

## Research Article

# A Generalized Algorithm for Blind Channel Identification with Linear Redundant Precoders

Borching Su and P. P. Vaidyanathan

*Department of Electrical Engineering, California Institute of Technology, Pasadena, CA 91125, USA*

Received 25 December 2005; Revised 19 April 2006; Accepted 11 June 2006

Recommended by See-May Phoong

It is well known that redundant filter bank precoders can be used for blind identification as well as equalization of FIR channels. Several algorithms have been proposed in the literature exploiting trailing zeros in the transmitter. In this paper we propose a generalized algorithm of which the previous algorithms are special cases. By carefully choosing system parameters, we can jointly optimize the system performance and computational complexity. Both time domain and frequency domain approaches of channel identification algorithms are proposed. Simulation results show that the proposed algorithm outperforms the previous ones when the parameters are optimally chosen, especially in time-varying channel environments. A new concept of generalized signal richness for vector signals is introduced of which several properties are studied.

Copyright © 2007 B. Su and P. P. Vaidyanathan. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

## 1. INTRODUCTION

Wireless communication systems often suffer from a problem due to multipath fading which makes the channels frequency-selective. Channel coefficients are often unknown to the receiver so that channel identification needs to be done before equalization can be performed. Among techniques for identifying unknown channel coefficients, blind methods have long been of great interest. In the literature many blind methods have been proposed based on the knowledge of second-order statistics (SOS) or higher-order statistics of the transmitted symbols [1, 2]. These methods often need to accumulate a large number of received symbols until channel coefficients can be estimated accurately. This requirement leads to a disadvantage when the system is working over a fast-varying channel.

A deterministic blind method using redundant filterbank precoders was proposed by Scaglione et al. [3] by exploiting trailing zeros introduced at the transmitter. Figure 1 shows a typical linear redundant precoded system. Source symbols are divided into blocks with size  $M$  and linearly precoded into  $P$ -symbol blocks which are then transmitted on the channel. It is well known that when  $P \geq M + L$ , where  $L$  is the maximum order of the FIR channel, interblock interference (IBI) can be completely eliminated in absence of

noise. When the block size  $M$  increases, the bandwidth efficiency  $\eta = (M + L)/M$  approaches unity asymptotically. The deterministic method proposed in [3] (which we will call the SGB method) exploits trailing zeros with length  $L$  introduced in each transmitted block and assumes the input sequence is *rich*. That is, the matrix composed of finite source blocks achieves full rank.

The method in [3] requires the receiver to accumulate at least  $M$  blocks before channel coefficients can be identified. This prevents the system from identifying channel coefficients accurately when the channel is fast-varying, especially when the block size  $M$  is large. More recently, Manton and Neumann pointed out that the channel could be identifiable with only two received blocks [4]. An algorithm based on viewing the channel identification problem as finding the greatest common divisor (GCD) of two polynomials is proposed in [5] (which we will call the MNP method). Even though it greatly reduces the number of received blocks needed for channel identification, the algorithm has much more computational complexity especially when the block size  $M$  is large.

In this paper, we propose a generalized algorithm of which the SGB algorithm proposed in [3] and the MNP algorithm in [5] are both special cases. By carefully choosing parameters, the system performance and computational

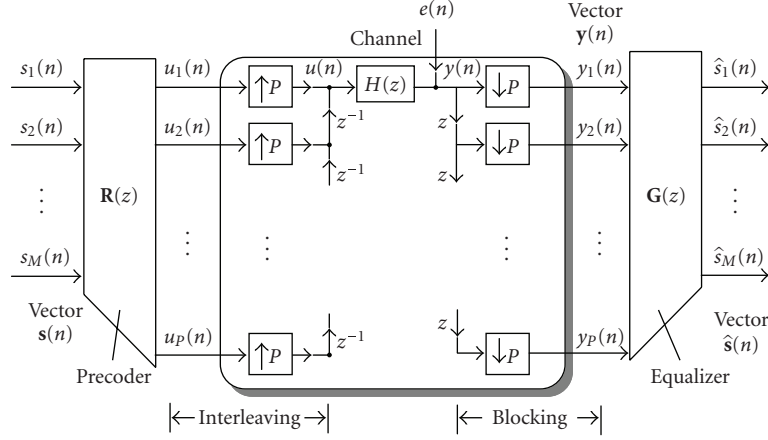


FIGURE 1: Communication system with redundant filter bank precoders.

complexity can be jointly optimized. The rest of the paper is organized as follows. Section 2 describes the system structure with linear precoder filter banks and reviews several existing blind algorithms. In Section 3 we present the generalized algorithm and derive the conditions on the input sequence under which the algorithm operates properly. In Section 4 we propose a frequency domain version of the generalized algorithm. The concept of *generalized signal richness* is introduced in Section 5 and some properties thereof are studied in detail. Simulation results and complexity analysis of both time and frequency domain approaches are presented in Section 6. In particular, simulations under time-varying channel environments are presented to demonstrate the strength of the proposed algorithm against channel variation. Finally, conclusions are made in Section 7. Some of the results in the paper have been presented at a conference [6].

### 1.1. Notations

Boldfaced lower-case letters represent column vectors. Boldfaced upper-case letters and calligraphic upper case letters are reserved for matrices. Superscripts as in  $\mathbf{A}^T$  and  $\mathbf{A}^\dagger$  denote the transpose and transpose-conjugate operations, respectively, of a matrix or a vector. All the vectors and matrices in this paper are complex-valued. In the figures “ $\uparrow P$ ” represents an expander and “ $\downarrow P$ ” a decimator [7].

If  $\mathbf{v} = [v_1 \ v_2 \ \cdots \ v_M]^T$  is an  $M \times 1$  column vector, then  $\mathcal{T}(\mathbf{v}, q)$  denotes an  $(M + q - 1) \times q$  Toeplitz matrix whose first row and first column are  $[v_1 \ 0 \ \cdots \ 0]$  and  $[v_1 \ v_2 \ \cdots \ v_M \ 0 \ \cdots \ 0^T]$ , respectively. For example,

$$\mathcal{T}\left(\begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \end{bmatrix}, 3\right) = \begin{bmatrix} a_1 & 0 & 0 \\ a_2 & a_1 & 0 \\ a_3 & a_2 & a_1 \\ a_4 & a_3 & a_2 \\ 0 & a_4 & a_3 \\ 0 & 0 & a_4 \end{bmatrix}. \quad (1)$$

## 2. PROBLEM FORMULATION AND LITERATURE REVIEW

### 2.1. Redundant filter bank precoders

Consider the multirate communication system [8] depicted in Figure 1. The source symbols  $s_1(n), s_2(n), \dots, s_M(n)$  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 a  $P \times M$  matrix  $\mathbf{R}(z)$  where  $P > M$ . The information with redundancy is then sent over the channel  $H(z)$ . We assume  $H(z)$  is an FIR channel with a maximum order  $L$ , that is,

$$H(z) = \sum_{k=0}^L h_k z^{-k}. \quad (2)$$

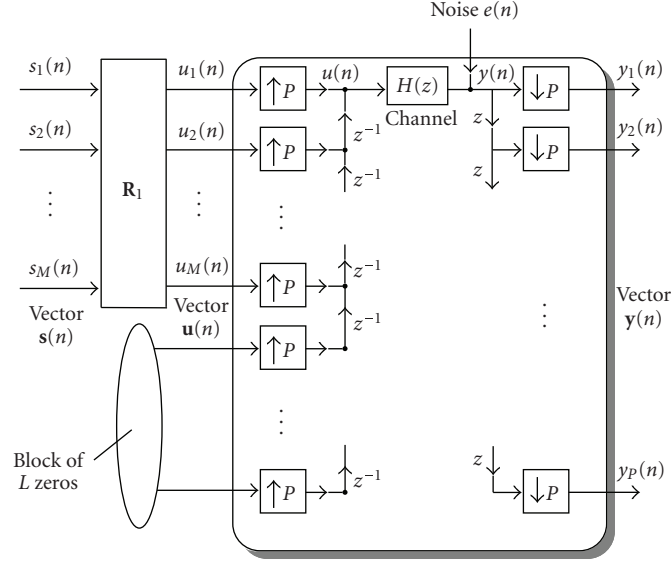
The signal is corrupted by channel noise  $e(n)$ . The received symbols  $y(n)$  are divided into  $P \times 1$  block vectors  $\mathbf{y}(n)$ . The  $M \times P$  matrix  $\mathbf{G}(z)$  is the channel equalizer and  $\hat{s}_1(n), \hat{s}_2(n), \dots, \hat{s}_M(n)$  are the recovered symbol streams. Also, for simplicity we define  $\mathbf{h}$  as the column vector  $[h_0 \ h_1 \ \cdots \ h_L^T]$ . We set

$$P = M + L, \quad (3)$$

that is, the redundancy introduced in a block is equal to the maximum channel order.

### 2.2. Trailing zeros as transmitter guard interval

Suppose we choose the precoder  $\mathbf{R}(z) = \begin{bmatrix} \mathbf{R}_1 \\ \mathbf{0} \end{bmatrix}$  where  $\mathbf{R}_1$  is an  $M \times M$  constant invertible matrix and the  $L \times M$  zero matrix  $\mathbf{0}$  represents zero-padding with length  $L$  in each transmitted block, as indicated in Figure 2. For simplicity of describing the algorithms, in this section we assume the noise is absent.

FIGURE 2: The zero-padding system with precoder  $\mathbf{R}_1$ .

Now, the received blocks can be written as

$$\begin{aligned} & \underbrace{\begin{bmatrix} \mathbf{y}(1) & \mathbf{y}(2) & \cdots & \mathbf{y}(J) \end{bmatrix}}_{\mathbf{Y} \text{ matrix; size } P \times J} \\ &= \mathcal{H}_M \mathbf{R}_1 \underbrace{\begin{bmatrix} \mathbf{s}(1) & \mathbf{s}(2) & \cdots & \mathbf{s}(J) \end{bmatrix}}_{\mathbf{S} \text{ matrix; size } M \times J}, \end{aligned} \quad (4)$$

where  $\mathcal{H}_M = \mathcal{T}(\mathbf{h}, M)$  is the full-banded Toeplitz channel matrix. As long as vector  $\mathbf{h}$  is nonzero, the matrix  $\mathcal{H}_M$  has full column rank  $M$ . Now, we assume the signal  $\mathbf{s}(n)$  is *rich*, that is, there exists an integer  $J$  such that the matrix  $\mathbf{S}$  has full row rank  $M$ . Since  $\mathbf{R}_1$  is an  $M \times M$  invertible matrix, we conclude that the  $P \times J$  matrix  $\mathbf{Y}$  has rank  $M$ . So there exist  $L$  linearly independent vectors that are left annihilators of  $\mathbf{Y}$ . In other words, there exists a  $P \times L$  matrix  $\mathbf{U}_0$  such that  $\mathbf{U}_0^\dagger \mathbf{Y} = \mathbf{U}_0^\dagger \mathcal{H}_M \mathbf{R}_1 \mathbf{S} = \mathbf{0}$ . Now that  $\mathbf{R}_1 \mathbf{S}$  has rank  $M$ , this implies

$$\mathbf{U}_0^\dagger \mathcal{H}_M = \mathbf{0}. \quad (5)$$

The channel coefficients  $\mathbf{h}$  can then be determined by solving (5). In practice where *channel noise* is present, the computation of the annihilators is replaced with the computation of the eigenvectors corresponding to the smallest  $L$  singular values of  $\mathbf{Y}$ . In this and the following sections, the channel noise term is not shown explicitly.

Note that this algorithm [3] works under the assumption that  $\mathbf{S}$  has full row rank  $M$ . Obviously  $J \geq M$  is a necessary condition for this assumption. This means the receiver must accumulate at least  $M$  blocks (i.e., a duration of  $M(M+L)$  symbols) before channel identification can be performed. This could be a disadvantage when the system is working over a fast-varying channel.

### 2.3. The GCD approach

Another approach proposed in [5] requires only two received blocks for blind channel identification. Recall that the channel is described by  $\mathbf{y} = \mathcal{H}_M \mathbf{u} = \mathcal{T}(\mathbf{h}, M) \mathbf{u}$ , or

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_P \end{bmatrix} = \begin{bmatrix} h_0 & \mathbf{0} \\ h_1 & \ddots \\ \vdots & h_0 \\ h_L & h_1 \\ \vdots & \ddots \\ \mathbf{0} & h_L \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_M \end{bmatrix}. \quad (6)$$

By multiplying  $[1 \ x \ x^2 \ \cdots \ x^{P-1}]$  to both sides of (6), we obtain

$$y(x) = h(x)u(x), \quad (7)$$

where

$$\begin{aligned} y(x) &\triangleq \sum_{k=0}^{P-1} y_{k+1} x^k, & h(x) &\triangleq \sum_{k=0}^L h_k x^k, \\ u(x) &\triangleq \sum_{k=0}^{M-1} u_{k+1} x^k \end{aligned} \quad (8)$$

are polynomial representations of the output vector, channel vector, and input vector, respectively. This means, (6) is nothing but a polynomial multiplication. Now, suppose we have two received blocks  $\mathbf{y}(1)$  and  $\mathbf{y}(2)$ , and let  $y_1(x) = h(x)u_1(x)$  and  $y_2(x) = h(x)u_2(x)$  represent the polynomial forms of these. Then the channel polynomial  $h(x)$  can be found as the GCD of  $y_1(x)$  and  $y_2(x)$ , given that the input polynomials  $u_1(x)$  and  $u_2(x)$  are coprime to each other.

To compute the GCD of  $y_1(x)$  and  $y_2(x)$ , we first construct a  $(2P - 1) \times 2P$  matrix [9]

$$\mathbf{Y}_P \triangleq \begin{bmatrix} y_{11} & 0 & \cdots & 0 & y_{21} & 0 & \cdots & 0 \\ y_{12} & y_{11} & \ddots & \vdots & y_{22} & y_{21} & \ddots & \vdots \\ \vdots & y_{12} & \ddots & 0 & \vdots & y_{22} & \ddots & 0 \\ y_{1P} & \vdots & & y_{11} & y_{2P} & \vdots & & y_{21} \\ 0 & y_{1P} & & y_{12} & 0 & y_{2P} & & y_{22} \\ \vdots & \ddots & \ddots & \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & y_{1P} & 0 & \cdots & 0 & y_{2P} \end{bmatrix}. \quad (9)$$

One can verify that

$$\mathbf{Y}_P = \underbrace{\begin{bmatrix} h_0 & \mathbf{0} \\ h_1 & \ddots \\ \vdots & h_0 \\ h_L & h_1 \\ \vdots & \vdots \\ \mathbf{0} & h_L \end{bmatrix}}_{\substack{\text{matrix } \mathcal{H}_{M+P-1} \\ \text{size } (2P-1) \times (M+P-1)}} \underbrace{\begin{bmatrix} u_{11} & \mathbf{0} & u_{21} & \mathbf{0} \\ u_{12} & \ddots & u_{22} & \ddots \\ \vdots & u_{11} & \vdots & u_{21} \\ u_{1M} & u_{12} & u_{2M} & u_{22} \\ \vdots & \vdots & \vdots & \vdots \\ \mathbf{0} & u_{1M} & \mathbf{0} & u_{2M} \end{bmatrix}}_{\substack{\text{matrix } \mathcal{U} \\ \text{size } (M+P-1) \times 2P}}. \quad (10)$$

When  $u_1(x)$  and  $u_2(x)$  are coprime to each other, it can be shown that the matrix  $\mathcal{U}$  has full rank  $M + P - 1$  (see Section 5). Since  $\mathcal{H}_{M+P-1} = \mathcal{T}(\mathbf{h}, M + P - 1)$  also has rank  $M + P - 1$ ,  $\text{rank}(\mathbf{Y}_P) = M + P - 1$  and hence  $\mathbf{Y}_P$  has  $L$  left annihilators (i.e., there exists a  $(2P - 1) \times L$  matrix  $\mathbf{U}_0$  such that  $\mathbf{U}_0^\dagger \mathbf{Y} = \mathbf{0}$ ). These annihilators are also annihilators of each column of matrix  $\mathcal{H}_{M+P-1}$ , and we can therefore, in absence of noise, identify channel coefficients  $h_0, h_1, \dots, h_L$  up to a scalar ambiguity. In presence of noise, the columns of  $\mathbf{U}_0$  would be selected as the eigenvectors associated with the smallest singular values of  $\mathbf{Y}_P$ .

## 2.4. Connection to the earlier literature

The MNP method described above can be viewed as a dual version of the subspace methods proposed in the earlier literature in multichannel blind identification [10, 11]. In the subspace method in [11], the single source can be estimated as the GCD of the received data from two (more generally  $N$ ) different antennas. The MNP method [5] swaps the roles of data blocks and multichannel coefficients.

## 3. A GENERALIZED ALGORITHM

In this section we propose a generalized algorithm of which each of the two algorithms described in the previous section is a special case. Comparing the two algorithms described above, we find that the MNP approach needs much fewer received blocks for blind identifiability. However, it has more computational complexity. Each received block is repeated  $P$  times to build a big matrix. Using the generalized algorithm, we can choose the number of repetitions and the number of

received blocks freely as long as they satisfy a certain constraint.

### 3.1. Algorithm description

Observe (6) again and note that it can be rewritten as

$$\mathcal{T}(\mathbf{y}, Q) = \mathcal{T}(\mathbf{h}, M + Q - 1) \mathcal{T}(\mathbf{u}, Q), \quad (11)$$

where  $\mathcal{T}(\cdot, \cdot)$  is defined as in (1). Here  $Q$  can be any positive integer. Note that in the MNP method  $Q$  is chosen as  $P$ , as described in the previous section. Suppose the receiver gathers  $J$  blocks with  $J \geq 2$ . Then we have  $\mathbf{Y}_Q^{(J)} = \mathcal{H}_{M+Q-1} \mathbf{U}_Q^{(J)}$ , where

$$\mathbf{Y}_Q^{(J)} = \begin{bmatrix} \mathcal{T}(\mathbf{y}(1), Q) & \mathcal{T}(\mathbf{y}(2), Q) & \cdots & \mathcal{T}(\mathbf{y}(J), Q) \end{bmatrix},$$

$$\mathcal{H}_{M+Q-1} = \mathcal{T}(\mathbf{h}, M + Q - 1), \quad (12)$$

$$\mathbf{U}_Q^{(J)} = \begin{bmatrix} \mathcal{T}(\mathbf{u}(1), P) & \cdots & \mathcal{T}(\mathbf{u}(J), P) \end{bmatrix}. \quad (13)$$

Note that  $\mathbf{U}_Q^{(J)}$  has size  $(M + Q - 1) \times QJ$  and  $\mathbf{Y}_Q^{(J)}$  has size  $(P + Q - 1) \times QJ$ . For notational simplicity, from now on we will use subscript  $Q$  as in  $N_Q$  to denote  $N_Q = N + Q - 1$  where  $N$  denotes a positive integer. In particular,

$$M_Q = M + Q - 1, \quad (14)$$

$$P_Q = P + Q - 1.$$

Notice that they still have the relationship  $P_Q = M_Q + L$ .

Assume now the matrix  $\mathbf{U}_Q^{(J)}$  has full row rank  $M_Q$ . Taking singular-value decomposition (SVD) of  $\mathbf{Y}_Q^{(J)}$  we have

$$\mathbf{Y}_Q^{(J)} = \begin{bmatrix} \mathbf{U}_r & \mathbf{U}_0 \end{bmatrix} \begin{bmatrix} \boldsymbol{\Sigma} \\ \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{V}_r & \mathbf{V}_0 \end{bmatrix}^\dagger. \quad (15)$$

The size of  $\boldsymbol{\Sigma}$  is  $M_Q \times M_Q$  since both  $\mathcal{H}_{M_Q}$  and  $\mathbf{U}_Q^{(J)}$  have full rank  $M_Q$ . The columns of the  $M_Q \times L$  matrix  $\mathbf{U}_0$  are left annihilators of matrix  $\mathbf{Y}^{(J)}$  and also of  $\mathbf{H}$  since  $\mathbf{U}^{(J)}$  has full row rank. Suppose

$$\mathbf{U}_0^\dagger = \begin{bmatrix} u_{11} & u_{12} & \cdots & u_{1,P+Q-1} \\ u_{21} & u_{22} & \cdots & u_{2,P+Q-1} \\ \vdots & & & \vdots \\ u_{L1} & u_{L2} & \cdots & u_{L,P+Q-1} \end{bmatrix}. \quad (16)$$

Form the Hankel matrices

$$\mathcal{U}_k \triangleq \begin{bmatrix} u_{k1} & u_{k2} & \cdots & u_{k,L+1} \\ u_{k2} & u_{k3} & \cdots & u_{k,L+2} \\ \vdots & & & \vdots \\ u_{k,M_Q} & u_{k,M_Q+1} & \cdots & u_{k,P_Q} \end{bmatrix} \quad (17)$$

for  $k, 1 \leq k \leq L$ . Then we have

$$\underbrace{\begin{bmatrix} \mathcal{U}_1 \\ \mathcal{U}_2 \\ \vdots \\ \mathcal{U}_L \end{bmatrix}}_{\mathcal{U} \text{ matrix; size } LM_Q \times (L+1)} \mathbf{h} = \mathbf{0}. \quad (18)$$

Vector  $\mathbf{h}$  can thus be identified up to a scalar ambiguity.

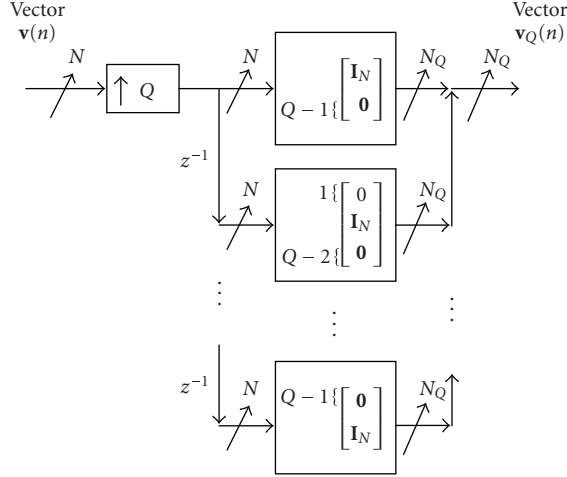


FIGURE 3: Q-repetition and shifting operation.

### 3.2. Q-repetition and shifting operation

As we can see in the previous section, the repetition and shifting operation on a vector signal is crucial in the generalized algorithm. Figure 3 gives a block diagram of this operation. For future notational convenience, the subscript  $Q$  as in  $\mathbf{v}_Q(n)$  denotes the result of this operation on a vector signal. By viewing (11) and applying this operation on  $\mathbf{y}(n)$  and  $\mathbf{u}(n)$ , we obtain the relationship

$$\mathbf{y}_Q(n) = \mathcal{H}_{M+Q-1} \mathbf{u}_Q(n)$$

for any positive integer  $Q$ .

### 3.3. Special cases of the algorithm

The blind channel identification algorithm described above uses two parameters: (a) the number of received blocks  $J$ ; (b) the number of repetitions per block  $Q$ . A number of points should be noted here:

- (1) the algorithm works for any  $J$  and  $Q$  as long as  $\mathbf{U}_Q^{(J)}$  has full row rank  $M_Q$ . This is the only constraint for choosing parameters  $J$  and  $Q$ ;
- (2) note that if we choose  $Q = 1$  and  $J \geq M$ , then the algorithm reduces to the SGB algorithm [3];
- (3) if we choose  $Q = P$  and  $J = 2$ , it becomes the MNP algorithm [5].

So both the SGB method and the MNP method are a special case of the proposed algorithm. Since  $\mathbf{U}_Q^{(J)}$  has size  $M_Q \times QJ$ ,  $\mathbf{U}_Q^{(J)}$  having full row rank implies  $QJ \geq M_Q = M + Q - 1$ , or

$$Q \geq \frac{M-1}{J-1}. \quad (19)$$

Also note that we cannot choose  $J = 1$  since  $\mathbf{U}_Q^{(J)}$  can never have full rank unless the block size  $M = 1$ . This is consistent with the theory that two blocks are required for blind channel identification [4]. While the inequality (19) is a necessary

condition for  $\mathbf{U}_Q^{(J)}$  to have full rank, it is not sufficient because it also depends on the values of entries of  $\mathbf{u}(n)$ . Nevertheless, when inequality (19) is satisfied, the probability of  $\mathbf{U}_Q^{(J)}$  having full rank is usually close to unity in practice, especially when a large symbol constellation is used. Thus,

$$Q = \left\lceil \frac{M-1}{J-1} \right\rceil \quad (20)$$

appears to be a selection that minimizes the computational cost given the number of received blocks  $J$ . A detailed study on the conditions for  $\mathbf{U}_Q^{(J)}$  to have full rank is presented in Section 5.

When  $J = 2$ ,  $Q$  can be chosen as small as  $M - 1$  rather than  $P$ . If we take  $J = 3$ ,  $Q = \lceil (M - 1/2) \rceil$  makes the matrix  $\mathbf{Y}$  twice smaller. We can choose  $Q = 1$  only when  $J \geq M$ . This coincides with the SGB algorithm which uses a richness assumption [3].

## 4. FREQUENCY DOMAIN APPROACH

In this section we slightly modify the blind identification algorithm and directly estimate the frequency responses of the channel at different frequency bins and equalize the channel in the frequency domain. We call the modified algorithm *frequency domain approach*. Some of the ideas come from [12]. The receiver structure for the frequency domain approach is shown in Figure 4. To demonstrate how this system works, observe the  $P_Q \times M_Q$  full-banded Toeplitz channel matrix

$$\mathcal{H}_{M_Q} = \mathcal{T}(\mathbf{h}, M_Q). \quad (21)$$

Define a row vector  $\mathbf{v}_\rho^T = [1 \ \rho^{-1} \ \dots \ \rho^{-(P_Q-1)}]$  with  $\rho$  a nonzero complex number. Due to full-banded Toeplitz structure of  $\mathcal{H}_{M_Q}$ , we have

$$\mathbf{v}_\rho^T \mathcal{H}_{M_Q} = [H(\rho) \ \rho^{-1} H(\rho) \ \dots \ \rho^{-(M_Q-1)} H(\rho)], \quad (22)$$

where  $H(\rho) = \sum_{k=0}^L h_k \rho^{-k}$  is the channel  $z$ -transform evaluated at  $z = \rho$ .

Let  $N$  be chosen as an integer greater than or equal to  $P_Q$ , and let  $\rho_1, \rho_2, \dots, \rho_N$  be distinct nonzero complex numbers. Consider an  $N \times P_Q$  matrix  $\mathbf{V}_{N \times P_Q}$  whose  $i$ th row is  $\mathbf{v}_{\rho_i}^T$ :

$$\mathbf{V}_{N \times P_Q} = \begin{bmatrix} 1 & \rho_1^{-1} & \rho_1^{-2} & \dots & \rho_1^{-(P_Q-1)} \\ 1 & \rho_2^{-1} & \rho_2^{-2} & \dots & \rho_2^{-(P_Q-1)} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \rho_N^{-1} & \rho_N^{-2} & \dots & \rho_N^{-(P_Q-1)} \end{bmatrix}. \quad (23)$$

It is easy to verify that

$$\mathbf{V}_{N \times P_Q} \mathcal{H}_{M_Q} = \mathbf{A}_N \underbrace{\begin{bmatrix} 1 & \rho_1^{-1} & \dots & \rho_1^{-(M_Q-1)} \\ 1 & \rho_2^{-1} & \dots & \rho_2^{-(M_Q-1)} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & \rho_N^{-1} & \dots & \rho_N^{-(M_Q-1)} \end{bmatrix}}_{\mathbf{V}_{N \times M_Q} \text{ matrix}}, \quad (24)$$

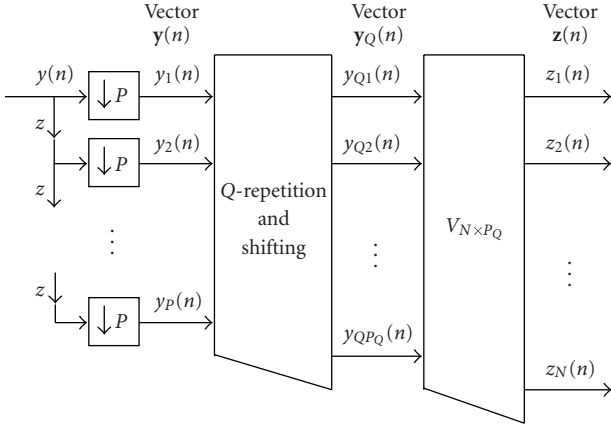


FIGURE 4: Receiver structure for frequency domain approach.

where

$$\Lambda_N = \text{diag} \left( \begin{bmatrix} H(\rho_1) & H(\rho_2) & \cdots & H(\rho_N) \end{bmatrix} \right) \triangleq \text{diag}(\tilde{\mathbf{h}}_N) \quad (25)$$

is a diagonal matrix with frequency domain channel coefficients as the diagonal entries. Now, when we gather receiving blocks and repeat them as in (12), we get the following matrix:

$$\mathbf{Y}_Q^{(J)} = \begin{bmatrix} \mathcal{T}(\mathbf{y}(1), Q) & \mathcal{T}(\mathbf{y}(2), Q) & \cdots & \mathcal{T}(\mathbf{y}(J), Q) \end{bmatrix}. \quad (26)$$

Since we have  $\mathbf{Y}_Q^{(J)} = \mathcal{H}_{M_Q} \mathbf{U}_Q^{(J)}$  in absence of noise, by multiplying  $\mathbf{V}_{N \times P_Q}$  and  $\mathbf{Y}_Q^{(J)}$ , we have

$$\mathbf{Z} = \mathbf{V}_{N \times P_Q} \mathbf{Y}_Q^{(J)} = \mathbf{V}_{N \times P_Q} \mathcal{H}_{M_Q} \mathbf{U}_Q^{(J)} = \Lambda_N \mathbf{V}_{N \times M_Q} \mathbf{U}_Q^{(J)}. \quad (27)$$

Recall that  $\text{rank}(\mathbf{Y}_Q^{(J)}) = \text{rank}(\mathbf{U}_Q^{(J)}) = M_Q$ . Since  $\rho_1, \rho_2, \dots, \rho_N$  are all distinct, the matrix  $\mathbf{Z}$  has the same rank as  $\mathbf{Y}_Q^{(J)}$ . The dimension of the null space of matrix  $\mathbf{Z}$  is hence  $N - M_Q$ . By performing SVD on  $\mathbf{Z}$ , we can find these  $N - M_Q$  left annihilators of  $\mathbf{Z}$ , which are also annihilators of  $\Lambda_N \mathbf{V}_{N \times M_Q}$ . There exists an  $(N - M_Q) \times N$  matrix  $\mathbf{U}_0^\dagger$  such that  $\mathbf{U}_0^\dagger \mathbf{Z} = \mathbf{0}$ . Since  $\mathbf{U}_Q^{(J)}$  has full rank, this implies

$$\mathbf{U}_0^\dagger \Lambda_N \mathbf{V}_{N \times M_Q} = \mathbf{0}. \quad (28)$$

Suppose

$$\mathbf{U}_0^\dagger = \begin{bmatrix} u_{11} & u_{12} & \cdots & u_{1N} \\ u_{21} & u_{22} & \cdots & u_{2N} \\ \vdots & \vdots & & \vdots \\ u_{N-M_Q,1} & u_{N-M_Q,2} & \cdots & u_{N-M_Q,N} \end{bmatrix}. \quad (29)$$

Then by observing the  $ij$ th entry of (28), we have

$$\mathbf{u}_{ij}^\dagger \tilde{\mathbf{h}}_N^\dagger = 0 \quad (30)$$

for all  $i, j$ ,  $1 \leq i \leq N - M_Q$  and  $1 \leq j \leq M_Q$ , where  $\mathbf{u}_{ij} = [u_{i1}\rho_1^{-(j-1)} \ u_{i2}\rho_2^{-(j-1)} \ \cdots \ u_{iN}\rho_N^{-(j-1)}]^\dagger$ . Here  $\tilde{\mathbf{h}}_N$  is the row

vector in (25). Form the  $M_Q \times N$  matrices

$$\mathcal{U}_i = \begin{bmatrix} u_{i1} & u_{i2} & \cdots & u_{iN} \\ u_{i1}\rho_1^{-1} & u_{i2}\rho_2^{-1} & \cdots & u_{iN}\rho_N^{-1} \\ u_{i1}\rho_1^{-2} & u_{i2}\rho_2^{-2} & \cdots & u_{iN}\rho_N^{-2} \\ \vdots & \vdots & \ddots & \vdots \\ u_{i1}\rho_1^{-(M_Q-1)} & u_{i2}\rho_2^{-(M_Q-1)} & \cdots & u_{iN}\rho_N^{-(M_Q-1)} \end{bmatrix}, \quad (31)$$

and let  $\mathcal{U} = [\mathcal{U}_1^T \ \mathcal{U}_2^T \ \cdots \ \mathcal{U}_{N-M_Q}^T]^T$ . Then from (30) we have  $\mathcal{U} \tilde{\mathbf{h}}_N = \mathbf{0}$ . Then the frequency domain channel coefficients  $\tilde{\mathbf{h}}_N$  can be estimated by solving this equation. After the frequency domain channel coefficients are estimated, the received symbols can be equalized directly in the frequency domain, as in DMT systems.

Recall that we have the freedom to choose  $N$  as any integer greater than or equal to  $P_Q$  and the values of  $\rho_i$ ,  $1 \leq i \leq N$  as any nonzero complex number in the  $z$ -domain. In this paper, we use  $N = P_Q$  and

$$\rho_k = \exp\left(\frac{j2k\pi}{N}\right), \quad k = 0, 1, \dots, N-1. \quad (32)$$

Note that since  $H(z)$  is an  $L$ th order system, there are at most  $L$  values among  $H(\rho_i)$  which can be zero (channel nulls). By choosing  $N \geq P_Q$ , there are at least  $M_Q$  nonzero values among  $H(\rho_i)$ ,  $i = 1, 2, \dots, P_Q$ . In practice we can choose to equalize the received symbols in frequency bins associated with the largest  $M_Q$  frequency responses  $H(\rho_i)$  to enhance the system performance. This provides resistance to channel nulls.

## 5. GENERALIZED SIGNAL RICHNESS

For the generalized blind channel identification method proposed in this paper to work properly, the matrix  $\mathbf{U}_Q^{(J)}$  defined in (13) must have full row rank for given parameters  $J$  and  $Q$ . An obvious necessary condition has been presented as inequality (19) in Section 3. The sufficiency, however, depends on the content of signal  $\mathbf{u}(n)$ . When  $Q = 1$  and  $\mathbf{u}(n)$  is *rich*, then there exists  $J$  such that  $\mathbf{U}_Q^{(J)} = [\mathbf{u}(0) \ \mathbf{u}(1) \ \cdots \ \mathbf{u}(J-1)]$  has full rank. When  $Q > 1$ ,  $\mathbf{u}(n)$  requires another kind of richness property so that  $\mathbf{U}_Q^{(J)}$  has full rank for a finite integer  $J$ . We call this property the *generalized signal richness* and define it as follows.

*Definition 1.* An  $M \times 1$  sequence  $\mathbf{u}(n)$ ,  $n \geq 0$  is said to be  $(1/Q)$ -rich if there exists a finite integer  $J$  such that the  $(M + Q - 1) \times JQ$  matrix

$$\mathbf{U}_Q^{(J)} = \begin{bmatrix} \mathcal{T}(\mathbf{s}(0), Q) & \mathcal{T}(\mathbf{s}(1), Q) & \cdots & \mathcal{T}(\mathbf{s}(J), Q) \end{bmatrix} \quad (33)$$

has full row rank  $M + Q - 1$ .

Several interesting properties of generalized signal richness will be presented in this section. The reason why we use the notation of  $(1/Q)$  will soon be clear when these properties are presented.



### 5.1. Measure of generalized signal richness

**Lemma 1.** *If an  $M \times 1$  sequence  $\mathbf{s}(n)$  is  $(1/Q)$ -rich, then  $\mathbf{s}(n)$  is  $(1/(Q+1))$ -rich.*

*Proof.* See the appendix.  $\square$

Lemma 1 states a basic property of generalized signal richness: the smaller the value of  $Q$  is, the “stronger” the condition of  $(1/Q)$ -richness is. For example, if an  $M \times 1$  sequence  $\mathbf{s}(n)$  is 1-rich, or simply *rich*, then it is  $(1/Q)$ -rich for any positive integer  $Q$ . On the contrary, a  $(1/2)$ -rich signal  $\mathbf{s}(n)$  is not necessarily 1-rich. We can thus define a measure of generalized signal richness for a given  $M \times 1$  sequence  $\mathbf{s}(n)$  as follows.

**Definition 2.** Given an  $M \times 1$  sequence  $\mathbf{s}(n)$ ,  $n \geq 0$ , the *degree of nonrichness* of  $\mathbf{s}(n)$  is defined as

$$Q_{\min} \triangleq \min \left( \mathbf{s}(n) \text{ is } \frac{1}{Q} \text{-rich} \right). \quad (34)$$

Recall that the larger the degree of nonrichness  $Q_{\min}$  is, the weaker the richness of the signal  $\mathbf{s}(n)$  is. If  $\mathbf{s}(n)$  is not  $(1/Q)$ -rich for any  $Q$ , then  $Q_{\min} = \infty$ . The property of an infinite degree of nonrichness can be described in the following lemma. We use the notation  $\mathbf{p}_M(x)$  to denote the column vector:

$$\mathbf{p}_M(x) = \begin{bmatrix} 1 & x & x^2 & \cdots & x^{M-1} \end{bmatrix}^T. \quad (35)$$

**Lemma 2.** *Consider an  $M \times 1$  sequence  $\mathbf{s}(n)$ . The following statements are equivalent:*

- (1)  $\mathbf{s}(n)$  is not  $(1/Q)$ -rich for any  $Q$ ;
- (2) the degree of nonrichness of  $\mathbf{s}(n)$  is infinity;
- (3) either there exists a complex number  $\alpha$  such that  $\begin{bmatrix} 1 & \alpha & \cdots & \alpha^{M-1} \end{bmatrix}$  is an annihilator of  $\mathbf{s}(n)$  or  $\begin{bmatrix} 0 & \cdots & 0 & 1 \end{bmatrix}$  is an annihilator of  $\mathbf{s}(n)$ ;
- (4) either polynomials  $p_n(x) = \mathbf{p}_M^T(x)\mathbf{s}(n)$ ,  $n \geq 0$  share a common zero (at  $\alpha$ ) or their orders are all less than  $M-1$ .

*Proof.* See the appendix.  $\square$

Note that the statement  $\begin{bmatrix} 0 & \cdots & 0 & 1 \end{bmatrix}$  is an annihilator of  $\mathbf{s}(n)$  in condition (3) and the statement that polynomials  $p_n(x)$  have orders less than  $M-1$  in condition (4) can be interpreted as the special situation when the common zero  $\alpha$  is at infinity.

If an  $M \times 1$  sequence  $\mathbf{s}(n)$  has a finite degree of nonrichness, or  $\mathbf{s}(n)$  is  $(1/Q)$ -rich for some integer  $Q$ , then it can be shown that the maximum possible value of  $Q_{\min}$  is  $M-1$ , as described in the following lemma.

**Lemma 3.** *If  $M > 1$  and an  $M \times 1$  sequence  $\mathbf{s}(n)$  is not  $(1/(M-1))$ -rich, then it is not  $(1/Q)$ -rich for any  $Q$ .*

*Proof.* See the appendix.  $\square$

With Lemma 3, we can see that for an  $M \times 1$  sequence  $\mathbf{s}(n)$ , the possible values of the degree of non-richness  $Q_{\min}$

are  $1, 2, \dots, M-1$ , and  $\infty$ .  $(1/(M-1))$ -richness is thus the weakest form of generalized richness. When using the MNP method [9], this weakest form of generalized richness is very crucial. If this weakest form of richness of  $\mathbf{s}(n)$  is not achieved, then by Lemma 2  $\mathbf{s}(n)$  has an infinite degree of non-richness and polynomials  $\mathbf{p}_M^T(x)\mathbf{s}(n)$  have a common factor  $(x - \alpha)$ . Then as in Section 2.3, when we take GCD of the polynomials representing the received blocks, the receiver would be unable to determine whether the factor  $(x - \alpha)$  belongs to the channel polynomial or is a common factor of the symbol polynomials. *Therefore, if the input signal  $\mathbf{s}(n)$  has infinite degree of non-richness, all methods proposed in this paper will fail for all  $Q$ .*

Furthermore, the MNP method proposed in [5] uses  $Q = P$ . Using Lemma 3, we see that using  $Q = M-1$  is sufficient if we are computing the GCD of polynomials representing received blocks and the following two conditions are true: (1) the GCD is known to have a degree less than or equal to  $L$ ; (2) the degree of each symbol polynomial is less than or equal to  $M-1$ . Using  $Q = P$  not only is computationally unnecessary, but also, as we will see in simulation results in Section 6, has sometimes a worse performance than using  $Q = M-1$  in presence of noise.

The sufficiency of  $Q = M-1$  can also be understood from the point of view of polynomial theory. Suppose polynomials  $a(x)$  and  $b(x)$  have degrees less than or equal to  $P-1$  and have a greatest common denominator  $d(x)$  whose degree is less than or equal to  $L$ . Suppose  $a(x) = d(x)a_1(x)$  and  $b(x) = d(x)b_1(x)$  and both  $a_1(x)$  and  $b_1(x)$  have degrees less than or equal to  $M-1$  and they are coprime to each other. Then there exists polynomials  $p(x)$  and  $q(x)$  whose degree are less than or equal to  $M-2$  such that  $1 = p(x)a_1(x) + q(x)b_1(x)$  and thus  $d(x) = p(x)a(x) + q(x)b(x)$ .

### 5.2. Connection to earlier literature

An earlier proposition mathematically equivalent to Lemma 3 has been presented in the single-input-multiple-output (SIMO) blind equalization literature [10, 13]. We review it here briefly.

**Proposition 1.** *Let  $\mathbf{h}[n]$  be  $J \times 1$  vectors. Suppose a  $QJ \times (Q+M-1)$  block Toeplitz matrix*

$$\mathcal{T}_Q(\mathbf{h}) = \begin{bmatrix} \mathbf{h}[0] & \mathbf{h}[1] & \cdots & \mathbf{h}[M-1] & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{h}[0] & \mathbf{h}[1] & \cdots & \mathbf{h}[M-1] & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \mathbf{0} \\ \mathbf{0} & \cdots & \mathbf{0} & \mathbf{h}[0] & \mathbf{h}[1] & \cdots & \mathbf{h}[M-1] \end{bmatrix} \quad (36)$$

*satisfies the following conditions:*

- (1)  $\mathbf{h}[0] \neq \mathbf{0}$  and  $\mathbf{h}[M-1] \neq \mathbf{0}$ ;
- (2)  $\mathbf{h}[n] = \mathbf{0}$  for  $n < 0$  and  $n \geq M$ ;
- (3)  $Q \geq M-1$ .

Then  $\mathcal{T}_Q(\mathbf{h})$  has full column rank if and only if

$$\mathbf{h}(z) \triangleq \sum_{i=0}^M \mathbf{h}[i]z^{-i} \neq \mathbf{0}, \quad \forall z. \quad (37)$$

Here  $\mathbf{h}[n]$  was used to refer to the impulse response of a  $J \times 1$  channel.  $Q$  stands for the observation period in the multiple-channel receiver end. Conditions (1) and (2) imply that the channel has finite impulse response. Condition (3) can be met by increasing the observation period  $Q$ . While this old proposition focuses on the coefficients of multiple channels rather than values of transmitted symbols, it is mathematically equivalent to the statement that  $\mathbf{s}(n)$  is  $(1/(M-1))$ -rich if and only if polynomials  $\mathbf{p}_M^T(x)\mathbf{s}(n)$  do not share common zeros. The case of  $Q < M - 1$ , however, has not been considered earlier in the literature, to the best of our knowledge.

### 5.3. Remarks on generalized signal richness

In this section we introduced the concept of generalized signal richness. Given an  $M \times 1$  signal  $\mathbf{s}(n)$ ,  $n \geq 0$ , the degree of non-richness  $Q_{\min}$  was defined. For an input signal with a degree of non-richness  $Q_{\min}$ , we can choose any

$$Q \geq Q_{\min} \quad (38)$$

and some finite  $J$  for the generalized algorithm proposed in Section 3 to work properly. The possible values of  $Q_{\min}$  are  $1, 2, \dots, M - 1$ , and  $\infty$ . If  $\mathbf{s}(n)$  has an infinite degree of non-richness, the algorithm proposed in this paper will fail for all  $Q$ . The degree of non-richness of a signal  $\mathbf{s}(n)$  directly depends on its content. A deeper study of degree of non-richness will be presented elsewhere [14].

## 6. SIMULATIONS AND DISCUSSIONS

In this section, several simulation results, comparisons, and discussions will be presented. We will first test our proposed method and compare it with the existing methods [3, 5] described in Section 2. Secondly, we will compare the performances of time domain versus frequency domain approaches and show that under some channel conditions the frequency domain approach outperforms the time domain approach. Finally, we will analyze and compare the computational complexity of algorithms proposed in this paper.

### 6.1. Simulations of time domain approaches

A Rayleigh fading channel of order  $L = 4$  is used. The size of transmitted blocks is  $M = 8$  and received block size is  $P = M + L = 12$ . 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:

$$E_{ch} = \frac{\|\hat{\mathbf{h}} - \mathbf{h}\|^2}{\|\mathbf{h}\|^2}, \quad (39)$$

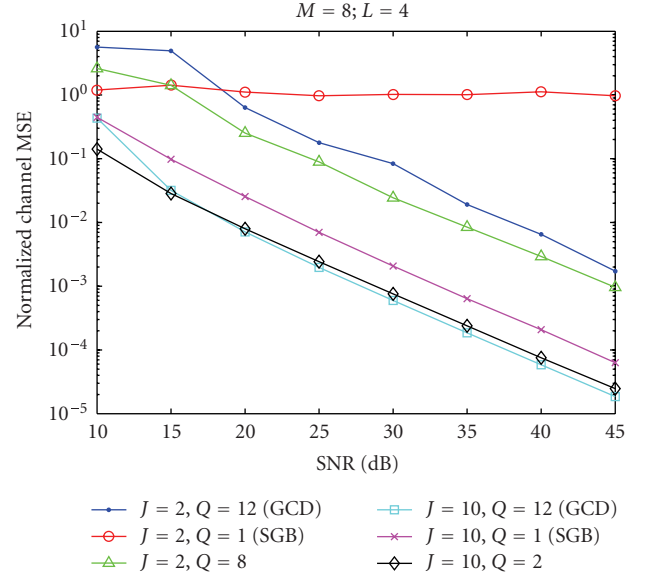


FIGURE 5: Normalized least squared channel error estimation.

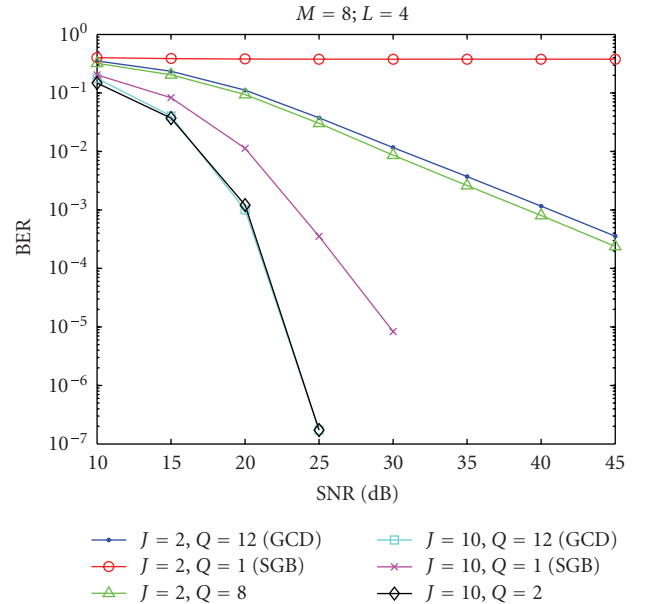


FIGURE 6: Bit error rate.

where  $\hat{\mathbf{h}}$  and  $\mathbf{h}$  are the estimated and the true channel vectors, respectively. The simulated normalized channel estimation error is shown in Figure 5 and the corresponding BER is presented in Figure 6. When the number of blocks  $J = 10$ , the MNP method (with the number of block repetitions  $Q = 12$ ) outperforms the SGB method ( $Q = 1$ ) by a considerable range. Taking  $Q = 2$  saves a lot of computation and yet yields a good performance as indicated. Furthermore, in the case of  $J = 2$ , the system with  $Q = 8$  even outperforms the original MNP method with  $Q = 12$ . This also strengthens our argument in Section 5 that choosing  $Q$  as large as  $P$  is unnecessary.



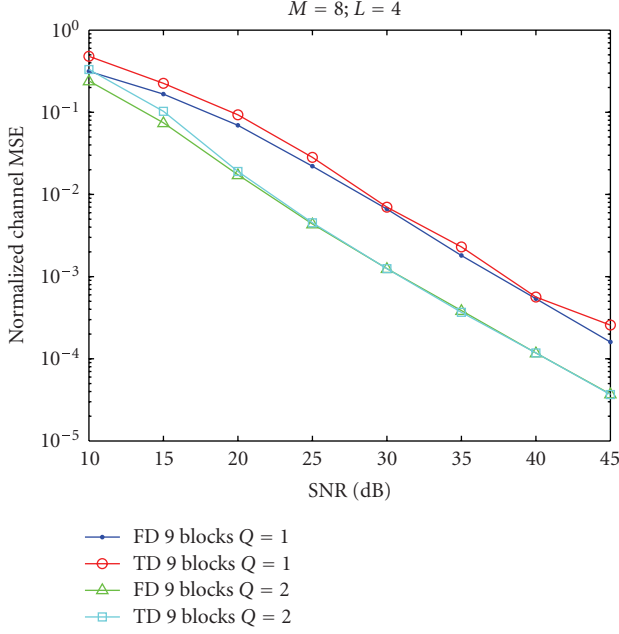


FIGURE 7: Normalized least squared channel error estimation.

### 6.2. Simulations of frequency domain approaches

Figure 7 shows the comparison of frequency domain approach and time domain approach under the channel coefficients  $H(z) = 1 - jz^{-1} + (-1 + 0.01j)z^{-2} + (0.01 + j)z^{-3} - 0.01jz^{-4}$ .

For frequency domain approach, the normalized least squared channel error is defined as

$$E_{ch} = \frac{\|\hat{\mathbf{h}} - \tilde{\mathbf{h}}\|^2}{\|\tilde{\mathbf{h}}\|^2}, \quad (40)$$

where

$$\tilde{\mathbf{h}} = [H(\rho_1) \ H(\rho_2) \ \cdots \ H(\rho_N)] \quad (41)$$

and  $\hat{\mathbf{h}}$  is the estimation of  $\tilde{\mathbf{h}}$ . Simulation results show that frequency domain approach outperforms time domain approach especially when the noise level is high. While the frequency domain approach does not in general beat the time domain approach for a random channel, it has been consistently observed that frequency domain approach performs better than time domain approach when the last channel coefficient  $h(L)$  has a small magnitude (i.e., at least one zero of  $H(z)$  is close to the origin).

Since we have the freedom to choose values of coefficients  $\rho_i$ , the receiver can adjust  $\rho_i$  dynamically according to the a priori knowledge of the approximated channel zero locations. This is especially useful when the channel coefficients are changing slowly from block to block.

### 6.3. Complexity analysis

For the algorithms presented in Section 3, the SVD computation dominates the computational complexity. The number of blocks  $J$ , the number of repetitions per block  $Q$ , and the received block size  $P$  decide the size of the matrix on which SVD is taken. The complexity of SVD operation on an  $n \times m$  matrix [15] is on the order of  $\mathcal{O}(mn^2)$  with  $m \geq n$ . Since  $\mathbf{Y}_Q^{(j)}$  has size  $(P+Q-1) \times QJ$ , the complexity is  $\mathcal{O}(QJ(P+Q-1)^2)$ . We can see that the complexity can be greatly reduced by choosing a smaller  $Q$ . Recall that the SGB method [3] uses  $Q = 1$  and the MNP method [5] uses  $Q = P$ . We thus have the following arguments:

- (i) the MNP method has a complexity around  $4P$  times the complexity of the SGB method for any  $J$ . A choice of  $Q$  between 1 and  $P$  could be seen as a compromise between system performance and complexity;
- (ii) when  $J$  is large, we have the freedom to choose a smaller  $Q$ , as explained in the previous section.

For the frequency domain approach presented in Section 4, an additional matrix multiplication is required. This demands extra computational complexity of the order of  $\mathcal{O}(JP_Q^2)$ . However, if the values  $\rho_i$  are chosen as equally spaced on the unit circle, an FFT algorithm can be exploited and the computational complexity will be reduced to  $\mathcal{O}(JP_Q \log P_Q)$  and is negligible compared to the complexity of SVD operations.

### 6.4. Simulations for time-varying channels

In this section, we demonstrate the capability of the proposed generalized blind identification algorithm in time-varying channels environments. The received symbols can be expressed as

$$y(n) = \sum_{k=0}^L h(n, k)x(n - k), \quad (42)$$

where the  $(L+1)$ -tap channel coefficients  $h(n, k)$  vary as the time index  $n$  changes. We generate the channel coefficients as follows. During a time interval  $T$ , the channel coefficients change from  $h_1(k)$  to  $h_2(k)$ , where  $h_1(k)$  and  $h_2(k)$ ,  $0 \leq k \leq L$  represent two sets of  $(L+1)$ -tap independent coefficients. The variation of the coefficient is done by linear interpolation such that

$$h(n, k) = \begin{cases} h_1(k), & \text{if } n = 0, \\ h_2(k), & \text{if } n = T, \\ \frac{T-n}{T}h_1(k) + \frac{n}{T}h_2(k) & \text{otherwise.} \end{cases} \quad (43)$$

In our simulation, we choose  $T = 180$ . Coefficients of  $h_1(k)$  and  $h_2(k)$  are given in Table 1. The size of transmitted blocks is  $M = 8$  and received block size is  $P = M + L = 12$  (so the channel coefficients completely change after 15 blocks). Simulations are performed under different choices of  $J$  and  $Q$ , as indicated in Figures 8 and 9. The normalized least squared

TABLE 1: Coefficients for the time-varying channel.

$k$	$h_1(k)$	$h_2(k)$
0	$-0.6563 + 0.7059i$	$-1.2519 + 0.2295i$
1	$-0.6534 + 1.1774i$	$0.9347 + 0.1237i$
2	$-0.4229 - 0.2362i$	$0.0346 - 0.6180i$
3	$0.2145 - 0.2207i$	$0.7272 - 1.4084i$
4	$-0.1478 + 0.2802i$	$0.8612 + 0.3455i$

channel error is defined as

$$E_{ch} = \frac{\|\hat{\mathbf{h}} - \bar{\mathbf{h}}\|^2}{\|\bar{\mathbf{h}}\|^2}, \quad (44)$$

where  $\hat{\mathbf{h}}$  is the estimated channel and  $\bar{\mathbf{h}}$  is the averaged coefficients during the time the channel is being estimated:

$$\bar{\mathbf{h}} = \frac{1}{JP} \sum_{n=n_0}^{n_0+JP-1} [h(n,0) \ h(n,1) \ \cdots \ h(n,L)]^T. \quad (45)$$

In Figure 8 we see that when  $J = 10$ , the time range is too large for the algorithm to estimate the time-varying channel accurately. The performance for  $J = 2$  is much better in high SNR region because the channel does not vary too much during the time of two blocks. However, in low SNR region the performance for  $J = 2$  becomes bad. The case for  $J = 4$  has the best performance among all other choices because the channel does not vary too much during the duration of four receiving blocks, and more data are available for accurate estimation. This simulation result provides clues about how we can choose the optimal  $J$ : if the channel variation is fast ( $T$  is smaller) we need a smaller  $J$  while we can use a larger  $J$  when  $T$  is larger.

### 6.5. Remarks on choosing the optimal parameters

According to the simulations results above, we summarize here a general guideline to choose a set of optimal parameters in practice.

- (1) When the channel is constant and for a fixed  $Q$ , a larger  $J$  appears to have a better performance (as shown in Figure 5) since more data are available for accurate estimation.
- (2) When the channel is time-varying, the optimal choice of  $J$  depends on the speed of channel variation. Simulation results in Figures 8 and 9 suggest when the channel coefficients completely change in  $N$  blocks, a choice of  $J \approx N/4$  could be appropriate.
- (3) Suppose  $J$  is given, a choice of  $Q$  as the smallest integer that satisfies inequality (19) often has a satisfactory performance. A slightly larger  $Q$  can sometimes be better (see Figure 5 for  $J = 10$ ) at the expense of a slightly increased complexity. However, if  $Q$  is too large, the performance could be even worse (see Figure 5 for  $J = 2$ ,  $Q = 12$ ).

The guidelines above are given by observing the simulation results. An analytically optimal set of  $J$  and  $Q$  is still under investigation.

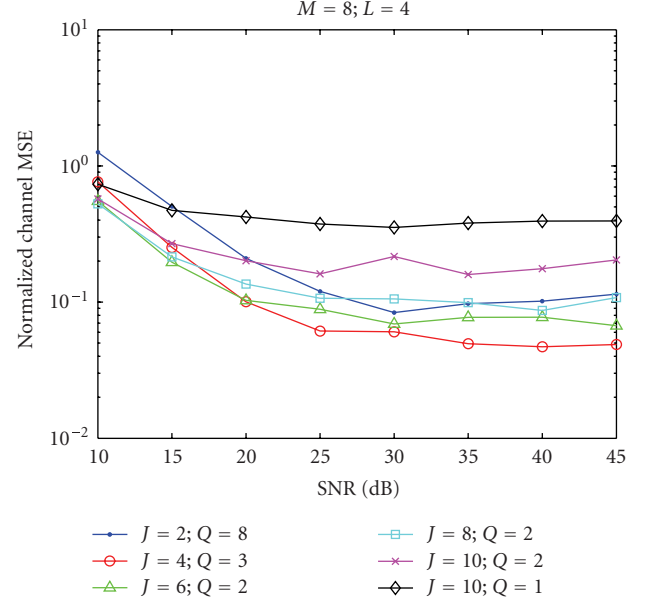


FIGURE 8: Normalized channel MSE performance for a time-varying channel.

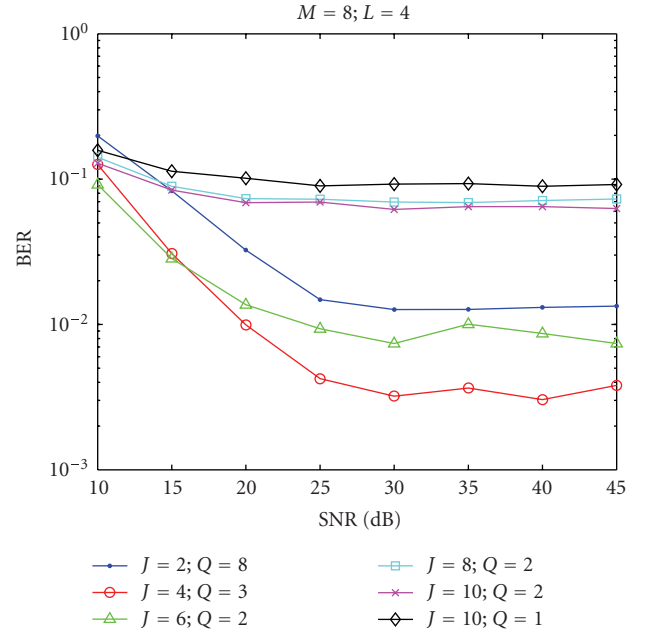


FIGURE 9: Bit error rate performance for a time-varying channel.

### 6.6. Noise handling for large $J$

It should be noted that when  $J$  is very large (and  $Q = 1$ ), the proposed method behaves like a traditional subspace method using second-order statistics. Suppose

$$\mathbf{Y}^{(J)} = \mathbf{H}\mathbf{U}^{(J)} + \mathbf{E}^{(J)}, \quad (46)$$

where  $\mathbf{E}^{(J)}$  is composed of  $J$  columns of noise vectors  $\mathbf{e}(n)$ . The autocorrelation matrix of received blocks can be estimated as

$$\mathbf{R}_{yy} = E[\mathbf{y}(n)\mathbf{y}^\dagger(n)] \approx \frac{1}{J} \mathbf{Y}^{(J)} \mathbf{Y}^{(J)\dagger}. \quad (47)$$

If the input signal and channel noise are uncorrelated, we can write  $\mathbf{R}_{yy}$  as

$$\mathbf{R}_{yy} = \mathbf{H} \mathbf{R}_{uu} \mathbf{H}^\dagger + \mathbf{R}_{ee}, \quad (48)$$

where  $\mathbf{R}_{uu} = E[\mathbf{u}(n)\mathbf{u}^\dagger(n)]$  and  $\mathbf{R}_{ee} = E[\mathbf{e}(n)\mathbf{e}^\dagger(n)]$  are autocorrelation matrices of input blocks and noise vectors, respectively. If  $\mathbf{R}_{ee}$  is known (e.g., if the noise is white and noise variance is  $N_0$ , then  $\mathbf{R}_{ee} = N_0 \mathbf{I}_p$ ), an improved estimation of annihilators of matrix  $\mathbf{H}$  can be performed by taking eigen-decomposition of  $\mathbf{R}_{yy} - \mathbf{R}_{ee}$ , which results in better channel estimation [3]. This technique, however, does not apply when  $J$  is small.

## 7. CONCLUDING REMARKS

In this paper we proposed a generalized algorithm for blind channel identification with linear redundant precoders. The number of received blocks  $J \geq 2$  can be chosen freely depending on the speed of channel variation. The minimum number of repetitions  $Q$  of each received block is derived to optimize the computational complexity while retaining good performance. Simulation shows that when the system parameter  $Q$  is properly chosen, the generalized algorithm outperforms previously reported special cases, especially in a time-varying channel environments.

A *frequency domain* version of the generalized algorithm is also presented. Simulation result shows that it outperforms time domain approach at low SNR region for certain types of channels, for example, channels with a zero close to the origin. Since we have the freedom to choose different frequency parameters in the frequency domain approach, certain choices other than equally spaced grids on the unit circle can be used to improve the system performance for different channel zero locations. An even more challenging problem might be to analytically derive the optimal frequency points for a specific type of channel.

The concept of *generalized signal richness* for a vector signal is introduced. With the *degree of non-richness* of the input signal decided, we can determine the minimum number of repetitions theoretically. A complete set of necessary and sufficient conditions for signals satisfying generalized signal richness is still under investigation. The study of effect of a linear precoder on the property of generalized signal richness could also be a challenging problem.

## APPENDIX

*Proof of Lemma 1.* Suppose  $\mathbf{s}(n)$  is  $(1/Q)$ -rich but not  $(1/(Q+1))$ -rich, then there exists a  $1 \times (M+Q)$  nonzero vector  $\mathbf{v}^T = [v_1 \ v_2 \ \dots \ v_{M+Q}]$  such that

$$\mathbf{v}^T \mathcal{T}(\mathbf{s}(n), Q+1) = \mathbf{0}_{1 \times (Q+1)}, \quad \forall n. \quad (A.1)$$

Observing the first  $Q$  elements of the vector equation above, we obtain

$$[v_1 \ v_2 \ \dots \ v_{M+Q-1}] \mathcal{T}(\mathbf{s}(n), Q) = \mathbf{0}_{1 \times Q}, \quad \forall n. \quad (A.2)$$

Without loss of generality, assume  $[v_1 \ v_2 \ \dots \ v_{M+Q-1}]$  to be nonzero and it is an annihilator of  $\mathcal{T}(\mathbf{s}(n), Q)$ . This violates the assumption that  $\mathbf{s}(n)$  is  $(1/Q)$ -rich.  $\square$

*Proof of Lemma 2.* Conditions (1) and (2) are equivalent by definition. The equivalence of conditions (3) and (4) can also be easily examined. If condition (3) is true, then either  $\mathbf{p}_{M+Q-1}^T(\alpha)$  or  $[0 \ \dots \ 0 \ 1]$  is an annihilator of  $\mathbf{s}_Q(n)$  (as defined in Section 3.2) for all  $Q$  and hence condition (1) is also true. In the case condition (1) is true, assume there exists  $n \geq 0$  such that the degree of the polynomial  $\mathbf{p}_M^T(x)\mathbf{s}(n)$  is  $M-1$ . Then for any  $Q$ , there exists a row vector  $\mathbf{v}^T = [v_1 \ v_2 \ \dots \ v_{M+Q-1}]$  such that  $\mathbf{v}^T \mathbf{s}_Q(n) = 0$ , for all  $n$ . This implies

$$\sum_{l=1}^M v_{k+l} [\mathbf{s}(n)]_l = 0, \quad \forall n, k \geq 0, \quad (A.3)$$

where  $[\cdot]_l$  represents the  $l$ th element of a column vector. So the series  $\{v_k\}_{k=1}^{M+Q-1}$  must satisfy the recurrence (A.3) for any  $n \geq 0$ . This requires the characteristic polynomials  $\mathbf{p}_M^T(x)\mathbf{s}(n)$ ,  $n \geq 0$  to share at least one zero. So condition (4) must be true. By the arguments above, these four conditions are equivalent.  $\square$

*Proof of Lemma 3.* If  $\mathbf{s}(n)$  is proportional to a same nonzero vector  $\mathbf{x}$  for all  $n$ , then it is obviously not  $(1/Q)$ -rich for any  $Q$ . We thus assume without loss of generality that  $\mathbf{s}(0)$  and  $\mathbf{s}(1)$  are linearly independent. Suppose polynomials  $\mathbf{p}_M^T(x)\mathbf{s}(0)$  and  $\mathbf{p}_M^T(x)\mathbf{s}(1)$  have two sets of *distinct* zeros  $\{\alpha_{01}, \alpha_{02}, \dots, \alpha_{0,M-1}\}$  and  $\{\alpha_{11}, \alpha_{12}, \dots, \alpha_{1,M-1}\}$ , respectively. Since  $\mathbf{s}(n)$  is not  $(1/Q)$ -rich, there exists a  $(2M-2)$ -row vector  $\mathbf{v}^T = [v_1 \ v_2 \ \dots \ v_{2M-2}]$  such that  $\mathbf{v}^T \mathcal{T}(\mathbf{s}(n), M-1) = \mathbf{0}_{1 \times (M-1)}$ . We have that the nonzero row vector  $\mathbf{v}^T$  must have the form of

$$\begin{aligned} \mathbf{v}^T &= \sum_{k=1}^{M-1} c_k \begin{bmatrix} 1 & \alpha_{0,k}^{-1} & \alpha_{0,k}^{-2} & \dots & \alpha_{0,k}^{-(M-2)} \end{bmatrix} \\ &= \sum_{k=1}^{M-1} d_k \begin{bmatrix} 1 & \alpha_{1,k}^{-1} & \alpha_{1,k}^{-2} & \dots & \alpha_{1,k}^{-(M-2)} \end{bmatrix} \end{aligned} \quad (A.4)$$

for some coefficients  $c_1, c_2, \dots, c_{M-1}$ ,  $d_1, d_2, \dots, d_{M-1}$ . This implies

$$[\mathbf{c}^T \ -\mathbf{d}^T] \mathbf{V} = \mathbf{0}^T, \quad (A.5)$$

where  $\mathbf{c}^T = [c_1 \ c_2 \ \dots \ c_{M-1}]$ ,  $\mathbf{d}^T = [d_1 \ d_2 \ \dots \ d_{M-1}]$ , and

$$\mathbf{V} = \begin{bmatrix} \mathbf{p}_{2M-2}^T(\alpha_{01}) \\ \vdots \\ \mathbf{p}_{2M-2}^T(\alpha_{0,M-1}) \\ \mathbf{p}_{2M-2}^T(\alpha_{11}) \\ \vdots \\ \mathbf{p}_{2M-2}^T(\alpha_{1,M-1}) \end{bmatrix} \quad (\text{A.6})$$

is a Vandermonde matrix. If all zeros  $\{\alpha_{ij}\}$  are distinct,  $\mathbf{V}$  is a  $(2M-2) \times (2M-2)$  invertible matrix and (A.5) implies  $\mathbf{c}^T = \mathbf{d}^T = \mathbf{0}^T$  and hence  $\mathbf{v}^T = \mathbf{0}^T$ . This contradicts the assumption that  $\mathbf{s}(n)$  is not  $(1/(M-1))$ -rich. Therefore, if  $\mathbf{s}(n)$  is not  $(1/(M-1))$ -rich, there must be a common zero shared by  $\mathbf{p}_{2M-2}^T(x)\mathbf{s}(0)$  and  $\mathbf{p}_{2M-2}^T(x)\mathbf{s}(1)$ . Similarly, we can obtain that there exists an  $\alpha$  such that  $\mathbf{p}_{2M-2}^T(\alpha)\mathbf{s}(n) = 0$  for all  $n$ . Using Lemma 2, this implies that  $\mathbf{s}(n)$  is not  $(1/Q)$ -rich for all  $Q$ .

In the case where the polynomial  $\mathbf{p}_{2M-2}^T(x)\mathbf{s}(n)$  has multiple zeros for some  $n$ , the matrix  $\mathbf{V}$  in (A.5) can be replaced with a *confluent Vandermonde matrix* [15] which is still invertible.  $\square$

## ACKNOWLEDGMENTS

This work was supported in part by the NSF Grant CCF-0428326, ONR Grant N00014-06-1-0011, and the Moore Fellowship of the California Institute of Technology.

## REFERENCES

- [1] B. Porat and B. Friedlander, "Blind equalization of digital communication channels using high-order moments," *IEEE Transactions on Signal Processing*, vol. 39, no. 2, pp. 522–526, 1991.
- [2] L. Tong, G. Xu, and T. Kailath, "Blind identification and equalization based on second-order statistics: a time domain approach," *IEEE Transactions on Information Theory*, vol. 40, no. 2, pp. 340–349, 1994.
- [3] A. Scaglione, G. B. Giannakis, and S. Barbarossa, "Redundant filter bank precoders and equalizers part II: blind channel estimation, synchronization, and direct equalization," *IEEE Transactions on Signal Processing*, vol. 47, no. 7, pp. 2007–2022, 1999.
- [4] J. H. Manton and W. D. Neumann, "Totally blind channel identification by exploiting guard intervals," *Systems and Control Letters*, vol. 48, no. 2, pp. 113–119, 2003.
- [5] 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 *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP '03)*, vol. 4, pp. 317–320, Hong Kong, April 2003.
- [6] B. Su and P. P. Vaidyanathan, "A generalization of deterministic algorithm for blind channel identification with filter bank precoders," in *Proceedings of IEEE International Symposium on Circuits and Systems (ISCAS '06)*, Kos Island, Greece, May 2006.

- [7] P. P. Vaidyanathan, *Multirate Systems and Filter Banks*, Prentice-Hall, Englewood Cliffs, NJ, USA, 1993.
- [8] Y.-P. Lin and S.-M. Phoong, "Perfect discrete multitone modulation with optimal transceivers," *IEEE Transactions on Signal Processing*, vol. 48, no. 6, pp. 1702–1711, 2000.
- [9] W. Qiu, Y. Hua, and K. Abed-Meraim, "A subspace method for the computation of the GCD of polynomials," *Automatica*, vol. 33, no. 4, pp. 741–743, 1997.
- [10] L. Tong, G. Xu, and T. Kailath, "A new approach to blind identification and equalization of multipath channels," in *Proceedings of the 25th Asilomar Conference on Signals, Systems, & Computers*, vol. 2, pp. 856–860, Pacific Grove, Calif, USA, November 1991.
- [11] E. Moulines, P. Duhamel, J.-F. Cardoso, and S. Mayrargue, "Subspace methods for the blind identification of multichannel FIR filters," *IEEE Transactions on Signal Processing*, vol. 43, no. 2, pp. 516–525, 1995.
- [12] P. P. Vaidyanathan and B. Vrcelj, "A frequency domain approach for blind identification with filter bank precoders," in *Proceedings of IEEE International Symposium on Circuits and Systems (ISCAS '04)*, vol. 3, pp. 349–352, Vancouver, BC, Canada, May 2004.
- [13] Y. Li and Z. Ding, "Blind channel identification based on second order cyclostationary statistics," in *Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP '93)*, vol. 4, pp. 81–84, Minneapolis, Minn, USA, April 1993.
- [14] B. Su and P. P. Vaidyanathan, "Generalized signal richness preservation problem and Vandermonde-form preserving matrices," to appear in *IEEE Transactions on Signal Processing*.
- [15] G. H. Golub and C. F. Van Loan, *Matrix Computations*, Johns Hopkins University Press, Baltimore, MD, USA, 3rd edition, 1996.

**Borching Su** 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 California Institute of Technology (Caltech). In 2003, he was awarded the Moore Fellowship from Caltech. His current research interests include multirate systems and their applications on digital communications.



**P. P. Vaidyanathan** received the B.Tech. and M.Tech. degrees in radiophysics and electronics, from the University of Calcutta, and the Ph.D. degree in electrical and computer engineering from the University of California at Santa Barbara, in 1982. Since then he has been with the Faculty of Electrical Engineering at the California Institute of Technology. He has authored many papers in the signal processing area. He has received several awards for excellence in teaching at the California Institute of Technology. In 1989, he received the IEEE ASSP Senior Paper Award. In 1990, he was 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 is a Fellow of the IEEE. He received the 1995 F. E. Terman Award of the American Society for Engineering Education. He has served as a Distinguished Lecturer for the IEEE Signal Processing Society (1996-1997). In 1999 he was chosen to receive the IEEE CAS Society's Golden Jubilee Medal. He was a recipient of the IEEE Signal Processing Society's Technical Achievement Award for the year 2002.