Resolving Multidimensional Ambiguity in Blind Channel Estimation of MIMO-FIR Systems via Block Precoding

Feifei Gao, *Student Member, IEEE*, and Arumugam Nallanathan, *Senior Member, IEEE*

***Abstract*—In this paper, we consider the identification problem of multiinput–multioutput (MIMO) finite-impulse-response systems via second-order statistics only. By assigning different block precoders to different transmitters, we develop a new technique that allows blind MIMO channel identification up to a scalar ambiguity for each transmitter. We provide sufficient conditions for removal of the matrix ambiguity for a specific set of precoding matrices and derive a general theorem for other kinds of precoding matrices based on a reasonable conjecture. This theorem is firmly tested via numerical examples. Two potential precoding schemes are proposed, considering different ways of eliminating interblock interference. Finally, numerical results are provided to verify our analysis.**

***Index Terms*—Ambiguity, blind channel estimation, block transmissions, multiinput multioutput (MIMO), precoding, subspace (SS) method.**

I. Introduction

MULTIANTENNA transmission over multiinput–multioutput (MIMO) channels has recently been proven effective in combating fading, as well as enhancing data rates [1], [2]. Since coherent detection requires accurate channel state information, channel estimation has become a critical component in a variety of modern wireless communication systems. Many communication systems identify the channel coefficients by transmitting pilot symbols that are known to both transmitters and receivers. These pilot-aided schemes, however, reduce the transmission-bandwidth efficiency [3]. Therefore, blind-channel-estimation algorithm has received considerable attention during the past few decades.

The subspace (SS)-based blind-channel-estimation algorithm has been developed in [4]–[7] for either single-input–single-output systems (SISO) or single-input–multioutput systems, where the channel coefficients between the transmitter, and all the receivers could be identified within a complex scalar ambiguity. This scalar ambiguity could easily be removed by further transmission of one or a few amount of pilot symbols. However, when the SS method is directly applied to the MIMO system [8]–[10], the channels could only be estimated up to an unknown matrix ambiguity. This matrix ambiguity is a multidimensional problem and is not acceptable in most applications. In [11], this matrix ambiguity is resolved under the assumption that sources from different transmitters are non-Gaussian and statistically independent. The property of cyclostationary is exploited in [12] and [13]. Other works discussing linear prefiltering via z-domain polynomial analysis can be found in [14]–[16] and the references therein. In a recent work [17], a linear-precoding technique is applied at the inputs for MIMO orthogonal frequency-division multiplexing (OFDM)-modulation transmissions. This method needs the transmitted symbols to be strictly white and is mainly proposed for blind channel estimation in a multiinput–single-output (MISO) scenario.

In this paper, we propose a new way of resisting the matrix ambiguity when the SS algorithm is applied to the MIMO systems. We find that by dividing the data sequences into blocks and assigning different block precoders to different transmitters, it is possible to reduce the matrix ambiguity to a scalar ambiguity for each transmitter. Note that block precoding is a well-studied topic through the literatures [18], [19] to improve the performance of the detection. However, few works on block precoding have been proposed regarding the resistance of the channel-estimation ambiguity in the MIMO frequency-selective channels. We provide strict conditions on removing the matrix ambiguity for a specific set of precoding matrices, namely, zero-padding matrix. Then, a more general conclusion on resisting the matrix ambiguity is derived based on a reasonable conjecture. Various numerical examples are provided to demonstrate the effectiveness of our proposed algorithms.

This paper is organized as follows. Section II presents the system model of MIMO transmissions. Section III presents our proposed precoding schemes. Section IV discusses the ambiguity and the identifiability issues. Numerical examples are provided in Section V to exhibit the effectiveness of our algorithms. Finally, conclusion is presented in Section VI, and the proof for the theorem is given in the Appendix.

The following notations are used in this paper. Transpose, complex conjugate, Hermitian, inverse, and pseudoinverse of matrix $\mathbf{A}$ are denoted by $\mathbf{A}^{\mathrm{T}}$, $\mathbf{A}^{*}$, $\mathbf{A}^{\mathrm{H}}$, $\mathbf{A}^{-1}$, and $\mathbf{A}^{\dagger}$, respectively. $[\mathbf{A}]_{ij}$ stands for the (i,j)th entry of the matrix $\mathbf{A}$, $\mathrm{tr}(\mathbf{A})$ denotes the trace operation, $\otimes$ represents the Kronecker product, $\mathbf{I}$ is the identity matrix, and $\mathrm{E}\{\cdot\}$ is the statistical expectation. The MATLAB notations for rows and columns are used in this paper. For example, $\mathbf{A}(:,p)$ represents the

Manuscript received March 16, 2005; revised November 10, 2005, July 20, 2006, and November 24, 2006; accepted April 5, 2007. The review of this paper was coordinated by Prof. R. Heath.

The authors are with the Department of Electrical and Computer Engineering, National University of Singapore, Singapore 119260 (e-mail: feifeigao@nus.edu.sg; elena@nus.edu.sg).

Digital Object Identifier 10.1109/TVT.2007.904542

pth column of the matrix $\mathbf{A}$, and $\mathbf{A}(:, p_1 : p_2)$ represents the submatrix obtained by extracting columns p_1 to p_2 from the matrix $\mathbf{A}$, respectively.

II. System Model

Let us consider an N_t-input N_r-output linear finite-impulse-response (FIR) MIMO system. The data streams from transmit antennas are denoted as $\{b_{i\tau}\}, i = 1, \ldots, N_\mathrm{t}$, and $\tau = 0, 1, \ldots, M-1$, where i and τ are the transmitter and time indexes, respectively. The data sequences are then divided into consecutive blocks with the length of K, namely

$$\mathbf{s}_i(m) = \left[b_{i(mK)}, b_{i(mK+1)}, \ldots, b_{i(mK+K-1)}\right]^\mathrm{T} \qquad m = 0, 1, \ldots, M/K - 1 \quad (1)$$

where m is the block index. Without loss of generality, M is taken as an integer multiple of K. Assume that perfect synchronization is achieved at the receivers. The data stream obtained from the jth receive antenna is denoted as $\{d_{j\tau}\}$, $j = 1, \ldots, N_\mathrm{r}$, which could be divided into blocks as

$$\mathbf{r}_{j(m)} = \left[d_{j(mK)}, d_{j(mK+1)}, \ldots, d_{j(mK+K-1)}\right]^\mathrm{T}. \quad (2)$$

For convenience, we assume that all channel responses between different pairs of transmitters and receivers have the same channel order L. Let

$$\mathbf{h}_{ij} = [h_{ij,0}, h_{ij,1}, \ldots, h_{ij,L}]^\mathrm{T} \quad (3)$$

$$\boldsymbol{\mathcal{M}}_K(\mathbf{h}_{ij}) = \begin{bmatrix} h_{ij,L} & \cdots & h_{ij,0} & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & h_{ij,L} & \cdots & h_{ij,0} \end{bmatrix} \quad (4)$$

represent the $(L+1) \times 1$ equivalent discrete channel vector and the $K \times (K+L)$ channel matrix from the ith transmitter to the jth receiver, respectively. The subscript K denotes the number of rows in $\boldsymbol{\mathcal{M}}_K(\mathbf{h}_{ij})$. The combined channel matrix from the ith transmitter to all the receivers can be represented by the $KN_\mathrm{r} \times (K+L)$ matrix

$$\boldsymbol{\mathcal{T}}_K(\mathbf{h}_i) = \left[\boldsymbol{\mathcal{M}}_K(\mathbf{h}_{i1})^\mathrm{T}, \boldsymbol{\mathcal{M}}_K(\mathbf{h}_{i2})^\mathrm{T}, \ldots, \boldsymbol{\mathcal{M}}_K(\mathbf{h}_{iN_\mathrm{r}})^\mathrm{T}\right]^\mathrm{T} \quad (5)$$

where

$$\mathbf{h}_i = \left[\mathbf{h}_{i1}^\mathrm{T}, \mathbf{h}_{i2}^\mathrm{T}, \ldots, \mathbf{h}_{iN_\mathrm{r}}^\mathrm{T}\right]^\mathrm{T} \quad (6)$$

is the combination of channel vectors from the ith transmitter to all the receivers. The overall $KN_\mathrm{r} \times (K+L)N_\mathrm{t}$ channel matrix for the MIMO system is

$$\boldsymbol{\Pi}_K(\mathbf{h}_{ij}) = \left[\boldsymbol{\mathcal{T}}_K(\mathbf{h}_1), \boldsymbol{\mathcal{T}}_K(\mathbf{h}_2), \ldots, \boldsymbol{\mathcal{T}}_K(\mathbf{h}_{N_\mathrm{t}})\right]. \quad (7)$$

The overall received signal block is then modeled as

$$\begin{aligned}\mathbf{r}(m) &= \left[\mathbf{r}_1(m)^\mathrm{T}, \ldots, \mathbf{r}_{N_\mathrm{r}}(m)^\mathrm{T}\right]^\mathrm{T} \\ &= \sum_{i=1}^{N_\mathrm{t}} \boldsymbol{\mathcal{T}}_K(\mathbf{h}_i)\bar{\mathbf{s}}_i(m) + \mathbf{n}(m) \\ &= \boldsymbol{\Pi}_K(\mathbf{h}_{ij})\tilde{\mathbf{s}}(m) + \mathbf{n}(m) \end{aligned} \quad (8)$$

with

$$\bar{\mathbf{s}}_i(m) = \left[\mathbf{s}_{iL}(m-1)^\mathrm{T}, \mathbf{s}_i(m)^\mathrm{T}\right]^\mathrm{T} \quad (9)$$

$$\tilde{\mathbf{s}}_i(m) = \left[\bar{\mathbf{s}}_1(m)^\mathrm{T}, \ldots, \bar{\mathbf{s}}_{N_\mathrm{t}}(m)^\mathrm{T}\right]^\mathrm{T} \quad (10)$$

where $\mathbf{n}(m)$ is the $KN_\mathrm{r} \times 1$ vector whose entry represents the independent identically distributed white Gaussian noise with variance σ^2. $\mathbf{s}_{iL}(m-1)$ denotes the last L entries of the vector $\mathbf{s}_i(m-1)$, $\bar{\mathbf{s}}_i(m)$ is the $(K+L) \times 1$ vector denoting the combination of $\mathbf{s}_{iL}(m-1)$ and $\mathbf{s}_i(m)$ due to the interblock interference (IBI), and $\tilde{\mathbf{s}}(m)$ is the $(K+L)N_\mathrm{t} \times 1$ vector representing the effective transmitted symbol vector for the mth received block $\mathbf{r}(m)$.

An alternative representation of the channel matrix is also provided for later use. Let $\mathbf{H}^l$ denote the $N_\mathrm{r} \times N_\mathrm{t}$ matrix with its (j, i)th entry given by $h_{ij,l}$. After properly permutating the rows and columns of $\boldsymbol{\Pi}_K(\mathbf{h}_{ij})$, the channel matrix can be expressed as

$$\boldsymbol{\Gamma}_K(\mathbf{H}^l) = \begin{bmatrix} \mathbf{H}^L & \mathbf{H}^{L-1} & \cdots & \mathbf{H}^0 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{H}^L & \mathbf{H}^{L-1} & \cdots & \mathbf{H}^0 & \cdots & \mathbf{0} \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \mathbf{0} & \cdots & \mathbf{0} & \mathbf{H}^L & \mathbf{H}^{L-1} & \cdots & \mathbf{H}^0 \end{bmatrix}. \quad (11)$$

Remark 1: The algorithms in this paper are presented based on the channel model $\boldsymbol{\Pi}_K(\mathbf{h}_{ij})$. However, theorems or lemmas are provided based on the second channel model $\boldsymbol{\Gamma}_K(\mathbf{H}^i)$. One can easily establish the equivalence between $\boldsymbol{\Pi}_K(\mathbf{h}_{ij})$ and $\boldsymbol{\Gamma}_K(\mathbf{H}^l)$.

Lemma 1 [20], [21]: The channel identifiability up to an ambiguity matrix by using only the second-order statistics (SOS) of the received signals is ensured if the following conditions are satisfied.

1) $\boldsymbol{\Pi}_K(\mathbf{h}_{ij})$ and $\boldsymbol{\Gamma}_K(\mathbf{H}^l)$ are tall matrices.
2) $\mathbf{H}(z) \triangleq \sum_{l=0}^{L} \mathbf{H}^l z^{-l}$ is irreducible and column-reduced.
3) $\mathbf{H}^L$ is of full column rank.

Since different channel vectors are random vectors, these conditions can be satisfied for most of the practical MIMO systems [16]. Therefore, the MIMO channel discussed in this paper is always considered to satisfy the conditions previously listed, which also indicates that $\boldsymbol{\Pi}_K(\mathbf{h}_{ij})$ and $\boldsymbol{\Gamma}_K(\mathbf{H}^l)$ are full-column-rank matrices.

III. Proposed Precoding Techniques

We found that, by applying different precoders to different transmitters, it is possible to reduce the matrix ambiguity to one scalar ambiguity for each transmitter. The intuitive explanation is that, since we artificially introduce certain "differences" among different transmitters, the channel can be discriminated from each other by exploiting these "differences."

A. Normal Precoding

An illustration of normal precoding scheme at the ith transmitter is shown in Fig. 1, where the original data sequence is

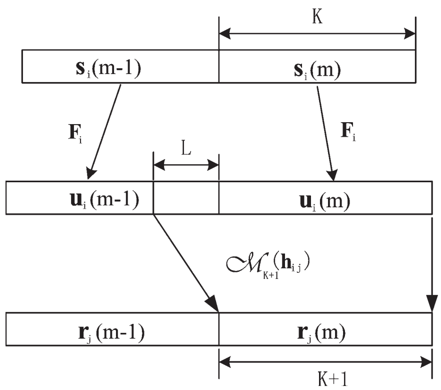


Fig. 1. Normal precoding.

first divided into continuous block $\mathbf{s}_i(m)$ and is then precoded by a $(K+1)\times K$ matrix $\mathbf{F}_i$, resulting in a new data block $\mathbf{u}_i(m)$ with the length $K+1$, namely

$$\mathbf{u}_i(m)=\mathbf{F}_i\mathbf{s}_i(m). \tag{12}$$

At the receiver, the received data stream $\{d_{j\tau}\}$, $\tau=0,\ldots,M(1+1/K)-1$ is divided into blocks of length $K+1$, namely

$$\mathbf{r}_j(m)=\left[d_{j(m(K+1))},\ldots,d_{j(m(K+1)+K)}\right]^{\mathrm{T}} \qquad m=0,\ldots,M/K-1 \tag{13}$$

as shown in Fig. 1. The overall received signal vector $\mathbf{r}(m)$ can be written as

$$\mathbf{r}(m)=\sum_{i=1}^{N_{\mathrm{t}}}\boldsymbol{\mathcal{T}}_{K+1}(\mathbf{h}_i)\bar{\mathbf{u}}_i(m)+\mathbf{n}(m) \tag{14}$$

where $\bar{\mathbf{u}}_i(m)=[\mathbf{u}_{iL}^{\mathrm{T}}(m-1),\ \mathbf{u}_i^{\mathrm{T}}(m)]^{\mathrm{T}}$, and $\mathbf{u}_{iL}(m-1)$ denotes the last L entries of the block vector $\mathbf{u}_i(m-1)$. Note that $\boldsymbol{\mathcal{T}}_{K+1}(\mathbf{h}_i)$ is of dimension $N_{\mathrm{r}}(K+1)\times(K+L+1)$ but possesses a similar structure as (5). By denoting $\mathbf{F}_{iL}=\mathbf{F}(K+2-L:K+1,:)$ as the last L rows of the matrix $\mathbf{F}_i$, the received signals can be rewritten as

$$\begin{aligned}\mathbf{r}(m)&=\sum_{i=1}^{N_{\mathrm{t}}}\boldsymbol{\mathcal{T}}_{K+1}(\mathbf{h}_i)\begin{bmatrix}\mathbf{F}_{iL}&\mathbf{0}\\\mathbf{0}&\mathbf{F}_i\end{bmatrix}\begin{bmatrix}\mathbf{s}_i(m-1)\\\mathbf{s}_i(m)\end{bmatrix}+\mathbf{n}(m)\\&=\sum_{i=1}^{N_{\mathrm{t}}}\boldsymbol{\mathcal{T}}_{K+1}(\mathbf{h}_i)\breve{\mathbf{F}}_i\breve{\mathbf{s}}_i(m)+\mathbf{n}(m)\end{aligned} \tag{15}$$

where $\breve{\mathbf{F}}_i$ is the corresponding $(K+1+L)\times 2K$ matrix, and $\breve{\mathbf{s}}_{i(m)}$ is constructed by two consecutive blocks of source vectors. In order to guarantee the recoverability of $\breve{\mathbf{s}}_{i(m)}$, $\breve{\mathbf{F}}_i$ has to be either a full-rank tall or a square matrix. We then design matrix $\mathbf{F}_i$ such that all the first $K-L$ columns of $\mathbf{F}_{iL}$ are zero, namely

$$\mathbf{F}_{iL}=[\,\mathbf{0}\quad\mathbf{F}_i^{LL}\,] \tag{16}$$

where $\mathbf{F}_i^{LL}=\mathbf{F}(K+2-L:K+1,K+1-L:K)$ is constructed by the last L rows and the last L columns of $\mathbf{F}_i$. The signal vector $\mathbf{r}(m)$ can be reexpressed as

$$\begin{aligned}\mathbf{r}(m)&=\sum_{i=1}^{N_{\mathrm{t}}}\boldsymbol{\mathcal{T}}_{K+1}(\mathbf{h}_i)\begin{bmatrix}\mathbf{F}_i^{LL}&\mathbf{0}\\\mathbf{0}&\mathbf{F}_i\end{bmatrix}\begin{bmatrix}\mathbf{s}_{iL}(m-1)\\\mathbf{s}_i(m)\end{bmatrix}+\mathbf{n}(m)\\&=\sum_{i=1}^{N_{\mathrm{t}}}\boldsymbol{\mathcal{T}}_{K+1}(\mathbf{h}_i)\bar{\mathbf{F}}_i\bar{\mathbf{s}}_i(m)+\mathbf{n}(m)\end{aligned} \tag{17}$$

where $\bar{\mathbf{F}}_i$ is the corresponding $(K+L+1)\times(K+L)$ matrix. Define

$$\tilde{\mathbf{F}}=\mathrm{diag}\{\bar{\mathbf{F}}_1,\bar{\mathbf{F}}_2,\ldots,\bar{\mathbf{F}}_{N_{\mathrm{t}}}\} \tag{18}$$

as the $(K+1)N_{\mathrm{t}}\times KN_{\mathrm{t}}$ block diagonal matrix representing the overall precoding matrix. The received signal covariance matrix is given by

$$\begin{aligned}\mathbf{R}&=\mathrm{E}\left\{\mathbf{r}(m)\mathbf{r}^{\mathrm{H}}(m)\right\}\\&=\boldsymbol{\Pi}_{K+1}(\mathbf{h}_{ij})\tilde{\mathbf{F}}\mathbf{R}_{\mathrm{s}}\tilde{\mathbf{F}}^{\mathrm{H}}\boldsymbol{\Pi}_{K+1}(\mathbf{h}_{ij})^{\mathrm{H}}+\sigma^2\mathbf{I}_{N_{\mathrm{r}}(K+1)}\end{aligned} \tag{19}$$

where $\mathbf{R}_{\mathrm{s}}=\mathrm{E}\{\tilde{\mathbf{s}}(m)\tilde{\mathbf{s}}^{\mathrm{H}}(m)\}$ is the $(K+L)\times(K+L)$ source covariance matrix. Since, normally, no two sources are fully correlated with each other,[1] $\mathbf{R}_{\mathrm{s}}$ can always be considered as a nonsingular matrix. Thus, $\boldsymbol{\Pi}_{K+1}(\mathbf{h}_{ij})\tilde{\mathbf{F}}$ is a full-column-rank matrix, and the covariance matrix $\mathbf{R}$ can be eigen-decomposed as

$$\mathbf{R}=\mathbf{E}\boldsymbol{\Lambda}\mathbf{E}^{\mathrm{H}}+\sigma^2\mathbf{G}\mathbf{G}^{\mathrm{H}} \tag{20}$$

where the $(K+L)N_{\mathrm{t}}\times(K+L)N_{\mathrm{t}}$ diagonal matrix $\boldsymbol{\Lambda}$ contains the signal-SS eigenvalues of $\mathbf{R}$. In turn, the columns of the $(K+1)N_{\mathrm{r}}\times(K+L)N_{\mathrm{t}}$ matrix $\mathbf{E}$ contain the signal-SS eigenvectors of $\mathbf{R}$, whereas the $(K+1)N_{\mathrm{r}}\times((K+1)N_{\mathrm{r}}-(K+L)N_{\mathrm{t}})$ matrix $\mathbf{G}$ is composed of the noise-SS eigenvectors of $\mathbf{R}$.

From the SS detection theory [4], we know that the noise SS spanned by $\mathbf{G}$ is the orthogonal complement space spanned by $\boldsymbol{\Pi}_K(\mathbf{h}_{ij})\tilde{\mathbf{F}}$. Hence, for different $\boldsymbol{\mathcal{T}}_{K+1}(\mathbf{h}_i)$, the following equation holds:

$$\begin{aligned}&\mathbf{G}^{\mathrm{H}}\boldsymbol{\mathcal{T}}_{K+1}(\mathbf{h}_i)\bar{\mathbf{F}}_i(:,p)=\mathbf{0},\qquad p=1,\ldots,K+L\\\Rightarrow\quad&(\boldsymbol{\mathcal{G}}_{ip})^{\mathrm{H}}\mathbf{h}_i=\mathbf{0}\\\Rightarrow\quad&\mathbf{h}_i^{\mathrm{H}}\boldsymbol{\mathcal{G}}_{ip}(\boldsymbol{\mathcal{G}}_{ip})^{\mathrm{H}}\mathbf{h}_i=0\end{aligned} \tag{21}$$

where $\boldsymbol{\mathcal{G}}_{ip}$ can straightforwardly be calculated from $\mathbf{G}^{\mathrm{H}}$ and $\bar{\mathbf{F}}_i(:,p)$.

Let

$$\Phi_i=\sum_{p=1}^{K+L}\boldsymbol{\mathcal{G}}_{ip}(\boldsymbol{\mathcal{G}}_{ip})^{\mathrm{H}}. \tag{22}$$

The estimate of $\mathbf{h}_i$, which is denoted as $\hat{\mathbf{h}}_i$, can be obtained from the eigenvector of $\boldsymbol{\Phi}_i$ that corresponds to the smallest eigenvalue; therefore, different $\mathbf{h}_i$ can be determined from its corresponding $\boldsymbol{\Phi}_i$.

[1] Partly correlated sources do not affect the rank of $\mathbf{R}_{\mathrm{s}}$.

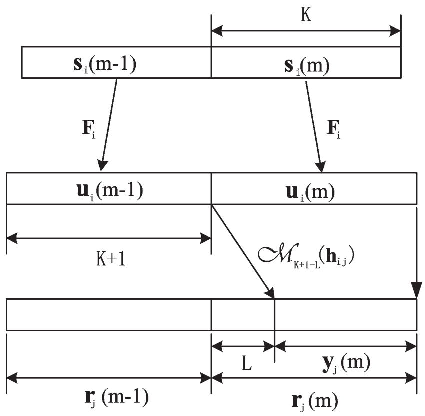


Fig. 2. Simplified precoding.

Remark 2: To resolve the matrix ambiguity, it is necessary that $\mathbf{\Phi}_i$ has only one zero eigenvalue for each i. The related discussion is provided in Section IV. A necessary condition is quoted here, which states that different $\bar{\mathbf{F}}_i$ should span different SSs from each other.

B. Simplified Precoding

Due to the existence of IBI, the proposed algorithm forces an $L \times (K-L)$ submatrix of $\mathbf{F}_i$ to be zero, thus reducing the flexibility of precoder design. An alternative way of eliminating the IBI can simply be implemented by deleting the first L elements in each received block $\mathbf{r}_j(m)$. The remaining symbol vector at the jth receive antenna can be expressed as

$$\mathbf{y}_j(m) = \left[d_{j(m(K+1)+L)}, \ldots, d_{j(m(K+1)+K)}\right]^{\mathrm{T}}. \tag{23}$$

Then, the additional constraint (16) is removed. The transmission scheme is shown in Fig. 2, and the overall signal vector $\mathbf{y}(m)$ can be written as

$$\begin{aligned}\mathbf{y}(m) &= \left[\mathbf{y}_1(m)^{\mathrm{T}}, \mathbf{y}_2(m)^{\mathrm{T}}, \ldots, \mathbf{y}_{N_{\mathrm{r}}}(m)^{\mathrm{T}}\right]^{\mathrm{T}} \\ &= \mathbf{\Pi}_{K+1-L}(\mathbf{h}_{ij})\tilde{\mathbf{F}}\tilde{\mathbf{s}}(m) + \mathbf{n}(m) \end{aligned} \tag{24}$$

where we still use $\tilde{\mathbf{F}}$ to represent the overall precoding matrix and $\tilde{\mathbf{s}}(m)$ to represent the overall source vectors, namely

$$\tilde{\mathbf{F}} = \mathrm{diag}\{\mathbf{F}_1, \mathbf{F}_2, \ldots, \mathbf{F}_{N_{\mathrm{t}}}\} \tag{25}$$

$$\tilde{\mathbf{s}}(m) = \left[\mathbf{s}_1^{\mathrm{T}}(m), \ldots, \mathbf{s}_{N_{\mathrm{t}}}^{\mathrm{T}}(m)\right]^{\mathrm{T}}. \tag{26}$$

The new noise SS matrix $\mathbf{G}_{\mathbf{y}}$ is obtained from the eigen decomposition of

$$\mathbf{R}_{\mathbf{y}} = \mathrm{E}\left\{\mathbf{y}(m)\mathbf{y}^{\mathrm{H}}(m)\right\} \tag{27}$$

and is of dimension $(K+1-L)N_{\mathrm{r}} \times ((K+1-L)N_{\mathrm{r}} - KN_{\mathrm{t}})$. Define

$$\Phi_{i\mathbf{y}} = \sum_{p=1}^{K} \boldsymbol{\mathcal{G}}_{\mathbf{y},ip}(\boldsymbol{\mathcal{G}}_{\mathbf{y},ip})^{\mathrm{H}} \tag{28}$$

where $\boldsymbol{\mathcal{G}}_{\mathbf{y},ip}$ is obtained from

$$(\boldsymbol{\mathcal{G}}_{\mathbf{y},ip})^{\mathrm{H}}\mathbf{h}_i = \mathbf{G}_{\mathbf{y}}^{\mathrm{H}}\mathcal{T}_{K+1-L}(\mathbf{h}_i)\mathbf{F}_i(:,p), \qquad p = 1, \ldots, K. \tag{29}$$

Similar to normal precoding, the channel vector $\mathbf{h}_i$ can be estimated from the eigenvector of $\mathbf{\Phi}_{i\mathbf{y}}$ that corresponds to the smallest eigenvalue.

Remark 3: Note that the way of generating precoders now becomes much easier. We can randomly generate a full-rank $(K+1) \times (K+1)$ matrix $\mathbf{F}$ and then select K different columns to form $\mathbf{F}_i$, by which means we can guarantee that different $\mathbf{F}_i$ is a full rank and spans different SS from each other. However, since the simplified algorithm does not consider the contribution of the first L symbols in the received block for each receiver, it cannot make full use of the received symbol blocks.

Remark 4: Let us take a look at the maximum value of N_{t}. If the precoders of dimension $(K+1) \times K$ are used, as previously assumed, the maximum value of N_{t} is $K+1$ since we can at most construct the $K+1$ precoders that span different SSs from each other. Moreover, we can also use a taller precoding matrix. If the precoders of dimension $(K+q) \times K$ are used, the maximum value of N_{t} can be calculated as C_{k+1}^{q}, where C is the notation of combination.

IV. Ambiguity Resistance and Channel Identifiability

It is known that the channel matrices $\mathbf{\Pi}_{\zeta}(\mathbf{h}_{ij})$ and $\mathbf{\Gamma}_{\zeta}(\mathbf{H}^l)$ that satisfy Lemma 1 are full-rank-tall matrices, where the value of ζ takes $K+1$ for normal precoding or $K+1-L$ for simplified precoding. If $\mathbf{F}_i$ and $\bar{\mathbf{F}}_i$ are full ranks for all i's, then $\mathbf{\Pi}_{\zeta}(\mathbf{h}_{ij})\tilde{\mathbf{F}}$ is also a full column rank. Therefore, the SS-based detection [4] could be applied. Denote the estimates of channel matrices as $\mathbf{\Pi}_{\zeta}(\hat{\mathbf{h}}_{ij})$ and $\mathbf{\Gamma}_{\zeta}(\hat{\mathbf{H}}^l)$, respectively. From the SS detection theory, we know that span$(\mathbf{\Pi}_{\zeta}(\hat{\mathbf{h}}_{ij})\tilde{\mathbf{F}} \in$ span$(\mathbf{\Pi}_{\zeta}(\mathbf{h}_{ij})\tilde{\mathbf{F}})$. Therefore

$$\mathbf{\Pi}_{\zeta}(\hat{\mathbf{h}}_{ij})\tilde{\mathbf{F}} = \mathbf{\Pi}_{\zeta}(\mathbf{h}_{ij})\tilde{\mathbf{F}}\boldsymbol{\mathcal{A}} \tag{30}$$

where $\boldsymbol{\mathcal{A}}$ is an unknown $N_{\mathrm{t}}(\zeta+L-1) \times N_{\mathrm{t}}(\zeta+L-1)$ matrix.

A. Ambiguity Resistance

Theorem 1: Suppose that the estimate of $\mathbf{H}(z)$ is achieved within an ambiguity matrix, namely, $\hat{\mathbf{H}}(z) = \mathbf{H}(z)\mathbf{B}$, where $\mathbf{B}$ is an $N_{\mathrm{t}} \times N_{\mathrm{t}}$ unknown matrix. If $\bar{\mathbf{F}}_i$ (for normal precoding) or $\mathbf{F}_i$ (for simplified precoding) is a full rank and spans different spaces from each other, then $\mathbf{B}$ must be a diagonal matrix, which indicates a scalar ambiguity for each transmit antenna.

Proof: Denote $\mathbf{H}_{\mathrm{c}} = [\mathbf{h}_1, \mathbf{h}_2, \ldots, \mathbf{h}_{N_{\mathrm{t}}}]$. The condition $\hat{\mathbf{H}}(z) = \mathbf{H}(z)\mathbf{B}$ can equivalently be expressed as

$$\hat{\mathbf{H}}_{\mathrm{c}} = \mathbf{H}_{\mathrm{c}}\mathbf{B} \quad \text{or} \quad \hat{\mathbf{H}}^l = \mathbf{H}^l\mathbf{B} \tag{31}$$

where $\hat{\mathbf{H}}_{\mathrm{c}} = [\hat{\mathbf{h}}_1, \hat{\mathbf{h}}_2, \ldots, \hat{\mathbf{h}}_{N_{\mathrm{t}}}]$ is the estimate of $\mathbf{H}_{\mathrm{c}}$, and $\hat{\mathbf{H}}^l$ is the estimate of $\mathbf{H}^l$. We only need to prove that $\mathbf{B}$ is a diagonal matrix.

If (31) holds, $\mathbf{\Pi}_\zeta(\hat{\mathbf{h}}_{ij})$ can be rewritten as

$$\mathbf{\Pi}_\zeta(\hat{\mathbf{h}}_{ij}) = \mathbf{\Pi}_\zeta(\mathbf{h}_{ij})\boldsymbol{\mathcal{B}} \tag{32}$$

where

$$\boldsymbol{\mathcal{B}} = \begin{bmatrix} [\mathbf{B}]_{11}\mathbf{I}_{\zeta+L} & \cdots & [\mathbf{B}]_{1N_\mathrm{t}}\mathbf{I}_{\zeta+L} \\ \vdots & \ddots & \vdots \\ [\mathbf{B}]_{N_\mathrm{t}1}\mathbf{I}_{\zeta+L} & \cdots & [\mathbf{B}]_{N_\mathrm{t}N_\mathrm{t}}\mathbf{I}_{\zeta+L} \end{bmatrix} \tag{33}$$

where $[\mathbf{B}]_{pq}$ is the (p,q)th entry of $\mathbf{B}$. By substituting (32) into (30), we obtain

$$\mathbf{\Pi}_\zeta(\mathbf{h}_{ij})\boldsymbol{\mathcal{B}}\tilde{\mathbf{F}} = \mathbf{\Pi}_\zeta(\mathbf{h}_{ij})\tilde{\mathbf{F}}\boldsymbol{\mathcal{A}}. \tag{34}$$

From Lemma 1, we know that $\mathbf{\Pi}_\zeta(\mathbf{h}_{ij})$ is a full-rank-tall matrix. Then, (34) indicates that

$$\boldsymbol{\mathcal{B}}\tilde{\mathbf{F}} = \tilde{\mathbf{F}}\boldsymbol{\mathcal{A}}. \tag{35}$$

Divide $\boldsymbol{\mathcal{A}}$ into blocks as

$$\boldsymbol{\mathcal{A}} = \begin{bmatrix} \boldsymbol{\mathcal{A}}_{11} & \cdots & \boldsymbol{\mathcal{A}}_{1N_\mathrm{t}} \\ \vdots & \ddots & \vdots \\ \boldsymbol{\mathcal{A}}_{N_\mathrm{t}1} & \cdots & \boldsymbol{\mathcal{A}}_{N_\mathrm{t}N_\mathrm{t}} \end{bmatrix}. \tag{36}$$

From (35) and (36), we obtain

$$\begin{aligned} &\begin{bmatrix} [\mathbf{B}]_{11}\mathbf{F}_1 & \cdots & [\mathbf{B}]_{1N_\mathrm{t}}\mathbf{F}_{N_\mathrm{t}} \\ \vdots & \ddots & \vdots \\ [\mathbf{B}]_{N_\mathrm{t}1}\mathbf{F}_1 & \cdots & [\mathbf{B}]_{N_\mathrm{t}N_\mathrm{t}}\mathbf{F}_{N_\mathrm{t}} \end{bmatrix} \\ &\qquad = \begin{bmatrix} \mathbf{F}_1\boldsymbol{\mathcal{A}}_{11} & \cdots & \mathbf{F}_1\boldsymbol{\mathcal{A}}_{1N_\mathrm{t}} \\ \vdots & \ddots & \vdots \\ \mathbf{F}_{N_\mathrm{t}}\boldsymbol{\mathcal{A}}_{N_\mathrm{t}1} & \cdots & \mathbf{F}_{N_\mathrm{t}}\boldsymbol{\mathcal{A}}_{N_\mathrm{t}N_\mathrm{t}} \end{bmatrix} \end{aligned} \tag{37}$$

where $\mathbf{F}_i$ should be changed to $\bar{\mathbf{F}}_i$ for normal precoding. Therefore

$$[\mathbf{B}]_{ij}\mathbf{F}_j = \mathbf{F}_i\boldsymbol{\mathcal{A}}_{ij} \tag{38}$$

for all pairs of (i,j). Since $\mathbf{F}_i$ is a full rank and spans different SSs from each other, it can easily be known that

$$\begin{cases} \boldsymbol{\mathcal{A}}_{ii} = [\mathbf{B}]_{ii}\mathbf{I}_{\zeta+L-1}, & \text{for } i=j \\ [\mathbf{B}]_{ij} = 0, \quad \boldsymbol{\mathcal{A}}_{ij} = \mathbf{0}, & \text{for } i \neq j \end{cases}. \tag{39}$$

Therefore, $\mathbf{B}$ must be a diagonal matrix, and the multidimensional ambiguity is converted to a scalar ambiguity for each transmit antenna. ■

B. Identifiability

Now, the question is whether (31) holds, which is normally categorized as the identifiability problem. Usually, the proof resorts to the z-domain polynomial analysis [20]. However, since the convolution property of the system is broken by block precoding, it is quite hard to implement the polynomial analysis here.

Conjecture 1: Suppose that the conditions listed in Lemma 1 are satisfied. If $\mathbf{F}_i$ (for simplified precoding) or $\bar{\mathbf{F}}_i$ (for normal precoding) is a full-rank matrix, then the channel identifiability $\hat{\mathbf{H}}(z) = \mathbf{H}(z)\mathbf{B}$ for the proposed algorithms can be guaranteed.

An intuition explanation is obtained from [5], [22]–[24], where the redundant precoding is applied to guarantee the channel identifiability, even if the overall channel matrix may not be a full rank. Therefore, it makes no sense that the application of full-rank redundant precoders may destroy the identification of the system whose identifiability is originally guaranteed (under Lemma 1).

We now provide a firm study on channel-estimation identifiability for a specific set of precoders, namely, zero-padding precoders introduced in [25, ch. 10]. Discussion on more general precoders is still an open problem and is currently under investigation.

Theorem 2: For simplified precoding algorithm, if $\mathbf{F}_i$ is obtained by deleting one column from $\mathbf{I}_{K+1}$ (zero-padding precoders), then sufficient conditions for (31) to hold are the following.

1) $\lfloor (K-1)/2 \rfloor > L+1$, where $\lfloor \cdot \rfloor$ denotes the largest integer that is no bigger than the specified value.
2) $N_\mathrm{r} > \lfloor (N_\mathrm{t}(\lfloor (K-1)/2 \rfloor - 1)/(\lfloor (K-1)/2 \rfloor - L - 1)) \rfloor$.

Proof: See the Appendix.

If N_r and K are sufficiently large, then the conditions listed in Theorem 2 could always be satisfied. However, since Theorem 2 is proved based only on several columns in $\mathbf{\Gamma}_\zeta(\mathbf{H})$, the sufficient conditions obtained are rather loose. Normally, even for smaller values of N_r and K that do not satisfy the condition in Theorem 2, the identification could still be achieved.

Remark 5: Although Theorem 2 only considers a simplified precoding algorithm, a similar proof can straightforwardly be applied for zero-padding precoders in the normal precoding algorithm. Note that the size and the structure of $\bar{\mathbf{F}}_i$ will vary if zeros are inserted into the different positions. The proof is rather tedious, and therefore, it is omitted here. However, similar conclusion can be made that, if N_r and K are sufficiently large or are larger than the certain values, the channel identifiability can be guaranteed.

V. Simulation Results

In this section, we provide numerical examples to verify the theorems/algorithms developed in previous sections. As previously claimed, the sufficient conditions listed in Theorem 2 are quite loose; therefore, we will not restrict ourselves in testing only those big K and N_r. For simplicity, the parameters are chosen as $K=7$, $N_\mathrm{t}=2$, $N_\mathrm{r}=3$, and $L=2$. One can see through the simulations that channel identification can still be guaranteed within an unknown complex scalar for each

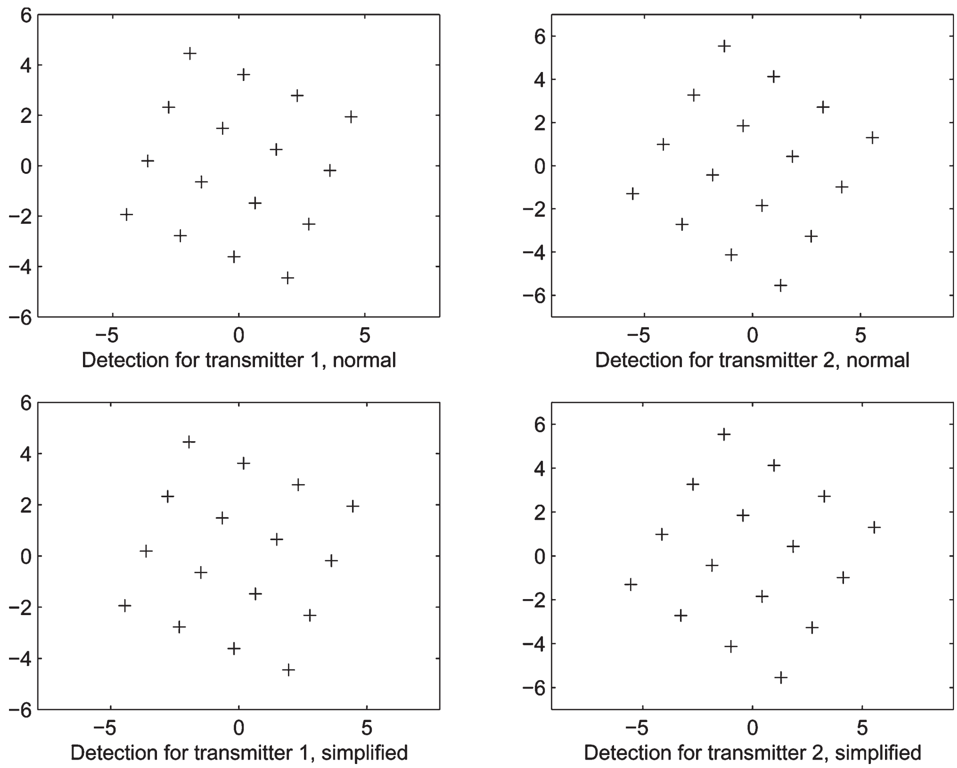


Fig. 3. Detected constellation diagram for zero-padding precoders under noise-free case.

transmit antenna. The channel coefficients are randomly generated, which, in this simulation, are

$$\mathbf{h}_{11} = [0.4608+1.8903i, 0.4574+0.2622i, 0.4507-0.8794i]^{\mathrm{T}}$$
$$\mathbf{h}_{12} = [0.4122-0.0678i, 0.9016-1.4460i, 0.0056+0.6126i]^{\mathrm{T}}$$
$$\mathbf{h}_{13} = [0.2974+1.3720i, 0.0492-0.4288i, 0.6932+0.1289i]^{\mathrm{T}}$$
$$\mathbf{h}_{21} = [0.6501-0.5702i, 0.9830+0.5521i, 0.5527-0.4084i]^{\mathrm{T}}$$
$$\mathbf{h}_{22} = [0.4001-1.1703i, 0.1988-0.5576i, 0.6252-0.2016i]^{\mathrm{T}}$$
$$\mathbf{h}_{23} = [0.7334-2.5679i, 0.3759-0.2724i, 0.0099-0.0160i]^{\mathrm{T}}$$

unless otherwise mentioned. The normalized-estimation mean-square-error (NMSE) of $\mathbf{h}_i$ is defined as [21]

$$\mathrm{NMSE} = \frac{1}{N_{\mathrm{q}}} \sum_{i=1}^{N_{\mathrm{q}}} \frac{\|\hat{\mathbf{h}}_i \alpha_i - \mathbf{h}_i\|^2}{\|\mathbf{h}\|_i^2} \tag{40}$$

where α_i is chosen such that $\|\hat{\mathbf{h}}_i \alpha_i - \mathbf{h}_i\|^2$ is minimized, and $N_{\mathrm{q}} = 100$ is the number of Monte Carlo runs. The symbols from each transmitters are independently generated from 16-quadratic-amplitude modulation (16-QAM). The signal-to-noise ratio (SNR) is defined as the ratio between the symbol power at the transmitter and the noise power at the receiver, which does not take the channel effect into account.

A. *Capability of Ambiguity Resistance*

We adopt the way in [16] to demonstrate the capability of the ambiguity resistance here. We first consider the zero-padding precoding matrix with $\mathbf{F}_1 = \mathbf{I}_{8\times 8}(:, 1 : 7)$ and $\mathbf{F}_2 = \mathbf{I}_{8\times 8}(:, 2 : 8)$. Totally, 30 blocks of signals are assumed to be sent from each transmit antenna, and the channel estimation is purely conducted based on the correspondingly received 30 blocks. For the noise-free case, the data patterns detected for each transmitter applying both algorithms are shown in Fig. 3. The constellations are drawn by placing the real parts of all the received 7×30 symbols onto the x-coordinate, whereas the corresponding imaginary parts are placed onto the y-coordinate. Clearly, the shape of the 16-QAM constellation is kept, but it is rotated and scaled compared with the standard 16-QAM constellation. Therefore, the matrix ambiguity reduces to one scalar ambiguity for each transmitter. Then, we consider the noisy case. Since we only wish to demonstrate the capability of ambiguity resistance of the proposed algorithms, the SNR is taken relatively higher as 25 dB. The detected data patterns are shown in Fig. 4. It can be seen that the proposed algorithms are still able to resolve the multidimensional ambiguity.

It is also important to numerically test the capability of the ambiguity resistance for more general precoding matrices. We consider 10^3 simulation runs and assume the noiseless environment for test purposes only. In each simulation run, the precoding matrices are randomly generated.[2] In this specific

[2]The randomly generated matrix has a full rank with a probability of one. Since the SNR is infinite, we could even adopt an ill-conditioned precoding matrix without affecting the detection results.

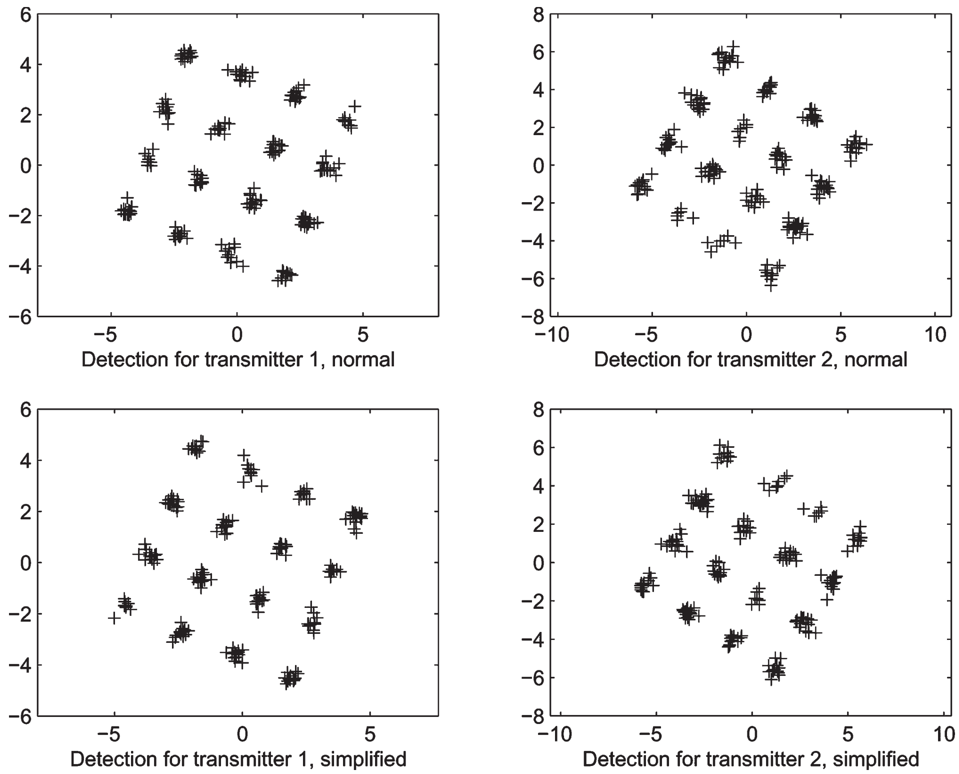


Fig. 4. Detected constellation diagram for zero-padding precoders under noisy case SNR $= 25$ dB.

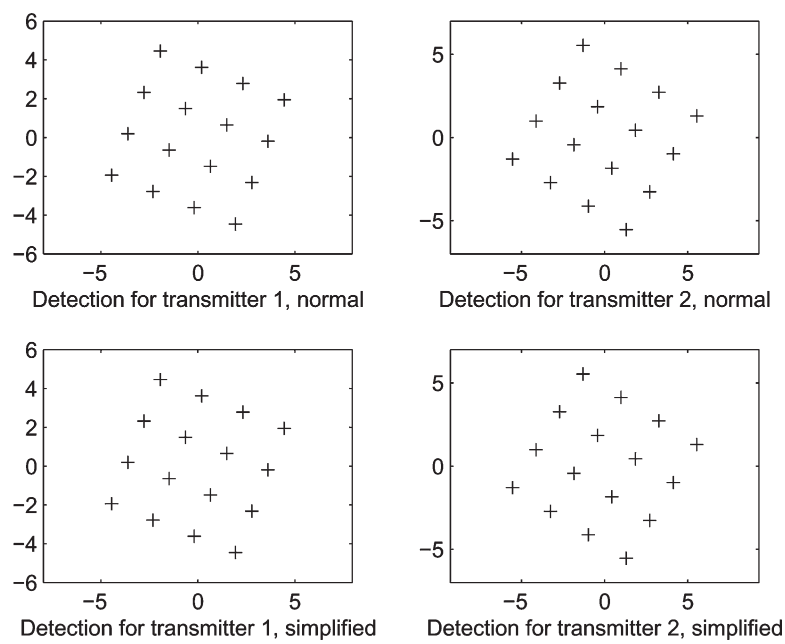


Fig. 5. Detected constellation diagram for random precoders under noise-free case.

example, we also allow the channel to randomly change for different simulation runs. It is found that the detected symbol points exactly form the true 16-QAM constellation, except that the constellation patterns have different scaling factors and rotation angles for different simulation runs. Therefore, we only show the data pattern for the specific channel realization (40) in Fig. 5. By this example, we numerically prove that the matrix ambiguity could be resolved by using more general precoders.

B. Performance of the Proposed Algorithms

In this example, we demonstrate the performance comparison between the normal- and simplified-precoding techniques. The precoders for the two transmitters are taken as $\mathbf{F}_1 = \mathbf{I}_{8\times 8}(:, 1 : 7)$ and $\mathbf{F}_2 = \mathbf{I}_{8\times 8}(:, 2 : 8)$. Each precoder is then scaled to keep the power of the precoded signal unchanged. We choose to compare with the algorithm in [17], where the block

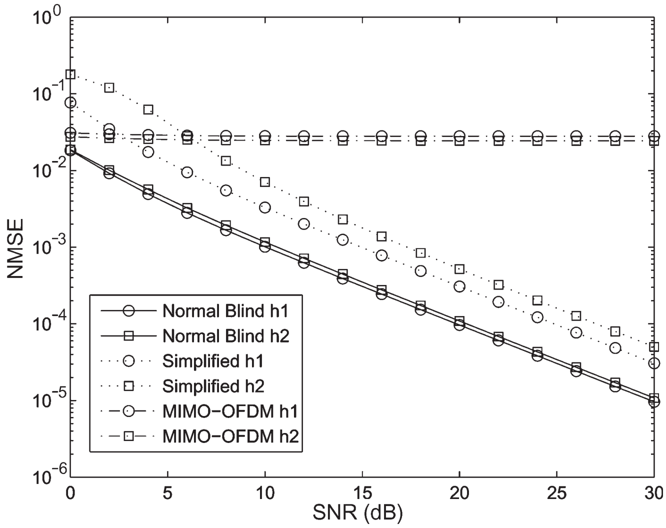


Fig. 6. Channel-estimation nmses versus the SNR: Comparing two algorithms and the existing work [17].

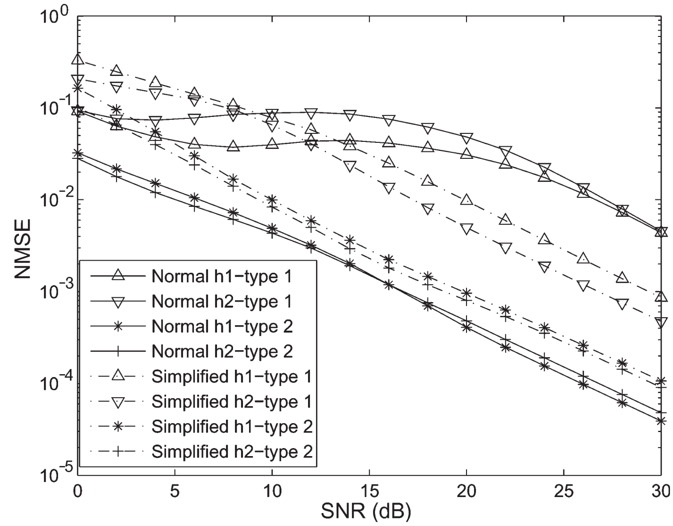


Fig. 7. Performance comparison for different ways of generating $\bar{\mathbf{F}}_i$.

transmissions are also adopted for fair comparisons. In Fig. 6, the NMSEs of individual channel estimation versus the SNRs of these two estimators, as well as the algorithm in [17], are shown. The number of the transmitted symbol blocks is taken as 200 for all algorithms. From Fig. 6, we see that the performance of the normal blind algorithm is better than the simplified algorithm. The reason lies in that it exploits more information from the received signal. It is also seen that both the proposed algorithms outperform the one in [17] at relatively higher SNR region. The reason is that the latter method, although it is applied to block transmissions and possesses scalar ambiguity, is mainly targeted for blind channel estimation in MISO system.

C. Different Precoder Design for Normal Precoding

The precoder $\mathbf{F}_i$ can be generated by extracting different columns from an arbitrary full-rank square matrix. In order to apply both algorithms, $\mathbf{F}_i$ could be modified according to the following two ways.

1) Type 1: Force the first K entries of the last L rows in $\mathbf{F}_i$ to be zero, and then, scale to keep the power of signals unchanged, as indicated in (16).
2) Type 2: In addition to the operation in type 1), force the last L entries of the first $K+1-L$ rows to be zero, and then, scale to keep the power of signals unchanged. In this case, $\mathbf{F}_i$ becomes a block diagonal matrix.

Fig. 7 shows the NMSEs for the channel estimation versus the SNR of both algorithms. It can be seen that normal precoding, as well as simplified precoding by type 2) precoders, gives a satisfactory performance. However, both algorithms by type 1) precoders perform much worse than the other two cases. The reason is that forcing one corner in $\mathbf{F}_i$ to be zero only makes it "imbalanced" and ill-conditioned. Therefore, type 2) construction of linear precoders is preferred for the normal precoding algorithm.

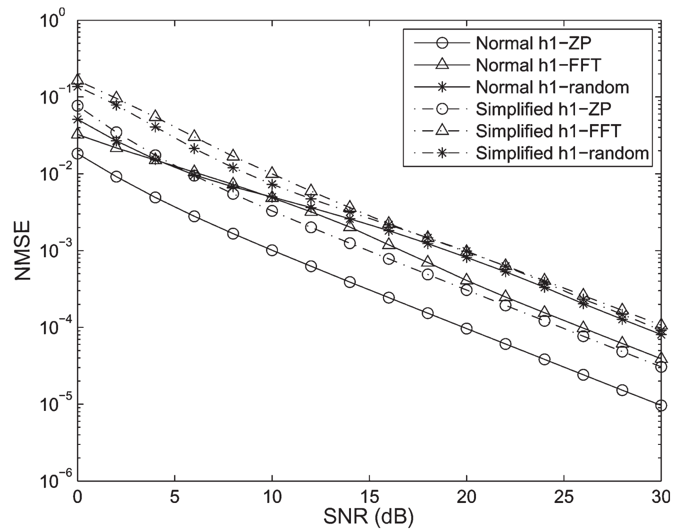


Fig. 8. Performance comparison for different linear precoders.

D. Performance for Different Types of Precoders

We compare the performance of different types of precoders in this example. The performances of the NMSEs versus the SNR for zero-padding precoders, the type 2) precoders extracted from the fast-Fourier-transformation (FFT) matrix, and the type 2) randomly generated precoders are shown in Fig. 8. For simplicity, only the curves for $\mathbf{h}_1$ are plotted. It is seen that the zero-padding precoders perform better than the other linear precoders. This phenomenon was also observed in [25], where the precoding is applied purely for SISO transmissions. Several possible reasons are provided in [25]. Intuitively, the zero-padding precoders do not introduce intersymbol interference, and the followed data detection undergoes less noise enhancement.

VI. Conclusion

We have investigated a new way to identify blind channel estimation for MIMO-FIR systems based on the SOS.

By assigning different precoders to different transmitters, the proposed algorithms can eliminate the higher dimensional ambiguity and are able to estimate the channel coefficient for different transmitters within a complex scalar ambiguity only. Strict conditions on identifiability are provided for zero-padding precoders, whereas identification for other precoding matrices is numerically tested. Other numerical examples on the performance of the proposed algorithms, as well as the comparison with the existing works, are also provided.

APPENDIX
PROOF OF THEOREM 2

An equivalent expression for (30) is

$$\mathbf{\Gamma}_{K+1-L}(\hat{\mathbf{H}}^l)\boldsymbol{\mathcal{F}} = \mathbf{\Gamma}_{K+1-L}(\mathbf{H}^l)\boldsymbol{\mathcal{F}}\boldsymbol{\mathcal{A}} \tag{41}$$

where $\boldsymbol{\mathcal{F}}$ is a row permutation matrix from $\tilde{\mathbf{F}}$. For convenience, we first provide the proof for $N_{\mathrm{t}} = 2$. Then, we assume that $\mathbf{F}_i$ is obtained by deleting the v_ith column from a $(K+1) \times (K+1)$ identity matrix, with $v_1 \neq v_2$, and we divide the discussion into two cases.

Case 1

$v_1 \leq \lfloor K/2 \rfloor + 1$, and $v_2 > \lfloor K/2 \rfloor + 1$. The matrices $\mathbf{\Gamma}_{K+1-L}(\hat{\mathbf{H}}^l)\boldsymbol{\mathcal{F}}$ and $\mathbf{\Gamma}_{K+1-L}(\mathbf{H}^l)\boldsymbol{\mathcal{F}}$ can be partitioned as

$$\mathbf{\Gamma}_{K+1-L}(\hat{\mathbf{H}}^l)\boldsymbol{\mathcal{F}} = \begin{bmatrix} \hat{\mathbf{P}}_1 & \mathbf{0} & \mathbf{0} \\ \hat{\mathbf{P}}_2 & \hat{\boldsymbol{\mathcal{H}}} & \hat{\mathbf{P}}_3 \\ \mathbf{0} & \mathbf{0} & \hat{\mathbf{P}}_4 \end{bmatrix} \tag{42}$$

$$\mathbf{\Gamma}_{K+1-L}(\mathbf{H}^l)\boldsymbol{\mathcal{F}} = \begin{bmatrix} \mathbf{P}_1 & \mathbf{0} & \mathbf{0} \\ \mathbf{P}_2 & \boldsymbol{\mathcal{H}} & \mathbf{P}_3 \\ \mathbf{0} & \mathbf{0} & \mathbf{P}_4 \end{bmatrix} \tag{43}$$

where $\hat{\mathbf{P}}_1$ and $\mathbf{P}_1$ are the $N_{\mathrm{r}}(\lfloor K/2 \rfloor - L) \times (2\lfloor K/2 \rfloor - 1)$ matrices, whereas $\hat{\mathbf{P}}_4$ and $\mathbf{P}_4$ are the $N_{\mathrm{r}}(\lfloor (K-1)/2 \rfloor - L) \times (2\lfloor (K-1)/2 \rfloor - 1)$ matrices. Define $\mathbf{h}_i^l$ as the ith column of $\mathbf{H}^l$, namely, $\mathbf{h}_i^l = [h_{i1,l}, h_{i2,l}, \ldots, h_{iN_{\mathrm{r}},l}]^{\mathrm{T}}$. Then, $\hat{\boldsymbol{\mathcal{H}}}$ and $\boldsymbol{\mathcal{H}}$ can be expressed as two $N_{\mathrm{r}}(L+2) \times 4$ matrices with the form

$$\hat{\boldsymbol{\mathcal{H}}} = \begin{bmatrix} \hat{\mathbf{h}}_1^0 & \mathbf{0} & \hat{\mathbf{h}}_2^0 & \mathbf{0} \\ \hat{\mathbf{h}}_1^1 & \hat{\mathbf{h}}_1^0 & \hat{\mathbf{h}}_2^1 & \hat{\mathbf{h}}_2^0 \\ \vdots & \vdots & \vdots & \vdots \\ \hat{\mathbf{h}}_1^L & \hat{\mathbf{h}}_1^{L-1} & \hat{\mathbf{h}}_2^L & \hat{\mathbf{h}}_2^{L-1} \\ \mathbf{0} & \hat{\mathbf{h}}_1^L & \mathbf{0} & \hat{\mathbf{h}}_2^L \end{bmatrix} \tag{44}$$

$$\boldsymbol{\mathcal{H}} = \begin{bmatrix} \mathbf{h}_1^0 & \mathbf{0} & \mathbf{h}_2^0 & \mathbf{0} \\ \mathbf{h}_1^1 & \mathbf{h}_1^0 & \mathbf{h}_2^1 & \mathbf{h}_2^0 \\ \vdots & \vdots & \vdots & \vdots \\ \mathbf{h}_1^L & \mathbf{h}_1^{L-1} & \mathbf{h}_2^L & \mathbf{h}_2^{L-1} \\ \mathbf{0} & \mathbf{h}_1^L & \mathbf{0} & \mathbf{h}_2^L \end{bmatrix}. \tag{45}$$

The $2\lfloor K/2 \rfloor$th to the $(2\lfloor K/2 \rfloor + 3)$th column of (41) could be rewritten as

$$\begin{bmatrix} \mathbf{0} \\ \hat{\boldsymbol{\mathcal{H}}} \\ \mathbf{0} \end{bmatrix} = \begin{bmatrix} \mathbf{P}_1 & \mathbf{0} & \mathbf{0} \\ \mathbf{P}_2 & \boldsymbol{\mathcal{H}} & \mathbf{P}_3 \\ \mathbf{0} & \mathbf{0} & \mathbf{P}_4 \end{bmatrix} \begin{bmatrix} \tilde{\boldsymbol{\mathcal{A}}}_1 \\ \tilde{\boldsymbol{\mathcal{A}}}_2 \\ \tilde{\boldsymbol{\mathcal{A}}}_3 \end{bmatrix}. \tag{46}$$

Note that $\mathbf{P}_1$ is obtained from the columns of $\mathbf{\Gamma}_{(\lfloor K/2 \rfloor - L)}(\mathbf{H}^l)$ and that $\mathbf{P}_4$ is obtained from the columns of $\mathbf{\Gamma}_{(\lfloor (K-1)/2 \rfloor - L)}(\mathbf{H}^l)$. From Lemma 1, we know that $\mathbf{P}_1$ and $\mathbf{P}_4$ are of full column rank if $\mathbf{\Gamma}_{(\lfloor K/2 \rfloor - L)}(\mathbf{H}^l)$ and $\mathbf{\Gamma}_{(\lfloor (K-1)/2 \rfloor - L)}(\mathbf{H}^l)$ are tall matrices, namely, $N_{\mathrm{r}}(\lfloor K/2 \rfloor - L) > 2\lfloor K/2 \rfloor$ and $N_{\mathrm{r}}(\lfloor (K-1)/2 \rfloor - L) > 2\lfloor (K-1)/2 \rfloor$. Then, $\tilde{\boldsymbol{\mathcal{A}}}_1$ and $\tilde{\boldsymbol{\mathcal{A}}}_3$ can be calculated as zero matrices. Therefore

$$\hat{\boldsymbol{\mathcal{H}}} = \boldsymbol{\mathcal{H}}\tilde{\boldsymbol{\mathcal{A}}}_2. \tag{47}$$

By properly rearranging the columns of $\hat{\boldsymbol{\mathcal{H}}}$, we obtain

$$\begin{bmatrix} \mathbf{0} & \hat{\mathbf{H}}^0 \\ \hat{\mathbf{H}}^0 & \hat{\mathbf{H}}^1 \\ \vdots & \vdots \\ \hat{\mathbf{H}}^{L-1} & \hat{\mathbf{H}}^L \\ \hat{\mathbf{H}}^L & \mathbf{0} \end{bmatrix} = \begin{bmatrix} \mathbf{0} & \mathbf{H}^0 \\ \mathbf{H}^0 & \mathbf{H}^1 \\ \vdots & \vdots \\ \mathbf{H}^{L-1} & \mathbf{H}^L \\ \mathbf{H}^L & \mathbf{0} \end{bmatrix} \begin{bmatrix} \boldsymbol{\mathcal{C}}_1 & \boldsymbol{\mathcal{C}}_2 \\ \boldsymbol{\mathcal{C}}_3 & \boldsymbol{\mathcal{C}}_4 \end{bmatrix}. \tag{48}$$

From the last block row, we can get

$$\mathbf{H}^L \boldsymbol{\mathcal{C}}_2 = \mathbf{0}. \tag{49}$$

Since $\mathbf{H}^l$ is, by assumption, a full-column-rank matrix, then $\boldsymbol{\mathcal{C}}_2$ is a zero matrix. By substituting this result into (48), it can be further derived that

$$\boldsymbol{\mathcal{C}}_1 = \boldsymbol{\mathcal{C}}_4, \quad \boldsymbol{\mathcal{C}}_3 = \mathbf{0}. \tag{50}$$

Taking $\mathbf{B} = \boldsymbol{\mathcal{C}}_1$ gives (31).

Case 2

$v_1 \leq \lfloor K/2 \rfloor + 1$, and $v_2 \leq \lfloor K/2 \rfloor + 1$. Matrices $\mathbf{\Gamma}(\hat{\mathbf{H}}^l)\boldsymbol{\mathcal{F}}$ and $\mathbf{\Gamma}(\mathbf{H}^l)\boldsymbol{\mathcal{F}}$ can be divided as

$$\hat{\tilde{\mathbf{H}}}\boldsymbol{\mathcal{F}} = \begin{bmatrix} \hat{\mathbf{Q}}_1 & \mathbf{0} & \mathbf{0} \\ \hat{\mathbf{Q}}_2 & \hat{\boldsymbol{\mathcal{H}}} & \hat{\mathbf{Q}}_3 \\ \mathbf{0} & \mathbf{0} & \hat{\mathbf{Q}}_4 \end{bmatrix} \tag{51}$$

$$\tilde{\mathbf{H}}\boldsymbol{\mathcal{F}} = \begin{bmatrix} \mathbf{Q}_1 & \mathbf{0} & \mathbf{0} \\ \mathbf{Q}_2 & \boldsymbol{\mathcal{H}} & \mathbf{Q}_3 \\ \mathbf{0} & \mathbf{0} & \mathbf{Q}_4 \end{bmatrix} \tag{52}$$

where $\hat{\mathbf{Q}}_1$ and $\mathbf{Q}_1$ are taken as the $N_{\mathrm{r}}(\lfloor K/2 \rfloor - L + 2) \times 2\lfloor K/2 \rfloor$ matrices, and $\hat{\mathbf{Q}}_4$ and $\mathbf{Q}_4$ are taken as the

$N_\mathrm{r}(\lfloor (K-1)/2 \rfloor - L - 1) \times (2\lfloor (K-1)/2 \rfloor - 2)$ matrices. $\hat{\boldsymbol{\mathcal{H}}}$ and $\boldsymbol{\mathcal{H}}$ are now $N_\mathrm{r}(L+1) \times 2$ matrices, with the form

$$\hat{\boldsymbol{\mathcal{H}}} = \begin{bmatrix} \hat{\mathbf{H}}^0 \\ \hat{\mathbf{H}}^1 \\ \vdots \\ \hat{\mathbf{H}}^L \end{bmatrix}, \quad \boldsymbol{\mathcal{H}} = \begin{bmatrix} \mathbf{H}^0 \\ \mathbf{H}^1 \\ \vdots \\ \mathbf{H}^L \end{bmatrix}. \tag{53}$$

The $(2\lfloor K/2 \rfloor + 1)$th and the $(2\lfloor K/2 \rfloor + 2)$th columns of (41) could be rewritten as

$$\begin{bmatrix} \mathbf{0} \\ \hat{\boldsymbol{\mathcal{H}}} \\ \mathbf{0} \end{bmatrix} = \begin{bmatrix} \mathbf{Q}_1 & \mathbf{0} & \mathbf{0} \\ \mathbf{Q}_2 & \boldsymbol{\mathcal{H}} & \mathbf{Q}_3 \\ \mathbf{0} & \mathbf{0} & \mathbf{Q}_4 \end{bmatrix} \begin{bmatrix} \bar{\boldsymbol{\mathcal{A}}}_1 \\ \bar{\boldsymbol{\mathcal{A}}}_2 \\ \bar{\boldsymbol{\mathcal{A}}}_3 \end{bmatrix}. \tag{54}$$

Note that $\mathbf{Q}_1$ is obtained from the columns of $\mathbf{\Gamma}_{(\lfloor K/2 \rfloor - L + 2)}(\mathbf{H}^l)$ and that $\mathbf{Q}_4$ is obtained from the columns of $\mathbf{\Gamma}_{(\lfloor (K-1)/2 \rfloor - L - 1)}(\mathbf{H}^l)$. From Lemma 1, we know that $\mathbf{Q}_1$ and $\mathbf{Q}_4$ are of full column rank if $N_\mathrm{r}(\lfloor K/2 \rfloor - L + 2) > 2\lfloor K/2 \rfloor + 4$ and $N_\mathrm{r}(\lfloor (K-1)/2 \rfloor - L - 1) > 2\lfloor (K-1)/2 \rfloor - 2$. Similar to case 1, we could arrive at (31).

For cases $\{v_1 \leq \lfloor K/2 \rfloor + 1, v_2 > \lfloor K/2 \rfloor + 1\}$ and $\{v_1 > \lfloor K/2 \rfloor + 1, v_2 > \lfloor K/2 \rfloor + 1\}$, a similar discussion can be made. By combining all cases, N_r should satisfy

$$N_\mathrm{r} > \left\lfloor \frac{2\lfloor (K-1)/2 \rfloor - 2}{(\lfloor (K-1)/2 \rfloor - L - 1)} \right\rfloor. \tag{55}$$

For $N_\mathrm{t} > 2$, basically, the proof can be divided into two cases.

1) All v_i's are smaller than $\lfloor K/2 \rfloor + 1$, or all v_i's are bigger than $\lfloor K/2 \rfloor + 1$. Then, similar equations as (52) and (53) can be obtained.
2) Otherwise, similar equations as (43) and (48) can be obtained.

The remaining discussion is the same as that for $N_\mathrm{t} = 2$. Note that $\mathbf{P}_1$ and $\mathbf{P}_4$ are still obtained from the columns of $\mathbf{\Gamma}_{(\lfloor K/2 \rfloor - L)}(\mathbf{H}^l)$ and $\mathbf{\Gamma}_{(\lfloor (K-1)/2 \rfloor - L)}(\mathbf{H}^l)$ by keeping in mind that column selection of $\mathbf{P}_1$ and $\mathbf{P}_4$ from $\mathbf{\Gamma}_{(\lfloor K/2 \rfloor - L)}(\mathbf{H}^l)$ and $\mathbf{\Gamma}_{(\lfloor (K-1)/2 \rfloor - L)}(\mathbf{H}^l)$ varies for different combinations of v_i. Meanwhile, $\mathbf{Q}_1$ is still obtained from the columns of $\mathbf{\Gamma}_{(\lfloor K/2 \rfloor - L + 2)}(\mathbf{H}^l)$, and $\mathbf{Q}_4$ is from the columns of $\mathbf{\Gamma}_{(\lfloor (K-1)/2 \rfloor - L - 1)}(\mathbf{H}^l)$. Therefore, N_r should satisfy

$$N_\mathrm{r} > \left\lfloor \frac{N_\mathrm{t}\left(\lfloor (K-1)/2 \rfloor - 1\right)}{(\lfloor (K-1)/2 \rfloor - L - 1)} \right\rfloor. \tag{56}$$

A detailed discussion for $N_\mathrm{t} > 2$ is omitted because it is quite straightforward.

References

[1] I. E. Telatar, *Capacity of Multi-Antenna Gaussian Channels*, 1995. Bell Labs Tech. Memo.

[2] G. J. Foschini and M. J. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," *Wirel. Pers. Commun.*, vol. 6, no. 3, pp. 311–335, Mar. 1998.

[3] L. Tong and S. Perreau, "Multichannel blind identification: From subspace to maximum likelihood methods," *Proc. IEEE*, vol. 86, no. 10, pp. 1951–1968, Oct. 1998.

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

[5] B. Muquet, Z. Wang, G. B. Giannakis, M. de Courville, and P. Duhamel, "Cyclic prefixing or zero padding for wireless multicarrier transmissions?" *IEEE Trans. Commun.*, vol. 50, no. 12, pp. 2136–2148, Dec. 2002.

[6] S. Roy and C. Li, "A subspace blind channel estimation method for OFDM systems without cyclic prefix," *IEEE Trans. Wireless Commun.*, vol. 1, no. 4, pp. 572–579, Oct. 2002.

[7] K. Wei and B. Champagne, "Subspace-based blind channel estimation: Generalization and performance analysis," *IEEE Trans. Signal Process.*, vol. 53, no. 3, pp. 1151–1162, Mar. 2005.

[8] D. Reynolds, X. Wang, and H. V. Poor, "Blind adaptive space-time multiuser detection with multiple transmitter and receiver antennas," *IEEE Trans. Signal Process.*, vol. 50, no. 6, pp. 1261–1276, Jun. 2002.

[9] Y. Hua and J. K. Tugnait, "Blind identifiability of FIR-MIMO systems with colored inputs using second order statistics," *IEEE Signal Process. Lett.*, vol. 7, no. 12, pp. 348–350, Dec. 2000.

[10] H. Luo, R. Liu, X. T. Ling, and X. Li, "The autocorrelation matching method for distributed MIMO communications over unknown FIR channels," in *Proc. IEEE ICASSP*, Salt Lake City, UT, May 2001, vol. 4, pp. 2161–2164.

[11] S. Visuri and V. Koivunen, "Resolving ambiguities in subspace-based blind receiver for MIMO channels," in *Proc. 36th Asilomar Conf.*, Pacific Grove, CA, Nov. 2002, vol. 1, pp. 589–593.

[12] I. Bradaric, A. P. Petropulu, and K. I. Diamantaras, "Blind MIMO FIR channel identification based on second order spectra correlations," *IEEE Trans. Signal Process.*, vol. 51, no. 6, pp. 1668–1674, Jun. 2003.

[13] H. Bolcskei, R. W. Heath, Jr., and A. J. Paulraj, "Blind channel estimation in spatial multiplexing systems using nonredundant antenna precoding," in *Proc. 33th Asilomar Conf.*, Pacific Grove, CA, Oct. 1999, vol. 2, pp. 1127–1132.

[14] A. Medles and D. T. M. Slock, "Linear precoding for spatial multiplexing MIMO systems: Blind channel estimation aspects," in *Proc. IEEE ICC*, New York, Apr. 28–May 2, 2002, vol. 1, pp. 401–405.

[15] H. Luo and R. W. Liu, "Blind equalization for MIMO for channels based only on second order statistics by use of pre-filters," in *Proc. IEEE SPAWC*, Annapolis, MD, May 1999, pp. 106–109.

[16] X. G. Xia, W. Su, and H. Liu, "Filterbank precoders for blind equalization: Polynomial ambiguity resistant precoders (PARP)," *IEEE Trans. Circuits Syst. I*, vol. 48, no. 2, pp. 19–209, Feb. 2001.

[17] F. Gao and A. Nallanathan, "Blind channel estimation for MIMO-OFDM systems via a nonredundant linear precoding," *IEEE Trans. Signal Process.*, vol. 55, no. 2, pp. 784–789, Jan. 2007.

[18] Z.-Q. Luo, T. N. Davidson, G. B. Giannakis, and K. M. Wong, "Transceiver optimization for block-based multiple access through ISI channels," *IEEE Trans. Signal Process.*, vol. 52, no. 4, pp. 1037–1052, Apr. 2004.

[19] D. J. Love and R. W. Heath, Jr., "Limited feedback unitary precoding for orthogonal space-time block codes," *IEEE Trans. Signal Process.*, vol. 53, no. 1, pp. 64–73, Jan. 2005.

[20] Y. Li and K. J. R. Liu, "Blind identification and equalization for multiple-input/multiple-output channels," in *Proc. GLOBECOM*, London, U.K., Nov. 1996, vol. 3, pp. 1789–1793.

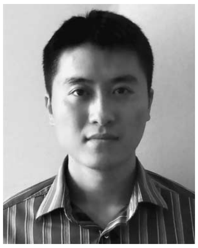
[21] A. Safavi and K. Abed-Meraim, "Orthogonal minimum noise subspace for multiple-input multiple-output system identification," in *Proc. IEEE SSP*, Singapore, Aug. 2001, pp. 285–288.

[22] S. Zhou, B. Muquet, and G. B. Giannakis, "Subspace-based (semi-) blind channel estimation for block precoded space-time OFDM," *IEEE Trans. Signal Process.*, vol. 50, no. 5, pp. 1215–1228, May 2002.

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

[24] O. Shuichi and G. B. Giannakis, "Optimal training and redundant precoding for block transmissions with application to wireless OFDM," *IEEE Trans. Commun.*, vol. 50, no. 12, pp. 2113–2123, Dec. 2002.

[25] G. B. Giannakis, Y. Hua, P. Stoica, and L. Tong, *Signal Processing Advances in Wireless and Mobile Communications*, vol. I. Englewood Cliffs, NJ: Prentice-Hall, Sep. 2000.



Feifei Gao (S'05) received the B.Eng. degree in information engineering from Xi'an Jiaotong University, Xi'an, China, in 2002 and the M.Sc. degree from the McMaster University, Hamilton, ON, Canada, in 2004. He is currently working toward the Ph.D. degree with the Department of Electrical and Computer Engineering, National University of Singapore, Singapore.

His research interests are in communication theory, broadband wireless communications, signal processing for communications, multiinput–multioutput systems, and array signal processing.

Mr. Gao was a recipient of a President Scholarship from the National University of Singapore.

Arumugam Nallanathan (S'97–M'00–SM'05) received the B.Sc. degree in electrical engineering (with honors) from the University of Peradeniya, Peradeniya, Sri Lanka, in 1991, the Certificate of Postgraduate Study in electrical engineering from Cambridge University, Cambridge, U.K., in 1994, and the Ph.D. degree in electrical engineering from the University of Hong Kong, Kowloon, Hong Kong, in 2000.

Since then, he has been an Assistant Professor with the Department of Electrical and Computer Engineering, National University of Singapore, Singapore. His research interests include orthogonal frequency-division-multiplexing systems, ultrawideband (UWB) communication and localization, multiinput–multioutput systems, and cooperative diversity techniques. He has published over 100 journal and conference papers in these areas.

Dr. Nallanathan currently serves on the editorial board of the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, the IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY, *Wireless Communications and Mobile Computing*, and the *EURASIP Journal of Wireless Communications and Networking* as an Associate Editor. He also serves as a Guest Editor for the *EURASIP Journal of Wireless Communications and Networking: Special Issue on UWB Communication Systems—Technology and Applications*. He served as a Technical Program Cochair and as a Technical Program Committee Member for more than 25 IEEE international conferences. He currently serves as the General Track Chair for IEEE VTC 2008-Spring and a Cochair for the IEEE GLOBECOM 2008 Signal Processing for Communications Symposium and the IEEE ICC 2009 Wireless Communications Symposium. He is a corecipient of the Best Paper Award presented at the 2007 IEEE International Conference on Ultra-Wideband.