Correlation Matching Approaches for Blind OSTBC Channel Estimation

Javier Vía, *Member, IEEE*, and Ignacio Santamaría, *Senior Member, IEEE*

***Abstract*—In this paper, the problem of blind channel estimation under orthogonal space–time block coded (OSTBC) transmissions is solved by minimizing some distance measure between the theoretical and estimated correlation matrices of the observations. Specifically, the minimization of the Euclidean distance and the Kullback–Leibler divergence leads, respectively, to the Euclidean correlation matching (ECM) and Kullback correlation matching (KCM) criteria. The proposed techniques exploit the knowledge of the source correlation matrix to unambiguously recover the multiple-input multiple-output (MIMO) channel. Furthermore, due to the orthogonality properties of OSTBCs, both the ECM and KCM criteria result in closed form solutions. In particular, the channel estimate is given by the principal eigenvector of a matrix, which is obtained from the estimated correlation matrix of the observations modified by the code matrices and a set of weights. In the ECM case, the weights are fixed and equal to the eigenvalues of the source correlation matrix, whereas the KCM weights depend on both the signal-to-noise ratio (SNR) and the source eigenvalues. Additionally, we show that the proposed approaches are equivalent in the low SNR regime, whereas in the high SNR regime the KCM criterion is asymptotically equivalent to the relaxed blind maximum-likelihood (ML) decoder. Finally, the performance of the proposed criteria is illustrated by means of some numerical examples.**

***Index Terms*—Blind channel estimation, correlation matching, information geometry, Kullback–Leibler divergence, maximum-likelihood (ML), orthogonal space–time block coding (OSTBC).**

I. Introduction

In the last years, OSTBC [1], [2] has emerged as one of the most promising techniques to exploit spatial diversity in MIMO systems [3], [4]. Specifically, if the channel is perfectly known at the receiver, OSTBC provides full diversity and reduces the complexity of the optimal maximum-likelihood (ML) decoder to a simple linear receiver followed by a symbol by symbol detector. However, in a practical scenario, the channel state information (CSI) is usually obtained through the transmission of a pilot sequence [5], which incurs in a penalty in terms of bandwidth efficiency. In order to avoid this penalty, several blind channel estimation and decoding techniques have been recently proposed [6]–[13]. These techniques, which are able to avoid the 3-dB penalty associated to differential approaches [14]–[18], can be divided into two groups. On one hand, some approaches exploit the finite alphabet of the sources for obtaining exact or approximated solutions to the optimal blind ML decoder [6]–[8], which translates into algorithms dependent on the specific signal constellation and with a relatively high computational complexity. On the other hand, the criteria solely based on the second-order statistics (SOS) of the observations [9]–[13] are independent of the signal constellation and result in low-complexity algorithms. However, in some practical cases, including the Alamouti code [1] and most of the multiple-input single-output (MISO) cases, the channel can not be unambiguously identified [19]–[21] without assuming some additional property, such as the transmission of correlated or linearly precoded sources [10].

In this paper, we propose two new blind OSTBC channel estimation techniques which are able to exploit the previous knowledge of the source correlation matrix.[1] Both techniques are based on the general idea of correlation matching, which amounts to minimize the divergence between the theoretical and finite sample estimate of the correlation matrix associated to the received signals. In particular, we consider two different measures of divergence, namely, the Euclidean distance and the Kullback–Leibler divergence, which lead to the Euclidean correlation matching (ECM) and Kullback correlation matching (KCM) criteria, respectively. The ECM criterion has been previously applied to other blind channel estimation and equalization problems [22]–[24], and KCM, which can be considered as the basis of information geometry [25], is closely related to the ML estimation problem [26]–[29]. However, in general both criteria result in nonlinear optimization problems, which must be solved by means of numerical methods.

Unlike most of the blind channel estimation problems [22]–[24], we show that for blind OSTBC channel estimation both the ECM and KCM criteria yield closed-form solutions. Concretely, in both cases the channel estimate is given by the principal eigenvector of a matrix, which is obtained from the correlation matrix of the observations, the OSTBC code matrices, and a set of weights. The ECM weights are directly given by the eigenvalues of the source correlation matrix, whereas the KCM weights also depend on the signal-to-noise ratio (SNR). This permits a straightforward interpretation of the proposed techniques. On one hand, in the case of white sources, both criteria reduce to the relaxed blind ML decoder [7], [10]. On the other hand, regardless of the source correlation matrix, the KCM technique reduces to the ECM criterion in the low SNR

Manuscript received November 27, 2007; revised July 14, 2008. First published August 19, 2008; current version published November 19, 2008. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Kostas Berberidis. This work was supported by the Spanish Government under projects TEC2007-68020-C04-02/TCM and CONSOLIDER-INGENIO 2010 CSD2008-00010 (COMONSENS).

The authors are with the Communications Engineering Department (DICOM), University of Cantabria, Santander 39005, Spain (e-mail: jvia@gtas.dicom.unican.es; nacho@gtas.dicom.unican.es).

Digital Object Identifier 10.1109/TSP.2008.929661

[1] Although the proposed techniques are not strictly blind due to the assumption of a known source correlation matrix. We refer to them as blind methods to point out that they do not require the transmission of any pilot sequence.

regime, and to the relaxed blind ML decoder in the high SNR case. In other words, the KCM technique uses the most reliable information, i.e., in the low SNR regime the channel estimate is extracted from the previous knowledge of the source correlation matrix, whereas for high SNRs, the channel information is recovered, almost exclusively, from the congruence between the observations and the OSTBC data model.

The structure of the paper is as follows. In Section II, the OSTBC data model and its main properties are summarized. A brief review of previous approaches to blind OSTBC channel estimation is presented in Section III. The ECM and KCM techniques for the blind extraction of the channel, up to a real scalar, are presented in Section IV. Although this is the key point in the blind channel estimation process, in Section V, it is shown that the ECM and KCM criteria also provide estimates of the channel energy and noise variance. The main properties of the proposed criteria, and their relationship with other approaches are summarized in Section VI. Finally, the performance of the proposed techniques is illustrated in Section VII by means of some numerical examples, and the main conclusions of the paper are summarized in Section VIII.

II. Some Background on OSTBCs

Notation

1) Vectors/Matrices: Throughout this paper, we will use bold-faced upper case letters to denote matrices, e.g., $\mathbf{X}$, with elements $x_{i,j}$; bold-faced lower case letters for column vector, e.g., $\mathbf{x}$, and light-face lower case letters for scalar quantities. The superscripts $(\cdot)^T$ and $(\cdot)^H$ denote transpose and Hermitian, respectively. The real and imaginary parts will be denoted as $\Re(\cdot)$ and $\Im(\cdot)$, and superscript $\hat{(\cdot)}$ will denote estimated matrices, vectors or scalars. The trace, Frobenius norm and column-wise vectorized version of a matrix $\mathbf{A}$ will be denoted as $\mathrm{Tr}(\mathbf{A})$, $\|\mathbf{A}\|$ and $\mathrm{vec}(\mathbf{A})$, respectively. The diagonal matrix obtained from the elements of a vector $\mathbf{a}$ is denoted as $\mathrm{diag}(\mathbf{a})$. We use $\mathbf{A} \in \mathbb{C}^{M\times N}$ and $\mathbf{A} \in \mathbb{R}^{M\times N}$ to denote that $\mathbf{A}$ is a complex or real matrix of dimension $M \times N$. Finally, the identity matrix of dimensions $p \times p$ will be denoted as $\mathbf{I}_p$, $\mathbf{0}$ will denote the zero matrix of the required dimensions, and $\mathcal{E}[\cdot]$ will denote the expectation operator.

2) MIMO Parameters: In this paper, a flat-fading MIMO system with n_T transmit and n_R receive antennas is assumed. The $n_T \times n_R$ complex channel matrix will be written as

$$\sqrt{E}\mathbf{H} = \begin{bmatrix} h_{1,1} & \dots & h_{1,n_R} \\ \vdots & \ddots & \vdots \\ h_{n_T,1} & \dots & h_{n_T,n_R} \end{bmatrix}$$

where $E = \sum_{i,j} |h_{i,j}|^2$ is the channel energy, $\mathbf{H}$ is the normalized (to $\|\mathbf{H}\| = 1$) MIMO channel, and $h_{i,j}$ denotes the channel response between the ith transmit and the jth receive antennas. Additionally, the complex noise at the receive antennas is considered independent and identically distributed (i.i.d.) Gaussian with variance σ^2.

A. Data Model for STBCs

Let us consider a linear space–time block code (STBC) transmitting M symbols during L time slots with n_T antennas at the transmitter side. The transmission rate is defined as $R = M/L$, and the number of real symbols M' transmitted in each block is

$$M' = \begin{cases} M & \text{for real codes} \\ 2M & \text{for complex codes.} \end{cases}$$

For a STBC, the nth block of data can be expressed as

$$\mathbf{S}(\mathbf{s}[n]) = \sum_{k=1}^{M'} \mathbf{C}_k s_k[n]$$

where $\mathbf{s}[n] = [s_1[n], \dots, s_{M'}[n]]^T$ contains the M' real information symbols transmitted in the nth block, and $\mathbf{C}_k \in \mathbb{C}^{L\times n_T}$, $k = 1, \dots, M'$, are the code matrices. In the case of real STBCs, the transmitted matrix $\mathbf{S}(\mathbf{s}[n])$ and the code matrices $\mathbf{C}_k$ are real.

The complex signal at the n_R receive antennas can be written as

$$\begin{aligned} \mathbf{Y}[n] &= \sqrt{E}\mathbf{S}(\mathbf{s}[n])\mathbf{H} + \mathbf{N}[n] \\ &= \sqrt{E}\sum_{k=1}^{M'} \mathbf{W}_k(\mathbf{H}) s_k[n] + \mathbf{N}[n] \end{aligned} \quad (1)$$

where $\mathbf{N}[n] \in \mathbb{C}^{L\times n_R}$ is the white Gaussian complex noise with variance σ^2, and

$$\mathbf{W}_k(\mathbf{H}) = \mathbf{C}_k\mathbf{H}, \quad k = 1, \dots, M'$$

represent the composite effect of the MIMO channel and the STBC code. Here, we must note that the data model in (1) can be seen as a particular case of a complex system with a noncircular (improper) source [30]–[33]. Furthermore, in Section II-C it will be shown that we can assume, without loss of generality, a diagonal correlation matrix $\mathbf{R}_\mathbf{s} = \mathcal{E}[\mathbf{s}[n]\mathbf{s}^T[n]]$. Thus, in order to exploit the fact that a set of real information symbols are observed through a complex equivalent channel given by $\sqrt{E}\mathbf{W}_k(\mathbf{H})$, we define the real vector $\tilde{\mathbf{y}}[n] = \mathrm{vec}\left(\left[\Re(\mathbf{Y}^T[n]) \quad \Im(\mathbf{Y}^T[n])\right]^T\right)$, and rewrite (1) in terms of real variables as

$$\tilde{\mathbf{y}}[n] = \sqrt{E}\tilde{\mathbf{W}}(\mathbf{H})\mathbf{s}[n] + \tilde{\mathbf{n}}[n] \quad (2)$$

where $\tilde{\mathbf{n}}[n]$ is defined analogously to $\tilde{\mathbf{y}}[n]$ and $\tilde{\mathbf{W}}(\mathbf{H})$ can be seen as the equivalent channel, whose kth column is given by $\mathrm{vec}\left(\left[\Re(\mathbf{W}_k^T(\mathbf{H})) \quad \Im(\mathbf{W}_k^T(\mathbf{H}))\right]^T\right) = \tilde{\mathbf{D}}_k\tilde{\mathbf{h}}$, with $\tilde{\mathbf{h}} = \mathrm{vec}\left(\left[\Re(\mathbf{H}^T) \quad \Im(\mathbf{H}^T)\right]^T\right)$,

$$\tilde{\mathbf{C}}_k = \underbrace{\begin{bmatrix} \Re(\mathbf{C}_k) & -\Im(\mathbf{C}_k) \\ \Im(\mathbf{C}_k) & \Re(\mathbf{C}_k) \end{bmatrix}}_{2L\times 2n_T},$$

and

$$\tilde{\mathbf{D}}_k = \underbrace{\begin{bmatrix} \tilde{\mathbf{C}}_k & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \tilde{\mathbf{C}}_k & \ddots & \vdots \\ \vdots & \ddots & \ddots & \mathbf{0} \\ \mathbf{0} & \dots & \mathbf{0} & \tilde{\mathbf{C}}_k \end{bmatrix}}_{2Ln_R\times 2n_T n_R}, \quad k = 1,\dots,M'.$$

If the MIMO channel is known at the receiver, the coherent ML decoder minimizes the following cost function [34]

$$\hat{\mathbf{s}}[n] = \underset{\mathbf{s}[n]}{\mathrm{argmin}} \left\| \tilde{\mathbf{y}}[n] - \sqrt{E}\tilde{\mathbf{W}}(\mathbf{H})\mathbf{s}[n] \right\|^2$$

subject to the constraint that the elements of $\mathbf{s}[n]$ belong to a finite set $\mathcal{S}$. Unfortunately, in a general case this is a NP-hard problem and optimal algorithms to solve it, such as *sphere decoding*, can be computationally expensive [35]–[38].

B. OSTBC Properties

In the OSTBC case, the equivalent channel matrix $\tilde{\mathbf{W}}(\mathbf{H})$ satisfies

$$\tilde{\mathbf{W}}^T(\mathbf{H})\tilde{\mathbf{W}}(\mathbf{H}) = \mathbf{I}_{M'} \tag{3}$$

which allows us to reduce the complexity of the ML receiver to find the closest symbols to the estimated signal

$$\hat{\mathbf{s}}_{\mathrm{ML}}[n] = \frac{\tilde{\mathbf{W}}^T(\mathbf{H})\tilde{\mathbf{y}}[n]}{\sqrt{E}}$$

i.e., the optimal receiver reduces to a matched filter followed by a symbol by symbol detector.

The necessary and sufficient conditions to satisfy (3) can be found in [34]. Specifically, the code matrices must satisfy, for $k = 1,\dots,M'$,

$$\mathbf{C}_k^H\mathbf{C}_l = \begin{cases} \mathbf{I}_{n_T} & k = l, \\ -\mathbf{C}_l^H\mathbf{C}_k & k \neq l, \end{cases} \tag{4}$$

and it is straightforward to prove that the above condition also holds for the code matrices with real elements

$$\tilde{\mathbf{C}}_k^T\tilde{\mathbf{C}}_l = \begin{cases} \mathbf{I}_{2n_T} & k = l \\ -\tilde{\mathbf{C}}_l^T\tilde{\mathbf{C}}_k & k \neq l. \end{cases}$$

Finally, the following properties are direct consequences of (3) and (4).

Property 1: The transmitted signals using an OSTBC satisfy $\mathbf{S}^H(\mathbf{s}[n])\mathbf{S}(\mathbf{s}[n]) = \|\mathbf{s}[n]\|^2\mathbf{I}_{n_T}$.

Remark 1: Considering a source correlation matrix $\mathbf{R}_\mathbf{s} = \mathcal{E}[\mathbf{s}[n]\mathbf{s}^T[n]]$, Property 1 implies that the energy transmitted by each antenna over the L channel uses is $\mathrm{Tr}(\mathbf{R}_\mathbf{s})$, and the total transmitted power per channel use is

$$P_T = \frac{n_T}{L}\mathrm{Tr}(\mathbf{R}_\mathbf{s}).$$

Furthermore, the signal power at the receiver is $P_R = P_T E/n_T$, which yields a received SNR

$$\mathrm{SNR}_\mathbf{H} = \frac{P_R}{n_R\sigma^2} = P_T\frac{E}{n_T n_R\sigma^2} = \frac{E}{n_R\sigma^2}\frac{\mathrm{Tr}(\mathbf{R}_\mathbf{s})}{L}.$$

Property 2: Given the OSTBC code matrices $\mathbf{C}_k \in \mathbb{C}^{L\times n_T}$, $k = 1,\dots,M'$, and an orthogonal matrix $\mathbf{Q} \in \mathbb{R}^{M'\times M'}$ with elements $q_{k,l}$, the matrices

$$\mathbf{B}_k = \sum_{l=1}^{M'} q_{k,l}\mathbf{C}_l, \quad k = 1,\dots,M'$$

define a new OSTBC with the same parameters n_T, L and M'.

Remark 2: Considering the eigenvalue decomposition $\mathbf{R}_\mathbf{s} = \mathbf{Q}_\mathbf{s}\mathbf{\Lambda}_\mathbf{s}\mathbf{Q}_\mathbf{s}^T$ (with $\mathbf{Q}_\mathbf{s}$ an orthogonal matrix), Property 2 allows us to rewrite the data model in (2) as

$$\tilde{\mathbf{y}}[n] = \sqrt{E}\tilde{\mathbf{W}}'(\mathbf{H})\mathbf{s}'[n] + \tilde{\mathbf{n}}[n]$$

where $\tilde{\mathbf{W}}'(\mathbf{H}) = \tilde{\mathbf{W}}(\mathbf{H})\mathbf{Q}_\mathbf{s}$ is the equivalent channel of a modified OSTBC, and $\mathbf{s}'[n] = \mathbf{Q}_\mathbf{s}^T\mathbf{s}[n]$ is a rotated information vector with $\mathcal{E}\left[\mathbf{s}'[n]\mathbf{s}'^T[n]\right] = \mathbf{\Lambda}_\mathbf{s}$. Thus, assuming that the correlation matrix $\mathbf{R}_\mathbf{s}$ is *a priori* known, we can easily obtain the modified code matrices $\mathbf{C}_k$ (as well as their extended versions $\tilde{\mathbf{C}}_k$ and $\tilde{\mathbf{D}}_k$). Therefore, from now on we will assume, without loss of generality, known and diagonal correlation matrices $\mathbf{R}_\mathbf{s} = \mathbf{\Lambda}_\mathbf{s}$ with elements $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_{M'}$ in its diagonal.

Finally, it should be noted that $\mathbf{R}_\mathbf{s}$ can differ from a scaled version of the identity matrix due to three different reasons.

1) The information symbols can be correlated due to the specific type of modulation format (duobinary signaling techniques for instance), or due to the channel encoder [39].
2) The power imbalance among the different information symbols can be a consequence of the specific OSTBC design [40], [41].
3) As will be shown in Section III, the sources $\mathbf{s}[n]$ could have been linearly precoded [10] in order to avoid the indeterminacies associated to some OSTBCs, such as the Alamouti code [1].

III. Previous Works on Blind OSTBC Channel Estimation

In the last years, several blind and semiblind techniques for channel estimation under STBC transmissions have been proposed [6]–[13]. These techniques can be divided into two groups, depending on whether they try to approximate the optimal joint estimator of the channel and sources [6]–[8], or they are solely based on the second-order statistics (SOS) of the observations [9]–[13]. The main advantage of SOS-based techniques resides in their independence of the specific signal constellation and reduced computational complexity. In this section, we review the basic approaches to blind OSTBC channel estimation, pointing out their main properties and drawbacks.

A. Relaxed Blind ML Decoder (Subspace Method)

In order to simplify the blind channel estimation problem, previous works [7], [9]–[13] have suggested to relax the finite-alphabet constraint on the information symbols. Thus, assuming that the MIMO channel remains constant during the transmission of N OSTBC data blocks, the relaxed blind ML decoder [7], [10] amounts to minimize

$$\left\{\hat{\mathbf{H}}^{\mathrm{SS}}, \hat{\mathbf{s}}^{\mathrm{SS}}[n]\right\} = \underset{\mathbf{H},\mathbf{s}[n]}{\arg\min} \sum_{n=0}^{N-1} \left\|\tilde{\mathbf{y}}[n] - \sqrt{E}\tilde{\mathbf{W}}(\mathbf{H})\mathbf{s}[n]\right\|^2 \tag{5}$$

and its solution $\hat{\tilde{\mathbf{h}}}^{\mathrm{SS}}$ can be directly obtained as the principal eigenvector of the matrix

$$\mathbf{\Phi}^{\mathrm{SS}} = \sum_{k=1}^{M'} \tilde{\mathbf{D}}_k^T \hat{\mathbf{R}}_{\tilde{\mathbf{y}}} \tilde{\mathbf{D}}_k$$

where $\hat{\mathbf{R}}_{\tilde{\mathbf{y}}}$ is the finite sample estimate of the correlation matrix of the observations $\mathbf{R}_{\tilde{\mathbf{y}}} = \mathcal{E}[\tilde{\mathbf{y}}[n]\tilde{\mathbf{y}}^H[n]]$, i.e.,

$$\hat{\mathbf{R}}_{\tilde{\mathbf{y}}} = \frac{1}{N}\sum_{n=0}^{N-1} \tilde{\mathbf{y}}[n]\tilde{\mathbf{y}}^T[n].$$

Obviously, the above criterion is affected by a real scale factor in the channel and signal estimates, which can be considered as a minor problem easy to solve in a later step. However, in some practical scenarios, which include the well-known Alamouti code [1] and most of the MISO cases, the relaxed blind ML decoder is affected by additional indeterminacies, which preclude the unambiguous blind recovery of the MIMO channel [20]. In particular, the identifiability problems are due to the existence of a not null subspace of spurious channels $\hat{\mathbf{H}}^{\mathrm{SS}} \neq \pm\mathbf{H}$ and symbol sequences $\hat{\mathbf{s}}^{\mathrm{SS}}[n] \neq \pm\mathbf{s}[n]$ satisfying

$$\tilde{\mathbf{W}}(\hat{\mathbf{H}}^{\mathrm{SS}})\hat{\mathbf{s}}^{\mathrm{SS}}[n] = \tilde{\mathbf{W}}(\mathbf{H})\mathbf{s}[n] \quad n = 0,\ldots,N-1 \tag{6}$$

which, from a practical point of view, translates into a multiplicity $P > 1$ of the largest eigenvalue of $\mathbf{\Phi}^{\mathrm{SS}}$.

Finally, we must point out that (5) is solely based on the congruence between the estimates ($\hat{\mathbf{H}}^{\mathrm{SS}}$ and $\hat{\mathbf{s}}^{\mathrm{SS}}[n]$) and the observations $\tilde{\mathbf{y}}[n]$. Therefore, in the absence of noise, and assuming that (6) is only satisfied by the true channel and symbol sequence, the channel can be exactly recovered by means of the relaxed blind ML decoder with a finite sample size. In the context of blind identification this is known as the finite sample convergence property [42], and any method satisfying this property is referred to as deterministic.

B. Weighted Version of the Relaxed Blind ML Decoder

In order to remove the indeterminacies pointed out in the previous subsection, in [10] the authors propose a linear precoding approach, which introduces a correlation $\mathbf{\Lambda}_{\mathbf{s}}$ on the information symbols, and leads to a weighted version of the relaxed blind ML estimator. Specifically, the normalized channel estimate is obtained as the eigenvector associated to the largest eigenvalue of

$$\mathbf{\Phi} = \sum_{k=1}^{M'} \rho_k \tilde{\mathbf{D}}_k^T \hat{\mathbf{R}}_{\tilde{\mathbf{y}}} \tilde{\mathbf{D}}_k \tag{7}$$

where the weights $\rho_1 \geq \rho_2 \geq \cdots \geq \rho_{M'}$ are free parameters to be selected by the user. Furthermore, given a correlation matrix $\mathbf{\Lambda}_{\mathbf{s}}$ such that one of its eigenvalues λ_k has multiplicity one, we could identify the associated eigenvector $\mathbf{v} = \tilde{\mathbf{D}}_{\mathbf{k}}\tilde{\mathbf{h}}$ in the correlation matrix $\mathbf{R}_{\tilde{\mathbf{y}}}$, and unambiguously extract the channel as $\tilde{\mathbf{h}} = \tilde{\mathbf{D}}_k^T\mathbf{v}$.

Unlike the unweighted version, (7) is based on the source correlation matrix $\mathbf{R}_{\mathbf{s}} = \mathbf{\Lambda}_{\mathbf{s}}$. Therefore, although this technique is able to recover the channel regardless of which OSTBC is used, i.e., even when there exists spurious solutions of (6), it is affected by a noise floor due to the difference between the theoretical correlation matrix $\mathbf{\Lambda}_{\mathbf{s}}$, and its finite sample version.[2]

C. Channel Energy and Noise Variance Estimation

As previously pointed out, the above criteria only recover the channel up to a real scalar. However, the analysis of the eigenvalues $\mu_1 \geq \cdots \geq \mu_{2Ln_R}$ of $\hat{\mathbf{R}}_{\tilde{\mathbf{y}}}$ provides the following estimates of the channel energy[3] and noise variance

$$\hat{E} = \frac{\sum_{k=1}^{M'} \left(\frac{\mu_k - \hat{\sigma}^2}{2}\right)}{\mathrm{Tr}\left(\mathbf{\Lambda}_{\mathbf{s}}\right)},$$
$$\hat{\sigma}^2 = \frac{2}{2Ln_R - M'} \sum_{k=M'+1}^{2Ln_R} \mu_k \tag{8}$$

where we are assuming that the noise subspace is not null, i.e., $2Ln_R > M'$, which is satisfied for all the OSTBCs excluding the MISO $(n_R = 1)$ Alamouti case [1].

IV. Blind OSTBC Channel Estimation Through Correlation Matching

In this work, we propose two new SOS-based blind channel estimation techniques that exploit our knowledge about the

[2]Obviously, the noise floor decreases with the number of observations N.

[3]See [10] for an alternative estimator based on an estimate of the normalized channel.

source correlation matrix $\mathbf{R_s}$. The proposed correlation matching criteria aim at adjusting the theoretical correlation matrix of the observations

$$\begin{aligned}\mathbf{R}_{\tilde{\mathbf{y}}}(\mathbf{H},E,\sigma^2) &= \mathcal{E}\left[\tilde{\mathbf{y}}[n]\tilde{\mathbf{y}}^T[n]\right]\\ &= E\tilde{\mathbf{W}}(\mathbf{H})\mathbf{\Lambda_s}\tilde{\mathbf{W}}^T(\mathbf{H}) + \frac{\sigma^2}{2}\mathbf{I}_{2Ln_R} \quad (9)\end{aligned}$$

which depends on the parameters $\mathbf{H}$, E and σ^2; and its finite sample estimate $\hat{\mathbf{R}}_{\tilde{\mathbf{y}}}$.

Different measures of the divergence between the two correlation matrices yield different criteria. Specifically, the use of the Euclidean distance leads to the ECM criterion, whereas the use of the Kullback–Leibler divergence between Gaussian pdfs provides the KCM criterion. In other estimation scenarios, both criteria result in nonlinear optimization problems, which must be solved by means of numerical methods. However, due to the orthogonality property in (3), both correlation matching criteria lead to closed-form channel estimates, which are obtained by solving an eigenvalue (EV) problem.

The key point in the channel estimation process consists in the estimate of the normalized channel $\mathbf{H}$, whereas the estimate of the channel energy only translates into a scale factor in the recovered signals. Therefore, the estimation of E and σ^2 will be relegated to Section V, and here we will focus on the blind estimation of the normalized channel $\mathbf{H}$.

A. Euclidean Correlation Matching

The classical correlation matching criterion for parameter estimation is based on the minimization of the Euclidean distance between the theoretical and estimated SOS of the observations. Under mild assumptions [22], [23], [43], this criterion provides, asymptotically as $N \to \infty$, the unbiased estimator with minimum variance [22], [23], [43]. However, although ECM has been successfully applied to blind channel estimation and equalization problems [22]–[24], it usually results in nonlinear optimization problems, which have to be solved by means of numerical methods.

Omitting, for notational simplicity, the dependence w.r.t. E and σ^2, the ECM criterion amounts to solve the following optimization problem,

$$\hat{\mathbf{H}}^{\mathrm{ECM}} = \arg\min_{\mathbf{H}} \left\|\hat{\mathbf{R}}_{\tilde{\mathbf{y}}} - \mathbf{R}_{\tilde{\mathbf{y}}}(\mathbf{H})\right\|^2$$

which can be rewritten as

$$\hat{\mathbf{H}}^{\mathrm{ECM}} = \arg\max_{\mathbf{H}} \left[2\mathrm{Tr}\left(\mathbf{R}_{\tilde{\mathbf{y}}}(\mathbf{H})\hat{\mathbf{R}}_{\tilde{\mathbf{y}}}\right) - \|\mathbf{R}_{\tilde{\mathbf{y}}}(\mathbf{H})\|^2\right]. \quad (10)$$

Let us now analyze separately the terms into the brackets in (10). By using the property $\mathrm{Tr}(\mathbf{AB}) = \mathrm{Tr}(\mathbf{BA})$ we obtain

$$\begin{aligned}&\mathrm{Tr}\left(\mathbf{R}_{\tilde{\mathbf{y}}}(\mathbf{H})\hat{\mathbf{R}}_{\tilde{\mathbf{y}}}\right)\\ &\quad = E\mathrm{Tr}\left(\mathbf{\Lambda_s}\tilde{\mathbf{W}}^T(\mathbf{H})\hat{\mathbf{R}}_{\tilde{\mathbf{y}}}\tilde{\mathbf{W}}(\mathbf{H})\right) + \frac{\sigma^2}{2}\mathrm{Tr}\left(\hat{\mathbf{R}}_{\tilde{\mathbf{y}}}\right).\end{aligned}$$

Furthermore, taking into account the orthogonality of the equivalent channel matrix $\tilde{\mathbf{W}}(\mathbf{H})$ (see (3)), we can write

$$\|\mathbf{R}_{\tilde{\mathbf{y}}}(\mathbf{H})\|^2 = E^2\|\mathbf{\Lambda_s}\|^2 + \sigma^2 E\mathrm{Tr}(\mathbf{\Lambda_s}) + \frac{Ln_R\sigma^4}{2}.$$

Thus, the ECM criterion reduces to

$$\hat{\mathbf{H}}^{\mathrm{ECM}} = \arg\max_{\mathbf{H}} \Psi^{\mathrm{ECM}}$$

where

$$\begin{aligned}\Psi^{\mathrm{ECM}} &= E\mathrm{Tr}\left(\mathbf{\Gamma}^{\mathrm{ECM}}\tilde{\mathbf{W}}^T(\mathbf{H})\hat{\mathbf{R}}_{\tilde{\mathbf{y}}}\tilde{\mathbf{W}}(\mathbf{H})\right) - \frac{E^2\|\mathbf{\Gamma}^{\mathrm{ECM}}\|^2}{2}\\ &\quad -\frac{\sigma^2}{2}E\mathrm{Tr}\left(\mathbf{\Gamma}^{\mathrm{ECM}}\right) + \frac{\sigma^2}{2}\mathrm{Tr}\left(\hat{\mathbf{R}}_{\tilde{\mathbf{y}}}\right) - \frac{Ln_R\sigma^4}{4} \quad (11)\end{aligned}$$

and we have defined $\mathbf{\Gamma}^{\mathrm{ECM}} = \mathrm{diag}\left(\left[\rho_1^{\mathrm{ECM}},\ldots,\rho_{M'}^{\mathrm{ECM}}\right]\right) = \mathbf{\Lambda_s}$.

Here, we must note that in the above equation only the first term depends on $\mathbf{H}$, and therefore the ECM criterion amounts to maximize $\mathrm{Tr}\left(\mathbf{\Gamma}^{\mathrm{ECM}}\tilde{\mathbf{W}}^T(\mathbf{H})\hat{\mathbf{R}}_{\tilde{\mathbf{y}}}\tilde{\mathbf{W}}(\mathbf{H})\right)$. Now, taking into account that the kth column of $\tilde{\mathbf{W}}(\mathbf{H})$ is given by $\tilde{\mathbf{D}}_k\tilde{\mathbf{h}}$, we can write

$$\mathrm{Tr}\left(\mathbf{\Gamma}^{\mathrm{ECM}}\tilde{\mathbf{W}}^T(\mathbf{H})\hat{\mathbf{R}}_{\tilde{\mathbf{y}}}\tilde{\mathbf{W}}(\mathbf{H})\right) = \tilde{\mathbf{h}}^T\mathbf{\Phi}^{\mathrm{ECM}}\tilde{\mathbf{h}} = \beta^{\mathrm{ECM}}$$

where

$$\mathbf{\Phi}^{\mathrm{ECM}} = \sum_{k=1}^{M'}\rho_k^{\mathrm{ECM}}\tilde{\mathbf{D}}_k^T\hat{\mathbf{R}}_{\tilde{\mathbf{y}}}\tilde{\mathbf{D}}_k$$

can be seen as a modified correlation matrix [44], [45], and the weights ρ_k^{ECM} are directly given by the source eigenvalues, i.e., $\rho_k^{\mathrm{ECM}} = \lambda_k$. Thus, since $\|\tilde{\mathbf{h}}\| = 1$, the ECM estimate $\hat{\tilde{\mathbf{h}}}^{\mathrm{ECM}}$ is obtained by solving the EV problem

$$\mathbf{\Phi}^{\mathrm{ECM}}\hat{\tilde{\mathbf{h}}}^{\mathrm{ECM}} = \hat{\beta}^{\mathrm{ECM}}\hat{\tilde{\mathbf{h}}}^{\mathrm{ECM}} \quad (12)$$

where $\hat{\beta}^{\mathrm{ECM}}$ is the largest eigenvalue of $\mathbf{\Phi}^{\mathrm{ECM}}$.

B. Kullback Correlation Matching

In this subsection, we consider a different measure of the divergence between two correlation matrices. Specifically, we minimize the Kullback–Leibler divergence between the empirical $\hat{p}(\tilde{\mathbf{y}}[n])$ and theoretical $p(\tilde{\mathbf{y}}[n]|\mathbf{H},E,\sigma^2)$ pdfs of the observations, which is closely related to the notion of information geometry [25] and to the ML estimation of the parameters [25]–[27].

Omitting, for notational simplicity, the dependence w.r.t. E and σ^2, the Kullback–Leibler divergence is defined as

$$D\left(\hat{p}\left(\tilde{\mathbf{y}}[n]\right);p\left(\tilde{\mathbf{y}}[n]|\mathbf{H}\right)\right)=\int_{\tilde{\mathbf{y}}[n]}\hat{p}(\tilde{\mathbf{y}}[n])\log\frac{\hat{p}(\tilde{\mathbf{y}}[n])}{p(\tilde{\mathbf{y}}[n]|\mathbf{H})}d\tilde{\mathbf{y}}[n]$$

and in the case of Gaussian distributed observations it reduces to [27]

$$\begin{aligned}D\left(\hat{p}\left(\tilde{\mathbf{y}}[n]\right);p\left(\tilde{\mathbf{y}}[n]|\mathbf{H}\right)\right)&\\=\frac{1}{2}\mathrm{Tr}\left(\mathbf{R}_{\tilde{\mathbf{y}}}^{-1}(\mathbf{H})\hat{\mathbf{R}}_{\tilde{\mathbf{y}}}-\mathbf{I}_{2Ln_R}\right)&\\-\frac{1}{2}\log\det\left(\mathbf{R}_{\tilde{\mathbf{y}}}^{-1}(\mathbf{H})\hat{\mathbf{R}}_{\tilde{\mathbf{y}}}\right).&\end{aligned}$$

Based on the previous definitions, the proposed KCM criterion amounts to solve the following optimization problem

$$\hat{\mathbf{H}}^{\mathrm{KCM}}=\arg\min_{\mathbf{H}}\left[\mathrm{Tr}\left(\mathbf{R}_{\tilde{\mathbf{y}}}^{-1}(\mathbf{H})\hat{\mathbf{R}}_{\tilde{\mathbf{y}}}\right)+\log\det\left(\mathbf{R}_{\tilde{\mathbf{y}}}(\mathbf{H})\right)\right]\tag{13}$$

which reduces to the ML estimator under Gaussian distributed nuisance parameters $\mathbf{s}[n]$ [26], [28]. Although the Gaussian assumption is only strictly correct in the asymptotic cases of $\sigma^2\to\infty$ or $M'\to\infty$ independent sources, it has been recently proven [29] that, under multilevel constellations, (13) provides (asymptotically as $\sigma^2\to 0$) the optimum second-order estimator [28].

In order to simplify (13), let us introduce the following lemmas:

Lemma 1: Under OSTBC transmissions, the inverse of the theoretical correlation matrix $\mathbf{R}_{\tilde{\mathbf{y}}}(\mathbf{H})$ is given by

$$\mathbf{R}_{\tilde{\mathbf{y}}}^{-1}(\mathbf{H})=\frac{1}{\sigma^2/2}\mathbf{I}_{2Ln_R}-\frac{E}{(\sigma^2/2)^2}\tilde{\mathbf{W}}(\mathbf{H})\mathbf{\Gamma}^{\mathrm{KCM}}\tilde{\mathbf{W}}^T(\mathbf{H})\tag{14}$$

where $\mathbf{\Gamma}^{\mathrm{KCM}}=\mathrm{diag}\left(\left[\rho_1^{\mathrm{KCM}},\dots,\rho_{M'}^{\mathrm{KCM}}\right]\right)$, and

$$\rho_k^{\mathrm{KCM}}=\frac{\lambda_k}{1+\lambda_k\frac{E}{\sigma^2/2}},\quad k=1,\dots,M'.\tag{15}$$

Proof: From (9), and taking into account (3), we can write

$$\mathbf{R}_{\tilde{\mathbf{y}}}(\mathbf{H})=E\tilde{\mathbf{W}}(\mathbf{H})\left(\mathbf{\Lambda}_{\mathbf{s}}+\frac{\sigma^2}{2E}\mathbf{I}_{M'}\right)\tilde{\mathbf{W}}^T(\mathbf{H})+\frac{\sigma^2}{2}\tilde{\mathbf{W}}_\perp(\mathbf{H})\tilde{\mathbf{W}}_\perp^T(\mathbf{H})$$

where $\tilde{\mathbf{W}}_\perp(\mathbf{H})$ is the orthogonal complement of $\tilde{\mathbf{W}}(\mathbf{H})$ and we have used $\mathbf{I}_{2Ln_R}=\tilde{\mathbf{W}}(\mathbf{H})\tilde{\mathbf{W}}^T(\mathbf{H})+\tilde{\mathbf{W}}_\perp(\mathbf{H})\tilde{\mathbf{W}}_\perp^T(\mathbf{H})$. Thus, the inverse is

$$\mathbf{R}_{\tilde{\mathbf{y}}}^{-1}(\mathbf{H})=\frac{1}{E}\tilde{\mathbf{W}}(\mathbf{H})\left(\mathbf{\Lambda}_{\mathbf{s}}+\frac{\sigma^2}{2E}\mathbf{I}_{M'}\right)^{-1}\tilde{\mathbf{W}}^T(\mathbf{H})+\frac{1}{\sigma^2/2}\tilde{\mathbf{W}}_\perp(\mathbf{H})\tilde{\mathbf{W}}_\perp^T(\mathbf{H})$$

and rewriting $\tilde{\mathbf{W}}_\perp(\mathbf{H})\tilde{\mathbf{W}}_\perp^T(\mathbf{H})=\mathbf{I}_{2Ln_R}-\tilde{\mathbf{W}}(\mathbf{H})\tilde{\mathbf{W}}^T(\mathbf{H})$, we obtain (14), where

$$\begin{aligned}\mathbf{\Gamma}^{\mathrm{KCM}}&=\frac{(\sigma^2/2)^2}{E}\left(\frac{1}{\sigma^2/2}\mathbf{I}_{M'}-\frac{1}{E}\left(\mathbf{\Lambda}_{\mathbf{s}}+\frac{\sigma^2}{2E}\mathbf{I}_{M'}\right)^{-1}\right)\\&=\frac{\sigma^2/2}{E}\left(\mathbf{I}_{M'}-\left(\mathbf{I}_{M'}+\frac{E}{\sigma^2/2}\mathbf{\Lambda}_{\mathbf{s}}\right)^{-1}\right)\\&=\mathbf{\Lambda}_{\mathbf{s}}\left(\mathbf{I}_{M'}+\frac{E}{\sigma^2/2}\mathbf{\Lambda}_{\mathbf{s}}\right)^{-1}\end{aligned}$$

is a diagonal matrix defined by the weights in (15). ■

Lemma 2: Under OSTBC transmissions, the determinant of the theoretical correlation matrix $\mathbf{R}_{\tilde{\mathbf{y}}}(\mathbf{H})$ is given by

$$\det\left(\mathbf{R}_{\tilde{\mathbf{y}}}(\mathbf{H})\right)=\left(\frac{\sigma^2}{2}\right)^{2Ln_R-M'}\prod_{k=1}^{M'}\left(E\lambda_k+\frac{\sigma^2}{2}\right).$$

Proof: The proof follows directly from the product of the eigenvalues of $\mathbf{R}_{\tilde{\mathbf{y}}}(\mathbf{H})$ [see (9)]. ■

Now, applying the results of Lemmas 1 and 2, the criterion in (13) can be rewritten as

$$\hat{\mathbf{H}}^{\mathrm{KCM}}=\arg\max_{\mathbf{H}}\Psi^{\mathrm{KCM}}$$

where

$$\begin{aligned}\Psi^{\mathrm{KCM}}=&\frac{E}{(\sigma^2/2)^2}\mathrm{Tr}\left(\mathbf{\Gamma}^{\mathrm{KCM}}\tilde{\mathbf{W}}^T(\mathbf{H})\hat{\mathbf{R}}_{\tilde{\mathbf{y}}}\tilde{\mathbf{W}}(\mathbf{H})\right)\\&-\frac{1}{\sigma^2/2}\mathrm{Tr}\left(\hat{\mathbf{R}}_{\tilde{\mathbf{y}}}\right)-\sum_{k=1}^{M'}\log\left(\lambda_kE+\frac{\sigma^2}{2}\right)\\&-(2Ln_R-M')\log\frac{\sigma^2}{2}.\end{aligned}\tag{16}$$

Analogously to the ECM case, (16) only depends on $\mathbf{H}$ through the term $\mathrm{Tr}\left(\mathbf{\Gamma}^{\mathrm{KCM}}\tilde{\mathbf{W}}^T(\mathbf{H})\hat{\mathbf{R}}_{\tilde{\mathbf{y}}}\tilde{\mathbf{W}}(\mathbf{H})\right)$, which can be rewritten as

$$\mathrm{Tr}\left(\mathbf{\Gamma}^{\mathrm{KCM}}\tilde{\mathbf{W}}^T(\mathbf{H})\hat{\mathbf{R}}_{\tilde{\mathbf{y}}}\tilde{\mathbf{W}}(\mathbf{H})\right)=\tilde{\mathbf{h}}^T\mathbf{\Phi}^{\mathrm{KCM}}\tilde{\mathbf{h}}=\beta^{\mathrm{KCM}}$$

where

$$\mathbf{\Phi}^{\mathrm{KCM}}=\sum_{k=1}^{M'}\rho_k^{\mathrm{KCM}}\tilde{\mathbf{D}}_k^T\hat{\mathbf{R}}_{\tilde{\mathbf{y}}}\tilde{\mathbf{D}}_k$$

and the weights ρ_k^{KCM} are given by (15). Finally, the KCM estimate of $\tilde{\mathbf{h}}$ is obtained as the eigenvector $\hat{\tilde{\mathbf{h}}}^{\mathrm{KCM}}$ associated to the largest eigenvalue $\hat{\beta}^{\mathrm{KCM}}$ of

$$\mathbf{\Phi}^{\mathrm{KCM}}\hat{\tilde{\mathbf{h}}}^{\mathrm{KCM}}=\hat{\beta}^{\mathrm{KCM}}\hat{\tilde{\mathbf{h}}}^{\mathrm{KCM}}.$$

V. Channel Energy and Noise Variance Estimation

Although the estimation of the channel energy E and noise variance σ^2 can be considered as a secondary problem, the weights of the KCM estimate depend on the ratio E/σ^2. In this section, we obtain the estimates of E and σ^2 under the ECM and KCM criteria.

A. ECM Estimates

In order to find the ECM estimates of E and σ^2, and following the derivation in Section IV-A, we obtain the optimization problem

$$\left\{\hat{E}, \hat{\sigma}^2\right\} = \arg\max_{E,\sigma^2} \Psi^{\mathrm{ECM}}.$$

Equating to zero the derivative of Ψ^{ECM} w.r.t. E and σ^2 the following system of linear equations is obtained

$$E = \frac{\beta^{\mathrm{ECM}} - \frac{\sigma^2}{2}\mathrm{Tr}\left(\mathbf{\Lambda_s}\right)}{||\mathbf{\Lambda_s}||^2} \tag{17}$$

$$\sigma^2 = \frac{\mathrm{Tr}\left(\hat{\mathbf{R}}_{\tilde{\mathbf{y}}}\right) - E\,\mathrm{Tr}\left(\mathbf{\Lambda_s}\right)}{Ln_R} \tag{18}$$

whose solutions are given by

$$\hat{E} = \frac{2Ln_R\beta^{\mathrm{ECM}} - \mathrm{Tr}\left(\mathbf{\Lambda_s}\right)\mathrm{Tr}\left(\hat{\mathbf{R}}_{\tilde{\mathbf{y}}}\right)}{2Ln_R||\mathbf{\Lambda_s}||^2 - \mathrm{Tr}^2\left(\mathbf{\Lambda_s}\right)} \tag{19}$$

$$\hat{\sigma}^2 = 2\frac{||\mathbf{\Lambda_s}||^2\mathrm{Tr}\left(\hat{\mathbf{R}}_{\tilde{\mathbf{y}}}\right) - \beta^{\mathrm{ECM}}\mathrm{Tr}\left(\mathbf{\Lambda_s}\right)}{2Ln_R||\mathbf{\Lambda_s}||^2 - \mathrm{Tr}^2\left(\mathbf{\Lambda_s}\right)}. \tag{20}$$

As can be seen, the estimates of the channel energy and noise variance only depend on the normalized channel through the term β^{ECM}, which is independent of E and σ^2. Therefore, the exact ECM solution for the joint estimate of $\mathbf{H}$, E and σ^2 can be obtained by first solving the EV in (12) and then substituting β^{ECM} in (19) and (20) by its estimate $\hat{\beta}^{\mathrm{ECM}} = \mathrm{Tr}\left(\mathbf{\Gamma}^{\mathrm{ECM}}\tilde{\mathbf{W}}^T(\hat{\mathbf{H}}^{\mathrm{ECM}})\hat{\mathbf{R}}_{\tilde{\mathbf{y}}}\tilde{\mathbf{W}}(\hat{\mathbf{H}}^{\mathrm{ECM}})\right)$.

Finally, we must point out that the necessary and sufficient condition for the solvability of the linear system given by (17) and (18) is

$$2Ln_R||\mathbf{\Lambda_s}||^2 - \mathrm{Tr}^2\left(\mathbf{\Lambda_s}\right) \neq 0$$

and it is easy to prove that

$$2Ln_R||\mathbf{\Lambda_s}||^2 - \mathrm{Tr}^2\left(\mathbf{\Lambda_s}\right) = 2Ln_R\sum_{k=1}^{M'}\lambda_k^2 - \left(\sum_{k=1}^{M'}\lambda_k\right)^2 \geq 0$$

where the equality is satisfied iff

$$M' = 2Ln_R, \text{ and } \lambda_k = \lambda \quad \forall\, k.$$

In other words, the linear system in (17) and (18) is always solvable with the only exception of transmitting a white source ($\mathbf{\Lambda_s} = P_T/n_T\mathbf{I}$) by means of the complex Alamouti code ($M' = 2L$) [1], and receiving with only one receive antenna. Clearly, this is a very special case for which

$$\mathbf{R}_{\tilde{\mathbf{y}}} = \frac{P_T E + \sigma^2}{2}\mathbf{I}_{2Ln_R}.$$

B. KCM Estimates

Analogously to the ECM case, the KCM estimates of E and σ^2 are obtained as

$$\left\{\hat{E}, \hat{\sigma}^2\right\} = \arg\max_{E,\sigma^2} \Psi^{\mathrm{KCM}}.$$

Thus, defining γ_k $(k = 1, \ldots, M')$ as the kth element along the diagonal of $\tilde{\mathbf{W}}^T(\mathbf{H})\hat{\mathbf{R}}_{\tilde{\mathbf{y}}}\tilde{\mathbf{W}}(\mathbf{H})$, and after some tedious but straightforward algebra, we obtain the following system of nonlinear equations

$$E = \frac{\sum_{k=1}^{M'}\rho_k^{\mathrm{KCM}}\left(\gamma_k - \frac{\sigma^2}{2}\right)}{\sum_{k=1}^{M'}\rho_k^{\mathrm{KCM}}\lambda_k} \tag{21}$$

$$\sigma^2 = \frac{\mathrm{Tr}\left(\hat{\mathbf{R}}_{\tilde{\mathbf{y}}}\right) - \frac{E}{\sigma^2/2}\sum_{k=1}^{M'}\gamma_k\rho_k^{\mathrm{KCM}}}{Ln_R} \tag{22}$$

which can be seen as the equivalent to the linear system in (17) and (18). Specifically, taking into account that $\rho_k^{\mathrm{ECM}} = \lambda_k$, the right hand side term in (17) can be rewritten as

$$\frac{\beta^{\mathrm{ECM}} - \frac{\sigma^2}{2}\mathrm{Tr}\left(\mathbf{\Lambda_s}\right)}{||\mathbf{\Lambda_s}||^2} = \frac{\sum_{k=1}^{M'}\rho_k^{\mathrm{ECM}}\left(\gamma_k - \frac{\sigma^2}{2}\right)}{\sum_{k=1}^{M'}\rho_k^{\mathrm{ECM}}\lambda_k}$$

which is analogous to (21), i.e., both estimates can be seen as the ratio between weighted versions (with weights ρ_k^{ECM} or ρ_k^{KCM}) of the total received and transmitted signal energy.

On the other hand, unlike the estimate in (8), which is solely based on the noise subspace of $\hat{\mathbf{R}}_{\tilde{\mathbf{y}}}$, the ECM and KCM (18) and (22) can be easily interpreted as an average of the received noise energy over the n_R receive antennas and L channel uses, i.e., they take into account both the signal and noise subspaces. Furthermore, in the asymptotic case of $N \to \infty$, we have $\gamma_k = \lambda_k E + \sigma^2/2$, and the term $2E/\sigma^2\sum_{k=1}^{M'}\gamma_k\rho_k^{\mathrm{KCM}}$ in (22) can be rewritten as

$$\frac{E}{\sigma^2/2}\sum_{k=1}^{M'}\gamma_k\rho_k^{\mathrm{KCM}} = E\sum_{k=1}^{M'}\lambda_k = E\,\mathrm{Tr}\left(\mathbf{\Lambda_s}\right)$$

which shows the equivalence between (22) and (18).

Finally, we must point out that, unlike the ECM case, the estimates of $\mathbf{H}$, E and σ^2 are coupled through the weights ρ_k^{KCM}, and the exact KCM solutions should be obtained by means of an iterative technique. However, in Section VII, it is shown, by means of numerical examples, that the estimates of $\mathbf{H}$ using either the exact or estimated ratios E/σ^2 are practically identical. Therefore, from a practical point of view we propose to obtain $\hat{\mathbf{H}}^{\mathrm{KCM}}$ using the subspace or ECM estimates of E and σ^2, and then solve the nonlinear system in (21) and (22) by means of a few iterations.

VI. Analysis of the Proposed Techniques and Further Discussion

The proposed ECM and KCM criteria, as well as the relaxed blind ML decoder and its weighted version, lead to similar EV problems, which only differ in the selection of the weights ρ_k, ρ_k^{ECM} or ρ_k^{KCM}. The analysis of these weights help us to establish clear relationships among the four different approaches to blind OSTBC channel estimation. The main conclusions of this analysis can be summarized as follows.

- In the case of equal power and uncorrelated sources ($\lambda_k = \lambda,\ \forall\ k$), the ECM, the KCM and the relaxed blind ML decoder are equivalent. Therefore, they are affected by the same ambiguity problems [see (6)].
- Both the ECM and KCM matrices $\mathbf{\Phi}^{\mathrm{ECM}}$ and $\mathbf{\Phi}^{\mathrm{KCM}}$ can be viewed as particular cases of $\mathbf{\Phi}$ in (7). Therefore, the proposed methods provide two different criteria for the selection of the weights and shed some light into how to choose the optimal weights for the linear precoding technique proposed in [10].
- As previously pointed out, the KCM weights ρ_k^{KCM} do not only depend on the source eigenvalues, but also on E/σ^2, which is proportional to the instantaneous signal to noise ratio. Thus, the behavior of the KCM solution can be easily interpreted.
  - In the low SNR regime, $\rho_k^{\mathrm{KCM}} = \lambda_k$ and the KCM criterion is equivalent to the ECM method, i.e., both techniques try to extract the channel state information from the prior knowledge of $\mathbf{\Lambda}_{\mathbf{s}}$.
  - In the high SNR regime, the KCM criterion is asymptotically equivalent to the relaxed blind ML decoder ($\rho_k^{\mathrm{KCM}} \simeq \sigma^2/(2E),\ \forall\ k$). Roughly speaking, this means that, instead of exploiting the information about the source correlation matrix (which is inaccurate due to the finite number of observations), the channel is (almost) exclusively extracted from the congruence between the observations and the OSTBC data model.
- Due to the asymptotic equivalence (for $\sigma \rightarrow 0$) between the KCM approach and the relaxed blind ML decoder, the performance of the KCM technique in the high SNR regime can be summarized as follows.
  - In the identifiable case[4] [19], [20], the KCM technique is able to exactly recover the MIMO channel. Unlike the weighted version of the relaxed blind ML decoder and the ECM criterion, it is not affected by the noise floor due to the difference between the theoretical source correlation matrix $\mathbf{R}_{\mathbf{s}} = \mathcal{E}[\mathbf{s}[n]\mathbf{s}^T[n]]$ and its finite sample version
  $$\hat{\mathbf{R}}_{\mathbf{s}} = \frac{1}{N}\sum_{n=0}^{N-1}\mathbf{s}[n]\mathbf{s}^T[n].$$
  - In the nonidentifiable case, the KCM criterion will remove the ambiguity by exploiting the knowledge of the source correlation matrix $\mathbf{R}_{\mathbf{s}}$.[5] Thus, unlike the relaxed blind ML decoder, the KCM criterion will be able to unambiguously recover the channel. Furthermore, although the differences between theoretical and empirical matrices now provoke the above referred noise floor problem, it will be shown by means of simulations that, in general, the noise floor in the KCM case is lower than that of the ECM approach.
- Finally, in [44], it has been shown that the weighted matrix $\mathbf{\Phi}$ can be interpreted as a modified correlation matrix, and therefore the blind channel estimation criterion can be reformulated as a principal component analysis (PCA) problem [46]. Thus, in [44] and [45] we have proposed adaptive blind channel estimation algorithms based on the direct application of the well-known Oja's rule [47].

[4]We refer to the identifiable case when (6) is only satisfied by scaled versions of the actual MIMO channel.

[5]Note that, as shown in the simulations section, the small differences in the weights ρ_k^{KCM} still avoid the identifiability problem pointed out in (6).

VII. Simulation Results

In this section, the performance of the proposed techniques is illustrated by means of some numerical examples. In all the cases the information symbols belong to a QPSK constellation ($s_k[n]$ are BPSK signals) and they are transmitted with unit power by channel use ($P_T = 1$)), which implies an instantaneous signal to noise ratio $\mathrm{SNR}_{\mathbf{H}} = E/\left(n_T n_R \sigma^2\right)$. The MIMO channel follows a Rayleigh distribution, i.e., each element $h_{i,j}$ is a complex Gaussian random variable with zero mean and unit variance. Therefore, the average SNR is defined as $\mathrm{SNR} = 1/\sigma^2$.

The performance of the ECM and KCM approaches has been evaluated with the well known Alamouti code [1] and with the complex OSTBC proposed in [40] (see also [41, Ch. 3]), whose parameters are $n_T = L = 8$ and $M = 4$ ($M' = 8$ and $R = 1/2$). This code was designed to provide better peak to average power ratio (PAR) than that of the conventional designs for $n_T = 8$ [34], and its transmission matrix is given by

$$\mathbf{S}[n] = \begin{bmatrix} \mathbf{S}_1[n] & \mathbf{S}_2[n] \\ \mathbf{S}_3[n] & -\mathbf{S}_1^H[n] \end{bmatrix} \tag{23}$$

where, omitting the temporal index $[n]$

$$\mathbf{S}_1 = \frac{1}{2}\begin{bmatrix} s_1+js_2 & s_1+js_2 & s_1+js_2 & s_1+js_2 \\ s_1+js_2 & -s_1+js_2 & s_1+js_2 & -s_1+js_2 \\ s_1+js_2 & s_1+js_2 & -s_1+js_2 & -s_1+js_2 \\ s_1+js_2 & -s_1+js_2 & -s_1+js_2 & s_1+js_2 \end{bmatrix}$$
$$\mathbf{S}_2 = \begin{bmatrix} s_3+js_4 & 0 & s_7+js_6 & s_5+js_8 \\ 0 & s_3+js_4 & -s_5+js_8 & s_7-js_6 \\ -s_7+js_6 & s_5+js_8 & s_3-js_4 & 0 \\ -s_5+js_8 & -s_7-js_6 & 0 & s_3-js_4 \end{bmatrix}$$
$$\mathbf{S}_3 = \begin{bmatrix} s_3-js_8 & s_5-js_6 & s_7-js_4 & 0 \\ -s_5-js_6 & s_3+js_8 & 0 & s_7-js_4 \\ -s_7-js_4 & 0 & s_3+js_8 & -s_5+js_6 \\ 0 & -s_7-js_4 & s_5+js_6 & s_3-js_8 \end{bmatrix}.$$

For this code, it has been proven in [41, Ch. 4] that, in order to optimize the bit error rate (BER) under QPSK constellations and Rayleigh channels, the source correlation matrix should be

$$\mathbf{\Lambda}_{\mathbf{s}} = \frac{1}{12}\mathrm{diag}([3,3,1,1,1,1,1,1])$$

i.e., the energy of the first complex symbol is three times higher than that of the three remaining symbols. As follows from the

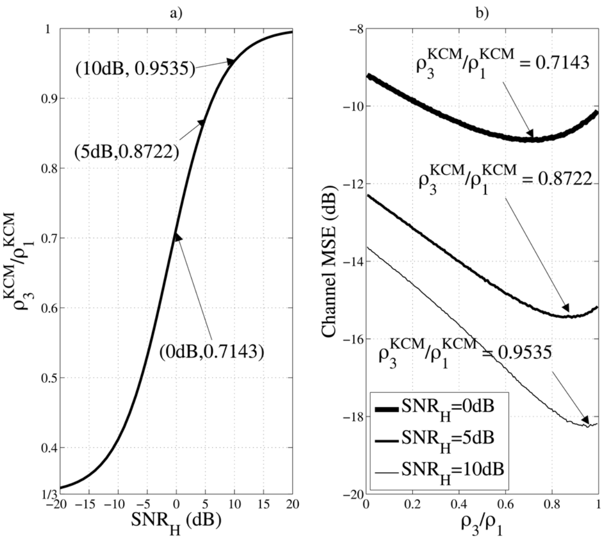

Fig. 1. Analysis of the KCM criterion with the OSTBC of [40], $n_T = 8$ and $n_R = 1$. a) Evolution of the ratio $\rho_3^{\mathrm{KCM}}/\rho_1^{\mathrm{KCM}}$ with the instantaneous SNR. b) MSE in the channel estimate for different ratios ρ_3/ρ_1, $N = 50$.

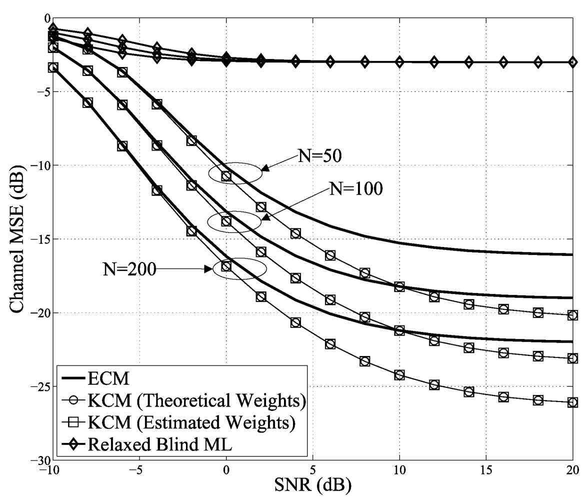

Fig. 2. MSE in the estimate of the normalized channel. Nonidentifiable case (complex OSTBC in (23), $n_T = 8, n_R = 1$).

ECM and KCM techniques, this suggests the use of only two different weights, i.e., in all the cases (ECM, KCM, and the weighted version of the relaxed blind ML decoder) we fixed

$$\rho_1 = \rho_2, \quad \rho_3 = \cdots = \rho_8.$$

A. Complex OSTBC With $n_T = 8$ and $n_R = 1$

In the first set of experiments we consider the complex OSTBC in (23). Here, we must point out that in the MISO case $(n_R = 1)$ this code is affected by the indeterminacy pointed out in (6), i.e., the channel can not be unambiguously recovered exclusively from the congruence between the observations and the OSTBC data model.

The results of the first experiment can be seen in Fig. 1(a), which shows the evolution of the ratio $\rho_3^{\mathrm{KCM}}/\rho_1^{\mathrm{KCM}}$ with the instantaneous SNR. As can be seen, this value ranges from 1/3 (low $\mathrm{SNR}_{\mathbf{H}}$), which is equivalent to the ECM approach, to 1 (high $\mathrm{SNR}_{\mathbf{H}}$), which matches the relaxed blind ML decoder. Additionally, Fig. 1(b) represents the MSE in the estimate of $\mathbf{H}$ as a function of the ratio ρ_3/ρ_1, where we can see that the minimum MSE is obtained with the value $\rho_3^{\mathrm{KCM}}/\rho_1^{\mathrm{KCM}}$ provided by the KCM criterion.

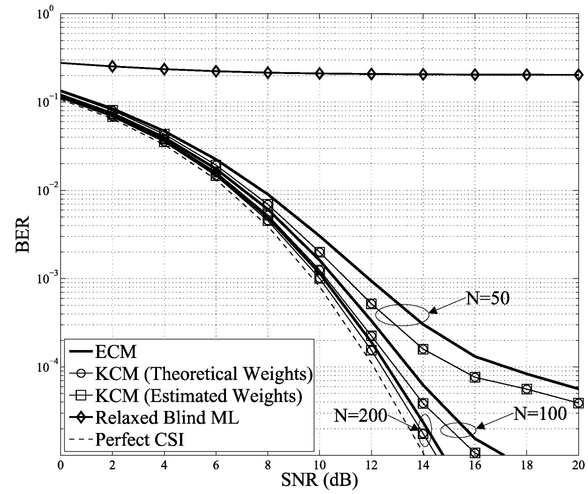

Fig. 3. BER after decoding. Nonidentifiable case (complex OSTBC in (23), $n_T = 8, n_R = 1$).

In the second example, the ECM and KCM techniques have been compared with the ML receiver with perfect channel state information (CSI) and the relaxed blind ML decoder. The performance of the KCM approach has been evaluated with the exact knowledge of the ratio E/σ^2 and with its estimate based on the subspace method,[6] obtaining almost identical results. Fig. 2 shows the mean square error (MSE) in the channel estimate for different numbers (N) of available blocks at the receiver. As can be seen, the relaxed blind ML decoder is not able to recover the channel, whereas the ECM and KCM approaches are affected by a noise floor, which rapidly decreases with N, due to the difference between the theoretical and empirical source correlation matrices. On the other hand, since the KCM wisely combines the prior information about the source correlation with the congruence between the observations and the data model, its noise floor is lower than that of the ECM criterion.

Finally, Fig. 3 shows the BER after blind channel estimation and decoding, where we can see that, for a moderate number of available blocks at the receiver and practical SNR (or BER) values, the performance of the proposed blind schemes is very close to that of the receiver with perfect CSI, avoiding the 3-dB penalty associated to differential or training-based approaches with one block of pilots. Furthermore, we must point out that in a practical scenario, the noise floor in the previous figures could be easily avoided by refining the channel estimates using the decoded symbols as training sequences.

B. Complex OSTBC With $n_T = 8$ and $n_R = 2$

In this subsection we present the results obtained, for the complex OSTBC in (23), with $n_R = 2$ receive antennas. Unlike the previous case, when the number of receive antennas is $n_R \geq 2$,

[6] Similar results have been obtained with the ECM estimate of E/σ^2.

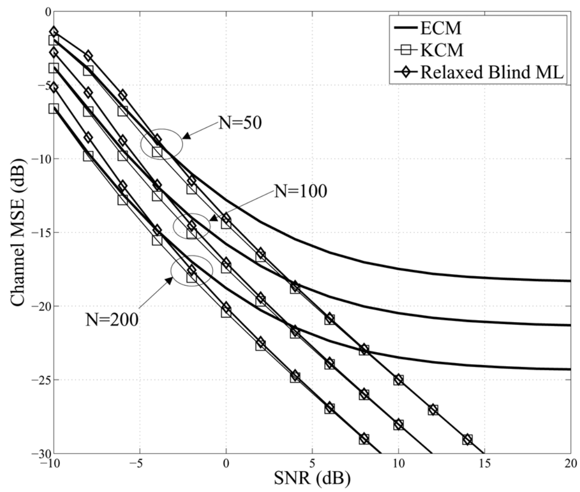

Fig. 4. MSE in the estimate of the normalized channel. Identifiable case (complex OSTBC in (23), $n_T = 8, n_R = 2$).

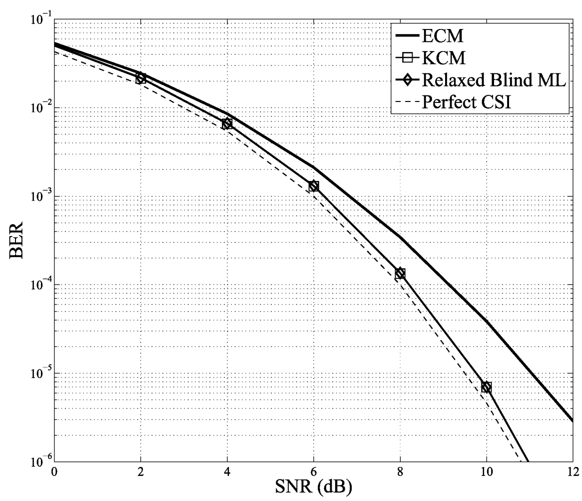

Fig. 5. BER after decoding ($N = 50$). Identifiable case (complex OSTBC in (23), $n_T = 8, n_R = 2$).

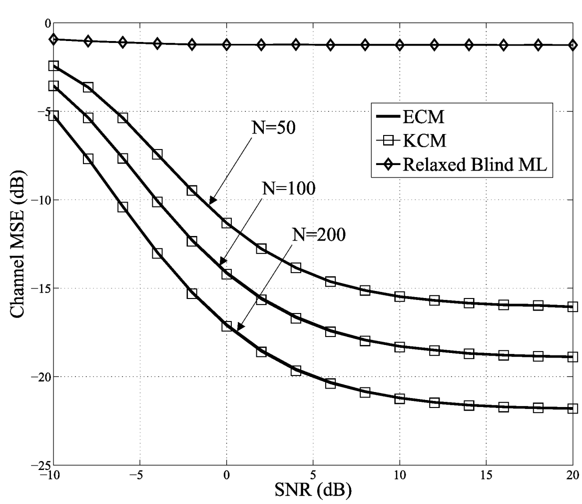

Fig. 6. MSE in the estimate of the normalized channel. Alamouti code with $n_T = n_R = 2$.

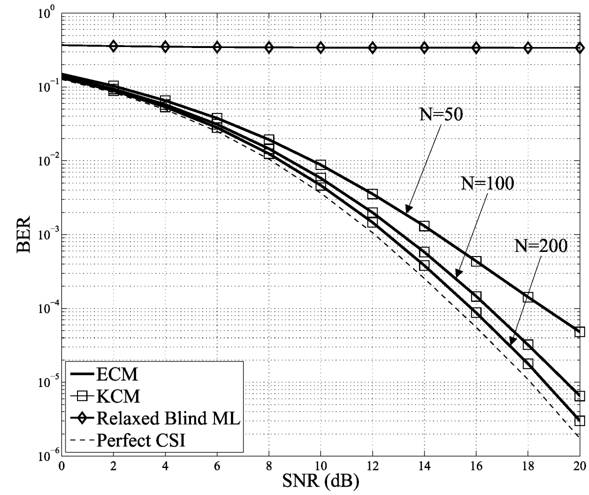

Fig. 7. BER after decoding. Alamouti code with $n_T = n_R = 2$.

it can be proven that the channel can be unambiguously recovered by means of the relaxed blind ML decoder [19], [20], i.e., the equality in (6) is only satisfied by scaled versions of the true MIMO channel. The MSE in the channel estimate and the BER after decoding are shown in Figs. 4 and 5, where we can see that the KCM approach, which is asymptotically equivalent to the relaxed blind ML decoder, is not affected by the noise floor.

C. Alamouti Code With $n_T = n_R = 2$

In the final set of examples, the proposed techniques have been evaluated for the Alamouti code [1] with $n_R = 2$ receive antennas. As previously pointed out, the Alamouti code does not allow the unambiguous recovery of the channel without exploiting the correlation or other properties of the sources. Here, in order to avoid the ambiguity problems we have linearly precoded the sources to obtain the following source correlation matrix

$$\mathbf{\Lambda_s} = \frac{1}{7}\mathrm{diag}\left([4,1,1,1]\right)$$

i.e., the first BPSK symbol is transmitted with four times more power than the remaining ones. The results are shown in Figs. 6 and 7, where we can see that, for this particular source correlation matrix, the ECM and KCM approaches provide almost identical results.

Finally, the ECM and KCM estimates of the channel energy and noise variance have been compared with those of the subspace-based approach summarized in Section III-A. In particular, the KCM technique is initialized with the estimate of E/σ^2 provided by the subspace method, which is used to obtain $\hat{\mathbf{H}}^{\mathrm{KCM}}$, and E and σ^2 have been estimated by means of only three iterations of the expressions in (21) and (22). The results are shown in Fig. 8, where we can see that the best results are provided by the KCM technique. Specifically, in the low SNR regime the best estimates are obtained by the ECM and KCM approaches, which exploit the prior knowledge of the correlation matrix $\mathbf{\Lambda_s}$. On the other hand, in the high SNR regime, the ECM technique is affected by a noise floor, which is avoided by the KCM and subspace methods. In summary, the KCM estimates of the noise variance and channel norm exploit

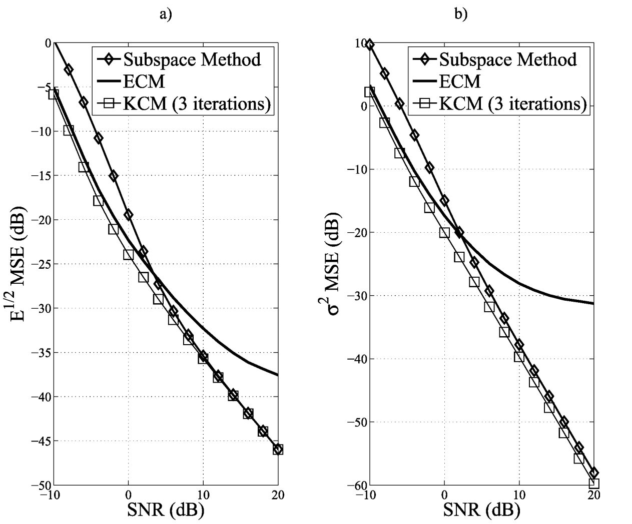

Fig. 8. Performance of the proposed techniques in the Alamouti case, $n_T = n_R = 2, N = 50$. a) MSE in the estimate of $\sqrt{E}$ b) MSE in the estimate of the noise variance σ^2.

all the information provided by the matrix $\hat{\mathbf{R}}_{\tilde{\mathbf{y}}}$, which translates into more accurate results than those of the subspace-based approach.

VIII. Conclusion

In this paper, the correlation matching criterion has been applied to the problem of blind channel estimation under orthogonal space–time block coded (OSTBC) transmissions. The proposed techniques are based on the minimization of the divergence between the theoretical and estimated correlation matrices of the observations. In particular, we have considered the Euclidean distance and the Kullback–Leibler divergence, which lead, respectively, to the Euclidean (ECM) and Kullback correlation matching (KCM) approaches. Interestingly, due to the particular OSTBC structure, the solutions of both criteria can be obtained in closed form, and the channel estimate is obtained as the solution of an eigenvalue problem, which is formed from the correlation matrix of the observations modified by the code matrices and a set of weights. In general, the KCM technique, whose weights depend on the SNR, outperforms the ECM approach, which has fixed weights. Finally, it has also been proven that, in the low and high SNR cases, the KCM method is asymptotically equivalent to the ECM criterion and the relaxed blind ML decoder, respectively.

References

[1] S. Alamouti, "A simple transmit diversity technique for wireless communications," *IEEE J. Sel. Areas Commun.*, vol. 45, no. 8, pp. 1451–1458, Oct. 1998.

[2] 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.

[3] G. Foschini and M. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," *Wireless Personal Commun.*, vol. 6, pp. 311–335, 1998.

[4] I. E. Telatar, "Capacity of multi-antenna Gaussian channels," *Eur. Trans. Telecommun.*, vol. 10, no. 6, pp. 585–595, Nov.–Dec. 1999.

[5] A. F. Naguib, V. Tarokh, N. Seshadri, and A. R. Calderbank, "A space–time coding modem for high-data-rate wireless communications," *IEEE J. Sel. Areas Commun.*, vol. 16, no. 8, pp. 1459–1478, Oct. 1998.

[6] 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.

[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] 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, no. 2, pp. 738–751, Feb. 2006.

[9] 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.

[10] 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.

[11] N. Ammar and Z. Ding, "Channel identifiability under orthogonal space–time coded modulations without training," *IEEE Trans. Wireless Commun.*, vol. 5, no. 5, pp. 1003–1013, May 2006.

[12] S. Shahbazpanahi, A. B. Gershman, and G. B. Giannakis, "Semiblind multiuser MIMO channel estimation using capon and music techniques," *IEEE Trans. Signal Process.*, vol. 54, no. 9, pp. 3581–3591, Sep. 2006.

[13] N. Ammar and Z. Ding, "Blind channel identifiability for generic linear space–time block codes," *IEEE Trans. Signal Process.*, vol. 55, no. 1, pp. 202–217, Jan. 2007.

[14] V. Tarokh and H. Jafarkhani, "A differential detection scheme for transmit diversity," *IEEE J. Sel. Areas Commun.*, vol. 18, no. 7, pp. 1169–1174, Jul. 2000.

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

[16] B. Hochwald and W. Sweldens, "Differential unitary space–time modulation," *IEEE Trans. Commun.*, vol. 48, pp. 2041–2052, Dec. 2000.

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

[18] G. Ganesan and P. Stoica, "Differential modulation using space–time block codes," *IEEE Signal Process. Lett.*, vol. 9, pp. 57–60, Feb. 2002.

[19] J. Vía and I. Santamaría, "Some results on the blind identifiability of orthogonal space–time block codes from second order statistics," presented at the IEEE Int. Conf. Acoustic, Speech, Signal Processing (ICASSP), Honolulu, HI, Apr. 2007.

[20] J. Vía and I. Santamaría, "On the blind identifiability of orthogonal space–time block codes from second order statistics," *IEEE Trans. Inf. Theory*, vol. 54, pp. 709–722, Feb. 2008.

[21] W.-K. Ma, "Blind ML detection of orthogonal space–time block codes: Identifiability and code construction," *IEEE Trans. Signal Process.*, vol. 55, no. 7, pp. 3312–3324, Jul. 2007.

[22] G. Giannakis and S. Halford, "Asymptotically optimal blind fractionally-spaced channel estimation and performance analysis," *IEEE Trans. Signal Process.*, vol. 45, no. 7, pp. 1815–1830, Jul. 1997.

[23] H. H. Zeng and L. Tong, "Blind channel estimation using the second-order statistics: Asymptotic performance and limitations," *IEEE Trans. Signal Process.*, vol. 45, no. 8, pp. 2060–2071, Aug. 1997.

[24] H. Gazzah, P. A. Regalia, J.-P. Delmas, and K. Abed-Meraim, "A blind multichannel identification algorithm robust to order overestimation," *IEEE Trans. Signal Process.*, vol. 50, no. 6, pp. 1449–1458, Jun. 2002.

[25] I. Csiszar, "I-divergence, geometry of probability distributions and minimization problems," *Ann. Probab.*, vol. 3, no. 1, pp. 146–158, 1975.

[26] B. Ottersten, M. Viberg, and P. Stoica, "Exact and large sample maximum likelihood techniques for parameter estimation and detection," in *Radar Array Processing*. New York: Springer-Verlag, 1993, ch. 4.

[27] J. Cardoso, *Unsupervised Adaptive Filters*. New York: Wiley, 2000, vol. 1, ch. Entropic Contrasts for Source Separation: Geometry snd Stability, pp. 139–190.

[28] J. Villares and G. Vazquez, "Second-order parameter estimation," *IEEE Trans. Signal Process.*, vol. 53, no. 7, pp. 2408–2420, Jul. 2005.

[29] J. Villares and G. Vazquez, "The Gaussian assumption in second-order estimation problems in digital communications," *IEEE Trans. Signal Process.*, vol. 55, no. 10, pp. 4994–5002, Oct. 2007.
[30] B. Picinbono, "On circularity," *IEEE Trans. Signal Process.*, vol. 42, no. 12, pp. 3473–3482, Dec. 1994.
[31] B. Picinbono and P. Chevalier, "Widely linear estimation with complex data," *IEEE Trans. Signal Process.*, vol. 43, no. 8, pp. 2030–2033, Aug. 1995.
[32] P. Schreier, L. Scharf, and C. Mullis, "Detection and estimation of improper complex random signals," *IEEE Trans. Inf. Theory*, vol. 51, no. 1, pp. 306–312, Jan. 2005.
[33] J.-P. Delmas, "Asymptotically minimum variance second-order estimation for noncircular signals with application to doa estimation," *IEEE Trans. Signal Process.*, vol. 52, no. 5, pp. 1235–1241, May 2004.
[34] E. G. Larsson and P. Stoica*, Space-Time Block Coding for Wireless Communications*. New York: Cambridge Univ. Press, 2003.
[35] U. Fincke and M. Pohst, "Improved methods for calculating vectors of short length in a lattice, including a complexity analysis," *Math. Comput.*, vol. 44, pp. 463–471, Apr. 1985.
[36] D. Gesbert, M. Shafi, D.-S. 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.
[37] O. Damen, A. Chkeif, and J. Belfiore, "Lattice code decoder for space–time codes," *IEEE Commun. Lett.*, vol. 4, pp. 161–163, May 2000.
[38] J. Jaldén, C. Martin, and B. Ottersten, "Semidefinite programming for detection in linear systems—Optimality conditions and space–time decoding," in *Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing (ICASSP)*, Apr. 2003, vol. 4, pp. 9–12.
[39] J. Mannerkoski and V. Koivunen, "Autocorrelation properties of channel encoded sequences—applicability to blind equalization," *IEEE Trans. Signal Process.*, vol. 48, no. 12, pp. 3501–3506, Dec. 2000.
[40] L. Tran, J. Seberry, Y. Wang, B. Wysocki, T. Wysocki, X. Tianbing, and Z. Ying, "Two complex orthogonal space–time codes for eight transmit antennas," *IEE Electron. Lett.*, vol. 40, no. 1, pp. 55–57, Jan. 2004.
[41] L. C. Tran, T. A. Wysocki, A. Mertins, and J. Seberry*, Complex Orthogonal Space-Time Processing in Wireless Communications*. Secaucus, NJ: Springer-Verlag, 2006.
[42] L. Tong and S. Perrau, "Multichannel blind identification: from subspace to maximum likelihood methods," *Proc. IEEE*, vol. 86, pp. 1951–1968, Oct. 1998.
[43] B. Porat*, Digital processing of random signals: theory and methods*. Upper Saddle River, NJ: Prentice-Hall, 1994.
[44] J. Vía, I. Santamaría, J. Pérez, and D. Ramírez, "Blind decoding of MISO-OSTBC systems based on principal component analysis," presented at the IEEE Int. Conf. Acoustic, Speech, Signal Processing (ICASSP), Toulouse, France, May 2006.
[45] J. Vía, I. Santamaría, and J. Pérez, "Blind identification of MIMO-OSTBC channels combining second and higher order statistics," presented at the Eur. Signal Processing Conf. (EUSIPCO 2006), Florence, Italy, Sep. 2006.
[46] K. I. Diamantaras and S. Y. Kung*, Principal Component Neural Networks, Theory and Applications*. New York: Wiley, 1996.
[47] E. Oja, "Principal components, minor components, and linear neural networks," *Neural Netw.*, vol. 5, no. 6, pp. 927–935, 1992.

Javier Vía (M'08) received the Telecommunication Engineer degree and the Ph.D. degree in electrical engineering from the University of Cantabria, Spain, in 2002 and 2007, respectively.

In 2002, he joined the Department of Communications Engineering, University of Cantabria, Spain, where he is currently Assistant Professor. In 2006, he spent a visiting period at the Smart Antennas Research Group (SARG), Stanford University Stanford, CA. His current research interests include blind channel estimation and equalization in wireless communication systems, multivariate statistical analysis, and kernel methods.

Ignacio Santamaría (M'96–SM'05) received the Telecommunication Engineer degree and the Ph.D. degree in electrical engineering from the Polytechnic University of Madrid, Spain, in 1991 and 1995, respectively.

In 1992, he joined the Departamento de Ingeniería de Comunicaciones, Universidad de Cantabria, Spain, where he is currently Full Professor. In 2000 and 2004, he spent visiting periods at the Computational NeuroEngineering Laboratory (CNEL), University of Florida. He has more than 100 publications in refereed journals and international conference papers. His current research interests include signal processing algorithms for wireless communication systems, MIMO systems, multivariate statistical techniques and machine learning theories. He has been involved in several national and international research projects on these topics.