Matrix Outer-Product Decomposition Method for Blind Multiple Channel Identification

Zhi Ding, *Senior Member, IEEE*

***Abstract*— Blind channel identification and equalization have recently attracted a great deal of attention due to their potential application in mobile communications and digital TV systems. In this paper, we present a new algorithm that utilizes second-order statistics for multichannel parameter estimation. The algorithm is simple and relies on an outer-product decomposition. Its implementation requires no adjustment for single- or multiple-user systems. This new algorithm can be viewed as a generalization of a recently proposed linear prediction algorithm. It is capable of generating more accurate channel estimates and is more robust to overmodeling errors in channel order estimate. The superior performance of this new algorithm is demonstrated through simulation examples.**

I. Introduction

In popular data communication systems such as the digital mobile systems and digital television systems, data signals are often transmitted through unknown channels that may introduce severe linear distortion. In order to improve system performance, it is important for receivers to remove channel distortions through equalization or sequence estimation. Because the available input training signal may be too short or even nonexistent for channel identification, blind channel identification can play useful roles in these systems.

Blind channel identification relies solely on the received channel output signal and some *a priori* statistical knowledge (such as whiteness) of the original channel input signal. While blind equalization (deconvolution) is often investigated to directly identify the effective channel inverse, the possible existence of frequency nulls can result in undesirable noise enhancement for linear filter equalizers. One different path, which we adopt here, is to first identify the unknown system and then design receiver equalizers or sequence estimators accordingly to recover the original channel input.

Traditionally, blind channel identification and equalization are based on exploiting higher order statistics of baud-rate sampled channel output signals. The algorithm presented by Tong *et al.* [1], which is known as the TXK algorithm, is one of the first subspace based methods exploiting only second-order statistics for fractionally sampled channel identification. Using the subchannel representation of the fractionally sampled QAM channels, Xu *et al*. [2] derived a subchannel matching algorithm that also relies on the subspace separability of signal and noise. Another elegant subspace method for channel estimation similar to the well-known MUSIC algorithm in array application was presented by Moulines *et al*. [3]. Since subspace separability requires the knowledge of channel model orders, subspace methods tend to be sensitive to errors in channel order estimates.

A linear prediction-based approach was first presented by Slock [4], [5] and was later generalized and refined by Abe-Meriam *et al*. [6]. Unlike many of the subspace methods that tend to be very unreliable when the channel order is over-estimated, the linear prediction approach is found to be rather robust. However, as will become clear in this paper, the linear prediction algorithm (LPA) tends to discard some useful second-order statistical information of the channel output. In essence, the linear prediction algorithm is based only on the estimate of the first few columns of the channel parameter outer-product matrix, which depend critically on the leading coefficients of the unknown multi-channel impulse responses. Hence, the estimation error can be very large if the channel has a weak precursor impulse response. In order to derive a more robust algorithm, the focus of this paper is to attempt to derive the estimate based on a full outer-product decomposition of the channel parameter vector. Our results will show that based on the complete outer-product decomposition, performance of channel identification can be significantly improved.

This paper is organized as follows. In Section II, we first describe the statistical model of the blind multichannel identification problem. Spectral diversities achieved from oversampled channel output and multiple sensors (antennas) are considered in the multiuser channel estimation problem. We show that rational fractional sampling achieves an equivalent multiuser system. In Section III, a new outer-product decomposition method is developed. Its practical implementation is fully described. In Section IV, we consider the special case of single-user channel identification and the subsequent simplification of the new algorithm. Finally, simulation results are provided in Section V to illustrate the performance improvement of the new method.

II. Channel Identification Based on Second-Order Statistics

A. Problem Formulation

A multiuser quadrature amplitude modulation (QAM) data communication system can be described using a baseband rep-

Manuscript received April 30, 1996; revised May 21, 1997. This work was supported by NSF Grant MIP-9210100 by the U.S. Army Research Office. The associate editor coordinating the review of this paper and approving it for publication was Dr. Andreas S. Spanias.

The author is with The Department of Electrical and Electronic Engineering, Hong Kong University of Technology, Clear Water Bay, Kowloon, Hong Kong, on leave from the Department of Electrical Engineering, Auburn University, Auburn, AL 36839-5201 USA.

Publisher Item Identifier S 1053-587X(97)08559-0.

resentation. Assuming that the N user channels are all linear and causal with impulse response $\{h_u(t), u = 1, 2, \cdots N\}$, the received output signal can be written as

$$x(t) = \sum_{u=1}^{N} \sum_{k=-\infty}^{\infty} s_{k,u} h_u(t - kT) + w(t), \qquad s_{k,u} \in \mathcal{A}_u \tag{2.1}$$

where T is the symbol baud period, and $\mathcal{A}_u$ is the input signal set of user u. The channel input sequences $\{s_{k,u}\}$ are typically independent for different users and are i.i.d as well. The noise $w(t)$ is stationary, white, and independent of channel input sequences $s_{k,u}$ but not necessarily Gaussian.

Note that $h_u(t)$ is a "composite" channel impulse response that includes transmitter and receiver filters as well as the physical channel response. In a typical multiuser system, multiple channels of observations will be available from multiple sensors. If J subchannels (sensors or antennas) exist, then $x(t), h_u(t)$, and $w(t)$ are all $J \times 1$ vectors.

In blind channel identification, the objective is to identify the unknown channel responses $h_u(t)$ based on the channel output $x(t)$ only. Only the statistical knowledge of the channel input sequences is known but not their actual values. In blind equalization, the desired objective is to recover each channel input without knowing channel responses.

The problem of single-user $(N = 1)$ and single channel $(J = 1)$ blind identification and equalization has received a great deal of attention recently. Various methods utilizing higher order statistics have been proposed in the literature in works such as [7]–[14] and references therein.

B. Channel Diversity from Oversampling

It has been shown by Tong *et al*. [1] that blind channel identification benefits from oversampling the channel outputs. In fact, single channel identification based on second-order statistics is possible only for oversampled channel output. This essentially arises from the spectral diversity available when the channel has bandwidth higher than $1/2T$ [15].

Let the sampling interval be $\Delta = T/p$, where p is an integer. The oversampled discrete signals and responses are

$$x_i \triangleq x(i\Delta), \quad h_u[i] \triangleq h_u(i\Delta) \quad \text{and} \quad w_i \triangleq w(i\Delta) \tag{2.2}$$

each of which is a $J \times 1$ vector. The channel output samples are hence

$$x_n = \sum_{u=1}^{N} \sum_{k=-\infty}^{\infty} s_{k,u} h_u[n - kp] + w_n.$$

Suppose $\{h_u(t)\}$ has joint finite support $[0, T_h)$ that spans $m_0 + 1$ integer baud periods. Let Mp be the number sampled channel output to be collected in a block, and let superscript $\{\cdot\}'$ represent matrix transpose. By defining notations

$$\begin{aligned}
s_k &\triangleq [s_{k,1} \quad s_{k,2} \quad \cdots \quad s_{k,N}] \\
\boldsymbol{s}[k] &\triangleq [s_k \quad s_{k-1} \quad \cdots \quad s_{k-m_0-M+1}]' \\
\boldsymbol{w}[k] &\triangleq [w'_{kp} \quad w'_{kp+1} \quad \cdots \quad w'_{kp-Mp+1}] \\
\boldsymbol{h}_u[i] &\triangleq \begin{bmatrix} h_u[ip] \\ h_u[ip+1] \\ \vdots \\ h_u[ip+p-1] \end{bmatrix} \\
\boldsymbol{H}_i &\triangleq [\boldsymbol{h}_1[i] \quad \boldsymbol{h}_2[i] \quad \cdots \quad \boldsymbol{h}_N[i]]
\end{aligned}$$

it is evident that

$$\begin{bmatrix} x_{kp} \\ x_{kp+1} \\ \vdots \\ x_{kp+p-1} \end{bmatrix} = \sum_{i=0}^{m_0} \boldsymbol{H}_i s'_{k-i} + \begin{bmatrix} w_{kp} \\ w_{kp+1} \\ \vdots \\ w_{kp+p-1} \end{bmatrix}.$$

Now, form an $MpJ \times (m_0 + M)N$ block Toeplitz matrix with $(M-1)N$ trailing zeros in the first pJ rows

$$\boldsymbol{H} = \begin{bmatrix} \boldsymbol{H}_0 & \boldsymbol{H}_1 & \cdots & \boldsymbol{H}_{m_0} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \boldsymbol{H}_0 & \boldsymbol{H}_1 & \cdots & \boldsymbol{H}_{m_0} & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \mathbf{0} \\ \mathbf{0} & \cdots & \mathbf{0} & \boldsymbol{H}_0 & \boldsymbol{H}_1 & \cdots & \boldsymbol{H}_{m_0} \end{bmatrix}. \tag{2.3}$$

Clearly, m_0 is the order of the N dynamic FIR channels. There are a total of $(m_0+1)NpJ$ unknown parameters to identify in the blind identification of FIR channels. With these notations, a sampled channel output signal vector of length Mp can be written as

$$\boldsymbol{x}[k] \triangleq \begin{bmatrix} x_{kp} \\ x_{kp+1} \\ \vdots \\ x_{kp+p-1} \\ x_{(k-1)p} \\ x_{(k-1)p+1} \\ \vdots \\ x_{kp-Mp+1} \end{bmatrix} = \boldsymbol{H}\boldsymbol{s}[k] + \boldsymbol{w}[k]. \tag{2.4}$$

C. Fractional Oversampling

Historically, there has been a belief that a noninteger oversampling factor may be more beneficial. In fact, it is heuristically plausible that an oversampling period of $\Delta = qT/p$ in which q is an integer will yield fewer nonzero channel samples than using $\Delta = T/p$. Hence, argument persists that a rational oversampling factor may help reduce the dimensionality of channel identification and simplify the problem. Here, we show that, in fact, a noninteger fractional sampling generates an equivalent multiuser system whose dimensionality is not reduced and may be more difficult to identify.

Let $\Delta = qT/p$, where p and q are co-prime integers. The noiseless received signal becomes

$$\begin{aligned} x_n &= \sum_{u=1}^{N} \sum_{L=-\infty}^{\infty} s_{k,u} h_u(n\Delta - LT) \\ &= \sum_{u=1}^{N} \sum_{i=0}^{q-1} \sum_{k=-\infty}^{\infty} s_{kq+i,u} h_u(n\Delta - kqT - iT). \end{aligned}$$

By defining equivalent signal sequences

$$\bar{s}_{k,i+1+(u-1)q} \triangleq s_{kq+i,u}, \quad i = 0, 1, \cdots q-1 \quad u = 1, 2, \cdots N \tag{2.5}$$

and equivalent subchannel responses

$$\bar{h}_{i+1+(u-1)q}[k] = h_u(k\Delta - iT), \quad u = 1, 2, \cdots N \quad i = 0, 1, \cdots q-1 \tag{2.6}$$

the received signal can be viewed as an output of Nq user channels

$$x_n = \sum_{v=1}^{Nq} \sum_{k=-\infty}^{\infty} \bar{s}_{k,v} \bar{h}_v[n-kp] + w_n. \tag{2.7}$$

It is therefore clear that an N user system sampled at interval of $\Delta = qT/p$ is equivalent to an Nq user system. We can thus formulate the rationally sampled multiuser system accordingly.

Notice that all channel impulse responses are assumed to be finite such that $h(mT) = 0$ for $m \geq m_0$. Hence, for $\Delta = T/p$

$$\begin{aligned} h_u[mp] &= 0, \qquad m > m_0 \quad \text{and} \\ \boldsymbol{H}_m &= \boldsymbol{0}, \qquad m > m_0. \end{aligned} \tag{2.8}$$

There are a total of $(m_0 + 1)NpJ$ unknown parameters to identify in $\boldsymbol{H}$. For noninteger oversampling that generates an equivalent Nq user system, we have

$$\bar{h}_v[mp] = 0, \qquad m > m_1$$

where

$$m_1 \triangleq \left\lceil \frac{m_0}{q} \right\rceil. \tag{2.9}$$

Hence, the blind identification problem of the fractionally sampled multiuser system can also be described by (2.4), where $\boldsymbol{H}$ is an $MpJ \times (M + m_1)Nq$ block Toeplitz matrix. The total number unknown parameters for identification is

$$\begin{aligned} (m_1 + 1)qNpJ &= (m_1 q + q)NpJ \geq (m_0 + q)NpJ \\ &\geq (m_0 + 1)NpJ. \end{aligned}$$

It is therefore clear that the number of unknown parameters for identification using noninteger fractional sampling is no less than that using integer fractional sampling. We have thus established that noninteger fractional sampling offers no reduction in identification cost and in the cost of implementing Viterbi algorithm.

If noninteger fractional sampling is utilized, it has been shown in [1] and [5] that the sufficient and necessary identification condition for $\boldsymbol{H}$ to be identifiable from second-order statistics is that $\boldsymbol{H}$ must be full rank. The necessary dimensional condition for $\boldsymbol{H}$ to be full rank requires that

$$MJp \geq (m_1 + M)Nq \tag{2.10}$$

i.e.,

$$M(Jp - qN) \geq m_1 qN. \tag{2.11}$$

Hence, unless m_1 is zero, which means that the sampled channel is trivial and has no memory, the fractional sampling must satisfy

$$Jp \geq Nq.$$

This implication is simple: The number of equivalent multichannels (Jp) must be no less than the number of equivalent users (Nq). This also shows that when a single user is present for a single channel, any amount of oversampling $(p > q)$ will satisfy the necessary dimensional condition.

Overall, the use of noninteger fractional sampling results in an additional identification ambiguity in that channels can only be identified subject to an $Nq \times Nq$ constant unitary matrix, as will be shown later. Our derivation clearly shows that there is neither computational nor algorithmic advantage in the use of noninteger fractional sampling.

D. Channel Identification

The additional channel zero condition for $\boldsymbol{H}$ to be full rank has been characterized in [5] and is not the focus of our work. We shall assume from here on that $\boldsymbol{H}$ has full column rank and is identifiable. Moreover, we shall also assume, without loss of generality, that the oversampling factor is an integer p while $q = 1$.

Assume that both the channel input signal and channel noise are white with zero mean. Let their respective covariance matrix be

$$R_s = E\{\boldsymbol{s}[k]\boldsymbol{s}[k]^H\} = \sigma_s^2 I$$

and

$$R_w = E\{\boldsymbol{w}[k]\boldsymbol{w}[k]^H\} = \sigma_w^2 I.$$

Based on (2.4), the channel output covariance matrix becomes

$$R_{\boldsymbol{x}} = E\{\boldsymbol{x}[k]\boldsymbol{x}[k]^H\} = \sigma_s^2 \boldsymbol{H}\boldsymbol{H}^H + \sigma_w^2 I. \tag{2.12}$$

Our objective is to identify the channel $\boldsymbol{H}$ from the second-order statistics of the channel output signal $\boldsymbol{x}[k]$ given in $R_{\boldsymbol{x}}$ under the identifiability condition [1] that both $\boldsymbol{H}$ and R_s are full rank. The use of second-order statistics for single user blind channel identification $(N = 1)$ was first exploited by Tong *et al.* [1]. The basic concept hinges on the signal and noise subspace separation through singular value decomposition (SVD) of the auto-covariance matrix $R_{\boldsymbol{x}}$.

The sub-channel matching (SCM) method presented in [2] and the subspace method of [3] can both be posed as a minimum eigenvector problem under proper channel length constraints. The special block Toeplitz structure is utilized in both algorithms. When the channel length is overestimated, common zeros must be factorized out from the subchannel estimates. As a result, both algorithms are very sensitive to channel length mismatching.

A nonlinear maximum likelihood method was presented by Hua [16] that utilized the SCM as the first step of a two-step maximum likelihood (TSML) optimization method. Given a good initial estimate from SCM, this TSML method was shown to provide improved performance.

In [4] and [6], a linear prediction algorithm (LPA) was presented for channel estimation. It is shown to be more robust to overestimated channel length. Still, as will become evident later in this paper, the LPA only uses part of the overall information because the channel estimate is based on the first pJ columns of the estimated channel parameter vector outer-product matrix. As a more robust and accurate channel estimation algorithm, the outer-product decomposition algorithm we propose will exploit second-order statistics more effectively. In addition, unlike many existing works such as the SCM and TSML, our method is virtually unchanged for both single and multiuser systems, as for LPA [6].

III. Algorithm Development

A. Outer-Product Construction

We will form an outer-product of the channel parameter matrix

$$\mathcal{H} \triangleq \begin{bmatrix} \boldsymbol{H}_0 \\ \boldsymbol{H}_1 \\ \vdots \\ \boldsymbol{H}_{m_0} \end{bmatrix}. \tag{3.1}$$

Our objective is to derive a method that would allow us to form an outer-product of the channel parameter vector $\mathcal{H}$ based on the second-order channel output statistics.

First, assume that the channel order m_0 is known. Let a $(m_0+1)pJ \times (m_0+M)N$ block Hankel matrix be denoted as

$$H_a \triangleq \begin{bmatrix} \boldsymbol{H}_0 & \boldsymbol{H}_1 & \cdots & \boldsymbol{H}_{m_0} & \mathbf{0} & \cdots & \mathbf{0} \\ \boldsymbol{H}_1 & \boldsymbol{H}_2 & \cdots & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots & \vdots & & \vdots \\ \boldsymbol{H}_{m_0} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \end{bmatrix}. \tag{3.2}$$

Notice that the first pJ rows of H_a and $\boldsymbol{H}$ are identical. Denote superscript $\{\cdot\}^H$ as conjugate transpose. It can be verified that

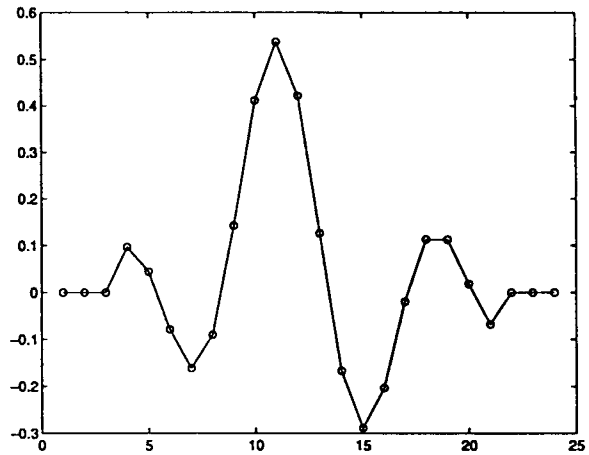

Fig. 1. Overall channel impulse response $h(t)$.

we have (3.3), shown at the bottom of the page, in which we must recall that $\boldsymbol{H}_m = 0, m > m_0$.

If we define $pJ \times pJ$ block matrices as

$$D_{i,j} \triangleq \sum_{k=i-1}^{m_0} \boldsymbol{H}_k \boldsymbol{H}_{k+j-i}^H, \quad 1 \le i,j \le m_0+1 \tag{3.4}$$

then

$$\boldsymbol{D}_1 \triangleq \begin{bmatrix} D_{1,1} & D_{1,2} & \cdots & D_{1,m_0+1} \\ D_{2,1} & D_{2,2} & \cdots & D_{2,m_0+1} \\ \vdots & \vdots & \ddots & \vdots \\ D_{m_0+1,1} & D_{m_0+1,2} & \cdots & D_{m_0+1,m_0+1} \end{bmatrix} = H_a H_a^H. \tag{3.5}$$

This is an $(m_0+1)pJ \times (m_0+1)pJ$ Hermitian matrix. Now, define a new matrix from the lower right block of $H_a H_a^H$ as

$$\boldsymbol{D}_2 = \begin{bmatrix} D_{2,2} & \cdots & D_{2,m_0+1} & \mathbf{0} \\ D_{3,2} & \cdots & D_{3,m_0+1} & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ D_{m_0+1,2} & \cdots & D_{m_0+1,m_0+1} & \mathbf{0} \\ \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \end{bmatrix}. \tag{3.6}$$

$$H_a H_a^H = \begin{bmatrix} \sum_{k=0}^{m_0} \boldsymbol{H}_k \boldsymbol{H}_k^H & \sum_{k=0}^{m_0} \boldsymbol{H}_k \boldsymbol{H}_{k+1}^H & \cdots & \sum_{k=0}^{m_0} \boldsymbol{H}_k \boldsymbol{H}_{k+m_0}^H \\ \sum_{k=1}^{m_0} \boldsymbol{H}_k \boldsymbol{H}_{k-1}^H & \sum_{k=1}^{m_0} \boldsymbol{H}_k \boldsymbol{H}_k^H & \cdots & \sum_{k=1}^{m_0} \boldsymbol{H}_k \boldsymbol{H}_{k+m_0-1}^H \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{k=m_0}^{m_0} \boldsymbol{H}_k \boldsymbol{H}_{k-m_0}^H & \sum_{k=m_0}^{m_0} \boldsymbol{H}_k \boldsymbol{H}_{k-m_0+1}^H & \cdots & \sum_{k=m_0}^{m_0} \boldsymbol{H}_k \boldsymbol{H}_k^H \end{bmatrix} \tag{3.3}$$

We can form another Hermitian matrix from

$$\begin{aligned}\Delta D &\triangleq \boldsymbol{D}_1 - \boldsymbol{D}_2\\ &= \begin{bmatrix} \boldsymbol{H}_0\boldsymbol{H}_0^H & \boldsymbol{H}_0\boldsymbol{H}_1^H & \cdots & \boldsymbol{H}_0\boldsymbol{H}_{m_0}^H \\ \boldsymbol{H}_1\boldsymbol{H}_0^H & \boldsymbol{H}_1\boldsymbol{H}_1^H & \cdots & \boldsymbol{H}_1\boldsymbol{H}_{m_0}^H \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{H}_{m_0}\boldsymbol{H}_0^H & \boldsymbol{H}_{m_0}\boldsymbol{H}_1^H & \cdots & \boldsymbol{H}_{m_0}\boldsymbol{H}_{m_0}^H \end{bmatrix}\\ &= \begin{bmatrix} \boldsymbol{H}_0 \\ \boldsymbol{H}_1 \\ \vdots \\ \boldsymbol{H}_{m_0} \end{bmatrix} [\boldsymbol{H}_0^H \quad \boldsymbol{H}_1^H \quad \cdots \quad \boldsymbol{H}_{m_0}^H] = \mathcal{H}\mathcal{H}^H.\end{aligned}$$

Hence, matrix ΔD forms the outer-product of the channel parameter matrix $\mathcal{H}$. The singular value decomposition of this outer-product matrix can be used to generate an estimate

$$\mathcal{H}Q$$

where Q is an $N \times N$ unitary matrix. Recall from [13] and [17] that this memoryless ambiguity is intrinsic to the multiuser blind identification problem and cannot be resolved unless additional information is available. For signal recovery, if a perfect multichannel equalizer is designed according to the channel estimate $\mathcal{H}Q$, then the N receiver outputs will be memoryless combinations of the N channel inputs and will need to be separated, as discussed in works such as [17]. This ambiguity would also add to the cost of channel identification using noninteger fractional sampling.

B. Outer-Product Estimation

Based on the above derivation, the key step in the algorithm is to obtain the matrix product $H_a H_a^H$ that can be used to define an estimate of the channel parameter vector outer-product. Hence, the crucial step in our algorithm development is to find an estimate of the matrix product $\boldsymbol{D_1}$ from the statistics of the channel output signal $\boldsymbol{x}[k]$. Since we focus on the use of second-order statistics, our task is to find an estimate of the matrix product $\boldsymbol{D}_1$ given $R_{\boldsymbol{x}}$.

Let

$$X[k] \triangleq \begin{bmatrix} x_{kp} \\ x_{kp+1} \\ \vdots \\ x_{kp+p-1} \end{bmatrix} = \sum_{i=0}^{m_0} \boldsymbol{H}_i s'_{k-i}. \tag{3.7}$$

For notational convenience, define

$$R(n) \triangleq E\{X[k]X[k-n]^H\} = E\{|s_k|^2\} \sum_{i=n}^{m_0} \boldsymbol{H}_i \boldsymbol{H}_{i-n}^H. \tag{3.8}$$

The channel output covariance covariance matrix can be written as

$$\begin{aligned}R_{\boldsymbol{x}} &\triangleq E\{\boldsymbol{x}[k]\boldsymbol{x}^H[k]\}\\ &= \begin{bmatrix} R(0) & R(1) & \cdots & R(m_0) \\ R(1)^H & R(0) & \cdots & R(m_0-1) \\ \vdots & \vdots & \ddots & \vdots \\ R(m_0)^H & R(m_0-1)^H & \cdots & R(0) \end{bmatrix}\\ &= \sigma_s^2 \boldsymbol{H}\boldsymbol{H}^H + \sigma_w^2 I.\end{aligned} \tag{3.9}$$

First, it is easy to verify that another block Hankel matrix satisfies the relationship

$$\begin{aligned}R_a &= \begin{bmatrix} R(0) - \sigma_w^2 I & R(1) & \cdots & R(m_0) \\ R(1) & R(2) & \cdots & \boldsymbol{0} \\ \vdots & \vdots & \ddots & \vdots \\ R(m_0) & \boldsymbol{0} & \cdots & \boldsymbol{0} \end{bmatrix}\\ &= \sigma_s^2 H_a \boldsymbol{H}^H.\end{aligned} \tag{3.10}$$

In addition, it is also evident that

$$R_{\boldsymbol{x}} - \sigma_w^2 I = \sigma_s^2 \boldsymbol{H}\boldsymbol{H}^H. \tag{3.11}$$

In order to estimate the product $H_a H_a^H$, it is important to note that when $\boldsymbol{H}$ has full column rank

$$\boldsymbol{H}^H(\boldsymbol{H}\boldsymbol{H}^H)^{\#}\boldsymbol{H} = I. \tag{3.12}$$

Note that $(\boldsymbol{H}\boldsymbol{H}^H)^{\#}$ denotes the pseudo-inverse of $\boldsymbol{H}\boldsymbol{H}^H$.

Recall that the sufficient and necessary identification condition for $\boldsymbol{H}$ to be identifiable from second-order statistics is that $\boldsymbol{H}$ must be full rank [1], [5]. As a result, if the multichannel system is identifiable from second-order statistics, the matrix product $\boldsymbol{D}_1$ can be estimated from

$$\begin{aligned}\hat{\boldsymbol{D}}_{\boldsymbol{1}} &= R_a(R_{\boldsymbol{x}} - \sigma_w^2 I)^{\#} R_a^H\\ &= \sigma_s^2 H_a \boldsymbol{H}^H \sigma_s^{-2}(\boldsymbol{H}\boldsymbol{H}^H)^{\#} \sigma_s^2 \boldsymbol{H} H_a^H = \sigma_s^2 H_a H_a^H.\end{aligned} \tag{3.13}$$

In many digital communication systems, σ_s^2 is known, and hence, we can obtain the estimate of $H_a H_a^H$ via

$$\hat{\boldsymbol{D}}_{\boldsymbol{1}} = \sigma_s^{-2} R_a (R_{\boldsymbol{x}} - \sigma_w^2 I)^{\#} R_a^H. \tag{3.14}$$

Consequently, the channel impulse response matrix $\mathcal{H}$ can be estimated from the singular value decomposition of the estimate of outer-product matrix $\widehat{\Delta D}$.

$$\hat{\mathcal{H}} = \text{SVD}(\widehat{\Delta D}) = \mathcal{H}Q. \tag{3.15}$$

We thus name the method "outer-product decomposition algorithm" (OPDA).

C. Practical Considerations and Implementation

Based on the algorithm derivations in the previous section, we can summarize the algorithm into the following steps.

1) Given K baud samples of the channel output data $\{X[1], \cdots, X[K]\}$, form the auto-correlation submatrices

$$\hat{R}(n) = \frac{1}{K-n} \sum_{k=1}^{K-n} X[n+k]X[k]^H, \quad n = 0, 1, \cdots N. \tag{3.16}$$

and form the estimate of the auto-covariance matrix

$$\hat{R}_{\boldsymbol{x}}[M] = \begin{bmatrix} \hat{R}(0) & \hat{R}(1) & \cdots & \hat{R}(M-1) \\ \hat{R}(1)^H & \hat{R}(0) & \cdots & \hat{R}(M-2) \\ \vdots & \vdots & \ddots & \vdots \\ \hat{R}(M-1)^H & \hat{R}(M-2)^H & \cdots & \hat{R}(0) \end{bmatrix}. \tag{3.17}$$

2) Estimate the channel order m_0 from $\hat{R}_{\boldsymbol{x}}$ by first applying the MDL signal rank test [18] and then determine

$$\hat{m}_0 = \frac{1}{N} \text{ signal rank } (\hat{R}_{\boldsymbol{x}}[M]) - M$$

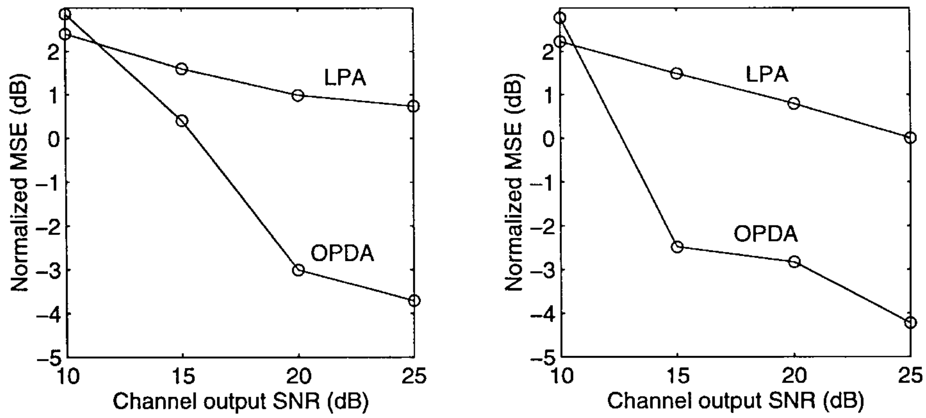

Fig. 2. Normalized MSE of channel estimate given different SNR levels.

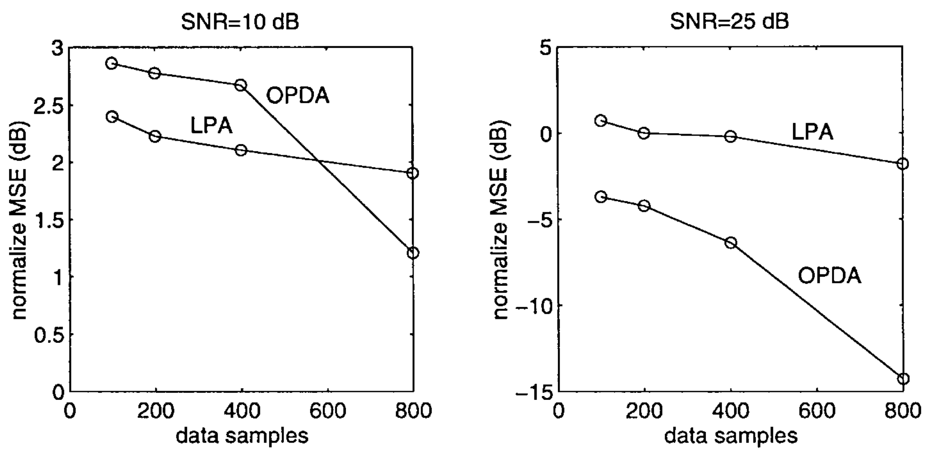

Fig. 3. Normalized MSE of channel estimate given different data lengths.

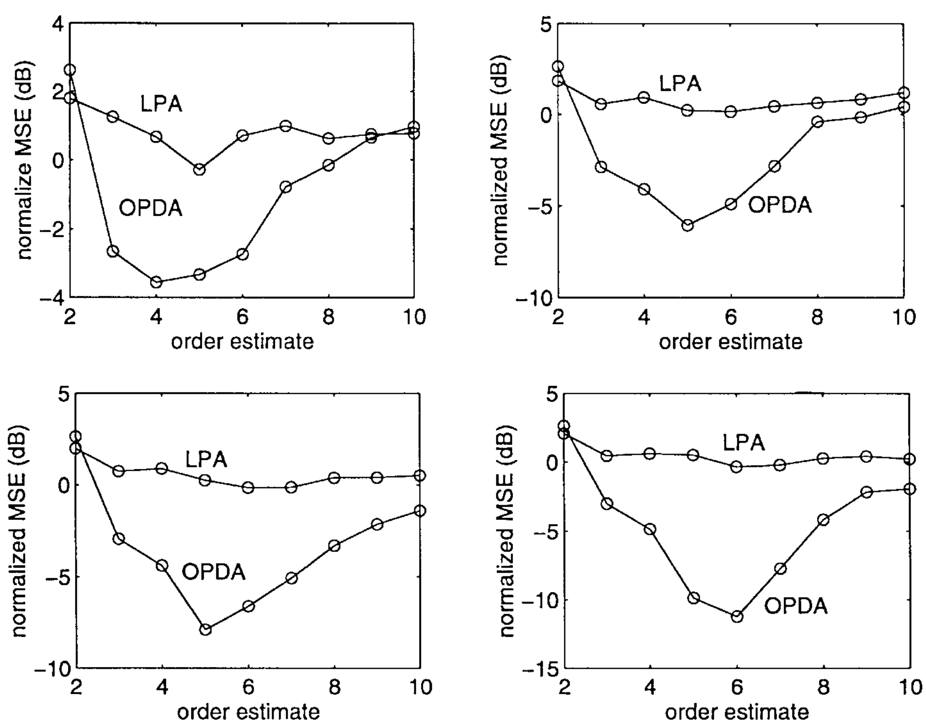

Fig. 4. Normalized MSE of channel estimate given channel length mismatch.

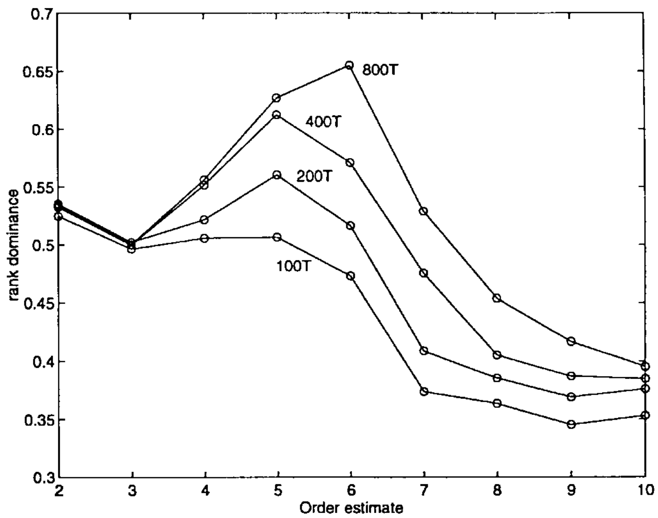

Fig. 5. Rank dominance factor as a function of the channel length estimate.

and estimate the noise variance as the average of the $MpJ-(\hat{m}_0+M)N$ smallest eigenvalues

$$\widehat{\sigma_w^2} = \frac{1}{MpJ-(\hat{m}_0+M)N} \sum_{i=\text{signal rank}(\hat{R}_{\boldsymbol{x}}[M])+1}^{MpJ} \lambda_i.$$

3) Based on the estimated channel order m_0, form matrices

$$\hat{R}_a = \begin{bmatrix} \hat{R}(0)-\sigma_w^2 I & \hat{R}(1) & \cdots & \hat{R}(\hat{m}_0) \\ \hat{R}(1) & \hat{R}(2) & \cdots & \hat{R}(\hat{m}_0+1) \\ \vdots & \vdots & \ddots & \vdots \\ \hat{R}(\hat{m}_0) & \hat{R}(\hat{m}_0+1) & \cdots & \hat{R}(2\hat{m}_0) \end{bmatrix} \tag{3.18}$$

and

$$\hat{\boldsymbol{D}}_{\mathbf{1}}[m_0] = \hat{R}_a(\hat{R}_{\boldsymbol{x}}[\hat{m}_0] - \widehat{\sigma_w^2} I)^{\#} \hat{R}_a^H. \tag{3.19}$$

4) Find the channel impulse response matrix $\mathcal{H}$ as the N eigenvectors corresponding to the N largest eigenvalues of the matrix $\widehat{\Delta D}$ using standard algorithms for eigendecomposition or singular value decomposition [21].

D. Information for Order Estimation from Rank Dominance

When the channel order is underestimated or overestimated, one immediate impact is the significant departure of $\hat{\Delta D}$ from the actual matrix product. Consequently, the outer-product estimate tends to be perturbed away from a rank N (dominated) matrix. When the channel order is underestimated, the mismodeling error tends to greatly reduce the rank dominance of the N largest eigenvectors. When the channel order is overestimated, however, the additional noise intrusion in R_a and $R_{\boldsymbol{x}}$ estimates, together with a higher dimensional $\hat{\Delta D}$, will also lessen the dominance of the channel parameter vectors.

Based on these observations, OPDA may be further enhanced in the channel order estimation stage by checking the rank dominance of the largest eigenvalue in $\hat{\Delta D}[m_0]$ for different order estimate m_0. Define the dominance factor of the outer-product estimate as

$$f_d(m_0) \triangleq \frac{\sum_{i=1}^{N} \lambda_i}{\text{Trace }(\widehat{\Delta D}[m_0])}. \tag{3.20}$$

The dominance information of the first N ranks of the outer-product estimate can be used to assist in the estimation of channel order by selecting m_0 that maximizes the dominance function. It can be used to signal the reliability of the channel identification results.

IV. Oversampled Single User Channel Identification

A. Maximum Eigenvector Solution

For a single user whose channel output is oversampled by an integer p, the effective user is one $(N=1)$, and the outer product decomposition can be uniquely determined as ΔD is ideally a rank one matrix. Hence, the channel impulse response vector can be estimated via

$$\hat{\mathcal{H}} = \arg\max_v v^H(\widehat{\Delta D})v. \tag{4.1}$$

In other words, $\mathcal{H}$ is estimated as the maximum eigenvector of the outer-product estimate. Although, theoretically, the outer-product matrix is a rank one matrix, the practical estimate is likely to have higher dimensions. This explains why we would prefer to use eigendecomposition for channel estimate. Alternatively, QR decomposition [21] may be a faster approach to channel estimation. An even faster but less accurate method is to postmultiply $\hat{\Delta D}$ with a random vector.

Notice that OPDA requires two singular value (or eigenvalue) decompositions in its implementation. Its order of complexity is therefore similar to that of the linear prediction algorithm (LPA) presented by Meriam *et al*. [6], the TXK method [1], and the subchannel matching method [2]. However, LPA estimates the channel only from the first pJ columns of the outer-product matrix. If the channel impulse response has weak precursor samples such that its leading coefficients are small, then LPA is likely to be highly inaccurate since noise and numerical error will likely dominate the first few columns of $\widehat{\Delta D}$. Therefore, OPDA is expected to provide more accurate result than LPA.

B. Simplified Outer-Product Decomposition Algorithm (SOPDA)

For $N=1$, the last step of OPDA can also be simplified by estimating $\boldsymbol{H}$ as the first column of the faster QR decomposition of $\widehat{\Delta D}[\hat{m}_0]$. In fact, if the receiver computation power is severely limited in practical systems such that it becomes impossible to perform the entire four steps of OPDA, a less accurate and simpler method can be implemented. Here, we summarize a simplified OPDA algorithm.

1) Complete step 1) of OPDA by selecting a large enough M.

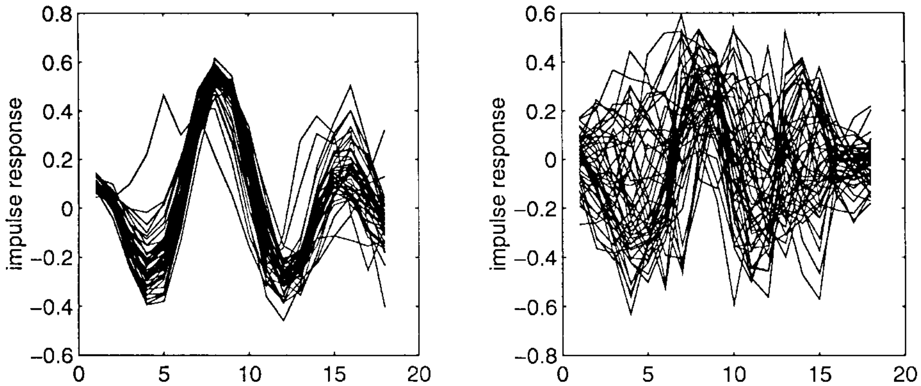

Fig. 6. Fifty independent channel estimates.

2) Form a matrix

$$\hat{R}_a = \begin{bmatrix} \hat{R}(0) & \hat{R}(1) & \cdots & \hat{R}(M-1) \\ \hat{R}(1) & \hat{R}(2) & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \hat{R}(M-1) & \mathbf{0} & \cdots & \mathbf{0} \end{bmatrix}. \quad (4.2)$$

3) Form

$$\hat{\boldsymbol{D}}_{\mathbf{1}} = \hat{R}_a R_{\boldsymbol{x}}[M]^{-1} \hat{R}_a^H \quad (4.3)$$

and $\widehat{\Delta D}$.

4) Estimate channel $\boldsymbol{H}$ as the first column of the Q matrix in the QR decomposition [21] of $\widehat{\Delta D}$.

V. Simulation Results

We now present simulation results to illustrate the channel identification performance of the proposed OPDA. Our experiments are based on a multipath channel model with a single sensor, i.e., $J = 1$. We consider a raised-cosine pulse $P(t)$ limited in $6T$ with roll-off factor 0.10 and a two-ray multipath channel

$$c(t) = \delta(t) - 0.7\delta(t - T/3).$$

The overall channel impulse response

$$h(t) = c(t) * P(t) = P(t) - 0.7P(t - T/3)$$

is shown in Fig. 1. A single user is assumed. The data input signal is i.i.d. BPSK, and the oversampling factor is $p = 3$. In all our simulations, M is chosen to be twice as long as $P(t)$.

In the first set of simulation tests, we compare the two methods OPDA and LPA based on 100 and 200 bauds of channel output samples. The channel order is unknown and is estimated using the MDL criterion. The normalized mean square error (MSE) is defined as

$$E\left\{ \frac{\sum_k |h[k] - \hat{h}[k]|^2}{\sum_k |h[k]|^2} \right\}.$$

The channel estimate under different channel SNR levels is shown in Fig. 2. It is apparent from the simulation results that OPDA outperforms LPA in most cases. When the SNR is very low, the two algorithms are comparable and perform equally poorly.

To show the effect of data length on the accuracy of channel estimation, we implement OPDA and LPA for several different data lengths. The resulting normalized MSE is shown in Fig. 3. Once again, the results show that OPDA and LPA are equally ineffective when SNR is low. The primary reason is the inaccurate channel order estimation using MDL. However, when the SNR is higher, the channel order estimates are more reliable, and subsequently, the OPDA outperforms LPA significantly. The performance improvement is more pronounced when a large amount of data are available for statistical approximation.

It is apparent from the estimation results that this particular channel is difficult to estimate. The main difficulty lies in the estimation of channel length. Since the channel impulse response has very small tails on both sides, accurate length determination based on noisy and short data collection is very hard to obtain. Since LPA was presented as an algorithm that is less sensitive to length mismatching, we would like to test the comparative sensitivity of the two algorithms when channel mismatching is present; see Fig. 4. Fixing SNR $= 20$ dB, we manually varied the channel length estimate from 2–10. Notice from Fig. 1 that the true channel length is $6T$. The results clearly show that while LPA is less sensitive to errors in channel order estimate, its performance is generally much worse compared with that of OPDA. When the channel order estimate deviates modestly from the true channel order, OPDA generates a much smaller normalized MSE.

Fig. 5 illustrates the dominance factor as a function of various channel length estimates. It is apparent that when the order estimate is close to the real channel order, the dominance factor is near its peak. This is a strong indication that when computation power permits, the dominance factor can be used to assist in channel order estimation.

Finally, we compare a group of typical impulse responses estimated from 50 independent trials of the OPDA and LPA under 20 dB SNR and data length of $L = 400T$. Assuming the channel length is correctly estimated, the estimated impulse responses are shown in Fig. 6.

VI. Conclusions

We present a new robust and accurate blind channel identification algorithm OPDA based on matrix outer-product decomposition. This new algorithm can be viewed as a generalized method of the recently proposed linear prediction algorithm (LPA). The new OPDA is capable of generating superior identification results. Its application to multiuser and rationally oversampled systems are simple and direct. For single-user channel identification, its implementation can also be approximated using far less computation power in exchange for less accurate estimates. Furthermore, the implementation of OPDA also provides a rank dominance factor test that can either be used as an indication of output reliability or as additional information for more accurate estimation of the unknown channel order.

REFERENCES

[1] L. Tong *et al*., "A new approach to blind identification and equalization of multipath channels," *IEEE Trans. Inform. Theory*, vol. IT-40, pp. 340–349, Mar. 1994.

[2] G. Xu, H. Liu, L. Tong, and T. Kailath "A lest-squares approach to blind channel identification," *IEEE Trans. Signal Processing*, vol. 43, pp. 2982–2993, Dec. 1995.

[3] 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, pp. 516–525, Feb. 1995.

[4] D. Slock, "Blind fractionally-spaced equalization, perfect-reconstruction filter banks and multichannel linear prediction," in *Proc. 1994 IEEE ICASSP*, Adelaide, Australia, May 1994, pp. IV:585–588.

[5] ——, "Blind joint equalization of multiple synchronous mobile users using oversampling and/or multiple antennas," in *Proc. 1994 Asilomar Conf. Signals, Syst., Comput.*, Nov. 1994, pp. 1154–1158.

[6] K. Abed-Meriam *et al.*, "Prediction error methods for time-domain blind identification of multichannel FIR filters," in *Proc. 1995 IEEE ICASSP*, Detroit, MI, May 1995, pp. 1968–1971.

[7] A. Benveniste, M. Goursat, and G. Ruget. "Robust identification of a nonminimum phase system" *IEEE Trans. Automat. Contr.*, vol. AC-25, pp. 385–399, June 1980.

[8] D. N. Godard, "Self-recovering equalization and carrier tracking in two-dimensional data communication systems," *IEEE Trans. Commun.*, vol. COMM-28, pp. 1867–1875, 1980.

[9] O. Shalvi and E. Weinstein, "New criteria for blind deconvolution of nonminimum phase systems (channels)," *IEEE Trans. Inform. Theory*, vol. IT-36, pp. 312–321, Mar. 1990.

[10] J. R. Treichler and B. G. Agee, "A new approach to multipath correction of constant modulus signals," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-31, pp. 349–372, Apr. 1983.

[11] G. B. Giannakis, Y. Inouye, and J. M. Mendel, "Cumulant based identification of multichannel moving-averaging methods" *IEEE Trans. Automat. Contr.*, vol. 34, pp. 783–787, 1989.

[12] S. Mayrargue, "Spatial equalization of a radio-mobile channel without beamforming using the constant modulus algorithm (CMA)," in *Proc. IEEE Intl. Conf. Acoust., Speech, Signal Processing*, May 1993, pp. III:344–347.

[13] J.-F. Cardoso, "Source separation using high order moments," in *Proc. IEEE ICASSP*, 1989, pp. 2109–2112.

[14] Y. Li and Z. Ding, "A new nonparametric cepstral method for blind channel identification from cyclostationary statistics," in *Proc. 1993 MILCOM*, Boston, MA, Oct. 1993, pp. 648–652.

[15] Z. Ding and Y. Li, "On channel identification based on second order cyclic spectra," *IEEE Trans. Signal Processing*, vol. 42, pp. 1260–1264, May 1994.

[16] Y. Hua, "Fast maximum likelihood for blind identification of multiple FIR channels," *IEEE Trans. Signal Processing*, vol. 44, pp. 661–672, Mar. 1996.

[17] X.-R. Cao and R.-W. Liu, "General approach to blind source separation," *IEEE Trans. Signal Processing*, vol. 44, pp. 562–571, Mar. 1996.

[18] M. Wax and T. Kailath, "Detection of signals by information theoretic criteria," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-33, pp. 387–392, Apr. 1985.

[19] S. Shamsunder and G. B. Giannakis, "Detection and parameter estimation of multiple nongaussian sources via higher order statistics," *IEEE Trans. Signal Processing*, vol. 42, pp. 1145–1155, May 1994.

[20] D. T. M. Slock and C. B. Papadias, "Further results on blind identification and equalization in multiple FIR channels," in *Proc. IEEE Intl. Conf. Acoust., Speech, Signal Processing*, May 1995, pp. 1964–1967.

[21] G. H. Golub and C. F. Van Loan, *Matrix Computations*, 2nd ed. Baltimore, MD: Johns Hopkins Univ. Press, 1989.

Zhi Ding (SM'95) was born in Harbin, China. He received the B.Eng. degree in July 1982 from the Department of Wireless Engineering, Nanjing Institute of Technology, Nanjing, China, and the M.A.Sc. degree from the Department of Electrical Engineering, University of Toronto, Toronto, Ont., Canada, in May 1987. He received the Ph.D. degree from the School of Electrical Engineering, Cornell University, Ithaca, NY, in August 1990.

He joined the faculty of Auburn University, Auburn, AL, in September 1990, where he is currently an associate professor with the Department of Electrical Engineering. He has held visiting positions in the Australian National University, Canberra, NASA Lewis Research Center, the USAF Wright Laboratory, and Hong Kong University of Science and Technology. His research interests include digital communications, signal processing, adaptive signal processing, blind equalization, and cyclostationary signal processing.

Dr. Ding is a member of the IEEE SSAP Technical Committee.