Blind Channel Estimation for MIMO-OFDM Systems

Changyong Shin, *Member, IEEE*, Robert W. Heath, Jr., *Senior Member, IEEE*, and Edward J. Powers, *Fellow, IEEE*

***Abstract*—By combining multiple-input multiple-output (MIMO) communication with the orthogonal frequency division multiplexing (OFDM) modulation scheme, MIMO-OFDM systems can achieve high data rates over broadband wireless channels. In this paper, to provide a bandwidth-efficient solution for MIMO-OFDM channel estimation, we establish conditions for channel identifiability and present a blind channel estimation technique based on a subspace approach. The proposed method unifies and generalizes the existing subspace-based methods for blind channel estimation in single-input single-output OFDM systems to blind channel estimation for two different MIMO-OFDM systems distinguished according to the number of transmit and receive antennas. In particular, the proposed method obtains accurate channel estimation and fast convergence with insensitivity to overestimates of the true channel order. If virtual carriers (VCs) are available, the proposed method can work with no or insufficient cyclic prefix (CP), thereby potentially increasing channel utilization. Furthermore, it is shown under specific system conditions that the proposed method can be applied to MIMO-OFDM systems without CPs, regardless of the presence of VCs, and obtains an accurate channel estimate with a small number of OFDM symbols. Thus, this method improves the transmission bandwidth efficiency. Simulation results illustrate the mean-square error performance of the proposed method via numerical experiments.**

***Index Terms*—Blind channel estimation, cyclic prefix, multiple-input multiple-output (MIMO) system, orthogonal frequency division multiplexing (OFDM), subspace method, virtual carrier.**

I. Introduction

RECENTLY, increasing interest has been concentrated on modulation techniques that provide high data rates over broadband wireless channels for applications, including wireless multimedia, wireless Internet access, and future-generation mobile communication systems. Orthogonal frequency division multiplexing (OFDM) is a promising digital modulation scheme to simplify the equalization in frequency-selective channels [1], [2]. The main benefit is that it simplifies implementation, and it is robust against the frequency-selective fading channels. Multiple-input multiple-output (MIMO) communication, enabled by multiple transmit and receive antennas, can increase significantly the channel capacity (see, e.g., [3]–[6] and references therein). Thus, MIMO-OFDM systems, which combine the OFDM with MIMO communication, can provide a high-performance transmission (see, e.g., [3], [4], and [7]–[9] and references therein).

In a MIMO-OFDM system, a coherent signal detection requires a reliable estimate of the channel impulse responses between the transmit and receive antennas. These channels can be estimated by sending training sequences. The training requirements, however, are significant [10]–[18]. Furthermore, transmitting the training sequences is undesirable for certain communication systems [19], [20]. Thus, blind channel estimation for MIMO-OFDM systems has been an active area of research in recent years. Zhou *et al.* [21] proposed a subspace-based blind channel estimation method for space-time coded MIMO-OFDM systems using properly designed redundant linear precoding and the noise subspace method [22]–[24]. Bölcskei *et al.* [25] proposed an algorithm for blind channel estimation and equalization for MIMO-OFDM systems using second-order cyclostationary statistics induced by employing a periodic nonconstant-modulus antenna precoding. Yatawatta and Petropulu [26] presented a blind channel estimation method based on a nonredundant linear block precoding and cross-correlation operations. Zeng and Ng [27] proposed a subspace technique based on the noise subspace method for estimating the MIMO channels in the uplink of multiuser multiantenna zero-padding OFDM systems [28], [29].

A variety of second-order statistics (SOS)-based blind estimators (see, e.g., [22]–[24], and [30]–[51] and references therein) have been presented since Tong *et al.* [52] introduced a SOS-based technique for the blind identification of single-input multiple-output systems. Among those methods, the noise subspace method is believed to be one of the most promising due to its simple structure and good performance. Thus, by exploiting the fundamental structure of the noise subspace method, subspace methods [53]–[55] for single-input single-output (SISO) OFDM systems have been proposed and achieved good estimation performance. Muquet *et al.* [53] developed a subspace method for SISO-OFDM systems by utilizing the redundancy introduced by the cyclic prefix (CP) insertion, and derived a condition for channel identifiability. For shaping of the transmit spectrum, practical OFDM systems are not fully loaded [56]. The subcarriers that are set to zero without any information are referred to as virtual carriers (VCs) [57]. Other than the CP, the presence of VCs provides another useful resource that can be used for channel estimation. Li and Roy [54] proposed a subspace blind channel estimator for SISO-OFDM systems by considering the existence of VCs and provided a condition for

Manuscript received April 15, 2005; revised April 19, 2006 and April 25, 2006. This paper was presented in part at the IEEE Global Telecommunications Conference, Dallas, TX, November 2004. The review of this paper was coordinated by Prof. Z. Wang.

C. Shin is with the Communications and Network Laboratory, Samsung Advanced Institute of Technology (SAIT), 446-712 Korea (e-mail: cyshin3@dreamwiz.com).

R. W. Heath, Jr. and E. J. Powers are with the Wireless Networking and Communications Group (WNCG), Department of Electrical and Computer Engineering, The University of Texas at Austin, Austin, TX 78712-0240 USA (e-mail: rheath@mail.utexas.edu; ejpowers@mail.utexas.edu).

Digital Object Identifier 10.1109/TVT.2007.891429

channel identifiability. They showed that their estimator based on the noise subspace method can achieve better estimation performance than other blind estimation techniques [55], [58]. In particular, they demonstrated that the CPs are more useful for their estimator than VCs. In [54], however, the reason why the utilization of CPs rather than VCs increases the accuracy of channel estimation was not discussed. In addition, it was not considered how the channel estimation performance can be improved in cases with no or insufficient CPs.

In this paper, we unify and generalize the SISO-OFDM subspace methods in [53] and [54] to the case of blind channel estimation for spatial multiplexing MIMO-OFDM systems with any number of transmit and receive antennas. We present a new blind channel estimator based on the noise subspace method and establish conditions for blind channel identification in spatially multiplexed MIMO-OFDM systems. The proposed method works regardless of the presence of VCs, can use as little as one received OFDM symbol for a filtering matrix, and operates with any number of transmit or receive antennas. Considering the presence of VCs, the proposed method can be applied to MIMO-OFDM systems without CPs, where blind estimation techniques based on CPs cannot be employed, thereby providing the systems with the potential to achieve a higher channel utilization. For MIMO-OFDM systems with CPs, the proposed method can provide an additional performance gain with respect to the existing blind channel estimation methods. We provide numerical results that illustrate tradeoffs in mean-square error as a function of the CP length, number of VCs, and number of OFDM symbols used in the estimate.

Compared with [25] and [26], we do not use a transmit precoding to aid our blind channel estimator. Furthermore, by virtue of our subspace approach, we need fewer OFDM symbols to obtain a reliable estimate. Our estimator, however, has an additional full rank requirement that is not present in the precoding-based methods. The combination of our estimator and the precoding-based methods is an interesting topic for future work. Our approach generalizes the studies in [53] and [54] to operate with multiple transmit and multiple receive antennas under the assumption that spatial multiplexing is used at the transmitter. With one transmit and receive antenna, our approach and identifiability conditions are simplified to those presented in [53] and [54]. A subspace-based method for MIMO-OFDM systems with spatial multiplexing was proposed in [59]. Compared with [59], we also consider the case of excess transmit antennas. Furthermore, the identifiability conditions provided in [59] are not complete. Specifically, their full rank requirement does not appear to be sufficient, and the channel ambiguity condition does not appear to be comprehensive.

The rest of this paper is organized as follows. In Section II, we briefly describe a MIMO-OFDM system model. In Section III, we establish conditions for the MIMO-OFDM channel identifiability by generalizing the conditions presented in [54] and develop a blind channel estimation scheme based on the noise subspace method. Section IV contains simulation results demonstrating the performance of the proposed method. Finally, a conclusion is provided in Section V.

The notations used in this paper are as follows. Matrices and vectors are denoted by symbols in boldface, and $(\cdot)^*$, $(\cdot)^{\mathrm{T}}$, and $(\cdot)^{\mathrm{H}}$ represent complex conjugate, transpose, and Hermitian, respectively. $\mathrm{rank}(\mathbf{X})$ and $\mathrm{span}(\mathbf{X})$ mean the rank of a matrix $\mathbf{X}$ and the subspace spanned by the column vectors of a matrix $\mathbf{X}$, respectively. $*$ and $\otimes$ stand for the convolution and the Kronecker product, respectively. w_N is equal to $e^{j2\pi/N}$. $\mathbf{I}_m$ denotes the $m \times m$ identity matrix, and $\mathbf{0}$ stands for the all-zeros matrix of appropriate dimensions. $\mathrm{diag}(\mathbf{x})$ denotes a diagonal matrix with $\mathbf{x}$ on its main diagonal. $E\{\cdot\}$ denotes statistical expectation. $\mathbf{x}[1:k]$ denotes the first k elements of a vector $\mathbf{x}$. $\mathbf{X}[i_1:i_2, j_1:j_2]$ denotes a submatrix obtained by extracting rows i_1 through i_2 and columns j_1 through j_2 from a matrix $\mathbf{X}$. If no specific range appears at the row or column position in notation, then all rows or columns are considered to constitute the submatrix. $[\mathbf{X}]_{i,j}$ denotes the (i,j)th element of a matrix $\mathbf{X}$. $\lfloor x \rfloor$ is the largest integer less than or equal to x. Also, $\lceil x \rceil$ indicates the nearest integer that is not smaller than x. $\langle x \rangle_y$ means the integer remainder after x is divided by y. $\det(\mathbf{X})$ denotes the determinant of a square matrix $\mathbf{X}$. $\|\cdot\|_2$ and $\|\cdot\|_F$ mean the l^2-norm and the Frobenius matrix norm, respectively. $\mathcal{CN}(0,\sigma^2)$ denotes a circular symmetric complex Gaussian distribution with zero mean and variance σ^2.

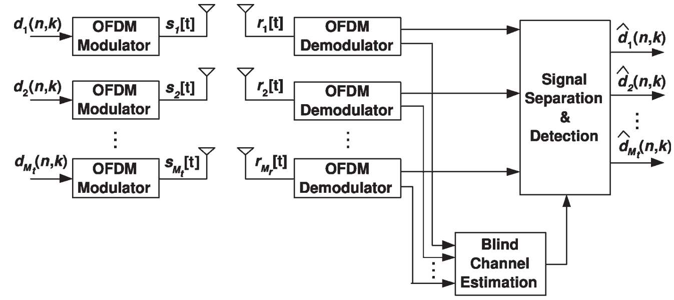


Fig. 1. MIMO-OFDM system model with M_t transmit and M_r receive antennas.

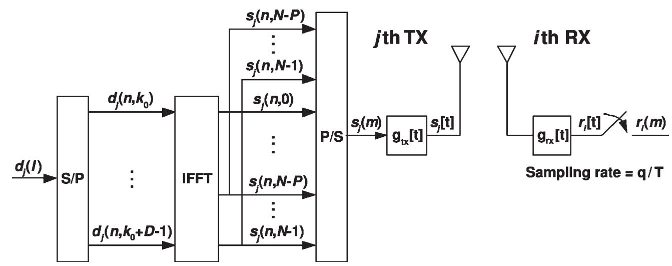


Fig. 2. Baseband OFDM system with virtual carriers for the jth transmit antenna and the ith receive antenna.

II. MIMO-OFDM SYSTEM MODEL

In this section, we describe the MIMO-OFDM system model with M_t transmit and M_r receive antennas considered in this paper, as illustrated in Fig. 1. Fig. 2 shows the baseband model of an OFDM system for the jth transmit and ith receive antennas in Fig. 1, which has N subcarriers and uses the subcarriers numbered k_0 to $k_0 + D - 1$ for information data. The remaining $N - D$ unmodulated carriers are referred to as VCs that are needed for the input signal pulse shaping by a transmit filter such as the raised cosine filter with a roll-off factor [57], [60]. If we set k_0 to 0 and D to N, the system no longer has VCs. Thus, our system model can be applied to both systems with and without VCs. Let the nth block of the frequency-domain information symbols in the jth transmit antenna be written as

$$\mathbf{d}_j(n) = [d_j(n,k_0),\ d_j(n,k_0+1),\ldots,d_j(n,k_0+D-1)]^{\mathrm{T}} \tag{1}$$

where the subscript j is the transmit antenna index with $1 \leq j \leq M_t$. Assuming that the length of the CP is P, each OFDM modulator adds $N-D$ zeros for VCs to the data block in (1), applies an N-point inverse fast Fourier transform (IFFT) to this block, and inserts the CP in front of the IFFT output vector, which is a copy of the last P samples of the IFFT output. This results in the time-domain sample vector of the nth OFDM symbol written as

$$\begin{aligned}\mathbf{s}_j(n) = [&s_j(n,N-P),\ldots,s_j(n,N-1),\\ &s_j(n,0),\ldots,s_j(n,N-1)]^{\mathrm{T}}.\end{aligned} \tag{2}$$

To generate the continuous-time signal to be sent on the channel, each element in the vector $\mathbf{s}_j(n)$ is pulse-shaped by a transmit filter $g_{tx}[t]$

$$s_j[t] = \sum_{n=-\infty}^{\infty}\sum_{k=0}^{Q-1} s_j\left(n,\langle N-P+k\rangle_N\right) g_{tx}\left[t-(k+nQ)T\right] \tag{3}$$

where $Q = N+P$, and T is the sample duration in the time domain. By denoting $s_j(n,\langle N-P+k\rangle_N)$, meaning the kth sample of the nth OFDM symbol in the time domain, as $s_j((k+nQ)T)$, and $k+nQ$ as a, the transmitted signal $s_j[t]$ can be concisely expressed as

$$s_j[t] = \sum_{a=-\infty}^{\infty} s_j(aT) g_{tx}[t-aT]. \tag{4}$$

Then, M_t transmit antennas simultaneously transmit the signals $s_1[t], s_2[t],\ldots,s_{M_t}[t]$.

During the transmission, the transmitted signal $s_j[t]$ from the jth transmit antenna passes through a dispersive channel with an impulse response $c_{ij}[t]$, it gets corrupted by a spatially uncorrelated additive white Gaussian noise $v_i[t]$, and it finally enters into a front-end receive filter $g_{rx}[t]$ at the ith receive antenna. When we denote the composite channel impulse response between the jth transmit antenna and the ith receive antenna as $h_{ij}[t] = g_{tx}[t] * c_{ij}[t] * g_{rx}[t]$, and the filtered noise at the ith receive antenna as $\eta_i[t] = v_i[t] * g_{rx}[t]$, the received signal $r_i[t]$ at the ith receive antenna is expressed as

$$r_i[t] = \sum_{j=1}^{M_t}\sum_{a=-\infty}^{\infty} s_j(aT)h_{ij}[t-aT] + \eta_i[t]. \tag{5}$$

We suppose that the composite channel impulse responses $h_{ij}[t]$ have the finite support $[0,(L+1)T)$ with $L \leq P$, which guarantee that $r_i[t]$ is not contaminated by previous OFDM symbols. By sampling $r_i[t]$ at a rate $1/T_s = q/T$ with a positive integer q, the sampled received signal $r_i[\epsilon_i + mT_s] = r_i[\epsilon_i + (mT/q)]$ at the ith receive antenna is given as

$$\begin{aligned} r_i\left[\epsilon_i + \frac{mT}{q}\right] &= \sum_{j=1}^{M_t}\sum_{a=\lfloor m/q\rfloor - L}^{\lfloor m/q\rfloor} s_j(aT) \\ &\quad\times h_{ij}\left[\epsilon_i + \frac{mT}{q} - aT\right] + \eta_i\left[\epsilon_i + \frac{mT}{q}\right] \\ &= \sum_{j=1}^{M_t}\sum_{l=0}^{L} s_j\left(\left\lfloor\frac{m}{q}\right\rfloor T - lT\right) \\ &\quad\times h_{ij}\left[\epsilon_i + \frac{\langle m\rangle_q T}{q} + lT\right] + \eta_i\left[\epsilon_i + \frac{mT}{q}\right] \end{aligned} \tag{6}$$

where $\epsilon_i \in [0, T_s)$ is the sample timing error at the ith receive antenna.

III. Subspace-Based Blind Channel Estimation

In this section, we establish conditions for the channels to be identifiable and present a subspace method for blind channel estimation for two MIMO-OFDM system structures. We distinguish between two different MIMO-OFDM system structures: one with $M_t \leq M_r$ and the other with $M_t > M_r$.

A. MIMO-OFDM System With $M_t \leq M_r$

Based on the systems illustrated in Figs. 1 and 2, we denote the information symbols before an OFDM modulation as

$$\mathbf{d}(n,k) = [d_1(n,k),\ d_2(n,k),\ldots,d_{M_t}(n,k)]^{\mathrm{T}} \tag{7}$$

$$\mathbf{d}_n = \left[\mathbf{d}(n,k_0)^{\mathrm{T}},\ \mathbf{d}(n,k_0+1)^{\mathrm{T}},\ldots,\mathbf{d}(n,k_0+D-1)^{\mathrm{T}}\right]^{\mathrm{T}} \tag{8}$$

where $d_j(n,k)$ is an information symbol loaded on the kth subcarrier in the nth OFDM symbol to be transmitted from the jth transmit antenna. By collecting J consecutive OFDM symbols from M_t transmit antennas, the information symbol vector $\mathbf{d}(n)$ is constructed as given in

$$\mathbf{d}(n) = \left[\mathbf{d}_n^{\mathrm{T}}, \mathbf{d}_{n-1}^{\mathrm{T}},\ldots,\mathbf{d}_{n-J+1}^{\mathrm{T}}\right]^{\mathrm{T}}. \tag{9}$$

When we define the matrices $\mathbf{W}(i)$, $\mathbf{W}$, and $\boldsymbol{\mathcal{W}}$ associated with the IFFT as, respectively

$$\mathbf{W}(i) \triangleq \frac{1}{\sqrt{N}}\left[w_N^{ik_0},\ w_N^{i(k_0+1)},\ldots,w_N^{i(k_0+D-1)}\right] \tag{10}$$

$$\begin{aligned}\mathbf{W} \triangleq \Big[&\mathbf{W}(N-1)^{\mathrm{T}},\ldots,\mathbf{W}(0)^{\mathrm{T}},\\ &\mathbf{W}(N-1)^{\mathrm{T}},\ldots,\mathbf{W}(N-P)^{\mathrm{T}}\Big]^{\mathrm{T}}\end{aligned} \tag{11}$$

$$\boldsymbol{\mathcal{W}} \triangleq \mathbf{I}_J \otimes \mathbf{W} \otimes \mathbf{I}_{M_t} \tag{12}$$

and denote the time-domain signal vector $\mathbf{s}(n)$ to be transmitted after the OFDM modulation as

$$\mathbf{s}(n,k) = [s_1(n,k), s_2(n,k), \dots, s_{M_t}(n,k)]^{\mathrm{T}} \tag{13}$$

$$\mathbf{s}_n = \left[\mathbf{s}(n,N-1)^{\mathrm{T}}, \dots, \mathbf{s}(n,0)^{\mathrm{T}}, \mathbf{s}(n,N-1)^{\mathrm{T}}, \dots, \mathbf{s}(n,N-P)^{\mathrm{T}}\right]^{\mathrm{T}} \tag{14}$$

$$\mathbf{s}(n) = \left[\mathbf{s}_n^{\mathrm{T}}, \mathbf{s}_{n-1}^{\mathrm{T}}, \dots, \mathbf{s}_{n-J+1}^{\mathrm{T}}\right]^{\mathrm{T}} \tag{15}$$

we obtain the relationship given as

$$\mathbf{s}(n) = \boldsymbol{\mathcal{W}}\mathbf{d}(n). \tag{16}$$

In (13), $s_j(n,k)$ means the $(k+P+1)$th element of the vector $\mathbf{s}_j(n)$ in (2).

By sampling a received signal at each receive antenna with a rate $1/T$, meaning $q=1$ in (6), we can consider the discrete composite channels as given in (6) instead of continuous channels at M_r receive antennas. We assume that the discrete composite channels between M_t transmit antennas and M_r receive antennas are modeled as an $M_r \times M_t$ finite impulse response filter with L as the upper bound on the orders of these channels. When we denote $h_{ij}[\epsilon_i + lT]$ in (6) as $h_{ij}(l)$, and the lth lag of the MIMO channel as

$$\mathbf{h}(l) = \begin{bmatrix} h_{11}(l) & h_{12}(l) & \cdots & h_{1M_t}(l) \\ h_{21}(l) & h_{22}(l) & \cdots & h_{2M_t}(l) \\ \vdots & \vdots & \vdots & \vdots \\ h_{M_r1}(l) & h_{M_r2}(l) & \cdots & h_{M_rM_t}(l) \end{bmatrix} \tag{17}$$

respectively, the matrix transfer function $\mathbf{H}(z)$ is given as

$$\mathbf{H}(z) = \sum_{l=0}^{L} \mathbf{h}(l) z^{-l}. \tag{18}$$

Denoting $r_i[\epsilon_i + mT]$ in (6) as $r_i(m)$, and rearranging $r_i(m)$ according to $r_i(n,k) = r_i(k+nQ)$, we express the received signal at M_r receive antennas as

$$\mathbf{r}(n,k) = [r_1(n,k), r_2(n,k), \dots, r_{M_r}(n,k)]^{\mathrm{T}} \tag{19}$$

$$\mathbf{r}_n = \left[\mathbf{r}(n,Q-1)^{\mathrm{T}}, \mathbf{r}(n,Q-2)^{\mathrm{T}}, \dots, \mathbf{r}(n,0)^{\mathrm{T}}\right]^{\mathrm{T}}. \tag{20}$$

By collecting J consecutively received OFDM symbols, the received signal vector $\mathbf{r}(n)$ is given as

$$\mathbf{r}(n) = \left[\mathbf{r}_n^{\mathrm{T}}, \mathbf{r}_{n-1}^{\mathrm{T}}, \dots, \mathbf{r}_{n-J+1}\left[1:(Q-L)M_r\right]^{\mathrm{T}}\right]^{\mathrm{T}}. \tag{21}$$

Similarly, denoting $\eta_i[\epsilon_i + mT]$ in (6) as $\eta_i(m)$, and rearranging $\eta_i(m)$ according to $\eta_i(n,k) = \eta_i(k+nQ)$, we write the additive noise at M_r receive antennas as

$$\boldsymbol{\eta}(n,k) = [\eta_1(n,k), \eta_2(n,k), \dots, \eta_{M_r}(n,k)]^{\mathrm{T}} \tag{22}$$

$$\boldsymbol{\eta}_n = \left[\boldsymbol{\eta}(n,Q-1)^{\mathrm{T}}, \boldsymbol{\eta}(n,Q-2)^{\mathrm{T}}, \dots, \boldsymbol{\eta}(n,0)^{\mathrm{T}}\right]^{\mathrm{T}} \tag{23}$$

$$\boldsymbol{\eta}(n) = \left[\boldsymbol{\eta}_n^{\mathrm{T}}, \boldsymbol{\eta}_{n-1}^{\mathrm{T}}, \dots, \boldsymbol{\eta}_{n-J+1}[1:(Q-L)M_r]^{\mathrm{T}}\right]^{\mathrm{T}}. \tag{24}$$

When we define a $(JQ-L)M_r \times JQM_t$ channel matrix $\boldsymbol{\mathcal{H}}$ as

$$\boldsymbol{\mathcal{H}} \triangleq \begin{bmatrix} \mathbf{h}(0) & \cdots & \mathbf{h}(L) & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{h}(0) & \cdots & \mathbf{h}(L) & \cdots & \mathbf{0} \\ \vdots & & \ddots & & \ddots & \\ \mathbf{0} & \cdots & \mathbf{0} & \mathbf{h}(0) & \cdots & \mathbf{h}(L) \end{bmatrix} \tag{25}$$

the received signal vector $\mathbf{r}(n)$ in (21) can be written in a matrix form as

$$\mathbf{r}(n) = \boldsymbol{\mathcal{H}}\mathbf{s}(n) + \boldsymbol{\eta}(n) = \boldsymbol{\mathcal{H}}\boldsymbol{\mathcal{W}}\mathbf{d}(n) + \boldsymbol{\eta}(n) \triangleq \boldsymbol{\Xi}\mathbf{d}(n) + \boldsymbol{\eta}(n). \tag{26}$$

By assuming that the Nyquist pulse shaping is employed, $\boldsymbol{\eta}(n)$ is considered as a spatially and temporally uncorrelated complex Gaussian noise vector with the zero mean vector and the covariance matrix $\sigma_\eta^2 \mathbf{I}_{(JQ-L)M_r}$.

For the MIMO channel to be identified by the noise subspace method [22], the matrix $\boldsymbol{\Xi}$ in (26) should have a full column rank. The following Theorem 1 gives a necessary and sufficient condition for the full column rank requirement.

Theorem 1: In the case of $M_t \le M_r$ and $L \le (Q-D)$, the matrix $\boldsymbol{\Xi}$ in (26) has a full column rank, if and only if $\mathrm{rank}(\mathbf{H}(w_N^i)) = M_t$ for all $i \in \{k\}_{k=k_0}^{k_0+D-1}$.

Proof: Refer to Appendix I. ■

This theorem generalizes the identifiability condition for SISO-OFDM systems in [54] to the condition for MIMO-OFDM systems. As we can see from the above theorem, the identification of the MIMO-OFDM channel based on the noise subspace method requires the frequency-domain MIMO channel matrices at the subcarriers, where information symbols are loaded, to have a full column rank. In addition, the identifiability condition needs an upper bound for the MIMO channel order rather than accurate knowledge of the MIMO channel order. Since the length of the CP is usually set to be greater than the channel delay spread in the practical MIMO-OFDM systems, we can consider the CP length as an upper bound of the MIMO channel order.

In our derivation, we consider the MIMO channels satisfying the condition of $\mathrm{rank}(\mathbf{H}(w_N^i)) = M_t$ for all $i \in \{k\}_{k=k_0}^{k_0+D-1}$ as stated in Theorem 1. In addition, we suppose that the additive noise is uncorrelated with the transmitted signal, and the autocorrelation matrix $\mathbf{R}_{\mathbf{dd}} = E\{\mathbf{d}(n)\mathbf{d}(n)^{\mathrm{H}}\}$ of the information symbol vector $\mathbf{d}(n)$ in (9) has full rank. When the autocorrelation matrix $\mathbf{R}_{\mathbf{rr}} = E\{\mathbf{r}(n)\mathbf{r}(n)^{\mathrm{H}}\}$ of the received signal vector $\mathbf{r}(n)$ in (21) is diagonalized through the eigenvalue decomposition, we can partition the eigenvectors $\mathbf{U}$ into the vectors $\mathbf{U}_s$ spanning a signal subspace $\mathrm{span}(\mathbf{U}_s)$ and the vectors $\mathbf{U}_n$ spanning a noise subspace $\mathrm{span}(\mathbf{U}_n)$ [22] as

$$\mathbf{U} = [\mathbf{U}_s | \mathbf{U}_n] = \left[\mathbf{u}_1, \dots, \mathbf{u}_{JDM_t} | \mathbf{u}_{JDM_t+1}, \dots, \mathbf{u}_{(JQ-L)M_r}\right]. \tag{27}$$

Since $\mathrm{span}(\boldsymbol{\Xi})$ and $\mathrm{span}(\mathbf{U}_s)$ share the same JDM_t-dimensional space and are orthogonal to $\mathrm{span}(\mathbf{U}_n)$, we have the following orthogonal relationship [22]:

$$\mathbf{u}_k^{\mathrm{H}} \boldsymbol{\Xi} = \mathbf{0} \text{ for all } k \in \{n\}_{n=JDM_t+1}^{(JQ-L)M_r}. \tag{28}$$

Let us define the $(L+1)M_r \times 1$ channel response vector $\mathbf{h}_i$ associated with the channel impulse responses between the ith transmit antenna and M_r receive antennas, and the channel coefficient matrix $\mathbf{H}$ consisting of $\mathbf{h}_i$ as, respectively

$$\mathbf{h}_i \triangleq \left[\mathbf{h}(0)[:,i]^{\mathrm{T}},\ \mathbf{h}(1)[:,i]^{\mathrm{T}},\ldots,\mathbf{h}(L)[:,i]^{\mathrm{T}}\right]^{\mathrm{T}}, \quad 1 \le i \le M_t \tag{29}$$

$$\mathbf{H} \triangleq [\mathbf{h}_1,\mathbf{h}_2,\ldots,\mathbf{h}_{M_t}] = \left[\mathbf{h}(0)^{\mathrm{T}},\mathbf{h}(1)^{\mathrm{T}},\ldots,\mathbf{h}(L)^{\mathrm{T}}\right]^{\mathrm{T}}. \tag{30}$$

Under the appropriate conditions detailed in Theorem 2 and Lemma 1 to be given below, the noise subspace can determine the channel coefficient matrix $\mathbf{H}$ up to an $M_t \times M_t$ multiplicative matrix associated with the number of transmit antennas.

Let $\mathbf{H}'$ be a matrix that has the same dimension as that of $\mathbf{H}$. Let $\boldsymbol{\mathcal{H}}'$ be a nonzero matrix constructed from $\mathbf{H}'$ in the same manner as the matrix $\boldsymbol{\mathcal{H}}$ is constructed from $\mathbf{H}$. In addition, we denote $\boldsymbol{\mathcal{H}}'\boldsymbol{\mathcal{W}}$ as $\boldsymbol{\Xi}'$, and $\sum_{l=0}^{L}\mathbf{h}'(l)z^{-l}$ as $\mathbf{H}'(z)$. By using these notations, we state Theorem 2 and Lemma 1 associated with the ambiguity of an estimated MIMO channel as follows.

Theorem 2: Assume that the matrix $\boldsymbol{\Xi}$ in (26) has a full column rank with $J \ge 2$, $M_r \ge M_t$, and $(Q-D) \ge L$. Then, $\mathbf{H}'$ is equal to $\mathbf{H}\boldsymbol{\Omega}$ with an $M_t \times M_t$ invertible matrix $\boldsymbol{\Omega}$, if and only if span$(\boldsymbol{\Xi}')$ is equal to span$(\boldsymbol{\Xi})$.

Proof: Refer to Appendix II. ■

By Theorem 2, a scalar channel ambiguity for SISO-OFDM systems in [54] is extended to a matrix channel ambiguity for MIMO-OFDM systems. The channel ambiguity is inherent to blind estimation schemes and can be resolved by exploiting the techniques based on independent component analysis [61] (and references therein) and/or a small number of pilot symbols as given in [62] and [27]. Furthermore, since practical MIMO-OFDM systems provide pilot symbols for the purpose of synchronization, these pilot symbols can be used to resolve the ambiguity matrix.

In particular, when a MIMO-OFDM system structure possesses the specific conditions given in the following Lemma 1, we can estimate a MIMO channel with $J \le 2$.

Lemma 1: In the case of $M_t < M_r$ and $L \le \lfloor (JD-1)/(M_t+1) \rfloor$ with $J \le 2$, assume that $\mathbf{h}(0)$, $\mathbf{h}(L)$, and $\mathbf{H}(z)$ have a full column rank for all z. Then, $\mathbf{H}'$ is equal to $\mathbf{H}\boldsymbol{\Omega}$ with an $M_t \times M_t$ invertible matrix $\boldsymbol{\Omega}$, if and only if $\mathbf{H}'$ has a full column rank and span$(\boldsymbol{\Xi}')$ is equal to span$(\boldsymbol{\Xi})$.

Proof: Refer to Appendix III. ■

Since Lemma 1 allows a MIMO channel to be estimated with $J \le 2$, we can obtain a MIMO channel estimate by exploiting a small number of OFDM symbols with $J = 1$. Furthermore, although the channel conditions required by Lemma 1 are stricter than those in Theorem 2, we note that the conditions in Lemma 1 do not impose any constraints on the number of CPs associated with the number of VCs. Thus, we can estimate a MIMO channel without CPs as long as the conditions are satisfied, thereby increasing the transmission bandwidth efficiency. Obviously, the above theorems and lemma are still valid for a MIMO-OFDM system with no VCs by setting k_0 to 0 and D to N.

To find the signal and noise subspaces, the true $\mathbf{R_{rr}}$ is required. In practice, $\mathbf{R_{rr}}$ is estimated over N_b blocks by

$$\widehat{\mathbf{R}}_{\mathbf{rr}} = \frac{1}{N_b}\sum_{n=0}^{N_b-1}\mathbf{r}(n)\mathbf{r}(n)^{\mathrm{H}}. \tag{31}$$

Thus, when a MIMO channel is estimated by the orthogonal relationship in (28), only estimates $\widehat{\mathbf{U}}_n$ of the eigenvectors spanning the noise subspace, which are obtained by the eigenvalue decomposition of $\widehat{\mathbf{R}}_{\mathbf{rr}}$, are available in practice. In this case, we can obtain the channel matrix estimate $\widehat{\boldsymbol{\mathcal{H}}}$ by minimizing a quadratic cost function $C(\boldsymbol{\mathcal{H}})$ given as

$$C(\boldsymbol{\mathcal{H}}) = \sum_{k=JDM_t+1}^{(JQ-L)M_r}\left\|\widehat{\mathbf{u}}_k^{\mathrm{H}}\boldsymbol{\Xi}\right\|_2^2 = \sum_{k=JDM_t+1}^{(JQ-L)M_r}\left\|\widehat{\mathbf{u}}_k^{\mathrm{H}}\boldsymbol{\mathcal{H}}\boldsymbol{\mathcal{W}}\right\|_2^2. \tag{32}$$

Partitioning the eigenvector estimate $\widehat{\mathbf{u}}_k$ with dimension $(JQ-L)M_r$ into $JQ-L$ equal segments as given in

$$\widehat{\mathbf{u}}_k = \begin{bmatrix}\widehat{\mathbf{v}}_1^{(k)}\\ \widehat{\mathbf{v}}_2^{(k)}\\ \vdots\\ \widehat{\mathbf{v}}_{JQ-L}^{(k)}\end{bmatrix} \tag{33}$$

constructing the $(L+1)M_r \times JQ$ matrix $\widehat{\boldsymbol{\mathcal{V}}}_k$ as

$$\widehat{\boldsymbol{\mathcal{V}}}_k = \begin{bmatrix}\widehat{\mathbf{v}}_1^{(k)} & \widehat{\mathbf{v}}_2^{(k)} & \cdots & \widehat{\mathbf{v}}_{JQ\text{–}L}^{(k)} & \mathbf{0} & \cdots & \mathbf{0}\\ \mathbf{0} & \widehat{\mathbf{v}}_1^{(k)} & \widehat{\mathbf{v}}_2^{(k)} & \cdots & \widehat{\mathbf{v}}_{JQ\text{–}L}^{(k)} & \cdots & \mathbf{0}\\ \vdots & & \ddots & \ddots & \ddots & \ddots & \\ \mathbf{0} & \cdots & \mathbf{0} & \widehat{\mathbf{v}}_1^{(k)} & \widehat{\mathbf{v}}_2^{(k)} & \cdots & \widehat{\mathbf{v}}_{JQ\text{–}L}^{(k)}\end{bmatrix} \tag{34}$$

and defining the matrix $\boldsymbol{\Psi}$ as

$$\boldsymbol{\Psi} \triangleq \sum_{k=JDM_t+1}^{(JQ-L)M_r}\widehat{\boldsymbol{\mathcal{V}}}_k(\mathbf{I}_J \otimes \mathbf{W}^*\mathbf{W}^{\mathrm{T}})\widehat{\boldsymbol{\mathcal{V}}}_k^{\mathrm{H}} \tag{35}$$

we can write a cost function $\sum_{i=1}^{M_t}\mathrm{h}_i^{\mathrm{H}}\boldsymbol{\Psi}\mathrm{h}_i$ equivalent to $C(\boldsymbol{\mathcal{H}})$. By imposing the constraints such as $\|\mathbf{h}_i\|_2 = 1$ for $1 \le i \le M_t$ to avoid trivial solutions, the estimate $\widehat{\mathbf{H}}$ of the channel coefficient matrix $\mathbf{H}$ in (30) is obtained by

$$\widehat{\mathbf{H}} = \left[\widehat{\mathbf{h}}_1,\ \widehat{\mathbf{h}}_2,\ldots,\widehat{\mathbf{h}}_{M_t}\right] = \arg\min_{\|\mathbf{h}_i\|_2=1}\left(\sum_{i=1}^{M_t}{\mathbf{h}_i}^{\mathrm{H}}\boldsymbol{\Psi}\mathbf{h}_i\right). \tag{36}$$

When we find $\widehat{\mathbf{h}}_i$ satisfying $\partial(\sum_{i=1}^{M_t}(\mathbf{h}_i^{\mathrm{H}}\boldsymbol{\Psi}\mathbf{h}_i + \lambda_i(1-\|\mathbf{h}_i\|_2^2)))/\partial\mathbf{h}_i^{\mathrm{H}} = \mathbf{0}$ with a Lagrange multiplier λ_i, the estimates $\widehat{\mathbf{h}}_i$ of the channel response vectors $\mathbf{h}_i$, $1 \le i \le M_t$ in (29) are the eigenvectors associated with the smallest M_t eigenvalues of $\boldsymbol{\Psi}$. Since the orthogonal relationship $\mathbf{u}_k^{\mathrm{H}}\boldsymbol{\Xi} = \mathbf{0}$ in (28) can be rewritten as $(\mathbf{I}_J \otimes \mathbf{W}^{\mathrm{T}})\boldsymbol{\mathcal{V}}_k^{\mathrm{H}}\mathbf{h}_i = \mathbf{0}$ for $1 \le i \le M_t$, we should find $\mathbf{h}_i$ closely orthogonal to column vectors of $\widehat{\boldsymbol{\mathcal{V}}}_k(\mathbf{I}_J \otimes \mathbf{W}^*)$ with an estimate $\widehat{\boldsymbol{\mathcal{V}}}_k$ of $\boldsymbol{\mathcal{V}}_k$. In addition, $\mathbf{h}_1, \mathbf{h}_2, \ldots, \mathbf{h}_{M_t}$ should be linearly independent to satisfy the condition in Theorem 1. Thus, the solution of (36) satisfies these orthogonality and linear independence conditions. Although the vectors $\mathbf{h}_1, \mathbf{h}_2, \ldots, \mathbf{h}_{M_t}$ and the solution of (36) span the same M_t-dimensional space in the ideal case with the knowledge of true $\mathbf{R}_{\mathbf{rr}}$, we do not have information about the direction and magnitude of each $\mathbf{h}_i$ in the space. This causes the channel ambiguity stated in Theorem 2 and Lemma 1. Thus, if the eigenvectors associated with the smallest M_t eigenvalues of $\boldsymbol{\Psi}$ are denoted as $\widehat{\mathbf{h}}'_1, \widehat{\mathbf{h}}'_2, \ldots, \widehat{\mathbf{h}}'_{M_t}$, respectively, we can express the estimated channel coefficient matrix $\widehat{\mathbf{H}}$ as

$$\widehat{\mathbf{H}} = \left[\widehat{\mathbf{h}}_1, \widehat{\mathbf{h}}_2, \ldots, \widehat{\mathbf{h}}_{M_t}\right] = \left[\widehat{\mathbf{h}}'_1, \widehat{\mathbf{h}}'_2, \ldots, \widehat{\mathbf{h}}'_{M_t}\right]\boldsymbol{\Omega} \tag{37}$$

where $\boldsymbol{\Omega}$ is an $M_t \times M_t$ channel ambiguity matrix.

In summary, insofar as the condition in Theorem 1 is satisfied, the proposed subspace method can be applied to the blind channel estimation for a MIMO-OFDM system with $M_t \le M_r$. In addition, we note that the condition requires an upper bound of a true MIMO channel order rather than the exact knowledge of the MIMO channel order. The estimated MIMO channel has an ambiguity corresponding to an $M_t \times M_t$ invertible matrix given in Theorem 2. Furthermore, without relying on the presence of VCs under the specific condition given in Lemma 1, we can apply the proposed method with a smaller J to blind channel estimation for the MIMO-OFDM system without CPs. This increases the bandwidth efficiency and makes it possible to obtain an accurate channel estimate by utilizing a small number of OFDM symbols.

Next, we extend the above results obtained in a MIMO-OFDM system with $M_t \le M_r$ to the blind channel estimation for a MIMO-OFDM system with $M_t > M_r$.

B. MIMO-OFDM System With $M_t > M_r$

To perform the blind channel estimation for a MIMO-OFDM system with $M_t > M_r$, we set the sampling rate at the receiver to q/T with $q \ge \lceil M_t/M_r \rceil$ in the system shown in Fig. 2. By considering the discrete composite channel impulse response between the jth transmit antenna and the ith receive antenna in (6), and defining $h_{ij}^{(\xi)}(l)$ with $\xi = \langle m \rangle_q$ as

$$h_{ij}^{(\xi)}(l) \triangleq h_{ij}\left[\epsilon_i + \left(\frac{\xi}{q} + l\right)T\right] \tag{38}$$

we denote the lth lag of the oversampled MIMO channel as

$$\widetilde{\mathbf{h}}(l) = \begin{bmatrix} h_{11}^{(0)}(l) & h_{12}^{(0)}(l) & \cdots & h_{1M_t}^{(0)}(l) \\ \vdots & \vdots & \vdots & \vdots \\ h_{11}^{(q-1)}(l) & h_{12}^{(q-1)}(l) & \cdots & h_{1M_t}^{(q-1)}(l) \\ h_{21}^{(0)}(l) & h_{22}^{(0)}(l) & \cdots & h_{2M_t}^{(0)}(l) \\ \vdots & \vdots & \vdots & \vdots \\ h_{21}^{(q-1)}(l) & h_{22}^{(q-1)}(l) & \cdots & h_{2M_t}^{(q-1)}(l) \\ \vdots & \vdots & \vdots & \vdots \\ h_{M_r1}^{(0)}(l) & h_{M_r2}^{(0)}(l) & \cdots & h_{M_rM_t}^{(0)}(l) \\ \vdots & \vdots & \vdots & \vdots \\ h_{M_r1}^{(q-1)}(l) & h_{M_r2}^{(q-1)}(l) & \cdots & h_{M_rM_t}^{(q-1)}(l) \end{bmatrix}. \tag{39}$$

Assuming that the discrete composite MIMO channel has L as the upper bound on the order of the channel, we construct the $(JQ-L)qM_r \times JQM_t$ channel matrix $\widetilde{\boldsymbol{\mathcal{H}}}$, in the same way as given in (25), as

$$\widetilde{\boldsymbol{\mathcal{H}}} = \begin{bmatrix} \widetilde{\mathbf{h}}(0) & \cdots & \widetilde{\mathbf{h}}(L) & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \widetilde{\mathbf{h}}(0) & \cdots & \widetilde{\mathbf{h}}(L) & \cdots & \mathbf{0} \\ \vdots & & \ddots & & \ddots & \\ \mathbf{0} & \cdots & \mathbf{0} & \widetilde{\mathbf{h}}(0) & \cdots & \widetilde{\mathbf{h}}(L) \end{bmatrix}. \tag{40}$$

Recalling $m/q = \lfloor m/q \rfloor + (\langle m \rangle_q/q)$ and defining m' and ξ as $m' \triangleq \lfloor m/q \rfloor$ and $\xi \triangleq \langle m \rangle_q$, respectively, we denote $r_i[\epsilon_i + (mT/q)]$ in (6) as $r_i^{(\xi)}(m')$. Rearranging $r_i^{(\xi)}(m')$ according to $r_i^{(\xi)}(n,k) = r_i^{(\xi)}(k+nQ)$, we express the oversampled received signal at M_r receive antennas as

$$\begin{aligned}\widetilde{\mathbf{r}}(n,k) = \Big[& r_1^{(0)}(n,k), \ldots, r_1^{(q-1)}(n,k), r_2^{(0)}(n,k), \ldots, \\ & r_2^{(q-1)}(n,k), \ldots, r_{M_r}^{(0)}(n,k), \ldots, r_{M_r}^{(q-1)}(n,k)\Big]^{\mathrm{T}}\end{aligned} \tag{41}$$

$$\widetilde{\mathbf{r}}_n = \left[\widetilde{\mathbf{r}}(n,Q-1)^{\mathrm{T}}, \widetilde{\mathbf{r}}(n,Q-2)^{\mathrm{T}}, \ldots, \widetilde{\mathbf{r}}(n,0)^{\mathrm{T}}\right]^{\mathrm{T}} \tag{42}$$

$$\widetilde{\mathbf{r}}(n) = \left[\widetilde{\mathbf{r}}_n^{\mathrm{T}}, \widetilde{\mathbf{r}}_{n-1}^{\mathrm{T}}, \ldots, \widetilde{\mathbf{r}}_{n-J+1}[1:(Q-L)qM_r]^{\mathrm{T}}\right]^{\mathrm{T}}. \tag{43}$$

Similarly, denoting $\eta_i[\epsilon_i + (mT/q)]$ in (6) as $\eta_i^{(\xi)}(m')$, and rearranging $\eta_i^{(\xi)}(m')$ according to $\eta_i^{(\xi)}(n,k) = \eta_i^{(\xi)}(k+nQ)$, we write the oversampled additive noise at M_r receive antennas as

$$\begin{aligned}\widetilde{\boldsymbol{\eta}}(n,k) = \Big[& \eta_1^{(0)}(n,k), \ldots, \eta_1^{(q-1)}(n,k), \eta_2^{(0)}(n,k), \ldots, \\ & \eta_2^{(q-1)}(n,k), \ldots, \eta_{M_r}^{(0)}(n,k), \ldots, \eta_{M_r}^{(q-1)}(n,k)\Big]^{\mathrm{T}}\end{aligned} \tag{44}$$

$$\widetilde{\boldsymbol{\eta}}_n = \left[\widetilde{\boldsymbol{\eta}}(n,Q-1)^{\mathrm{T}}, \widetilde{\boldsymbol{\eta}}(n,Q-2)^{\mathrm{T}}, \ldots, \widetilde{\boldsymbol{\eta}}(n,0)^{\mathrm{T}}\right]^{\mathrm{T}} \tag{45}$$

$$\widetilde{\boldsymbol{\eta}}(n) = \left[\widetilde{\boldsymbol{\eta}}_n^{\mathrm{T}}, \widetilde{\boldsymbol{\eta}}_{n-1}^{\mathrm{T}}, \ldots, \widetilde{\boldsymbol{\eta}}_{n-J+1}[1:(Q-L)qM_r]^{\mathrm{T}}\right]^{\mathrm{T}}. \tag{46}$$

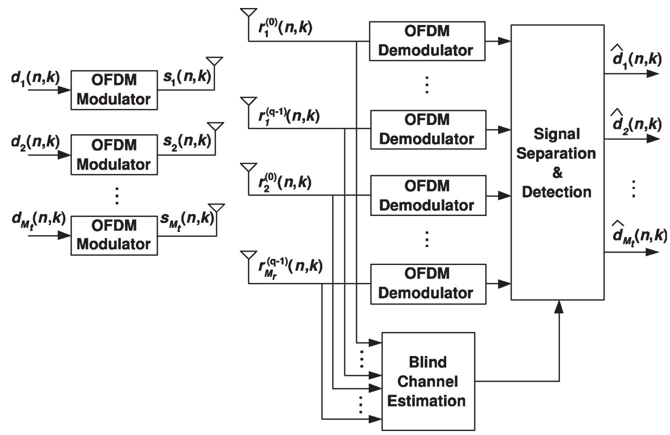


Fig. 3. Equivalent MIMO-OFDM system model with M_t transmit and qM_r receive antennas.

Then, the oversampled received signal vector $\widetilde{\mathbf{r}}(n)$ in (43) can be written in a matrix form as

$$\widetilde{\mathbf{r}}(n)=\widetilde{\boldsymbol{\mathcal{H}}}\mathbf{s}(n)+\widetilde{\boldsymbol{\eta}}(n)=\widetilde{\boldsymbol{\mathcal{H}}}\boldsymbol{\mathcal{W}}\mathbf{d}(n)+\widetilde{\boldsymbol{\eta}}(n)\triangleq\widetilde{\boldsymbol{\Xi}}\mathbf{d}(n)+\widetilde{\boldsymbol{\eta}}(n). \tag{47}$$

When we consider (39) through (47), we can model a MIMO-OFDM system with M_t transmit and M_r receive antennas, where the received signals are oversampled by a factor of q, as an equivalent MIMO-OFDM system with M_t transmit and qM_r receive antennas. This equivalent system model is shown in Fig. 3, where the received signal at each receive antenna is sampled at the rate $1/T$. Since the equivalent system has $qM_r \geq \lceil M_t/M_r \rceil M_r \geq M_t$, we can apply Theorem 1 for the channel identifiability of the system. Let us denote $\sum_{l=0}^{L}\widetilde{\mathbf{h}}(l)z^{-l}$ as $\widetilde{\mathbf{H}}(z)$. According to Theorem 1, if the condition in (48) is satisfied

$$\operatorname{rank}\left(\widetilde{\mathbf{H}}(w_N^i)\right)=M_t \text{ for all } i\in\{k\}_{k=k_0}^{k_0+D-1} \tag{48}$$

the matrix $\widetilde{\boldsymbol{\Xi}}$ has a full column rank, which implies that the oversampled MIMO channel can be identified through the noise subspace method.

In our derivation, we consider MIMO channels satisfying the condition in (48). Furthermore, we note that since the sampling rate q/T is higher than the Nyquist rate, the oversampled noise $\eta_i^{(\xi)}(n,k)$ is not necessarily temporally uncorrelated. Although it is possible to design a front-end receive filter $g_{rx}[t]$ with a wider bandwidth to whiten the oversampled noise [63], we suppose that the oversampled noise vector $\widetilde{\boldsymbol{\eta}}\ (n)$ is generally colored with the covariance matrix $\mathbf{R}_{\widetilde{\boldsymbol{\eta}}\widetilde{\boldsymbol{\eta}}}$ that has full rank. By decomposing $\mathbf{R}_{\widetilde{\boldsymbol{\eta}}\widetilde{\boldsymbol{\eta}}}$ as $\mathbf{R}_{\widetilde{\boldsymbol{\eta}}\widetilde{\boldsymbol{\eta}}}=\mathbf{R}_{\widetilde{\boldsymbol{\eta}}\widetilde{\boldsymbol{\eta}}}^{1/2}\mathbf{R}_{\widetilde{\boldsymbol{\eta}}\widetilde{\boldsymbol{\eta}}}^{\mathrm{H}/2}$, and whitening the oversampled received signal vector $\widetilde{\mathbf{r}}(n)$ in (47) by the inverse of $\mathbf{R}_{\widetilde{\boldsymbol{\eta}}\widetilde{\boldsymbol{\eta}}}^{1/2}$, denoted as $\mathbf{R}_{\widetilde{\boldsymbol{\eta}}\widetilde{\boldsymbol{\eta}}}^{-1/2}$, we obtain the whitened received signal vector $\widetilde{\mathbf{r}}_w(n)$ as given in

$$\widetilde{\mathbf{r}}_w(n)=\mathbf{R}_{\widetilde{\boldsymbol{\eta}}\widetilde{\boldsymbol{\eta}}}^{-\frac{1}{2}}\widetilde{\mathbf{r}}(n)=\mathbf{R}_{\widetilde{\boldsymbol{\eta}}\widetilde{\boldsymbol{\eta}}}^{-\frac{1}{2}}\widetilde{\boldsymbol{\Xi}}\mathbf{d}(n)+\mathbf{R}_{\widetilde{\boldsymbol{\eta}}\widetilde{\boldsymbol{\eta}}}^{-\frac{1}{2}}\widetilde{\boldsymbol{\eta}}(n). \tag{49}$$

We assume that the additive noise is uncorrelated with the transmitted signal, and the autocorrelation matrix $\mathbf{R}_{\mathbf{dd}}$ of the information symbol vector $\mathbf{d}(n)$ has full rank. When the autocorrelation matrix $\mathbf{R}_{\widetilde{\mathbf{r}}_w\widetilde{\mathbf{r}}_w}=E\{\widetilde{\mathbf{r}}_w(n)\widetilde{\mathbf{r}}_w(n)^{\mathrm{H}}\}$ of the whitened received signal vector $\widetilde{\mathbf{r}}_w(n)$ is diagonalized through the eigenvalue decomposition, we can partition the eigenvectors $\widetilde{\mathbf{U}}$ into the vectors $\widetilde{\mathbf{U}}_s$ spanning a signal subspace $\operatorname{span}(\widetilde{\mathbf{U}}_s)$ and the vectors $\widetilde{\mathbf{U}}_n$ spanning a noise subspace $\operatorname{span}(\widetilde{\mathbf{U}}_n)$ as

$$\begin{aligned}\widetilde{\mathbf{U}}&=\left[\widetilde{\mathbf{U}}_s|\widetilde{\mathbf{U}}_n\right]\\&=\left[\widetilde{\mathbf{u}}_1,\ldots,\widetilde{\mathbf{u}}_{JDM_t}|\widetilde{\mathbf{u}}_{JDM_t+1},\ldots,\widetilde{\mathbf{u}}_{(JQ-L)qM_r}\right].\end{aligned} \tag{50}$$

Since $\operatorname{span}(\mathbf{R}_{\widetilde{\boldsymbol{\eta}}\widetilde{\boldsymbol{\eta}}}^{-1/2}\widetilde{\boldsymbol{\Xi}})$ and $\operatorname{span}(\widetilde{\mathbf{U}}_s)$ share the same JDM_t-dimensional space and are orthogonal to $\operatorname{span}(\widetilde{\mathbf{U}}_n)$, we obtain the orthogonal relationship given as

$$\widetilde{\mathbf{u}}_k^{\mathrm{H}}\mathbf{R}_{\widetilde{\boldsymbol{\eta}}\widetilde{\boldsymbol{\eta}}}^{-\frac{1}{2}}\widetilde{\boldsymbol{\Xi}}=\mathbf{0} \text{ for all } k\in\{n\}_{n=JDM_t+1}^{(JQ-L)qM_r}. \tag{51}$$

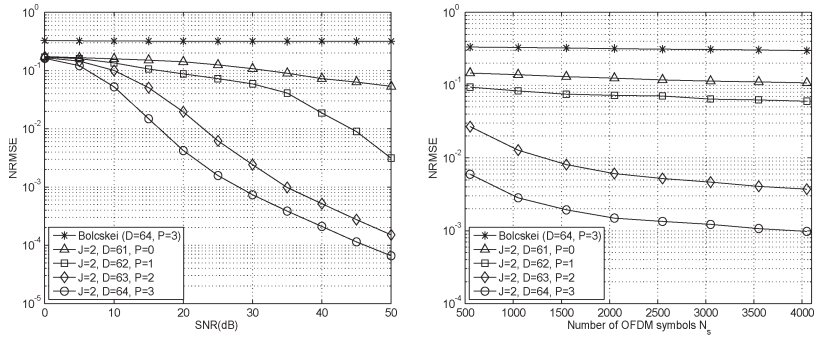
Defining the $(L+1)qM_r\times 1$ channel response vector $\widetilde{\mathbf{h}}_i$ associated with the channel impulse responses between the ith transmit antenna and qM_r receive antennas, and the channel coefficient matrix $\widetilde{\mathbf{H}}$ consisting of $\widetilde{\mathbf{h}}_i$ as, respectively

$$\widetilde{\mathbf{h}}_i\triangleq\left[\widetilde{\mathbf{h}}(0)[:,i]^{\mathrm{T}},\widetilde{\mathbf{h}}(1)[:,i]^{\mathrm{T}},\ldots,\widetilde{\mathbf{h}}(L)[:,i]^{\mathrm{T}}\right]^{\mathrm{T}},\quad 1\leq i\leq M_t \tag{52}$$

$$\widetilde{\mathbf{H}}\triangleq\left[\widetilde{\mathbf{h}}_1,\widetilde{\mathbf{h}}_2,\ldots,\widetilde{\mathbf{h}}_{M_t}\right]=\left[\widetilde{\mathbf{h}}(0)^{\mathrm{T}},\widetilde{\mathbf{h}}(1)^{\mathrm{T}},\ldots,\widetilde{\mathbf{h}}(L)^{\mathrm{T}}\right]^{\mathrm{T}} \tag{53}$$

and replacing M_r, $\widehat{\mathbf{u}}_k$, $\boldsymbol{\Xi}$, $\boldsymbol{\mathcal{H}}$, $\mathbf{h}_i$, and $\mathbf{H}$ in (32) through (37) with qM_r, $\mathbf{R}_{\widetilde{\boldsymbol{\eta}}\widetilde{\boldsymbol{\eta}}}^{-\mathrm{H}/2}\widehat{\widetilde{\mathbf{u}}}_k$, $\widetilde{\boldsymbol{\Xi}}$, $\widetilde{\boldsymbol{\mathcal{H}}}$, $\widetilde{\mathbf{h}}_i$, and $\widetilde{\mathbf{H}}$, respectively, we can construct a cost function in the same manner as given in Section III-A. By minimizing this cost function, we can estimate the channel coefficient matrix $\widetilde{\mathbf{H}}$ up to an $M_t\times M_t$ channel ambiguity matrix stated in Theorem 2 and Lemma 1. Furthermore, the channel ambiguity matrix can be resolved by using the schemes given in [62] and [27].

In summary, when the received signal at each receive antenna is oversampled by a factor of $q\geq\lceil M_t/M_r\rceil$ and the condition in (48) according to Theorem 1 is satisfied, we can still apply the proposed method to the blind channel estimation for a MIMO-OFDM system with $M_t > M_r$. Again, the condition depends on an upper bound of a true MIMO channel order rather than the exact knowledge of the MIMO channel order. A MIMO channel is estimated up to an $M_t\times M_t$ ambiguity matrix given in Theorem 2. If the matrices $\widetilde{\mathbf{h}}(0)$, $\widetilde{\mathbf{h}}(L)$, and $\widetilde{\mathbf{H}}(z)$ satisfy the same conditions as those required for $\mathbf{h}(0)$, $\mathbf{h}(L)$, and $\mathbf{H}(z)$ in Lemma 1, respectively, the proposed method with $J\leq 2$ is still applicable to the blind channel estimation for the MIMO-OFDM system without CPs, regardless of the existence of VCs. This increases the bandwidth efficiency and enables accurate channel estimation by exploiting a small number of OFDM symbols.



(a) (b)

Fig. 4. Comparison of NRMSE performance when the sum of the number of virtual carriers $(N-D)$ and the number of CPs (P) is fixed to three. (a) NRMSE versus SNR for the number of OFDM symbols used for channel estimation $N_s = 2000$. (b) NRMSE versus the number of OFDM symbols used for channel estimation N_s for SNR $= 25$ dB.

IV. Simulation Results

To evaluate the performance of the proposed method, we consider a MIMO-OFDM system with two transmit antennas $(M_t = 2)$ and two receive antennas $(M_r = 2)$. The number of subcarriers N is set to 64. Information symbols $d_i(n,k)$ are independent and identically distributed (i.i.d.) 16-quadrature amplitude modulation symbols. Each channel tap $h_{ij}(l)$ is i.i.d. and randomly generated from $\mathcal{CN}(0,\sigma_h^2)$. The order of the MIMO channel is considered to be $L = 3$. We suppose that the channel is time-invariant during each channel estimation. For the fairness of performance comparison, the transmit power per OFDM symbol is fixed to E_s for all simulations, and the additive noise at each receive antenna is a spatially uncorrelated complex white Gaussian noise with zero mean and variance σ_η^2 determined by the SNR defined as

$$\text{SNR} \triangleq 10\log_{10}\frac{M_t(L+1)\sigma_h^2 E_s}{(N+P_o)\sigma_\eta^2}\ (\text{dB}) \tag{54}$$

where P_o is the maximum CP length used throughout the simulations, which is set to three.

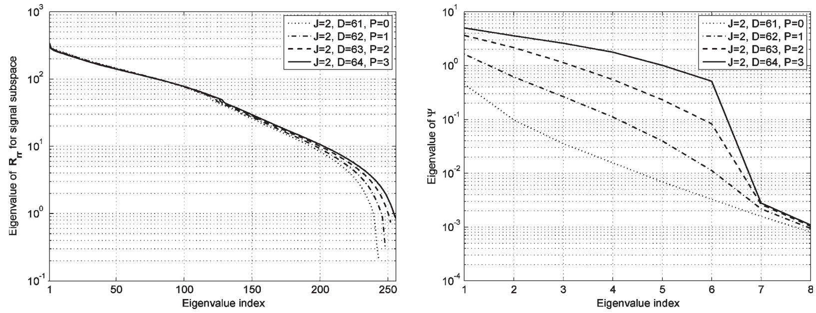
As a measure of performance, we consider the normalized root mse (NRMSE) given as

$$\text{NRMSE} = \sqrt{\frac{1}{N_m M_t M_r (L+1)}\sum_{k=1}^{N_m}\sum_{i=1}^{M_t}\frac{\left\|\mathbf{h}_i^{(k)} - \widehat{\mathbf{h}}_i^{(k)}\right\|_2^2}{\left\|\mathbf{h}_i^{(k)}\right\|_2^2}} \tag{55}$$

where N_m is the number of Monte Carlo trials, the superscript k refers to the kth Monte Carlo trial, and $\mathbf{h}_i^{(k)}$ and $\widehat{\mathbf{h}}_i^{(k)}$ represent the true channel response vector and the estimated channel response vector after resolving a channel ambiguity, respectively. All the results are obtained by averaging $N_m = 500$ independent Monte Carlo trials. To isolate the impact of a scheme for resolving a channel ambiguity on channel estimation in computing NRMSE, we calculate the ambiguity $\mathbf{\Omega}$ by minimizing $\|[\mathbf{h}_1^{(k)}, \mathbf{h}_2^{(k)}, \ldots, \mathbf{h}_{M_t}^{(k)}] - [\widehat{\mathbf{h}}_1^{\prime(k)}, \widehat{\mathbf{h}}_2^{\prime(k)}, \ldots, \widehat{\mathbf{h}}_{M_t}^{\prime(k)}]\mathbf{\Omega}\|_F^2$ as used in [38] and [32]. $[\widehat{\mathbf{h}}_1^{\prime(k)}, \widehat{\mathbf{h}}_2^{\prime(k)}, \ldots, \widehat{\mathbf{h}}_{M_t}^{\prime(k)}]$ is the estimated channel coefficient matrix by the proposed method. By using this approach, the NRMSE provides a measure of how well the true MIMO channel and the estimated MIMO channel by the proposed method span the same M_t-dimensional space.

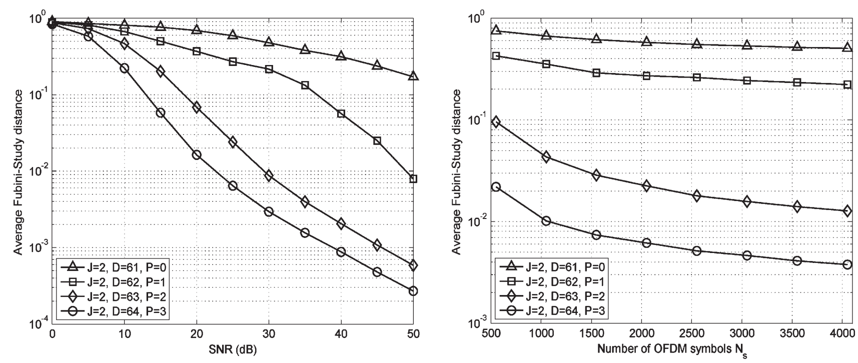
In Fig. 4, the NRMSE performance of the proposed method with different combinations of the number of information symbols (D) and the number of CPs (P) is compared with that of the method in [25] that is marked with "Bölcskei." In the cases for the proposed method, the redundancy $(N-D+P)$ is fixed to three through various combinations of VCs and CPs. An observed OFDM symbol block J associated with the dimensions of subspaces is fixed to two. To obtain the NRMSE performance as a function of SNR shown in Fig. 4(a), we use 2000 OFDM symbols. The NRMSE performance as a function of the number of OFDM symbols N_s used for channel estimation in Fig. 4(b) is obtained by setting the SNR to 25 dB. For a fair comparison, the complex scalar ambiguities α_{ij} from the method in [25] are also resolved by minimizing $\|[h_{ij}^{(k)}(0), h_{ij}^{(k)}(1), \ldots, h_{ij}^{(k)}(L)] - \alpha_{ij}[\widehat{h}_{ij}^{\prime(k)}(0), \widehat{h}_{ij}^{\prime(k)}(1), \ldots, \widehat{h}_{ij}^{\prime(k)}(L)]\|_2^2$. $[\widehat{h}_{ij}^{\prime(k)}(0), \widehat{h}_{ij}^{\prime(k)}(1), \ldots, \widehat{h}_{ij}^{\prime(k)}(L)]$ is the estimated channel impulse response vector between the jth transmit antenna and the ith receive antenna by the method in [25].

As we can see from Fig. 4, the estimator errors of all the cases decrease with increasing SNR and OFDM symbol record length N_s. Furthermore, the proposed method demonstrates much better performance than the method in [25], which reveals the fast convergence property of the noise subspace method for a small data record. In addition, there are performance gaps among the non-CP system ($J = 2$, $D = 61$, $P = 0$), the insufficient CP systems ($J = 2$, $D = 62$, $P = 1$ and $J = 2$, $D = 63$, $P = 2$) and the CP-only system ($J = 2$, $D = 64$, $P = 3$). This demonstrates that CPs are more useful for the noise subspace-based estimation method than VCs. We discuss this benefit of CPs by referring to Fig. 5 that is obtained by using 2000 OFDM symbols at the SNR of 25 dB and averaging 500 independent trials. Fig. 5(a) shows the estimated eigenvalues of the autocorrelation matrix $\mathbf{R_{rr}}$ corresponding to the estimated eigenvectors spanning the signal subspace in a descending order of the eigenvalues. From Fig. 5(a), we note that as fewer CPs are used in the presence of VCs, the eigenvalues rapidly decrease. Although the boundary between the signal subspace and the



(a) (b)

Fig. 5. Comparison of eigenvalue distributions when the sum of the number of virtual carriers $(N-D)$ and the number of CPs (P) is fixed to three, and 2000 OFDM symbols are used for channel estimation at the SNR of 25 dB. (a) Distributions of estimated eigenvalues of the autocorrelation matrix $\mathbf{R_{rr}}$ corresponding to estimated eigenvectors spanning the signal subspace in a descending order of the eigenvalues. (b) Distributions of eigenvalues of the matrix $\boldsymbol{\Psi}$ in (35) in a descending order.



(a) (b)

Fig. 6. Comparison of average Fubini–Study distances when the sum of the number of virtual carriers $(N-D)$ and the number of CPs (P) is fixed to three. (a) Average Fubini–Study distance versus SNR for the number of OFDM symbols used for channel estimation $N_s = 2000$. (b) Average Fubini–Study distance versus the number of OFDM symbols used for channel estimation N_s for SNR $= 25$ dB.

noise subspace is theoretically given in (27), this boundary with the rapidly decreasing eigenvalues is usually indistinguishable in the presence of additive noise. That is, it is difficult to accurately differentiate eigenvectors spanning the signal subspace and eigenvectors spanning the noise subspace from the estimated eigenvalues and eigenvectors in the presence of noise. This causes the performance degradation in the cases exploiting fewer CPs. In addition, the more rapid the eigenvalues decrease, the poorer the estimation performance becomes. Furthermore, since the eigenvectors corresponding to the seventh and eighth smallest eigenvalues of the matrix $\boldsymbol{\Psi}$ in (35) provide the estimated MIMO channel with an ambiguity for this simulation example, a distinctive boundary between the sixth smallest eigenvalue and the seventh smallest eigenvalue is desirable. As demonstrated in Fig. 5(b), which shows the eigenvalues of $\boldsymbol{\Psi}$ in a descending order, however, this boundary is not clear in the cases with a reduced number of CPs. Also, CPs increase the dimension of each eigenvector estimate spanning the noise subspace, thereby imposing more constraints on the estimates of the channel impulse responses. Thus, although the subspace dimension extended by larger CPs increases the computational complexity of the proposed method, increasing CPs rather than VCs can significantly improve the performance of the subspace method.

As another measure evaluating the closeness between two M_t-dimensional spaces spanned by the true MIMO channel and the estimated MIMO channel by the proposed method, we can consider the Fubini–Study distance [64] given as

$$d_{\mathrm{FS}}\left(\mathbf{U}_h^{(k)}, \widehat{\mathbf{H}}^{(k)}\right) = \arccos\left|\det\left(\mathbf{U}_h^{(k)\mathrm{H}}\widehat{\mathbf{H}}^{(k)}\right)\right|. \tag{56}$$

In (56), $\mathbf{U}_h^{(k)}$ is a matrix consisting of eigenvectors associated with nonzero eigenvalues of $\mathbf{H}^{(k)}{\mathbf{H}^{(k)}}^{\mathrm{H}}$, where $\mathbf{H}^{(k)}$ is the kth realization of the true channel coefficient matrix and $\widehat{\mathbf{H}}^{(k)}$ is an estimate of $\mathbf{H}^{(k)}$ obtained by the proposed method. By comparing the average Fubini–Study distances defined as $(1/N_m)\sum_{k=1}^{N_m} d_{\mathrm{FS}}(\mathbf{U}_h^{(k)}, \widehat{\mathbf{H}}^{(k)})$ in Fig. 6, we evaluate the performance of the proposed method for the systems considered in Fig. 4. To compute the average Fubini–Study distances, we use $N_m = 500$ independent trials. Fig. 6(a) shows the average Fubini–Study distance of the proposed method as a function of SNR with the utilization of 2000 OFDM symbols for channel estimation, whereas Fig. 6(b) demonstrates the average Fubini–Study distance of the proposed method as a function of the number of OFDM symbols used for channel estimation with the SNR fixed to 25 dB. As expected, the Fubini–Study distances of all the cases still decrease with increasing SNR and OFDM symbol record length N_s. Furthermore, we can see that the decreasing trends in the distances are similar to those in the NRMSE performance in Fig. 4, which reconfirms that

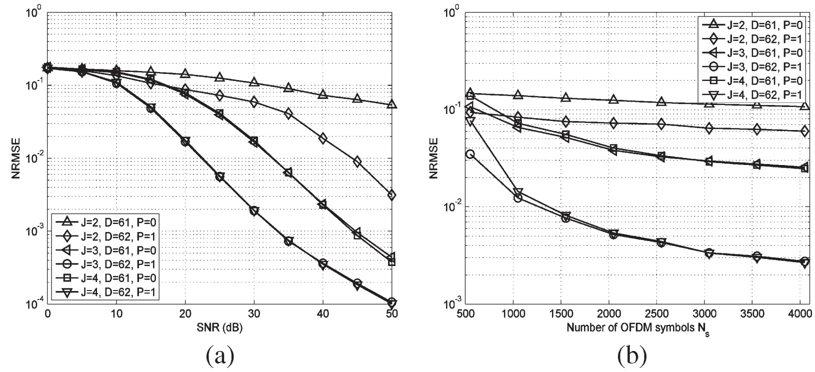


Fig. 7. Comparison of NRMSE performance when an observed OFDM symbol block J increases. (a) NRMSE versus SNR for the number of OFDM symbols used for channel estimation $N_s = 2000$. (b) NRMSE versus the number of OFDM symbols used for channel estimation N_s for SNR $= 25$ dB.

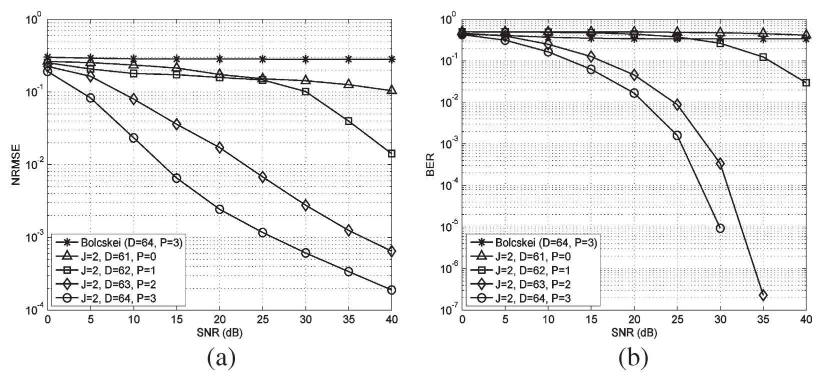


Fig. 8. Comparison of NRMSE and BER performances when the MIMO channel in (57) is estimated by using 2000 OFDM symbols. (a) NRMSE versus SNR. (b) BER versus SNR.

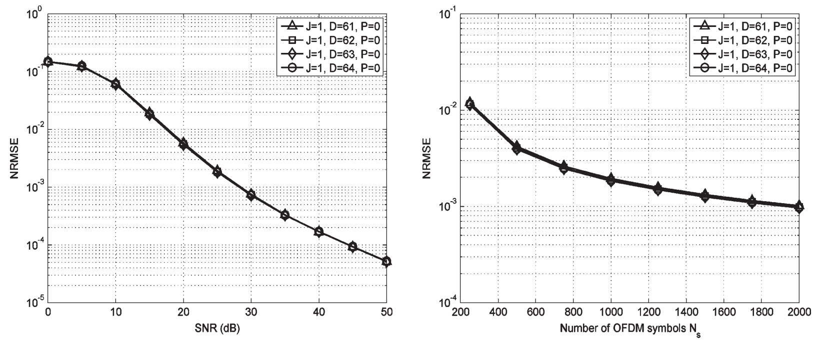
exploitation of CPs rather than VCs for the proposed method improves the closeness of the distance. Due to the similarity between the NRMSE and the Fubini–Study distance in our simulations, we consider only the NRMSE performance in the simulations hereafter.

In the cases having no or insufficient CPs, Fig. 7 shows the NRMSE performance of the proposed method obtained by increasing an observed OFDM symbol block J in two cases of $D = 61$, $P = 0$ and $D = 62$, $P = 1$. We obtain the NRMSE performance as a function of SNR in Fig. 7(a) by using 2000 OFDM symbols. The NRMSE performance as a function of the number of OFDM symbols N_s in Fig. 7(b) is obtained with the SNR fixed to 25 dB. From Fig. 7, we notice that the channel estimation performance in both cases is improved by increasing J which is associated with the dimensions of subspaces. In particular, increasing J from 2 to 3 significantly improves the estimation performance, whereas increasing J from 3 to 4 results in trivial performance improvement. This illustrates that there might be an adequate dimension for the channel estimation based on the noise subspace method and increasing J more than the adequate dimension might not enhance remarkably the performance. Thus, even if the dimension extended by a larger J increases the computational complexity of the eigenstructure-based method, the proposed method with an adequate dimension is applicable to the MIMO-OFDM system with no or insufficient CPs and can achieve the improved performance, thereby potentially leading to higher channel utilization.

To demonstrate that the proposed method is insensitive to a true channel order, we consider the following MIMO channel given in [25].

$$\mathbf{H}(z) = \begin{bmatrix} 0.4851 & 0.3200 \\ -0.3676 & 0.2182 \end{bmatrix} + \begin{bmatrix} -0.4851 & 0.9387 \\ 0.8823 & 0.8729 \end{bmatrix} z^{-1} + \begin{bmatrix} 0.7276 & -0.1280 \\ 0.2941 & -0.4364 \end{bmatrix} z^{-2}. \quad (57)$$

We assume that the upper bound of the channel order L is equal to three even if the true channel order is two, and the zero-forcing detection based on the estimated MIMO channel is used for symbol recovery. Fig. 8(a) and (b) shows the NRMSE and bit error rate (BER) performances as functions of SNR when the MIMO channel in (57) is estimated by using 2000 OFDM symbols, respectively. As we can see from Fig. 8(a), the proposed method still achieves a good estimation performance, which demonstrates its insensitivity to a true channel order. In addition, the proposed method outperforms the method in [25]. Even in this example, we observe that the utilization of CPs rather than VCs increases the accuracy of channel estimation. The BER performance in Fig. 8(b) reflects an influence of the channel estimation accuracy shown in Fig. 8(a) on symbol recovery. In particular, we note that although the two cases of $J = 2$, $D = 61$, $P = 0$ and $J = 2$, $D = 62$, $P = 1$ using the proposed method achieve lower estimation errors than the method in [25], they exhibit poorer BER performance than the



(a) (b)

Fig. 9. Comparison of NRMSE performance for a MIMO-OFDM system without CPs when the observed OFDM symbol block $J = 1$ is used. The system has two transmit and three receive antennas, and the MIMO channel in (58) is considered. (a) NRMSE versus SNR for the number of OFDM symbols used for channel estimation $N_s = 1000$. (b) NRMSE versus the number of OFDM symbols used for channel estimation N_s for SNR $= 25$ dB.

method in [25]. This is due to the fact that the estimated channel matrices in these cases, which are obtained by combining block Toeplitz matrices constructed from the estimated MIMO channels with the IFFT matrices, tend to be ill-conditioned. In this case, the zero-forcing detection by the inversion of the channel matrices may increase the detrimental effects of the channel estimation error and the additive noise on the symbol recovery. Thus, this ill-conditioning can result in poor BER performance even with a small channel estimation error or a small amount of additive noise. On the other hand, as the length of CPs increases, the BER performance of the proposed method is significantly improved and much better than that of the method in [25].

Finally, we consider a MIMO-OFDM system with two transmit antennas $(M_t = 2)$ and three receive antennas $(M_r = 3)$ and the MIMO channel in (58) to evaluate the performance of the proposed method with an observed OFDM symbol block $J = 1$

$$\mathbf{H}(z) = \begin{bmatrix} 0.2200 + j0.0850 & 0.0904 - j0.2397 \\ 0.0397 - j0.5745 & 0.0172 - j0.1402 \\ -0.0096 - j0.0873 & 0.4328 + j0.2893 \end{bmatrix} + \begin{bmatrix} -0.1739 - j0.7379 & -0.1858 + j0.3828 \\ 0.7121 + j0.0601 & -0.0209 - j0.1041 \\ -0.1419 + j0.4276 & -0.2524 + j0.6386 \end{bmatrix} z^{-1} + \begin{bmatrix} 0.1984 - j0.4376 & 0.6829 + j0.4328 \\ -0.1216 + j0.2128 & 0.2530 + j0.3916 \\ -0.0433 - j0.6528 & 0.2312 + j0.1217 \end{bmatrix} z^{-2} + \begin{bmatrix} 0.3704 - j0.0400 & 0.0855 - j0.3039 \\ 0.2967 - j0.0977 & 0.6466 + j0.5773 \\ -0.2574 - j0.5432 & 0.3431 - j0.2673 \end{bmatrix} z^{-3}. \tag{58}$$

To achieve a high bandwidth efficiency, we do not insert a CP to each OFDM symbol to be transmitted. Fig. 9(a) shows the NRMSE performance as a function of SNR that is obtained by using 1000 OFDM symbols. Fig. 9(b) demonstrates the NRMSE performance as a function of the number of OFDM symbols N_s that is obtained with the SNR fixed to 25 dB. The channel estimation errors of all the cases still decrease with increasing SNR and OFDM symbol record length N_s, which demonstrates that the proposed method can achieve an accurate channel estimation by using a smaller number of OFDM symbols with $J = 1$. In particular, the estimation performance in all the cases is almost identical. This indicates that the dimensions of the subspaces in all the cases reach an adequate dimension for the proposed subspace method.

Furthermore, by applying the proposed method to a MIMO-OFDM system with four transmit and two receive antennas, we also confirmed the similar results to those of the MIMO-OFDM system with two transmit and two receive antennas given above.

V. Conclusion

In this paper, we established the conditions for blind channel identifiability in a MIMO-OFDM system and presented a blind channel estimation scheme based on the noise subspace method. The proposed method unifies and generalizes the existing SISO-OFDM blind channel estimators to the case of MIMO-OFDM with any number of transmit and receive antennas. Furthermore, the proposed method achieves accurate channel estimation and fast convergence. This method also demonstrates insensitivity to the exact knowledge of a true MIMO channel order, which implies that it only requires an upper bound on the MIMO channel order. In terms of both channel estimation accuracy and convergence speed, increasing the length of CPs rather than the number of VCs for the proposed method was found to significantly improve the performance in the simulations. In addition, by increasing an observed OFDM symbol block to an adequate dimension for channel estimation, the proposed method can achieve an accurate channel estimation in an MIMO-OFDM system with no or insufficient CPs, thereby potentially increasing channel utilization. Finally, when a system configuration is satisfied with the specific conditions given in Lemma 1, the proposed method can be applied to a MIMO-OFDM system without CPs, regardless of the presence of VCs, thereby achieving a higher bandwidth efficiency.

APPENDIX I
PROOF OF THEOREM 1

First, we show that if $\mathrm{rank}(\mathbf{H}(w_N^i)) = M_t$ for all $i \in \{k\}_{k=k_0}^{k_0+D-1}$, the matrix Ξ has a full column rank. We define $\boldsymbol{\lambda}(i)$ and $\boldsymbol{\Lambda}$ as

$$\boldsymbol{\lambda}(i) \triangleq w_N^{(k_0+i-1)L}\mathbf{H}\left(w_N^{(k_0+i-1)}\right) \tag{59}$$

$$\boldsymbol{\Lambda} \triangleq \mathrm{diag}\left(\boldsymbol{\lambda}(1), \boldsymbol{\lambda}(2), \ldots, \boldsymbol{\lambda}(D)\right). \tag{60}$$

By choosing the rows $(j-1)QM_r+1$ to $(j-1)QM_r+DM_r$ for $1 \le j \le J$ from the matrix Ξ, we can generate a $JDM_r \times JDM_t$ submatrix $\widehat{\Xi}$ that is a block diagonal matrix with $(\mathbf{W}\otimes\mathbf{I}_{M_r})[LM_r+1:(L+D)M_r,:]\boldsymbol{\Lambda}$s on its main diagonal [54]. By using the structure of $(\mathbf{W}\otimes\mathbf{I}_{M_r})[LM_r+1:(L+D)M_r,:]$ and a Vandermonde matrix property [65], we can easily show that $(\mathbf{W}\otimes\mathbf{I}_{M_r})[LM_r+1:(L+D)M_r,:]$ is a nonsingular matrix. In addition, if $\mathrm{rank}(\mathbf{H}(w_N^i)) = M_t$ for all $i \in \{k\}_{k=k_0}^{k_0+D-1}$, the matrix $\boldsymbol{\Lambda}$ has a full column rank. Thus, $\mathrm{rank}((\mathbf{W}\otimes\mathbf{I}_{M_r})[LM_r+1:(L+D)M_r,:]\boldsymbol{\Lambda}) = \mathrm{rank}(\boldsymbol{\Lambda}) = DM_t$, and consequently $\mathrm{rank}(\widehat{\Xi}) = JDM_t$. Since the submatrix $\widehat{\Xi}$ is generated from the matrix Ξ by removing some rows, $\mathrm{rank}(\Xi)$ becomes JDM_t, which means that the matrix Ξ has a full column rank.

Conversely, to prove that if the matrix Ξ has a full column rank, then $\mathrm{rank}(\mathbf{H}(w_N^i)) = M_t$ for all $i \in \{k\}_{k=k_0}^{k_0+D-1}$, we show that if $\mathrm{rank}(\mathbf{H}(w_N^i)) \ne M_t$ for some $i \in \{k\}_{k=k_0}^{k_0+D-1}$, then the matrix Ξ does not have a full column rank. We can construct a proper matrix $\overline{\Xi}$ column equivalent to Ξ by applying elementary column operations [65] to the matrix Ξ, so that the submatrix $\overline{\Xi}[:,1:DM_t]$ can be expressed as the product of a matrix having a full column rank of DM_r and the matrix $\boldsymbol{\Lambda}$ [54]. From the structure of this submatrix, it is observed that if any diagonal element of the block diagonal matrix $\boldsymbol{\Lambda}$ does not have a full column rank, equivalently $\mathrm{rank}(\mathbf{H}(w_N^i)) \ne M_t$ for some $i \in \{k\}_{k=k_0}^{k_0+D-1}$, then $\mathrm{rank}(\overline{\Xi}[:,1:DM_t]) < DM_t$. Thus, the matrix $\overline{\Xi}$ does not have a full column rank. Since $\mathrm{rank}(\overline{\Xi}) = \mathrm{rank}(\Xi)$ [65], the matrix Ξ does not have a full column rank if $\mathrm{rank}(\mathbf{H}(w_N^i)) \ne M_t$ for some $i \in \{k\}_{k=k_0}^{k_0+D-1}$.

APPENDIX II
PROOF OF THEOREM 2

First, we prove that if $\mathrm{span}(\Xi')$ is equal to $\mathrm{span}(\Xi)$, $\mathbf{H}' = \mathbf{H}\boldsymbol{\Omega}$, where $\boldsymbol{\Omega}$ is an $M_t \times M_t$ invertible matrix. From the condition of $\mathrm{span}(\Xi') = \mathrm{span}(\Xi)$, we know that there exists a nonsingular matrix $\boldsymbol{\mathcal{A}}$ satisfying $\Xi' = \Xi\boldsymbol{\mathcal{A}}$. Let us partition the matrix $\boldsymbol{\mathcal{A}}$ as

$$\begin{aligned}\boldsymbol{\mathcal{A}}\,[&(iD+m)M_t+1:(iD+m+1)M_t,\\ &(jD+n)M_t+1:(jD+n+1)M_t] = \boldsymbol{\Omega}_{m,n}^{(i,j)}.\end{aligned} \tag{61}$$

In (61), i, j, m, and n are integers satisfying $0 \le i, j \le J-1$ and $0 \le m,\ n \le D-1$. Performing some mathematical manipulations to $\Xi' = \Xi\boldsymbol{\mathcal{A}}$ and considering the submatrix $\Xi'\,[iQM_r+1:((i+1)Q-L)\,M_r, iDM_t+1:(i+1)DM_t]$ for $0 \le i \le J-1$, we can obtain

$$(\mathbf{K}_1\boldsymbol{\Theta}\otimes\mathbf{I}_{M_r})\begin{bmatrix}\mathbf{H}'(w_N^{k_0})\boldsymbol{\Omega}_{0,n}^{(i,i)}\\ \mathbf{H}'(w_N^{(k_0+1)})\boldsymbol{\Omega}_{1,n}^{(i,i)}\\ \vdots\\ \mathbf{H}'(w_N^{(k_0+n-1)})\boldsymbol{\Omega}_{(n-1),n}^{(i,i)}\\ \mathbf{H}'(w_N^{(k_0+n+1)})\boldsymbol{\Omega}_{(n+1),n}^{(i,i)}\\ \vdots\\ \mathbf{H}'(w_N^{(k_0+D-1)})\boldsymbol{\Omega}_{(D-1),n}^{(i,i)}\end{bmatrix} = \mathbf{0},\quad 0 \le n \le D-1 \tag{62}$$

$$\begin{aligned}\boldsymbol{\Theta} = \mathrm{diag}\,(&\theta(k_0),\ \theta(k_0+1), \ldots,\ \theta(k_0+n-1),\\ &\theta(k_0+n+1), \ldots, \theta(k_0+D-1))\end{aligned} \tag{63}$$

$$\theta(k) = w_N^{-k} - w_N^{-(k_0+n)}. \tag{64}$$

In (62), the matrix $\mathbf{K}_1$ is generated by removing the $(n+1)$th column from the matrix $[\mathbf{W}(N-1)^{\mathrm{T}}, \mathbf{W}(N-2)^{\mathrm{T}}, \ldots, \mathbf{W}(L-P+1)^{\mathrm{T}}]^{\mathrm{T}}$. From the condition $(Q-D) \ge L$ and the structure of $\mathbf{K}_1$ based on a Vandermonde matrix, we can show that the matrix $\mathbf{K}_1$ is a tall matrix with a full column rank. Since the matrix Ξ has a full column rank by the assumption and $\mathrm{span}(\Xi') = \mathrm{span}(\Xi)$, Theorem 1 states that the matrix $\mathbf{H}'(w_N^i)$ for $k_0 \le i \le k_0+D-1$ has a full column rank. Thus, we obtain $\boldsymbol{\Omega}_{m,n}^{(i,i)} = \mathbf{0}$ for $0 \le i \le J-1$, $0 \le m$, $n \le D-1$, and $m \ne n$. By using this result, we can easily show that $\boldsymbol{\Omega}_{m,m}^{(0,0)} = \boldsymbol{\Omega}_{m,m}^{(1,1)} = \cdots = \boldsymbol{\Omega}_{m,m}^{(J-1,J-1)}$ for $0 \le m \le D-1$ as well.

On the other hand, by considering the submatrix $\Xi'[iQM_r+1:((i+1)Q-L)M_r, jDM_t+1:(j+1)DM_t]$ for $0 \le i$, $j \le J-1$ and $i \ne j$, we obtain, for $0 \le n \le D-1$

$$\left(\left[\mathbf{W}(N-1)^{\mathrm{T}}\ \mathbf{W}(N-2)^{\mathrm{T}}, \ldots, \mathbf{W}(L-P)^{\mathrm{T}}\right]^{\mathrm{T}} \otimes \mathbf{I}_{M_r}\right)\begin{bmatrix}\mathbf{H}'(w_N^{k_0})\boldsymbol{\Omega}_{0,n}^{(i,j)}\\ \mathbf{H}'(w_N^{(k_0+1)})\boldsymbol{\Omega}_{1,n}^{(i,j)}\\ \vdots\\ \mathbf{H}'(w_N^{(k_0+D-1)})\boldsymbol{\Omega}_{(D-1),n}^{(i,j)}\end{bmatrix} = \mathbf{0}. \tag{65}$$

Since the matrix $[\mathbf{W}(N-1)^{\mathrm{T}}, \mathbf{W}(N-2)^{\mathrm{T}}, \ldots, \mathbf{W}(L-P)^{\mathrm{T}}]^{\mathrm{T}}$ is a tall matrix with a full column rank and the matrix $\mathbf{H}'(w_N^i)$ for $k_0 \le i \le k_0+D-1$ has a full column rank, we obtain $\boldsymbol{\Omega}_{m,n}^{(i,j)} = \mathbf{0}$ for $0 \le i$, $j \le J-1$, $0 \le m$, $n \le D-1$, and $i \ne j$. Thus, the submatrix $\boldsymbol{\Omega}_{m,m}^{(i,i)}$ for $0 \le i \le J-1$ and $0 \le m \le D-1$ is invertible. By using the above results and rearranging $\boldsymbol{\Xi}'[(Q-L-1)M_r+1:QM_r,1:DM_t] = (\boldsymbol{\Xi}\boldsymbol{\mathcal{A}})[(Q-L-1)M_r+1:QM_r,1:DM_t]$, we obtain

$$(\mathbf{K}_2 \otimes \mathbf{I}_{M_r})(\mathbf{H}' - \mathbf{H}\boldsymbol{\Omega}^{(0,0)}_{k-k_0,\,k-k_0}) = \mathbf{0}, \qquad k_0 \le k \le k_0+D-1 \tag{66}$$

$$\mathbf{K}_2 = \begin{bmatrix} w_N^{(N-Q+L)k} & w_N^{(N-Q+L-1)k} & \cdots & w_N^{(N-P+1)k} & w_N^{(N-P)k} \\ w_N^{(N-Q+L-1)k} & w_N^{(N-Q+L-2)k} & \cdots & w_N^{(N-P)k} & 0 \\ & \iddots & \iddots & \iddots & \vdots \\ w_N^{(N-P)k} & 0 & \cdots & 0 & 0 \end{bmatrix} \tag{67}$$

thereby deriving $\mathbf{H}' = \mathbf{H}\boldsymbol{\Omega}^{(0,0)}_{m,m}$ for $0 \le m \le D-1$. Since the matrix $\mathbf{H}(w_N^i)$ for $k_0 \le i \le k_0+D-1$ has a full column rank, $\operatorname{rank}(([1, w_N^{-i}, \ldots, w_N^{-iL}] \otimes \mathbf{I}_{M_r})\mathbf{H}) = \operatorname{rank}(\mathbf{H}(w_N^i)) = M_t \le \operatorname{rank}(\mathbf{H}) \le M_t$ [65], which means that $\mathbf{H}$ has a full column rank. Thus, we have $\boldsymbol{\Omega}^{(0,0)}_{0,0} = \boldsymbol{\Omega}^{(0,0)}_{1,1} = \cdots = \boldsymbol{\Omega}^{(0,0)}_{D-1,D-1}$. By denoting $\boldsymbol{\Omega}^{(i,i)}_{m,m}$ for $0 \le i \le J-1$ and $0 \le m \le D-1$ as an invertible matrix $\boldsymbol{\Omega}$, we conclude $\mathbf{H}' = \mathbf{H}\boldsymbol{\Omega}$.

Next, we show that if $\mathbf{H}' = \mathbf{H}\boldsymbol{\Omega}$ with an invertible matrix $\boldsymbol{\Omega}$, span$(\boldsymbol{\Xi}')$ is equal to span$(\boldsymbol{\Xi})$. By using $\mathbf{H}' = \mathbf{H}\boldsymbol{\Omega}$, we can write

$$\begin{aligned} \boldsymbol{\Xi}' &= \boldsymbol{\mathcal{H}}'\boldsymbol{\mathcal{W}} \\ &= \boldsymbol{\mathcal{H}}(\mathbf{I}_J \otimes \mathbf{I}_Q \otimes \boldsymbol{\Omega})(\mathbf{I}_J \otimes \mathbf{W} \otimes \mathbf{I}_{M_t}) \\ &= \boldsymbol{\mathcal{H}}(\mathbf{I}_J \otimes \mathbf{W} \otimes \mathbf{I}_{M_t})(\mathbf{I}_J \otimes \mathbf{I}_D \otimes \boldsymbol{\Omega}) \\ &= \boldsymbol{\mathcal{H}}\boldsymbol{\mathcal{W}}(\mathbf{I}_J \otimes \mathbf{I}_D \otimes \boldsymbol{\Omega}) \\ &= \boldsymbol{\Xi}(\mathbf{I}_J \otimes \mathbf{I}_D \otimes \boldsymbol{\Omega}). \end{aligned} \tag{68}$$

Since $\boldsymbol{\Omega}$ is invertible, the matrix $(\mathbf{I}_J \otimes \mathbf{I}_D \otimes \boldsymbol{\Omega})$ is a nonsingular matrix. Thus, this means that span$(\boldsymbol{\Xi}') = $ span$(\boldsymbol{\Xi})$.

Appendix III
Proof of Lemma 1

First, we show that if rank$(\mathbf{H}') = M_t$ and span$(\boldsymbol{\Xi}') = $ span$(\boldsymbol{\Xi})$, then $\mathbf{H}' = \mathbf{H}\boldsymbol{\Omega}$ with an $M_t \times M_t$ invertible matrix $\boldsymbol{\Omega}$. Once we obtain a proof for the case with $J = 2$, the same approach can be applied to a proof for the case with $J = 1$, which is simpler than that of the case with $J = 2$. Thus, our proof is focused on the case with $J = 2$ by setting J to two.

For brevity of notation, we define several matrices as follows:

$$\begin{aligned} \mathbf{B}_1 &\triangleq \boldsymbol{\mathcal{H}}[1:(Q-D)M_r, 1:(Q-D)M_t] \\ \mathbf{B}_2 &\triangleq \boldsymbol{\mathcal{H}}[1:(Q-D)M_r, (Q-D)M_t+1:QM_t] \\ \mathbf{B}_3 &\triangleq \boldsymbol{\mathcal{H}}[(Q-D)M_r+1:(Q+D-L)M_r, \\ &\qquad (Q-D)M_t+1:QM_t] \\ \mathbf{B}_4 &\triangleq \boldsymbol{\mathcal{H}}[(Q-D)M_r+1:(Q+D-L)M_r, \\ &\qquad QM_t+1:(Q+D)M_t] \\ \mathbf{B}_5 &\triangleq \boldsymbol{\mathcal{H}}[(Q+D-L)M_r+1:(2Q-L)M_r, \\ &\qquad QM_t+1:(Q+D)M_t] \\ \mathbf{B}_6 &\triangleq \boldsymbol{\mathcal{H}}[(Q+D-L)M_r+1:(2Q-L)M_r, \\ &\qquad (Q+D)M_t+1:2QM_t] \\ \mathbf{W}_1 &\triangleq \boldsymbol{\mathcal{W}}[1:(Q-D)M_t, 1:DM_t] \\ \mathbf{W}_2 &\triangleq \boldsymbol{\mathcal{W}}[(Q-D)M_t+1:QM_t, 1:DM_t] \\ \mathbf{W}_3 &\triangleq \boldsymbol{\mathcal{W}}[QM_t+1:(Q+D)M_t, DM_t+1:2DM_t] \\ \mathbf{W}_4 &\triangleq \boldsymbol{\mathcal{W}}[(Q+D)M_t+1:2QM_t, DM_t+1:2DM_t]. \end{aligned} \tag{69}$$

By using the notations in (69), we can write the matrix $\boldsymbol{\Xi}$ with $J = 2$ as

$$\begin{aligned} \boldsymbol{\Xi} &= \begin{bmatrix} \mathbf{B}_1 & \mathbf{B}_2 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{B}_3 & \mathbf{B}_4 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{B}_5 & \mathbf{B}_6 \end{bmatrix} \begin{bmatrix} \mathbf{W}_1 & \mathbf{0} \\ \mathbf{W}_2 & \mathbf{0} \\ \mathbf{0} & \mathbf{W}_3 \\ \mathbf{0} & \mathbf{W}_4 \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{Z}_1\mathbf{W}_2 & \mathbf{0} \\ \mathbf{B}_3\mathbf{W}_2 & \mathbf{B}_4\mathbf{W}_3 \\ \mathbf{0} & \mathbf{Z}_2\mathbf{W}_3 \end{bmatrix}. \end{aligned} \tag{70}$$

In (70), the matrices $\mathbf{Z}_1$ and $\mathbf{Z}_2$ represent $\mathbf{B}_1\mathbf{W}_1\mathbf{W}_2^{-1} + \mathbf{B}_2$ and $\mathbf{B}_5 + \mathbf{B}_6\mathbf{W}_4\mathbf{W}_3^{-1}$, respectively. From the assumption of rank$(\mathbf{h}(0)) = M_t$ and rank$(\mathbf{H}(z)) = M_t$ for all z, the polynomial matrix $\mathbf{H}(z)$ is irreducible [39], [66], [67]. In addition, $\mathbf{H}(z)$ is column reduced [39], [66], [67] by the assumption of rank$(\mathbf{h}(L)) = M_t$. Thus, the column polynomial vectors of $\mathbf{H}(z)$ form a minimal polynomial basis [39], [66], [67] for span$(\mathbf{H}(z))$.[1] Since $\mathbf{h}(L)$ has a full column rank, $\deg(\mathbf{H}(z)[:, i]) = L$[2] for all $i \in \{k\}_{k=1}^{M_t}$. Therefore, by Theorem 1 in [39], the submatrix $[\mathbf{B}_3,\ \mathbf{B}_4]$ has a full column rank with the conditions of $M_t < M_r$ and $L \le \lfloor (2D-1)/(M_t+1) \rfloor$. In addition, the submatrix $\boldsymbol{\mathcal{W}}[(Q-D)M_t+1:(Q+D)M_t, :]$ is a block diagonal matrix with $\mathbf{W}_2$ and $\mathbf{W}_3$ on its main diagonal, and $\mathbf{W}_2$ and $\mathbf{W}_3$ are all nonsingular. Thus, the submatrix $\boldsymbol{\Xi}[(Q-D)M_r+1:(Q+D-L)M_r, :]$ has a full column rank, which implies that the matrix $\boldsymbol{\Xi}$ has a full column rank as well.

After we perform the eigenvalue decomposition of the autocorrelation matrix $\mathbf{R}_{\mathbf{rr}}$ of the received signal vector $\mathbf{r}(n)$ in (26) and reorder its eigenvectors and the corresponding eigenvalues, we partition the unitary matrix $\boldsymbol{\mathcal{U}}$ consisting of the eigenvectors and the diagonal matrix $\boldsymbol{\mathcal{D}}$ composed of the eigenvalues as in the expressions in (71) and (72), respectively, shown at the top of the next page.

[1] span$(\mathbf{H}(z))$ denotes the linear subspace over the field of scalar rational functions spanned by the column vectors of $\mathbf{H}(z)$, i.e., the set of rational function vectors written as $\sum_{i=1}^{M_t} c_i(z)\mathbf{H}(z)[:, i]$, where $c_i(z)$ is a scalar rational function (see [50] and references therein).

[2] The degree of a $M_r \times 1$ polynomial vector $\mathbf{H}(z)[:, i] = [\,[\mathbf{H}(z)]_{1,i},\ [\mathbf{H}(z)]_{2,i}, \ldots, [\mathbf{H}(z)]_{M_r,i}]^{\mathrm{T}}$ is defined as the greatest degree of its components as given in $\deg(\mathbf{H}(z)[:, i]) = \max_{1\le k\le M_r} \deg([\mathbf{H}(z)]_{k,i})$ (see [23], [39], [66], and [67] and references therein).

$$\boldsymbol{\mathcal{U}} = \begin{bmatrix} \mathbf{U}_{11} & \mathbf{U}_{12} & \mathbf{U}_{13} & \mathbf{U}_{14} \\ \mathbf{U}_{21} & \mathbf{U}_{22} & \mathbf{U}_{23} & \mathbf{U}_{24} \\ \mathbf{U}_{31} & \mathbf{U}_{32} & \mathbf{U}_{33} & \mathbf{U}_{34} \end{bmatrix} \tag{71}$$

$$\boldsymbol{\mathcal{D}} = \begin{bmatrix} \sigma_\eta^2\mathbf{I}_{(Q-D)M_r} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Sigma} + \sigma_\eta^2\mathbf{I}_{2DM_t} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \sigma_\eta^2\mathbf{I}_{(2D-L)M_r-2DM_t} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \sigma_\eta^2\mathbf{I}_{(Q-D)M_r} \end{bmatrix} \tag{72}$$

In (71), the submatrices $\mathbf{U}_{i1}$, $\mathbf{U}_{i2}$, $\mathbf{U}_{i3}$, and $\mathbf{U}_{i4}$ for $1 \le i \le 3$ have $(Q-D)M_r$, $2DM_t$, $(2D-L)M_r - 2DM_t$, and $(Q-D)M_r$ columns, respectively. The submatrices $\mathbf{U}_{1j}$, $\mathbf{U}_{2j}$, and $\mathbf{U}_{3j}$ for $1 \le j \le 4$ have $(Q-D)M_r$, $(2D-L)M_r$, and $(Q-D)M_r$ rows, respectively. The submatrix $\boldsymbol{\Sigma}$ in (72) is a $2DM_t \times 2DM_t$ diagonal matrix with positive diagonal elements. From (26) and (70) through (72), we obtain the following relationship:

$$[\mathbf{B}_3,\ \mathbf{B}_4] \begin{bmatrix} \mathbf{W}_2 & \mathbf{0} \\ \mathbf{0} & \mathbf{W}_3 \end{bmatrix} \mathbf{R}_{\mathbf{dd}}^{\frac{1}{2}} = \mathbf{U}_{22}\boldsymbol{\Sigma}^{\frac{1}{2}}\mathbf{V}^{\mathrm{H}} \tag{73}$$

where $E\{\mathbf{d}(n)\mathbf{d}(n)^{\mathrm{H}}\} = \mathbf{R}_{\mathbf{dd}} = \mathbf{R}_{\mathbf{dd}}^{1/2}\mathbf{R}_{\mathbf{dd}}^{\mathrm{H}/2}$, $\boldsymbol{\Sigma} = \boldsymbol{\Sigma}^{1/2}\boldsymbol{\Sigma}^{\mathrm{H}/2}$, and $\mathbf{V}$ is an unitary matrix. Assuming that $\mathbf{R}_{\mathbf{dd}}$ has full rank, we obtain $\mathrm{span}([\mathbf{B}_3,\ \mathbf{B}_4]) = \mathrm{span}(\mathbf{U}_{22})$ from (73). Therefore, the matrix $\mathbf{U}_{22}$ has a full column rank. Defining a $(2D-L)M_r \times (2D-L)M_r - 2DM_t$ matrix $\mathbf{U}_{22}^{\perp}$ whose column vectors are linearly independent and span the subspace orthogonal to $\mathrm{span}(\mathbf{U}_{22})$, we can find an invertible matrix $\mathbf{G}$ satisfying the following condition [65]:

$$\boldsymbol{\mathcal{U}}_n = \begin{bmatrix} \mathbf{U}_{11} & \mathbf{U}_{13} & \mathbf{U}_{14} \\ \mathbf{U}_{21} & \mathbf{U}_{23} & \mathbf{U}_{24} \\ \mathbf{U}_{31} & \mathbf{U}_{33} & \mathbf{U}_{34} \end{bmatrix} = \begin{bmatrix} \widetilde{\mathbf{U}}_{11} & \mathbf{0} & \widetilde{\mathbf{U}}_{14} \\ \widetilde{\mathbf{U}}_{21} & \mathbf{U}_{22}^{\perp} & \widetilde{\mathbf{U}}_{24} \\ \widetilde{\mathbf{U}}_{31} & \mathbf{0} & \widetilde{\mathbf{U}}_{34} \end{bmatrix} \mathbf{G} \tag{74}$$

where $\widetilde{\mathbf{U}}_{ij}$ has the same dimension as that of $\mathbf{U}_{ij}$ for $1 \le i \le 3$ and $j = 1, 4$. By defining the matrices $\mathbf{B}_3'$ and $\mathbf{B}_4'$ as, respectively

$$\begin{aligned} \mathbf{B}_3' &\triangleq \boldsymbol{\mathcal{H}}'\,[(Q-D)M_r + 1 : \\ &\qquad (Q+D-L)M_r, (Q-D)M_t + 1 : QM_t] \\ \mathbf{B}_4' &\triangleq \boldsymbol{\mathcal{H}}'\,[(Q-D)M_r + 1 : \\ &\qquad (Q+D-L)M_r, QM_t + 1 : (Q+D)M_t] \end{aligned} \tag{75}$$

and using the condition of $\mathrm{span}(\boldsymbol{\Xi}') = \mathrm{span}(\boldsymbol{\Xi})$ and the orthogonal relationship in (28), we derive

$$\begin{aligned} \boldsymbol{\mathcal{U}}_n^{\mathrm{H}}\boldsymbol{\Xi}' = \mathbf{0} &\Longrightarrow \left(\mathbf{U}_{22}^{\perp}\right)^{\mathrm{H}} [\mathbf{B}_3', \mathbf{B}_4'] \begin{bmatrix} \mathbf{W}_2 & \mathbf{0} \\ \mathbf{0} & \mathbf{W}_3 \end{bmatrix} \\ &= \mathbf{0} \Longleftrightarrow \left(\mathbf{U}_{22}^{\perp}\right)^{\mathrm{H}} [\mathbf{B}_3', \mathbf{B}_4'] = \mathbf{0}. \end{aligned} \tag{76}$$

Furthermore, we know that the column vectors of $\mathbf{U}_{22}^{\perp}$ constitute a basis for the left-hand nullspace of the matrix $[\mathbf{B}_3,\ \mathbf{B}_4]$. By relying on Theorem 2 in [50] along with the relationship in (76) and the condition of $\mathrm{rank}(\mathbf{H}') = M_t$, we obtain $\mathbf{H}'(z) = \mathbf{H}(z)\boldsymbol{\Omega}$ with an $M_t \times M_t$ invertible matrix $\boldsymbol{\Omega}$. That is, $\mathbf{H}' = \mathbf{H}\boldsymbol{\Omega}$.

On the other hand, by following the same procedure as given in (68), we can prove that if $\mathbf{H}' = \mathbf{H}\boldsymbol{\Omega}$ with an invertible matrix $\boldsymbol{\Omega}$, then $\mathrm{span}(\boldsymbol{\Xi}') = \mathrm{span}(\boldsymbol{\Xi})$. In addition, since $\mathbf{H}' = \mathbf{H}\boldsymbol{\Omega}$, we have $\mathbf{h}'(0) = \mathbf{h}(0)\boldsymbol{\Omega}$. By the assumption of $\mathrm{rank}(\mathbf{h}(0)) = M_t$, we can conclude that $\mathbf{H}'$ has a full column rank.

ACKNOWLEDGMENT

The authors would like to thank the anonymous reviewers for their constructive comments, which helped improve this paper.

REFERENCES

[1] L. J. Cimini, Jr., "Analysis and simulation of a digital mobile channel using orthogonal frequency division multiplexing," *IEEE Trans. Commun.*, vol. COM-33, no. 7, pp. 665–675, Jul. 1985.

[2] J. A. C. Bingham, "Multicarrier modulations for data transmission: An idea whose time has come," *IEEE Commun. Mag.*, vol. 28, no. 5, pp. 5–14, May 1990.

[3] G. L. Stüber, J. R. Barry, S. W. Mclaughlin, Y. Li, M. A. Ingram, and T. G. Pratt, "Broadband MIMO-OFDM wireless communications," *Proc. IEEE*, vol. 92, no. 2, pp. 271–294, Feb. 2004.

[4] A. J. Paulraj, D. A. Gore, R. U. Nabar, and H. Bölcskei, "An overview of MIMO communications—A key to gigabit wireless," *Proc. IEEE*, vol. 92, no. 2, pp. 198–218, Feb. 2004.

[5] A. J. Paulraj, R. U. Nabar, and D. A. Gore, *Introduction to Space-Time Wireless Communications*. Cambridge, U.K.: Cambridge Univ. Press, 2003.

[6] D. Gesbert, M. Shafi, D. Shiu, P. J. Smith, and A. Naguib, "From theory to practice: An overview of MIMO space-time coded wireless systems," *IEEE J. Sel. Areas Commun.*, vol. 21, no. 3, pp. 281–302, Apr. 2003.

[7] H. Sampath, S. Talwar, J. Tellado, V. Erceg, and A. Paulraj, "A fourth-generation MIMO-OFDM broadband wireless system: Design, performance, and field trial results," *IEEE Commun. Mag.*, vol. 40, no. 9, pp. 143–149, Sep. 2002.

[8] C. Dubuc, D. Starks, T. Creasy, and Y. Hou, "A MIMO-OFDM prototype for next-generation wireless WANs," *IEEE Commun. Mag.*, vol. 42, no. 12, pp. 82–87, Dec. 2004.

[9] A. van Zelst and T. C. W. Schenk, "Implementation of a MIMO OFDM-based wireless LAN system," *IEEE Trans. Signal Process.*, vol. 52, no. 2, pp. 483–494, Feb. 2004.

[10] Y. Li, N. Seshadri, and S. Ariyavisitakul, "Channel estimation for OFDM systems with transmitter diversity in mobile wireless channels," *IEEE J. Sel. Areas Commun.*, vol. 17, no. 3, pp. 461–471, Mar. 1999.

[11] Y. Li, J. H. Winters, and N. R. Sollenberger, "MIMO-OFDM for wireless communications: Signal detection with enhanced channel estimation," *IEEE Trans. Commun.*, vol. 50, no. 9, pp. 1471–1477, Sep. 2002.

[12] Y. Li, "Simplified channel estimation for OFDM systems with multiple transmit antennas," *IEEE Trans. Wireless Commun.*, vol. 1, no. 1, pp. 67–75, Jan. 2002.

[13] H. Minn, D. I. Kim, and V. K. Bhargava, "A reduced complexity channel estimation for OFDM systems with transmit diversity in mobile wireless channels," *IEEE Trans. Commun.*, vol. 50, no. 5, pp. 799–807, May 2002.

[14] J. Du and Y. Li, "MIMO-OFDM channel estimation based on subspace tracking," in *Proc. IEEE VTC*, Jeju, Korea, Apr. 2003, vol. 2, pp. 1084–1088.

[15] I. Barhumi, G. Leus, and M. Moonen, "Optimal training design for MIMO OFDM systems in mobile wireless channels," *IEEE Trans. Signal Process.*, vol. 51, no. 6, pp. 1615–1624, Jun. 2003.

[16] Y. Gong and K. B. Letaief, "Low complexity channel estimation for space-time coded wideband OFDM systems," *IEEE Trans. Wireless Commun.*, vol. 2, no. 5, pp. 876–882, Sep. 2003.

[17] M. Shin, H. Lee, and C. Lee, "Enhanced channel-estimation technique for MIMO-OFDM systems," *IEEE Trans. Veh. Technol.*, vol. 53, no. 1, pp. 261–265, Jan. 2004.

[18] H. Minn and N. Al-Dhahir, "Optimal training signals for MIMO OFDM channel estimation," in *Proc. IEEE GLOBECOM*, Dallas, TX, Nov. 2004, vol. 1, pp. 219–224.

[19] D. N. Godard, "Self-recovering equalization and carrier tracking in two-dimensional data communication systems," *IEEE Trans. Commun.*, vol. COM-28, no. 11, pp. 1867–1875, Nov. 1980.

[20] M. Ghosh, "Blind decision feedback equalization for terrestrial television receivers," *Proc. IEEE*, vol. 86, no. 10, pp. 2070–2081, Oct. 1998.

[21] 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.

[22] E. Moulines, P. Duhamel, J. 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.

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

[24] W. Qiu and Y. Hua, "Performance analysis of the subspace method for blind channel identification," *Signal Process.*, vol. 50, no. 1/2, pp. 71–81, Apr. 1996.

[25] 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.

[26] S. Yatawatta and A. P. Petropulu, "Blind channel estimation in MIMO OFDM systems," in *Proc. IEEE Workshop Statistical Signal Process.*, St. Louis, MO, Sep. 2003, pp. 363–366.

[27] Y. Zeng and T. S. Ng, "A semi-blind channel estimation method for multiuser multiuser multiantenna OFDM systems," *IEEE Trans. Signal Process.*, vol. 52, no. 5, pp. 1419–1429, May 2004.

[28] 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.

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

[30] J. K. Tugnait and W. Luo, "Blind identification of time-varying channels using multistep linear predictors," *IEEE Trans. Signal Process.*, vol. 52, no. 6, pp. 1739–1749, Jun. 2004.

[31] E. de Carvalho and D. T. M. Slock, "Blind and semi-blind FIR multichannel estimation: (Global) identifiability conditions," *IEEE Trans. Signal Process.*, vol. 52, no. 4, pp. 1053–1064, Apr. 2004.

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

[33] J. K. Tugnait, "A multidelay whitening approach to blind identification and equalization of SIMO channels," *IEEE Trans. Wireless Commun.*, vol. 1, no. 3, pp. 456–467, Jul. 2002.

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

[35] J. K. Tugnait, "Blind estimation and equalization of MIMO channels via multidelay whitening," *IEEE J. Sel. Areas Commun.*, vol. 19, no. 8, pp. 1507–1519, Aug. 2001.

[36] L. Perros-Meilhac, E. Moulines, K. Abed-Meraim, P. Chevalier, and P. Duhamel, "Blind identification of multipath channels: A parametric subspace approach," *IEEE Trans. Signal Process.*, vol. 49, no. 7, pp. 1468–1480, Jul. 2001.

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

[38] T. P. Krauss and M. D. Zoltowski, "Bilinear approach to multiuser second-order statistics-based blind channel estimation," *IEEE Trans. Signal Process.*, vol. 48, no. 9, pp. 2473–2486, Sep. 2000.

[39] P. Loubaton and E. Moulines, "On blind multiuser forward link channel estimation by the subspace method: Identifiability results," *IEEE Trans. Signal Process.*, vol. 48, no. 8, pp. 2366–2376, Aug. 2000.

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

[41] A. Gorokhov and P. Loubaton, "Blind identification of MIMO-FIR systems: A generalized linear prediction approach," *Signal Process.*, vol. 73, no. 1, pp. 105–124, Jan. 1999.

[42] J. K. Tugnait, "On linear predictors for MIMO channels and related blind identification and equalization," *IEEE Signal Process. Lett.*, vol. 5, no. 11, pp. 289–291, Nov. 1998.

[43] 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.

[44] K. Abed-Meraim, Y. Hua, P. Loubaton, and E. Moulines, "Subspace method for blind identification of multichannel FIR systems in noise field with unknown spatial covariance," *IEEE Signal Process. Lett.*, vol. 4, no. 5, pp. 135–137, May 1997.

[45] Y. Hua, K. Abed-Meraim, and M. Wax, "Blind system identification using minimum noise subspace," *IEEE Trans. Signal Process.*, vol. 45, no. 3, pp. 770–773, Mar. 1997.

[46] K. Abed-Meraim, E. Moulines, and P. Loubaton, "Prediction error method for second-order blind identification," *IEEE Trans. Signal Process.*, vol. 45, no. 3, pp. 694–705, Mar. 1997.

[47] D. T. M. Slock, "Blind fractionally-spaced equalization, perfect-reconstruction filter banks and multichannel linear prediction," in *Proc. IEEE ICASSP Conf.*, Adalaide, Australia, Apr. 1994, vol. 4, pp. 585–588.

[48] H. Liu, G. Xu, L. Tong, and T. Kailath, "Recent developments in blind channel equalization: From cyclostationarity to subspaces," *Signal Process.*, vol. 50, no. 1/2, pp. 83–99, Apr. 1996.

[49] 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.

[50] G. B. Giannakis, Y. Hua, P. Stoica, and L. Tong, *Signal Processing Advances in Wireless and Mobile Communications*, vol. 1. Englewood Cliffs, NJ: Prentice-Hall, 2001.

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

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

[53] B. Muquet, M. de Courville, and P. Duhamel, "Subspace-based blind and semi-blind channel estimation for OFDM systems," *IEEE Trans. Signal Process.*, vol. 50, no. 7, pp. 1699–1712, Jul. 2002.

[54] C. Li and S. Roy, "Subspace-based blind channel estimation for OFDM by exploiting virtual carriers," *IEEE Trans. Wireless Commun.*, vol. 2, no. 1, pp. 141–150, Jan. 2003.

[55] X. Cai and A. N. Akansu, "A subspace method for blind channel identification in OFDM systems," in *Proc. IEEE ICC*, New Orleans, LA, Jun. 2000, vol. 2, pp. 929–933.

[56] R. van Nee and R. Prasad, *OFDM for Wireless Multimedia Communications*. Norwood, MA: Artech House, 2000.

[57] H. Sari, G. Karam, and I. Jeanclaude, "Transmission techniques for digital terrestrial TV broadcasting," *IEEE Commun. Mag.*, vol. 33, no. 2, pp. 100–109, Feb. 1995.

[58] R. W. Heath, Jr. and G. B. Giannakis, "Exploiting input cyclostationarity for blind channel identification in OFDM systems," *IEEE Trans. Signal Process.*, vol. 47, no. 3, pp. 848–856, Mar. 1999.

[59] W. Bai, C. He, L. Jiang, and H. Zhu, "Blind channel estimation in MIMO-OFDM systems," in *Proc. IEEE GLOBECOM*, Taipei, Taiwan, R.O.C., Nov. 2002, vol. 1, pp. 317–321.

[60] M. Speth, S. A. Fechtel, G. Fock, and H. Meyr, "Optimum receiver design for wireless broad-band systems using OFDM—Part I," *IEEE Trans. Commun.*, vol. 47, no. 11, pp. 1668–1677, Nov. 1999.

[61] J.-F. Cardoso, "Blind signal separation: Statistical principles," *Proc. IEEE*, vol. 86, no. 10, pp. 2009–2025, Oct. 1998.

[62] S. Visuri and V. Koivunen, "Resolving ambiguities in subspace-based blind receiver for MIMO channels," in *Proc. Asilomar Conf.*, Pacific Grove, CA, Nov. 2002, vol. 1, pp. 589–593.
[63] D. K. Borah, R. A. Kennedy, Z. Ding, and I. Fijalkow, "Sampling and prefiltering effects on blind equalizer design," *IEEE Trans. Signal Process.*, vol. 49, no. 1, pp. 209–218, Jan. 2001.
[64] A. Edelman, T. A. Arias, and S. T. Smith, "The geometry of algorithms with orthogonality constraints," *SIAM J. Matrix Anal. Appl.*, vol. 20, no. 2, pp. 303–353, 1998.
[65] C. D. Meyer, *Matrix Analysis and Applied Linear Algebra*. Philadelphia, PA: SIAM, 2000.
[66] G. D. Forney, Jr., "Minimal bases of rational vector spaces, with applications to multivariable linear systems," *SIAM J. Control*, vol. 13, no. 3, pp. 493–520, May 1975.
[67] T. Kailath, *Linear Systems*. Englewood Cliffs, NJ: Prentice-Hall, 1980.

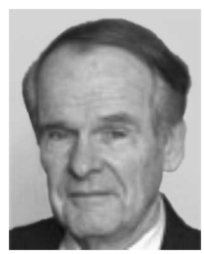
Changyong Shin (S'04–M'07) received the B.S. and M.S. degrees from Yonsei University, Seoul, Korea, in 1993 and 1995, respectively, and the Ph.D. degree from the University of Texas at Austin in 2006, all in electrical engineering.

From 1995 to 2001, he was a Senior Research Engineer at LG Electronics Inc., Seoul, where he worked on digital video signal processing and very large scale integration design for digital signal processing. He is currently with Samsung Advanced Institute of Technology, Gyunggi-do, Korea. His research interests cover multi-input–multi-output communications, multicarrier modulation, and signal processing for communications including channel estimation, signal detection, interference cancellation, space–time processing, and synchronization.

Robert W. Heath, Jr. (S'96–M'01–SM'06) received the B.S. and M.S. degrees from the University of Virginia, Charlottesville, in 1996 and 1997, respectively, and the Ph.D. degree from Stanford University, Stanford, CA, in 2002, all in electrical engineering.

From 1998 to 2001, he was a Senior Member of the Technical Staff and then a Senior Consultant with Iospan Wireless Inc., San Jose, CA, where he worked on the design and implementation of the physical and link layers of the first commercial MIMO-OFDM communication system. In 2003, he founded MIMO Wireless Inc.: a consulting company dedicated to the advancement of MIMO technology. Since January 2002, he has been with the Department of Electrical and Computer Engineering, University of Texas at Austin, where he is currently an Assistant Professor and a member of the Wireless Networking and Communications Group. His research interests cover a broad range of MIMO communication, including limited feedback techniques, multihop networking, multiuser MIMO, antenna design, and scheduling algorithms, as well as 60-GHz communication techniques.

Dr. Heath serves as an Editor for the IEEE TRANSACTIONS ON COMMUNICATIONS, and an Associate Editor for the IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY, and he is a member of the Signal Processing for Communications Technical Committee of the IEEE Signal Processing Society.



Edward J. Powers (F'83) received the B.S., M.S., and Ph.D. degrees from Tufts University, Medford, MA, the Massachusetts Institute of Technology, Cambridge, and Stanford University, Stanford, CA, respectively, all in electrical engineering.

He is the Texas Atomic Energy Research Foundation Professor in engineering and Professor of electrical and computer engineering with the University of Texas at Austin. His primary professional interests lie in the innovative application of digital higher order statistical signal processing in the analysis, interpretation, and modeling of time series data characterizing nonlinear physical phenomena, and the utilization of the wavelet transform and time–frequency analysis to detect and identify transient events in various physical systems.

Dr. Powers served as a chair of his department from 1981 to 1988. He is a past Editor of the IEEE TRANSACTIONS ON PLASMA SCIENCE.