Data-Efficient Blind OFDM Channel Estimation Using Receiver Diversity

Hao Wang, *Student Member, IEEE*, Ying Lin, and Biao Chen, *Member, IEEE*

***Abstract*—In this paper, we investigate non data-aided channel estimation for cyclically prefixed orthogonal frequency division multiplexing (OFDM) systems. By exploiting channel diversity using only two receive antennas, a blind deterministic algorithm is proposed. Identifiability conditions are derived that guarantee the perfect channel retrieval in the absence of noise. In the presence of noise, the proposed method has the desired property of being data efficient—only a single OFDM block is needed to achieve good estimation performance for a wide range of SNR values. The algorithm is also robust to input symbols as it does not have any restriction on the input symbols with regard to their constellation or their statistical properties. In addition, this diversity-based algorithm is computationally efficient, and its performance compares favorably to most existing blind algorithms.**

***Index Terms*—Channel estimation, Cramér–Rao lower bound (CRLB), orthogonal frequency division multiplexing (OFDM), receiver diversity.**

I. INTRODUCTION

CURRENT and future broadband wireless communication systems aim to provide high data rate services. As a result, multipath fading becomes a major concern as systems with high data rate are more liable to intersymbol interference (ISI) due to the increase in the normalized delay spread. It is therefore imperative to use modulation schemes that are robust to multipath fading. Orthogonal frequency division multiplexing (OFDM), because of the split of a high rate data stream into a parallel of lower rate data streams, provides the much-needed resistance to multipath fading and has attracted increasing interest in recent years. In Europe, OFDM was first standardized for digital audio broadcasting (DAB) in 1995 [1] and terrestrial digital video broadcasting in 1997 [2]. It has also been proposed for high data rate packet transmission, *a.k.a.*, wireless LAN, as standardized in IEEE 802.11a [3] and HIPERLAN/2 [4].

To maximize the performance advantage of OFDM system, coherent detection is desired, and in this case, reliable estimation of the time-dispersive channel is required. Training symbol-based OFDM systems usually require an extra $\sim 20\%$ bandwidth, thereby consuming too many precious resources. Blind channel estimation algorithms have the advantage of being bandwidth efficient as they do not require the transmission of training symbols. Many existing blind OFDM channel estimation methods are statistical in nature (e.g., second-order statistics based as in [5]–[10]) and usually require large number of data blocks. For example, in [5], [7], and [8], statistical subspace approaches were proposed for blind channel identification that utilize cyclostationarity inherent in the OFDM systems with cyclic prefix (CP). Clearly, these methods require that the channel be constant for a large number of blocks and, thus, have limited applicability in wireless channels involving high mobility when the channel may vary from block to block. Deterministic blind OFDM channel estimation, on the other hand, is usually more data efficient. For example, the finite alphabet property has been exploited in [11]–[13] for blind channel estimation. A decision-directed iterative algorithm was proposed in [11] for joint symbol and channel estimation. The performance, however, largely depends on the initial point and is subject to error propagation. The proposed identifiability also heavily hinges on the signal constellation. For example, for 16QAM, the subcarrier number should be at least 52 times the channel length; therefore, it has limited application in practice. In [12] and [13], the finite alphabet was explicitly exploited to obtain an estimate of the channel in the frequency domain. The channel identifiability of this approach is independent of the presence of channel nulls at subcarrier frequencies. For PSK modulation, this method can work with a single OFDM block. However, for the QAM modulation scheme, multiple data blocks are still needed along with some statistical assumptions on the input sequence. The major concern, however, is its computational complexity, which can be prohibitively high when the channel length is large. Even with the suboptimal phase directed (PD) algorithm, a blind initialization using the modified minimum distance (MMD) method can still be computationally expensive. Some heuristics have been proposed to relieve the computation burden [14].

Receiver diversity is another important resource that can be exploited in OFDM channel estimation. In [15] and [16], multiple receive antennas are used for channel estimation for OFDM systems *without* CP. The proposed algorithms are statistical subspace methods and require a large number of blocks. It is interesting to note that transmitter diversity has also been utilized for blind channel estimation of OFDM systems [17]. In this paper, we exploit receiver diversity for CP-based OFDM systems. With only two receive antennas, we demonstrate that a single OFDM data block is sufficient to obtain reasonable channel estimation for a wide range of SNR values. At low SNR, several OFDM blocks can be processed together to obtain the desired performance for a quasistationary channel. The number of blocks required, however, is considerably smaller than that needed by

Manuscript received May 20, 2002; revised February 24, 2003. This work was supported in part by the CASE Center of Syracuse University. This paper was presented in part at the Annual Conference on Information Sciences and Systems, Princeton, NJ, March 2002. The associate editor coordinating the review of this paper and approving it for publication was Prof. Nicholas Sidiropoulos.

The authors are with the Department of Electrical Engineering and Computer Science, Syracuse University, Syracuse, NY 13244 USA (e-mail: hwang08@ecs.syr.edu; ylin20@ecs.syr.edu; bichen@ecs.syr.edu).

Digital Object Identifier 10.1109/TSP.2003.816879

the statistical subspace methods. Channel identifiability conditions are developed that guarantee perfect channel retrieval in the absence of noise using only a single OFDM block. These conditions are derived for the general case when virtual carriers are present, which is compatible with most practical wireless OFDM systems. The Cramér–Rao lower bound (CRLB) is derived for performance evaluation, and it is found through numerical examples that the proposed algorithm is efficient for large SNR.

The proposed algorithm is intimately related to the cross relation (CR) method [18] for single carrier system with channel diversity. As pointed out in Section III-E, the new algorithm can be derived using the CR principle applied to circular convolution instead of linear convolution. However, because of the CP, the conversion from linear convolution to circular convolution has simplified the algorithm as well as the identifiablity conditions, especially with regard to the input symbols. Furthermore, the proposed algorithm, along with its identifiability conditions, can be easily adapted to the cases when virtual carriers are present.

The organization of the paper is as follows. In the next section, we introduce the signal model with receiver diversity. In Section III, we present a simple yet effective blind channel estimation algorithm using channel diversity. Section IV gives the identifiability conditions. Performance evaluation including both CRLB analysis and simulation results are contained in Section V, and we conclude in Section VI.

The following notation is frequently used in this paper. The DFT matrix $\mathbf{W}$ is defined as

$$\mathbf{W} = \frac{1}{\sqrt{N}} \begin{bmatrix} W_N^0 & W_N^0 & W_N^0 & \cdots & W_N^0 \\ W_N^0 & W_N^1 & W_N^2 & \cdots & W_N^{N-1} \\ W_N^0 & W_N^2 & W_N^4 & \cdots & W_N^{2(N-1)} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ W_N^0 & W_N^{(N-1)} & W_N^{2(N-1)} & \cdots & W_N^{(N-1)^2} \end{bmatrix} \tag{1}$$

where N is the number of subcarriers, and $W_N = e^{-j2\pi/N}$. The factor $1/\sqrt{N}$ is introduced to make $\mathbf{W}$ unitary for ease of notation. The normalized DFT matrix $\mathbf{W}$ can be partitioned as $\mathbf{W} = [\mathbf{W}_{L+1} \quad \mathbf{W}_{N-L-1}]$, where $L+1$ is the length of channel impulse response and is assumed known *a priori* in this paper, and $\mathbf{W}_{L+1}$ is the matrix composed of the first $L+1$ columns of $\mathbf{W}$. Further, we write

$$\mathbf{W}_{L+1} = \begin{bmatrix} \mathbf{u}_1^H \\ \vdots \\ \mathbf{u}_N^H \end{bmatrix} \tag{2}$$

where each $\mathbf{u}_k$ is an $L+1$ by 1 vector. We use bold face capital letters to denote matrices, whereas bold-face lowercase letters denote vectors. Finally, we denote circular convolution by $\otimes$ and linear convolution by $*$.

II. OFDM Signal Model with Receiver Diversity

In OFDM systems with N subcarriers, N information symbols are used to construct one OFDM block.[1] Specifically, each of the N symbols is used to modulate a subcarrier, and the N modulated subcarriers are added together to form an OFDM block. The orthogonality among subcarriers is achieved by carefully selecting the carrier frequencies such that each OFDM symbol interval contains an integer number of periods for all subcarriers. Guard time, which is cyclically extended to maintain the intercarrier orthogonality in the presence of the time-dispersive channel, is inserted and is assumed longer than the maximum delay spread of the channel to totally eliminate interblock interference [19]. The CP insertion effectively converts convolutional channel in the time domain into a multiplicative channel in the frequency domain, thus alleviating the need for channel equalization. This greatly simplifies the transceiver design [20].

The discrete-time complex baseband OFDM signal, including the CP part, is

$$s(n) = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} d_k e^{j2\pi kn/N}, n = -P, \ldots, -1, 0, \ldots, N-1 \tag{3}$$

where P is the length of the CP, and $d_{-k} = d_{N-k}$. In the presence of a time-dispersive channel and additive white Gaussian noise, assuming two receive antennas are employed, the discrete-time baseband signals at the two receivers, after timing and frequency synchronization, are

$$\begin{aligned} x_1(n) &= \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} H_1(k) d_k e^{j2\pi kn/N} + v_1(n) \\ x_2(n) &= \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} H_2(k) d_k e^{j2\pi kn/N} + v_2(n) \end{aligned}$$

for $n = 0, \ldots, N-1$, where $H_i(k)$ is the channel frequency response corresponding to the ith channel at subcarrier k, and $v_1(n)$ and $v_2(n)$ are both additive white complex Gaussian noise and are uncorrelated with each other. The CP part is truncated in the above equations because it contains a contribution from the previous OFDM block. The above signal model can be written in a compact matrix form as

$$\begin{aligned} \mathbf{x}_1 &= \mathbf{W}^H \mathbf{H}_1 \mathbf{d} + \mathbf{v}_1 \\ \mathbf{x}_2 &= \mathbf{W}^H \mathbf{H}_2 \mathbf{d} + \mathbf{v}_2 \end{aligned}$$

where $\mathbf{W}$ is the normalized DFT matrix as in (1), $\mathbf{H}_i = \text{diag}(\mathbf{h}_i)$ with

$$\mathbf{h}_i = [H_i(0), \ldots, H_i(N-1)]^T \tag{4}$$

that is, $\mathbf{H}_i$ is a diagonal matrix with diagonal element $H_i(k)$, and $\mathbf{d} = [d_0, \ldots, d_{N-1}]^T$ is the symbol vector. Taking the DFT at the receiver, we have the equivalent frequency domain observation

$$\begin{aligned} \mathbf{y}_1 &= \mathbf{W}\mathbf{x}_1 = \mathbf{H}_1 \mathbf{d} + \mathbf{z}_1 \\ \mathbf{y}_2 &= \mathbf{W}\mathbf{x}_2 = \mathbf{H}_2 \mathbf{d} + \mathbf{z}_2 \end{aligned}$$

where $\mathbf{z}_1$ and $\mathbf{z}_2$ are statistically identical to $\mathbf{v}_1$ and $\mathbf{v}_2$ because of the unitary property of $\mathbf{W}$, i.e., $\mathbf{z}_1$ and $\mathbf{z}_2$ are both white complex Gaussian and are uncorrelated with each other. Blind channel estimation aims to retrieve both $\mathbf{H}_1$ and $\mathbf{H}_2$ without any knowledge about $\mathbf{d}$. Clearly, without extra information, a direct approach to estimate the frequency response matrix $\mathbf{H}_i$ is not feasible—the number of unknowns ($6N$ from the three

[1]Extension to virtual carrier present systems will be addressed in Section III-C.

$N \times 1$ complex vectors $\mathbf{h}_1$, $\mathbf{h}_2$, and $\mathbf{d}$) exceeds the number of observations ($4N$ from the two observation vectors $\mathbf{y}_1$ and $\mathbf{y}_2$). However, we note that the actual degrees of freedom associated with $\mathbf{H}_1$ and $\mathbf{H}_2$ are far smaller than N. This is because the frequency response is simply the DFT of the channel impulse response that is usually assumed to be shorter than the cyclic prefix and, hence, is far smaller than N. With this observation, we can rewrite the signal model as

$$\begin{aligned}\mathbf{y}_1 =& \mathbf{D}\mathbf{h}_1 + \mathbf{z}_1 = \mathbf{D}\mathbf{W}_{L+1}\mathbf{g}_1 + \mathbf{z}_1\\ \mathbf{y}_2 =& \mathbf{D}\mathbf{h}_2 + \mathbf{z}_2 = \mathbf{D}\mathbf{W}_{L+1}\mathbf{g}_2 + \mathbf{z}_2\end{aligned} \tag{5}$$

where $\mathbf{D} = \text{diag}(\mathbf{d})$, i.e., it is a diagonal matrix with diagonal element d_k, $\mathbf{h}_i$ is as in (4), $\mathbf{W}_{L+1}$ is the partition of DFT matrix as in (2), and $\mathbf{g}_i$ is the impulse response for the ith channel; hence, we have the relation $\mathbf{h}_i = \mathbf{W}_{L+1}\mathbf{g}_i$. Here, we have a total of $N + 2(L+1)$ unknown complex variables with $2N$ complex observations. For most practical mobile channels. the number of observations is always greater than the number of unknowns. It is interesting to note that due to the inherent scalar ambiguity, channel identifiability does not require that the number of observations is always greater than that of the unknowns. This will be elaborated on further in Section IV. We remark again that in this paper, we deal exclusively with the case of known channel length.

III. Blind Channel Estimation

A. Noiseless Case

Given the signal model in (5), we first consider the channel estimation in the noiseless case. We emphasize again that by converting the channel estimation from frequency domain to time domain, we have reduced the degrees of freedom [21] of the estimation problem. Using the above model, we now devise a simple algorithm that can perfectly retrieve the time domain channel impulse response in the absence of noise.

Without the channel noise, (5) can be written as, in an element-by-element fashion

$$\begin{aligned}y_1(k) =& d_k \cdot \mathbf{u}_k^H \mathbf{g}_1\\ y_2(k) =& d_k \cdot \mathbf{u}_k^H \mathbf{g}_2\end{aligned}$$

where $\mathbf{u}_k^H$ is given in (2). Therefore, for $d_k \neq 0$

$$y_1(k)\mathbf{u}_k^H\mathbf{g}_2 = y_2(k)\mathbf{u}_k^H\mathbf{g}_1.$$

The matrix form of the above equation is

$$\mathbf{Y}_1\mathbf{W}_{L+1}\mathbf{g}_2 = \mathbf{Y}_2\mathbf{W}_{L+1}\mathbf{g}_1$$

where $\mathbf{Y}_1 = \text{diag}(\mathbf{y}_1)$ and $\mathbf{Y}_2 = \text{diag}(\mathbf{y}_2)$ are diagonal matrices constructed using $\mathbf{y}_1$ and $\mathbf{y}_2$, respectively. Equivalently, we have

$$[\,\mathbf{Y}_2\mathbf{W}_{L+1} \quad -\mathbf{Y}_1\mathbf{W}_{L+1}\,]\begin{bmatrix}\mathbf{g}_1\\ \mathbf{g}_2\end{bmatrix} \triangleq \mathbf{V}\mathbf{g} = 0 \tag{6}$$

where $\mathbf{V} = [\,\mathbf{Y}_2\mathbf{W}_{L+1} \quad -\mathbf{Y}_1\mathbf{W}_{L+1}\,]$, and $\mathbf{g} = [\mathbf{g}_1, \mathbf{g}_2]^T$. Therefore, in the noiseless case, the channel can be retrieved up to a scalar ambiguity by simply finding a solution to the above homogeneous equations. We will address the existence and uniqueness of the solution to the above equation (*a.k.a.*, identifiability issue) in the next section. We first discuss below the implementation of the algorithm in the noisy case as well as its extension to the virtual carrier present case.

B. Noisy Case

In the presence of channel noise, it is clear that (6) will not hold. Instead of finding an exact solution, we may seek to minimize the quadratic form:

$$\min_{\mathbf{g}} \quad \mathbf{g}^H\mathbf{U}\mathbf{g} = \mathbf{g}^H\mathbf{V}^H\mathbf{V}\mathbf{g} \quad s.t. \quad |\mathbf{g}| = 1$$

where $\mathbf{U} = \mathbf{V}^H\mathbf{V}$ is a $2(L+1) \times 2(L+1)$ Hermitian matrix and is positive definite with probability one in the presence of channel noise. The constraint on $\mathbf{g}$ is to avoid the trivial solution $\mathbf{g} \equiv 0$. The solution to the above constrained minimization problem is the eigenvector corresponding to the smallest eigenvalue of $\mathbf{U}$. Equivalently, we can find the right singular vector corresponding to the smallest singular value of the matrix $\mathbf{V}$.

Next, we present the extension to OFDM systems involving virtual carriers.

C. Virtual Carrier Present Case

In most wireless OFDM systems, virtual carriers are inserted for anti-aliasing after the D/A converter at the transmitter. For example, in both IEEE802.11a and HIPERLAN/2, 12 out of 64 carriers are virtual carriers [3], [4]. The proposed algorithm can be easily modified to adapt to cases when virtual carriers are present.

Let us assume $M \leq N$, and, without loss of generality, the last $N - M$ carriers are virtual carriers, that is, these $N - M$ carriers are inactive in a sense that they are not used to modulate information symbols. Implemented in baseband using DFT, this is analogous to zero padding for DFT. The input time domain signal, corresponding to (3) but with virtual carriers, is

$$s(n) = \frac{1}{\sqrt{N}}\sum_{k=0}^{M-1} d_k e^{j2\pi kn/N}.$$

Thus, $d_k = 0$ for $k \geq M$. Then, the received signal can be easily modified as

$$\begin{aligned}\bar{\mathbf{y}}_1 =& \overline{\mathbf{D}}\,\overline{\mathbf{W}}_{L+1}\mathbf{g}_1 + \mathbf{z}_1\\ \bar{\mathbf{y}}_2 =& \overline{\mathbf{D}}\,\overline{\mathbf{W}}_{L+1}\mathbf{g}_2 + \mathbf{z}_2\end{aligned} \tag{7}$$

where $\bar{\mathbf{y}}_i = [y_i(0), \ldots, y_i(M-1)]^T$, $\overline{\mathbf{Y}}_i = diag(\bar{\mathbf{y}}_i)$ for $i = 1, 2$, $\overline{\mathbf{D}} = diag(d_0, \ldots, d_{M-1})^T$, and $\overline{\mathbf{W}}_{L+1}$ is the first M rows of $\mathbf{W}_{L+1}$. The channel estimate can be obtained by taking SVD on $\overline{\mathbf{V}}$, where this new $\overline{\mathbf{V}}$ matrix is constructed using the newly defined matrices $\overline{\mathbf{Y}}_i$ and $\overline{\mathbf{W}}_{L+1}$ in a fashion similarly to that of (6),

Notice that the number of virtual carriers is considered known *a priori*. If the actual number of virtual carriers is larger than that is assumed in the algorithm, the proposed method will not work because the $\overline{\mathbf{V}}$ matrix in (6) is more than rank-one deficient in the noiseless case.

D. Multiple OFDM Data Blocks

If the channel is quasistationary and, therefore, can be assumed constant during several OFDM blocks, channel estimation can be improved by utilizing multiple OFDM blocks. Assuming that K blocks are used for channel estimation, it is straightforward to extend the algorithm to the following minimization problem:

$$\min_{\mathbf{g}} \mathbf{g}^H \left[\sum_{k=1}^{K} \mathbf{V}_k^H \mathbf{V}_k\right] \mathbf{g} \quad s.t. \quad |\mathbf{g}| = 1$$

where $\mathbf{V}_k$ is constructed for each OFDM block. Notice that this extension does not have any substantial increase in complexity—only one eigendecomposition is required no matter how many blocks are used.

E. Connection With the Cross Relation Approach

In this section, we investigate the connection between the proposed method and the CR approach in [18] that was proposed for deterministic blind channel estimation for a single carrier system with receiver diversity. Assuming that there is no virtual carrier, the new method amounts to solving (6), which we repeat here for convenience:

$$[\mathbf{Y}_2\mathbf{W}_{L+1} \quad -\mathbf{Y}_1\mathbf{W}_{L+1}] \begin{bmatrix} \mathbf{g}_1 \\ \mathbf{g}_2 \end{bmatrix} \triangleq \mathbf{V}\mathbf{g} = 0.$$

Note $\mathbf{Y}_i = \text{diag}(\mathbf{W}\mathbf{x}_i), i = 1, 2$ and $\mathbf{x}_i = [x_i(0), x_i(1), \ldots, x_i(N-1)]^T$. Then, we have the following [22]:

$$\tilde{\mathbf{X}}_i = \mathbf{W}^{-1}\mathbf{Y}_i\mathbf{W}$$

where $\tilde{\mathbf{X}}_i$ is a circulant matrix constructed with $\mathbf{x}_i$, i.e.,

$$\tilde{\mathbf{X}}_i = \begin{bmatrix} x_i(0) & x_i(N-1) & \cdots & x_i(1) \\ x_i(1) & x_i(0) & \cdots & x_i(2) \\ \vdots & \vdots & \vdots & \vdots \\ x_i(N-1) & x_i(N-2) & \vdots & x_i(0) \end{bmatrix} \quad i = 1, 2.$$

Clearly

$$\mathbf{Y}_i\mathbf{W} = \mathbf{W}\tilde{\mathbf{X}}_i.$$

Using matrix partitioning, we get $\mathbf{Y}_i\mathbf{W}_{L+1} = \mathbf{W}\mathbf{X}_i$, where $\mathbf{X}_i$ is the first $L+1$ columns of $\tilde{\mathbf{X}}_i$, i.e.,

$$\mathbf{X}_i = \begin{bmatrix} x_i(0) & x_i(N-1) & \cdots & x_i(N-L-1) \\ x_i(1) & x_i(0) & \cdots & x_i(N-L) \\ \vdots & \vdots & \vdots & \vdots \\ x_i(N-1) & x_i(N-2) & \vdots & x_i(N-L) \end{bmatrix}.$$

Therefore

$$[\mathbf{Y}_2\mathbf{W}_{L+1} \quad -\mathbf{Y}_1\mathbf{W}_{L+1}] \begin{bmatrix} \mathbf{g}_1 \\ \mathbf{g}_2 \end{bmatrix} = 0 \overset{(a)}{\Longrightarrow}$$
$$[\mathbf{W}\mathbf{X}_2 \quad -\mathbf{W}\mathbf{X}_1] \begin{bmatrix} \mathbf{g}_1 \\ \mathbf{g}_2 \end{bmatrix} = 0 \overset{(b)}{\Longrightarrow}$$
$$[\mathbf{X}_2 \quad -\mathbf{X}_1] \begin{bmatrix} \mathbf{g}_1 \\ \mathbf{g}_2 \end{bmatrix} = 0$$

which is

$$\mathbf{x}_1 \otimes \mathbf{g}_2 = \mathbf{x}_2 \otimes \mathbf{g}_1 \tag{8}$$

where $\otimes$ denotes circular convolution. This is closely related to the CR method in [18], where linear convolution, instead of circular convolution, is used, i.e.,

$$\mathbf{x}_1 * \mathbf{g}_2 = \mathbf{x}_2 * \mathbf{g}_1.$$

Notice that steps a) and b) are all reversible, and therefore, the proposed method can be considered as an application of the CR principle to circular convolution. Specifically, since $\mathbf{x}_1 = \mathbf{d} \otimes \mathbf{g}_1$ and $\mathbf{x}_2 = \mathbf{d} \otimes \mathbf{g}_2$, we have

$$\mathbf{d} \otimes \mathbf{g}_1 \otimes \mathbf{g}_2 = \mathbf{d} \otimes \mathbf{g}_2 \otimes \mathbf{g}_1$$

which yields (8). The circular convolution is a direct result of the insertion of the CP. We also note that in the presence of virtual carriers, this direct application of the CR principle does not work—the constructed $\mathbf{V}$ matrix is more than rank one deficient. In this case, direct frequency domain approach as in Section III-C should be used.

IV. Identifiability

The channels are said to be identifiable if, in the absence of noise, there is a unique solution (up to a scalar ambiguity) that satisfies the signal model (5) or (7). In this section, we discuss the identifiability issues based on OFDM system with virtual carriers therefore we will use signal model (7). Since OFDM system without virtual carriers is just a special case with the number of virtual carriers being 0, all results derived in this section automatically apply directly to cases without virtual carriers.

In the following, we give a sufficient **and** a necessary condition (though *not sufficient and necessary* condition) for channel identifiability using receiver diversity.

Theorem 1 (Sufficient Condition): The channel impulse responses $\mathbf{g}_1$ and $\mathbf{g}_2$ can be identified up to a scalar factor if the following conditions hold:

1) $H_1(z)$ and $H_2(z)$ do not share common zeros.
2) $M \geq 2L + 1$.

The proof is lengthy, and we defer it to Appendix A.

Theorem 2 (Necessary Condition): If the channel impulse responses $\mathbf{g}_1$ and $\mathbf{g}_2$ are identifiable up to a scalar factor, then $M \geq 2L + 1$.

Proof: If the system is identifiable, there will be a unique (up to a scalar ambiguity) solution $\mathbf{g}_1$ and $\mathbf{g}_2$ for (6). Therefore, the rank of $\mathbf{V}$ must be $2L + 1$, i.e., its null space must have a dimension equal to 1. Since $\mathbf{V}$ is an M by $2(L+1)$ matrix, we must have $M \geq 2L + 1$. Q.E.D.

A. Discussion

Condition $M \geq 2L + 1$ guarantees that we have enough observations in order for (7) to be solvable. Notice in the boundary situation that when $M = 2L + 1$, we actually have $2M$ complex observations, which is less than the number of unknown complex parameters $M + 2(L+1) = 2M + 1$. In this case, we can still estimate the channels because the inherent scalar ambiguity relieves us from identifying the exact channel responses. Since in most practical wireless channels, M, which is the total number of nonvirtual carriers, is far greater than the length of channel impulse response $L + 1$, it is fair enough to consider that this necessary condition and the second part of the sufficient condition are automatically satisfied.

Another observation is that the sufficient condition is reminiscent of that of the identifiability condition with receiver diversity for single carrier systems (see, e.g., [18] and [23]). The linear complexity of the input sequence, which is defined as the number of modes of the polynomial constructed from the input sequence, is now translated into the number of nonvirtual carriers in the frequency domain. This equivalence is indeed a direct consequence of the CP insertion (otherwise, infinite length input is required). Since different modes correspond to different subcarrier frequencies, the sufficient condition is also simplified in the frequency domain as we only require those d_k's to be nonzero. The necessary condition, however, appears to be stronger than that obtained for single carrier case. The stronger condition is easily obtained in the frequency domain, again due to the CP insertion that translates time domain convolution into the frequency domain multiplication.

V. Performance Evaluation

In this section, performance evaluation of the proposed algorithm is carried out both analytically (using CRLB) and numerically. We start from analysis by deriving the CRLB and compare it with the mean squared error (MSE).

A. Cramér–Rao Lower Bound

In this section, we evaluate the performance of the blind estimation algorithm by deriving the CRLB. The unknown parameter vector is

$$\theta = [\mathrm{Re}(\mathbf{g}_1), \mathrm{Re}(\mathbf{g}_2), \mathrm{Re}(\mathbf{d}), \mathrm{Im}(\mathbf{g}_1), \mathrm{Im}(\mathbf{g}_2), \mathrm{Im}(\mathbf{d})]^T .$$

Based on (5) and given that $\mathbf{z}_1$ and $\mathbf{z}_1$ are uncorrelated with each other, the negative log likelihood function can be obtained as, after discarding any irrelevant constant

$$-\ln\Lambda = (\mathbf{y}_1 - \mathbf{D}\mathbf{W}_{L+1}\mathbf{g}_1)^H (\mathbf{y}_1 - \mathbf{D}\mathbf{W}_{L+1}\mathbf{g}_1) + (\mathbf{y}_2 - \mathbf{D}\mathbf{W}_{L+1}\mathbf{g}_2)^H (\mathbf{y}_2 - \mathbf{D}\mathbf{W}_{L+1}\mathbf{g}_2).$$

Under the assumption of complex Gaussian noise, the Fisher information matrix (FIM) can be derived as [24]–[26]

$$\mathbf{F} = 2\begin{bmatrix} \mathrm{Re}(\mathbf{F_c}) & -\mathrm{Im}(\mathbf{F_c}) \\ \mathrm{Im}(\mathbf{F_c}) & \mathrm{Re}(\mathbf{F_c}) \end{bmatrix} \quad (9)$$

where

$$\mathbf{F}_c = \frac{1}{\sigma^2}\begin{bmatrix} \mathbf{Q}^H\mathbf{Q} & 0 & \mathbf{Q}^H\mathbf{H}_1 \\ 0 & \mathbf{Q}^H\mathbf{Q} & \mathbf{Q}^H\mathbf{H}_2 \\ \mathbf{H}_1^H\mathbf{Q} & \mathbf{H}_2^H\mathbf{Q} & \mathbf{H}_1^H\mathbf{H}_1 + \mathbf{H}_2^H\mathbf{H}_2 \end{bmatrix}$$

and σ^2 is the noise power, and $\mathbf{Q} = \mathbf{D}\mathbf{W}_{L+1}$. A detailed derivation of the $\mathbf{F}_c$ matrix is provided in Appendix B. Note that matrix $\mathbf{F_c}$ is at least rank 1 deficient due to the scalar ambiguity of the channel. To evaluate the CRLB, we often consider one element of the channel (e.g., the first element of $\mathbf{g_1}$), as known. After deleting the column and row associated with the known parameter, the remaining matrix will be full rank, and the CRLB can be evaluated by inverting this full rank matrix [27].

It is interesting to consider the situation when $\mathbf{g}_1$ and $\mathbf{g}_2$ share a common zero at a subcarrier frequency, say, k_0. In this case, matrix $\mathbf{F_c}$ will have an all-zero row and an all-zero column at the corresponding input symbol location, i.e., the row and column corresponding to d_{k_0}. Therefore, even if we assume $g_1(0)$ is known and its corresponding row and column is deleted from the FIM, the remaining FIM is still not full rank. One explanation could be that because of the common zero at a subcarrier frequency, the corresponding symbol d_{k_0} is clearly not identifiable. However, it is found numerically that after getting rid of the row and column corresponding to d_{k_0}, the remaining FIM is still rank deficient—which implies that the channel itself may not be identifiable. This observation suggests that a possible necessary condition for channel identifiability is that $\mathbf{g}_1$ and $\mathbf{g}_2$ do not share common zeros at subcarrier frequencies. Notice that this condition is weaker than that stated in the sufficient condition where $\mathbf{g}_1$ and $\mathbf{g}_2$ do not share any common zeros without regard to their possible locations.

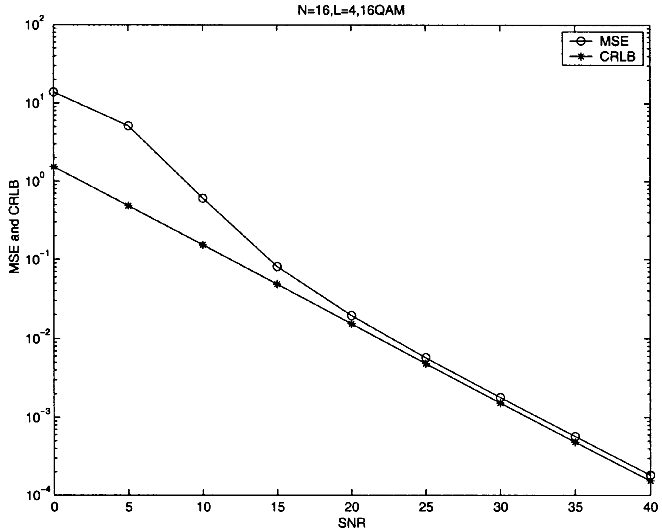

Fig. 1. Example of MSE for the blind OFDM channel estimation using diversity scheme and the corresponding CRLB for the channel pair specified in (10). The MSE and CRLB are computed for $\mathbf{g}_2$ under the assumption that $g_1(0)$ is known.

We compare the MSE of the proposed algorithm to the corresponding CRLB using a single OFDM block. The following channel pair is generated using Rayleigh fading model for each tap.

$$\begin{aligned}\mathbf{g}_1 =& [0.158 - 0.664i, -0.198 + 0.267i, -0.325 + 0.197i \\ & - 0.378 - 0.245i, . - 0.278 - 0.003i]^T \\ \mathbf{g}_2 =& [-0.531 + 0.189i, -0.409 + 0.176i \\ & - 0.035 - 0.313i, 0.147 + 0.218i, 0.557 + 0.076i]^T.\end{aligned} \quad (10)$$

The scalar ambiguity is removed by assuming that $g_1(0)$ is known, and we compute the MSE and CRLB for $\mathbf{g}_2$. Correspondingly, the CRLB is calculated by removing the row and column of the FIM corresponding to that of $g_1(0)$. The results are given in Fig. 1. We also plot the bias of the estimator in Fig. 2, where the bias is defined as $\sum_k |g_2(k) - \hat{g}_2(k)|$, where $\hat{\mathbf{g}}_2$ is the estimated channel. Both the MSE and bias are obtained using 800 Monte Carlo runs. Clearly, the proposed algorithm is close to efficient and unbiased for large SNR.

Finally, we discuss the effect of virtual carriers on the performance of the estimator and, in particular, on the CRLB. It is straightforward to verify that the CRLB derived here applies to the case with virtual carriers except that $\mathbf{D}$ and $\mathbf{W}_{L+1}$ are replaced with $\overline{\mathbf{D}}$ and $\overline{\mathbf{W}}_{L+1}$, as defined in Section III-C. The pres-

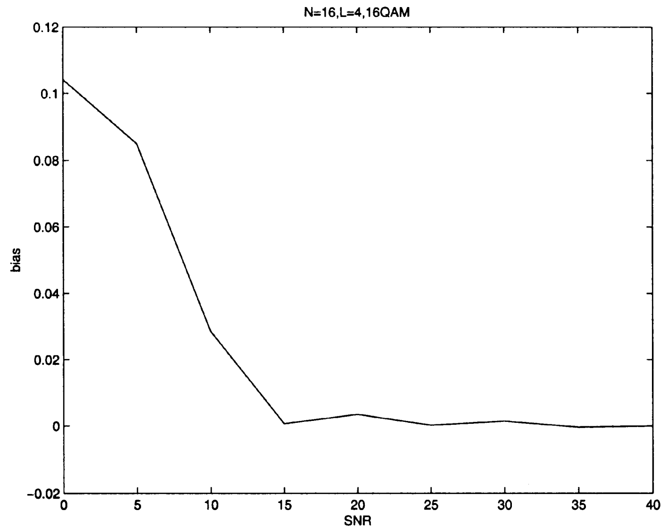

Fig. 2. Example of bias for the blind OFDM channel estimation using diversity scheme for the channel pair specified in (10). The bias is computed for $\mathbf{g}_2$ under the assumption that $g_1(0)$ is known.

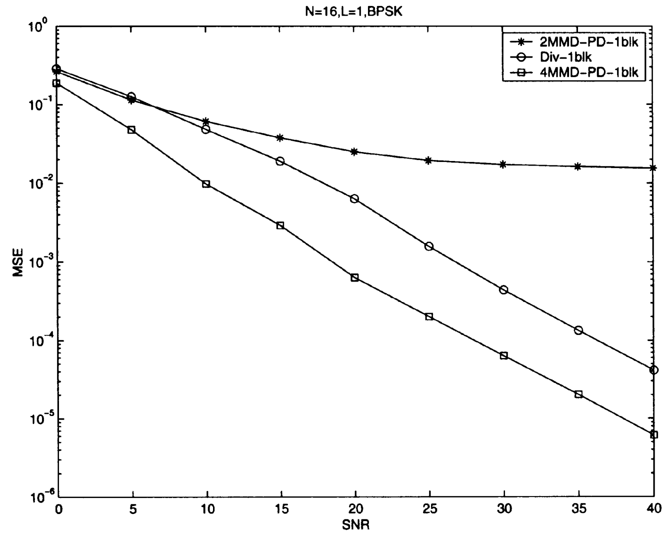

Fig. 3. MSE for the blind OFDM channel estimation using diversity and the finite-alphabet method for BPSK. The MSE are computed by averaging over 200 randomly generated channels and using 800 Monte Carlo runs. Four iterations were performed for the PD algorithm. Both methods use one OFDM data block.

ence of virtual carriers will affect the performance of the proposed algorithm. Numerical results indicate that as the number of virtual carriers increases (yet the identifiability conditions are still met), the CRLB, as well as the MSE of the channel estimate, becomes larger. This can be explained as follows. For the receiver diversity scheme, if the number of virtual carriers increases by one (say, with index k), the number of unknown parameters to be estimated decreases by one (corresponding to d_k), whereas the number of observations decreases by two ($y_1(k)$ and $y_2(k)$).

B. Numerical Results

In this section, we compare the MSE performance of the proposed method with the finite-alphabet method in [12] and [13] and the subspace method in [7] and [8]. Throughout this section, we choose the number of carriers $N = 16$ and assume that no virtual carrier is present. The CP length is chosen to be 5. The scalar ambiguity is eliminated as follows. Assume that the channel estimate for $\mathbf{g}_1$ is $\tilde{\mathbf{g}}_1$. Defining $\alpha = \tilde{\mathbf{g}}_1^H \mathbf{g}_1/|\tilde{\mathbf{g}}_1|^2$, we use $\hat{\mathbf{g}}_1 = \alpha\tilde{\mathbf{g}}_1$ as the final channel estimate. The reason is that in the simulation, we average the performance over 200 randomly generated channels. The approach of assuming a known channel coefficient may occasionally lead to unreasonable results for *any blind channel estimation algorithm* if the coefficient happens to be of a very small magnitude. The MSE is defined using

$$\text{MSE} = \frac{1}{M_c}\sum_{m=1}^{M_c} ||\hat{\mathbf{g}} - \mathbf{g}||^2.$$

Diversity Method versus Finite-Alphabet Method: We first compare the diversity method with the finite-alphabet method proposed in [12] and [13]. The finite-alphabet method first obtains an estimate of $H^J(k)$. Then, it tries to resolve the phase ambiguity to determine $H(k)$ from $H^J(k)$. Here, k is the subcarrier index, and J is determined by symbol constellation. For example, $J = 4$ for QAM modulation, and J equals the constellation size for any PSK modulation. The optimum minimum distance (MD) algorithm requires a search of J^N possible channels, which is computationally prohibitive. Instead, initialization using the MMD algorithm followed by the PD algorithm can be used [13]. We use a two-ray channel (i.e., $L = 1$) for all comparisons between the diversity scheme and the finite-alphabet method. We remark here that if L becomes large, even the suboptimal MMD-PD becomes computationally formidable. In this case, more heuristics need to be resorted to relieve the computational burden [14].

In Fig. 3, we compare the diversity method with the finite-alphabet method for BPSK modulation scheme, both using a single OFDM data block. The PSK modulation scheme enables the finite-alphabet method to be implemented using only one OFDM block [12], [13]. The MMD algorithm is used for initial channel estimation followed by the PD algorithm to perform a recursive search. In the legend of the figure, "$\overline{N}$-MMD" stands for $\overline{N}$-point MMD initialization, with a complexity of $J^{\overline{N}}$. We see that 4MMD initialization followed by the PD search outperforms the diversity method.

We also compare the diversity method with the finite-alphabet method for the QPSK modulation scheme, both using one OFDM block. From Fig. 4, it is clear that 4-MMD does not give a good MSE performance. Numerical analysis found out that of all the 200 channels being averaged, there are considerable cases when the 4-MMD initialization does not give an accurate enough estimate and eventually causes the PD algorithm to diverge. For QPSK, 6-MMD can yield good initial channel estimate to ensure the convergence of the PD algorithm, but the complexity goes up significantly. The diversity scheme, however, requires only a single SVD independent of the channel length or signal constellation. Further, while the diversity method has a similar MSE performance as the finite-alphabet method for low SNR, it outperforms the latter in the high SNR region.

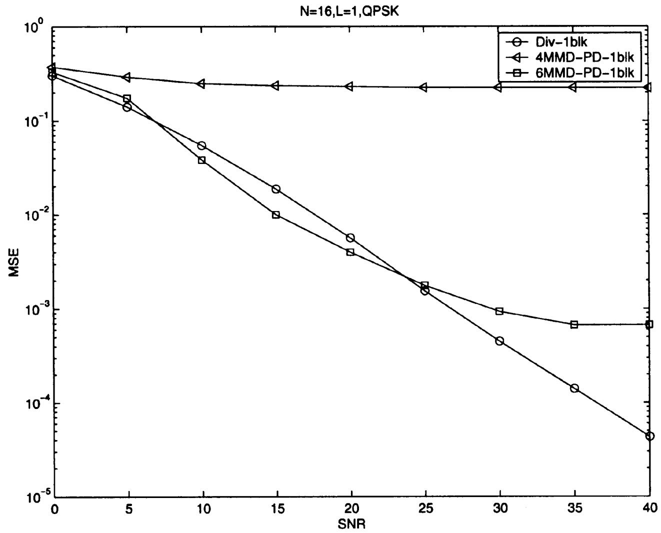

Fig. 4. MSE for the blind OFDM channel estimation using diversity and finite-alphabet method for QPSK. The MSE are computed by averaging 200 randomly generated channels and using 800 Monte Carlo runs. Four iterations were performed for the PD algorithm. Both methods use one OFDM data block.

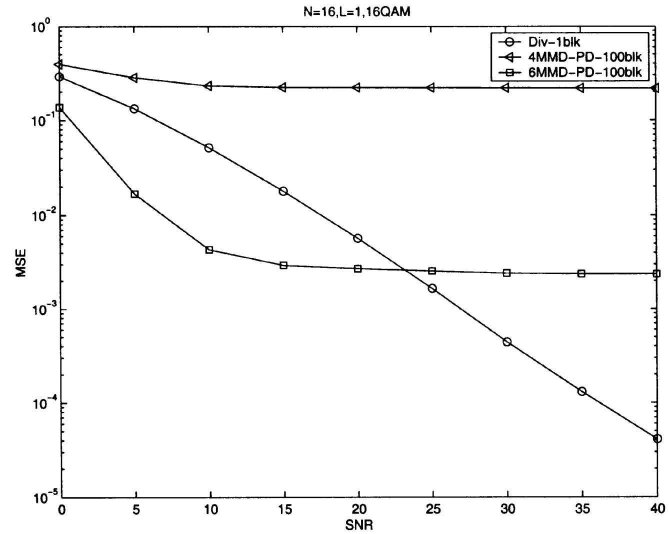

Fig. 5. MSE for the blind OFDM channel estimation using diversity and the finite-alphabet method for 16QAM. The MSE is computed by averaging over 200 randomly generated channels and using 800 Monte Carlo runs. Four iterations were performed for the PD algorithm. We use one OFDM block for the proposed method and 100 blocks for the finite-alphabet algorithm.

The performance degradation of the finite alphabet algorithm can be explained as follows. The finite alphabet algorithm needs to resolve the phase ambiguity—the direct available estimate is $H^J(k)$, from which we need to get back to $H(k)$ (or its corresponding time domain channel impulse response). The degrees of freedom associated with the above problem are $JL+1$ as the frequency domain multiplication is equivalent to time domain convolution. For BPSK, $J=2$, which is relatively small, and thus, 4-MMD can give reasonably good estimate for a channel with $L=1$ to allow the PD to converge. For large signal constellation, $J \geq 4$; therefore, significantly higher dimensional MMD is needed to get a reasonable initial point. For example, QPSK constellation with $L=1$ requires $\overline{N} \geq 5$, which explains why 4-MMD does not work well in the previous example. The problem could be even more prominent for large-dimension PSK modulation as J is identical to the size of the constellation for PSK modulation.

In Fig. 5, we compare the two methods for the 16QAM modulation scheme. We use one OFDM block for the diversity method and 100 blocks for the finite-alphabet method. Notice that for QAM modulation, J is fixed at 4 and is independent of the modulation size. However, it requires a large number of OFDM blocks to estimate $H^J(k)$ as no deterministic results exist for d^J with QAM modulation. Similar initialization issues exist for 16QAM modulation. From Fig. 5, we can see that a 100-block finite alphabet method using 6-MMD initialization has a better performance than diversity using a single block in a low SNR region, yet it exhibits an error-floor effect in the high SNR region.

For channels with $L \geq 2$, the complexity of the MMD will go even higher in order to obtain a reasonable initial estimate for the PD algorithm to converge. The heuristic approach in [14] could be applied to reduce the complexity of the finite-alphabet-based method. Notice that the proposed diversity scheme can also be used to initialize the finite-alphabet algorithm.

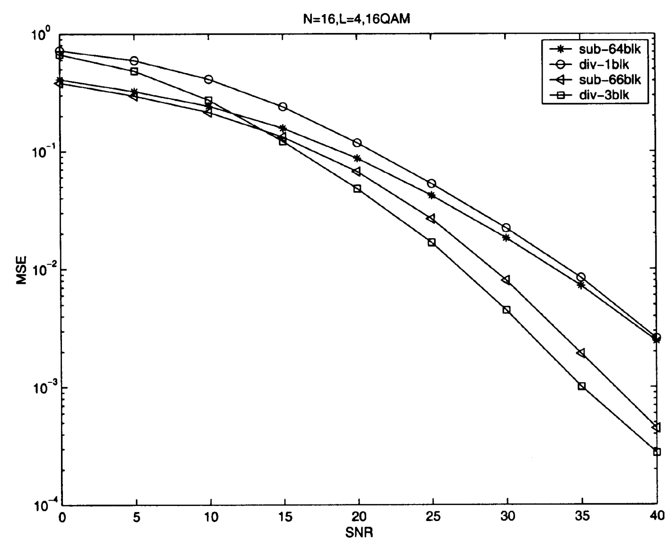

Fig. 6. MSE for the blind OFDM channel estimation using diversity and the subspace method for 16QAM. The MSE are computed by averaging over 200 randomly generated channels and using 800 Monte Carlo runs.

In this paper, we consider the scalar ambiguity known to the receiver. In practice, scalar ambiguity is resolved through either training symbols or pilot tones, which will introduce extra error for channel estimation [28]. The finite-alphabet property of the input symbols can be utilized to alleviate this effect. The computational complexity of the finite-alphabet method can also be reduced with the aid of training symbols or pilot tones.

Diversity Method versus Subspace Method: In Fig. 6, we compare the diversity method with the subspace method proposed in [7] and [8] using the 16QAM modulation scheme. From the plot, it is clear that the diversity method with only a few OFDM blocks can achieve equivalent or better performance than the subspace method using more than 60 blocks,

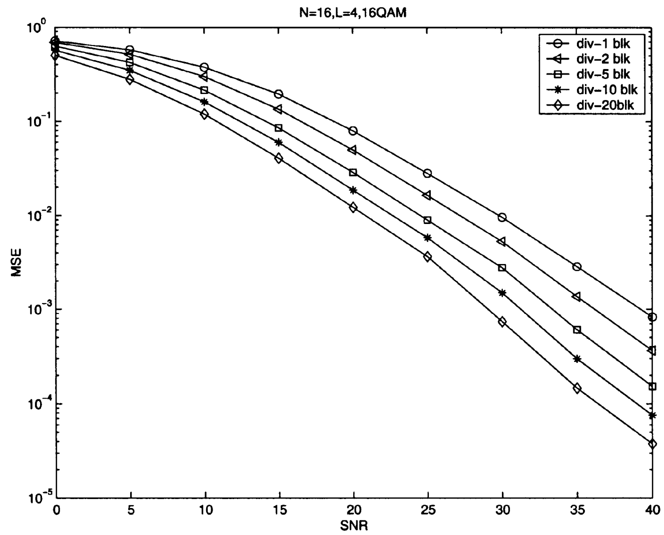

Fig. 7. MSE for the blind OFDM channel estimation using diversity scheme with 1, 2, 5, 10, and 20 blocks of data for 16QAM. The MSE are computed by averaging over 200 randomly generated channels and using 800 Monte Carlo runs.

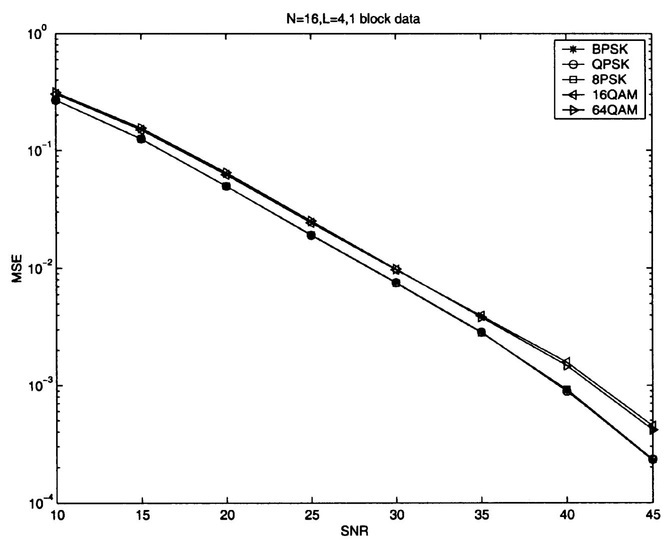

Fig. 8. MSE for the blind OFDM channel estimation using diversity with different modulation schemes. The MSE are computed by averaging over 200 randomly generated channels and using 800 Monte Carlo runs.

making it more appealing for high mobility applications. We remark here that through extensive numerical simulations, similar improvement can be observed for different modulation schemes and channel lengths.

Performance of the Diversity Method With Multiple Blocks: When the channel is assumed constant over a period of multiple OFDM blocks, the performance of the diversity method can be further improved by utilizing more data blocks. This is observed in Fig. 7. Note that the improvement does not grow linearly with the number of data blocks and will eventually saturate. Since a few blocks of data is enough for the diversity method to give a reasonably accurate channel estimate, users can choose a "cutoff" point in terms of the number of OFDM blocks, balancing performance and data efficiency.

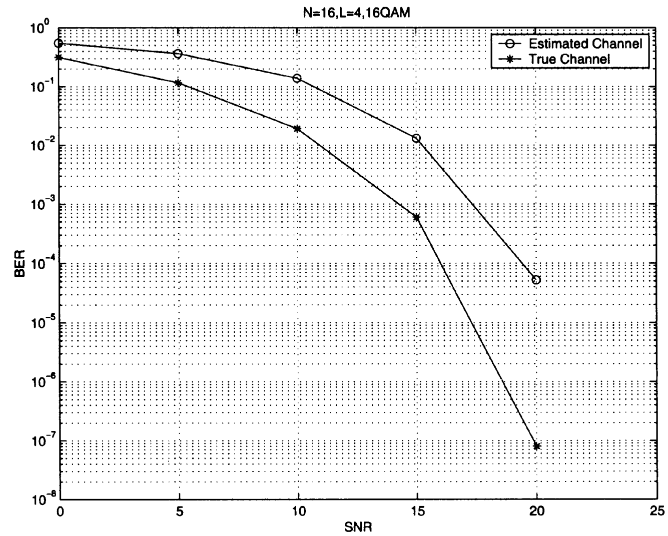

Fig. 9. Example of BER with estimated channel using diversity scheme and the true channel. BER are calculated using 12.5 million OFDM blocks (200 million bits).

Robustness to Different Modulation Schemes: Fig. 8 shows the performance of the diversity method for different modulation schemes. Clearly, the diversity method is fairly insensitive to the constellation of the input symbols.

BER of the Diversity Method: Fig. 9 shows the BER using both estimated channel and the true channel. A channel pair is randomly generated and fixed throughout the simulation. While it is expected that the true channel gives better BER performance, the SNR loss using the estimated channel becomes smaller as the channel SNR increases. This is due to the improved channel estimation performance at high SNR.

VI. CONCLUSIONS

In this paper, a data-efficient blind channel identification algorithm utilizing receiver diversity is proposed. The proposed algorithm utilizes frequency domain observation to estimate the channel impulse response and is intimately related to the CR method for blind channel estimation in single carrier systems. In the noiseless case, the algorithm can perfectly retrieve the channels up to a scalar factor. In the presence of noise, the algorithm has very low complexity—only a single eigenvalue decomposition is needed, no matter how many OFDM blocks are used. Some identifiability results are obtained, and the effect of channel nulls at subcarrier frequencies is also discussed. One of the most desirable merits of this new approach is its data efficiency—it can be implemented using a single data block to achieve reasonably accurate estimate, and this property is independent of the input signal constellation. Furthermore, the proposed algorithm imposes no restriction on the input symbol, i.e., no statistical assumption regarding the input symbols are needed. The Cramér–Rao lower bound of the channel estimate based on the diversity model is also derived, along with numerical simulation to evaluate the performance of the proposed algorithm.

Notice that in this work, we assumed that the channel length is perfectly known, as is with most blind methods. Extension of the diversity-based channel estimation with unknown channel length is under investigation and will be reported elsewhere.

APPENDIX A
PROOF OF THE SUFFICIENT CONDITION FOR IDENTIFIABILITY

In the noiseless case, model (7) yields

$$y_1(k) = d_k \cdot \mathbf{u}_k^H \mathbf{g}_1$$
$$y_2(k) = d_k \cdot \mathbf{u}_k^H \mathbf{g}_2.$$

Assuming we have another set of parameters $(\tilde{d}_k, \tilde{\mathbf{g}}_1, \tilde{\mathbf{g}}_2)$ that also satisfy the same system model, we then have

$$d_k \cdot \mathbf{u}_k^H \mathbf{g}_1 = \tilde{d}_k \cdot \mathbf{u}_k^H \tilde{\mathbf{g}}_1$$
$$d_k \cdot \mathbf{u}_k^H \mathbf{g}_2 = \tilde{d}_k \cdot \mathbf{u}_k^H \tilde{\mathbf{g}}_2. \quad (11)$$

From this, we get, through cross multiplication

$$d_k \tilde{d}_k \left(\mathbf{u}_k^H \mathbf{g}_1\right)\left(\mathbf{u}_k^H \tilde{\mathbf{g}}_2\right) = d_k \tilde{d}_k \left(\mathbf{u}_k^H \mathbf{g}_2\right)\left(\mathbf{u}_k^H \tilde{\mathbf{g}}_1\right).$$

Consider index k corresponding to a nonvirtual carrier, i.e., for k such that $d_k \neq 0$. If $\tilde{d}_k = 0$, then from (11), $\mathbf{g}_1$ and $\mathbf{g}_2$ must share a common zero. Thus, $\tilde{d}_k \neq 0$, and we have

$$H_1(k)\tilde{H}_2(k) = \tilde{H}_1(k) H_2(k) \quad (12)$$

for $k = 0, \ldots, M-1$. Notice that $H_i(k)$ and $\tilde{H}_i(k)$ are, respectively, the Z transform sampled at frequency $2\pi k/N$ for impulse response $\mathbf{g}_i$ and $\tilde{\mathbf{g}}_i$. This is equivalent to, for $z = e^{-j2\pi k/N}$,

$$\left[\sum_{n=0}^{L} g_1(n) z^{-n}\right]\left[\sum_{n=0}^{L} \tilde{g}_2(n) z^{-n}\right] = \left[\sum_{n=0}^{L} g_2(n) z^{-n}\right] \cdot \left[\sum_{n=0}^{L} \tilde{g}_1(n) z^{-n}\right].$$

Expanding the products on both sides, we get

$$\beta_0 + \beta_1 z^{-1} + \cdots + \beta_{2L+1} z^{-2L} \,|_{z=e^{-j2\pi kn/N}} = 0$$

for $k = 0, \ldots, M-1$, where

$$\beta_i = \left[\sum_{n=0}^{i} g_1(n)\tilde{g}_2(i-n)\right] - \left[\sum_{n=0}^{i} g_2(n)\tilde{g}_1(i-n)\right].$$

In matrix form

$$\mathbf{Z}_M \boldsymbol{\beta} = 0.$$

The rows of $\mathbf{Z}_M$ are the corresponding M rows of $\mathbf{W}_{2L+1}$, where $\mathbf{W}_{2L+1}$ is the first $2L+1$ columns of DFT matrix $\mathbf{W}$, and $\boldsymbol{\beta} = [\beta_0, \beta_1, \ldots \beta_{2L+1}]^T$. If $M \geq 2L+1$, then the Vandermonde matrix $\mathbf{Z}_M$ is full column rank. Therefore

$$\boldsymbol{\beta} = 0 \Longrightarrow \sum_{n=0}^{i} g_1(n)\tilde{g}_2(i-n) = \sum_{n=0}^{i} g_2(n)\tilde{g}_1(i-n).$$

The left- and right-hand sides of the above equation correspond to the linear convolution between $\mathbf{g}_1$ and $\mathbf{g}_2'$, $\mathbf{g}_2$ and $\mathbf{g}_1'$, respectively. Thus, we have $H_1(z)\tilde{H}_2(z) = \tilde{H}_1(z)H_2(z)$. Therefore, $\Re(H_1(z)) \subseteq \Re(\tilde{H}_1(z)) \cup \Re(H_2(z))$, where $\Re(H_1(z))$ is the set of roots of $H_1(z)$ [18]. Since the channels do not share any common zero, we must have

$$\Re\left(H_1(z)\right) \subseteq \Re\left(\tilde{H}_1(z)\right). \quad (13)$$

Since $\mathbf{g}_1$ and $\tilde{\mathbf{g}}_1$ are of the same length, their corresponding z-transforms have the same number of roots. Combined with (13), we have

$$H_1(z) = \alpha \tilde{H}_1(z) \Longrightarrow \mathbf{g}_1 = \alpha \tilde{\mathbf{g}}_1$$

From (12) and the above expression, we have

$$\frac{H_1(k)}{\tilde{H}_1(k)} = \frac{H_2(k)}{\tilde{H}_2(k)} = \alpha. \quad (14)$$

Hence, $\mathbf{g}_2 = \alpha\tilde{\mathbf{g}}_2$, i.e., the scalar factor α is common to $\mathbf{g}_1$ and $\mathbf{g}_2$.

Note that a nonvirtual carrier system is just a special case with $M = N$, and in this case, the proof can be simplified. Using (12) for $k = 0, \ldots, N-1$ and the fact that $H_i(k)$ and $\tilde{H}_i(k)$ are, respectively, the N-point DFT at frequency $2\pi k/N$ for impulse responses $\mathbf{g}_i$ and $\tilde{\mathbf{g}}_i$, we have in the time domain the following identity for N-point circular convolution:

$$\mathbf{g}_1 \otimes \tilde{\mathbf{g}}_2 = \tilde{\mathbf{g}}_1 \otimes \mathbf{g}_2.$$

Given that $\mathbf{g}_1, \mathbf{g}_2, \tilde{\mathbf{g}}_1, \tilde{\mathbf{g}}_2$ are all vectors of length $L+1$, if $N \geq 2L+1$, then the N-point circular convolution is equivalent to linear convolution. Therefore

$$\mathbf{g}_1 * \tilde{\mathbf{g}}_2 = \tilde{\mathbf{g}}_1 * \mathbf{g}_2$$

or

$$\mathbf{H}_1(z)\tilde{\mathbf{H}}_2(z) = \tilde{\mathbf{H}}_1(z)\mathbf{H}_2(z). \quad (15)$$

Given (15), it is shown in [18] that the channel can be identified up to a scalar factor if $\mathbf{H}_1(z)$ and $\mathbf{H}_2(z)$ do not share any common nulls. Therefore, we must have

$$\begin{bmatrix} \mathbf{g}_1 \\ \mathbf{g}_2 \end{bmatrix} = \alpha \begin{bmatrix} \tilde{\mathbf{g}}_1 \\ \tilde{\mathbf{g}}_2 \end{bmatrix}.$$

Q.E.D.

APPENDIX B
DERIVATION OF THE FISHER INFORMATION MATRIX

Consider the signal model as in (5). The unknown parameter vector is

$$\boldsymbol{\theta} = [\mathrm{Re}(\mathbf{g}_1), \mathrm{Re}(\mathbf{g}_2), \mathrm{Re}(\mathbf{d}), \mathrm{Im}(\mathbf{g}_1), \mathrm{Im}(\mathbf{g}_2), \mathrm{Im}(\mathbf{d})]^T.$$

Apparently, the FIM, denoted by $\mathbf{F}$, is of dimension $2N + 4(L+1)$ by $2N + 4(L+1)$. Define

$$\boldsymbol{\mu} = \begin{bmatrix} \mathbf{D}\mathbf{W}_{L+1}\mathbf{g}_1 \\ \mathbf{D}\mathbf{W}_{L+1}\mathbf{g}_2 \end{bmatrix}$$

to be the mean value of the observation vector $[\mathbf{y}_1, \mathbf{y}_2]^T$ that is otherwise Gaussian distributed. Each element of the FIM can be written, given that the noise covariance matrix is $\sigma^2\mathbf{I}$, as

$$\mathbf{F}(i,j) = \frac{2}{\sigma^2} Re\left[\left(\frac{\partial\boldsymbol{\mu}}{\partial\theta_i}\right)^H \left(\frac{\partial\boldsymbol{\mu}}{\partial\theta_j}\right)\right].$$

Define $\tilde{\boldsymbol{\theta}} = [\mathbf{g}_1, \mathbf{g}_2, \mathbf{d}]^T$. In matrix form, $\mathbf{F}$ can be written as [24]–[26]

$$\mathbf{F} = 2\begin{bmatrix} \mathrm{Re}(\mathbf{F}_c) & -\mathrm{Im}(\mathbf{F}_c) \\ \mathrm{Im}(\mathbf{F}_c) & \mathrm{Re}(\mathbf{F}_c) \end{bmatrix}$$

where each element of $\mathbf{F_c}$ is

$$\mathbf{F_c}(i,j) = \frac{1}{\sigma^2}\left[\left(\frac{\partial\boldsymbol{\mu}}{\partial\theta_i}\right)^H \quad \left(\frac{\partial\boldsymbol{\mu}}{\partial\theta_j}\right)\right].$$

Write $\mathbf{F_c}$ in partitioned matrix form as

$$\mathbf{F_c} = \frac{1}{\sigma^2}\begin{bmatrix}\mathbf{A}_{11} & \mathbf{A}_{12} & \mathbf{A}_{13}\\ \mathbf{A}_{21} & \mathbf{A}_{22} & \mathbf{A}_{23}\\ \mathbf{A}_{31} & \mathbf{A}_{32} & \mathbf{A}_{33}\end{bmatrix}.$$

Letting $\mathbf{Q} = \mathbf{DW}_{L+1}$, we can obtain each block of the $\mathbf{F}_c$ matrix as follows:

$$\begin{aligned}
\mathbf{A}_{11} =& \frac{\partial\boldsymbol{\mu}^H}{\partial\mathbf{g}_1}\frac{\partial\boldsymbol{\mu}}{\partial\mathbf{g}_1^H} = [\mathbf{W}_{L+1}^H\mathbf{D}^H \quad 0][\mathbf{DW}_{L+1} \quad 0]^T\\
=& \mathbf{Q}^H\mathbf{Q}\\
\mathbf{A}_{12} =& \frac{\partial\boldsymbol{\mu}^H}{\partial\mathbf{g}_1}\frac{\partial\boldsymbol{\mu}}{\partial\mathbf{g}_2^H} = [\mathbf{W}_{L+1}^H\mathbf{D}^H \quad 0][0 \quad \mathbf{DW}_{L+1}]^T = 0\\
\mathbf{A}_{13} =& \frac{\partial\boldsymbol{\mu}^H}{\partial\mathbf{g}_1}\frac{\partial\boldsymbol{\mu}}{\partial\mathbf{d}^H} = [\mathbf{W}_{L+1}^H\mathbf{D}^H \quad 0][\mathbf{H}_1 \quad 0]^T = \mathbf{Q}^H\mathbf{H}_1\\
\mathbf{A}_{21} =& \frac{\partial\boldsymbol{\mu}^H}{\partial\mathbf{g}_2}\frac{\partial\boldsymbol{\mu}}{\partial\mathbf{g}_1^H} = [0 \quad \mathbf{W}_{L+1}^H\mathbf{D}^H][\mathbf{DW}_{L+1} \quad 0]^T = 0\\
\mathbf{A}_{22} =& \frac{\partial\boldsymbol{\mu}^H}{\partial\mathbf{g}_2}\frac{\partial\boldsymbol{\mu}}{\partial\mathbf{g}_2^H} = [0 \quad \mathbf{W}_{L+1}^H\mathbf{D}^H][0 \quad \mathbf{DW}_{L+1}]^T\\
=& \mathbf{Q}^H\mathbf{Q}\\
\mathbf{A}_{23} =& \frac{\partial\boldsymbol{\mu}^H}{\partial\mathbf{g}_2}\frac{\partial\boldsymbol{\mu}}{\partial\mathbf{d}^H} = [0 \quad \mathbf{W}_{L+1}^H\mathbf{D}^H][0 \quad \mathbf{H}_2]^T = \mathbf{Q}^H\mathbf{H}_2\\
\mathbf{A}_{31} =& \frac{\partial\boldsymbol{\mu}^H}{\partial\mathbf{d}}\frac{\partial\boldsymbol{\mu}}{\partial\mathbf{g}_1^H} = [\mathbf{H}_1^H \quad \mathbf{H}_2^H][\mathbf{DW}_{L+1} \quad 0]^T = \mathbf{H}_1^H\mathbf{Q}\\
\mathbf{A}_{32} =& \frac{\partial\boldsymbol{\mu}^H}{\partial\mathbf{d}}\frac{\partial\boldsymbol{\mu}}{\partial\mathbf{g}_2^H} = [\mathbf{H}_1^H \quad \mathbf{H}_2^H][0 \quad \mathbf{DW}_{L+1}]^T = \mathbf{H}_2^H\mathbf{Q}\\
\mathbf{A}_{33} =& \frac{\partial\boldsymbol{\mu}^H}{\partial\mathbf{d}}\frac{\partial\boldsymbol{\mu}}{\partial\mathbf{d}^H} = [\mathbf{H}_1^H \quad \mathbf{H}_2^H][\mathbf{H}_1 \quad \mathbf{H}_2]^T\\
=& \mathbf{H}_1^H\mathbf{H}_1 + \mathbf{H}_2^H\mathbf{H}_2.
\end{aligned}$$

Finally

$$\mathbf{F}_c = \frac{1}{\sigma^2}\begin{bmatrix}\mathbf{Q}^H\mathbf{Q} & 0 & \mathbf{Q}^H\mathbf{H}_1\\ 0 & \mathbf{Q}^H\mathbf{Q} & \mathbf{Q}^H\mathbf{H}_2\\ \mathbf{H}_1^H\mathbf{Q} & \mathbf{H}_2^H\mathbf{Q} & \mathbf{H}_1^H\mathbf{H}_1 + \mathbf{H}_2^H\mathbf{H}_2\end{bmatrix}.$$

REFERENCES

[1] *Radio Broadcasting Systems: Digital Audio Broadcasting to Mobile, Portable and Fixed Receivers*, Eur. Telecommun. Stand., Feb. 1995.

[2] *Digital Video Broadcasting: Framing Structure, Channel Coding and Modulation for Digital Terrestrial Television*, Eur. Telecommun. Stand., Aug. 1997.

[3] *Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications*, IEEE Stand. 802.11a, 1999.

[4] "Broadband Radio Access Networks (BRAN): HIPERLAN Type 2 Technical Specification Part 1—Physical Layer," ETSI, DTS/BRAN030 003–1, 1999.

[5] R. W. Heath and G. B. Giannakis, "Exploiting input cyclostationarity for blind channel identification in OFDM systems," *IEEE Trans. Signal Processing*, vol. 47, pp. 846–856, Mar. 1999.

[6] B. Muquet and M. de Courville, "Blind and semi-blind channel identification methods using second order statistics for OFDM systems," in *Proc. ICASSP*, vol. 5, Phoenix, AZ, Mar. 1999, pp. 2745–2748.

[7] B. Muquet, M. de Courville, P. Duhamel, and V. Buenac, "A subspace based blind and semi-blind channel identification method for OFDM systems," in *Proc. IEEE Workshop Signal Processing Advances Wireless Commun.*, Annapolis, MD, May 1999, pp. 170–173.

[8] X. Cai and A. N. Akansu, "A subspace method for blind channel identification in OFDM systems," in *Proc. ICC*, vol. 2, New Brunswick, NJ, Mar. 2000, pp. 929–933.

[9] X. Zhuang, Z. Ding, and A. L. Swindlehurst, "A statistical subspace method for blind channel identification in OFDM communications," in *Proc. ICASSP*, vol. 5, Istanbul, Turkey, June 2000, pp. 2493–2496.

[10] C. Li and S. Roy, "Subspace based blind channel estimation for OFDM by exploiting virtual carrier," in *Proc. CLOBECOM*, vol. 1, San Antonio, TX, 2001, pp. 295–299.

[11] N. Chotikakamthorn and H. B. Suzuki, "On identifiability of OFDM blind channel estimation," in *Proc. IEEE Vehicular Technology Conf.*, Amsterdam, The Netherlands, Sept. 1999.

[12] S. Zhou and G. B. Giannakis, "Long codes for generalized FH-OFDMA through unknown multipath channels," *IEEE Trans. Commun.*, vol. 49, pp. 721–733, Apr. 2001.

[13] ——, "Finite-alphabet based channel estimation for OFDM and related multicarrier systems," *IEEE Trans. Commun.*, vol. 49, pp. 1402–1414, Aug. 2001.

[14] T. Petermann, S. Vogeler, D. Boss, and K. Kammeyer, "Blind turbo channel estimation in OFDM receivers," in *Proc. 35th Asilomar Conf. Signals, Syst., Comput.*, vol. 1, Pacific Grove, CA, Nov. 2001, pp. 1489–1493.

[15] H. Ali, J. H. Manton, and Y. Hua, "A SOS subspace method for blind channel identification and equalization in bandwidth efficient OFDM systems based on receive antenna diversity," in *Proc. 11th IEEE Signal Processing Workshop Statistical Signal Processing*, Singapore, Aug. 2001, pp. 401–404.

[16] C. Li and S. Roy, "A subspace blind channel estimation method for OFDM systems without cyclic prefix," in *Proc. Vehicular Technol. Conf.*, vol. 4, Atlantic City, NJ, Oct. 2001, pp. 2148–2152.

[17] Z. Liu, G. B. Giannakis, S. Barbarossa, and A. Scaglione, "Transmit-antennae space-time block coding for generalized OFDM in the presence of unknown multipath," *IEEE Trans. Commun.*, vol. 19, pp. 1352–1364, July 2001.

[18] G. Xu, H. Liu, L. Tong, and T. Kailath, "A least-squares approach to blind channel identification," *IEEE Trans. Signal Processing*, vol. 43, pp. 2982–2993, Dec. 1995.

[19] R. van Nee and R. Prasad, *OFDM For Multimedia Wireless Communications*. Boston, MA: Artech House, 2000.

[20] Z. Wang and G. B. Giannakis, "Wireless multicarrier communications: Where fourier meets Shannon," *IEEE Signal Processing Mag.*, vol. 47, pp. 29–48, May 2000.

[21] R. Negi and J. Cioffi, "Pilot tone selection for channel estimation in a mobile OFDM system," *IEEE Trans. Consumer Electron.*, vol. 44, pp. 1122–1128, Aug. 1998.

[22] G. H. Golub and C. F. Van Loan, *Matrix Computations*. Baltimore, MD: The Johns Hopkins Univ. Press, 1990.

[23] Y. Hua and M. Wax, "Strict identifibility of multiple FIR channels driven by an unknown arbitrary sequence," *IEEE Trans. Signal Processing*, vol. 44, pp. 756–759, Mar. 1996.

[24] P. Stoica and R. Moses, *Introduction to Spectral Analysis*. Upper Saddle River, NJ: Prentice-Hall, 1997.

[25] Y. Hua, "Fast maximum likelihood for blind identification of multiple FIR channels," *IEEE Trans. Signal Processing*, vol. 44, pp. 661–672, Mar. 1996.

[26] M. Morelli and U. Mengali, "A comparison of pilot-aided channel estimation methods for OFDM systems," *IEEE Trans. Signal Processing*, vol. 49, pp. 3065–3073, Dec. 2001.

[27] P. Stoica and B. C. Ng, "On the Cramer-Rao bound under parametric constraints," *IEEE Signal Processing Lett.*, vol. 5, pp. 177–179, July 1998.

[28] S. Zhou, B. Muquet, and G. B. Giannakis, "Subspace-based (semi-) blind channel estimation for block precoded space-time OFDM," *IEEE Trans. Signal Processing*, vol. 50, pp. 1215–1228, May 2002.

Hao Wang (S'03) received the B.E. degree from Tsinghua University, Beijing, China, in 1996 and the M.S. degree from the University of Connecticut, Storrs, in 1998, both in electrical engineering. Since 2001, he has been with the Department of Electrical Engineering and Computer Science, Syracuse University, Syracuse, NY, where he is pursuing the Ph.D. degree.

From 1999 to 2000, he was with Schlumberger Technologies Asian Ltd., Beijing. His current research interests focus on multicarrier communications, specifically, channel estimation, timing/frequency synchronization, and peak power analysis for OFDM systems.

Ying Lin received the B.S. and M.S. degrees in electrical engineering from Harbin Institute of Technology, Harbin, China, in 1995 and 1997, respectively. Since 2000, she has been pursuing the Ph.D. degree with the Department of Electrical Engineering and Computer Science, Syracuse University, Syracuse, NY.

From 1997 to 2000, she was a software engineer with CTC Telecom Company, Beijing, China. Her current research interests are related to wireless sensor networks, wireless communications, and statistical signal processing.

Biao Chen (S'97–M'99) received the M.S. degree in statistics and the Ph.D. degree in electrical engineering from the University of Conneticut, Storrs, in 1998 and 1999, respectively.

From 1999 to 2000, he was with Cornell University, Ithaca, NY, as a post-doctoral research associate. Since 2000, he has been with the Department of Electrical Engineering and Computer Science, Syracuse University, Syracuse, NY, as an assistant professor. His area of interest mainly focuses on statistical signal processing with applications to communication and radar systems.