EURASIP Journal on Applied Signal Processing 2002:12, 1427–1436

Multiuser Channel Estimation from Higher-Order Statistical Matrix Pencil

Jing Liang

Department of Electrical and Computer Engineering, University of California, Davis, CA 95616, USA
Email: jngliang@ece.ucdavis.edu

Zhi Ding

Department of Electrical and Computer Engineering, University of California, Davis, CA 95616, USA
Email: zding@ece.ucdavis.edu

Received 3 February 2002 and in revised form 11 August 2002

This paper presents a new statistical approach to the blind estimation of linear multiple-input multiple-output (MIMO) channels with finite impulse response. A matrix pencil is constructed from a set of fourth-order cumulant matrices of the channel output signals. The MIMO channel impulse responses can then be efficiently estimated from the generalized eigendecomposition of this cumulant matrix pencil. Random weighting is applied in the matrix pencil construction to improve the reliability of the algorithm. The proposed new method requires a relaxed channel identifiability condition and is robust in the sense that it does not require the exact knowledge of the MIMO channel order.

Keywords and phrases: blind system identification, higher-order statistics, multiple-input multiple-output linear systems, matrix pencil, identifiability condition.

1. INTRODUCTION

In recent years, blind estimation of multiple-input multiple-output (MIMO) linear channels has become a well-known research problem in multichannel communications and signal recovery. Satisfactory solutions of this problem can find diverse applications in areas such as multiuser detection, array signal processing, speech processing, and multichannel biomedical signal recovery.

The key objective of blind MIMO channel estimation is to determine the unknown matrix channel impulse response without direct training or knowledge of the channel input signals. The receiver must rely on the statistical information from the channel output signals. When the channel is a memoryless system, the problem is often known as blind source separation (or independent component analysis) with the goal of directly extracting source signals from the instantaneous mixtures without explicitly identifying the mixing matrix [1, 2, 3]. On the other hand, many multiuser systems must deal with dynamic channels that are characterized by a convolutive model. Once the dynamics of the MIMO system are estimated, techniques previously used in blind source separation of memoryless systems can be employed subsequently for individual signal separation.

Most existing approaches to blind channel estimation rely on the use of either second-order statistics (SOS) or higher-order statistics (HOS) of the channel output signals. As two equally important directions, second-order and higher-order methods have different characteristics suitable for different application scenarios. Compared with HOS methods, SOS methods may provide better performance for shorter data records [4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18]. They also require a subsequent source separation step after identifying the channel dynamics. On the other hand, the proper use of HOS information allows blind identification of a wider class of MIMO channels. HOS methods can directly resolve the inherent unitary ambiguity in SOS channel estimation. Considering channels without excessive diversity, in this work, we present an HOS approach to blind channel estimation.

There are many well-known HOS blind identification methods for single-input single-output (SISO) and MIMO systems. An incomplete list of works include [19, 20, 21, 22, 23, 24, 25, 26, 27, 28]. Specifically, Giannakis et al. [22] generalized their original “GM method” to MIMO systems. Swami and his colleagues [23] presented a unified framework to define cumulants of vector processes for arbitrary orders and developed parameter estimation algorithms for causal and noncausal multichannel AR, MA, and ARMA models. In [24], Tong proposed a new eigen-structure based idea under a very general channel identifiability condition. Using an indirect approach of inverse criteria, Tugnait [26, 27] proposed two nonlinear algorithms that iteratively recover one user signal at a time before estimating its corresponding channels.

In this paper, our goal is to develop a simple method that can estimate a large class of MIMO channels. A matrix pencil is a powerful tool and has been successfully applied in the fields of blind source separation and array processing [1, 29, 30, 31]. Motivated by the SOS matrix pencil algorithms for the blind separation of nonstationary or colored signals [30, 31], we develop an HOS approach to the blind estimation of MIMO channels driven by white stationary input signals. In a previous work [32], we developed a matrix pencil-based algorithm for blind identification of SISO systems. Here, the work in [32] is generalized to MIMO channel estimation. The channel identifiability condition of the new method is weaker than some existing higher-order methods, for example, [22, 26, 27], but is still stronger than the condition given in [24]. In addition, this matrix pencil method exhibits some robustness to errors in channel order estimation.

Our paper consists of the following parts: Section 2 describes the problem formulation and its basic assumptions. Section 3 outlines the principle of channel estimation from the cumulant matrix pencil, and Section 4 presents the algorithm in details. Section 5 provides simulation examples that illustrate the performance advantages of the proposed algorithm compared with some existing algorithms [26, 33].

2. PROBLEM FORMULATION

2.1. System model and basic assumptions

We consider the discrete-time model of a linear MIMO system with m input users and p outputs. The ith channel output signals $x_i(n)$ is given by

$$x_i(n) = \sum_{j=1}^{m}\sum_{k=0}^{q} h_{ij}(k)s_j(n-k) + w_i(n), \quad 1 \le i \le p, \tag{1}$$

where $h_{ij}(n)$ represents the channel impulse response between the jth channel input and the ith channel output and $w_i(n)$ denotes additive Gaussian noise at the ith channel output. We assume that the MIMO channel has finite impulse response (FIR) with maximum memory order of q.

For notational convenience, define the channel impulse response matrix as

$$H[n] \triangleq \begin{bmatrix} h_{11}(n) & \cdots & h_{1m}(n) \\ \vdots & \vdots & \vdots \\ h_{p1}(n) & \cdots & h_{pm}(n) \end{bmatrix} \tag{2}$$

and define the MIMO output signal vector, input signal vector, and noise vector as

$$\vec{x}[n] \triangleq \begin{bmatrix} x_1(n) \\ \vdots \\ x_p(n) \end{bmatrix}, \qquad \vec{s}[n] \triangleq \begin{bmatrix} s_1(n) \\ \vdots \\ s_m(n) \end{bmatrix}, \qquad \vec{w}[n] \triangleq \begin{bmatrix} w_1(n) \\ \vdots \\ w_p(n) \end{bmatrix}, \tag{3}$$

respectively. The input-output relationship for the linear MIMO system is given by

$$\vec{x}[n] = \sum_{k=0}^{q} H[k]\vec{s}[n-k] + \vec{w}[n]. \tag{4}$$

We can define the MIMO channel transfer function as

$$H(z) \triangleq \sum_{i=0}^{q} H[i]z^{-i}. \tag{5}$$

For convenience of algorithm derivation, the convolutional relationship (4) can also be written as

$$\mathbf{x}[n] = \mathbf{H}\mathbf{s}[n] + \mathbf{w}[n], \tag{6}$$

where the stacked signal vectors are defined as

$$\mathbf{x}[n] \triangleq \begin{bmatrix} \vec{x}[n] \\ \vdots \\ \vec{x}[n-L] \end{bmatrix}, \qquad \mathbf{s}[n] \triangleq \begin{bmatrix} \vec{s}[n] \\ \vdots \\ \vec{s}[n-L-q] \end{bmatrix}, \qquad \mathbf{w}[n] \triangleq \begin{bmatrix} \vec{w}[n] \\ \vdots \\ \vec{w}[n-L] \end{bmatrix} \tag{7}$$

and the channel convolution matrix $\mathbf{H}$ is an $(L+1)p \times (L+q+1)m$ block Toeplitz matrix

$$\mathbf{H} \triangleq \begin{bmatrix} H[0] & \cdots & H[q] & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & H[0] & \cdots & H[q] & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \mathbf{0} \\ \mathbf{0} & \cdots & \mathbf{0} & H[0] & \cdots & H[q] \end{bmatrix}. \tag{8}$$

We will present a method to estimate the channel impulse response $\{H[k]\}_{k=0}^{q}$ from the fourth-order cumulants of output signal $\mathbf{x}[n]$.

The discussion throughout this paper is based on several key assumptions. Here, we provide a list of these assumptions for later reference.

(A1) Independent source signals $\{s_i(n)\}$ are stationary, temporally i.i.d. processes with zero-means and nonzero fourth-order kurtoses $\{\gamma_i\}$.

(A2) Channel noises $\{w_i(n)\}$ are zero-mean stationary Gaussian processes. They are mutually independent, and are also independent of input signals.

(A3) The number of input signals is no more than the number of channel outputs, that is, $p \geq m$.
(A4) Matrix $H[0]$ only consists of nonzero columns.
(A5) There exists a nonzero z_0 (including ∞), such that $H(z_0)$ has full column rank.

The first three assumptions (A1), (A2), and (A3) are commonly used in higher-order methods. Assumption (A4) can be made without loss of generality since all-zero column corresponding to user $s_i(n)$ can be removed by renaming $\tilde{s}_i(n) = s_i(n-\tau)$ for some delay τ, as indicated in [24]. Note that assumption (A5) is the channel identifiability condition (see [33]) required in our channel estimation algorithm. This condition, though stronger than that proposed in [24], is less restrictive than the irreducibility condition required by second-order methods and the requirements of some other higher-order methods [22, 26, 27].

2.2. *Cumulant matrix construction*

As in [33], fourth-order cumulant matrix $\mathbf{C}_l[k]$, associated with the lth channel output signal $x_l(n-k)$, is defined as

$$\mathbf{C}_l[k] \triangleq \operatorname{cum}\left(\mathbf{x}[n], \mathbf{x}[n]^H, x_l(n-k), x_l^*(n-k)\right), \tag{9}$$

where $(\cdot)^*$ and $(\cdot)^H$ represent complex conjugate and conjugate transpose, respectively. Matrix $\mathbf{C}_l[k]$ has the same dimension as the covariance matrix of $\mathbf{x}[n]$. Based on assumptions (A1), (A2) and the properties of cumulants [34], the noise contribution to this matrix is zero and we have

$$\mathbf{C}_l[k] = \mathbf{H}\Lambda_l[k]\mathbf{H}^H, \tag{10}$$

in which

$$\Lambda_l[k] = \operatorname{diag}\Big(\underbrace{\mathbf{0},\ldots,\mathbf{0}}_{k \text{ blocks}}, D_l[0],\ldots,D_l[q], \underbrace{\mathbf{0},\ldots,\mathbf{0}}_{(L-k)\text{ blocks}}\Big),$$

$$D_l[j] = \operatorname{diag}\left(\gamma_1\left|h_{l1}(j)\right|^2,\ldots,\gamma_m\left|h_{lm}(j)\right|^2\right), \quad j = 0,\ldots,q. \tag{11}$$

We will describe a matrix pencil algorithm for MIMO channel estimation using a collection of these cumulant matrices.

3. CHANNEL ESTIMATION FROM MATRIX PENCIL

This section outlines the basic principle of the proposed *multiuser cumulant matrix pencil* (MCMP) channel estimation algorithm. It first describes the matrix pencil formation using a set of cumulant matrices $\{\mathbf{C}_l[k]\}$. Then, it gives key equations for channel identification that involves finding nontrivial generalized eigenvectors of the cumulant matrix pencil.

3.1. *Cumulant matrix pencil*

Our approach is motivated in part by [30, 31], where a matrix pencil, formed from output autocorrelation matrices at different time delays, was used to extract source signals that are nonstationary with colored power spectra. Here, we use cumulant matrices $\{\mathbf{C}_l[k]\}$ associated with the same time delay k, but different spatial indices l. Following our study on single-user system [32], the matrix pencil is constructed from linear combinations of these cumulant matrices.

Equations (10) and (11) describe the general form of cumulant matrix $\mathbf{C}_l[k]$ for given parameters $\{k, l, L\}$. Here, we consider the special case of

$$k = q, \qquad L = 2q. \tag{12}$$

Under this condition, the cumulant matrix can be specifically represented as (see [33])

$$\mathbf{C}_l[q] = \mathbf{H}_s\Lambda_l[q]\mathbf{H}_s^H, \tag{13}$$

with

$$\mathbf{H}_s = \begin{bmatrix} H[q] & \mathbf{0} & \cdots & \mathbf{0} \\ \vdots & H[q] & \ddots & \vdots \\ H[0] & \ddots & \ddots & \mathbf{0} \\ \mathbf{0} & H[0] & \ddots & H[q] \\ \vdots & \ddots & \ddots & \vdots \\ \mathbf{0} & \cdots & \mathbf{0} & H[0] \end{bmatrix}, \tag{14}$$

$$\Lambda_l[q] = \operatorname{diag}\left(D_l[0],\ldots,D_l[q]\right).$$

Notice that $\mathbf{H}_s$ is a convolution matrix with block Toeplitz structure, and can be a tall matrix with dimension $(L+1)p \times (L+1-q)m$, given that $p \geq m$.

Then, we define a cumulant matrix pencil $\{\mathbf{S}_1, \mathbf{S}_2\}$ in which

$$\mathbf{S}_i \triangleq \sum_{l=1}^{p} w_{il}\mathbf{C}_l[q] = \mathbf{H}_s\mathbf{\Gamma}_i\mathbf{H}_s^H, \quad i = 1, 2, \tag{15}$$

$$\mathbf{\Gamma}_i = \operatorname{diag}\left(\Sigma_i[0],\ldots,\Sigma_i[q]\right). \tag{16}$$

Two sets of weighting factors $\{w_{il}\}$ are i.i.d. random numbers with uniform distribution within interval $(0, 1)$. From (13) and (15), we have

$$\Sigma_i[j] = \operatorname{diag}\left(\gamma_1\sum_{l=1}^{p} w_{il}\left|h_{l1}(j)\right|^2,\ldots,\gamma_m\sum_{l=1}^{p} w_{il}\left|h_{lm}(j)\right|^2\right), \quad j = 0,\ldots,q. \tag{17}$$

It can be verified that $\Sigma_1[j]$ and $\Sigma_2[j]$ are nonsingular if and only if $H[j]$ has no all-zero column. On the other hand, if matrix $H[j]$ has an all-zero column corresponding to input signal $s_u(n)$, then both $\Sigma_1[j]$ and $\Sigma_2[j]$ have a zero element at the uth diagonal entry. This property will be used to single out trivial solutions of our channel estimation algorithm.

3.2. *Channel identification via matrix pencil*

We consider the generalized eigenvalue problem

$$\mathbf{S}_1\mathbf{v}_i = \lambda_i\mathbf{S}_2\mathbf{v}_i, \tag{18}$$

or equivalently

$$\mathbf{H}_s(\boldsymbol{\Gamma}_1 - \lambda_i\boldsymbol{\Gamma}_2)\mathbf{H}_s^H\mathbf{v}_i = \mathbf{0}. \tag{19}$$

It has a total of $(L+1)p$ generalized eigenvectors $\{\mathbf{v}_i\}$ with corresponding generalized eigenvalues $\{\lambda_i\}$. For MIMO systems, our matrix pencil algorithm requires that $\mathbf{H}_s$ have full column rank. As shown in [33], this requirement can be satisfied under assumption (A5). Given the full column rank matrix $\mathbf{H}_s$, every generalized eigenvector $\mathbf{v}_i$ should satisfy

$$\begin{bmatrix} \Sigma_1[0] - \lambda_i\Sigma_2[0] & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \Sigma_1[q] - \lambda_i\Sigma_2[q] \end{bmatrix} \mathbf{H}_s^H\mathbf{v}_i = \mathbf{0}. \tag{20}$$

We will show that the MIMO channel impulse response can be identified by finding the proper solutions $\{\lambda_i, \mathbf{v}_i\}$ of (20). For clarity, three different classes of generalized eigenvectors are discussed separately.

3.2.1 Trivial eigenvectors

A trivial solution of the generalized eigenvalue problem (18) is given by

$$\mathbf{S}_1\mathbf{v}_i = \mathbf{S}_2\mathbf{v}_i = \mathbf{0}. \tag{21}$$

Thus, λ_i can be of any complex value. Such a solution can occur in two cases.

First, the full row rank matrix $\mathbf{H}_s^H$ has a nonempty nullspace with a dimension of $(L+1)p - (L+1-q)m$. Any vector in this nullspace will give $\mathbf{H}_s^H\mathbf{v}_i = \mathbf{0}$. In addition, other trivial solutions may exist depending on the MIMO channel impulse response. Since an all-zero column in $H[j]$ ($j = 0, \ldots, q$) will lead to the singularity of $\Sigma_1[j]$ and $\Sigma_2[j]$, there is a one-to-one mapping between an all-zero column in $H[j]$ and a trivial generalized eigenvector $\mathbf{v}_i$ such that $\mathbf{H}_s^H\mathbf{v}_i \neq \mathbf{0}$, but $\boldsymbol{\Gamma}_1\mathbf{H}_s^H\mathbf{v}_i = \mathbf{0}$ and $\boldsymbol{\Gamma}_2\mathbf{H}_s^H\mathbf{v}_i = \mathbf{0}$.

All these trivial eigenvectors provide no information about the channel and thus will be discarded. They can be singled out based on the following two properties:

(1) $\mathbf{v}_i^H\mathbf{S}_1\mathbf{v}_i = 0$ and $\|\mathbf{S}_1\mathbf{v}_i\| = 0$.
(2) $\mathbf{v}_i^H\mathbf{C}_l[q]\mathbf{v}_i = 0$ and $\|\mathbf{C}_l[q]\mathbf{v}_i\| = 0$, for $l = 1, \ldots, p$.

Here, $\|\cdot\|$ represents the Euclidean norm of a vector.

3.2.2 Nontrivial eigenvectors of distinct eigenvalues

Nontrivial eigenvectors of distinct eigenvalues are the most useful ones because they possess the desired properties for channel estimation. When an eigenvalue is said to be distinct, it does not have any multiplicity and has only one corresponding eigenvector. It can be seen from (20) that, if λ_i is a distinct eigenvalue, then there is only one j ($j \in \{0, \ldots, q\}$), such that λ_i is also a unique eigenvalue of matrix pencil $\{\Sigma_1[j], \Sigma_2[j]\}$. In other words, there exists a vector $\vec{a}$ such that

$$\Sigma_1[j]\vec{a} = \lambda_i\Sigma_2[j]\vec{a}. \tag{22}$$

Because both $\Sigma_1[j]$ and $\Sigma_2[j]$ are diagonal matrices defined by (17), λ_i is given by

$$\lambda_i = \frac{\sum_{l=1}^{p} w_{1l}\left|h_{lu}(j)\right|^2}{\sum_{l=1}^{p} w_{2l}\left|h_{lu}(j)\right|^2}, \quad u \in \{1, 2, \ldots, m\}. \tag{23}$$

This implies that the uth element, corresponding to input signal $s_u(n)$, is the only zero diagonal element of $(\Sigma_1[j] - \lambda_i\Sigma_2[j])$. Meanwhile, the diagonal elements of

$$\Sigma_1[\ell] - \lambda_i\Sigma_2[\ell], \quad \ell \neq j \tag{24}$$

are nonzero except for trivial cases discussed above. Consequently, the generalized eigenvector $\mathbf{v}_i$ corresponding to λ_i must satisfy

$$\mathbf{H}_s^H\mathbf{v}_i = \alpha_i\mathbf{e}_u^j, \tag{25}$$

where α_i is an unknown scaling factor and $\mathbf{e}_u^j$ represents the $(jm+u)$th canonical vector in the identity matrix (i.e., $\mathbf{e}_u^j$ has 1 in the $(jm+u)$th entry and zeros elsewhere).

Therefore, by multiplication, we have

$$\mathbf{S}_1\mathbf{v}_i = \alpha_i\mathbf{H}_s\mathbf{e}_u^j = \alpha_i\mathbf{h}_u^j, \tag{26}$$

where $\mathbf{h}_u^j$ represents the $(jm+u)$th column of $\mathbf{H}_s$ and α_i absorbs all scaling factors. Similarly,

$$\mathbf{C}_l[q]\mathbf{v}_i = \alpha_i\mathbf{h}_u^j, \quad l = 1, \ldots, p. \tag{27}$$

Apparently, vector $\mathbf{S}_1\mathbf{v}_i$ (or $\mathbf{C}_l[q]\mathbf{v}_i$) provides an estimate of the channel impulse response for input $s_u(n)$.

3.2.3 Nontrivial eigenvectors of repeated eigenvalues

It should be noted that the proposed MCMP algorithm will fail if the cumulant matrix pencil $\{\mathbf{S}_1, \mathbf{S}_2\}$ has nontrivial repeated eigenvalues. Let λ_i represent a repeated eigenvalue with the corresponding eigenspace denoted by $\mathbf{V}_i$. We will have

$$\mathbf{H}_s^H\mathbf{V}_i = \Lambda P U_i \tag{28}$$

in which Λ is a diagonal matrix, P is a permutation matrix, and U_i is a unitary matrix. In this case, channel estimates can not be directly extracted from $\mathbf{S}_1\mathbf{V}_i$ because of the unknown mixture matrix U_i.

3.3. Discussions

From the above analysis, MCMP channel estimation algorithm requires that all nontrivial generalized eigenvalues be unique for the cumulant matrix pencil under consideration. When we adopt random weighting, it is *with probability one* that all nontrivial generalized eigenvalues are distinct, as in (23).

Equation (26) indicates that each nontrivial eigenvector can only provide an estimate of one user channel and we do not know which user it belongs to. Thus, the remaining problem is to ensure that each active user channel can be estimated. Since both $\Sigma_1[0]$ and $\Sigma_2[0]$ are nonsingular under assumption (A4), matrix pencil $\{\Sigma_1[0], \Sigma_2[0]\}$ should have m distinct nontrivial eigenvalues. Consequently, the corresponding nontrivial eigenvectors of $\{\mathbf{S}_1, \mathbf{S}_2\}$ will pick the first m columns of $\mathbf{H}_s$, which provide channel estimates of all user channels.

Another important issue is to note that the MCMP algorithm does not rely on the exact knowledge of channel order q. Instead, it needs an approximate channel order $\hat{q}$ such that $\hat{q} \geq q$. Suppose that we do not know the true channel order q and use $\hat{q}$ in the delay and length parameter setting (12) to compute cumulant matrix $\mathbf{C}_l[\hat{q}]$. Equations (13), (14), (15), (16), and (17) remain unchanged except that parameter q is replaced by $\hat{q}$, and the analysis in Section 3.2 is also valid. It is now equivalent to estimate a new channel impulse response $\{H[0], \ldots, H[\hat{q}]\}$ in which $\{H[q+1], \ldots, H[\hat{q}]\}$ are zero matrices. The major difference lies in that

$$\boldsymbol{\Gamma}_i = \operatorname{diag}\Big(\Sigma_i[0], \ldots, \Sigma_i[q], \underbrace{\mathbf{0}, \ldots, \mathbf{0}}_{(\hat{q}-q)\text{ blocks}}\Big), \quad i = 1, 2. \tag{29}$$

As a result, the number of trivial eigenvectors is increased, but the nontrivial solutions are *not* affected.

4. ALGORITHM IMPLEMENTATION AND CHANNEL TAGGING

This section focuses on a detailed procedure to implement the MCMP channel estimation algorithm and describes how to separate the estimates of different user channels.

4.1. Generalized eigendecomposition

First, cumulant matrices $\{\mathbf{C}_l[q]\}$ are estimated from the channel output signals to form a matrix pencil $\{\mathbf{S}_1, \mathbf{S}_2\}$. As shown above, symmetric matrices $\mathbf{S}_1$ and $\mathbf{S}_2$ are positive semidefinite and share a common nullspace. This means that $\{\mathbf{S}_1, \mathbf{S}_2\}$ is a symmetric, singular matrix pencil.

The QZ algorithm, such as the one used in MATLAB function eig(A, B), can be a standard method for solving small- to moderate-dimensional generalized eigenvalue problems. However, as pointed in [35], the QZ algorithm is generally unreliable when handling a singular matrix pencil. In [36], Cao specifically addressed the generalized eigenvalue problem of a symmetric matrix pencil $\{A, B\}$ (regular or singular), in which one of the matrices is positive (or negative) semidefinite. The Fix-Heiberger reduction was generalized to deflate the infinite and the singular structure from $\{A, B\}$. By applying this deflation method, it was shown that the Kronecker canonical form of $\{A, B\}$ is very special and the generalized eigendecomposition of $\{A, B\}$ can be considerably simplified.

4.2. Eigenvector selection

The matrix pencil $\{\mathbf{S}_1, \mathbf{S}_2\}$ has $(L+1)p$ generalized eigenvectors overall. When none of $H[n]$ $(n = 0, \ldots, q)$ has an all-zero column, the number of nontrivial eigenvectors reaches the maximum of $(L+1-q)m$. Our goal here is to determine these nontrivial eigenvectors. The selecting procedures are proposed as follows.

(i) Normalize all generalized eigenvectors. For the ith normalized eigenvector $\mathbf{v}_i$, find $\mathbf{C}_l[q]$ corresponding to

$$l_i = \arg\max_{l=1,\ldots,p} \mathbf{v}_i^H \mathbf{C}_l[q] \mathbf{v}_i. \tag{30}$$

(ii) Compute

$$f_{l_i} = \mathbf{v}_i^H \mathbf{C}_{l_i}[q] \mathbf{v}_i, \tag{31}$$

for $i = 1, \ldots, (L+1)p$.

(iii) Let $\mathcal{K}$ denote the set of index i for nontrivial eigenvectors. Since f_{l_i} for trivial eigenvectors are very small, set $\mathcal{K}$ consists of indices corresponding to $(L+1-q)m$ largest f_{l_i}. Although it is possible that a small number of trivial eigenvectors may be wrongly included, they will be discarded in the channel tagging step (Section 4.3).

(iv) Channel estimates can be obtained as shown in (26) and (27). For better estimation performance, we compute

$$\vec{y}_i = \mathbf{C}_{l_i}[q]\mathbf{v}_i = \beta_i \mathbf{h}_u^j, \quad i \in \mathcal{K}, \tag{32}$$

where β_i is a scaling factor and $\mathbf{h}_u^j$ represents the $(jm+u)$th column of $\mathbf{H}_s$.

Regarding the last step, the block Toeplitz structure of $\mathbf{H}_s$ determines that channel parameters are located in the middle of vector $\vec{y}_i$ with zero entries at one end or both ends. In other words, the estimation result $\vec{y}_i$ includes unknown delay ambiguity and scaling ambiguity β_i.

4.3. Channel tagging

Since different $\vec{y}_i$ carry channel information for different users, we need to classify all $\{\vec{y}_i\}$ into m groups, each corresponding to one user. This is known as channel tagging.

To construct a proper classification criterion, we first define a distance measure between two vectors $\vec{y}_{i_1}$ and $\vec{y}_{i_2}$. If they correspond to the same user, without loss of generality, they can be written as

$$\vec{y}_{i_1} = \beta_{i_1}\mathbf{h}_u^{j_1}, \quad \vec{y}_{i_2} = \beta_{i_2}\mathbf{h}_u^{j_2}, \quad i_1, i_2 \in \mathcal{K}, \ j_1 < j_2. \tag{33}$$

It is clear that $\mathbf{h}_u^{j_1}$ and $\mathbf{h}_u^{j_2}$ belong to different column blocks of matrix $\mathbf{H}_s$. Moreover, when $\vec{y}_{i_2}$ is circularly up shifted by $(j_2 - j_1)p$ elements, the resulting vector $\vec{y}_{i_2}^{(j_2-j_1)}$ is exactly the same as $\vec{y}_{i_1}$ except for a scalar difference. In this case, the normalized cross-correlation between $\vec{y}_{i_1}$ and $\vec{y}_{i_2}^{(j_2-j_1)}$ reaches the maximum value. Based on this property, we define a cross-correlation matrix $\mathbf{R}_{\vec{\mathbf{y}}}$ to measure the similarity between every pair of $\vec{y}_{i_1}$ and $\vec{y}_{i_2}$. We normalize each vector $\vec{y}_i$ into $\vec{\mathbf{y}}_i$. The (i_1, i_2)th entry of $\mathbf{R}_{\vec{\mathbf{y}}}$ is given as

$$[\mathbf{R}_{\vec{\mathbf{y}}}]_{i_1,i_2} = \max_{n=0,\ldots,L} \left| \vec{\mathbf{y}}_{i_1}^H \cdot \vec{\mathbf{y}}_{i_2}^{(n)} \right|, \quad i_1, i_2 \in \mathcal{K}, \tag{34}$$

where $\vec{\mathbf{y}}_{i_2}$ is circularly up shifted by np elements to form $\vec{\mathbf{y}}_{i_2}^{(n)}$. Matrix $\mathbf{R}_{\vec{\mathbf{y}}}$ provides an effective distance measure for channel tagging.

We then apply the simple *hierarchical dimensionality reduction* (HDR) approach to classify $\{\vec{\mathbf{y}}_i\}$ into m groups. HDR method was also adopted in [31] for signal grouping. Initially, we have $(L+1-q)m$ groups and each of them has only one vector $\vec{\mathbf{y}}_i$. In every iteration, two most correlated groups are merged together. For any two groups, the *group correlation* is defined as the median of cross-correlation values $[\mathbf{R}_{\vec{\mathbf{y}}}]_{i_1,i_2}$ of any two vectors taken from each of these two groups, respectively. The merging procedure is continued until only m groups are left.

After classification, every user group may have several $\vec{\mathbf{y}}_i$ providing channel estimates for the same user. The final estimation result is selected as the vector with maximum norm. By doing this, the results obtained from wrongly selected trivial eigenvectors will be discarded.

4.4. Summary

To summarize, the MCMP channel estimation algorithm contains the following major steps.

(1) Estimate cumulant matrices $\{\mathbf{C}_l[q]\}$ for $l = 1, \ldots, p$.
(2) Construct matrix pencil $\{\mathbf{S}_1, \mathbf{S}_2\}$, and solve all generalized eigenvectors $\{\mathbf{v}_i\}$ for $i = 1, \ldots, (L+1)p$.
(3) Determine set $\mathcal{K}$ for nontrivial generalized eigenvectors, and compute $\{\vec{y}_i \mid i \in \mathcal{K}\}$.
(4) Compute cross-correlation matrix $\mathbf{R}_{\vec{\mathbf{y}}}$ and apply HDR method for channel tagging.
(5) Select final estimation results from each user group.

5. SIMULATIONS

In this section, we present several simulation examples to demonstrate the feasibility and performance of the proposed MCMP channel estimation algorithm. For comparison, the simulation results of two existing MIMO channel estimation algorithms, namely the *MIMO Cumulant Subspace* (MIMOCS-MP) algorithm [33] and the iterative *Constant Modulus Algorithm* (IT-CMA) based approach [26], are also included.

In our simulation setup, channel inputs are mutually independent, i.i.d. QPSK signals. Additive channel noise is complex white Gaussian with zero mean, and its variance is determined by the averaged signal-to-noise ratio (SNR) over all channel outputs. We consider 2-input/2-output FIR MIMO channels, in which each subchannel is a random realization of COST207 typical urban (TU) propagation model [37] for GSM system. In addition to the examples where exact channel order q is used, some results illustrating the performance under channel order overestimation are also provided.

For MCMP algorithm, we set $\{k = q, L = 2q\}$ as in (12) and solve $(2q+1)p$ generalized eigenvectors of matrix pencil $\{\mathbf{S}_1, \mathbf{S}_2\}$. MIMOCS-MP algorithm chooses the delay and length parameters as $\{k_1 = q, k_2 = 2q-1, L = 3q-1\}$. It needs to perform singular value decompositions of a $(3q^2p \times 3qp)$ rectangular matrix and a $(q+1)p \times (q+1)p$ square matrix. Obviously, the implementation of MCMP algorithm is less costly than the MIMOCS-MP algorithm. On the other hand, IT-CMA algorithm involves adaptation of MIMO-CMA equalizer parameters and linear filtering operation to recover the input signals. The low-computational complexity of IT-CMA algorithm is proportional to the data length.

As a performance measure, we use the normalized mean square error (NMSE) criterion. For a given subchannel $h_{ij}(n)$, define

$$\mathrm{NMSE}_{ij} \triangleq E\left\{\frac{\sum_{n=0}^{q} \left| h_{ij}(n) - \rho_j \hat{h}_{ij}(n) \right|^2}{\sum_{n=0}^{q} \left| h_{ij}(n) \right|^2}\right\} \tag{35}$$

in which $\hat{h}_{ij}(n)$ is the estimated channel impulse response and ρ_j is used to compensate the scalar ambiguity associated with the estimation results for the jth user. The overall NMSE (ONMSE) is obtained by averaging over all subchannels

$$\mathrm{ONMSE} \triangleq \frac{1}{pm} \sum_{i=1}^{p} \sum_{j=1}^{m} \mathrm{NMSE}_{ij}. \tag{36}$$

In our simulations, all results are based on 100 Monte Carlo runs.

Example 1. We first consider a channel model CH1 with its impulse response given by

$$\begin{aligned} H[0] &= \begin{bmatrix} 0.5822 - 0.4495i & -0.4889 - 0.1731i \\ 0.3071 - 0.0295i & 0.0056 + 1.6768i \end{bmatrix}, \\ H[1] &= \begin{bmatrix} 0.3368 + 0.1992i & 0.0812 - 0.1652i \\ -0.0878 + 0.1669i & -0.0215 + 0.0973i \end{bmatrix}. \end{aligned} \tag{37}$$

Assume that channel order q is known. We select the length of MIMO-CMA equalizer to be $L_e = 4$ and the adaptation stepsize $\mu = 0.0002$.

Figure 1 illustrates the performance of these three algorithms at different SNR levels. Comparison is carried out for data lengths of 4000 and 8000 samples, respectively. In both cases, MCMP algorithm outperforms MIMOCS-MP and IT-CMA at various SNR levels. Figure 2 shows the comparative results for various data lengths when SNR is fixed at 10 dB and 20 dB, respectively. These results confirm the observation made from Figure 1. It is seen that MCMP algorithm demonstrates its performance advantage over IT-CMA and MIMOCS-MP algorithms, especially with shorter data lengths and at lower SNR levels.

Example 2. Next, we study the performance of the three algorithms when channel order is not exactly known. Recall that both MCMP algorithm and IT-CMA algorithm only need an approximate channel order $\hat{q}$ such that $\hat{q} \geq q$, while the subspace decomposition based MIMOCS-MP algorithm requires the exact knowledge of the channel order to guarantee the uniqueness of estimation results.

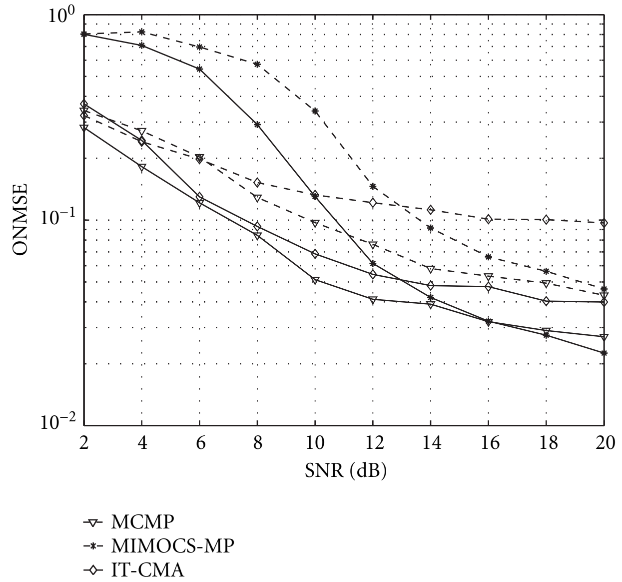


Figure 1: CH1: Performance comparison of three algorithms at different SNR levels, dashed lines for 4000 samples, and solid lines for 8000 samples.

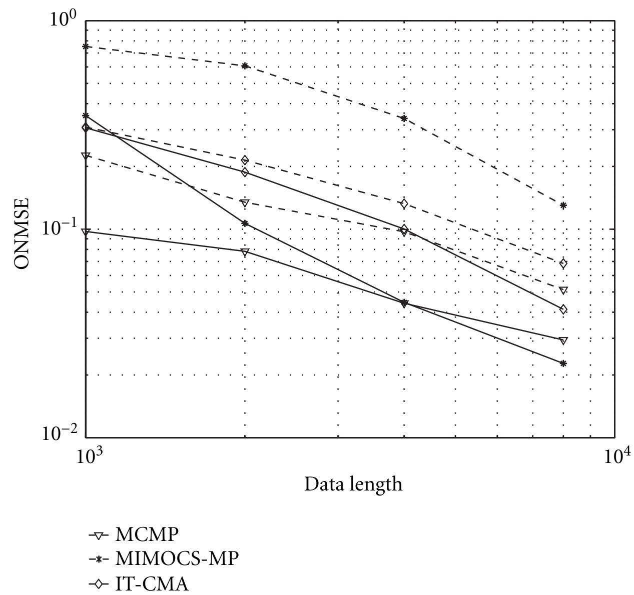


Figure 2: CH1: Performance comparison of three algorithms with different data lengths, dashed lines for SNR = 10 dB, and solid lines for SNR = 20 dB.

We test these algorithms on channel CH1 using 8000 data samples. In Figure 3, the dashed lines represent estimation results when true channel order q is known, and the solid lines represent estimation results when the channel order is overestimated as $\hat{q} = q + 1$. Clearly, both MCMP and IT-CMA algorithms introduce mild performance degradation, while MIMOCS-MP algorithm shows much severe performance loss. This example illustrates that MCMP algorithm is less sensitive to channel order overestimation.

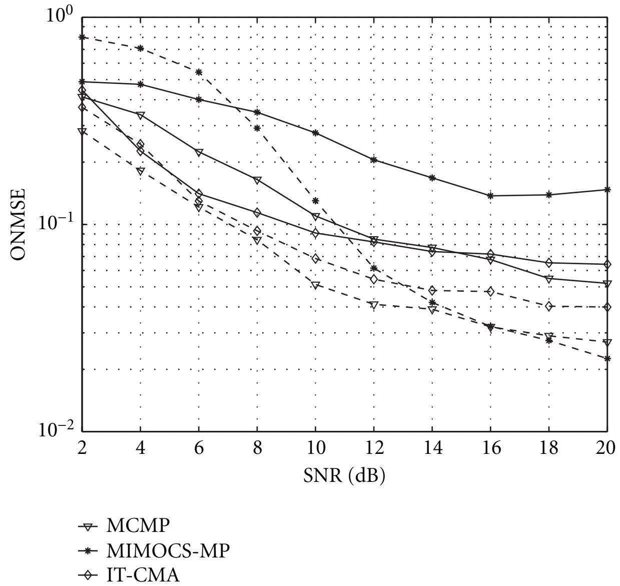


Figure 3: CH1: Performance comparison of three algorithms under channel order overestimation, data length is 8000 samples, dashed lines for $\hat{q} = q$, and solid lines for $\hat{q} = q + 1$.

Example 3. Here, we consider another channel model CH2, in which subchannels $h_{11}(z)$ and $h_{21}(z)$ share a common zero $z = -0.9252 - 0.3794i$ on the unit circle. The MIMO channel impulse response is given by

$$H[0] = \begin{bmatrix} 0.021440 + 0.548955i & -0.260932 + 0.108733i \\ 0.168106 - 0.273989i & -0.402841 - 0.640907i \end{bmatrix},$$
$$H[1] = \begin{bmatrix} -0.310738 + 0.829142i & -0.050598 - 0.224913i \\ 0.114620 - 0.268211i & 0.123198 - 0.141783i \end{bmatrix},$$
$$H[2] = \begin{bmatrix} -0.231945 + 0.243287i & -0.096203 + 0.098883i \\ -0.104259 - 0.127584i & 0.121224 - 0.249428i \end{bmatrix}. \tag{38}$$

Note that IT-CMA algorithm requires $H(z)$ to be of full column rank for $|z| = 1$ when a doubly infinite length equalizer is used. Clearly, channel CH2 violates the identifiability condition of IT-CMA algorithm, but it is still identifiable by the proposed MCMP algorithm.

Figure 4 shows the comparative results of the MCMP algorithm and the IT-CMA algorithm at different SNR levels. Assume that channel order q is known. The length of MIMO-CMA equalizer is selected as $L_e = 10$. We tested different adaptation stepsizes for fast and good convergence of IT-CMA algorithm, and chose $\mu = 0.0032$ for data length of 4000 samples and $\mu = 0.0016$ for data length of 8000 samples. It is observed that the performance of IT-CMA tends to stagnate, while MCMP can evidently improve its performance as SNR increases or data record length becomes longer. For the class of noninvertible channels as CH2, global convergence of MIMO-CMA equalizer is not guaranteed and the equalizer parameters may be trapped

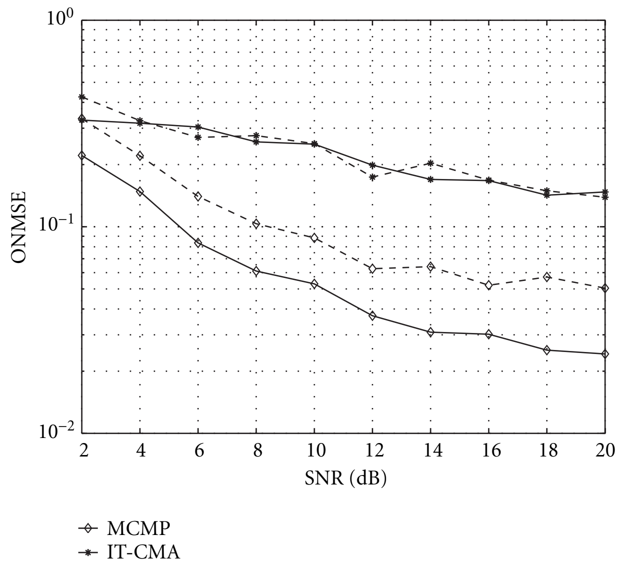


Figure 4: CH2: Performance comparison of MCMP with IT-CMA at different SNR levels, dashed lines for 4000 samples, and solid lines for 8000 samples.

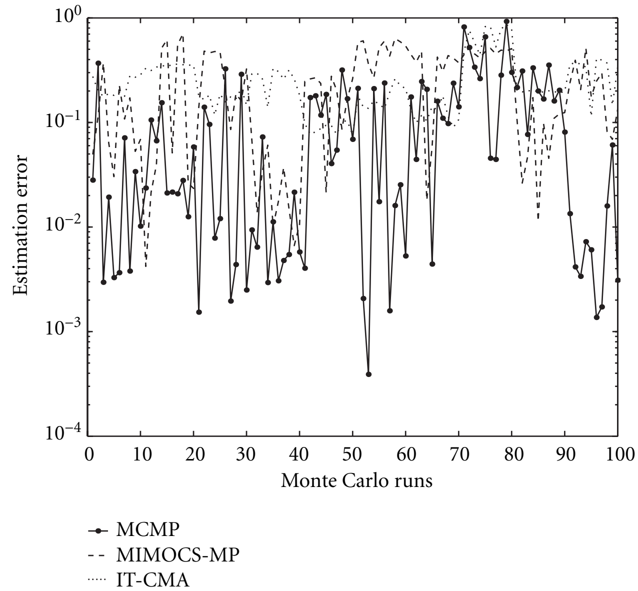


Figure 6: Random channel test: Estimation error versus Monte Carlo runs, data length is 8000 samples, SNR = 20 dB.

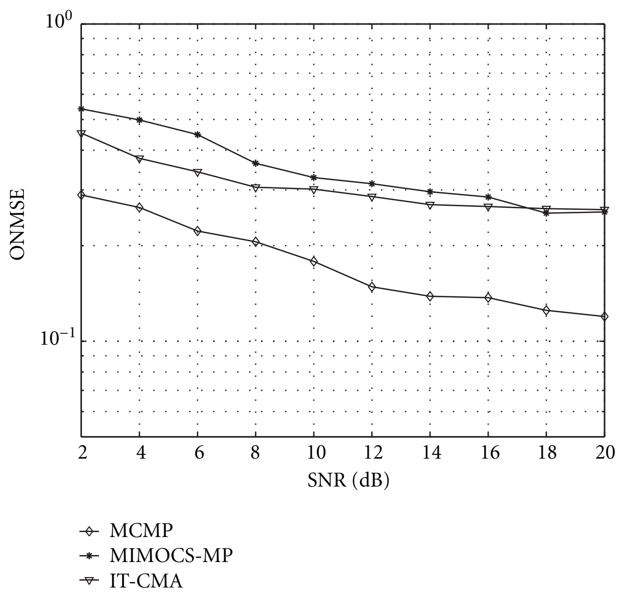


Figure 5: Random channel test: Averaged ONMSE performance, data length is 8000 samples.

at local minima. When the recovered input signal from equalizer is unreliable, it in turn affects the accuracy of channel estimation in IT-CMA algorithm. This example illustrates the advantage of the MCMP algorithm in the sense that it allows a less restrictive channel identifiability condition.

Example 4. All three examples above are based on a fixed channel test in which the same channel model is used in each Monte Carlo run. To investigate the robustness of the three algorithms to various channel conditions, we employ the following random channel test. We first randomly generate ten FIR MIMO channel models in the same way used to obtain CH1 and CH2. Then, ten Monte Carlo runs are performed for each random channel model. The ONMSE is obtained by averaging over 100 runs corresponding to these ten random channel models. Here, all subchannels have channel order $q = 1$ which is assumed to be known. The parameters of these three algorithms are chosen as in Example 1.

As illustrated in Figure 5, the proposed MCMP algorithm demonstrates better averaged performance than the other two algorithms in this random channel test, and the averaged performance of IT-CMA and MIMOCS-MP are very close. Next, in Figure 6, the results are shown versus Monte Carlo runs when data length is 8000 samples and SNR is 20 dB. It can be seen that, in most cases, MCMP algorithm generates more reliable estimates than IT-CMA and MIMOCS-MP. On the other hand, there do exist some “bad” TU channels, for example, Monte Carlo runs 70 to 90, for which all three algorithms perform poorly.

6. CONCLUSIONS

A new matrix pencil algorithm is developed for the blind estimation of FIR MIMO systems from fourth-order cumulants of the multiple channel outputs. It is shown that an unknown MIMO channel impulse responses can be identified by finding nontrivial generalized eigenvectors of a cumulant matrix pencil. The proposed new method requires a weaker channel identifiability condition than some existing methods and does not rely on the exact knowledge of the channel order. Numerical simulation examples are given to demonstrate its consistent performance.

ACKNOWLEDGMENTS

The authors would like to thank Professor Z. Bai of Computer Science Department at UC Davis for his help and valuable suggestions on solving the generalized eigenvalue problem of singular matrix pencils. This work is supported in part by the National Science Foundation (NSF) under grant CCR-0196364 and grant ECS-0121469.

REFERENCES

[1] L. Tong, R. W. Liu, V. C. Soon, and Y. F. Huang, "Indeterminacy and identifiability of blind identification," *IEEE Trans. Circuits and Systems*, vol. 38, no. 5, pp. 499–509, 1991.

[2] J. F. Cardoso, "Blind signal separation: statistical principles," *Proceedings of the IEEE*, vol. 86, pp. 2009–2025, 1998.

[3] D.-T. Pham and J. F. Cardoso, "Blind separation of instantaneous mixtures of nonstationary sources," *IEEE Trans. Signal Processing*, vol. 49, no. 9, pp. 1837–1848, 2001.

[4] L. Tong, G. Xu, and T. Kailath, "Blind identification and equalization based on second-order statistics: a time domain approach," *IEEE Transactions on Information Theory*, vol. 40, no. 2, pp. 340–349, 1994.

[5] E. Moulines, P. Duhamel, J. F. Cardoso, and S. Mayrargue, "Subspace methods for the blind identification of multichannel FIR filters," *IEEE Trans. Signal Processing*, vol. 43, no. 2, pp. 516–526, 1995.

[6] L. Tong and S. Perreau, "Multichannel blind identification: from subspace to maximum likelihood methods," *Proceedings of the IEEE*, vol. 86, no. 10, pp. 1951–1968, 1998.

[7] K. Abed-Meraim, P. Loubaton, and E. Moulines, "A subspace algorithm for certain blind identification problems," *IEEE Transactions on Information Theory*, vol. 43, no. 2, pp. 499–511, 1997.

[8] P. Loubaton and E. Moulines, "On blind multiuser forward link channel estimation by the subspace method: identifiability results," *IEEE Trans. Signal Processing*, vol. 48, no. 8, pp. 2366–2377, 2000.

[9] J. K. Tugnait and B. Huang, "Multistep linear predictors-based blind identification and equalization of multiple-input multiple-output channels," *IEEE Trans. Signal Processing*, vol. 48, no. 1, pp. 26–38, 2000.

[10] J. H. Gunther and A. L. Swindlehurst, "On the use of kernel structure for blind equalization," *IEEE Trans. Signal Processing*, vol. 48, pp. 799–809, March 2000.

[11] T. P. Krauss and R. D. Zoltowski, "Bilinear approach to multiuser second-order statistics-based blind channel estimation," *IEEE Trans. Signal Processing*, vol. 48, pp. 2473–2486, September 2000.

[12] X. G. Xia, W. Su, and H. Liu, "Filterbank precoders for blind equalization: polynomial ambiguity resistant precoders (PARP)," *IEEE Trans. on Circuits and Systems I: Fundamental Theory and Applications*, vol. 48, no. 2, pp. 193–209, 2001.

[13] A. Touzni, I. Fijalkow, M. G. Larimore, and J. R. Treichler, "A globally convergent approach for blind MIMO adaptive deconvolution," *IEEE Trans. Signal Processing*, vol. 49, no. 6, pp. 1166–1178, 2001.

[14] H. Ali, J. H. Manton, and Y. Hua, "A SOS subspace method for blind channel identification and equalization in bandwidth efficient OFDM systems based on receive antenna diversity," in *11th IEEE Workshop on Statistical Signal Processing*, pp. 401–404, Singapore, August 2001.

[15] Z. Xu and M. K. Tsatsanis, "Blind adaptive algorithms for minimum variance CDMA receivers," *IEEE Trans. Communications*, vol. 49, no. 1, pp. 180–194, 2001.

[16] J. Zhu, Z. Ding, and X. Cao, "Column-anchored zero-forcing blind equalization for multiuser wireless FIR channels," *IEEE Journal on Selected Areas in Communications*, vol. 17, no. 3, pp. 411–423, 1999.

[17] X. Wang and H. V. Poor, "Blind equalization and multiuser detection in dispersive CDMA channels," *IEEE Trans. Communications*, vol. 46, no. 1, pp. 91–103, 1998.

[18] M. Torlak and G. Xu, "Blind multiuser channel estimation in asynchronous CDMA systems," *IEEE Trans. Signal Processing*, vol. 45, no. 1, pp. 137–147, 1997.

[19] J. M. Mendel, "Tutorial on higher-order statistics (spectra) in signal processing and system theory: theoretical results and some applications," *Proceedings of the IEEE*, vol. 79, no. 3, pp. 278–305, 1991.

[20] C. L. Nikias and A. P. Petropulu, *Higher-Order Spectra Analysis*, Prentice-Hall, Englewood Cliffs, NJ, USA, 1993.

[21] A. Swami, G. B. Giannakis, and G. Zhou, "Bibliography on higher-order statistics," *Signal Processing*, vol. 60, no. 1, pp. 65–126, 1997.

[22] G. B. Giannakis, Y. Inouye, and J. M. Mendel, "Cumulant based identification of multichannel moving-average models," *IEEE Trans. Automatic Control*, vol. 34, no. 7, pp. 783–787, 1989.

[23] A. Swami, G. B. Giannakis, and S. Shamsunder, "Multichannel ARMA processes," *IEEE Trans. Signal Processing*, vol. 42, no. 4, pp. 898–914, 1994.

[24] L. Tong, "Identification of multichannel MA parameters using higher-order statistics," *Signal Processing*, vol. 53, no. 2, pp. 195–209, 1996.

[25] Y. Inouye and K. Hirano, "Cumulant-based blind identification of linear multi-input-multiple-output system driven by colored inputs," *IEEE Trans. Signal Processing*, vol. 45, no. 6, pp. 1543–1552, 1997.

[26] J. K. Tugnait, "Blind spatio-temporal equalization and impulse response estimation for MIMO channels using a Godard cost function," *IEEE Trans. Signal Processing*, vol. 45, no. 1, pp. 268–271, 1997.

[27] J. K. Tugnait, "Identification and deconvolution of multichannel linear non-Gaussian processes using higher order statistics and inverse filter criteria," *IEEE Trans. Signal Processing*, vol. 45, no. 3, pp. 658–672, 1997.

[28] B. Chen and A. P. Petropulu, "Frequency domain blind MIMO system identification based on second- and higher-order statistics," *IEEE Trans. Signal Processing*, vol. 49, no. 8, pp. 1677–1688, 2001.

[29] R. Roy and T. Kailath, "ESPRIT—Estimation of signal parameters via rotational invariance techniques," *IEEE Trans. Acoustics, Speech, and Signal Processing*, vol. 37, no. 7, pp. 984–995, 1989.

[30] C. Chang, Z. Ding, S. F. Yau, and F. H. Y. Chan, "A matrix-pencil approach to blind separation of colored nonstationary signals," *IEEE Trans. Signal Processing*, vol. 48, no. 3, pp. 900–907, 2000.

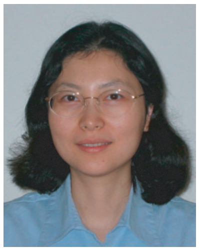
[31] C. T. Ma, Z. Ding, and S. F. Yau, "A two-stage algorithm for MIMO blind deconvolution of nonstationary colored signals," *IEEE Trans. Signal Processing*, vol. 48, no. 4, pp. 1187–1192, 2000.

[32] J. Liang and Z. Ding, "Higher order statistical approach for channel estimation using matrix pencils," in *Proc. IEEE International Conference on Communications*, New York, USA, April 2002.

[33] J. Liang and Z. Ding, "A cumulant subspace approach to FIR multiuser channel estimation," in *Proc. IEEE Workshop on Statistical Signal and Array Processing*, pp. 616–620, Pocono Manor, Pa, USA, August 2000.

[34] D. R. Brillinger, *Time Series: Data Analysis and Theory*, McGraw-Hill, New York, NY, USA, 1981.

[35] Z. Bai, J. Demmel, J. Dongarra, A. Ruhe, and H. van der Vorst, Eds., *Templates for the Solution of Algebraic Eigenvalue Problems: A Practical Guide*, SIAM, Philadelphia, Pa, USA, 2000.
[36] Z. Cao, "On a deflation method for the symmetric generalized eigenvalue problem," *Linear Algebra and Its Applications*, vol. 92, pp. 187–196, 1987.
[37] P. Hoeher, "A statistical discrete-time model for the WSSUS multipath channel," *IEEE Trans. Vehicular Technology*, vol. 41, no. 4, pp. 461–468, 1992.



Jing Liang received the B.E. and M.S. degrees in electrical engineering from Southeast University, Nanjing, China, in 1996 and 1998, respectively. She received the Ph.D. degree in electrical and computer engineering from the University of California, Davis, in 2002. From 1995 to 1998, she was a Research Assistant with the National Mobile Communications Research Laboratory, Southeast University, China. She was a Research Assistant with the Department of Electrical and Computer Engineering at the University of Iowa from 1999 to 2000, and at the University of California, Davis, from 2000 to 2002. Her general research interests include statistical signal processing, blind channel identification and equalization, digital communications, error-control codes, and iterative decoding algorithms.

Zhi Ding is a Professor of electrical and computer engineering at the University of California, Davis. He received his Ph.D. degree from the School of Electrical Engineering, Cornell University, Ithaca, NY, in 1990. From 1990 to 1998, he was a faculty member at Auburn University. He later served at the University of Iowa from 1999 till 2000. He has also held various visiting positions at the Australia National University, NASA Lewis Research Center, USAF Wright Laboratory (Eglin AFB), Hong Kong University of Science and Technology, and University of Hong Kong. He is a visiting Professor at Southeast University, China. His main research interests include digital wireless communications, signal detection, adaptive signal processing, blind equalization, and cyclostationary signal processing.