that

$$\gamma_1 \geq \gamma_2 \geq \cdots \geq \gamma_{2K};\;\; \gamma_k \neq \gamma_{n_k}, \quad (31)$$

for the $k \in \{1,\ldots,2K\}$ according to (30) such that $\mathrm{E}\{|s_k|^2\} \neq \mathrm{E}\{|s_{n_k}|^2\}$. The following lemma summarizes the proposed weighting strategy:

Lemma 3: If the signal powers $\{\mathrm{E}\{|s_k|^2\}\}_{k=1}^{2K}$ and weighting coefficients $\{\gamma_k\}_{k=1}^{2K}$ in (25) satisfy (30) and (31), respectively, then the principal eigenvalue multiplicity of matrix $\mathcal{X}(\boldsymbol{\gamma})$ coming from implementation of rotatable OSTBC as shown in (29), will be resolved.

The proof of Lemma 3 directly follows from Lemma 1 and the particular relations between (30) and (31). Recall from Lemma 1, that the true channel vector $\mathbf{h}$ and the transformed channel vector $\breve{\mathbf{h}}$ are orthogonal to each other and orthogonal to the remaining columns in $\tilde{\mathbf{A}}_k(\mathbf{h})$, for $k = 1,\ldots,2K$. Consequently, both $\mathbf{h}$ and $\breve{\mathbf{h}}$ are eigenvectors of $\mathcal{X}(\boldsymbol{\gamma})$ in (25) even in the case of nonuniform weighting and the corresponding eigenvalues are given by (26) and

$$\lambda_{\breve{\mathbf{h}}} \triangleq \|\breve{\mathbf{h}}\|^2\sum_{k=1}^{2K}\gamma_k\,\mathrm{E}\{|s_{n_k}|^2\} + \sum_{k=1}^{2K}\gamma_k\frac{\sigma^2}{2}, \quad (32)$$

respectively. From (19), we have that $\|\breve{\mathbf{h}}\| = \|\mathbf{h}\|$. Therefore, taking into account (26), the weighting coefficients corresponding to the symbol powers in (31) and (32), we can show that

$\sum_{k=1}^{2K}\gamma_k\,\mathrm{E}\{|s_k|^2\} > \sum_{k=1}^{2K}\gamma_k\,\mathrm{E}\{|s_{n_k}|^2\}$. As a result, $\lambda_{\mathbf{h}} > \lambda_{\tilde{\mathbf{h}}}$ and we conclude that the principal eigenvalue multiplicity of $\mathcal{X}(\gamma)$ presented in (29) has been resolved.

In the following, we show that the covariance matching approaches of [9] can be viewed as special cases of the weighted covariance approach proposed in this paper for specific choices of the weight coefficients that satisfy the conditions of (31).

3.1. Euclidean Covariance Matching Criterion

Taking into account that the symbol streams are mutually independent and independent of the sensor noise along with (6), we obtain the following covariance matrix [8]

$$\mathbf{R} \triangleq \mathrm{E}\{\mathbf{y}\mathbf{y}^T\} = \mathbf{A}(\mathbf{h})\boldsymbol{\Lambda}_{\underline{\mathbf{s}}}\mathbf{A}^T(\mathbf{h}) + \frac{\sigma^2}{2}\mathbf{I}_{2MT}, \tag{33}$$

where $\boldsymbol{\Lambda}_{\underline{\mathbf{s}}} \triangleq \mathrm{E}\{\underline{\mathbf{s}}\ \underline{\mathbf{s}}^T\}$. Multiplying (33) from the right by $\mathbf{A}(\mathbf{h})/\|\mathbf{h}\|$ and using (8), we have [8]

$$\mathbf{R}\,\frac{\mathbf{A}(\mathbf{h})}{\|\mathbf{h}\|} = \frac{\mathbf{A}(\mathbf{h})}{\|\mathbf{h}\|}\,\left(\boldsymbol{\Lambda}_{\underline{\mathbf{s}}}\|\mathbf{h}\|^2 + \frac{\sigma^2}{2}\mathbf{I}_{2K}\right). \tag{34}$$

Since $\mathbf{A}(\mathbf{h})/\|\mathbf{h}\|$ has orthonormal columns and $\boldsymbol{\Lambda}_{\underline{\mathbf{s}}}$ is diagonal in the case of mutually uncorrelated transmitted symbols, (34) can be viewed as the characteristic equation for $\mathbf{R}$. Hence, the signal subspace eigenvalues of $\mathbf{R}$ depend only on the norm of the channel vector $\mathbf{h}$ and not on its direction. In practice, $\mathbf{R}$ can be estimated as

$$\hat{\mathbf{R}} = \frac{1}{P}\sum_{p=1}^{P}\mathbf{y}(p)\mathbf{y}(p)^T, \tag{35}$$

where P is the number of data blocks that are used to estimate $\mathbf{R}$.

The key idea of the Euclidean covariance matching (ECM) approach is to estimate the channel vector $\mathbf{h}$, by minimizing the norm of difference between the true and sample covariance matrices as [9]

$$\begin{aligned}\hat{\mathbf{h}}_{\mathrm{ECM}} &= \arg\min_{\tilde{\mathbf{h}}}\ \|\hat{\mathbf{R}} - \mathbf{R}(\tilde{\mathbf{h}})\|^2 \\ &= \arg\max_{\tilde{\mathbf{h}}}\left\{2\,\mathrm{tr}(\hat{\mathbf{R}}\mathbf{R}(\tilde{\mathbf{h}})) - \|\mathbf{R}(\tilde{\mathbf{h}})\|^2\right\},\end{aligned} \tag{36}$$

where $\mathrm{tr}(\cdot)$ stands for the trace of matrix. Using the orthogonality property (8) and equation (33), we can rewrite the both terms in the right-hand side of (36) as

$$\begin{aligned}\mathrm{tr}(\hat{\mathbf{R}}\mathbf{R}(\tilde{\mathbf{h}})) &= \mathrm{tr}(\mathbf{A}^T(\tilde{\mathbf{h}})\hat{\mathbf{R}}\mathbf{A}(\tilde{\mathbf{h}})\boldsymbol{\Lambda}_{\underline{\mathbf{s}}}) + \frac{\sigma^2}{2}\mathrm{tr}(\hat{\mathbf{R}}) \\ \|\mathbf{R}(\tilde{\mathbf{h}})\|^2 &= \|\tilde{\mathbf{h}}\|^4\|\boldsymbol{\Lambda}_{\underline{\mathbf{s}}}\|^2 + \sigma^2\|\tilde{\mathbf{h}}\|^2\mathrm{tr}(\boldsymbol{\Lambda}_{\underline{\mathbf{s}}}) + \frac{MT\sigma^4}{2}.\end{aligned}$$

Using the latter two equations and dropping the terms which do not depend on $\tilde{\mathbf{h}}$, (36) can be expressed as

$$\begin{aligned}\hat{\mathbf{h}}_{\mathrm{ECM}} = \arg\max_{\tilde{\mathbf{h}}}&\left\{2\,\mathrm{tr}(\mathbf{A}^T(\tilde{\mathbf{h}})\hat{\mathbf{R}}\mathbf{A}(\tilde{\mathbf{h}})\boldsymbol{\Lambda}_{\underline{\mathbf{s}}})\right. \\ &\left.-\|\tilde{\mathbf{h}}\|^4\|\boldsymbol{\Lambda}_{\underline{\mathbf{s}}}\|^2 - \sigma^2\|\tilde{\mathbf{h}}\|^2\mathrm{tr}(\boldsymbol{\Lambda}_{\underline{\mathbf{s}}})\right\}.\end{aligned} \tag{37}$$

It is noteworthy to stress that the main issue in the blind channel estimation algorithm is the estimation of the channel vector direction while estimation of the channel norm boils down to finding a proper scaling factor and can be done, i.e., as in [8]. Hence, by assuming the norm constraint on the optimization variable in (37) as $\|\tilde{\mathbf{h}}\| = \|\mathbf{h}\|$, the terms $(\|\tilde{\mathbf{h}}\|^4\|\boldsymbol{\Lambda}_{\underline{\mathbf{s}}}\|^2)$ and $(\sigma^2\|\tilde{\mathbf{h}}\|^2\mathrm{tr}(\boldsymbol{\Lambda}_{\underline{\mathbf{s}}}))$ become constants and, therefore, they can be dropped. It can be shown [8] that using (12), we have that (37) is equivalent to

$$\hat{\mathbf{h}}_{\mathrm{ECM}} = \arg\max_{\tilde{\mathbf{h}}}\tilde{\mathbf{h}}^T\mathcal{X}(\gamma_{\mathrm{ECM}})\,\tilde{\mathbf{h}}, \tag{38}$$

where

$$\begin{aligned}\mathcal{X}(\gamma_{\mathrm{ECM}}) &\triangleq \boldsymbol{\Phi}^T(\boldsymbol{\Lambda}_{\underline{\mathbf{s}}}\otimes\hat{\mathbf{R}})\boldsymbol{\Phi} \\ \gamma_{\mathrm{ECM}} &\triangleq [\mathrm{E}\{|s_1|^2\}, \mathrm{E}\{|s_2|^2\}, \ldots, \mathrm{E}\{|s_{2K}|^2\}]^T,\end{aligned} \tag{39}$$

together with $\|\tilde{\mathbf{h}}\| = \|\mathbf{h}\|$. Obviously, comparing (39) and (31) considering the condition in (30) reveals that the ECM approach is the special case of the proposed weighting strategy.

3.2. Kullback Covariance Matching Criterion

The main idea of Kullback covariance matching (KCM) is to minimize the divergence between the true and sample covariance matrices of the received data based on Kullback-Leibler divergence. Using this measure in the case of Gaussian observations results in the following optimization problem to estimate the channel vector as [9]

$$\begin{aligned}\hat{\mathbf{h}}_{\mathrm{KCM}} &= \arg\min_{\tilde{\mathbf{h}}}\left\{\mathrm{tr}\left(\mathbf{R}^{-1}(\tilde{\mathbf{h}})\,\hat{\mathbf{R}} - \mathbf{I}_{2MT}\right)\right. \\ &\qquad\left.-\log\det\left(\mathbf{R}^{-1}(\tilde{\mathbf{h}})\,\hat{\mathbf{R}}\right)\right\} \\ &= \arg\min_{\tilde{\mathbf{h}}}\left\{\mathrm{tr}(\mathbf{R}^{-1}(\tilde{\mathbf{h}})\,\hat{\mathbf{R}}) + \log\det\mathbf{R}(\tilde{\mathbf{h}})\right\},\end{aligned} \tag{40}$$

where $\log(\cdot)$ and $\det(\cdot)$ stand for the logarithm and the matrix determinant, respectively. To simplify the first term in (40), let us apply the Woodbury identity to the true covariance matrix in (33) to obtain

$$\begin{aligned}\mathbf{R}^{-1}(\tilde{\mathbf{h}}) &= \frac{1}{\sigma^2/2}\mathbf{I}_{2MT} - \frac{1}{(\sigma^2/2)^2}\mathbf{A}(\tilde{\mathbf{h}})\boldsymbol{\Lambda}_{\underline{\mathbf{s}}} \\ &\quad\cdot\left(\mathbf{I}_{2K} + \frac{\|\tilde{\mathbf{h}}\|^2}{(\sigma^2/2)}\boldsymbol{\Lambda}_{\underline{\mathbf{s}}}\right)^{-1}\mathbf{A}^T(\tilde{\mathbf{h}}).\end{aligned} \tag{41}$$

Using (41) and dropping the term which does not depend on $\tilde{\mathbf{h}}$, we obtain

$$\begin{aligned}\hat{\mathbf{h}}_{\mathrm{KCM}} = \arg\max_{\tilde{\mathbf{h}}}&\left\{\mathrm{tr}\left(\mathbf{A}(\tilde{\mathbf{h}})\boldsymbol{\Lambda}_{\underline{\mathbf{s}}}\left(\mathbf{I}_{2K} + \frac{\|\tilde{\mathbf{h}}\|^2}{(\sigma^2/2)}\boldsymbol{\Lambda}_{\underline{\mathbf{s}}}\right)^{-1}\right.\right. \\ &\left.\left.\cdot\ \mathbf{A}^T(\tilde{\mathbf{h}})\hat{\mathbf{R}}\right) - \log\det\left(\mathbf{R}(\tilde{\mathbf{h}})\right)\right\}.\end{aligned} \tag{42}$$

The $\log\det\left(\mathbf{R}(\tilde{\mathbf{h}})\right)$ term in (42) depends on the product of the eigenvalues of the true covariance matrix which in turn depends on the norm of the channel vector and not its direction. Hence, this term become constant and can be dropped if we again consider the norm constraint as in (38). Hence, using (12), we have that (42) is equivalent to

$$\hat{\mathbf{h}}_{\mathrm{KCM}} = \arg\max_{\tilde{\mathbf{h}}}\ \tilde{\mathbf{h}}^T\mathcal{X}(\gamma_{\mathrm{KCM}})\,\tilde{\mathbf{h}}, \tag{43}$$

where

$$\begin{aligned}\mathcal{X}(\gamma_{\mathrm{KCM}}) &\triangleq \boldsymbol{\Phi}^T\left(\boldsymbol{\Lambda}_{\underline{\mathbf{s}}}\left(\mathbf{I}_{2K} + \frac{\|\tilde{\mathbf{h}}\|^2}{(\sigma^2/2)}\boldsymbol{\Lambda}_{\underline{\mathbf{s}}}\right)^{-1}\otimes\hat{\mathbf{R}}\right)\boldsymbol{\Phi} \\ [\gamma_{\mathrm{KCM}}]_k &\triangleq \frac{\mathrm{E}\{|s_k|^2\}}{1 + \mathrm{E}\{|s_k|^2\}\frac{\|\tilde{\mathbf{h}}\|^2}{(\sigma^2/2)}},\end{aligned} \tag{44}$$

with $\|\tilde{\mathbf{h}}\| = \|\mathbf{h}\|$. Again, comparison between (44) and (31) and taking into account (30) approves that the KCM approach is the special case of our proposed weighting strategy.

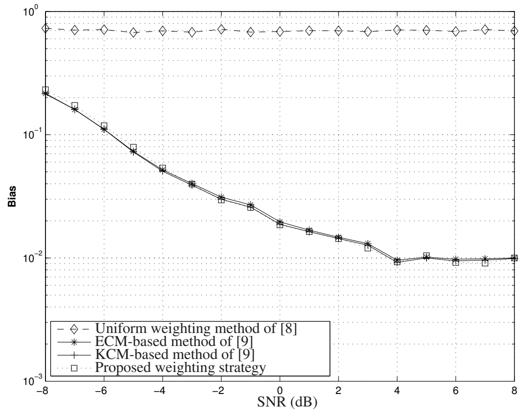

Fig. 1. Bias vs SNR, rotatable OSTBC.

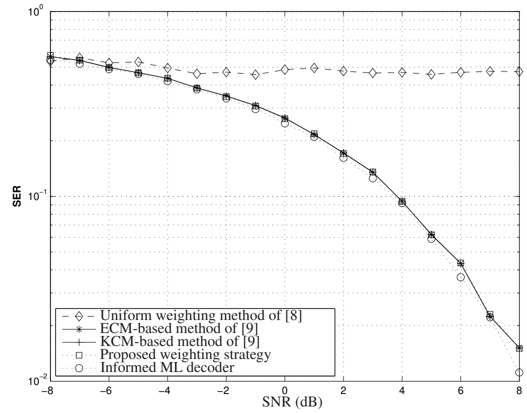

Fig. 2. SER vs SNR, rotatable OSTBC.

4. SIMULATIONS

In the simulations, the entries of $\mathbf{G}_l$ are kept fixed in each run and are independently drawn from a Gaussian distribution with zero mean and variance σ_g^2. The rotatable full rate Alamouti OSTBC [10] with $N = M = K = T = 2$ and QPSK symbols are used for encoding. Also, a frequency flat block fading channel is considered in the simulation setup in which $P = 100$ data blocks have been used to estimate covariance matrix according to (35) and the results are averaged over 200 Monte Carlo realizations. The proposed estimator is compared to the ECM and KCM methods of [9] and the relaxed maximum likelihood (RML) technique of [8] with uniform weighting. Fig. 1 displays the bias of the estimates, computed for a fixed channel vector $\mathbf{h}$ as the norm of the averaged channel estimation errors

$$\text{Bias} = \left\| \frac{1}{N_{\text{runs}}} \sum_{m=1}^{N_{\text{runs}}} \frac{\hat{\mathbf{h}}^{(m)}}{\|\hat{\mathbf{h}}^{(m)}\|} - \frac{\mathbf{h}}{\|\mathbf{h}\|} \right\|,$$

for the methods tested versus the signal-to-noise ratios (SNRs) where N_{runs} is the number of Monte-Carlo runs, and $\hat{\mathbf{h}}^{(m)}$ is the estimate of the $\mathbf{h}$ in the m-th run. Fig. 2 shows the symbol error rates (SERs) versus the SNR for the methods tested combined with the maximum likelihood (ML) decoder. Additionally, the results for the informed ML decoder are shown in this figure. The latter decoder is assumed to know the channel exactly. In the presented example, we have assumed $\mathbf{\Lambda}_{\underline{\mathbf{s}}} = \frac{2K}{8}\text{diag}([5, 1, 1, 1])$ which guarantees $\text{tr}(\mathbf{\Lambda}_{\underline{\mathbf{s}}}) = 2K$, i.e., the average transmit power per symbol is equal to that with equipower source. The weight vector $\boldsymbol{\gamma}$ corresponding to the proposed method is generated randomly and sorted according to (30) and (31). Also, $\boldsymbol{\gamma}_{\text{ECM}}$ and $\boldsymbol{\gamma}_{\text{KCM}}$ are selected based on (39) and (44), respectively. Aforementioned choices of weight vectors should be able to resolve the non-scalar ambiguities associated to the use of Alamouti as all satisfy the uniqueness condition presented in Lemma 3.

It can be observed from Fig. 1 that the RML approach of [8] with uniform weighting strategy is effectively not able to resolve the ambiguity corresponding to the use of rotatable OSTBC while the other methods resolve the ambiguities. Also, all other methods have quite the same performance. From Fig. 2, it follows that the same relationship between the performances can be observed in terms of SERs of the ML decoder. Also, it follows that the SER performance of all methods that satisfy conditions in Lemma 3 combined with the ML decoder closely achieves that of the informed ML detector.

5. 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, pp. 744-765, Mar. 1998.

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

[3] H. Jafarkhani and V. Tarokh, "Multiple transmit antenna differential detection from generalized orthogonal design," *IEEE Trans. Inf. Theory*, vol. 47, pp. 2626-2631, Sep. 2001.

[4] W.-K. Ma, B.-N. Vo, T. N. Davidson, and P.-C. Ching, "Blind ML detection of orthogonal space-time block codes: Efficient high-performance implementations," *IEEE Trans. Signal Process.*, vol. 54, pp. 738-751, Feb. 2006.

[5] D. S. Papailiopoulos and G. N. Karystinos, "Maximum-likelihood noncoherent OSTBC detection with polynomial complexity," *IEEE Trans. Wireless Commun.*, vol. 9, pp. 1935-1945, June 2010.

[6] M. Zia and Z. Ding, "A linear programming receiver for blind detection of full rate space-time block codes," *IEEE Trans. Signal Process.*, vol. 58, pp. 5819-5834, Nov. 2010.

[7] P. Stoica and G. Ganesan, "Space-time block codes: Trained, blind, and semi-blind detection," *Digital Signal Process.*, vol. 13, pp. 93-105, Jan. 2003.

[8] 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, pp. 4506-4517, Dec. 2005.

[9] J. Vía and I. Santamaría, "Correlation matching approaches for blind OSTBC channel estimation," *IEEE Trans. Signal Process.*, vol. 56, pp. 5950-5961, Dec. 2008.

[10] W.-K. Ma, "Blind ML detection of orthogonal space-time block codes: Identifiability and code construction," *IEEE Trans. Signal Process.*, vol. 55, pp. 3312-3324, July 2007.

[11] N. Sarmadi and M. Pesavento, "Closed-form blind channel estimation in orthogonally coded MIMO-OFDM systems: a simple strategy to resolve non-scalar ambiguities," in *Proc. SPAWC'11*, San francisco, CA, USA, June 2011, pp. 311-315.