Blind Identification of MIMO Channels Using Optimal Periodic Precoding

Ching-An Lin and Yi-Sheng Chen

***Abstract*—We propose a method for blind identification of multiple-input mutiple-out (MIMO) finite-impulse response (FIR) channels that exploits cyclostationarity of the received data induced at the transmitters by periodic precoding. It is shown that, by properly choosing the precoding sequence, the MIMO FIR transfer functions, with M_t inputs and M_r outputs, can be identified up to a unitary matrix ambiguity. The transfer functions need not be irreducible or column reduced, and there can be more outputs $(M_r \geq M_t)$ or more inputs $(M_r < M_t)$. The method exploits the linear relation between the covariance matrix of the received data and the "channel product matrices". The method is shown to be robust with respect to channel-order overestimation. The proposed algorithm requires solving linear equations and computing the nonzero eigenvalues and eigenvectors of a Hermitian positive semidefinite matrix. The performance of the algorithm, and indeed the identifiability, depends on the choice of the precoding sequence. We propose a method for optimal selection of the precoding sequence which takes into account the effect of additive channel noise and numerical error in covariance matrix estimation. Simulation results are used to demonstrate the performance of the algorithm.**

***Index Terms*—Blind identification, multiple-input mutiple-out (MIMO) channel, periodic precoding, transmitter induced cyclostationarity.**

I. Introduction

BLIND-channel identification is a technique that alleviates the need for training sequences to identify the unknown channel-impulse response from the received signal. Since the requirement of extra bandwidth for training overhead is reduced, this technique has received great research interest and many blind identification algorithms have been proposed (see [1]–[3] for a detailed review).

Blind identification of single-input single-output (SISO) frequency selective channels exploiting transmitter induced cyclostationarity of the second-order statistics of the received data is first proposed in [4], [5]. Since then, various schemes have been proposed to induce cyclostationarity at the transmitter and to blindly identify SISO [6]-[11] and multiple-input mutiple-out (MIMO) channels [12]–[14], [17]. One way to induce cyclostationarity at the transmitter is by periodic precoding, i.e., multiplying the source symbols with a periodic sequence before transmission [5], [8], [9], [11]–[14], [17]. For SISO channels, blind identification methods based on periodic precoding are shown to be robust with respective to channel-order overestimation and impose no restriction on the locations of channel zeros [5], [8], [9], [11].

In the MIMO context, Chevreuil and Loubaton [12] proposes a scheme that multiplies the input sequence by a constant modulus complex exponential precoding sequence to induce conjugate cyclostationarity at the transmitter. The scheme reduces the MIMO channel identification problem to several SIMO ones, which are then solved by the subspace method [18]. Each SIMO channel is required to be free from common zeros and only real symbols can be used. Bölcskei *et al.* [13] proposes a method for identifying each of the scalar channels individually up to a phase ambiguity using non-constant modulus periodic precoding sequences. The method imposes no restriction on channel zeros and is insensitivity to channel-order overestimation. However, no general procedure for the design of the precoding sequences is given. The method is extended to the multicarrier case in [14]. In [17], Ding and Ward regard the precoding sequences applied at the transmitters as a special kind of GSTBCs and they propose a subspace method for identification and equalization. Each element in the precoding sequences is random with modulus 1.

In this paper, we propose a method for blind identification of MIMO finite-impulse response (FIR) channels using periodic precoding as a means to induce cyclostationarity. We show that, by properly choosing the precoding sequence, the MIMO FIR transfer functions, with M_t inputs and M_r outputs, can be identified up to a unitary matrix ambiguity. The transfer functions need not be irreducible or column reduced [15], [16], and there can be more outputs $(M_r \geq M_t)$ or more inputs $(M_r < M_t)$. The method exploits the linear relation between the covariance matrix of the received data and the "channel product matrices". The method is shown to be robust with respect to channel-order overestimation. The proposed algorithm requires solving linear equations and computing the nonzero eigenvalues and eigenvectors of a Hermitian positive semidefinite matrix. The performance of the algorithm, and indeed the identifiability, depends on the choice of the precoding sequence. We propose a method for optimal selection of the precoding sequence which takes into account the effect of additive channel noise and numerical error in covariance matrix estimation. Simulation results are used to demonstrate the performance of the algorithm. The paper generalizes the results for the SISO case discussed in [11].

The paper is organized as follows. Section II is problem statement and formulation. In Section III, we derive the identification method and propose the blind identification algorithm. In Section IV, we discuss optimal selection of the precoding sequence. Simulation results are given in Section V. Section VI concludes the paper.

Notations used in this paper are quite standard. Bold uppercase is used for matrices, and bold lowercase is used for vectors.

Manuscript received January 3, 2006; revised April 25, 2006. This paper has been recommended by Associate Editor P. Regalia.

The authors are with the Department of Electrical and Control Engineering, National Chiao-Tung University, Hsinchu, Taiwan 300, R.O.C. (e-mail: calin@cc.nctu.edu.tw; yschen.ece90g@nctu.edu.tw).

Digital Object Identifier 10.1109/TCSI.2006.888762

Fig. 1. MIMO channel model.

$\mathbf{A}^T$ represents transpose of the matrix $\mathbf{A}$, and $\mathbf{A}^*$ represents conjugate transpose of the matrix $\mathbf{A}$. $\mathbf{A} \otimes \mathbf{B}$ is the Kronecker product of $\mathbf{A}$ and $\mathbf{B}$. $\mathbf{0}_{M\times N}$ is the zero matrix of dimension $M \times N$, and $\mathbf{I}_M$ is the identity matrix of dimension $M \times M$. The symbols $\mathbb{R}$ and $\mathbb{C}$ stand for the set of real number and the set of complex number, respectively.

II. Problem Statement and Formulation

We consider the linear MIMO baseband model of a communication channel with M_t transmitters and M_r receivers shown in Fig. 1, where each source symbol sequence is multiplied by an N-periodic sequence, $p(n)$, before transmission. The transmitted signal is

$$w_j(n) = p(n)s_j(n), \qquad j = 1, 2, \ldots, M_t \tag{2.1}$$

where $p(n+N) = p(n), \forall n$. The discrete time model describing the relation between the transmitted signal $w_j(n)$ and the received signal $x_i(n)$ has the form of an MIMO FIR filter with additive noise

$$x_i(n) = \sum_{j=1}^{M_t}\sum_{l=0}^{L_{ij}} h_{ij}(l)w_j(n-l) + v_i(n), \qquad i = 1, 2, \ldots, M_r \tag{2.2}$$

where $h_{ij}(0), h_{ij}(1), \ldots, h_{ij}(L_{ij})$, are the impulse responses of the channel between the jth transmitter and the ith receiver, and $v_i(n)$ is the channel noise seen at the input of the ith receiver. The (2.1) and (2.2) can be written more compactly as

$$\mathbf{w}(n) = p(n)\mathbf{s}(n); \quad \mathbf{x}(n) = \sum_{l=0}^{L}\mathbf{H}(l)\mathbf{w}(n-l) + \mathbf{v}(n) \tag{2.3}$$

where $\mathbf{w}(n), \mathbf{s}(n) \in \mathbb{C}^{M_t}$, and $\mathbf{x}(n), \mathbf{v}(n) \in \mathbb{C}^{M_r}$ are vector signals formed by stacking the respective scalar signals together, e.g., $\mathbf{x}(n) = [x_1(n)\ x_2(n)\ \ldots\ x_{M_r}(n)]^T$. The ijth element of $\mathbf{H}(l) \in \mathbb{C}^{M_r \times M_t}$ is $h_{ij}(l)$, and $L = \max_{i,j}\{L_{ij}\}$ is the order of the MIMO channel. Thus, $\mathbf{H}(L) \neq \mathbf{0}_{M_r\times M_t}$. We assume that the receivers are synchronized with the transmitters. In addition, the following assumptions are made throughout the paper.

A1) $\mathbf{s}(n)$ and $\mathbf{v}(n)$ are white zero-mean vector sequences, and $\mathbf{s}(n)$ and $\mathbf{v}(n)$ are temporally and spatially uncorrelated. More precisely, $E[\mathbf{s}(k)\mathbf{s}(j)^*] = \delta(k-j)\mathbf{I}_{M_t} \in \mathbb{R}^{M_t\times M_t}, E[\mathbf{v}(k)\mathbf{v}(j)^*] = \delta(k-j)\sigma_v^2\mathbf{I}_{M_r} \in \mathbb{R}^{M_r\times M_r}, E[\mathbf{s}(k)\mathbf{v}(j)^*] = \mathbf{0}_{M_t\times M_r}, \forall k, j$, where $\delta(\cdot)$ is the Kronecker delta function.

A2) An upper bound $\hat{L}$ of the channel order L is known and the period $N > \hat{L} + 1$.

A3) $\operatorname{rank}([\mathbf{H}(0)^T\ \mathbf{H}(1)^T\ \ldots\ \mathbf{H}(L)^T]^T) = M_t$.

Due to periodic precoding, the input–output relation between the source $\mathbf{s}(n)$ and the received signal $\mathbf{x}(n)$, described by (2.3), is periodically time-varying. In order to obtain a time-invariant representation, we consider input–output relation between block input and block output of size N [19]. Define block signal $\bar{\mathbf{x}}(n) = [\mathbf{x}(Nn)^T, \mathbf{x}(Nn+1)^T, \cdots, \mathbf{x}(Nn+N-1)^T]^T \in \mathbb{C}^{M_tN}$, and let $\bar{\mathbf{v}}(n), \bar{\mathbf{w}}(n), \bar{\mathbf{s}}(n)$ be similarly defined. Since $p(n)$ is periodic, $\bar{\mathbf{w}}(n) = \mathbf{G}\bar{\mathbf{s}}(n)$ for all n, where $\mathbf{G} = \operatorname{diag}[p(0)\mathbf{I}_{M_t}, p(1)\mathbf{I}_{M_t}, \ldots, p(N-1)\mathbf{I}_{M_t}] \in \mathbb{R}^{M_tN\times M_tN}$ is a diagonal matrix. In terms of block signals, (2.3) can be written as

$$\begin{aligned}\bar{\mathbf{x}}(n) &= \mathbf{H_0}\bar{\mathbf{w}}(n) + \mathbf{H_1}\bar{\mathbf{w}}(n-1) + \bar{\mathbf{v}}(n)\\ &= \mathbf{H_0}\mathbf{G}\bar{\mathbf{s}}(n) + \mathbf{H_1}\mathbf{G}\bar{\mathbf{s}}(n-1) + \bar{\mathbf{v}}(n)\end{aligned} \tag{2.4}$$

where $\mathbf{H_0}$ is an $M_rN \times M_tN$ block lower triangular Toeplitz matrix with $[\mathbf{H}(0)^T\ \mathbf{H}(1)^T\ \cdots\ \mathbf{H}(L)^T\ \mathbf{0}_{M_r\times M_t}^T\ \cdots\ \mathbf{0}_{M_r\times M_t}^T]^T \in \mathbb{C}^{M_rN\times M_t}$ as its first block column (i.e., the first M_t columns), and $\mathbf{H_1}$ is an $M_rN \times M_tN$ block upper triangular Toeplitz matrix with $[\mathbf{0}_{M_r\times M_t}\ \cdots\ \mathbf{0}_{M_r\times M_t}\ \mathbf{H}(L)\ \mathbf{H}(L-1)\ \cdots\ \mathbf{H}(1)] \in \mathbb{C}^{M_r\times M_tN}$ as its first block row (i.e., the first M_r rows).

The problem we study in this paper is blind identification of the MIMO channel matrix $\mathbf{H} = [\mathbf{H}(0)^T\mathbf{H}(1)^T\ \ldots\mathbf{H}(L)^T]^T$ using second-order statistics of the received data. We define the following operations that will be used in the derivation of the main result. First, for any $m\times m$ matrix $\mathbf{A} = [a_{k,l}]_{0\le k,l\le m-1}$, define $\Gamma_j(\mathbf{A}) = [a_{0,j}\ a_{1,j+1}\ \ldots\ a_{m-1-j,m-1}]^T$ for $0 \le j \le m-1$, i.e., $\Gamma_j(\mathbf{A})$ is the vector formed from the jth superdiagonal of $\mathbf{A}$. Second, for any $M_rn \times M_rn$ matrix $\mathbf{B} = [\mathbf{B}_{k,l}]_{0\le k,l\le n-1}$, where $\mathbf{B}_{k,l}$ is a block matrix of dimension $M_r\times M_r$, define $\Upsilon_j(\mathbf{B}) = [\mathbf{B}_{0,j}^T\ \mathbf{B}_{1,j+1}^T\ \ldots\ \mathbf{B}_{n-1-j,n-1}^T]^T$ for $0 \le j \le n-1$, i.e., $\Upsilon_j(\mathbf{B})$ is the matrix formed from the jth block superdiagonal of $\mathbf{B}$.

III. Channel Identification

We study channel identification in this section. In Section III-A, we derive the proposed method assuming the channel order is known and the noise is absent. We show that by appropriately selecting the periodic precoding sequence, any MIMO channel satisfying $(\mathbf{A3})$ is identifiable up to an $M_t \times M_t$ unitary matrix ambiguity. In Section III-B, we show that the proposed method is robust with respect to channel-order overestimation and we propose an identification algorithm in Section III-C. The effect of noise and optimal selection of the precoding sequence are discussed in Section IV.

A. The Identification Method

When the noise is absent and the channel order L is known, the (2.4) now becomes

$$\bar{\mathbf{x}}(n) = \mathbf{H_0}\mathbf{G}\bar{\mathbf{s}}(n) + \mathbf{H_1}\mathbf{G}\bar{\mathbf{s}}(n-1). \tag{3.1}$$

By assumption $(\mathbf{A1})$, the covariance matrix of $\bar{\mathbf{x}}(n)$ can be written as

$$\mathbf{R}_{\bar{\mathbf{x}}}(0) = E[\bar{\mathbf{x}}(n)\bar{\mathbf{x}}(n)^*] = \mathbf{H_0}\mathbf{G}^2\mathbf{H_0}^* + \mathbf{H_1}\mathbf{G}^2\mathbf{H_1}^*. \tag{3.2}$$

Let $\mathbf{J} \in \mathbb{R}^{N\times N}$ be the matrix whose first sub-diagonal are all one, i.e., $\Gamma_1(\mathbf{J}^T) = [11\cdots 1]^T \in \mathbb{R}^{(N-1)}$, and all remaining entries are zero. The block Toeplitz structures of $\mathbf{H_0}$ and $\mathbf{H_1}$ allow us to write $\mathbf{H_0} = \sum_{k=0}^{L} \mathbf{J}^k \otimes \mathbf{H}(k)$ and $\mathbf{H_1} = \sum_{k=0}^{L} (\mathbf{J}^T)^{N-k} \otimes \mathbf{H}(k)$, respectively. Besides, we define $\mathbf{G_p} = \mathrm{diag}[p(0), p(1), \ldots, p(N-1)] \in \mathbb{R}^{N\times N}$. Hence, $\mathbf{H_0}\mathbf{G}^2\mathbf{H_0^*}$ can be written as

$$\begin{aligned}
&\mathbf{H_0}\mathbf{G}^2\mathbf{H_0^*}\\
&\quad= \sum_{k=0}^{L} \mathbf{J}^k \otimes \mathbf{H}(k)\left(\mathbf{G_p^2} \otimes \mathbf{I}_{M_t}\right) \sum_{l=0}^{L} (\mathbf{J}^l \otimes \mathbf{H}(l))^*\\
&\quad= \sum_{k=0}^{L}\sum_{l=0}^{L} (\mathbf{J}^k \otimes \mathbf{H}(k))(\mathbf{G_p^2} \otimes \mathbf{I}_{M_t})((\mathbf{J}^T)^l \otimes \mathbf{H}(l)^*)\\
&\quad= \sum_{k=0}^{L}\sum_{l=0}^{L} (\mathbf{J}^k\mathbf{G_p^2}(\mathbf{J}^T)^l) \otimes (\mathbf{H}(k)\mathbf{H}(l)^*) \qquad (3.3)
\end{aligned}$$

where we have used the identities $(\mathbf{A}\otimes\mathbf{B})^* = \mathbf{A}^*\otimes\mathbf{B}^*$ and $(\mathbf{A}\otimes\mathbf{B})(\mathbf{C}\otimes\mathbf{D}) = (\mathbf{AC})\otimes(\mathbf{BD})$ [22, , p. 190]. Similarly, $\mathbf{H_1}\mathbf{G}^2\mathbf{H_1^*}$ can be written as

$$\mathbf{H_1}\mathbf{G}^2\mathbf{H_1^*} = \sum_{k=0}^{L}\sum_{l=0}^{L} \left((\mathbf{J}^T)^{N-k}\mathbf{G_p^2}\mathbf{J}^{N-l}\right) \otimes (\mathbf{H}(k)\mathbf{H}(l)^*). \qquad (3.4)$$

The following proposition shows that the matrices $\mathbf{J}^k\mathbf{G_p^2}(\mathbf{J}^T)^l$ and $(\mathbf{J}^T)^{N-k}\mathbf{G_p^2}\mathbf{J}^{N-l}$ have special structures that allow decomposition of (3.2) into a group of decoupled equations. Roughly speaking, the jth block superdiagonal part of (3.2) involves only the unknown "channel product matrices," $\mathbf{H}(k)\mathbf{H}(k+j)^*, k = 0, 1, \ldots, L-j$. For example, the equations corresponding to the diagonal blocks ($j = 0$) involve only $\mathbf{H}(k)\mathbf{H}(k)^*, k = 0, 1, \ldots, L$. In the proposed identification algorithm, these "channel product matrices" are computed first by solving linear equations, and then the channel-impulse response matrices $\mathbf{H}(k)$ are computed via eigenvalue-eigenvector decomposition.

Proposition 3.1: Let $0 \le k, l \le L$ be two non-negative integers. Then, the following are true.

a) For $l = k + j$, where $0 \le j \le L - k$, both $\mathbf{J}^k\mathbf{G_p^2}(\mathbf{J}^T)^l$ and $(\mathbf{J}^T)^{N-k}\mathbf{G_p^2}\mathbf{J}^{N-l}$ are upper triangular matrices with only the respective jth upper diagonals nonzero, as shown in (3.5) and (3.6) at the bottom of the page.

b) For $l < k$, both $\Gamma_j(\mathbf{J}^k\mathbf{G_p^2}(\mathbf{J}^T)^l)$ and $\Gamma_j((\mathbf{J}^T)^{N-k}\mathbf{G_p^2}\mathbf{J}^{N-l})$ are lower triangular with zero diagonal matrices.

Proof: See [11].

It follows from (3.5) and (3.6) that (3.7), shown at the bottom of the page, is true. Since

$$\begin{aligned}
&\Upsilon_j\left(\left(\mathbf{J}^k\mathbf{G_p^2}(\mathbf{J}^T)^l\right) \otimes \mathbf{H}(k)\mathbf{H}(l)^*\right)\\
&\quad= \Gamma_j\left(\mathbf{J}^k\mathbf{G_p^2}(\mathbf{J}^T)^l\right) \otimes \mathbf{H}(k)\mathbf{H}(l)^* \qquad (3.8)\\
&\qquad\text{and}\\
&\Upsilon_j\left(\left((\mathbf{J}^T)^{N-k}\mathbf{G_p^2}\mathbf{J}^{N-l}\right) \otimes \mathbf{H}(k)\mathbf{H}(l)^*\right)\\
&\quad= \Gamma_j\left((\mathbf{J}^T)^{N-k}\mathbf{G_p^2}\mathbf{J}^{N-l}\right) \otimes \mathbf{H}(k)\mathbf{H}(l)^* \qquad (3.9)
\end{aligned}$$

it follows from (3.2)–(3.4) and (3.7)–(3.9) that $\Upsilon_j(\mathbf{R}_{\bar{\mathbf{x}}}(0))$ can be derived as shown in (3.10), at the bottom of the next page.

The right-hand side of (3.10) is a linear combination of block columns with the channel product matrices $\mathbf{H}(k)\mathbf{H}(k+j)^*$ as coefficients. If we define $\mathbf{F}_j,\ 0 \le j \le L$, shown in (3.11), at the bottom of the next page, then (3.10) can be written in a more compact form as

$$\Upsilon_j(\mathbf{R}_{\bar{\mathbf{x}}}(0)) = \mathbf{M}_j\mathbf{F}_j \qquad \forall 0 \le j \le L \qquad (3.12)$$

where $\mathbf{M}_j \in \mathbb{R}^{M_r(N-j)\times M_r(L-j+1)}$ is defined as (3.13), shown at the bottom of the next page. We note that $\mathbf{M}_j, 1 \le j \le L$, is obtained from $\mathbf{M}_0$ by deleting its last jM_r rows and last jM_r columns.

$$\begin{aligned}
&\Gamma_j\left(\mathbf{J}^k\mathbf{G_p^2}(\mathbf{J}^T)^l\right)\\
&\quad= [\underbrace{0\ \cdots\ 0}_{k \text{ entries}}\underbrace{p(0)^2\ \ p(1)^2\ \ \cdots\ \ p(N-1-k-j)^2}_{N-k-j \text{ entries}}]^T \qquad (3.5)\\
&\Gamma_j\left((\mathbf{J}^T)^{N-k}\mathbf{G_p^2}\mathbf{J}^{N-l}\right)\\
&\quad= [\underbrace{p(N-k)^2\ \ \cdots\ \ p(N-1)^2}_{k \text{ entries}}\ \underbrace{0\ \cdots\ 0}_{N-k-j \text{ entries}}]^T. \qquad (3.6)
\end{aligned}$$

$$\begin{aligned}
&\Gamma_j\left(\mathbf{J}^k\mathbf{G_p^2}(\mathbf{J}^T)^l\right) + \Gamma_j\left((\mathbf{J}^T)^{N-k}\mathbf{G_p^2}\mathbf{J}^{N-l}\right)\\
&\qquad= \begin{cases} [\underbrace{p(N-k)^2\ \ \cdots\ \ p(N-1)^2}_{k \text{ entries}}\underbrace{p(0)^2\ \ \cdots\ \ p(N-1-k-j)^2}_{N-k-j \text{ entries}}]^T & \text{if } j = l-k \ge 0\\ \mathbf{0}_{(N-j)\times 1} & \text{if } j \ne l-k \end{cases} \qquad (3.7)
\end{aligned}$$

Since $N > L+1$, the $(L+1)$ equations in (3.12) are overdetermined and for the noise-free case, these equations are consistent. If $\mathbf{M}_j$ is full column rank, then the solution can be obtained as

$$\mathbf{F}_j = \left(\mathbf{M}_j^T\mathbf{M}_j\right)^{-1}\mathbf{M}_j^T\Upsilon_j\left(\mathbf{R}_{\bar{\mathbf{x}}}(0)\right). \tag{3.14}$$

If $\mathbf{F}_j, 0 \le j \le L$, are computed from (3.14), then we have the channel product matrices $\mathbf{H}(k)\mathbf{H}(l)^*$ for $0 \le k \le l \le L$. We now consider the computation required to determine the channel-impulse response matrices $\mathbf{H}(0), \mathbf{H}(1), \ldots, \mathbf{H}(L)$ from $\mathbf{F}_j$.

Let $\mathbf{Q}$ be the Hermitian matrix defined by $\Upsilon_j(\mathbf{Q}) = \mathbf{F}_j$ for $j = 0, 1, \ldots, L$, and let the channel matrix $\mathbf{H} = [\mathbf{H}(0)^T \ \mathbf{H}(1)^T \ \ldots \ \mathbf{H}(L)^T]^T$. Clearly we have

$$\mathbf{Q} = \mathbf{H}\mathbf{H}^*. \tag{3.15}$$

Since $\mathrm{rank}(\mathbf{H}) = M_t$ by assumption $(\mathbf{A3})$, $\mathbf{Q}$ has rank M_t. Since $\mathbf{Q}$ is Hermitian and positive semidefinite, $\mathbf{Q}$ has M_t positive eigenvalues, say, $\lambda_1, \ldots, \lambda_{M_t}$. We can expand $\mathbf{Q}$ as

$$\mathbf{Q} = \sum_{j=1}^{M_t}(\sqrt{\lambda_j}\mathbf{d}_j)(\sqrt{\lambda_j}\mathbf{d}_j)^* \tag{3.16}$$

where $\mathbf{d}_j$ is a unit norm eigenvector of $\mathbf{Q}$ associated with $\lambda_j > 0$. We can thus choose the channel matrix to be

$$\begin{aligned}\hat{\mathbf{H}} &= [\sqrt{\lambda_1}\mathbf{d}_1 \quad \sqrt{\lambda_2}\mathbf{d}_2 \quad \cdots \quad \sqrt{\lambda_{M_t}}\mathbf{d}_{M_t}] \\ &\in \mathbb{C}^{M_r(L+1)\times M_t}.\end{aligned} \tag{3.17}$$

We note $\mathbf{H}$ can only be identified up to a unitary matrix ambiguity $\mathbf{U} \in \mathbb{C}^{M_t\times M_t}$ [15], [16], i.e., $\hat{\mathbf{H}} = \mathbf{H}\mathbf{U}$, since $\hat{\mathbf{H}}\hat{\mathbf{H}}^* = \mathbf{H}\mathbf{H}^* = \mathbf{Q}$. The ambiguity matrix $\mathbf{U}$ is intrinsic to methods for blind identification of multiple input systems using only second-order statistics [15], [16].

We note that the matrix $\mathbf{M}_j, j = 0, 1, \ldots, L$, is completely determined by the precoding sequence. By appropriately selecting the precoding sequence, we can make each $\mathbf{M}_j$ full column rank.

We summarize what we have so far.

a) If the MIMO channel described by (2.3) satisfies $(\mathbf{A1})$ and $(\mathbf{A3})$ and the channel order L is known, then the channel matrix $\mathbf{H}$ can be identified up to a unitary matrix ambiguity.

b) The proposed identification method use the covariance matrix of the received signal $\mathbf{R}_{\bar{\mathbf{x}}}(0)$ as data, and the computations involved are solving linear (3.12) and performing eigenvalue-eigenvector decomposition of the Hermitian matrix $\mathbf{Q}$ in (3.16).

We note that in the proposed method, the channel matrix $\mathbf{H}$ is only assumed to be full column rank $(\mathbf{A3})$. Hence, the channel

$$\begin{aligned}\Upsilon_j\left(\mathbf{R}_{\bar{\mathbf{x}}}(0)\right) &= \Upsilon_j\left(\mathbf{H_0}\mathbf{G}^2\mathbf{H_0^*} + \mathbf{H_1}\mathbf{G}^2\mathbf{H_1^*}\right) \\ &= \sum_{k=0}^{L}\sum_{l=0}^{L}\Upsilon_j\left(\left(\mathbf{J}^k\mathbf{G}_\mathbf{p}^2(\mathbf{J}^T)^l\right)\otimes(\mathbf{H}(k)\mathbf{H}(l)^*)\right) + \Upsilon_j\left(\left((\mathbf{J}^T)^{N-k}\mathbf{G}_\mathbf{p}^2\mathbf{J}^{N-l}\right)\otimes(\mathbf{H}(k)\mathbf{H}(l)^*)\right) \\ &= \sum_{k=0}^{L}\sum_{l=0}^{L}\left\{\Gamma_j\left(\mathbf{J}^k\mathbf{G}_\mathbf{p}^2(\mathbf{J}^T)^l\right) + \Gamma_j\left((\mathbf{J}^T)^{N-k}\mathbf{G}_\mathbf{p}^2\mathbf{J}^{N-l}\right)\right\}\otimes\mathbf{H}(k)\mathbf{H}(l)^* \\ &= \sum_{k=0}^{L-j}[p(N-k)^2 \quad \cdots \quad p(N-1)^2 \quad p(0)^2 \quad \cdots \quad p(N-1-k-j)^2]^T\otimes\mathbf{H}(k)\mathbf{H}(k+j)^* \\ &= \sum_{k=0}^{L-j}[p(N-k)^2\mathbf{I}_{M_r} \quad \cdots \quad p(N-1)^2\mathbf{I}_{M_r} \quad p(0)^2\mathbf{I}_{M_r} \quad \cdots \quad p(N-1-k-j)^2\mathbf{I}_{M_r}]^T\mathbf{H}(k)\mathbf{H}(k+j)^*\end{aligned} \tag{3.10}$$

$$\mathbf{F}_j = [(\mathbf{H}(0)\mathbf{H}(j)^*)^T \quad (\mathbf{H}(1)\mathbf{H}(j+1)^*)^T \quad \cdots \quad (\mathbf{H}(L-j)\mathbf{H}(L)^*)^T]^T \in \mathbb{C}^{M_r(L-j+1)\times M_r} \tag{3.11}$$

$$\mathbf{M}_j = \begin{bmatrix} p(0)^2 & p(N-1)^2 & \cdots & p(N-L+j)^2 \\ p(1)^2 & p(0)^2 & \cdots & p(N-L+j+1)^2 \\ \vdots & \vdots & \vdots & \vdots \\ p(N-2-j)^2 & p(N-3-j)^2 & \cdots & p(N-L-2)^2 \\ p(N-1-j)^2 & p(N-2-j)^2 & \cdots & p(N-L-1)^2 \end{bmatrix} \otimes \mathbf{I}_{M_r} \tag{3.13}$$

needs not be irreducible or column reduced. If $M_r \geq M_t$ (more outputs), then $(\mathbf{A3})$ is generically satisfied [20, ch. 7]. If $M_t > M_r$ (more inputs), then $(\mathbf{A3})$ is generically satisfied provided $(L+1)M_r \geq M_t$. We note that if the channel has more inputs than outputs, channel equalization and source separation may be difficult even if accurate channel estimate is available.

B. Channel-Order Overestimation

So far, we have assumed that the channel order L is known. If only an upper bound $\hat{L} \geq L$ is available with $N > \hat{L}+1$, then following the same process given in Section III-A, the corresponding $M_r(\hat{L}+1)\times M_r(\hat{L}+1)$ matrix $\mathbf{Q}$ can be similarly constructed as in (3.15). The last $(\hat{L}-L)$ block columns (i.e., $(\hat{L}-L)M_r$ columns) of $\mathbf{Q}$ are zero, so are its last $(\hat{L}-L)$ block rows. Hence, again, $\mathbf{Q}$ is of rank M_t and has M_t positive eigenvalues with the associated eigenvectors all of the form $\hat{\mathbf{d}} = [\mathbf{d}^T\ 0\ \ldots\ 0]^T \in \mathbb{C}^{M_r(\hat{L}+1)}$ where $\mathbf{d} \in \mathbb{C}^{M_r(L+1)}$. Thus, we can determine impulse response matrices, up to a unitary matrix ambiguity, from the M_t eigenvectors associated with the M_t positive eigenvalues of $\mathbf{Q}$. In the noise-free case, we can, in theory, also determine the actual channel order.

C. Identification Algorithm

We summarize the proposed method as the following algorithm.

1) Select the precoding sequence $p(n)$ such that each matrix $\mathbf{M}_j$ defined in (3.13) is full column rank.
2) Estimate the covariance matrix $\mathbf{R}_{\bar{\mathbf{x}}}(0)$ via the time average

$$\hat{\mathbf{R}}_{\bar{x}}(0) = \frac{1}{K}\sum_{i=1}^{K} \bar{\mathbf{x}}(i)\bar{\mathbf{x}}(i)^* \tag{3.18}$$

   where K is the number of data block (i.e., KN is the number of samples for each transmitter).
3) Compute $\mathbf{F}_j$, formed by the channel product matrices, for $j = 0, 1, \ldots, L$, using (3.14).
4) Form the matrix $\mathbf{Q}$ as in (3.15), and obtain the channel-impulse response (3.17) by computing the M_t largest eigenvalues and the associated eigenvectors of $\mathbf{Q}$.

IV. Optimal Selection of Precoding Sequence

In Section III, we see that in order to identify the channel, the precoding sequence must be selected so that the resulting matrix $\mathbf{M}_j$ is full column rank such that $\mathbf{F}_j$ can be exactly solved as (3.14). However, when noise is present, the covariance matrix $\hat{\mathbf{R}}_{\bar{x}}(0)$ contains the contribution of noise and numerical error is present in the estimation of $\hat{\mathbf{R}}_{\bar{x}}(0)$ by (3.18). This implies that (3.12) usually has no solution and (3.14) becomes a least squares approximate solution. The choice of $\mathbf{M}_j$ will affect error in the computation of $\mathbf{F}_j$ since different $\mathbf{M}_j^T\mathbf{M}_j$ in (3.14) usually have different condition numbers. In this section, we discuss the optimal selection of the precoding sequence, which takes into account the effect of noise and numerical error in estimating $\hat{\mathbf{R}}_{\bar{x}}(0)$, so as to increase the accuracy of $\mathbf{F}_j$ and thus reduce the channel estimation error.

A. Optimality Criterion

Now we consider the general case that noise is present and discuss the design of the precoding sequence $p(n)$. From (2.4) and assumption $(\mathbf{A1})$, the covariance matrix of the received signal is

$$\mathbf{R}_{\bar{\mathbf{x}}}(0) = \mathbf{H}_0\mathbf{G}^2\mathbf{H}_0^* + \mathbf{H}_1\mathbf{G}^2\mathbf{H}_1^* + \sigma_v^2\mathbf{I}_{M_r}\otimes\mathbf{I}_N. \tag{4.1}$$

From (4.1) and (3.2), we see that noise has only contribution to the diagonal entries of $\mathbf{R}_{\bar{\mathbf{x}}}(0)$. Therefore, the $(L+1)$ decoupled groups of equations in (3.12) remain unchanged, except for the $j = 0$ group, which becomes

$$\begin{aligned}\Upsilon_0(\mathbf{R}_{\bar{\mathbf{x}}}(0)) &= \Upsilon_0(\mathbf{H}_0\mathbf{G}^2\mathbf{H}_0^* + \mathbf{H}_1\mathbf{G}^2\mathbf{H}_1^*) \\ &\quad + \sigma_v^2\Upsilon_0\left(\mathbf{I}_{M_r}\otimes\mathbf{I}_N\right) \\ &= \mathbf{M}_0\mathbf{F}_0 + \mathbf{Y}\end{aligned} \tag{4.2}$$

where $\mathbf{Y} = \sigma_v^2[\mathbf{I}_{M_r}\ \mathbf{I}_{M_r}\ \ldots\ \mathbf{I}_{M_r}]^T \in \mathbb{R}^{M_rN\times M_r}$. Thus, from (3.14), $\hat{\mathbf{F}}_0$, the least squares approximation of $\mathbf{F}_0$, can be written by

$$\begin{aligned}\hat{\mathbf{F}}_0 &= (\mathbf{M}_0^T\mathbf{M}_0)^{-1}\mathbf{M}_0^T\underbrace{(\mathbf{M}_0\mathbf{F}_0 + \mathbf{Y})}_{\Upsilon_0(\mathbf{R}_{\bar{\mathbf{x}}}(0))} \\ &= \mathbf{F}_0 + (\mathbf{M}_0^T\mathbf{M}_0)^{-1}\mathbf{M}_0^T\mathbf{Y} = \mathbf{F}_0 + \mathbf{Z}\end{aligned} \tag{4.3}$$

which is $\mathbf{F}_0$ plus a perturbation term due to noise. The perturbation term $\mathbf{Z}$ is the least squares solution of the equation $\mathbf{M}_0\mathbf{Z} = \mathbf{Y}$. We note that if every column of $\mathbf{Y}$ is orthogonal to every column of $\mathbf{M}_0$, then $\mathbf{Z} = \mathbf{0}$, which implies $\hat{\mathbf{F}}_0 = \mathbf{F}_0$. But that is impossible since the entries of $\mathbf{M}_0$ are positive and those of $\mathbf{Y}$ are nonnegative. Therefore, we seek to appropriately choose the precoding sequence $p(n)$ such that every column of $\mathbf{Y}$ is as close to being orthogonal to that of $\mathbf{M}_0$ as possible. To this end, we first define $\mathbf{q}_{ki}$ and $\mathbf{y}_i$ shown below as the columns of $\mathbf{M}_0$ and $\mathbf{Y}$, respectively, in (4.4) and (4.5), at the bottom of the next page. Then, due to the special structure of the block matrix $\mathbf{M}_0$ and $\mathbf{Y}$, it is easy to check that $\mathbf{q}_{ki}$ is orthogonal to $\mathbf{y}_j$, i.e., $\mathbf{q}_{ki}^T\mathbf{y}_j = 0$ for $j \neq i$, (e.g., $\mathbf{q}_{01}^T\mathbf{y}_2 = 0$ shown in

$$\mathbf{M}_0 = \left[\underbrace{\mathbf{q}_{01}\quad \mathbf{q}_{02}\quad \cdots\quad \mathbf{q}_{0M_r}}_{\mathbf{M}_0(:,1:M_r)}\quad \cdots\quad \underbrace{\mathbf{q}_{L1}\quad \mathbf{q}_{L2}\quad \cdots\quad \mathbf{q}_{LM_r}}_{\mathbf{M}_0(:,LM_r+1:(L+1)M_r)}\right] \tag{4.4}$$

$$\mathbf{Y} = \sigma_v^2\left[\mathbf{I}_{M_r}\quad \mathbf{I}_{M_r}\quad \cdots\quad \mathbf{I}_{M_r}\right]^T = [\mathbf{y}_1\quad \mathbf{y}_2\quad \cdots\quad \mathbf{y}_{M_r}] \tag{4.5}$$

the second equation at the bottom of this page), and each $\mathbf{q}_{ki}^T\mathbf{y}_i$ assumes the same value $\sigma_v^2\sum_{n=0}^{N-1}p(n)^2$, for $k=0,1,\ldots,L$, $i=1,2,\ldots,M_r$, (e.g., $\mathbf{q}_{01}^T\mathbf{y}_1=\sigma_v^2\sum_{n=0}^{N-1}p(n)^2$ shown in the last equation at the bottom of the this page). Thus, we only need to consider the relation between columns of $\mathbf{q}_{01}$ and $\mathbf{y}_1$ (the case of $k=0$ and $i=1$). Define the correlation coefficient

$$\gamma=\frac{\mathbf{q}_{01}^T\mathbf{y}_1}{\|\mathbf{q}_{01}\|_2\|\mathbf{y}_1\|_2}. \tag{4.6}$$

Since γ is nonnegative and by Cauchy–Schwarz inequality, $0\le\gamma\le 1$. In order to make the perturbation term $\mathbf{Z}$ small, we choose $\mathbf{q}_{01}$ so that the correlation coefficient γ is as small as possible. Based on this point of view, we formulate the optimal selection problem as minimizing γ subject to

$$\frac{1}{N}\sum_{n=0}^{N-1}|p(n)|^2=1 \tag{4.7}$$

$$|p(n)|^2\ge\tau>0 \qquad \forall 0\le n\le N-1. \tag{4.8}$$

Roughly, constraint (4.7) normalizes the power gain of the precoding sequence of each transmitter to 1; constraint (4.8) requires that at each instant, the power gain is no less than τ. Note that the problem of selecting the precoding sequence is identical to the SISO case considered in [11]. Thus, the optimal precoding sequence $p(n)$ is a two-level sequence with a single peak in one period [11]. More specifically, for each $m, 0\le m\le N-1$

$$p(n)=\begin{cases}\sqrt{N(1-\tau)+\tau}, & n=m\\ \sqrt{\tau}, & n\ne m, \qquad 0\le n\le N-1\end{cases} \tag{4.9}$$

is an optimal precoding sequence. Because the precoding sequence is periodic with period N, the single peak can be placed at any one of the N positions which yield the same $\gamma=(1/\sqrt{N(1-\tau)^2+\tau(2-\tau)})$. Note that γ decreases as τ decreases, which implies that the noise effect in the estimation of covariance matrix $\mathbf{R}_{\bar{\mathbf{x}}}(0)$ is minimized and thus identification performance improves. However the peak location m does significantly affect the numerical condition of the linear (3.12). We discuss the selection of m next.

B. On Selection of m

We now consider the selection of m. We know that different choices of m result in different matrix $\mathbf{M}_j$ and affect the numerical computation of $\mathbf{F}_j, j=1,2,\ldots,L$, in (3.14) and $\hat{\mathbf{F}}_0$ in (4.3), since different $\mathbf{M}_j^T\mathbf{M}_j$ may have different condition number. If the condition number is large, then the matrix $\mathbf{M}_j^T\mathbf{M}_j$ is ill-conditioned and the computations in (3.14) and (4.3) are sensitive to data error. Let

$$\mu=\max_{0\le j\le L}\kappa\left(\mathbf{M}_j^T\mathbf{M}_j\right) \tag{4.10}$$

where $\kappa(\mathbf{A})$ is the condition number of $\mathbf{A}$. Our goal is to choose m so as to minimize the largest condition number of the corresponding matrices $\mathbf{M}_j^T\mathbf{M}_j, j=0,1,\ldots,L$. Since the peak appears at one of the N possible positions in the periodic precoding sequence, there are N precoding sequences which may result in N different μ. The following result shows that some choices of m are to be avoided since they result in some $\mathbf{M}_j$ being rank deficient and thus $\mu=\infty$.

Proposition 4.1: At least one $\mathbf{M}_j, 0\le j\le L$, is not full column rank if and only if $N-L+1\le m\le N-2$.

Proof: See Appendix A.

Hence if we choose, either $0\le m\le N-L$ or $m=N-1$, then each $\mathbf{M}_j$ is full column rank and the channel is identifiable. The following result shows that we can classify the remaining choices into 2 groups that are relevant to the optimal choice of m.

Proposition 4.2:

a) Each of the $(N-L)$ choices, $m=0, m=1,\ldots,m=N-L-1$, results in the same μ denoted by μ_1.
b) The two choices $m=N-L$ and $m=N-1$ result in the same μ denoted by μ_2. Also $\mu_2\ge\mu_1$.

Proof: See Appendix A.

From Proposition 4.2, we know if $\mu_2>\mu_1$, then we choose case (a); if $\mu_2=\mu_1$, we proceed to compare the second largest condition numbers of the set of matrices $\{\mathbf{M}_j^T\mathbf{M}_j\}_{j=0}^L$ for these two cases and choose the case whose value is smaller. If they are again equal, the same procedure can be done by comparing the third largest condition numbers and so on. Moreover, for $0\le m\le N-L-1$ [case (a)], since the condition numbers of $\mathbf{M}_j^T\mathbf{M}_j$ are the same for each fixed $j, j=0,1,\ldots,L$, (see Appendix A), we can use $m=0$ to represent case (a). Similarly, $m=N-1$ can be used to represent case (b). Hence, the optimal selection of m reduces to one of two cases: $m=0$ or $m=N-1$. In other words, the optimal precoding sequence has a peak either at the beginning or at the end.

$$\mathbf{q}_{01}^T\mathbf{y}_2=[\underbrace{p(0)^2 0 \ \cdots \ 0}_{M_r \text{ entries}} \ \cdots \ \underbrace{p(N-1)^2 \ 0 \ \cdots \ 0}_{M_r \text{ entries}}][\underbrace{0 \ \sigma_v^2 \ 0 \ \cdots \ 0}_{M_r \text{ entries}} \ \cdots \ \underbrace{0 \ \sigma_v^2 \ 0 \ \cdots \ 0}_{M_r \text{ entries}}]^T=0$$

$$\mathbf{q}_{01}^T\mathbf{y}_1=[\underbrace{p(0)^2 \ 0 \ \cdots \ 0}_{M_r \text{ entries}} \ \cdots \ \underbrace{p(N-1)^2 \ 0 \ \cdots \ 0}_{M_r \text{ entries}}][\underbrace{\sigma_v^2 \ 0 \ \cdots \ 0}_{M_r \text{ entries}} \ \cdots \ \underbrace{\sigma_v^2 \ 0 \ \cdots \ 0}_{M_r \text{ entries}}]^T=\sigma_v^2\sum_{n=0}^{N-1}p(n)^2$$

V. Simulation Results

In this section, we use several examples to demonstrate the performance of the proposed method. The channel normalized root-mean-square error (NRMSE) is defined as

$$\mathrm{NRMSE} = \frac{1}{\|\mathbf{H}\|_F}\sqrt{\frac{1}{I}\sum_{i=1}^{I}\left\|\hat{\mathbf{H}}^{(i)} - \mathbf{H}\right\|_F^2} \tag{5.1}$$

where $\|\cdot\|_F$ denotes the Frobenius norm. $\hat{\mathbf{H}}^{(i)} = [\hat{\mathbf{H}}^{(i)}(0)^T \;\; \hat{\mathbf{H}}^{(i)}(1)^T \;\; \cdots \;\; \hat{\mathbf{H}}^{(i)}(L)^T]^T$ is the estimate of channel-impulse response matrix $\mathbf{H}$ after removing the unitary matrix ambiguity by the least squares method [16]. $I = 100$ is the number of Monte Carlo runs. The input source symbols are i.i.d. quadrature phase shift key (QPSK) signals. The channel noises are white Gaussian. The signal-to-noise ratio (SNR) at the output is defined as $\mathrm{SNR} = ((1/N)\sum_{n=0}^{N-1} E[\|\mathbf{z}(n)\|_2^2]/E[\|\mathbf{v}(n)\|_2^2])$, where $\mathbf{z}(n) = [z_1(n) \;\; \cdots \;\; z_{M_r}(n)]^T$ is the signal component of the received signal (see Fig. 1).

1) Simulation 1—Optimal Selection of Precoding Sequences: In this simulation, we use the following model

$$\mathbf{H}(z) = \underbrace{\begin{bmatrix} 1.34-0.55i & 1.67+0.12i \\ -0.69+0.25i & -0.51-0.33i \end{bmatrix}}_{\mathbf{H}(0)} + \underbrace{\begin{bmatrix} -1.45+0.21i & -1.35+0.21i \\ 0.62-0.31i & -0.76+0.43i \end{bmatrix}}_{\mathbf{H}(1)} z^{-1} + \underbrace{\begin{bmatrix} -0.31+0.15i & -0.41-0.16i \\ -0.29+0.21i & -0.25-0.14i \end{bmatrix}}_{\mathbf{H}(2)} z^{-2} \tag{5.2}$$

to demonstrate the effect of different precoding sequences on the performance of the proposed method. In experiment 1, the first sequence is chosen as $\{0.767\ 1.07\ 1.07\ 1.07\}$, which satisfies (4.7) and (4.8). The second and third sequences are chosen based on (4.9) for $N = 4$ and $\tau = 0.5878$ with the two possible peak positions: $m = 0$ and $m = 3$. By computation, the corresponding μ for the three cases are 40.0, 4.66, and 22.1, respectively. Thus, $m = 0$ is the optimal selection. Fig. 2 shows that for $\mathrm{SNR} = 10$ dB, there are about 5–7 dB and 5–9 dB difference in NRMSE between the optimal one and two others.

In experiment 2, we use the precoding sequences that satisfy (4.9) with $m = 0$, but with different τ to test the effect of τ on the identification performance. Fig. 3 shows that for each sequence, when the number of samples (for each transmitter) is fixed at 1000, the NRMSE decreases as SNR increases and is roughly constant for $\mathrm{SNR} \geq 20$ dB. A possible explanation is that for sufficiently large SNR, the NRMSE is contributed mainly by numerical error rather than by channel noise. Fig. 3 also shows that the identification performs better for smaller τ, which is consistent with the conclusion at the end of Section IV-A.

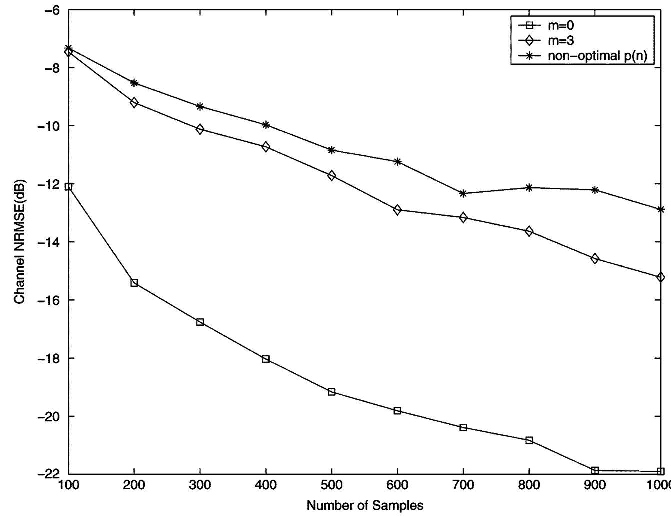


Fig. 2. Channel NRMSE versus number of samples.

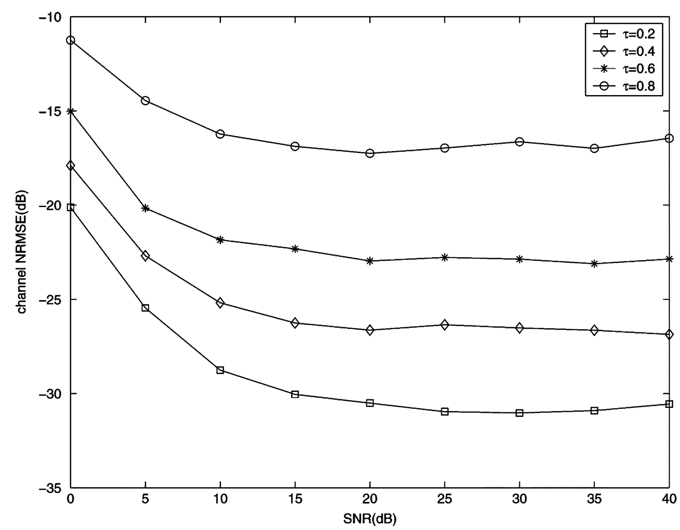


Fig. 3. Channel NRMSE versus output SNR.

2) Simulation 2—Robustness to Channel-Order Overestimation: In this simulation, we use the following channel model:

$$\mathbf{H}(z) = \underbrace{\begin{bmatrix} 0.4851 & 0.3200 \\ -0.3676 & 0.2182 \end{bmatrix}}_{\mathbf{H}(0)} + \underbrace{\begin{bmatrix} -0.4851 & 0.9387 \\ 0.8823 & 0.8729 \end{bmatrix}}_{\mathbf{H}(1)} z^{-1} + \underbrace{\begin{bmatrix} 0.7276 & -0.1280 \\ 0.2941 & -0.4364 \end{bmatrix}}_{\mathbf{H}(2)} z^{-2} \tag{5.3}$$

given in [14]. For each upper bound $\hat{L}, 0 \leq (\hat{L} - L) \leq 6$, we choose $N = \hat{L} + 2, \mathrm{SNR} = 10$ dB, and 1000 samples (for each transmitter) for simulation. The precoding sequences are chosen as (4.9) with $m = 0$ and $\tau = 0.2, 0.4, 0.6,$ and 0.8. Fig. 4 shows the NRMSE increases with increasing channel-order overestimation. We see the proposed method is quite robust to channel-order overestimation when τ is small. For example, with $\tau = 0.4$, when $(\hat{L} - L)$ increases from 0 to 3, the

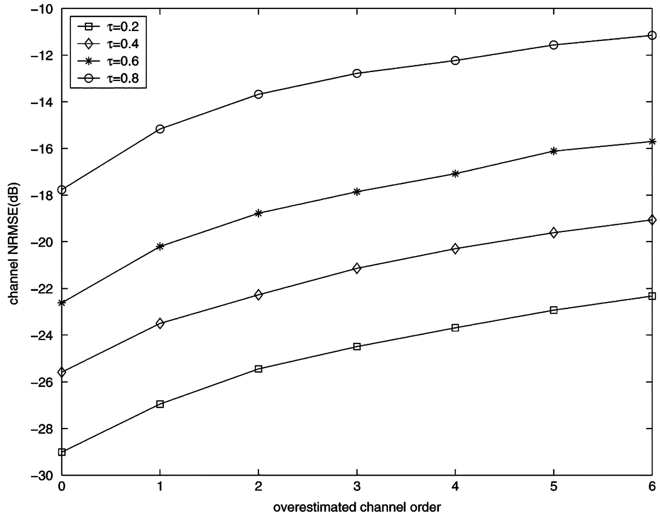


Fig. 4. Channel NRMSE versus $(\hat{L} - L)$.

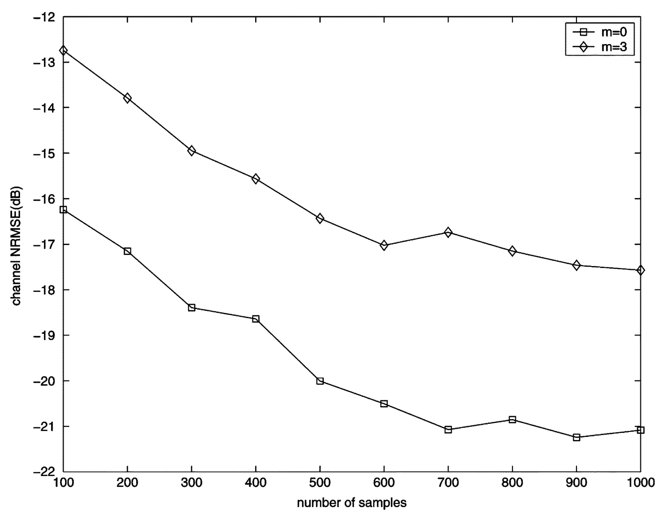


Fig. 5. 3-input 2-output model: Channel NRMSE versus number of samples.

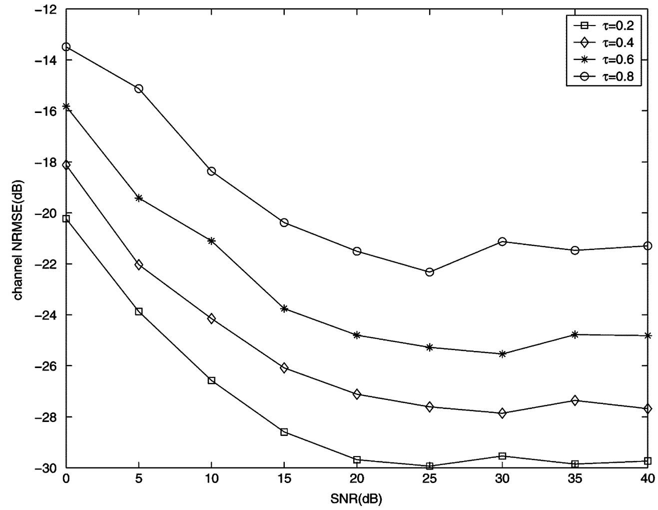


Fig. 6. 3-input 2-output model: Channel NRMSE versus output SNR.

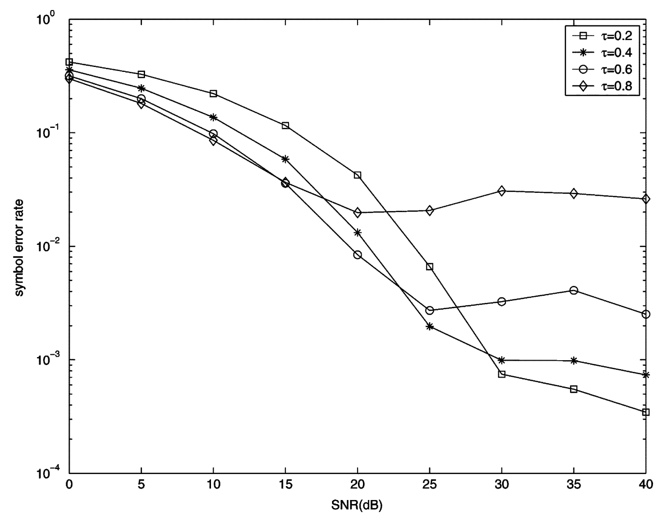


Fig. 7. Symbol error rate versus output SNR.

NRMSE increases from -25.5 dB to -21 dB, which is still a low value.

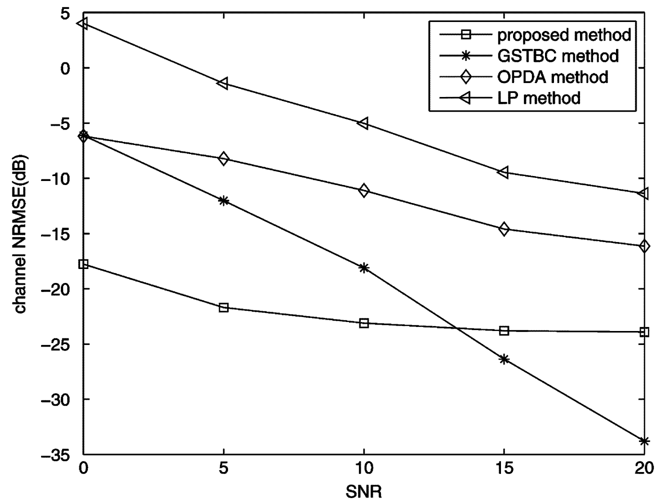
3) Simulation 3—A 3-Input 2-Output Channel: In this simulation, we use the 3-input 2-output model

$$\mathbf{H}(z) = \underbrace{\begin{bmatrix} 1.6 & 0.88 & 0.66 \\ 0.8 & 0.44 & 0.33 \end{bmatrix}}_{\mathbf{H}(0)} + \underbrace{\begin{bmatrix} -0.44 & 0.35 & 0.14 \\ -0.14 & 0.37 & 0.23 \end{bmatrix}}_{\mathbf{H}(1)} z^{-1} + \underbrace{\begin{bmatrix} 0.13 & 0.01 & 0.08 \\ 0.26 & 0.02 & 0.16 \end{bmatrix}}_{\mathbf{H}(2)} z^{-2} \quad (5.4)$$

to illustrate the performance of the proposed method for channel with more inputs than outputs. Note that $\mathbf{H}$ is full column rank, but the channel is not irreducible [16] because $\mathbf{H}(0)$ is not full rank, and it is not column reduced [16] either because $\mathbf{H}(2)$ is not full rank. In experiment 1, the precoding sequences $(N = 4)$ are given as in (4.9) with $m = 0$ and $m = 3$, respectively. Fig. 5 shows that the NRMSE decreases as the number of data samples increases for $\mathrm{SNR} = 10$ dB. As expected, $m = 0$ case (the optimal selection) is better than $m = 3$ case.

In experiment 2, we use the precoding sequences that satisfy (4.9) with $m = 0$, but with different τ to test the effect of τ on the identification performance. Fig. 6 shows that for each sequence, when the number of samples (for each transmitter) is fixed at 1000, the NRMSE decreases as SNR increases and is roughly constant for $\mathrm{SNR} \geq 25$ dB due to numerical error. Fig. 6 also shows the identification performs better for smaller τ.

4) Simulation 4—Channel Equalization Performance: In this simulation, we use the channel model given in (5.3) to demonstrate the performance of the proposed method for channel equalization. We use the precoding sequences that satisfy (4.9) with $m = 0$, but with different τ to test the effect of τ on the equalization performance. For simplicity, we use the minimum mean square error (MMSE) equalizer. The equalizer is a 17-tap Wiener filter with 12-tap reconstruction delay whose jth output $\hat{w}_j(k)$ is an estimate of $w_j(k)$ for $j = 1, 2, \ldots, M_t$. Since the precoding scheme is applied at the transmitter, we need to multiply $\hat{w}_j(k)$ by the corresponding $p(k)^{-1}$ to obtain an estimate of $s_j(k)$ for $j = 1, 2, \ldots, M_t$. The number of



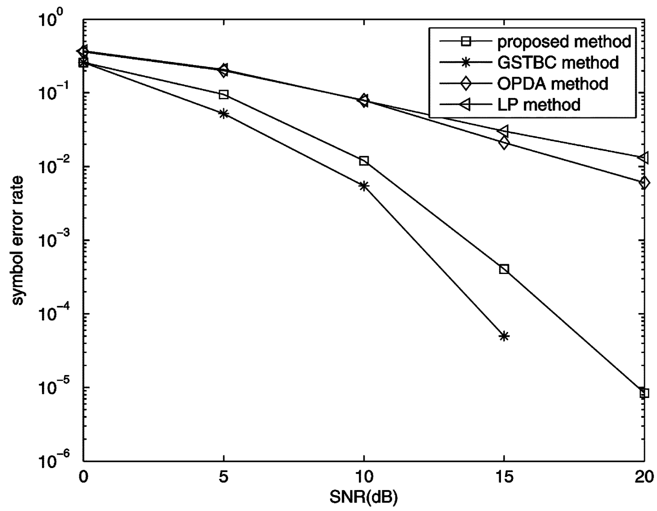


Fig. 8. Comparison of NRMSE and symbol error rate, number of input samples = 1200. (a) Channel NRMSE versus output SNR. (b) Symbol error rate versus output SNR.

samples is 1200. We first identify the channel using the first 400 samples and then do equalization. To obtain smoother curves, we use $I = 300$ as the number of Monte Carlo runs rather than 100.

Fig. 7 shows that under low SNR, the proposed method performs better when τ is large; however, under high SNR, the proposed method performs better when τ is low. A possible explanation is as follows.

Channel estimates become more accurate as τ becomes smaller, but the gains $p(k)^{-1} = (1/\sqrt{\tau}), k = 1, 2, \ldots, N-1$ become larger and result in larger noise amplification at the receiver. Both channel-estimation error and channel noise contribute to the (maximum likelihood) detection performance, i.e., the symbol error rate. In the low SNR region, the detrimental effect of noise amplification outweighs the benefit of small estimation error; whereas in the high SNR region, accurate channel estimation weighs more than the noise amplification effect. Hence, we choose a small τ when SNR is high and a large τ when SNR is low.

5) Simulation 5—Comparisons With Other Methods: In this simulation, we generate 100 2-input 4-output random channels with order $L = 2$; each element in the channel-impulse response matrix is a complex circular Gaussian random variable with unit variance. We compare the proposed method with a generalized space time block codes (GSTBCs) [17] based method. Both methods require periodic precoding sequences. For the proposed method, the precoding sequence is chosen as $\{1.500 \;\; 0.767 \;\; 0.767 \;\; 0.767\}$; whereas the entries in the precoding sequence for the GSTBC method is chosen as random entries with modulus 1 for each random channel simulation [17]. The performance of the proposed method is also compared with a linear prediction (LP) ([3], chap. 6) based method, and an outer product decomposition algorithm (OPDA) [15]. Both methods do not require a periodic precoder. MMSE equalizers are used for the proposed method, LP method, and OPDA method. For the GSTBC method, we use the customized equalizer proposed in [17]. Fig. 8(a) shows that when the number of samples is 1200 (for each transmitter), the identification performance of the proposed method is better than those of the other three methods excepting the GSTBC method for $\mathrm{SNR} \geq 13$ dB. However, Fig. 8(b) shows the equalization performance of the proposed method is only better than those of the LP and OPDA methods and worse than the GSTBC method. The inconsistency of the channel estimation and equalization performance of the proposed method and the GSTBC method for $\mathrm{SNR} \leq 13$ dB may be due to the different precoding sequences and equalizers used. Fig. 9 shows that when the number of samples is 200 (for each transmitter), the identification and equalization performance of the proposed method is better than that of the GSTBC method for $\mathrm{SNR} \leq 15$ dB. Fig. 9 shows that when the number of samples is small, the proposed method has better performance than the GSTBC method under low SNR.

VI. Conclusion

We propose a method for blind identification of FIR MIMO channels using periodic precoding sequence. Since the cyclostationarity is induced at the transmitter, the identifiability condition imposed on the channel is minimum: it only requires that channel-impulse response matrix $\mathbf{H} = [\mathbf{H}(0)^T \;\; \mathbf{H}(1)^T \;\; \cdots \;\; \mathbf{H}(L)^T]^T$ is full column rank. The channel transfer matrix is not required to be irreducible or column reduced. The channel can have more outputs or more inputs. The method is shown to be robust with respect to channel-order overestimation. The performance of the algorithm depends on the precoding sequence which is optimally designed to reduce the effect of noise and error in estimating the covariance matrix of the received data. Simulation results show that the method yields good performance.

Appendix

A. Proof of Proposition 4.1 and 4.2

Preliminary: For each j, let $\mathbf{N}_j \in \mathbb{R}^{(N-j)\times(L-j+1)}$ be similarly defined as (3.13), except that $\mathbf{I}_{M_r}$ is replaced by 1. It can be easily check that there exists permutation matrices $\mathbf{P}_{\mathbf{l_j}} \in \mathbb{R}^{M_r(N-j)\times M_r(N-j)}$ and $\mathbf{P}_{\mathbf{r_j}} \in \mathbb{R}^{M_r(L-j+1)\times M_r(L-j+1)}$ such that $\mathbf{P}_{\mathbf{l_j}}\mathbf{M}_j\mathbf{P}_{\mathbf{r_j}} =$

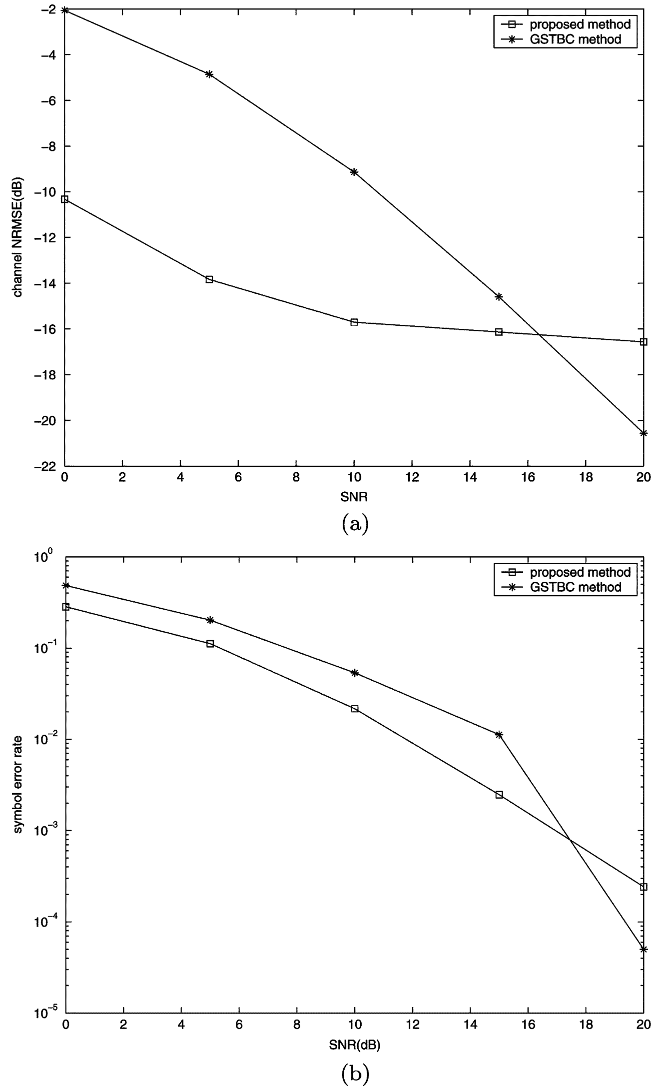


Fig. 9. Comparison of NRMSE and symbol error rate, number of input samples $= 200$. (a) Channel NRMSE versus output SNR. (b) Symbol error rate versus output SNR.

$\mathrm{diag}[\mathbf{N}_j, \mathbf{N}_j, \ldots, \mathbf{N}_j] = \mathbf{D}_j \in \mathbb{R}^{M_r(N-j)\times M_r(L-j+1)}$ is a block diagonal matrix with each block of dimension $(N-j)\times(L-j+1)$. Since ${\mathbf{P}_{\mathbf{l_j}}}^T = {\mathbf{P}_{\mathbf{l_j}}}^{-1}$ and ${\mathbf{P}_{\mathbf{r_j}}}^T = {\mathbf{P}_{\mathbf{r_j}}}^{-1}$ [21, p. 110], we have $\mathbf{M}_j = {\mathbf{P}_{\mathbf{l_j}}}^T \mathbf{D}_j {\mathbf{P}_{\mathbf{r_j}}}^T$. Hence, $\mathbf{M}_j$ is full column rank if and only if $\mathbf{N}_j$ is full column rank for $j = 0, 1, \ldots, L$.

Also, $\mathbf{M}_j^T \mathbf{M}_j = (\mathbf{P}_{\mathbf{r_j}} \mathbf{D}_j^T \mathbf{P}_{\mathbf{l_j}})({\mathbf{P}_{\mathbf{l_j}}}^T \mathbf{D}_j {\mathbf{P}_{\mathbf{r_j}}}^T) = \mathbf{P}_{\mathbf{r_j}} \mathbf{D}_j^T \mathbf{D}_j {\mathbf{P}_{\mathbf{r_j}}}^T = \mathbf{P}_{\mathbf{r_j}} \mathrm{diag}[\mathbf{N}_j^T\mathbf{N}_j, \ldots, \mathbf{N}_j^T\mathbf{N}_j] {\mathbf{P}_{\mathbf{r_j}}}^T$. Let $\lambda(\mathbf{A})$ denote the spectrum of $\mathbf{A}$ [21, p. 310], that is, the set of eigenvalues of $\mathbf{A}$. Then $\lambda(\mathbf{M}_j^T\mathbf{M}_j) = \lambda(\mathbf{N}_j^T\mathbf{N}_j)$.

Proof of Proposition 4.1: If at $N-L+1 \le m \le N-2$, it can be checked that $\mathbf{N}_j, j = 2, 3, \ldots, L-1$ is not of full column rank since it has two columns both equal to $[\tau\ \tau\ \cdots\ \tau]^T$ which implies that at least one $\mathbf{M}_j$ is rank deficient and vice versa.

Proof of Proposition 4.2: From the **Preliminary**, since $\lambda(\mathbf{M}_j^T\mathbf{M}_j) = \lambda(\mathbf{N}_j^T\mathbf{N}_j)$, the condition number of $\mathbf{M}_j^T\mathbf{M}_j$ is identical to that of $\mathbf{N}_j^T\mathbf{N}_j$, i.e., $\kappa(\mathbf{M}_j^T\mathbf{M}_j) = \kappa(\mathbf{N}_j^T\mathbf{N}_j)$. Thus, we need only compute the condition number of $\mathbf{N}_j^T\mathbf{N}_j$.

Case (a): For $m = 0, m = 1, \ldots$, and $m = N-L-1$, we know

$$\mathbf{N}_j^T\mathbf{N}_j = a \cdot \mathbf{I}_{L-j+1} + (2b + c_j)\cdot[1\cdots1]^T[1\cdots1] \quad \text{(A.1)}$$

where $a = N^2(1-\tau)^2, b = N\tau(1-\tau), c_j = (N-j)\tau^2$. Hence, the maximum and minimum eigenvalues are $a + (L-j+1)(2b+c_j)$ and a respectively. Thus, the condition number of $\mathbf{M}_j^T\mathbf{M}_j$ is $1 + [(L-j+1)(2b+c_j)/a]$ which is a decreasing function of j. Therefore, the corresponding μ is equal to $\mu_1 = 1 + [(L+1)(2b+c_0)/a]$.

Case (b): For $m = N-L$ and $m = N-1$, we consider the $j = 0$ case and $j \ne 0$ case for N_j separately. For $j = 0$ with $m = N-L$ or $m = N-1$, direct multiplication of $\mathbf{N}_0^T\mathbf{N}_0$ gives the same matrix as (A.1), and the condition number of $\mathbf{M}_0^T\mathbf{M}_0$ is μ_1. For $j \ne 0$ with $m = N-L$, direct multiplication of $\mathbf{N}_j^T\mathbf{N}_j$ yields (A.2), shown at the bottom of the page. The eigenvalues of $\mathbf{N}_j^T\mathbf{N}_j$ in ascending order, are α_j, a, β_j, where a has a multiplicity $L-j-1$, and $\beta_j = (1/2)\{(L-j)(2b+c_j) + (a+c_j) + \sqrt{[(L-j)(2b+c_j)+(a-c_j)]^2 + 4(L-j)(b+c_j)^2}\}$, $\alpha_j = (1/2)\{(L-j)(2b+c_j) + (a+c_j) - \sqrt{[(L-j)(2b+c_j)+(a-c_j)]^2 + 4(L-j)(b+c_j)^2}\}$. All of the eigenvalues are positive and real. (A proof is given in Appendix B). It can be similarly shown that for $j \ne 0$ with $m = N-1, \mathbf{N}_j^T\mathbf{N}_j$ has the same eigenvalues α_j, a, β_j. Hence, for $j = 1, 2, \ldots, L, \lambda(\mathbf{M}_j^T\mathbf{M}_j) = \{\alpha_j, a, \beta_j\}$ and the condition number is

$$\begin{aligned}\kappa\left(\mathbf{M}_j^T\mathbf{M}_j\right) &= \frac{\beta_j}{\alpha_j} = 1 \\ &\quad + \frac{\chi_j^2 - 4(N-L)b^2 + \chi_j\sqrt{\chi_j^2 - 4(N-L)b^2}}{2(N-L)b^2},\end{aligned} \quad \text{(A.3)}$$

where $\chi_j = (L-j)(2b+c_j) + a + c_j$. Since β_j/α_j is also a decreasing function of j, then the maximum value is β_1/α_1. Therefore, combining the two cases $(j = 0, j \ne 0)$, the corresponding μ is $\mu_2 = \max\{\mu_1, \beta_1/\alpha_1\} \ge \mu_1$.

$$\mathbf{N}_j^T\mathbf{N}_j = \begin{bmatrix} a+2b+c_j & 2b+c_j & 2b+c_j & \cdots & 2b+c_j & b+c_j \\ 2b+c_j & a+2b+c_j & 2b+c_j & \cdots & 2b+c_j & b+c_j \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 2b+c_j & 2b+c_j & 2b+c_j & \cdots & a+2b+c_j & b+c_j \\ b+c_j & b+c_j & b+c_j & \cdots & b+c_j & c_j \end{bmatrix} \in \mathbb{R}^{(L-j+1)\times(L-j+1)} \quad \text{(A.2)}$$

B. Eigenvalues of $\mathbf{N}_j^T\mathbf{N}_j$ for $m = N - L$

Proof: Let $\mathbf{A}_j = \mathbf{N}_j^T\mathbf{N}_j$ defined in (A.2), then $\mathbf{A}_j$ is positive definite since $\mathbf{N}_j$ is full column rank. It can be checked that the eigenvectors corresponding to $(L-j-1)$ multiple eigenvalue a are $[1,-1,0,0,\ldots,0]^T, [1,1,-2,0,\ldots,0]^T, \ldots, [1,1,\ldots,1,-(L-j-1),0]^T$. The remaining eigenvectors are $[1,1,\ldots,1,x]^T \in \mathbb{R}^{L-j+1}$. Hence,

$$\mathbf{A}_j \begin{bmatrix} 1 \\ \vdots \\ 1 \\ x \end{bmatrix} = \begin{bmatrix} a+(L-j)(2b+c_j)+(b+c_j)x \\ \vdots \\ a+(L-j)(2b+c_j)+(b+c_j)x \\ (L-j)(b+c_j)+c_j x \end{bmatrix} = \lambda_j \begin{bmatrix} 1 \\ \vdots \\ 1 \\ x \end{bmatrix} \tag{B.1}$$

which implies the following two equations

$$a+(L-j)(2b+c_j)+(b+c_j)x = \lambda_j \tag{B.2}$$

$$(L-j)(b+c_j)+c_j x = \lambda_j x. \tag{B.3}$$

Substitute (B.2) into (B.3), we can get an second-order equation of x. Solving this equation can lead to two solutions of x. Bring these two x into (B.2) and we can obtain the two eigenvalues β_j, α_j. In addition, $\beta_j \geq a$ because of (B.4)

$$\begin{aligned}\beta_j &= \frac{1}{2}\{(L-j)(2b+c_j)+(a+c_j) \\ &\quad + \sqrt{[(L-j)(2b+c_j)+a-c_j]^2+4(L-j)(b+c_j)^2}\} \\ &\geq \frac{1}{2}\{(L-j)(2b+c_j)+(a+c_j) \\ &\quad + \sqrt{[(L-j)(2b+c_j)+a-c_j]^2}\} \\ &= \frac{1}{2}\{[(L-j)(2b+c_j)+(a+c_j) \\ &\quad + [(L-j)(2b+c_j)+a-c_j]\} \\ &= a+(L-j)(2b+c_j) \\ &\geq a \end{aligned} \tag{B.4}$$

and $\alpha_j \leq a$ because of the interlacing property [21, p. 396].

REFERENCES

[1] K. Abed-Meraim, W. Qiu, and Y. Hua, "Blind system identification," *Proc. IEEE*, vol. 85, no. 8, pp. 1310–1332, Aug. 1997.

[2] Z. Ding and Y. Li, *Blind Equalization and Identification*. New York: Marcel Dekker, 2001.

[3] G. B. Giannakis, Y. Hua, P. Stoica, and L. Tong, *Signal Processing Advances in Wireless and Mobile Communications Volume I: Trends in Channel Identification and Equalization*. Englewood Cliffs, NJ: Prentice-Hall, 2001.

[4] M. K. Tsatsanis and G. B. Giannakis, "Transmitter induced cyclostationarity for blind channel equalization," *IEEE Trans. Signal Process.*, vol. 45, no. 7, pp. 1785–1794, Jul. 1997.

[5] E. Serpedin and G. B. Giannakis, "Blind channel identification and equalization with modulation induced cyclostationarity," *IEEE Trans. Signal Process.*, vol. 46, no. 7, pp. 1930–1944, Jul. 1998.

[6] A. Scaglione, G. B. Giannakis, and S. Barbarossa, "Redundant filter bank precoders and equalizers—I: Unification and optimal designs.," *IEEE Trans. Signal Process.*, vol. 47, no. 7, pp. 1988–2006, Jul. 1999.

[7] A. Scaglione, G. B. Giannakis, and S. Barbarossa, "Redundant filter bank precoders and equalizers—: Blind channel estimation, synchronization, and direct equalization," *IEEE Trans. Signal Process.*, vol. 47, no. 7, pp. 2007–2022, Jul. 1999.

[8] A. Chevreuil, E. Serpedin, P. Loubaton, and G. B. Giannakis, "Blind channel identification and equalization using periodic modulation precoders: Performance analysis," *IEEE Trans. Signal Process.*, vol. 48, no. 6, pp. 1570–1586, Jun. 2000.

[9] E. Serpedin, A. Chevreuil, G. B. Giannakis, and P. Loubaton, "Blind channel and carrier frequency offset estimation using periodic modulation precoders," *IEEE Trans. Signal Process.*, vol. 48, no. 8, pp. 2389–2405, Aug. 2000.

[10] P. Ciblat, A. Chevreuil, and P. Loubaton, "α-repetition/modulation and blind second-order identification," *IEEE Trans. Signal Process.*, vol. 48, no. 11, pp. 3153–3161, Nov. 2000.

[11] C. A. Lin and J. W. Wu, "Blind identification with periodic modulation: A time-domain approach," *IEEE Trans. Signal Process.*, vol. 50, no. 11, pp. 2875–2888, Nov. 2002.

[12] A. Chevreuil and P. Loubaton, "MIMO blind second-order equalization method and conjugate cyclostationarity," *IEEE Trans. Signal Process.*, vol. 47, no. 2, pp. 572–578, Feb. 1999.

[13] H. Bölcskei, R. W. Heath, Jr., and A. J. Paulraj, "Blind channel estimation in spatial multiplexing systems using nonredundant antenna precoding," in *Proc. 33rd Ann. IEEE Asilomar Conf. Signals, Syst., Comput.*, Oct. 1999, vol. 2, pp. 1127–1132.

[14] H. Bölcskei, R. W. Heath, Jr., and A. J. Paulraj, "Blind channel identification and equalization in OFDM-based multiantenna systems," *IEEE Trans. Signal Process.*, vol. 50, no. 1, pp. 96–109, Jan. 2002.

[15] Z. Ding, "Matrix outer-product decomposition method for blind multiple channel identification," *IEEE Trans. Signal Process.*, vol. 45, no. 12, pp. 3053–3061, Dec. 1997.

[16] Z. Ding and L. Qiu, "Blind MIMO channel identification from second-order statistics using rank deficient channel convolution matrix," *IEEE Trans. Signal Process.*, vol. 51, no. 2, pp. 535–544, Feb. 2003.

[17] Z. Ding and D. B. Ward, "Subspace approach to blind and semi-blind channel estimation for space-time block codes," *IEEE Trans. Wireless Commun.*, vol. 4, no. 2, pp. 357–362, Mar. 2005.

[18] E. Moulines, P. Duhamel, J.-F. Cardoso, and S. Mayrargue, "Subspace method for the blind identification of multichannel FIR filters," *IEEE Trans. Signal Process.*, vol. 43, no. 2, pp. 516–525, Feb. 1995.

[19] C. A. Lin and C. W. King, "Minimal periodic realization of transfer matrices," *IEEE Trans. Autom. Contr.*, vol. 38, no. 3, pp. 462–466, Mar. 1993.

[20] M. W. Hirsch and S. Smale, *Differential Equations, Dynamical Systems, and Linear Algebra*. New York: Academic , 1974.

[21] G. H. Golub and C. F. Van Loan, *Matrix Computions*, 3rd ed. Baltimore, MD: The Johns Hopkins University Press, 1996.

[22] F. Zhang, *Matrix Theory: Basic Results and Techniques*. New York: Springer-Verlag, 1999.

Ching-An Lin received the B.S. degree from the National Chiao Tung University, Hsinchu, Taiwan, R.O.C., in 1977, the M.S. degree from the University of New Mexico, Albuquerque, in 1980, and the Ph.D. degree from the University of California, Berkeley, in 1984, all in electrical engineering.

He was with the Chung Shan Institute of Science and Technology from 1977 to 1979 and with Integrated Systems Inc. from 1984 to 1986. Since June, 1986, he has been with the Department of Electrical and Control Engineering, the National Chiao Tung University, where he is a Professor. His current research interests are in control and signal processing.

Yi-Sheng Chen received the B.S. and the M.S. degrees, both in electrical engineering from the National Chung Hsing University, Taichung, Taiwan, R.O.C., in 1997 and 1999, respectively, and the Ph.D. degree in electrical and control engineering from National Chiao Tung University, Hsinchu, Taiwan, R.O.C., in 2006.

He was with the military service from 1999 to 2001, and with the Kaohsiung Rapid Transmit Corporation from 2001 to 2002. His current research interests are in signal processing and control.