Subspace-Based Blind Channel Identification for Cyclic Prefix Systems Using Few Received Blocks

Borching Su, *Student Member, IEEE*, and P. P. Vaidyanathan, *Fellow, IEEE*

***Abstract*—In this paper, a novel generalization of subspace-based blind channel identification methods in cyclic prefix (CP) systems is proposed. For the generalization, a new system parameter called *repetition index* is introduced whose value is unity for previously reported special cases. By choosing a repetition index larger than unity, the number of received blocks needed for blind identification is significantly reduced compared to all previously reported methods. This feature makes the method more realistic especially in wireless environments where the channel state is usually fast-varying. Given the number of received blocks available, the minimum value of repetition index is derived. Theoretical limit allows the proposed method to perform blind identification using only three received blocks in absence of noise. In practice, the number of received blocks needed to yield a satisfactory bit-error-rate (BER) performance is usually on the order of half the block size. Simulation results not only demonstrate the capability of the algorithm to perform blind identification using fewer received blocks, but also show that in some cases system performance can be improved by choosing a repetition index larger than needed. Simulation of the proposed method over time-varying channels clearly demonstrates the improvement over previously reported methods.**

***Index Terms*—Blind identification, cyclic prefix, orthogonal frequency division multiplexing (OFDM), repetition index, single-carrier cyclic prefix (SC-CP), subspace-based methods.**

I. Introduction

In recent years, linear redundant precoding (LRP) has become popular in digital communication systems due to its capability to facilitate block channel equalization of frequency-selective channels. By inserting in each transmitting block a redundant segment of a length greater than or equal to the channel order, the interblock interference (IBI) at the receiver can be eliminated [11]. Two major types of LRP techniques are zero-padding (ZP) and cyclic prefixing (CP). ZP systems guarantee symbol recovery regardless of channel null locations, but the CP technique is more widely used in many current standards such as orthogonal frequency division multiplexing (OFDM) and single-carrier cyclic prefix (SC-CP) systems.

Besides the capability of LRP to facilitate block equalization, the redundancy introduced in the transmitter is also useful for blind channel identification problems. Unlike earlier works in the blind identification literature which require either higher-order statistics (HOS) of the received data [13] or over-sampling at the receiver [6], [17], blind channel identification exploiting LRP requires only second-order statistics (SOS) of the received data and is robust to channel order overestimation. Scaglione *et al.* proposed a blind identification algorithm for ZP-based communication systems [12] based on subspace decomposition. The method requires only the persistency of excitation (p.o.e) property of the input signal (i.e., richness) to render the data covariance matrix to have full rank. This requirement demands the receiver to collect at least a number of blocks equal to the block size for one channel estimate and thus makes the approach less applicable when the channel is fast-varying. More recently, Pham and Manton proposed a subspace-based method in ZP systems which requires only *two* received blocks [10]. A recent generalization of subspace-based blind methods in ZP systems was proposed by Su and Vaidyanathan [15], [14]. There are some other nonsubspace-based blind methods for ZP systems which can perform blind identification using a small number of received blocks. These methods, however, often take advantage of finite alphabet property of the input symbols [19], and are not considered in this paper.

We consider blind identification methods for CP-based systems which are currently more popular than ZP-based systems in wireless standards such as OFDM systems. Similar to those for ZP systems, most existing blind identification methods for CP/OFDM systems fall into either subspace-based or nonsubspace-based categories [5], [9], [19]. We consider only subspace approaches in this paper since they require no knowledge of symbol constellations and hence the computational complexity does not grow as the constellation size increases. Since the guard interval is nonzero in CP systems, subspace-based blind identification methods in CP systems usually have a more sophisticated design than those in ZP systems. Heath *et al.* exploited cyclostationarity induced by cyclic prefixes to estimate channel coefficients blindly [2]. Muquet *et al.* proposed an algorithm using the second order statistics of the received data for OFDM systems of which a semiblind adaptation was also developed [7], [8]. Cai and Akansu proposed another deterministic algorithm of blind channel estimation for OFDM systems [1]. Li and Roy further exploited the presence of virtual carriers of OFDM systems [4]. Zhuang *et al.* proposed a statistical method that can estimate channels whose length is larger than the CP length. All these previously reported methods require the number of received blocks to be at least as large as *two times the block size* to satisfy the persistency of excitation (p.o.e) criterion of the input, which limits the application in a fast-varying channel environment.

Manuscript received August 10, 2006; revised January 4, 2007. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Petr Tichavsky. This work was supported in part by the NSF by Grant CCF-0428326, by the ONR by Grant N00014-06-1-0011, and by the Moore Fellowship of the California Institute of Technology.

The authors are with the California Institute for Technology, Pasadena, CA 91125 USA (e-mail: borching@gmail.com).

Digital Object Identifier 10.1109/TSP.2007.896262

In this paper, we propose a generalization to some of previously reported subspace-based blind methods for CP systems [1], [4], and [8] by introducing a new system parameter called *repetition index*, whose value is unity for these previously reported methods. When the repetition index is chosen to be greater than unity, the number of received blocks needed will be significantly reduced. The rest of the paper is organized as follows. In Section II, we review the basic ideas of blind identification methods in CP systems that have been known so far in [1], [4], [8]. In Section III, we present the generalized algorithm. In Section V, simulations of the proposed algorithm are performed both in static and time-varying channel environments and the results are presented. In Section IV the conditions under which the proposed algorithm works are studied in detail. Conclusions are made in Section VI. Some results contained in this paper have been submitted to a conference [16].

II. Problem Formulation

A. Notations

Boldfaced lower case letters represent column vectors. Boldfaced upper case letters and calligraphic upper case letters are reserved for matrices. Superscripts *, ^{T}, and † as in a^*, $\mathbf{A}^T$, and $\mathbf{A}^\dagger$ denote the conjugate, transpose, and transpose-conjugate operations, respectively. $[\mathbf{v}]_k$ denotes the kth entry of vector $\mathbf{v}$. All the vectors and matrices in this paper are complex-valued. The notation W_M denotes $e^{j2\pi/M}$, and $\mathbf{W}_M$ is the $M\times M$ normalized DFT matrix whose klth entry is $W_M^{-(k-1)(l-1)}/\sqrt{M}$. Column and row indices of all matrices and vectors begin at one. $\mathbf{A}_{k,l}$ is the entry at the kth row and the lth column of $\mathbf{A}$. $\mathbf{I}_M$ is the $M\times M$ identity matrix, and $\mathbf{0}_{m\times n}$ is the $m\times n$ zero matrix.

Notations for commonly used matrix structures in this paper are presented. If $\mathbf{v}=[v_1 \quad v_2 \quad \cdots \quad v_m]^T$ is an $m\times 1$ vector, we use $\mathcal{T}_n(\mathbf{v})$ to denote the $(m+n-1)\times n$ full-banded Toeplitz matrix

$$\mathcal{T}_n(\mathbf{v})=\begin{bmatrix} v_1 & 0 & \cdots & 0 \\ v_2 & v_1 & \ddots & \vdots \\ \vdots & v_2 & \ddots & 0 \\ v_m & \vdots & \ddots & v_1 \\ 0 & v_m & & v_2 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & v_m \end{bmatrix} \tag{1}$$

and $\mathcal{K}_l(\mathbf{v})$ to denote the $l\times(m-l+1)$ Hankel matrix

$$\mathcal{K}_l(\mathbf{v})=\begin{bmatrix} v_1 & v_2 & v_3 & \cdots & v_{m-l+1} \\ v_2 & v_3 & \iddots & \iddots & \vdots \\ \vdots & \iddots & \iddots & \iddots & v_{m-1} \\ v_l & \cdots & \cdots & v_{m-1} & v_m \end{bmatrix}. \tag{2}$$

Due to the special property of cyclic prefixes, we will use the following notation extensively in this paper. Suppose $\mathbf{y}$ is an $m\times 1$ column vector $\mathbf{y}=[y_1 \quad y_2 \cdots \quad y_m]^T$. Then the notation $[\mathbf{y}]_b^a$ denotes the $(b-a+1)\times 1$ vector

$$[\mathbf{y}]_b^a=[y_a \quad y_{a+1} \quad \cdots \quad y_b]^T$$

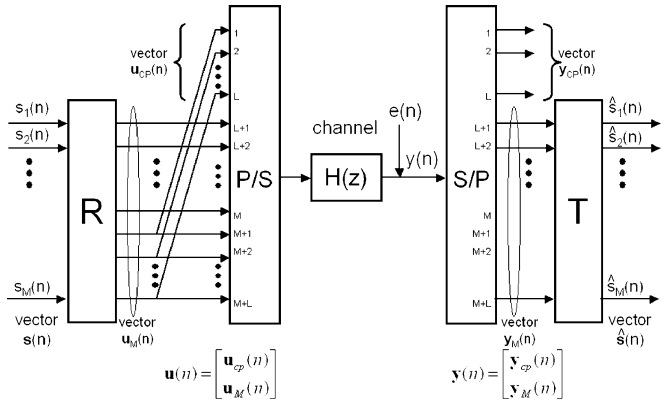

Fig. 1. A typical cyclic prefix system.

if $1\le a\le b\le m$. An extension of this definition to any arbitrary pair of integers a and b satisfying $a\le b$ is made by defining y_k as $y_{(k-1 \bmod m)+1}$ for any $k>m$ or $k<1$. For example, if $\mathbf{y}=[y_1 \quad y_2 \quad y_3]^T$, then $[\mathbf{y}]_7^{-1}$ denotes the vector $[y_2 \quad y_3 \quad y_1 \quad y_2 \quad y_3 \quad y_1 \quad y_2 \quad y_3 \quad y_1]^T$.

B. Cyclic Prefix System Overview

Consider the communication system using cyclic prefix (CP) depicted in Fig. 1. The source symbols $s_1(n), s_2(n), \ldots, s_M(n)$ may come from M different users or from a serial-to-parallel operation on data of a single user. For convenience, we consider the blocked version $\mathbf{s}(n)$ as indicated. The vector $\mathbf{s}(n)$ is precoded by an $M\times M$ constant matrix $\mathbf{R}$ and results in precoded data $\mathbf{u}_M(n)$. In particular, for OFDM or multicarrier (MC) systems, $\mathbf{R}=\mathbf{W}_M^\dagger$ is the normalized IDFT matrix; for single-carrier cyclic prefix (SC-CP) systems, $\mathbf{R}$ is chosen as $\mathbf{I}_M$. A cyclic prefix of length L, taking from the last L elements of $\mathbf{u}_M(n)$, is defined as $\mathbf{u}_{\rm cp}(n)=[\mathbf{0}_{L\times(M-L)} \quad \mathbf{I}_L]\mathbf{u}_M(n)$. We assume $L+1<M$. The cyclic prefix is appended to $\mathbf{u}_M(n)$, forming a vector

$$\mathbf{u}(n)=\begin{bmatrix}\mathbf{u}_{\rm cp}(n)\\ \mathbf{u}_M(n)\end{bmatrix}=[\mathbf{u}_M(n)]_M^{-L+1}$$

whose length is $P=M+L$. The vector $\mathbf{u}(n)$, after parallel-to-serial conversion, is sent over the channel $H(z)$. We assume $H(z)$ is an FIR channel with a maximum order L, i.e.

$$H(z)=\sum_{k=0}^{L}h_k z^{-k} \tag{3}$$

and define $\mathbf{h}$ as the $(L+1)$-column vector $[h_0 \quad h_1 \quad \cdots \quad h_L]^T$. The signal is corrupted by channel noise $e(n)$. The received symbols $y(n)$ are blocked into $P\times 1$ vectors $\mathbf{y}(n)$. We assume perfect block synchronization between the transmitter and receiver. Also let $\mathbf{e}(n)$ denote the blocked version of the noise $e(n)$. Denote $\mathbf{y}_{\rm cp}(n)$ as the first L entries and $\mathbf{y}_M(n)$ as the last M entries of $\mathbf{y}(n)$ so that $\mathbf{y}(n)=[\mathbf{y}_{\rm cp}(n)^T \quad \mathbf{y}_M(n)^T]^T$. It can be shown that

$$\mathbf{y}_M(n)=\mathbf{H}_{\rm cir}\mathbf{u}_M(n)+\mathbf{e}_M(n) \tag{4}$$

where

$$\mathbf{H}_{\text{cir}} = \begin{bmatrix} h_0 & \mathbf{0} & h_L & \cdots & h_1 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ h_L & & \ddots & \ddots & h_L \\ & \ddots & & \ddots & \mathbf{0} \\ \mathbf{0} & & h_L & \cdots & h_0 \end{bmatrix}$$

is an $M \times M$ circulant matrix [18] and $\mathbf{e}_M(n) = [\mathbf{e}(n)]_P^{L+1}$ is the noise vector. The $L \times 1$ vector $\mathbf{y}_{\text{cp}}(n)$ contains inter-block interference (IBI) and can be expressed as

$$\mathbf{y}_{\text{cp}}(n) = \mathbf{H}_l \mathbf{u}_{\text{cp}}(n) + \mathbf{H}_u \mathbf{u}_{\text{cp}}(n-1) + \mathbf{e}_{\text{cp}}(n) \tag{5}$$

where

$$\mathbf{H}_l \triangleq \begin{bmatrix} h_0 & & \mathbf{0} \\ \vdots & \ddots & \\ h_{L-1} & \cdots & h_0 \end{bmatrix} \quad \text{and} \quad \mathbf{H}_u \triangleq \begin{bmatrix} h_L & \cdots & h_1 \\ & \ddots & \vdots \\ \mathbf{0} & & h_L \end{bmatrix}$$

are $L \times L$ matrices and $\mathbf{e}_{\text{cp}}(n) = [\mathbf{e}(n)]_L^1$ is the noise component. For channel equalization, $\mathbf{y}_{\text{cp}}(n)$ is usually dropped and only $\mathbf{y}_M(n)$ passes the $M \times M$ equalizer $\mathbf{T}$ and results in recovered symbol $\hat{\mathbf{s}}(n)$. When the channel coefficients are known, the optimal equalizer $\mathbf{T}$ can be derived to minimize mean square error (MSE) of equalized symbols.

C. Subspace-Based Blind Channel Identification in CP Systems

While $\mathbf{y}_{\text{cp}}(n)$ is often dropped before equalization, the information from $\mathbf{y}_{\text{cp}}(n)$ is useful to estimate the channel coefficients. In this section we review the essences of blind identification algorithms which have been used in earlier methods reported in [1], [8], and [4]. For simplicity we first ignore the noise term $\mathbf{e}(n)$. Define a *composite block* $\bar{\mathbf{y}}(n)$ which has a length $2M + L$ and contains information from two consecutive blocks as follows:

$$\bar{\mathbf{y}}(n) = [\mathbf{y}_M(n-1)^T \quad \mathbf{y}_{\text{cp}}(n)^T \quad \mathbf{y}_M(n)^T]^T. \tag{6}$$

Then from (4) and (5) we have

$$\bar{\mathbf{y}}(n) = \begin{bmatrix} \mathbf{H}_{\text{cir}} \mathbf{u}_M(n-1) \\ \mathbf{H}_l \mathbf{u}_{\text{cp}}(n) + \mathbf{H}_u \mathbf{u}_{\text{cp}}(n-1) \\ \mathbf{H}_{\text{cir}} \mathbf{u}_M(n) \end{bmatrix} = \tilde{\mathbf{H}} \tilde{\mathbf{u}}(n) \tag{7}$$

where

$$\tilde{\mathbf{H}} = \left[\begin{array}{c|c} \mathbf{H}_{cir} & \mathbf{0}_{M\times M} \\ \hline \begin{array}{|c|c|}\hline \mathbf{0}_{L\times(M-L)} & \mathbf{H}_u \\ \hline \end{array} & \begin{array}{|c|c|}\hline \mathbf{0}_{L\times(M-L)} & \mathbf{H}_l \\ \hline \end{array} \\ \hline \mathbf{0}_{M\times M} & \mathbf{H}_{cir} \end{array}\right]$$

and $\tilde{\mathbf{u}}(n) = [\mathbf{u}_M(n-1)^T \quad \mathbf{u}_M(n)^T]^T$. Note that $\tilde{\mathbf{H}}$ is a $(2M + L) \times 2M$ matrix. A special case of (7) when $M = 4$ and $L = 2$ is shown in (8) at the bottom of the page. For notational convenience, we set $y_{0k} = [\mathbf{y}_M(n-1)]_k$, $y_{1k} = [\mathbf{y}_M(n)]_k$, and $y_{\text{cp}k} = [\mathbf{y}_{\text{cp}}(n)]_k$ in (8).

Theorem 1: If $H(z) = \sum_{k=0}^{L} h_k z^{-k}$ does not have any zero on the unit circle grid $W_M^l, 0 \le l \le M-1$, then $\tilde{\mathbf{H}}$ has full column rank $2M$.

Proof: See [8]. ■

We now review some of the key ideas in [8]. Suppose we gather J consecutive received blocks $\mathbf{y}(0), \mathbf{y}(1), \ldots, \mathbf{y}(J-1)$ at the receiver, then we have $J-1$ composite blocks $\bar{\mathbf{y}}(n)$ defined in (6) for $n = 1, 2, \ldots, J-1$. We can construct the $(2M + L) \times (J-1)$ matrix by placing these composite blocks together as

$$\mathbf{Y}^{(J)} = [\bar{\mathbf{y}}(1) \quad \bar{\mathbf{y}}(2) \quad \cdots \quad \bar{\mathbf{y}}(J-1)].$$

Then, we have

$$\mathbf{Y}^{(J)} = \tilde{\mathbf{H}} \mathbf{U}^{(J)} \tag{9}$$

where

$$\mathbf{U}^{(J)} = [\tilde{\mathbf{u}}(1) \quad \tilde{\mathbf{u}}(2) \quad \cdots \quad \tilde{\mathbf{u}}(J-1)]$$

$$\left[\begin{array}{c} y_{01} \\ y_{02} \\ y_{03} \\ y_{04} \\ \hline y_{cp1} \\ y_{cp2} \\ \hline y_{11} \\ y_{12} \\ y_{13} \\ y_{14} \end{array}\right] = \left[\begin{array}{cccc|cccc} h_0 & 0 & h_2 & h_1 & 0 & 0 & 0 & 0 \\ h_1 & h_0 & 0 & h_2 & 0 & 0 & 0 & 0 \\ h_2 & h_1 & h_0 & 0 & 0 & 0 & 0 & 0 \\ 0 & h_2 & h_1 & h_0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & h_2 & h_1 & 0 & 0 & h_0 & 0 \\ 0 & 0 & 0 & h_2 & 0 & 0 & h_1 & h_0 \\ \hline 0 & 0 & 0 & 0 & h_0 & 0 & h_2 & h_1 \\ 0 & 0 & 0 & 0 & h_1 & h_0 & 0 & h_2 \\ 0 & 0 & 0 & 0 & h_2 & h_1 & h_0 & 0 \\ 0 & 0 & 0 & 0 & 0 & h_2 & h_1 & h_0 \end{array}\right] \left[\begin{array}{c} u_{01} \\ u_{02} \\ u_{03} \\ u_{04} \\ \hline u_{11} \\ u_{12} \\ u_{13} \\ u_{14} \end{array}\right] \tag{8}$$

is a $2M \times (J-1)$ matrix. Assume there exists an integer $J \geq 2M+1$ such that $\mathbf{U}^{(J)}$ has full row rank $2M$. Then $\mathrm{rank}(\mathbf{Y}^{(J)}) = 2M$ and hence $\mathbf{Y}^{(J)}$ has L linearly independent left annihilators. Let $\mathbf{g}_k^\dagger$ be the kth annihilator of $\mathbf{Y}^{(J)}$, $1 \leq k \leq L$, i.e., $\mathbf{g}_k^\dagger \mathbf{Y}^{(J)} = \mathbf{0}$. Then $\mathbf{g}_k^\dagger \tilde{\mathbf{H}} = \mathbf{0}$ since $\mathbf{U}^{(J)}$ has full rank. Write $\mathbf{g}_k^\dagger$ as

$$\mathbf{g}_k^\dagger = [g_{01} \quad \cdots \quad g_{0M} | g_{c1} \quad \cdots \quad g_{cL} | g_{11} \quad \cdots \quad g_{1M}].$$

For notational simplicity, we ignore the index k in the contents of $\mathbf{g}_k^\dagger$. By observing the columns of $\tilde{\mathbf{H}}$, we can construct a $2M \times (L+1)$ matrix $\mathcal{G}_k$ as follows such that $\mathcal{G}_k \mathbf{h} = \mathbf{0}$.

$$\mathcal{G}_k = \left[\begin{array}{cccc} g_{01} & g_{02} & \cdots & g_{0,1+L} \\ g_{02} & g_{03} & \cdots & g_{0,2+L} \\ \vdots & \vdots & & \vdots \\ g_{0,M-L} & g_{0,M-L+1} & \cdots & g_{0M} \\ \hline g_{0,M-L+1} & \cdots & g_{0M} & g_{01}+g_{c1} \\ \vdots & \iddots & \iddots & \vdots \\ g_{0M} & g_{01}+g_{c1} & \cdots & g_{0L}+g_{cL} \\ \hline g_{11} & g_{12} & \cdots & g_{1,1+L} \\ g_{12} & g_{13} & \cdots & g_{1,2+L} \\ \vdots & \vdots & & \vdots \\ g_{1,M-L} & g_{1,M-L+1} & \cdots & g_{1M} \\ \hline g_{1,M-L+1}+g_{c1} & \cdots & g_{1M}+g_{cL} & g_{11} \\ \vdots & \iddots & \iddots & \vdots \\ g_{1M}+g_{cL} & g_{11} & \cdots & g_{1L} \end{array}\right]. \tag{10}$$

Define $\mathcal{G} = [\mathcal{G}_1^T \quad \mathcal{G}_2^T \quad \cdots \quad \mathcal{G}_L^T]^T$. Then the channel coefficients $\mathbf{h}$ can be recovered within a scalar ambiguity by finding the only right-annihilating vector of $\mathcal{G}$ [8].

Although the developments above are based on the assumption that $\tilde{\mathbf{H}}$ has full column rank (i.e., $H(z)$ has no zeros on DFT grid), a slight modification of the algorithm when this is not true can be found in [8]. Due to the length of the text, we do not elaborate this special case throughout this paper.

In presence of noise, (9) becomes

$$\mathbf{Y}^{(J)} = \tilde{\mathbf{H}} \mathbf{U}^{(J)} + \mathbf{N}^{(J)}$$

where the noise component $\mathbf{N}^{(J)}$ comes accordingly from (4) and (5). In this case, $\mathbf{Y}^{(J)}$ usually becomes full rank and no longer has L left annihilators. The left annihilators of $\tilde{\mathbf{H}}$, i.e., the noise space, can be estimated by taking singular value decomposition (SVD) of $\mathbf{Y}^{(J)}$. In the equation

$$\mathbf{Y}^{(J)} = [\mathbf{U}_s \quad \mathbf{U}_n] \begin{bmatrix} \mathbf{\Sigma}_s & \mathbf{0} \\ \mathbf{0} & \mathbf{\Sigma}_n \end{bmatrix} [\mathbf{V}_s \quad \mathbf{V}_n]^\dagger \tag{11}$$

$\mathbf{U}_n$ contains the singular-vectors associated with the smallest L singular values of $\mathbf{Y}^{(J)}$ and $\mathbf{g}_k$ is chosen as the kth column of $\mathbf{U}_n$.

Note that in (11) if the matrix $\mathbf{Y}^{(J)}$ is replaced with the estimated autocorrelation matrix

$$\mathbf{R}_{\bar{\mathbf{y}}\bar{\mathbf{y}}} = \mathbf{Y}^{(J)} \left[\mathbf{Y}^{(J)}\right]^\dagger$$

then the null space $\mathbf{U}_n$ obtained by singular value decomposition will remain unchanged. Since the size of $\mathbf{R}_{\bar{\mathbf{y}}\bar{\mathbf{y}}}$ is usually smaller than $\mathbf{Y}^{(J)}$, especially when J is large, taking SVD on $\mathbf{R}_{\bar{\mathbf{y}}\bar{\mathbf{y}}}$ rather than on $\mathbf{Y}^{(J)}$ actually saves computational complexity, although an additional computation will be needed for creating matrix $\mathbf{R}_{\bar{\mathbf{y}}\bar{\mathbf{y}}}$. However, the matrix $\mathbf{R}_{\bar{\mathbf{y}}\bar{\mathbf{y}}}$, once created, can be easily updated each time a new block is received (see [8, eq. (18)]). The idea of maintaining an autocorrelation matrix further develops into a strategy where newer blocks can be put a greater weighting than older blocks. Specifically, after an initial estimate of $\mathbf{R}_{\bar{\mathbf{y}}\bar{\mathbf{y}}}$ is established, $\mathbf{R}_{\bar{\mathbf{y}}\bar{\mathbf{y}}}$ is updated each time a new composite block $\bar{\mathbf{y}}(n)$ is obtained using

$$\hat{\mathbf{R}}_{\bar{\mathbf{y}}\bar{\mathbf{y}}}^{(N)} = \alpha \hat{\mathbf{R}}_{\bar{\mathbf{y}}\bar{\mathbf{y}}}^{(N-1)} + (1-\alpha) \bar{\mathbf{y}}(N) \bar{\mathbf{y}}(N)^\dagger. \tag{12}$$

The parameter $\alpha \in [0,1]$ is called the *forgetting factor*. The technique of using a forgetting factor has been applied especially in time-varying channel environments.

D. Limitations

In order for the aforementioned method to work, the $2M \times (J-1)$ matrix $\mathbf{U}^{(J)}$ must have full row rank $2M$. This is also known as the property of persistency of excitation [8]. Obviously, $\mathbf{U}^{(J)}$ has full row rank only when the number of columns is not smaller than the number of rows, i.e., $J-1 \geq 2M$. This requires the receiver to wait for at least $(2M+1)P$ symbol durations before a channel estimation can be performed. This limitation makes these previously reported algorithms unrealistic in environments with fast-fading channels since the channel coefficients may have changed significantly during accumulation of the data. Even though a forgetting factor can be used to give a larger weighting to newer blocks than to older blocks, the use of blocks as old as $2M+1$ blocks earlier is still unavoidable. The method we propose in Section III will overcome this fundamental limit present in previously reported methods [see (13), shown at the bottom of the next page].

III. Proposed Method

For a subspace method, it is always necessary to write an equation

$$\mathbf{Y} = \mathbf{H}\mathbf{U} + \mathbf{N} \tag{14}$$

or

$$\mathbf{R}_{\mathbf{y}} = \mathbf{H}\mathbf{R}_{\mathbf{u}}\mathbf{H}^\dagger + \mathbf{R}_{\mathbf{n}} \tag{15}$$

where $\mathbf{H}$ contains unknown information on the channel, $\mathbf{U}$ or $\mathbf{R}_{\mathbf{u}}$ contain unknown information of transmitted symbols, and $\mathbf{Y}$ or $\mathbf{R}_{\mathbf{y}}$ contain the noise-corrupted observation of received data. Note that (15) can always be obtained from (14) by setting $\mathbf{R}_{\mathbf{y}} = \mathbf{Y}\mathbf{Y}^\dagger$, $\mathbf{R}_{\mathbf{u}} = \mathbf{U}\mathbf{U}^\dagger$, and $\mathbf{R}_{\mathbf{n}} = \mathbf{N}\mathbf{N}^\dagger$, as long as the input symbols and the noise are uncorrelated. The following discussions will be focused on (14) only. In order to make the subspace method work, (14) must satisfy the following two conditions:

1) $\mathbf{H}$ must be a tall matrix. That is, if $\mathbf{H}$ has a size $p \times m$, then $p > m$;
2) $\mathbf{U}$ must have full row rank, i.e., $\mathrm{rank}(\mathbf{U}) = m$.

The idea of accumulating two consecutive blocks and keeping the ISI-containing CP between the two blocks, as reviewed previously, was actually intended to satisfy condition 1). To satisfy condition 2), the minimum number of blocks must be at least as large as the number of rows of $\mathbf{U}$, since each composite block $\bar{\mathbf{y}}(n)$ defined in (6) can at most increase the rank of $\mathbf{U}$ only by one [as can be seen in (9) and the equation after (9)].

In this section, we reformulate (14) in such a way that each new composite block $\bar{\mathbf{y}}(n)$ can increase the rank of $\mathbf{U}$ by more than one. By repeated use of the same blocks, the "speed" of rank growth of matrix $\mathbf{U}$ will be faster so that a smaller number of received blocks is needed to satisfy condition 2). The idea of repeated use of the same blocks originated in the work of Pham and Manton [10] and was later generalized by Su and Vaidyanathan [14]. These developments were originally for ZP systems. We now show that for CP systems, similar extensions are possible. The generalized method works well in situations in which the previously reported methods [1], [4], [8] either fail or do not perform well, as we shall demonstrate next.

A. The Repetition Index

In this subsection, we will present the idea of repetition index. We will first present the development using an example with small values M and L.

We first rewrite (7) so that the channel matrix has a more symmetric and "tidy" form. The rearranged version of (7) is

$$\bar{\mathbf{y}}(n) = \bar{\mathbf{H}}\bar{\mathbf{u}}(n) \tag{16}$$

where

$$\bar{\mathbf{H}} = \left[\begin{array}{c|c} \mathbf{H}_{cir} & \mathbf{0}_{M\times M} \\ \hline \begin{array}{cc|cc} \multicolumn{2}{c|}{\mathbf{0}_{L\times(M-L)}} & \mathbf{H}_u & \\ & & & \end{array} & \begin{array}{cc} \mathbf{H}_l & \multicolumn{2}{c}{\mathbf{0}_{L\times(M-L)}} \end{array} \\ \hline \mathbf{0}_{M\times M} & \mathbf{H}_{cir2} \end{array}\right]$$

$\mathbf{H}_{\mathrm{cir}2}$ is obtained by permuting columns of $\mathbf{H}_{\mathrm{cir}}$ and is still a circulant matrix. Note that this rewriting is simply to cut the last L columns of $\hat{\mathbf{H}}$ and insert them into the middle. Accordingly, we permute elements of $\mathbf{u}_M(n)$ such that $\mathbf{u}'_M(n) = [\mathbf{u}_M(n)]_{M-L}^{-L+1}$. A special case of (16) when $M = 4$ and $L = 2$ is shown as in (13). This might give a clearer view of the structure of the channel matrix $\bar{\mathbf{H}}$. Observe that $\bar{\mathbf{H}}$ is nearly a Toeplitz matrix except for some sparse terms present in the top and bottom L rows. This Toeplitz-like structure of $\bar{\mathbf{H}}$ will become very useful in the following development. For the sake of clarity, the following developments will start from (13).

We take advantage of the property of circulant matrices. Notice that since

$$\begin{bmatrix} y_{01} \\ y_{02} \\ y_{03} \\ y_{04} \end{bmatrix} = \begin{bmatrix} h_0 & 0 & h_2 & h_1 \\ h_1 & h_0 & 0 & h_2 \\ h_2 & h_1 & h_0 & 0 \\ 0 & h_2 & h_1 & h_0 \end{bmatrix} \begin{bmatrix} u_{01} \\ u_{02} \\ u_{03} \\ u_{04} \end{bmatrix}$$

we have

$$\begin{aligned} \begin{bmatrix} y_{03} \\ y_{04} \\ y_{01} \\ y_{02} \\ y_{03} \\ y_{04} \end{bmatrix} &= \begin{bmatrix} h_2 & h_1 & h_0 & 0 \\ 0 & h_2 & h_1 & h_0 \\ h_0 & 0 & h_2 & h_1 \\ h_1 & h_0 & 0 & h_2 \\ h_2 & h_1 & h_0 & 0 \\ 0 & h_2 & h_1 & h_0 \end{bmatrix} \begin{bmatrix} u_{01} \\ u_{02} \\ u_{03} \\ u_{04} \end{bmatrix} \\ &= \left[\begin{array}{cccc|cc} h_0 & 0 & h_2 & h_1 & 0 & 0 \\ h_1 & h_0 & 0 & h_2 & 0 & 0 \\ h_2 & h_1 & h_0 & 0 & 0 & 0 \\ 0 & h_2 & h_1 & h_0 & 0 & 0 \\ \hline 0 & 0 & h_2 & h_1 & h_0 & 0 \\ 0 & 0 & 0 & h_2 & h_1 & h_0 \end{array}\right] \begin{bmatrix} u_{03} \\ u_{04} \\ u_{01} \\ u_{02} \\ u_{03} \\ u_{04} \end{bmatrix}. \end{aligned} \tag{17}$$

In general, we can show that if $\mathbf{y}_M(n-1) = \mathbf{H}_{\mathrm{cir}}\mathbf{u}_M(n-1)$ is true, then we have

$$[\mathbf{y}_M(n-1)]_M^{1-k} = \begin{bmatrix} \mathbf{H}_{\mathrm{cir}} & \mathbf{0}_{M\times k} \\ \mathbf{0}_{k\times(M-L)} & \mathcal{H}_k \end{bmatrix} [\mathbf{u}_M(n-1)]_M^{1-k} \tag{18}$$

$$\left[\begin{array}{c} y_{01} \\ y_{02} \\ y_{03} \\ y_{04} \\ \hline y_{cp1} \\ y_{cp2} \\ \hline y_{11} \\ y_{12} \\ y_{13} \\ y_{14} \end{array}\right] = \left[\begin{array}{cccc|cccc} h_0 & 0 & h_2 & h_1 & 0 & 0 & 0 & 0 \\ h_1 & h_0 & 0 & h_2 & 0 & 0 & 0 & 0 \\ h_2 & h_1 & h_0 & 0 & 0 & 0 & 0 & 0 \\ 0 & h_2 & h_1 & h_0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & h_2 & h_1 & h_0 & 0 & 0 & 0 \\ 0 & 0 & 0 & h_2 & h_1 & h_0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & h_2 & h_1 & h_0 & 0 \\ 0 & 0 & 0 & 0 & 0 & h_2 & h_1 & h_0 \\ 0 & 0 & 0 & 0 & h_0 & 0 & h_2 & h_1 \\ 0 & 0 & 0 & 0 & h_1 & h_0 & 0 & h_2 \end{array}\right] \left[\begin{array}{c} u_{01} \\ u_{02} \\ u_{03} \\ u_{04} \\ \hline u_{13} \\ u_{14} \\ u_{11} \\ u_{12} \end{array}\right] \tag{13}$$

for any $k \geq 0$. Here

$$\mathcal{H}_k = \begin{bmatrix} h_L & h_{L-1} & \cdots & h_0 & 0 & \cdots & 0 \\ 0 & h_L & h_{L-1} & \cdots & h_0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & & \ddots & 0 \\ 0 & \cdots & 0 & h_L & h_{L-1} & \cdots & h_0 \end{bmatrix} \quad (19)$$

is a $k \times (L+k)$ Toeplitz matrix. Equation (17) was a special case when $k = 2$. Similarly if $\mathbf{y}_M(n) = \mathbf{H}_{\text{cir2}}\mathbf{u}'_M(n)$, then we have

$$[\mathbf{y}_M(n)]_{M+l}^{1} = \begin{bmatrix} \mathcal{H}_l & \mathbf{0}_{l\times(M-L)} \\ \mathbf{0}_{M\times l} & \mathbf{H}_{\text{cir2}} \end{bmatrix} [\mathbf{u}'_M(n)]_{M+l}^{1} \quad (20)$$

for any $l \geq 0$. Combining knowledge of (18) and (20), we can "expand" the composite block $\bar{\mathbf{y}}(n)$ in (16) by k symbols upward and l symbols downward for any nonnegative integers k and l. If we choose k and l such that $k + l = Q - 1$ for some positive integer Q, we will be able to write a new channel equation as follows:

$$\bar{\mathbf{y}}_{kl}(n) = \bar{\mathbf{H}}_Q \bar{\mathbf{u}}_{kl}(n) \quad (21)$$

where

$$\bar{\mathbf{y}}_{kl}(n) = \begin{bmatrix} [\mathbf{y}_M(n-1)]_M^{-k+1} \\ \mathbf{y}_{\text{cp}}(n) \\ [\mathbf{y}_M(n)]_{M+l}^{1} \end{bmatrix}$$

$$\bar{\mathbf{H}}_Q = \begin{bmatrix} \mathbf{H}_{\text{cir}} & & \mathbf{0}_{M\times(M+Q-1)} \\ \mathbf{0}_{(L+Q-1)\times(M-L)} & \mathcal{H}_{L+Q-1} & \mathbf{0}_{(L+Q-1)\times(M-L)} \\ \mathbf{0}_{M\times(M+Q-1)} & & \mathbf{H}_{\text{cir2}} \end{bmatrix}$$

and

$$\bar{\mathbf{u}}_{kl}(n) = \begin{bmatrix} [\mathbf{u}_M(n-1)]_M^{-k+1} \\ [\mathbf{u}'_M(n)]_{M+l}^{1} \end{bmatrix}. \quad (22)$$

Note that if we choose $Q = 1$, then $k = l = 0$ and (21) reduces to (16). Now, by combining cases when k is chosen from 0 to $Q-1$ (and so l from $Q-1$ to 0) in (21), we get

$$\mathbf{Y}_Q(n) = \bar{\mathbf{H}}_Q \mathbf{U}_Q(n) \quad (23)$$

where

$$\mathbf{Y}_Q(n) = [\bar{\mathbf{y}}_{0,Q-1}(n) \quad \bar{\mathbf{y}}_{1,Q-2}(n) \quad \cdots \quad \bar{\mathbf{y}}_{Q-1,0}(n)]$$

is a $(2M+Q+L-1) \times Q$ matrix and

$$\mathbf{U}_Q(n) = [\bar{\mathbf{u}}_{0,Q-1}(n) \quad \bar{\mathbf{u}}_{1,Q-2}(n) \quad \cdots \quad \bar{\mathbf{u}}_{Q-1,0}(n)] \quad (24)$$

is a $(2M+Q-1) \times Q$ matrix. A special case of (23) when $M = 4, L = 2$, and $Q = 3$ is shown in (25) at the bottom of this page. Note that (16) implies (23) without any additional assumptions. We can see this, for example, by verifying that (13) is equivalent to (25). This may provide more insight for (21). The new channel matrix $\bar{\mathbf{H}}_Q$ with a parameter Q maintains a Toeplitz-like structure plus some sparse components: two triangular-shaped "residues" in the top and bottom few rows. As Q increases, the Toeplitz component of $\bar{\mathbf{H}}_Q$ is elongated while the triangular-shaped components keep the same size. We call the parameter Q the *repetition index* since for each composite block $\bar{\mathbf{y}}(n)$ we can generate a matrix $\mathbf{Y}_Q(n)$ whose number of columns is Q.

Finally, if we accumulate J consecutive blocks $(J \geq 2)\mathbf{y}(n)$, $0 \leq n \leq J-1$, we have $J-1$ composite blocks $\bar{\mathbf{y}}(n), 1 \leq n \leq J-1$. Construct the $(2M+Q+L-1)\times Q(J-1)$ matrix

$$\mathbf{Y}_Q^{(J)} = [\mathbf{Y}_Q(1) \quad \mathbf{Y}_Q(2) \quad \cdots \quad \mathbf{Y}_Q(J-1)]. \quad (26)$$

Then we have

$$\mathbf{Y}_Q^{(J)} = \bar{\mathbf{H}}_Q \mathbf{U}_Q^{(J)}$$

$$\left[\begin{array}{ccc} y_{01} & y_{04} & y_{03} \\ y_{02} & y_{01} & y_{04} \\ y_{03} & y_{02} & y_{01} \\ y_{04} & y_{03} & y_{02} \\ \hline y_{cp1} & y_{04} & y_{03} \\ y_{cp2} & y_{cp1} & y_{04} \\ y_{11} & y_{cp2} & y_{cp1} \\ y_{12} & y_{11} & y_{cp2} \\ \hline y_{13} & y_{12} & y_{11} \\ y_{14} & y_{13} & y_{12} \\ y_{11} & y_{14} & y_{13} \\ y_{12} & y_{11} & y_{14} \end{array}\right] = \left[\begin{array}{cccc|cc|cccc} h_0 & 0 & h_2 & h_1 & 0 & 0 & 0 & 0 & 0 & 0 \\ h_1 & h_0 & 0 & h_2 & 0 & 0 & 0 & 0 & 0 & 0 \\ h_2 & h_1 & h_0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & h_2 & h_1 & h_0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & h_2 & h_1 & h_0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & h_2 & h_1 & h_0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & h_2 & h_1 & h_0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & h_2 & h_1 & h_0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & h_2 & h_1 & h_0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & h_2 & h_1 & h_0 \\ 0 & 0 & 0 & 0 & 0 & 0 & h_0 & 0 & h_2 & h_1 \\ 0 & 0 & 0 & 0 & 0 & 0 & h_1 & h_0 & 0 & h_2 \end{array}\right] \left[\begin{array}{ccc} u_{01} & u_{04} & u_{03} \\ u_{02} & u_{01} & u_{04} \\ u_{03} & u_{02} & u_{01} \\ u_{04} & u_{03} & u_{02} \\ \hline u_{13} & u_{04} & u_{03} \\ u_{14} & u_{13} & u_{04} \\ \hline u_{11} & u_{14} & u_{13} \\ u_{12} & u_{11} & u_{14} \\ u_{13} & u_{12} & u_{11} \\ u_{14} & u_{13} & u_{12} \end{array}\right] \quad (25)$$

where

$$\mathbf{U}_Q^{(J)} = [\mathbf{U}_Q(1) \quad \mathbf{U}_Q(2) \quad \cdots \quad \mathbf{U}_Q(J-1)] \tag{27}$$

is a $(2M+Q-1) \times Q(J-1)$ matrix.

Theorem 2: $\bar{\mathbf{H}}_Q$ has full column rank $2M+Q-1$ if and only if $H(z)$ as defined in (3) does not have any zero at $z = W_M^l$, $0 \le l \le M-1$.

Proof: If $H(W_M^l) = 0$ for some l, then

$$\begin{bmatrix} 1 & W_M^l & W_M^{2l} & \cdots & W_M^{l(2M+Q-2)} \end{bmatrix}^T$$

is a right annihilator of $\bar{\mathbf{H}}_Q$ and hence $\bar{\mathbf{H}}_Q$ does not have full rank. On the other hand, when $H(W_M^l) \neq 0$ for any l, suppose $\bar{\mathbf{H}}_Q$ does not have full rank. Then there exists a nonzero vector $\mathbf{v} = [\mathbf{v}_{M1}^T \quad \mathbf{v}_{Q-1}^T \quad \mathbf{v}_{M2}^T]^T$ such that $\bar{\mathbf{H}}_Q \mathbf{v} = \mathbf{0}$. The lengths of $\mathbf{v}_{M1}, \mathbf{v}_{Q-1}$, and $\mathbf{v}_{M2}$ are $M, Q-1$, and M, respectively. Note that when $Q = 1$, the segment $\mathbf{v}_{Q-1}$ has a zero length (i.e., this segment simply does not exist). Observe that $\mathbf{H}_{\text{cir}} \mathbf{v}_{M1} = \mathbf{0}$. Since

$$\det(\mathbf{H}_{\text{cir}}) = \prod_{l=0}^{M-1} H(W_M^l) \neq 0$$

we have $\mathbf{v}_{M1} = \mathbf{0}$. Similarly, $\mathbf{v}_{M2} = \mathbf{0}$ since $\det(\mathbf{H}_{\text{cir2}}) \neq 0$. If $Q = 1$, this already leads to a contradiction. In the case when $Q > 1, \bar{\mathbf{H}}_Q \mathbf{v} = \mathbf{0}$ implies $\mathcal{T}_{Q-1}(\mathbf{h}) \mathbf{v}_{Q-1} = \mathbf{0}$ [see (1) for definition of notation $\mathcal{T}_{Q-1}(\mathbf{h})$]. But $\mathcal{T}_{Q-1}(\mathbf{h})$ has full rank. So $\mathbf{v}_{Q-1}$ must also be zero. This contradicts the fact that $\mathbf{v}$ is nonzero, and so $\bar{\mathbf{H}}_Q$ must have full column rank. ■

Note that when $Q = 1$, Theorem 2 reduces to Theorem 1. Theorem 2 states that the necessary and sufficient conditions for $\bar{\mathbf{H}}_Q$ to have full column rank does not change whatever the repetition index Q we use. Assume the channel $H(z)$ does not have zeros at $z = W_M^l$ for any l. Then $\bar{\mathbf{H}}_Q$ has full column rank $2M+Q-1$. This assumption is usually reasonable since the probability that a channel $H(z)$ has a zero exactly at $z = W_M^l$ is zero. We also assume that there exists J such that $\mathbf{U}_Q^{(J)}$ achieves full row rank $2M+Q-1$. Under these two assumptions, we obtain that the $(2M+L+Q-1)$-row matrix $\mathbf{Y}_Q^{(J)}$ has rank $2M+Q-1$. This means there exist L linearly independent vectors $\mathbf{g}_k, 1 \le k \le L$ such that

$$\mathbf{g}_k^\dagger \mathbf{Y}_Q^{(J)} = \mathbf{0}^T. \tag{28}$$

Since $\mathbf{U}_Q^{(J)}$ has full row rank, these vectors $\mathbf{g}_k^\dagger$ are also annihilators of $\bar{\mathbf{H}}_Q$.

For each annihilator $\mathbf{g}_k^\dagger$ of $\bar{\mathbf{H}}_Q$, we can construct a $(2M+Q-1) \times (L+1)$ matrix $\mathcal{G}_k$ in a way similar to (10) in Section II such that

$$\mathcal{G}_k \mathbf{h} = \mathbf{0}. \tag{29}$$

The construction of $\mathcal{G}_k$ is conceptually easy. We simply inspect each column of $\bar{\mathbf{H}}_Q$ and find locations of each channel coefficient $h_i, 0 \le i \le L$. For example, in the special case where $M = 4, L = 2$, and $Q = 3$, the structure of $\mathcal{G}_k$ is given as

$$\mathcal{G}_k = \begin{bmatrix} g_{k1} & g_{k2} & g_{k3} \\ g_{k2} & g_{k3} & g_{k4} \\ g_{k3} & g_{k4} & g_{k5}+g_{k1} \\ g_{k4} & g_{k5}+g_{k1} & g_{k6}+g_{k2} \\ g_{k5} & g_{k6} & g_{k7} \\ g_{k6} & g_{k7} & g_{k8} \\ g_{k7}+g_{k,11} & g_{k8}+g_{k,12} & g_{k9} \\ g_{k8}+g_{k,12} & g_{k9} & g_{k,10} \\ g_{k9} & g_{k,10} & g_{k,11} \\ g_{k,10} & g_{k,11} & g_{k,12} \end{bmatrix}$$

where g_{kl} denotes the lth element of $\mathbf{g}_k^\dagger$. A systematic way of construction of $\mathcal{G}_k$ is given as follows.

First note that $\bar{\mathbf{H}}_Q = \mathcal{H}_{2M+Q+L-1} \mathbf{A}$, where the notation $\mathcal{H}_k$ was defined in (19) and $\mathbf{A}$ is a sparse matrix defined as follows.

$$\mathbf{A} = [\mathbf{A}_1^T \quad \mathbf{I}_{2M+Q-1} \quad \mathbf{A}_2^T]^T$$

where

$$\mathbf{A}_1 = [\mathbf{0}_{L \times (M-L)} \quad \mathbf{I}_L \quad \mathbf{0}_{L \times (M+Q-1)}]$$

and

$$\mathbf{A}_2 = [\mathbf{0}_{L \times (M+Q-1)} \quad \mathbf{I}_L \quad \mathbf{0}_{L \times (M-L)}].$$

Now, we have

$$\begin{aligned} \mathbf{0}^T = \mathbf{g}_k^\dagger \bar{\mathbf{H}}_Q &= \mathbf{g}_k^\dagger \mathcal{H}_{2M+Q+L-1} \mathbf{A} \\ &= [h_L \quad \cdots \quad h_0] \mathcal{T}_{L+1}^T (\mathbf{g}_k^\dagger) \mathbf{A} \\ &= \mathbf{h}^T \mathbf{G}_k \mathbf{A} \end{aligned}$$

where $\mathbf{G}_k = \mathcal{K}_{L+1}([\mathbf{0}_{1 \times L}, (\mathbf{g}_k^\dagger)^T, \mathbf{0}_{1 \times L}])$ is a Hankel matrix [see (2) for definition of the notation] composed of elements of $\mathbf{g}_k^\dagger$. Now, by simply choosing

$$\mathcal{G}_k = \mathbf{A}^T \mathbf{G}_k^T$$

Equation (29) is satisfied. By defining

$$\mathcal{G} = [\mathcal{G}_1^T \quad \mathcal{G}_2^T \quad \cdots \quad \mathcal{G}_L^T]^T \tag{30}$$

we now have $\mathcal{G}\mathbf{h} = \mathbf{0}$. The channel coefficients $\mathbf{h}$ can be identified within a scalar ambiguity.

In presence of noise, the estimated annihilators $\mathbf{g}_k^\dagger$ can be found by taking SVD on $\mathbf{Y}_Q^{(J)}$ and choosing the L singular vectors associated with the L smallest singular values [similar to the description after (11)]. Also, after constructing the $\mathcal{G}$ matrix, we use the vector $\mathbf{h}$ which minimizes the norm of $\mathcal{G}\mathbf{h}$ as the estimated channel coefficients. This optimal estimation can be written as

$$\hat{\mathbf{h}} = \arg \min_{\|\mathbf{h}\|=1} \|\mathcal{G}\mathbf{h}\|^2 = \arg \min_{\|\mathbf{h}\|=1} \mathbf{h}^\dagger (\mathcal{G}^\dagger \mathcal{G}) \mathbf{h}. \tag{31}$$

B. Necessary Condition for Persistency of Excitation

Recall that the matrix $\mathbf{U}_Q^{(J)}$ defined in (27) must have full row rank. If $\mathbf{U}_Q^{(J)}$ does not have full rank, some annihilators of $\mathbf{Y}_Q^{(J)}$ as defined in (28) may not be annihilators of $\bar{\mathbf{H}}_Q$ and will result in failure of the proposed algorithm. Since $\mathbf{U}_Q^{(J)}$ has size $(2M+Q-1)\times(J-1)Q$, it has full row rank only when

$$(J-1)Q \geq 2M+Q-1 \tag{32}$$

or

$$Q \geq \frac{2M-1}{J-2}. \tag{33}$$

This necessary condition for $\mathbf{U}_Q^{(J)}$ to have full row rank $(2M+Q-1)$ is not sufficient since it still depends on the values of transmitted symbols $\mathbf{u}_M(n)$. However, simulations in Section IV show that (for most choices of M and input constellations) once (33) is satisfied, the probability that $\mathbf{U}_Q^{(J)}$ has full rank is very close to unity. Thus

$$Q = \left\lceil \frac{2M-1}{J-2} \right\rceil \tag{34}$$

is usually a valid choice in practice. A detailed study on the conditions of $\mathbf{U}_Q^{(J)}$ having full rank is presented in Section IV. Now, if we choose

$$J \geq 3$$

then there exists Q such that $\mathbf{U}_Q^{(J)}$ can possibly have full rank. This suggests that the proposed algorithm is *potentially capable of identifying the channel from only three blocks*. In Section V we will demonstrate these with examples.

C. Repetition Index for the Forgetting Factor

The idea of using a repetition index Q can also be applied when a forgetting factor is used. The technique of using a forgetting factor has been reviewed in Section II right before (12). The "autocorrelation matrix" $\mathbf{R}_{\bar{\mathbf{y}}\bar{\mathbf{y}},Q}^{(0)}$ is initiated as $\mathbf{R}_{\bar{\mathbf{y}}\bar{\mathbf{y}},Q}^{(0)}=\mathbf{0}$ and updated each time when a new composite block $\bar{\mathbf{y}}(N-1)$ is received as

$$\mathbf{R}_{\bar{\mathbf{y}}\bar{\mathbf{y}},Q}^{(N)} = \alpha\mathbf{R}_{\bar{\mathbf{y}}\bar{\mathbf{y}},Q}^{(N-1)} + (1-\alpha)\mathbf{Y}_Q(N-1)[\mathbf{Y}_Q(N-1)]^\dagger \tag{35}$$

where $\alpha\in[0,1]$ is the forgetting factor. The SVD of $\mathbf{R}_{\bar{\mathbf{y}}\bar{\mathbf{y}},Q}^{(N)}$ is then taken, and the estimated annihilators $\mathbf{g}_k^\dagger$ chosen as the singular vectors associated with the smallest L singular values of $\mathbf{R}_{\bar{\mathbf{y}}\bar{\mathbf{y}},Q}^{(N)}$. Note that N must satisfy $N\geq(2M+Q-1)/Q$ to render $\mathbf{R}_{\bar{\mathbf{y}}\bar{\mathbf{y}},Q}^{(N)}$ full rank. This means the first channel estimation after initialization can be requested only when $N\geq(2M+Q-1)/Q$. After this, an estimation can be requested at any time instant N.

D. Summary of the Proposed Algorithm

The proposed algorithm can be summarized as follows.

1) Given M and the CP length L, choose J and the repetition index Q such that
$$Q \geq \frac{2M-1}{J-2}.$$
Some remarks on choosing a good pair of J and Q will be presented in Section V.
2) Collect J blocks $\mathbf{y}(n)$ at the receiver and construct a $(2M+L+Q-1)\times(J-1)Q$ matrix $\mathbf{Y}_Q^{(J)}$ as defined in (26). Let $\mathbf{Z}=\mathbf{Y}_Q^{(J)}\mathbf{Y}_Q^{(J)^\dagger}$.
3) Perform SVD on $\mathbf{Z}$ so that
$$\mathbf{Z}=[\mathbf{U}_s \quad \mathbf{U}_n]\begin{bmatrix}\mathbf{\Sigma}_s & \mathbf{0}\\ \mathbf{0} & \mathbf{\Sigma}_n\end{bmatrix}\begin{bmatrix}\mathbf{U}_s^\dagger\\ \mathbf{U}_n^\dagger\end{bmatrix}$$
where the diagonal entries of $\mathbf{\Sigma}_n$ are the L smallest singular values of $\mathbf{Z}$.
4) Let $\mathbf{g}_k$ be chosen as the kth column of $\mathbf{U}_n$. Construct the $(2M+Q-1)L\times(L+1)$ matrix $\mathcal{G}$ as in (30).
5) Let $\hat{\mathbf{h}}$ be the eigenvector of $\mathcal{G}^\dagger\mathcal{G}$ associated with the smallest eigenvalue. This is the estimated channel vector within a scalar ambiguity.

When a forgetting factor is used, steps 1 and 2 are modified as follows.

1) Choose $\alpha\in[0,1]$ and the repetition index Q. Some remarks of choosing a good α will be presented in Section V.
2) Update the "autocorrelation matrix" $\mathbf{R}_{\bar{\mathbf{y}}\bar{\mathbf{y}},Q}^{(N)}$ as received blocks are accumulated. Choose $\mathbf{Z}=\mathbf{R}_{\bar{\mathbf{y}}\bar{\mathbf{y}},Q}^{(N)}$ as defined in (35) where N is the block index when a channel estimation is requested.

E. System Complexity

The computational complexity of the proposed algorithm is dominated by the SVD of the matrix $\mathbf{Z}$, whose size is $2M+Q+L-1$. The computational complexity is proportional to $\mathcal{O}((2M+Q+L-1)^3)$. A larger repetition index Q leads to a greater complexity. However, when M and L are much larger than Q, this complexity increase due to increase of Q is not very serious. On the other hand, if Q is chosen as large as $2M-1$ (e.g., when $J=3$), the complexity increase can be significant.

F. Equalization and Resolving the Scalar Ambiguity

After estimating the channel coefficients, the receiver proceeds to equalize the effects of the frequency-selective channels. A standard linear minimum mean square error (L-MMSE) equalizer is used at the receiver. Fig. 2 depicts the equalizer structure of the system. Here $\mathbf{\Lambda}$ is a diagonal matrix whose kth diagonal entry is

$$\mathbf{\Lambda}_{k,k} = \frac{E_s\hat{H}^*\left(W_M^k\right)}{E_s|\hat{H}\left(W_M^k\right)|^2+N_0} \tag{36}$$

where E_s is the average energy of transmitted symbols, N_0 is the channel noise variance, and $\hat{H}(W_M^k)=\sum_{l=0}^{L}[\hat{\mathbf{h}}]_l W_M^{-kl}$ is the frequency response of the estimated channel. Since there is a scalar ambiguity in the estimated channel coefficients, all equalized symbols will be scaled by an unknown complex-valued scalar c. A usual way to resolve this scalar is to introduce *one*

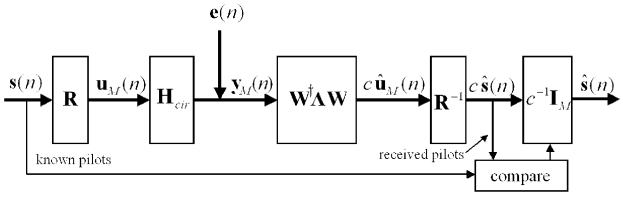

Fig. 2. The transceiver system equipped with a method to resolve scale-factor ambiguity.

extra pilot symbol and compare it with the corresponding received symbol. If several blocks are using the same channel estimate $\hat{\mathbf{h}}$, the scalar ambiguity can be estimated as follows:

$$\hat{c} = \arg\min_{c\in\mathbb{C}} \sum_n \|s_{\rm rec}(n) - cs_{\rm pil}(n)\|^2 \tag{37}$$

$$= \frac{\sum_n s^*_{\rm pil}(n)s_{\rm rec}(n)}{\sum_n |s_{\rm pil}(n)|^2} \tag{38}$$

where $s_{\rm pil}(n)$ is the pilot symbol of the nth block and $s_{\rm rec}(n)$ is the corresponding received pilot. We set the first symbol of each source block $\mathbf{s}(n)$ as the known symbol (i.e., $s_{\rm pil}(n) = [\mathbf{s}(n)]_1$) defined as

$$s_{\rm pil}(n) = \sqrt{E_s}p_{(n \bmod 4)}$$

where $[p_0 \quad p_1 \quad p_2 \quad p_3] = [1 \quad j \quad -j \quad -1]$. There are definitely many other alternative designs of these pilot symbols. The choice here is just to make sure that $\mathbf{U}_Q^{(J)}$ defined in (27) would not become rank deficient due to the introduction of these pilot symbols.

IV. On the Probability that $\mathbf{U}_Q^{(J)}$ Has Full Rank

Before presenting simulation results which demonstrate the performance of the above algorithm, we discuss the technical issue of rank requirement of the matrix $\mathbf{U}_Q^{(J)}$ defined in (27) in greater detail.

Recall that one assumption for the proposed algorithm is that the $(2M+Q-1)\times Q(J-1)$ matrix $\mathbf{U}_Q^{(J)}$ must have full row rank. Inequality (33) is a necessary condition but is not sufficient since whether $\mathbf{U}_Q^{(J)}$ has full rank or not ultimately depends on the content of $\mathbf{U}_Q^{(J)}$. As long as the contents of $\mathbf{U}_Q^{(J)}$ are chosen from a finite constellation, then there is always a nonzero probability that $\mathbf{U}_Q^{(J)}$ is rank-deficient. To see this, simply consider the extreme case where the contents of $\mathbf{U}_Q^{(J)}$ are always chosen as identical symbols. All subspace-based blind methods suffer from the possibility of rank deficiency of the data matrix. Here, we will study how this probability of rank deficiency changes when J and Q change. To facilitate our discussion, we formally define the probability of $\mathbf{U}_Q^{(J)}$ having full rank as follows.

Definition 1: Consider a constellation $\mathcal{S}$ (which has at least two elements) and an $M\times M$ nonsingular precoder $\mathbf{R}$. Let each element of the $M\times J$ matrix $\mathbf{S} = [\mathbf{s}(0) \quad \mathbf{s}(1) \quad \cdots \quad \mathbf{s}(J-1)]$ be independently selected from the constellation $\mathcal{S}$ with equal probabilities. Let $\mathbf{u}_M(n) = \mathbf{R}\mathbf{s}(n)$ and let $\mathbf{U}_Q^{(J)}$ be defined as in (27). For $J \geq 2, Q \geq 1$, the probability that $\mathbf{U}_Q^{(J)}$ has full rank will be denoted as $P_{\mathcal{S},\mathbf{R}}(J,Q)$. ■

Obviously, $P_{\mathcal{S},\mathbf{R}}(J,Q) = 0$ whenever $(J-2)Q < 2M-1$. Also, we have $P_{\mathcal{S},\mathbf{R}}(J+1,Q) \geq P_{\mathcal{S},\mathbf{R}}(J,Q)$ and $P_{\mathcal{S},\mathbf{R}}(J,Q+1) \geq P_{\mathcal{S},\mathbf{R}}(J,Q)$. The former inequality comes from the fact that the row rank of a matrix never decreases when additional columns are appended, and the latter can be verified by the following theorem. These inequalities show that both increasing J and increasing Q have the potential to increase the probability that $\mathbf{U}_Q^{(J)}$ has full rank.

Theorem 3: If $\mathbf{U}_Q^{(J)}$ has full row rank $(2M+Q-1)$, then $\mathbf{U}_{Q+1}^{(J)}$ also has full row rank $(2M+Q)$.

Proof: See Appendix A. ■

When J approaches infinity, it can be shown that $\lim_{J\to\infty} P_{\mathcal{S},\mathbf{R}}(J,Q) = 1$ for any constellation $\mathcal{S}$ and precoder $\mathbf{R}$ (and any $Q \geq 1$). However, this is not the case when we increase Q. The probability of full rank of $\mathbf{U}_Q^{(J)}$ always stops increasing when $Q \geq 2M-1$, which can be verified by the following theorem.

Theorem 4: If $\mathbf{U}_Q^{(J)}$ does not have full rank when $Q = 2M-1$, then $\mathbf{U}_Q^{(J)}$ does not have full rank for any Q.

Proof: See Appendix A. ■

Combining Theorems 3 and 4, we immediately have

$$P_{\mathcal{S},\mathbf{R}}(J,Q) = P_{\mathcal{S},\mathbf{R}}(J,2M-1)$$

for any $Q \geq 2M-1$.

We perform simulations with three commonly used constellations in communications: BPSK, QPSK, and 16-QAM. The $M\times M$ precoder $\mathbf{R}$ is chosen as $\mathbf{I}_M$ for SC-CP systems and $\mathbf{W}^\dagger$ for OFDM systems. Although the exact probability of $\mathbf{U}_Q^{(J)}$ having full rank can be actually obtained by testing all possible transmitted data, an exhaustive simulation is barely feasible. For each $J \geq 3$, the simulations are performed under two values of $Q = 2M-1$ and $Q = \lceil (2M-1)/(J-2) \rceil$. When $Q = 2M-1$, the simulation gives an upper bound of $P_{\mathcal{S},\mathbf{R}}(J,Q)$ for a given J and the simulation where $Q = \lceil (2M-1)/(J-2) \rceil$ gives a lower bound of nonzero $P_{\mathcal{S},\mathbf{R}}(J,Q)$. M is chosen as 16.

Figs. 3 and 4 show the results when the precoder is chosen as an identity matrix and an IDFT matrix, respectively. Some comments on these results are made below.

1) As expected, the probability of $\mathbf{U}_Q^{(J)}$ having full rank is smaller when a smaller constellation is used or when J is smaller. When $J \geq 12$, the probability becomes very close to unity for all combinations of constellations and precoders. When a 16-QAM constellation is used, the probability is already very high when $J = 5$.
2) It should be especially noted that the probability of $\mathbf{U}_Q^{(J)}$ having full rank is significantly smaller when $\mathbf{R}$ is chosen as the IDFT matrix than when $\mathbf{R}$ is an identity matrix. An explanation of this phenomenon can be found in Appendix B. This phenomenon suggests the proposed algorithm is more stable when operated in SC-CP systems than in OFDM systems when the constellation is small and/or when J is small.
3) Finally, although the theory suggests $P_{\mathcal{S},\mathbf{R}}(J,2M-1) \geq P_{\mathcal{S},\mathbf{R}}(J,\lceil (2M-1)/(J-2) \rceil)$, in simulation the above two quantities look almost the same so that a conjecture may be made that $P_{\mathcal{S},\mathbf{R}}(J,Q) = P_{\mathcal{S},\mathbf{R}}(J,\lceil (2M-1)/(J-2) \rceil)$ for

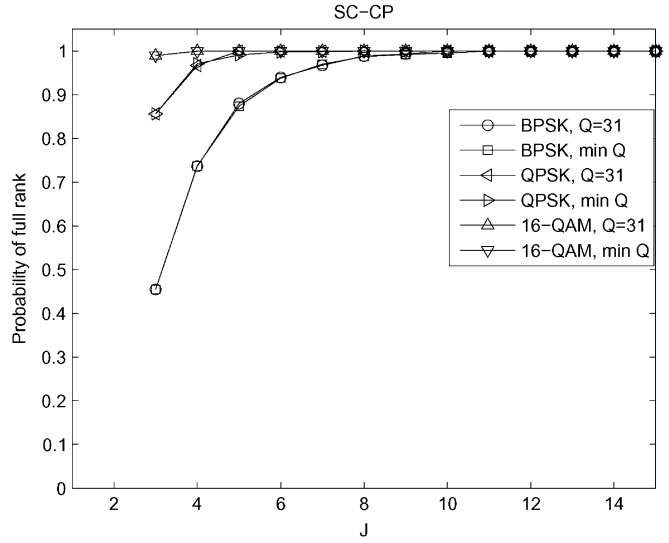

Fig. 3. The probability of $\mathbf{U}_Q^{(J)}$ having full rank in SC-CP systems.

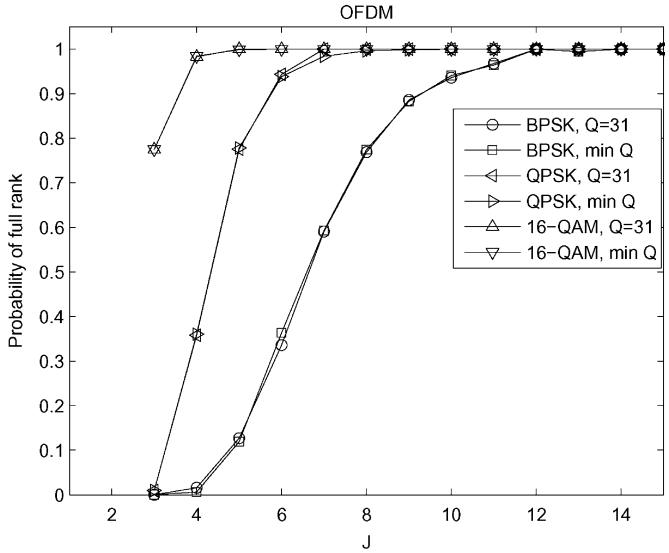

Fig. 4. The probability of $\mathbf{U}_Q^{(J)}$ having full rank in OFDM systems.

any $Q \geq \lceil (2M-1)/(J-2) \rceil$. This conjecture, however, has not yet been verified or disproved at the time of writing of this paper.

V. Simulation Results and Discussion

In this section, we conduct several Monte Carlo simulations to demonstrate the performance of the proposed method under different system parameters: the number of collected blocks J, the repetition index Q, and the forgetting factor α. The block size M is chosen as 64 and the length of cyclic prefix is $L = 16$. The sample period is 1 μs and so the block length is 80 μs. We assume perfect block synchronization in all simulations. Note that in practice a blind block synchronization must be done before blind channel identification can be performed. Recall that all previously reported algorithms in the literature use $Q = 1$.

TABLE I
Channel Model in Section IV

Tap	Delays (μs)	Avg. Power (dB)	Tap	Delays (μs)	Avg. Power (dB)
1	0	0.0	9	8	-6.9
2	1	-0.9	10	9	-7.8
3	2	-1.7	11	10	-4.7
4	3	-2.6	12	11	-7.3
5	4	-3.5	13	12	-9.9
6	5	-4.3	14	13	-12.5
7	6	-5.2	15	14	-13.7
8	7	-6.1	16	15	-18.0

A. Static Channels

We first test our methods in static channel environments. The channel is an FIR filter whose order is upper bounded by the CP length $L = 16$. The constellation of source symbols is QPSK and the precoder $\mathbf{R}$ is chosen as the identity matrix (i.e., an SC-CP system). The simulation is performed over 500 different channels generated by Rayleigh fading statistics according to Table I. The normalized least squared channel estimation error, denoted as E_{ch}, is used as the figure of merit for channel identification and is defined as follows:

$$E_{\text{ch}} = \frac{1}{N_{\text{ch}}} \left[\sum_{k=1}^{N_{\text{ch}}} \min_{c \in \mathbb{C}} \frac{\|c\hat{\mathbf{h}}_k - \mathbf{h}_k\|^2}{\|\mathbf{h}_k\|^2} \right]$$

where N_{ch} is the number of channel estimates performed, $\mathbf{h}_k$ is the true channel vector, and $\hat{\mathbf{h}}_k$ is the channel estimate with a scalar ambiguity as defined in (31).

The simulation results for normalized channel estimation error E_{ch} is shown in Fig. 5 and the corresponding bit-error-rate (BER) plot is presented in Fig. 6. When $J = 86$ and $Q = 1$, the algorithm simply does not work since inequality (33) is not satisfied. This means the previously reported methods are unable to perform blind channel identification using only 86 blocks. When we choose $Q = 2$, the algorithm works with a fairly satisfactory result. When $Q = 3$, the system performance further improves.

When the number of received blocks is $J = 129$, the algorithm works, but not very well, with $Q = 1$. In view of (33), this is the minimum number of blocks J needed for any previously reported algorithm ($Q = 1$). If we use $Q = 2$, the performance has a significant boost. This suggests that choosing Q larger than necessary sometimes yields a better performance. When $J = 257$, the performance is even better since more data are available for blind identification. Using $Q = 2$ stills slightly improves the system performance but the improvement is not as large as in the previous cases. It is worthy to note that the performance curves of three cases where "$J = 86; Q = 3$," "$J = 129; Q = 2$," and "$J = 257; Q = 1$" are very close to each other. Recognizing that $(J-1)Q$ are very close to each other in these three cases, this phenomenon suggests that the system performance could be directly proportional to the number of column of $\mathbf{Y}_Q^{(J)}((J-1)Q)$ as defined in (26) regardless of the actual number of accumulated received blocks (J).

We repeated the same simulation settings for other constellations and precoders $\mathbf{R}$. Fig. 7 depicts the BER performance where a 16-QAM constellation and a precoder $\mathbf{R} = \mathbf{I}_M$ are

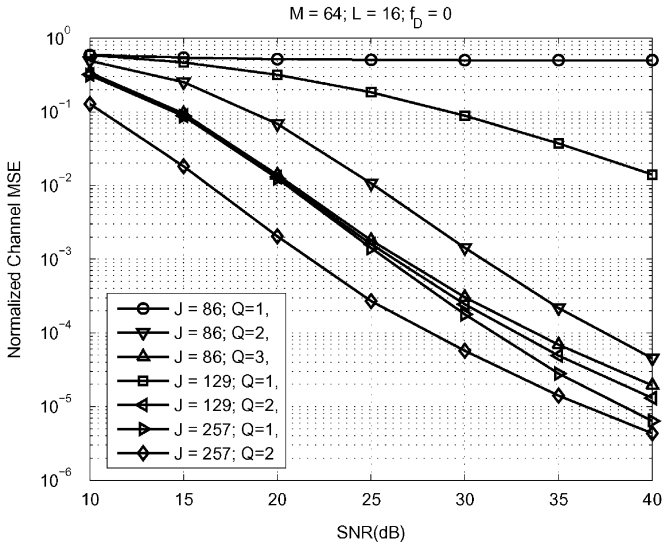

Fig. 5. Normalized MSE of channel estimation for static channels with the QPSK constellation in SC-CP systems.

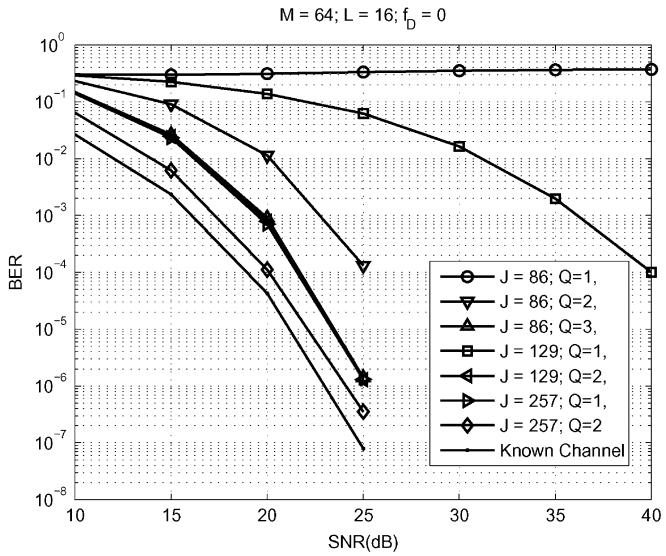

Fig. 6. BER performance for static channels with the QPSK constellation in SC-CP systems.

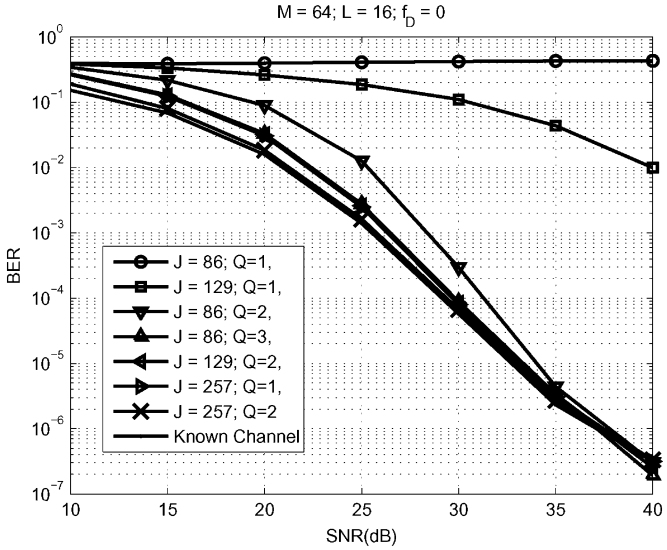

Fig. 7. BER performance for static channels with the 16-QAM constellation in SC-CP systems.

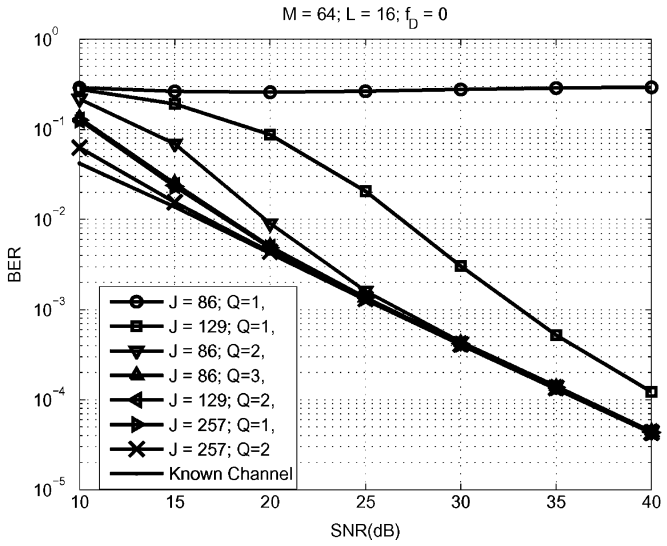

Fig. 8. BER performance for static channels with the QPSK constellation in OFDM systems.

used. The BER performance of the case where a QPSK constellation and a precoder $\mathbf{R} = \mathbf{W}^{\dagger}$ (i.e., an OFDM system) are used are shown in Fig. 8. All these results exhibit similar characteristics to the case described in the previous paragraph.

B. Simulations With Smaller J

We also test our algorithm when the number of available received blocks are smaller, with nine different values of J ranging from 3 to 64. Note that $J = 3$ is the smallest integer that satisfies (32). The repetition index Q is chosen as

$$Q = \left\lceil \frac{2M-1}{J-2} \right\rceil + 3$$

for each J. Here we choose repetition indices larger by three than needed, in order to achieve a better system performance. Other system parameters are the same as in the first simulation in Section V-A. The BER performance is shown in Fig. 9. When $J = 3$, the BER decreases slowly as SNR increases. This demonstrates the theoretical limit on the number of received blocks required for the proposed system as argued in Section III-B. However, when J is smaller than 10, the BER performances as shown in Fig. 9 are usually unrealistic in practice. Also, a small J requires a large Q, which imposes a very demanding computational complexity. These observations largely limit the applicability of the proposed algorithm with these extremely small J in practical situations.

When the number of available received blocks is larger, the BER performance is much better. When $J = 10$ and $Q = 19$, a BER of around 10^{-5} is achieved when SNR is 30 dB. When $J = 20$ and $Q = 11$, the BER is on the order of 10^{-5} when SNR is 25 dB. The SNR margin between the BER curves of this case $(J = 20)$ and of the case of known channel is around 5 dB at $\mathrm{BER} = 10^{-4}$. When $J = 30$ and $J = 40$, this margin reduces to around 4 and 3 dB, respectively. These results are

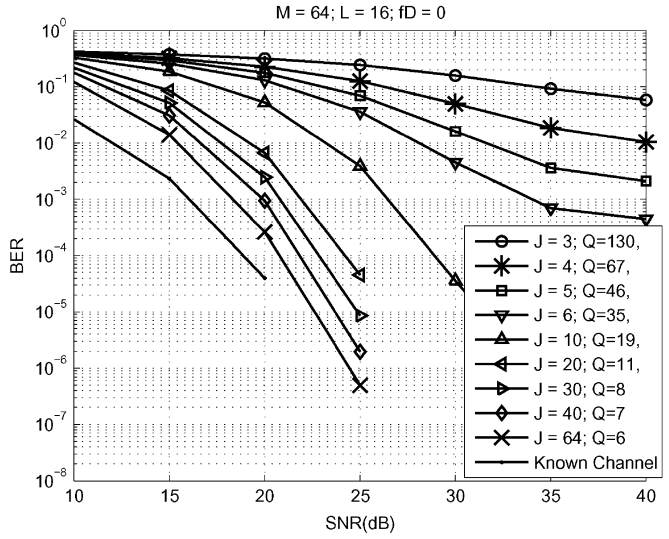

Fig. 9. BER performance for static channels with the QPSK constellation in SC-CP systems when J is small.

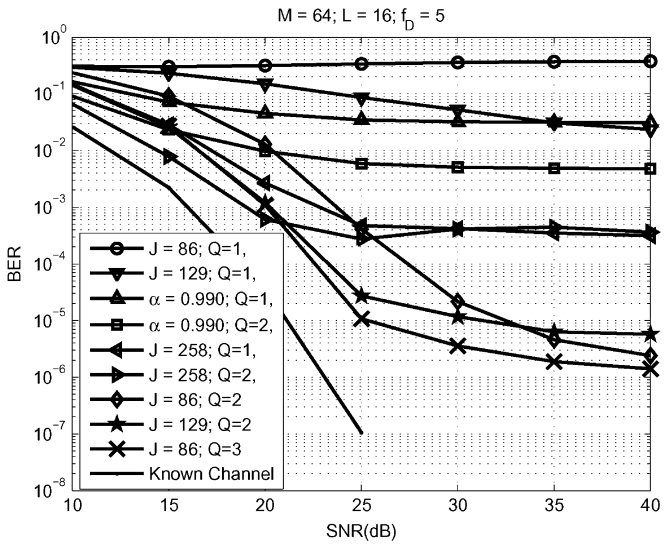

Fig. 10. BER performance for blind identification systems when the Doppler frequency is 5 Hz (5.4 km/h).

considered acceptable BER in some practical applications (note that the presented results are all uncoded BER). Since $J = 30$ is slightly less than half the block size $M = 64$, we can argue that *the minimum number of received blocks required in a practical situation is on the order of half block size*. Three more similar simulations results with $M = 32, M = 128, M = 256$ strengthen this argument. Due to high similarity and space limit, they are not shown here. Compared to previously reported subspace-based blind algorithms [1], [4], [8], which always require a number of received blocks larger than *twice the block size*, the introduction of repetition index indeed largely reduces the required number of received blocks.

C. Time-Varying Channels

We now test our algorithm in an environment of time-varying channels. For time-varying channels there is always a dilemma for subspace-based blind channel identification algorithms in choosing the number of accumulated blocks (J). When J is large, the channel state may have changed significantly during data accumulation so that the estimation results could be meaningless. When J is small, the performance would be poor due to very limited amount of available data. With the introduction of repetition index Q, this problem can be solved to a certain extent.

In our simulation, the channel model considered is a random FIR channel with an order upper bounded by the CP length whose characteristics is shown in Table I. A standard Jakes' Doppler spectrum is used and Rayleigh fading statistics are assumed for all taps [3]. A channel estimate is obtained using data of J consecutive blocks and then used to equalize the middle N_B blocks of the J blocks, where N_B is usually chosen as an integer small than or equal to J. One reason of doing this is, in the context of time-varying channels, the channel estimate obtained from J blocks may not be very accurate for the first few and the last few of the J blocks. In order to equalize each received block, a channel estimate is obtained every N_B blocks.

For the first simulation, the Doppler frequency is chosen as 5 Hz, which corresponds to an object speed 1.5 m/s (5.4 km/h) if the carrier frequency is 1 GHz. The symbol duration is 10^{-6} seconds. This setting implies that the channel coefficients become totally uncorrelated in around 0.08 s (i.e., coherence interval), equal to 80 000 symbol durations, or 1000 received blocks. A channel estimate is performed once for a time duration of 50 blocks (i.e., $N_B = 50$). The plot of BER performance is shown in Fig. 10. In the low SNR region, the case where $J = 258$ and $Q = 2$ has the best performance. However, in the high SNR region, the case where $J = 86$ and $Q = 3$ becomes the best. Note that in the high SNR region, except for a few cases [where (32) is not satisfied or is satisfied with a very small margin], the BER is greater when J is larger. This is because when channel noise is small, the channel estimation error comes solely from channel variation due to accumulation of a large number of blocks. In the low-SNR region, curves with similar values $(J-1)Q$ tend to have similar performances, just like what has been observed in static channel environments. We also compare an adaptive scheme where a forgetting factor $\lambda = 0.99$ is used. When $Q = 1$, the performance is not very good. Now if we choose $Q = 2$, a considerable improvement over $Q = 1$ is observed. Although the performance of forgetting factor schemes is not very good when SNR is high, they could be more promising than methods using a fixed J in the low-SNR region.

Due to channel variation, the channel estimation error does not converge to zero even when the SNR is very high. As a consequence, the linear MMSE receiver defined in (36) becomes inaccurate when the SNR is large since the channel estimation error constitutes a larger variance than channel noise. In the simulation for the BER plot, we slightly adjust the linear MMSE equalizer defined in (36) as

$$\mathbf{\Lambda}_{k,k} = \begin{cases} \dfrac{E_s \hat{H}^*\left(W_M^k\right)}{E_s\left|\hat{H}\left(W_M^k\right)\right|^2 + N_0} & \text{if } N_0 \geq N_t \\ \dfrac{E_s \hat{H}^*\left(W_M^k\right)}{E_s\left|\hat{H}\left(W_M^k\right)\right|^2 + N_t} & \text{if } N_0 < N_t \end{cases} \tag{39}$$

where N_t is the threshold noise level. In this case we choose $N_t = 10^{-3}$ since the channel MSE approaches at a value greater

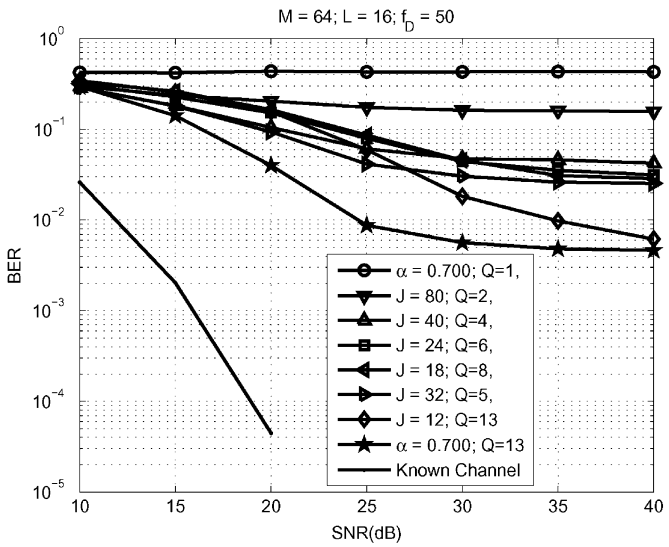

Fig. 11. BER performance for blind identification systems when the Doppler frequency is 50 Hz (54 km/h).

than or equal to 10^{-3} in most settings. (The plot for channel MSE is not shown due to space limit.)

For the second simulation, the Doppler frequency is chosen as 50 Hz, which corresponds to an object speed 15 m/s (54 km/h) if the carrier frequency is 1 GHz. The symbol duration is 10^{-6} seconds. This setting implies that the coherence interval is around 8×10^{-3} seconds, equal to 8000 symbol durations, or 100 received blocks. Since the channel is varying much faster than the previous case, we need to choose a much smaller J. The number of blocks J is ranging from 12 to 80, the parameter Q is chosen as the minimum value for each J, and N_B is chosen as $J/2$ for each J. The BER plot is shown in Fig. 11. A modified linear MMSE receiver as defined in (39) with $N_t = 10^{-2}$ is used when producing the BER plot. When $J = 80$, the performance is fairly poor since the estimated channel coefficients are hardly accurate due to channel variation. When the number of received blocks J is reduced, the performance becomes better and $J = 32$ yields the best performance in the low-SNR region among all values of J chosen in this simulation. When an even smaller J is chosen, performance in low-SNR region becomes worse again due to lack of data available for estimation. For high-SNR region, "$J = 12; Q = 13$" has the best performance. We also test the algorithm with a forgetting factor chosen as $\alpha = 0.7$ and repetition index as $Q = 13$. In this setting the data obtained 12 blocks earlier will be given a weighting of $\alpha^{12} \approx 0.0138$. If we use 1% as a threshold, we could say that the autocorrelation matrix [as defined in (35)] contains effective information from 12 composite blocks. This setting outperforms all other settings using a fixed J, which suggests the forgetting factor technique is more promising in a fast-varying channel environment. It should be especially noted that using a large repetition index $Q = 13$ makes it possible to choose a forgetting factor as small as 0.7. As shown in the plots, the same forgetting factor does not work at all for $Q = 1$.

In all our simulations here, we used $M = 64$. However, in some applications, M can have a much larger value (e.g., $M = 1024$). In this case, the task of blind estimation is more sensitive to time-varying channels. The number of blocks J needs to be chosen even smaller to fit in a coherence interval. Note that J can be chosen as small as three. This implies the requirement of a larger repetition index Q. As we learned in Section IV, the problem of rank deficiency of $\mathbf{U}_Q^{(J)}$ may arise. However, since M is large, the probability of rank deficiency would be much smaller. So the proposed algorithm has the potential to work well in the case of time-varying channels and a large M. The only concern here may be a high complexity as can be seen in Section III-E.

VI. Conclusion

In this paper, we proposed a generalized algorithm for subspace-based blind channel estimation in cyclic prefix systems. A new system parameter called the repetition index (Q) was introduced. By using a repetition index larger than unity, the number of received blocks (J) is significantly reduced compared to previously reported methods so that the proposed algorithm is more feasible in time-varying channel environments. A necessary condition on the system parameters J and Q for the algorithm to work is derived. The number of received blocks $J \geq 3$ can be chosen depending on the speed of channel variation to yield the best performance. The generalization can also applied to blind methods using a forgetting factor α. Simulation shows that when the number of received blocks J and the repetition index Q are properly chosen, the generalized algorithm outperforms previously reported special cases, especially in a time-varying channel environment. The proposed method can be directly applied to existing systems such as OFDM, SC-CP, etc., without any modification of the transmitter structure. In the future, developing the strategy to find the optimal J and Q or the optimal α and Q given knowledge of channel variation can be a challenging yet important problem. Extending this scheme for multiinput-multioutput (MIMO) channels is also of great interest.

Appendix A
Proofs of Theorems

Proof of Theorem 3: Assume $\mathbf{U}_{Q+1}^{(J)}$ does not have full row rank. Then there exists a nonzero row vector $\mathbf{v}^T = [v_1 \quad \cdots \quad v_{2M+Q}]$ such that $\mathbf{v}^T \mathbf{U}_{Q+1}^{(J)} = \mathbf{0}^T$. From the definition in (27), we obtain that $\mathbf{v}^T$ is a left annihilator of $\mathbf{U}_{Q+1}(n)$ for $1 \leq n \leq J-1$. The notation of $\mathbf{U}_Q(n)$ was defined in (24). Notice that $\mathbf{U}_Q(n)$ is a submatrix of $\mathbf{U}_{Q+1}(n)$ and can be obtained by removing the first row and the first column of $\mathbf{U}_{Q+1}(n)$, or by removing the last row and the last column of $\mathbf{U}_{Q+1}(n)$. This means that both $\mathbf{v}_1^T = [v_1 \quad \cdots \quad v_{2M+Q-1}]$ and $\mathbf{v}_2^T = [v_2 \quad \cdots \quad v_{2M+Q}]$ are left annihilators of $\mathbf{U}_Q(n)$ for $1 \leq n \leq J$. So $\mathbf{v}_1^T \mathbf{U}_Q^{(J)} = \mathbf{v}_2^T \mathbf{U}_Q^{(J)} = \mathbf{0}^T$. Since $\mathbf{v}^T$ is nonzero, at least one of $\mathbf{v}_1^T$ and $\mathbf{v}_2^T$ must also be nonzero. This implies that $\mathbf{U}_Q^{(J)}$ does not have full rank and contradicts the assumption. ■

Proof of Theorem 4: Let $\tilde{\mathbf{U}}_Q^{(J)} = \mathbf{K}\mathbf{U}_Q^{(J)}$ where

$$\mathbf{K} = \mathbf{I}_{2M+Q-1} - \begin{bmatrix} \mathbf{0}_{(M+Q-1)\times M} & \mathbf{I}_{M+Q-1} \\ \mathbf{0}_{M\times M} & \mathbf{0}_{(M+Q-1)\times M} \end{bmatrix}.$$

Then we have $\mathrm{rank}(\tilde{\mathbf{U}}_Q^{(J)}) = \mathrm{rank}(\mathbf{U}_Q^{(J)})$ since $\mathbf{K}$ is nonsingular. Also define $\mathbf{U}_Q(n) = \mathbf{K}\mathbf{U}_Q(n)$ where $\mathbf{U}_Q(n)$ is defined as in (24). It can be shown that $\mathbf{U}_Q(n)$ can be written as

$$\mathbf{U}_Q(n) = [[\mathbf{T}(n)]^T \quad [\mathbf{C}(n)]^T]^T$$

where $\mathbf{T}(n) = \mathcal{T}_Q\left(\mathbf{u}_M(n-1) - \mathbf{u}'_M(n)\right)$ is an $(M+Q-1)\times Q$ Toeplitz matrix and the $M \times Q$ matrix

$$\mathbf{C}(n) = \left[[\mathbf{u}'_M(n)]_M^1 \quad [\mathbf{u}'_M(n)]_{M+1}^2 \quad \cdots \quad [\mathbf{u}'_M(n)]_{2M+Q-2}^{M+Q-1}\right]$$

has a "circulant" structure. For simplicity, hereafter we denote $\mathbf{a}(n) = \mathbf{u}_M(n-1) - \mathbf{u}'_M(n)$ and $\mathbf{b}(n) = \mathbf{u}'_M(n)$. We also define polynomials in x as $A(x) = [1 \quad x \quad \cdots \quad x^{M-1}]\mathbf{a}(n)$ and $B(x) = [1 \quad x \quad \cdots \quad x^{M-1}]\mathbf{b}(n)$. $\tilde{\mathbf{U}}_Q(n)$ is a $(2M+Q-1)\times Q$ matrix and has at least $(2M-1)$ linearly independent left annihilators. These annihilators can always be written in the following forms, regardless of the value of Q.

$$\mathbf{v}_k^\dagger(n) = \left[1 \quad \alpha_k \cdots \alpha_k^{M+Q-2} \,\middle|\, \mathbf{0}_{1\times M}\right], \quad 1 \le k \le M-1 \quad (40)$$

and

$$\mathbf{v}_{M-1+k}^\dagger(n) = \left[B\left(W_M^{-k}\right)\mathbf{w}_k^\dagger \,\middle|\, -A\left(W_M^{-k}\right)\mathbf{x}_k^\dagger\right], \quad 1 \le k \le M \quad (41)$$

where $\{\alpha_1, \alpha_2, \ldots, \alpha_{M-1}\}$ are distinct roots of the polynomial $A(x)$, $\mathbf{w}_k^\dagger = [1 \quad W_M^{-k} \quad \cdots \quad W_M^{-k(M+Q-2)}]$, and $\mathbf{x}_k^\dagger = [1 \quad W_M^{-k} \quad \cdots \quad W_M^{-k(M-1)}]$. Please note that annihilators in the form of (40) come because of the Toeplitz structure of $\mathbf{T}(n)$ and annihilators in the form of (41) come because $A(W_M^{-k})$ and $B(W_M^{-k})$, the DFT coefficients of $\mathbf{a}(n)$ and $\mathbf{b}(n)$, respectively, cancel each other when $\tilde{\mathbf{U}}_Q(n)$ is multiplied by $\mathbf{v}_{M-1+k}^\dagger$ defined in (41). Here we omit the index n in polynomials $A(x)$ and $B(x)$ for the sake of notational simplicity. Also note that vectors $\mathbf{v}_k, 1 \le k \le 2M-1$, are always linearly independent as long as 1) the polynomial $A(x)$ has degree $M-1$; 2) all roots of $A(x)$ are distinct; and 3) none of roots of $A(x)$ is on the DFT grid. When any of these is not true, a slight modification of (40) and (41) can be found so that they are still linearly independent.

If $\tilde{\mathbf{U}}_Q^{(J)}$ is rank-deficient and there exists any left annihilator of $\tilde{\mathbf{U}}_Q^{(J)}$, in the form of either (40) or (41), then $\tilde{\mathbf{U}}_Q^{(J)}$ is rank-deficient for all Q, since the same form of vectors will continue to be annihilators of $\tilde{\mathbf{U}}_Q^{(J)}$. Now, we will prove that if $\tilde{\mathbf{U}}_{2M-1}^{(J)}$ is rank-deficient (as assumed in the theorem statement), then at least an annihilator in the form of either (40) or (41) will be a common annihilator for all $\tilde{\mathbf{U}}_Q(n)$. Suppose this is not the case and there exist two nonzero $\tilde{\mathbf{U}}_Q(n)$, say, $\tilde{\mathbf{U}}_Q(1)$ and $\tilde{\mathbf{U}}_Q(2)$, without loss of generality, which do not have common annihilators. Since $\tilde{\mathbf{U}}_Q^{(J)}$ is rank-deficient when $Q = 2M-1$ (as assumed in the theorem statement), there exists a nonzero $(4M-2)$-row vector $\mathbf{v}^\dagger$ such that $\mathbf{v}^\dagger\tilde{\mathbf{U}}_{2M-1}^{(J)} = \mathbf{0}^T$. Clearly, $\mathbf{v}^\dagger$ is also an annihilator of $\tilde{\mathbf{U}}_{2M-1}(1)$, and $\tilde{\mathbf{U}}_{2M-1}(2)$. Thus, $\mathbf{v}^\dagger$ can be decomposed into the following form:

$$\mathbf{v}^\dagger = \sum_{k=1}^{2M-1} c_k\mathbf{v}_k^\dagger(1) = \sum_{k=1}^{2M-1} d_k\mathbf{v}_k^\dagger(2)$$

where $\mathbf{v}_k^\dagger(n), 1 \le k \le 2M-1, n = 1, 2$ are as defined in (40) and (41) with $Q = 2M-1$. So we have

$$\mathbf{V}\begin{bmatrix}\mathbf{c}\\ -\mathbf{d}\end{bmatrix} = \mathbf{0} \quad (42)$$

where $\mathbf{c}$ and $\mathbf{d}$ are $(2M-1)$-column vectors containing coefficients c_k and d_k, respectively, and $\mathbf{V}$ is a $(4M-2)\times(4M-2)$ matrix whose columns are $\mathbf{v}_k(n), 1 \le k \le 2M-1, n = 1, 2$. Since the annihilators of $\tilde{\mathbf{U}}_Q(1)$ and $\tilde{\mathbf{U}}_Q(2)$ are linearly independent, $\mathbf{V}$ has full rank. Thus (42) implies $\mathbf{c} = \mathbf{d} = \mathbf{0}$ and, hence, $\mathbf{v}^\dagger = \mathbf{0}^T$. This contradicts the assumption that $\tilde{\mathbf{U}}_{2M-1}^{(J)}$ is rank-deficient. This completes the proof. ■

Appendix B
Probability of $\mathbf{U}_Q^{(J)}$ Having Full Rank For Different Precoders

We now explain why the probability of $\mathbf{U}_Q^{(J)}$ having full rank is much smaller when $\mathbf{R} = \mathbf{W}$ than $\mathbf{R} = \mathbf{I}$. As explained in the proof of Theorem 4, if $\mathbf{U}_Q^{(J)}$ does not have full rank for $Q \ge 2M-1$, then a row vector $\mathbf{v}^\dagger$ in the form of either (40) or (41) will be a common annihilator of $\tilde{\mathbf{U}}_Q(n)$. The probability of this depends on how many possible values of these vectors there are. Focusing on (41), since $\mathbf{w}_k^\dagger$ and $\mathbf{x}_k^\dagger$ are fixed, the variety of this form of annihilators comes from the values of $A(W_M^{-k})$ and $B(W_M^{-k})$, which are Fourier transforms of $\mathbf{a}(n)$ and $\mathbf{b}(n)$. If there is no precoding (i.e., $\mathbf{R} = \mathbf{I}$), the number of possible values of $A(W_M^{-k})$ and $B(W_M^{-k})$ can be quite large. On the contrary, when an IDFT precoder is used (i.e., $\mathbf{R} = \mathbf{W}^\dagger$), $A(W_M^{-k})$, and $B(W_M^{-k})$ can only be symbols in the constellation or the difference of two of them. Since the possible values of $A(W_M^{-k})$ and $B(W_M^{-k})$ are much fewer, it is more likely that a common annihilator of $\tilde{\mathbf{U}}_Q^{(J)}$ in this form exists. So the probability of $\mathbf{U}_Q^{(J)}$ having full rank is smaller in OFDM systems.

References

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

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

[3] M. C. Jeruchim, P. Balaban, and K. S. Shanmugan, *Simulation of Communication Systems*, 2nd ed. New York: Kluwer Academic/Plenum, 2000.

[4] 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.

[5] R. Lin and A. P. Petropulu, "Linear precoding assisted blind channel estimation for OFDM systems," *IEEE Trans. Veh. Technol.*, vol. 54, no. 3, pp. 983–995, May 2005.

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

[7] B. Muquet, M. de Courville, P. Duhamel, and V. Buzenac, "A subspace based blind and semi-blind channel identification method for OFDM systems," in *IEEE Work. Signal Process. Adv. Wireless Commun.*, May 1999, pp. 170–173.
[8] 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.
[9] A. P. Petropulu, R. Zhang, and R. Lin, "Blind OFDM channel estimation through simple linear precoding," *IEEE Trans. Wireless Commun.*, vol. 3, no. 2, pp. 647–655, Mar. 2004.
[10] D. H. Pham and J. H. Manton, "A subspace algorithm for guard interval based channel identification and source recovery requiring just two received blocks," in *Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing (ICASSP)*, , Hong Kong, China, 2003, pp. 317–320.
[11] A. Scaglione, G. B. Giannakis, and S. Barbarossa, "Redundant filter bank precoders and equalizers Part I: Unification and optimal designs," *IEEE Trans. Signal Process.*, vol. 47, no. 7, pp. 2007–2022, Jul. 1999.
[12] A. Scaglione, G. B. Giannakis, and S. Barbarossa, "Redundant filter bank precoders and equalizers Part II: Blind channel estimation, synchronization, and direct equalization," *IEEE Trans. Signal Process.*, vol. 47, no. 7, pp. 1988–2006, Jul. 1999.
[13] O. Shalvi and E. Weinstein, "New criteria for blind deconvolution of non-minimum phase systems (channels)," *IEEE Trans. Inf. Theory*, vol. 36, pp. 312–321, Mar. 1990.
[14] B. Su and P. P. Vaidyanathan, "A generalized algorithm for blind channel identification with linear redundant precoders," in *EURASIP J. Adv. Signal Process.*, 2007, vol. 2007, Article ID 25672, 13 pp..
[15] B. Su and P. P. Vaidyanathan, "A generalization of deterministic algorithm for blind channel identification with filter bank precoders," presented at the Int. Symp. Circuits Systems (ISCAS) 2006, Kos Island, Greece, May 2006.
[16] B. Su and P. P. Vaidyanathan, "Generalized subspace-based algorithms for blind channel estimation in cyclic prefix systems," presented at the 40th Asilomar Conf. Signals, Syst., Comp., Pacific Grove, CA, Nov. 2006.
[17] L. Tong, G. Xu, and T. Kailath, "A new approach to blind identification and equalization of multipath channels," in *Proc. 25th Asilomar Conf.*, 1991, pp. 856–860.
[18] P. P. Vaidyanathan, *Multirate Systems and Filter Banks*. Englewood Cliffs, NJ: Prentice-Hall, 1993.
[19] S. Zhou and G. B. Giannakis, "Finite-alphabet based channel estimation for OFDM and related multicarrier systems," *IEEE Trans. Commun.*, vol. 49, pp. 1402–1414, Aug. 2001.
[20] X. Zhuang, Z. Ding, and A. L. Swindlehurst, "A statistical subspace method for blind channel identification in OFDM communications," in *Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing (ICASSP)*, Istanbul, Turkey, Jun. 2000, vol. 5, pp. 2493–2496.

Borching Su (S'00) was born in Tainan, Taiwan, on October 8, 1978. He received the B.S. and M.S. degrees in electrical engineering and communication engineering, both from National Taiwan University (NTU), Taipei, Taiwan, in 1999 and 2001, respectively. He is currently pursuing the Ph.D. degree in the field of digital signal processing at the California Institute of Technology (Caltech), Pasadena.

His current research interests include multirate systems and their applications on digital communications.

Mr. Su was awarded the Moore Fellowship from Caltech in 2003.

P. P. Vaidyanathan (S'80–M'83–SM'88–F'91) was born in Calcutta, India, on October 16, 1954. He received the B.Sc. (Hons.) degree in physics and the B.Tech. and M.Tech. degrees in radio physics and electronics, all from the University of Calcutta, in 1974, 1977, and 1979, respectively, and the Ph.D. degree in electrical and computer engineering from the University of California at Santa Barbara in 1982.

He was a Postdoctoral Fellow at the University of California, Santa Barbara, from September 1982 to March 1983. In March 1983, he joined the Electrical Engineering Department, California Institute of Technology, Pasadena, as an Assistant Professor, and since 1993 has been Professor of electrical engineering. His main research interests are in digital signal processing, multirate systems, wavelet transforms, and signal processing for digital communications. He has authored a number of papers in IEEE journals, and is the author of the book *Multirate Systems and Filter Banks* (Englewood Cliffs, NJ: Prentice-Hall, 1993). He has written several chapters for various signal processing handbooks.

Dr. Vaidyanathan served as Vice-Chairman of the Technical Program Committee for the 1983 IEEE International symposium on Circuits and Systems, and Technical Program Chairman for the 1992 IEEE International symposium on Circuits and Systems. He was an Associate Editor for the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS during 1985–1987, and is currently an Associate Editor for the IEEE SIGNAL PROCESSING LETTERS, and a consulting editor for the *Applied and Computational Harmonic Analysis* journal. He was a Guest Editor in 1998 for special issues of the IEEE TRANSACTIONS ON SIGNAL PROCESSING and the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS II, on the topics of filter banks, wavelets, and subband coders. He was a recipient of the Award for Excellence in Teaching at the California Institute of Technology for the years 1983–1984, 1992–1993, and 1993–1994. He also received the NSF's Presidential Young Investigator award in 1986. In 1989, he received the IEEE ASSP Senior Award for his paper on multirate perfect-reconstruction filter banks. In 1990, he was the recipient of the S. K. Mitra Memorial Award from the Institute of Electronics and Telecommunications Engineers, India, for his joint paper in the *IETE Journal*. He was also the coauthor of a paper on linear-phase perfect reconstruction filter banks in the IEEE TRANSACTIONS ON SIGNAL PROCESSING, for which the first author (T. Nguyen) received the Young Outstanding Author award in 1993. He received the 1995 F. E. Terman Award of the American Society for Engineering Education, sponsored by Hewlett Packard Co., for his contributions to engineering education. He has given several plenary talks including at the Sampta'01, Eusipco'98, SPCOM'95, and Asilomar'88 conferences on signal processing. He has been chosen a distinguished lecturer for the IEEE Signal Processing Society for the year 1996–1997. In 1999, he received the IEEE CAS Society's Golden Jubilee Medal. He is a recipient of the IEEE Signal Processing Society's Technical Achievement Award for 2002.