Blind and Semiblind Channel and Carrier Frequency-Offset Estimation in Orthogonally Space-Time Block Coded MIMO Systems

Shahram Shahbazpanahi, *Member, IEEE*, Alex B. Gershman, *Fellow, IEEE*, and Georgios B. Giannakis, *Fellow, IEEE*

***Abstract*—In this paper, the problem of joint channel and carrier frequency offset (CFO) estimation is studied in the context of multiple-input multiple-output (MIMO) communications based on orthogonal space-time-block codes (OSTBCs). A new blind approach is proposed to jointly estimate the channel matrix and the CFO parameters using a relaxed maximum likelihood (ML) estimator that, for the sake of simplicity, ignores the finite alphabet constraint. Although the proposed technique can be applied to the majority of OSTBCs, there are, however, a few codes that suffer from an intrinsic ambiguity in the joint channel, CFO, and symbol estimates. For such specific OSTBCs, a semiblind modification of the proposed approach is developed that resolves the aforementioned estimation ambiguity. Our simulation results demonstrate that although the finite alphabet constraint is relaxed, the performance of the proposed techniques approaches that of the informed (fully frequency-synchronized and coherent) receiver, provided that a sufficient number of data blocks is available for each channel realization.**

***Index Terms*—Blind channel and carrier frequency offset estimation, multiple-input multiple-output (MIMO) communications, orthogonal space-time block codes.**

I. Introduction

SPACE-TIME coding has recently gained much interest because of its ability to combat fading by means of exploiting spatial diversity provided by multiple-input multiple-output (MIMO) communication channels [1]–[3].

Among different space-time coding techniques proposed so far, orthogonal space-time codes (OSTBCs) are of great interest as they collect full diversity at low decoding complexity. The optimal ML decoder for OSTBCs amounts to a simple linear matched filter (MF) receiver followed by a symbol-by-symbol decoder. It has recently been shown in the literature that for the majority of OSTBCs, the MIMO channel is blindly identifiable [4], [5]. Therefore, employing OSTBCs as the underlying space-time coding scheme can potentially reduce the system training requirements and, therefore, improve its bandwidth efficiency.

For flat fading links, a blind method has been developed in [5] to estimate the channel matrix. The approach of [5] has been shown to be equivalent to a joint channel and symbol ML estimator after relaxing the finite alphabet constraint on the information symbols. Interestingly, neglecting the finite alphabet property in this case does not seriously affect the performance of the resulting decoder. This surprising fact is due to the unique structural properties of the OSTBCs [6].

The technique of [5] assumes that no carrier frequency offset (CFO) is present between the transmitter and receiver. Unfortunately, the latter assumption may be often violated in practice. For example, even if the frequency synchronization between the transmitter and receiver is perfect, such frequency offsets can be caused by mobility-induced Doppler effects.

The problem of estimating CFO parameters has been studied for different communication schemes including training based MIMO systems [7], orthogonal frequency-division multiplexing (OFDM)-based SISO and MIMO systems [8], [9], and joint CFO and symbol timing recovery schemes exploiting cyclostationarity [10]. However, to the best of our knowledge, the CFO estimation problem has not been studied in application to OSTBC-based MIMO communication systems. Therefore, in this paper (see also [11] and [12]), we extend the work of [5] to the problem of joint blind channel and CFO estimation. Assuming constant modulus constellations, we use the ML approach to obtain joint estimates of the channel matrix and the CFO parameters, and to detect the information symbols based on these estimates. The proposed approach, in essence, relaxes the finite alphabet constraint and reformulates the original problem as a simpler non-linear optimization problem with a single unknown CFO parameter. It is shown that this parameter can be estimated by maximizing the principal eigenvalue of a certain data-dependent matrix and, once the CFO estimate is obtained, the channel matrix estimate can be easily recovered from the principal eigenvector of the same matrix. The obtained CFO and channel matrix estimates can then be used along with the conventional (coherent) OSTBC symbol detector to recover the transmitted symbols.

Although the proposed blind method is applicable to the majority of existing OSTBCs, there are, however, a few codes (including the celebrated Alamouti's code [2]) that suffer from an intrinsic ambiguity in joint channel, CFO, and symbol estimates. For these OSTBCs, a semiblind modification of the proposed method is developed that resolves the above-mentioned estimation ambiguity by means of using a few training blocks.

Manuscript received December 5, 2006; revised May 24, 2007. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Aleksandar Dogandzic. This work was supported in part by the Wolfgang Paul Award Program of the Alexander von Humboldt Foundation (Germany) and the German Ministry of Education and Research; Natural Sciences and Engineering Research Council (NSERC) of Canada; and the National Science Foundation (NSF) Grant CCF-0515032.

S. Shahbazpanahi is with the Faculty of Engineering and Applied Science, University of Ontario Institute of Technology, Oshawa, ON, Canada (e-mail: shahram.shahbazpanahi@uoit.ca).

A. B. Gershman is with the Communication Systems Group, Darmstadt University of Technology, D-64283 Darmstadt, Germany (e-mail: gershman@ieee.org).

G. B. Giannakis is with the Department of Electrical and Computer Engineering, University of Minnesota, Minneapolis, MN 55455 USA (e-mail: georgios@ece.umn.edu).

Digital Object Identifier 10.1109/TSP.2007.906668

Simulation results validate the performance of the proposed blind and semiblind methods as compared to the informed (i.e., perfectly synchronized channel-coherent) receiver and to the training-based technique of [7]. It is demonstrated that despite the relaxed finite alphabet constraint, the performances of the proposed techniques closely approach that of the informed receiver provided that there is a sufficient number of information-bearing data blocks available for each particular channel realization. It is also shown that since the proposed semiblind technique uses the information-bearing data in addition to the training data, it substantially outperforms the method of [7].

The remainder of the paper is organized as follows. In Section II, the data model is presented. Section III introduces the concept of time-varying OSTBC in the case when frequency offsets are present. Section IV develops our joint channel and CFO estimation techniques. Simulation results are presented in Section V and conclusions are drawn in Section VI.

II. Signal Model

Let us consider a MIMO system with N transmit and M receive antennas. In the flat block-fading channel case, its input-output relationship can be written as [13]

$$\mathbf{y}(t) = e^{j\omega_o t}\mathbf{x}(t)\mathbf{H} + \mathbf{v}(t) \tag{1}$$

where $\mathbf{H}$ is the $N \times M$ complex channel matrix, ω_o is the CFO between the transmitter and receiver, and

$$\begin{aligned}\mathbf{y}(t) &\triangleq [y_1(t)\ y_2(t)\ \ldots\ y_M(t)]\\ \mathbf{x}(t) &\triangleq [x_1(t)\ x_2(t)\ \ldots\ x_N(t)]\\ \mathbf{v}(t) &\triangleq [v_1(t)\ v_2(t)\ \ldots\ v_M(t)]\end{aligned}$$

are the complex row-vectors of the received signal, transmitted signal, and additive noise, respectively. It is assumed that the noise is temporally and spatially white complex Gaussian with variance σ^2.

Assuming a block transmission scheme with block length T, the nth received data block can be expressed as

$$\mathbf{Y}(n) = \mathbf{D}(n,\omega_o)\mathbf{X}(n)\mathbf{H} + \mathbf{V}(n) \tag{2}$$

where

$$\mathbf{Y}(n) \triangleq \begin{bmatrix}\mathbf{y}((n-1)T+1)\\ \mathbf{y}((n-1)T+2)\\ \vdots\\ \mathbf{y}(nT)\end{bmatrix} \quad \mathbf{X}(n) \triangleq \begin{bmatrix}\mathbf{x}((n-1)T+1)\\ \mathbf{x}((n-1)T+2)\\ \vdots\\ \mathbf{x}(nT)\end{bmatrix}$$

$$\mathbf{V}(n) \triangleq \begin{bmatrix}\mathbf{v}((n-1)T+1)\\ \mathbf{v}((n-1)T+2)\\ \vdots\\ \mathbf{v}(nT)\end{bmatrix}$$

are the nth blocks of the received signals, transmitted signals, and additive noise, respectively; the $T \times T$ complex diagonal matrix $\mathbf{D}(n,\omega_o)$ is defined as

$$\mathbf{D}(n,\omega_o) \triangleq \mathrm{diag}\left\{e^{j\omega_o((n-1)T+1)}\ \cdots\ e^{j\omega_o nT}\right\}$$

and $(\cdot)^T$ denotes the transpose. Hereafter, we assume a slow fading channel whose coherence time is much longer than the data block length T.

The matrix $\mathbf{X}(n) = \mathbf{X}(\mathbf{s}(n))$ can be viewed as a mapping that transforms the nth symbol block to a $T \times N$ complex matrix of transmit signals, where $\mathbf{s}(n) \triangleq [s_1(n)\ s_2(n)\ \ldots\ s_K(n)]^T$ is the nth symbol vector of length K. The entries of $\mathbf{s}(n)$ are assumed to be randomly drawn from a *constant modulus* constellation, that is, $|s_k(n)| = 1$. Using this assumption, the need for estimating the norm of the channel matrix can be alleviated [5].

Note that $\mathbf{s}(n) \in \mathcal{S}$ where $\mathcal{S} \triangleq \{\mathbf{s}^{(1)}\,\mathbf{s}^{(2)}\ldots\mathbf{s}^{(L)}\}$ is the symbol vector alphabet of the size L, that is, the set of all possible symbol vectors.

The $T \times N$ matrix $\mathbf{X}(\mathbf{s}(n))$ is called an OSTBC [3] if all elements of this matrix are linear functions of the K complex variables $s_1(n), s_2(n), \ldots, s_K(n)$ and their complex conjugates, and if for any arbitrary $\mathbf{s}(n)$, it satisfies the following property:

$$\mathbf{X}^H(\mathbf{s}(n))\mathbf{X}(\mathbf{s}(n)) = \|\mathbf{s}(n)\|^2\mathbf{I}_N$$

where $\mathbf{I}_N$ is the $N \times N$ identity matrix, $\|\cdot\|$ is the Euclidean norm, and $(\cdot)^H$ stands for Hermitian transposition.

From the definition of OSTBCs, it directly follows that the matrix $\mathbf{X}(\mathbf{s}(n))$ can be expressed as [6], [14], [15]

$$\mathbf{X}(\mathbf{s}(n)) = \sum_{k=1}^{K}\left(\mathbf{C}_k\mathrm{Re}\{s_k(n)\} + \mathbf{C}_{k+K}\mathrm{Im}\{s_k(n)\}\right) \tag{3}$$

where

$$\mathbf{C}_k \triangleq \begin{cases}\mathbf{X}(\mathbf{e}_k), & \text{for } 1 \le k \le K\\ \mathbf{X}(j\mathbf{e}_{k-K}), & \text{for } K+1 \le k \le 2K\end{cases} \tag{4}$$

$j \triangleq \sqrt{-1}$, $\mathrm{Re}\{\cdot\}$ and $\mathrm{Im}\{\cdot\}$ denote the real and imaginary parts, respectively, and the $K \times 1$ vector $\mathbf{e}_k$ is defined as the kth column of the identity matrix $\mathbf{I}_K$. It is worth noting that any OSTBC is completely defined by its *basis matrices* $\{\mathbf{C}_k\}_{k=1}^{2K}$, for example, see [6].

For any matrix $\mathbf{P}$, let us define the "underline" operator as

$$\underline{\mathbf{P}} \triangleq \begin{bmatrix}\mathrm{vec}\{\mathrm{Re}(\mathbf{P})\}\\ \mathrm{vec}\{\mathrm{Im}(\mathbf{P})\}\end{bmatrix} \tag{5}$$

where $\mathrm{vec}\{\cdot\}$ is the vectorization operator that stacks all columns of a matrix on top of each other.

If there is no CFO ($\omega_o = 0$), then, using (3) and (5), we can rewrite the model (2) in the following vectorized form [6], [15]

$$\mathbf{z}_n \triangleq \underline{\mathbf{Y}}(n) = \mathbf{A}(\mathbf{H})\mathbf{g}(n) + \boldsymbol{\nu}(n) \tag{6}$$

where, for the sake of notational simplicity, $\mathbf{g}(n) \triangleq \underline{\mathbf{s}}(n)$ and $\boldsymbol{\nu}(n) \triangleq \underline{\mathbf{V}}(n)$, and the $2MT \times 2K$ real-valued matrix

$$\mathbf{A}(\mathbf{H}) \triangleq [\underline{\mathbf{C}_1\mathbf{H}}\quad \underline{\mathbf{C}_2\mathbf{H}}\quad \cdots\quad \underline{\mathbf{C}_{2K}\mathbf{H}}]$$

captures both the effects of the OSTBC and the channel.

It is readily verifiable that, regardless of the value of the channel matrix $\mathbf{H}$, the matrix $\mathbf{A}(\mathbf{H})$ obeys the so-called *decoupling property*, that is, all its columns have identical norms and are orthogonal to each other [6]

$$\mathbf{A}^T(\mathbf{H})\mathbf{A}(\mathbf{H}) = \|\mathbf{H}\|_F^2\,\mathbf{I}_{2K} \tag{7}$$

where $\|\cdot\|_F$ denotes the Frobenius norm of a matrix.

III. Time-Varying OSTBCs

In the presence of CFO, the matrix

$$\tilde{\mathbf{X}}(n,\omega_o,\mathbf{s}(n)) \triangleq \mathbf{D}(n,\omega_o)\mathbf{X}(\mathbf{s}(n))$$

obeys the orthogonality property

$$\begin{aligned}\tilde{\mathbf{X}}^H&(n,\omega_o,\mathbf{s}(n))\tilde{\mathbf{X}}(n,\omega_o,\mathbf{s}(n))\\&=\mathbf{X}^H(\mathbf{s}(n))\underbrace{\mathbf{D}^H(n,\omega_o)\mathbf{D}(n,\omega_o)}_{\mathbf{I}_T}\mathbf{X}(\mathbf{s}(n))\\&=\mathbf{X}^H(\mathbf{s}(n))\mathbf{X}(\mathbf{s}(n))\\&=\|\mathbf{s}(n)\|^2\mathbf{I}_N\end{aligned}$$

and, therefore, is a *legitimate* OSTBC regardless of the value of ω_0. It is easy to verify that

$$\tilde{\mathbf{X}}(n,\omega_o,\mathbf{s}(n))=\sum_{k=1}^{K}\Big(\tilde{\mathbf{C}}_k(n,\omega_o)\mathrm{Re}\{s_k(n)\} + \tilde{\mathbf{C}}_{k+K}(n,\omega_o)\mathrm{Im}\{s_k(n)\}\Big) \quad (8)$$

where

$$\tilde{\mathbf{C}}_k(n,\omega_o) \triangleq \mathbf{D}(n,\omega_o)\mathbf{C}_k. \quad (9)$$

From (9), it follows that the basis matrices $\{\tilde{\mathbf{C}}_k\}_{k=1}^{2K}$ of $\tilde{\mathbf{X}}(n,\omega_o,\mathbf{s}(n))$ are time varying. Because of this, we will refer to $\tilde{\mathbf{X}}(n,\omega_o,\mathbf{s}(n))$ as *time-varying OSTBC*.

Using (8), we have that in the presence of CFO, (6) should be modified as

$$\mathbf{z}(n)=\mathbf{A}(n,\omega_o,\mathbf{H})\mathbf{g}(n)+\boldsymbol{\nu}(n) \quad (10)$$

where

$$\mathbf{A}(n,\omega_o,\mathbf{H}) \triangleq [\,\underline{\tilde{\mathbf{C}}_1(n,\omega_o)\mathbf{H}} \quad \cdots \quad \underline{\tilde{\mathbf{C}}_{2K}(n,\omega_o)\mathbf{H}}\,].$$

As $\tilde{\mathbf{X}}(n,\omega_o,\mathbf{s}(n))$ is a legitimate OSTBC, the matrix $\mathbf{A}(n,\omega_o,\mathbf{H})$ satisfies the decoupling property regardless of the values of n and ω_o, that is

$$\mathbf{A}^T(n,\omega_o,\mathbf{H})\mathbf{A}(n,\omega_o,\mathbf{H})=\|\mathbf{H}\|_F^2\,\mathbf{I}_{2K}. \quad (11)$$

Introducing the equivalent *channel vector* $\mathbf{h}$ as

$$\mathbf{h} \triangleq \underline{\mathbf{H}} \quad (12)$$

let us hereafter with a small abuse of notation replace $\mathbf{A}(n,\omega_o,\mathbf{H})$ with $\mathbf{A}(n,\omega_o,\mathbf{h})$. As $\mathbf{A}(n,\omega_o,\mathbf{h})$ is linear in $\mathbf{h}$, we have

$$\mathrm{vec}\{\mathbf{A}(n,\omega_o,\mathbf{h})\}=\boldsymbol{\Phi}(n,\omega_o)\mathbf{h} \quad (13)$$

where $\boldsymbol{\Phi}(n,\omega_o)$ is a $4KMT\times 2MN$ matrix whose kth column can be defined as

$$[\boldsymbol{\Phi}(n,\omega_o)]_k \triangleq \mathrm{vec}\{\mathbf{A}(n,\omega_o,\mathbf{e}_k)\} \quad (14)$$

where $[\cdot]_k$ stands for the kth column of a matrix and $\mathbf{e}_k$ is the kth column of the identity matrix $\mathbf{I}_{2MN}$. We stress here that in (14), the size of $\mathbf{e}_k$ is different as compared to $\mathbf{e}_k$ used in (4). However, for the sake of simplicity the same notation $\mathbf{e}_k$ is used in (4) and (14).

IV. Joint Blind Channel and Frequency Offset Estimation

Let n_B data blocks be available for each channel realization. Treating the channel vector $\mathbf{h}$, the CFO parameter ω_o, and the information symbols $\{\mathbf{g}(n)\}_{n=1}^{n_B}$ as unknown deterministic parameters, let us use the ML approach to jointly estimate these parameters. To obtain the ML estimates of all these parameters, the log-likelihood (LL) function needs to be maximized. Hence, the parameter estimates can be obtained by solving the following optimization problem:

$$\max_{\omega_o,\mathbf{h}}\max_{\mathbf{S}\in\Omega}\log f\big(\mathbf{z}(1),\ldots,\mathbf{z}(n_B)\,|\,\mathbf{S},\,\mathbf{h},\omega_o\big) \quad (15)$$

where $f\big(\mathbf{z}(1),\ldots,\mathbf{z}(n_B)\,|\,\mathbf{S},\,\mathbf{h},\omega_o\big)$ is the likelihood function computed for n_B snapshots $\{\mathbf{z}(n)\}_{n=1}^{n_B}$, $\mathbf{S} \triangleq [\mathbf{g}(1)\ \mathbf{g}(2)\ \ldots\ \mathbf{g}(n_B)]$, and Ω is the finite set of all possible values of $\mathbf{S}$. Note, however, that it is extremely difficult to solve (15) because its computational cost grows exponentially in n_B. To simplify the optimization problem in (15), let us relax the finite alphabet constraint $\mathbf{S}\in\Omega$, that is, assume that $\mathbf{S}\in\mathbb{R}^{2K\times n_B}$. Then, the optimization problem in (15) can be rewritten as

$$\max_{\omega_o,\mathbf{h},\mathbf{S}}\log f\big(\mathbf{z}(1),\ldots,\mathbf{z}(n_B)\,|\,\mathbf{S},\,\mathbf{h},\omega_o\big). \quad (16)$$

As the noise vectors $\{\boldsymbol{\nu}(n)\}_{n=1}^{n_B}$ are zero-mean independent identically distributed (i.i.d.) Gaussian with the covariance matrix

$$\mathrm{E}\{\boldsymbol{\nu}(n)\,\boldsymbol{\nu}^T(n)\}=\frac{\sigma^2}{2}\,\mathbf{I}_{2TM}$$

the likelihood function for any $\mathbf{z}(n)$ can be expressed as

$$f\big(\mathbf{z}(n)\,|\,\mathbf{g}(n),\,\mathbf{h},\omega_o\big)=\frac{1}{(\pi\sigma^2)^{MT}}\cdot\exp\left(-\frac{\|\mathbf{z}(n)-\mathbf{A}(n,\omega_o,\mathbf{h})\mathbf{g}(n)\|^2}{\sigma^2}\right) \quad (17)$$

where $\mathrm{E}\{\cdot\}$ denotes the statistical expectation. Taking into account that all $\{\mathbf{z}(n)\}_{n=1}^{n_B}$ are independent random vectors, we obtain

$$f\big(\mathbf{z}(1),\ldots,\mathbf{z}(n_B)\,|\,\mathbf{S},\,\mathbf{h},\omega_o\big)=\prod_{n=1}^{n_B}f\big(\mathbf{z}(n)\,|\,\mathbf{g}(n),\mathbf{h},\omega_o\big). \quad (18)$$

Using (17) and (18), the problem in (16) can be reformulated as

$$\min_{\omega_o,\mathbf{h},\mathbf{S}}\sum_{n=1}^{n_B}\big\|\mathbf{z}(n)-\mathbf{A}(n,\omega_o,\mathbf{h})\mathbf{g}(n)\big\|^2. \quad (19)$$

Note that the nth term of the sum in (19) is minimized with

$$\mathbf{g}(n)=\frac{1}{\|\mathbf{h}\|^2}\mathbf{A}^T(n,\omega_o,\mathbf{h})\,\mathbf{z}(n) \quad (20)$$

where (20) follows from the fact that $\mathbf{A}(n,\omega_o,\mathbf{H})$ satisfies the decoupling property (11).

Using (20), the objective function in (19) can be concentrated with respect to $\{\mathbf{g}(n)\}_{n=1}^{n_B}$ and, after such concentration, the latter optimization problem can be expressed as

$$\min_{\omega_o,\mathbf{h}} \sum_{n=1}^{n_B} \left\| \mathbf{z}(n) - \frac{\mathbf{A}(n,\omega_o,\mathbf{h})\mathbf{A}^T(n,\omega_o,\mathbf{h})\mathbf{z}(n)}{\|\mathbf{h}\|^2} \right\|^2 . \quad (21)$$

The objective function in (21) can be further simplified as (22), shown at the bottom of the page, where $\mathrm{tr}(\cdot)$ denotes the trace of a matrix. We also note that

$$\begin{aligned} &\mathrm{tr}(\mathbf{A}^T(n,\omega_o,\mathbf{h})\mathbf{z}(n)\mathbf{z}^T(n)\mathbf{A}(n,\omega_o,\mathbf{h})) \\ &= \mathrm{vec}^T\{\mathbf{A}(n,\omega_o,\mathbf{h})\}\left(\mathbf{I}_{2K}\otimes \mathbf{z}(n)\mathbf{z}^T(n)\right)\mathrm{vec}\{\mathbf{A}(n,\omega_o,\mathbf{h})\} \\ &= \mathbf{h}^T\mathbf{\Phi}^T(n,\omega_o)\left(\mathbf{I}_{2K}\otimes \mathbf{z}(n)\mathbf{z}^T(n)\right)\mathbf{\Phi}(n,\omega_o)\mathbf{h} \quad (23) \end{aligned}$$

where (13) was used, and $\otimes$ denotes the Kronecker matrix product. Inserting (23) into (22), the concentrated optimization problem in (21) can be expressed as

$$\max_{\omega_o,\mathbf{h}} \frac{\mathbf{h}^T\mathbf{\Psi}(\omega_o)\mathbf{h}}{\|\mathbf{h}\|^2} \quad (24)$$

where

$$\mathbf{\Psi}(\omega_o) \triangleq \sum_{n=1}^{n_B} \mathbf{\Phi}^T(n,\omega_o)\left(\mathbf{I}_{2K}\otimes \mathbf{z}(n)\mathbf{z}^T(n)\right)\mathbf{\Phi}(n,\omega_o) \quad (25)$$

is a $2MN \times 2MN$ real matrix which depends on the received data vectors $\{\mathbf{z}_n\}_{n=1}^{n_B}$ and the CFO parameter ω_o.

For any value of ω_o, the maximization over $\mathbf{h}$ yields

$$\mathbf{h} = \mathcal{P}\{\mathbf{\Psi}(\omega_o)\} \quad (26)$$

where $\mathcal{P}\{\cdot\}$ denotes the *normalized* principal eigenvector of a matrix, and it is assumed that there is no multiplicity in the largest eigenvalue of $\mathbf{\Psi}(\omega_o)$. Substituting (26) into (24), we obtain that the CFO parameter can be estimated as

$$\hat{\omega}_o = \arg\max_{\omega_o} \lambda_{\max}\{\mathbf{\Psi}(\omega_o)\} \quad (27)$$

where $\lambda_{\max}\{\cdot\}$ denotes the largest eigenvalue of a matrix. Substituting (27) into (26), we obtain that, given the estimate of ω_o, the channel vector estimate can be expressed as

$$\hat{\mathbf{h}} = \mathcal{P}\{\mathbf{\Psi}(\hat{\omega}_o)\}. \quad (28)$$

The CFO estimate in (27) has to be obtained through a one-dimensional search, that is, the matrix $\mathbf{\Psi}(\omega_0)$ and its principal eigenvalues and eigenvectors have to be calculated for all possible values of ω_0. The value of ω_0 which results in the maximal principal eigenvalue is referred to as the CFO estimate, and the corresponding eigenvector in (28) gives the channel estimate.

We stress that the channel estimate in (28) suffers from a real scalar ambiguity because the norm of the channel remains unknown. However, as we assumed that the symbol constellation is constant modulus, the symbol decoder will not be affected by such an ambiguity. As the data are real-valued, there is no phase ambiguity, but the channel estimate in (28) may suffer from the sign ambiguity. However, the latter type of ambiguity is common to all blind detectors and can be resolved by appropriate decoding of each symbol sequence [16].

It should be also stressed that we have not been able to prove that the CFO estimate in (27) is unique. Note that the non-uniqueness of blind CFO estimates in wireless communications is a common issue [17]. However, throughout extensive simulations, we have not observed any single case where the CFO estimate is not unique.

Once the CFO parameter and the channel vector are estimated using (27) and (28), respectively, the information symbols can be straightforwardly decoded by replacing $\mathbf{h}$ and ω_o in (20) with their estimates $\hat{\mathbf{h}}$ and $\hat{\omega}_o$ to obtain the estimate $\hat{\mathbf{g}}(n)$ of $\mathbf{g}(n)$, and then by computing the estimate

$$\hat{\mathbf{s}}(n) = [\mathbf{I}_K \;\; j\mathbf{I}_K]\hat{\mathbf{g}}(n) \quad (29)$$

of the symbol vector $\mathbf{s}(n)$. Finally, the kth information symbol can be recovered by comparing the kth entry of $\hat{\mathbf{s}}(n)$ with all the points in the corresponding constellation and using nearest neighbor decoding.

V. Semiblind Extension of the Proposed Estimator

In the previous section, we have assumed that the largest eigenvalue of $\mathbf{\Psi}(\hat{\omega}_o)$ has no multiplicity (i.e., its multiplicity

$$\begin{aligned} &\sum_{n=1}^{n_B} \left\| \mathbf{z}(n) - \frac{\mathbf{A}(n,\omega_o,\mathbf{h})\mathbf{A}^T(n,\omega_o,\mathbf{h})\mathbf{z}(n)}{\|\mathbf{h}\|^2} \right\|^2 \\ &= \sum_{n=1}^{n_B} \Bigg(-\frac{2}{\|\mathbf{h}\|^2}\mathbf{z}^T(n)\mathbf{A}(n,\omega_o,\mathbf{h})\mathbf{A}^T(n,\omega_o,\mathbf{h})\mathbf{z}(n) + \frac{1}{\|\mathbf{h}\|^4}\mathbf{z}^T(n)\mathbf{A}(n,\omega_o,\mathbf{h})\underbrace{\mathbf{A}^T(n,\omega_o,\mathbf{h})\mathbf{A}(n,\omega_o,\mathbf{h})}_{\|\mathbf{h}\|^2\mathbf{I}_{2K}} \\ &\quad \cdot \mathbf{A}^T(n,\omega_o,\mathbf{h})\mathbf{z}(n) + \|\mathbf{z}(n)\|^2\Bigg) \\ &= -\frac{1}{\|\mathbf{h}\|^2}\sum_{n=1}^{n_B}\mathbf{z}^T(n)\mathbf{A}(n,\omega_o,\mathbf{h})\mathbf{A}^T(n,\omega_o,\mathbf{h})\mathbf{z}(n) + \mathrm{constant} \\ &= -\frac{1}{\|\mathbf{h}\|^2}\sum_{n=1}^{n_B}\mathrm{tr}\left(\mathbf{A}^T(n,\omega_o,\mathbf{h})\mathbf{z}(n)\mathbf{z}^T(n)\mathbf{A}(n,\omega_o,\mathbf{h})\right) + \mathrm{constant} \quad (22) \end{aligned}$$

order is equal to one). This assumption[1] holds true for the majority of OSTBCs with a few exceptions that include the Alamouti's code [5]. For those specific OSTBCs that result in $\boldsymbol{\Psi}(\hat{\omega}_o)$ with multiple largest eigenvalues, $\mathbf{h}$ belongs to the subspace spanned by the corresponding multiple principal eigenvectors of $\boldsymbol{\Psi}(\hat{\omega}_o)$ and, as a result, the blind technique proposed in the previous section is not applicable. Therefore, to enable joint channel and CFO estimation in the latter case, we will develop a semiblind modification of the proposed approach[2] that uses a small number of training blocks. Let the multiplicity order of the largest eigenvalue of $\boldsymbol{\Psi}(\hat{\omega}_o)$ be $n_o > 1$ and the corresponding orthonormal principal eigenvectors be $\{\mathbf{u}_l\}_{l=1}^{n_o}$. As $\mathbf{h}$ belongs to the subspace spanned by $\{\mathbf{u}_l\}_{l=1}^{n_o}$, we have

$$\mathbf{h} = \sum_{l=1}^{n_o} \alpha_l \mathbf{u}_l = \mathbf{U}\boldsymbol{\alpha} \tag{30}$$

where $\mathbf{U} \triangleq [\mathbf{u}_1 \; \mathbf{u}_2 \; \dots \; \mathbf{u}_{n_o}]$ and $\boldsymbol{\alpha} \triangleq [\alpha_1 \; \alpha_2 \; \dots \; \alpha_{n_o}]^T$. The key idea of the proposed semiblind method is to obtain the estimate of $\mathbf{U}$ in a blind way, while estimating the vector $\boldsymbol{\alpha}$ using a few training symbols. As the number of entries in $\boldsymbol{\alpha}$ is much less than that in $\mathbf{h}$, such a *semiblind* estimator will require much less training data than the direct training-based channel estimator obtaining all entries of $\mathbf{h}$ in a nonblind way.

Assuming that the first n_T blocks contain training symbols and using (30), for any ith training block ($i = 1 \dots, n_T$) we have

$$\begin{aligned}\mathbf{z}(i) &= \mathbf{A}(i, \omega_0, \mathbf{h})\mathbf{g}(i) + \boldsymbol{\nu}(i) \\ &= \mathbf{A}(i, \omega_0, \mathbf{U}\boldsymbol{\alpha})\mathbf{g}(i) + \boldsymbol{\nu}(i) \\ &= \sum_{l=1}^{n_o} \alpha_l \mathbf{A}(i, \omega_0, \mathbf{u}_l)\mathbf{g}(i) + \boldsymbol{\nu}(i) \\ &= \sum_{l=1}^{n_o} \alpha_l \mathbf{B}(i, \omega_0, \mathbf{g}(i))\mathbf{u}_l + \boldsymbol{\nu}(i) \\ &= \mathbf{B}(i, \omega_0, \mathbf{g}(i))\mathbf{U}\boldsymbol{\alpha} + \boldsymbol{\nu}(i) \end{aligned} \tag{31}$$

where the third equality in (31) follows from the fact that $\mathbf{A}(i, \omega_0, \mathbf{h})$ is linear in $\mathbf{h}$, and the kth column of a $2MT \times 2MN$ real-valued matrix $\mathbf{B}(i, \omega_0, \mathbf{g}(i))$ is defined as

$$[\mathbf{B}(i, \omega_0, \mathbf{g}(i))]_k \triangleq \mathbf{A}(i, \omega_0, \mathbf{e}_k)\mathbf{g}(i). \tag{32}$$

Defining

$$\mathbf{r} \triangleq \begin{bmatrix} \mathbf{z}(1) \\ \mathbf{z}(2) \\ \vdots \\ \mathbf{z}(n_T) \end{bmatrix} \qquad \boldsymbol{\xi} \triangleq \begin{bmatrix} \boldsymbol{\nu}(1) \\ \boldsymbol{\nu}(2) \\ \vdots \\ \boldsymbol{\nu}(n_T) \end{bmatrix}$$

$$\mathbf{Q} \triangleq \begin{bmatrix} \mathbf{B}(1, \omega_0, \mathbf{g}(1))\mathbf{U} \\ \mathbf{B}(2, \omega_0, \mathbf{g}(2))\mathbf{U} \\ \vdots \\ \mathbf{B}(n_T, \omega_0, \mathbf{g}(n_T))\mathbf{U} \end{bmatrix}$$

we can rewrite (31) for all $i = 1, \dots, n_T$ as

$$\mathbf{r} = \mathbf{Q}\boldsymbol{\alpha} + \boldsymbol{\xi}. \tag{33}$$

[1]According to our simulations, the occurrence of having multiplicity of principal eigenvalue does not depend on the value of the CFO, but only depends on the structure of the underlying OSTBC.

[2]For the perfect CFO case, a semiblind approach to MIMO channel estimation has recently been developed in [18].

Using (33), the ML estimate of the vector $\boldsymbol{\alpha}$ can be written as

$$\hat{\boldsymbol{\alpha}} = (\mathbf{Q}^T\mathbf{Q})^{-1}\mathbf{Q}^T\mathbf{r}. \tag{34}$$

This estimate can be used to obtain the coefficients $\{\alpha_l\}_{l=1}^{n_o}$ from a few training symbols to resolve the ambiguity in the channel vector estimate.

To ensure that the ML estimate in (34) is unique, it is required that $2MTn_T \geq n_0$. However, for known nonidentifiable OSTBCs $n_0 = 4$ holds true and, therefore, as $T \geq 2$, the condition $2MTn_T \geq n_0$ is satisfied for any number of receive antennas.

The following lemma is useful for simplifying (34).

Lemma 1: The matrix $\mathbf{Q}$ satisfies the following property:

$$\mathbf{Q}^T\mathbf{Q} = \left(\sum_{n=1}^{n_T} \|\mathbf{g}(n)\|^2\right)\mathbf{I}_{n_o}. \tag{35}$$

Proof: See Appendix A. ■

The orthogonality property of $\mathbf{Q}$ that follows from this lemma can be used to reduce the complexity of computing the ML estimator in (34). Using (35), we can rewrite (34) as

$$\hat{\boldsymbol{\alpha}} = \frac{1}{\sum_{n=1}^{n_T} \|\mathbf{g}(n)\|^2}\mathbf{Q}^T\mathbf{r} \tag{36}$$

which enables to estimate $\hat{\boldsymbol{\alpha}}$ in a very simple way that is devoid of matrix inversion.

VI. Simulation Results

We have considered three scenarios with three different OSTBCs and different numbers of transmit and receive antennas. Throughout our simulations, the SNR is defined as σ_h^2/σ^2 where σ_h^2 is the variance of the elements of the channel matrix $\mathbf{H}$. In all figures except for Figs. 9 and 10, the elements of $\mathbf{H}$ are independently drawn in each run from a Gaussian distribution with variance σ_h^2 and are assumed to be fixed during that run (i.e., the channel remains constant over the number of data blocks that are used to estimate the CFO and the channel matrix).

In all examples, we compare our blind algorithm (or its semiblind modification) with the informed (clairvoyant) ML decoder that enjoys perfect knowledge of the MIMO channel and the CFO parameter. Note that the latter decoder does not correspond to any practical case, but it is used in our simulations as a benchmark. In all figures but Figs. 7 and 8, it is assumed that $\omega_o = 0.9$. Throughout our simulations, the QPSK modulation is used.

To quantify the performance of the methods tested in estimating the CFO parameter, we use the CFO estimation mean squared error (MSE)

$$\text{CFO-MSE} \triangleq \mathrm{E}\{|\hat{\omega}_o - \omega_o|^2\}. \tag{37}$$

To quantify the performance of the methods tested in terms of the channel estimation accuracy, we use the normalized MSE (NMSE) of channel estimates that is defined as

$$\text{C-NMSE} \triangleq \mathrm{E}\left\{\left\|\hat{\mathbf{h}} - \frac{\mathbf{h}}{\|\mathbf{h}\|}\right\|^2\right\}. \tag{38}$$

Note that in (38), we compare our channel estimate with the normalized true channel vector $\mathbf{h}/\|\mathbf{h}\|$ because our channel estimate $\hat{\mathbf{h}}$ in (28) corresponds to one of the orthonormal eigenvectors of $\boldsymbol{\Psi}(\hat{\omega}_o)$ and, therefore, the norm of $\hat{\mathbf{h}}$ is equal to one.

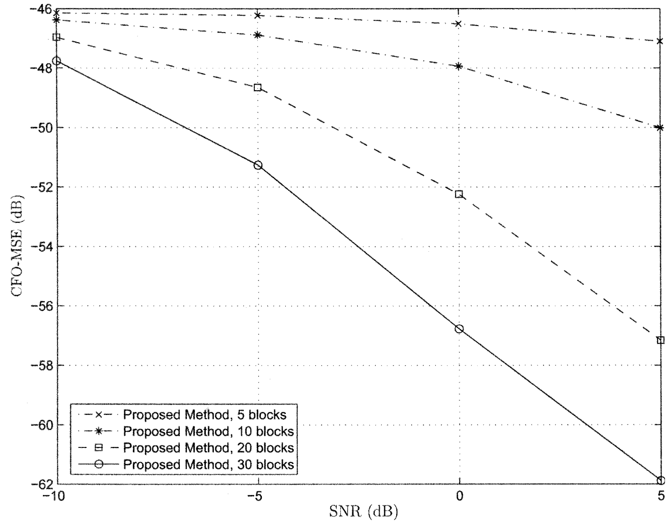


Fig. 1. CFO-MSE versus SNR for different values of n_B; first example.

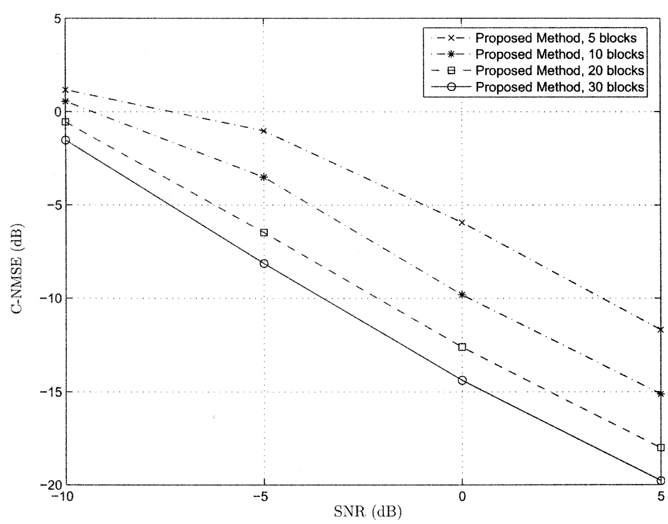


Fig. 2. C-NMSE versus SNR for different values of n_B; first example.

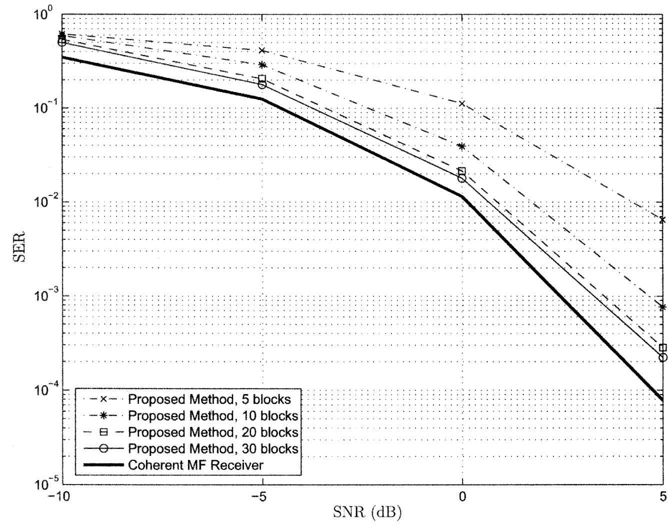


Fig. 3. SER versus SNR for different values of n_B; first example.

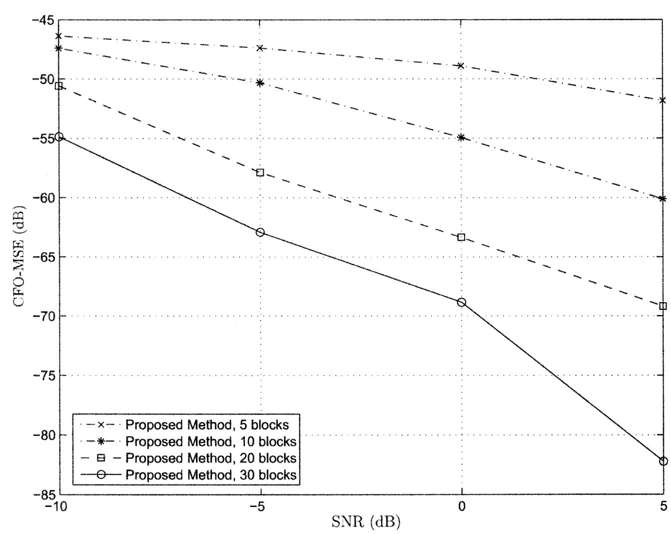


Fig. 4. CFO-MSE versus SNR for different values of n_B; second example.

In the first example, we assume the 3/4-rate OSTBC of [13] with $K = 3, T = N = 4$, and $M = 2$. As the largest eigenvalue of $\mathbf{\Psi}(\hat{\omega}_o)$ has no multiplicity for this code, the proposed blind estimator is compared with the informed receiver in this example. Figs. 1–3 show the CFO-MSE, C-NMSE, and symbol error rate (SER), respectively, versus SNR for different values of n_B.

In the second example, we assume the half-rate OSTBC of [3] with $K = 4, T = 8, N = 3$, and $M = 4$. Similar to the first example, the largest eigenvalue of $\mathbf{\Psi}(\hat{\omega}_o)$ has no multiplicity for this OSTBC and, therefore, the proposed blind receiver and the informed receiver are compared. Figs. 4–6 display the CFO-MSE, C-NMSE, and SER, respectively, versus SNR for different values of n_B. Figs. 7 and 8 show the CFO-MSE and C-NMSE, respectively, versus ω_0 for for $\mathrm{SNR} = 0$ dB and $n_B = 30$. Figs. 9 and 10 study the performance of the proposed method in the time-varying channel case. In contrast to all the other figures, in these two plots it is assumed that the channel fluctuates in each data block around its mean value with the variance σ_e^2. Figs. 9 and 10 display the CFO-MSE and C-NMSE, respectively, versus σ_e^2/σ_h^2 for $\mathrm{SNR} = 5$ dB and $\omega_0 = 0.9$.

As can be observed from Figs. 1–6, our blind channel and CFO estimator provides a very good performance that rapidly improves with increasing n_B or SNR. In particular, Figs. 1 and 4 demonstrate that the CFO estimates remain very accurate even for low values of SNR and/or n_B.

Figs. 3 and 6 show that the SER performance of our blind receiver approaches that of the informed ML receiver when increasing n_B. For instance, even for $n_B = 20$ the performance of our technique in both examples is approximately within 1 dB from that of the informed ML receiver.

Figs. 7 and 8 clearly demonstrate that the performance of the CFO and channel estimation in our method does not significantly depend on the value of ω_0.

From Figs. 9 and 10, it follows that even though the proposed method is developed under the assumption that the channel is fixed at the interval of n_B blocks, its performance remains acceptable in the time-varying channel case, provided that the parameter σ_e^2/σ_h^2 is sufficiently small.

In the last example, we consider the half-rate OSTBC of [3] with $K = 4, T = 8, N = 4$, and $M = 2$. This code is known to

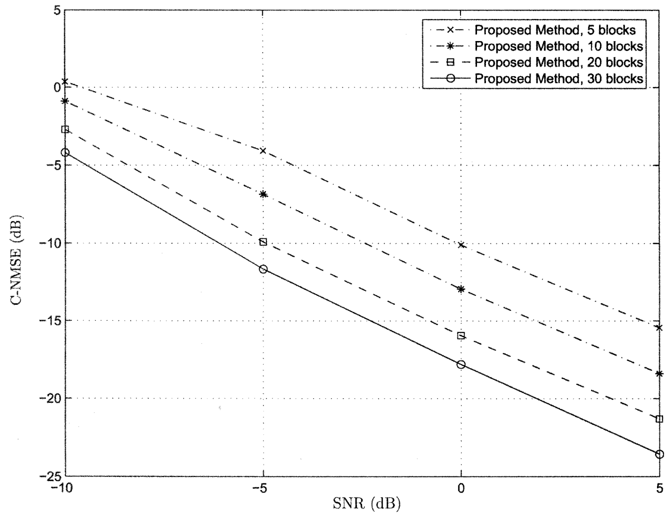


Fig. 5. C-NMSE versus SNR for different values of n_B; second example.

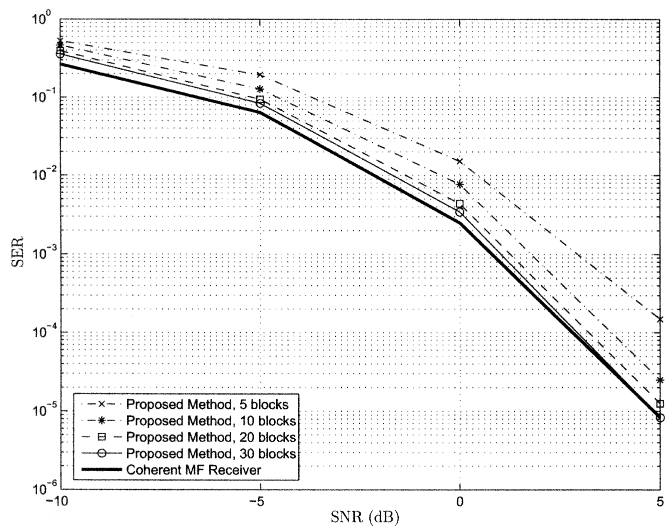


Fig. 6. SER versus SNR for different values of n_B; second example.

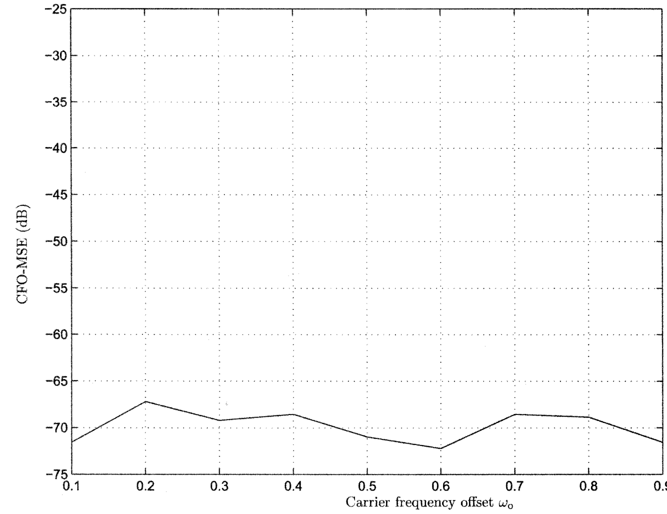


Fig. 7. CFO-MSE versus ω_o for $\mathrm{SNR} = 0$ dB and $n_B = 30$; second example.

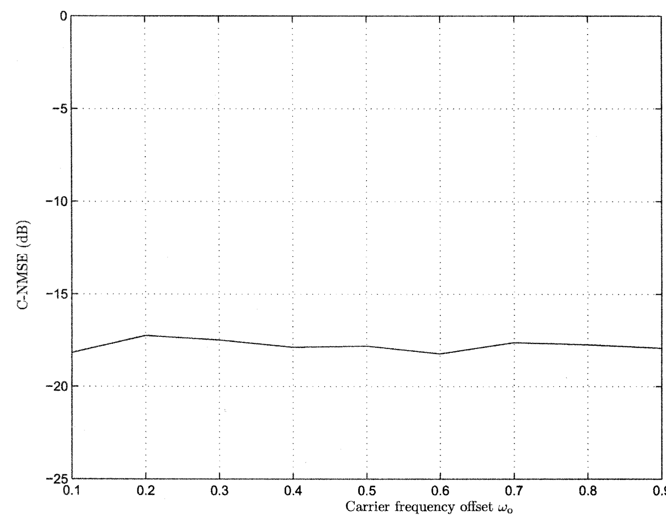


Fig. 8. C-NMSE versus ω_o for $\mathrm{SNR} = 0$ dB and $n_B = 30$; second example.

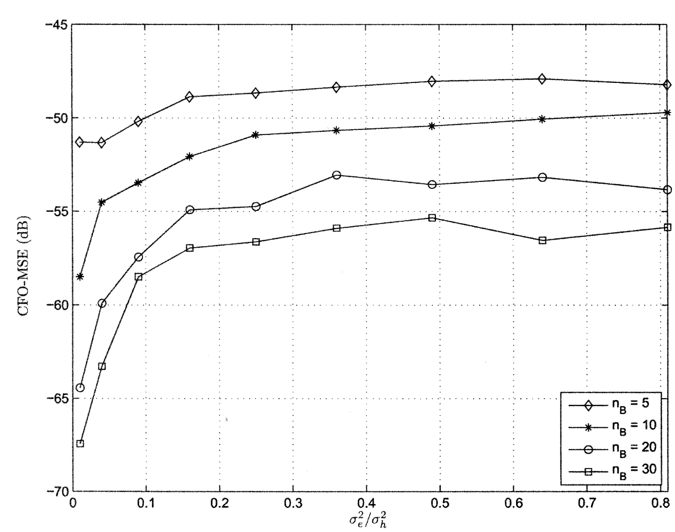


Fig. 9. CFO-MSE versus $\sigma_\epsilon^2/\sigma_h^2$ for different values of n_B; second example.

suffer from ambiguity [5] because the multiplicity of the principal eigenvalue of $\Psi(\omega_0)$ is equal to four ($n_0 = 4$). Therefore, the proposed blind channel estimation scheme is not applicable to this scenario and we have to resort to our semiblind method. We compare the latter method with a single training block to the training-based technique of [7]. Once the CFO parameter and the channel vector are estimated either using our semiblind technique or the method of [7], the resulting channel and CFO estimates are substituted to (20) in lieu of $\mathbf{h}$ and ω_o to obtain $\hat{\mathbf{g}}(n)$, and then the symbols are estimated using (29). Figs. 11–13 display the CFO-MSE, C-NMSE, and SER, respectively, versus SNR for different values of n_B.

It can be observed from the latter three figures that our semiblind technique substantially outperforms the method of [7]. As follows from Fig. 13, the SER performance of the latter approach is much worse than that of the informed ML receiver at high SNRs, whereas the SER penalty of the proposed semiblind estimator with respect to the informed receiver is quite moderate for any SNR value tested. For example, for $n_B = 30$ the latter SER penalty does not exceed 1.6 dB for any SNR value in this figure, while the penalty of the method of [7] may exceed 5 dB. These performance improvements over the technique of [7] can be explained by the fact that our approach is

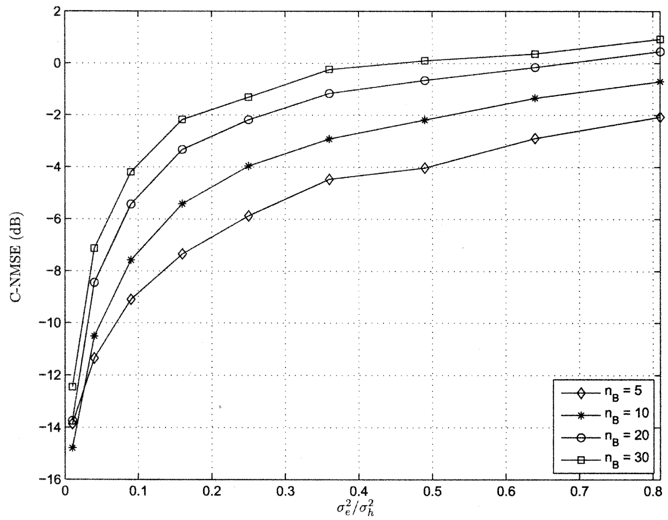


Fig. 10. C-NMSE versus σ_e^2/σ_h^2 for different values of n_B; second example.

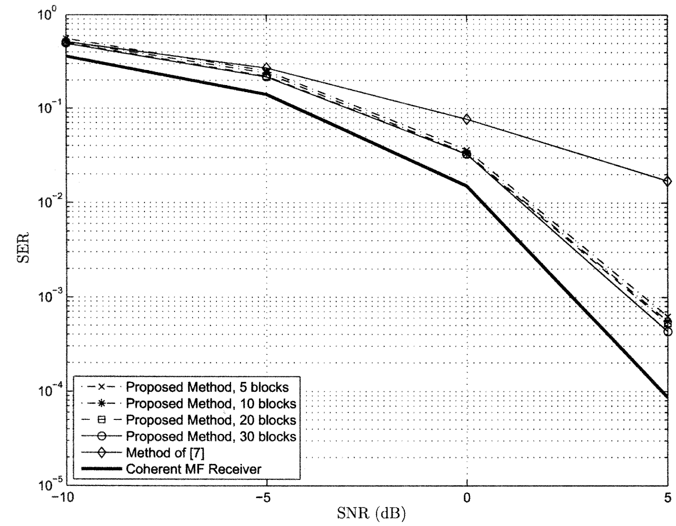


Fig. 13. SER versus SNR for different values of n_B; third example.

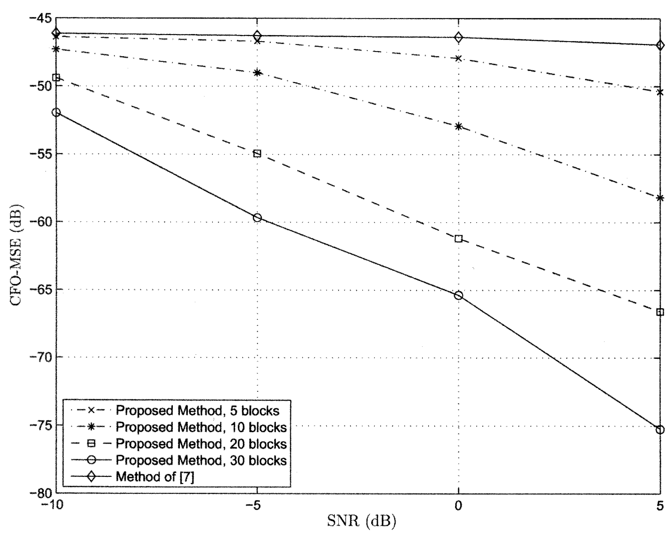


Fig. 11. CFO-MSE versus SNR for different values of n_B; third example.

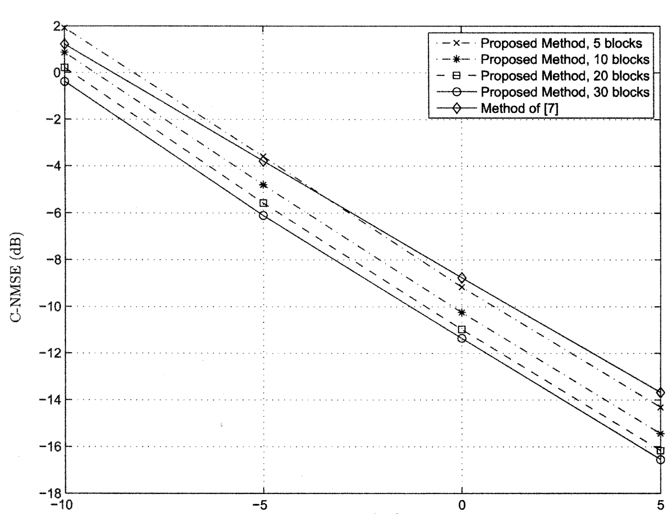


Fig. 12. C-NMSE versus SNR for different values of n_B; third example.

based on the channel model which is more parsimonious than that of [7]. Moreover, the proposed technique uses both the information-bearing and training data, whereas the method of [7] is based only on the training data.

VII. Conclusion

A new joint blind channel and carrier frequency offset estimator has been developed for orthogonally space-time block coded MIMO systems. The proposed technique is based on the relaxed maximum likelihood estimator that does not take into account the finite alphabet constraint. Although the proposed technique can be applied to the majority of orthogonal space-time codes, there are several codes that suffer from an intrinsic ambiguity in the joint channel, CFO, and symbol estimation. For such codes, a semiblind modification of our technique has been developed that estimates the "channel subspace" in a blind way, while extracting the channel matrix from this subspace using a few training symbols.

Appendix
Proof of Lemma 1

First, let us show that for any value of n, the columns of the matrix $\mathbf{B}(n, \omega_0, \mathbf{g}(n))$ are orthogonal. More specifically, we show that

$$\mathbf{B}^T(n, \omega_0, \mathbf{g}(n))\mathbf{B}(n, \omega_0, \mathbf{g}(n)) = \|\mathbf{g}(n)\|^2 \mathbf{I}_{2MN}. \quad (39)$$

To prove (39), let us recall that the matrix $\mathbf{A}(n, \omega_0, \mathbf{h})$ is linear in $\mathbf{h}$. Using this fact, we can write the lth column of $\mathbf{A}(n, \omega_0, \mathbf{h})$ as

$$[\mathbf{A}(n, \omega_0, \mathbf{h})]_l = \mathbf{\Phi}_l(n, \omega_0)\mathbf{h} \quad \text{for} \quad l = 1, \ldots, 2K$$

where for any fixed n and ω_0, $\mathbf{\Phi}_l(n, \omega_0)$ $(l = 1, \ldots, 2K)$ are $2MT \times 2MN$ real-valued matrices that satisfy:

$$\mathbf{\Phi}_l^T(n, \omega_0)\mathbf{\Phi}_m(n, \omega_0) = \begin{cases} \mathbf{I}_{2MN}, & \text{if } l = m, \\ -\mathbf{\Phi}_m^T(n, \omega_0)\mathbf{\Phi}_l(n, \omega_0), & \text{if } l \neq m. \end{cases} \quad (40)$$

To prove (40), we use the decoupling property (11) which implies that for any channel vector $\mathbf{h}$, and regardless of the values of n and ω_0

$$[\mathbf{A}(n,\omega_0,\mathbf{h})]_l^T[\mathbf{A}(n,\omega_0,\mathbf{h})]_l = \|\mathbf{h}\|^2 \tag{41}$$

or, equivalently

$$\mathbf{h}^T\mathbf{\Phi}_l^T(n,\omega_0)\mathbf{\Phi}_l(n,\omega_0)\mathbf{h} = \mathbf{h}^T\mathbf{h}. \tag{42}$$

As (42) holds true for any $\mathbf{h}$ and $\mathbf{\Phi}_l^T(n,\omega_0)\mathbf{\Phi}_l(n,\omega_0)$ is a symmetric matrix, we conclude that

$$\mathbf{\Phi}_l^T(n,\omega_0)\mathbf{\Phi}_l(n,\omega_0) = \mathbf{I}_{2MN}.$$

To prove the second part of (40), we use the fact that different columns of $\mathbf{A}(n,\omega_0,\mathbf{h})$ are orthogonal to each other, that is, for any $l \neq m$

$$\begin{aligned}[\mathbf{A}(n,\omega_0,\mathbf{h})]_l^T[\mathbf{A}(n,\omega_0,\mathbf{h})]_m &= \mathbf{h}^T\mathbf{\Phi}_l^T(n,\omega_0)\mathbf{\Phi}_m(n,\omega_0)\mathbf{h}\\ &= 0 \qquad (43)\\ [\mathbf{A}(n,\omega_0,\mathbf{h})]_m^T[\mathbf{A}(n,\omega_0,\mathbf{h})]_l &= \mathbf{h}^T\mathbf{\Phi}_m^T(n,\omega_0)\mathbf{\Phi}_l(n,\omega_0)\mathbf{h}\\ &= 0. \qquad (44)\end{aligned}$$

As the matrices $\mathbf{\Phi}_l^T(n,\omega_0)\mathbf{\Phi}_m(n,\omega_0)$ and $\mathbf{\Phi}_m^T(n,\omega_0)\mathbf{\Phi}_l(n,\omega_0)$ in (43) and (44) are not symmetric, the latter equations do not necessarily imply that $\mathbf{\Phi}_l^T(n,\omega_0)\mathbf{\Phi}_m(n,\omega_0) = \mathbf{\Phi}_m^T(n,\omega_0)\mathbf{\Phi}_l(n,\omega_0) = \mathbf{0}$ where $\mathbf{0}$ stands for a zero matrix of compatible dimension.

Adding (43) and (44) side-by-side yields

$$\mathbf{h}^T\left(\mathbf{\Phi}_l^T(n,\omega_0)\mathbf{\Phi}_m(n,\omega_0) + \mathbf{\Phi}_m^T(n,\omega_0)\mathbf{\Phi}_l(n,\omega_0)\right)\mathbf{h} = 0. \tag{45}$$

As (45) holds true for any vector $\mathbf{h}$ and $\mathbf{\Phi}_l^T(n,\omega_0)\mathbf{\Phi}_m(n,\omega_0) + \mathbf{\Phi}_m^T(n,\omega_0)\mathbf{\Phi}_l(n,\omega_0)$ is a symmetric matrix, we conclude that

$$\mathbf{\Phi}_l^T(n,\omega_0)\mathbf{\Phi}_m(n,\omega_0) = -\mathbf{\Phi}_m^T(n,\omega_0)\mathbf{\Phi}_l(n,\omega_0).$$

This completes the proof of (40). To prove (39), we note that the (l,m)th element of $\mathbf{B}^T(n,\omega_0,\mathbf{g}(n))\mathbf{B}(n,\omega_0,\mathbf{g}(n))$ is given by

$$\begin{aligned}&[\mathbf{B}^T(n,\omega_0,\mathbf{g}(n))\mathbf{B}(n,\omega_0,\mathbf{g}(n))]_{lm}\\ &\quad = \mathbf{g}^T(n)\mathbf{A}^T(n,\omega_0,\mathbf{e}_l)\mathbf{A}(n,\omega_0,\mathbf{e}_m)\mathbf{g}(n).\end{aligned} \tag{46}$$

Note that for $l = m$, the right-hand side (RHS) of (46) can be written as

$$\mathbf{g}^T(n)\underbrace{\mathbf{A}^T(n,\omega_0,\mathbf{e}_l)\mathbf{A}(n,\omega_0,\mathbf{e}_l)}_{\|\mathbf{e}_l\|^2\mathbf{I}_{2K}}\mathbf{g}(n) = \|\mathbf{g}(n)\|^2 \tag{47}$$

which follows from the decoupling property. Let us denote the rth element of $\mathbf{g}(1)$ as g_r. Then, for $l \neq m$, the RHS of (46) can be written as

$$\begin{aligned}&\mathbf{g}^T(n)\mathbf{A}^T(n,\omega_0,\mathbf{e}_l)\mathbf{A}(n,\omega_0,\mathbf{e}_m)\mathbf{g}(n)\\ &= \sum_{r=1}^{2K}\sum_{s=1}^{2K} g_r g_s\,[\mathbf{A}(n,\omega_0,\mathbf{e}_l)]_r^T[\mathbf{A}(n,\omega_0,\mathbf{e}_m)]_s\\ &= \sum_{r=1}^{2K}\sum_{\substack{s=1\\ s\neq r}}^{2K} g_r g_s \mathbf{e}_l^T\mathbf{\Phi}_r^T(n,\omega_0)\mathbf{\Phi}_s(n,\omega_0)\mathbf{e}_m\\ &\qquad + \sum_{r=1}^{2K}\underbrace{g_r^2\mathbf{e}_l^T\overbrace{\mathbf{\Phi}_r^T(n,\omega_0)\mathbf{\Phi}_r(n,\omega_0)}^{\mathbf{I}_{2MN}}\mathbf{e}_m}_{g_r^2\mathbf{e}_l^T\mathbf{e}_m=0}\\ &= \sum_{r=1}^{2K}\sum_{\substack{s=1\\ s\neq r}}^{2K} g_r g_s \mathbf{e}_l^T\mathbf{\Phi}_r^T(n,\omega_0)\mathbf{\Phi}_s(n,\omega_0)\mathbf{e}_m\\ &= -\sum_{r=1}^{2K}\sum_{\substack{s=1\\ s\neq r}}^{2K} g_r g_s \mathbf{e}_l^T\mathbf{\Phi}_s^T(n,\omega_0)\mathbf{\Phi}_r(n,\omega_0)\mathbf{e}_m\\ &= -\sum_{r=1}^{2K}\sum_{s=1}^{2K} g_r g_s\,[\mathbf{A}(n,\omega_0,\mathbf{e}_l)]_s^T[\mathbf{A}(n,\omega_0,\mathbf{e}_m)]_r\\ &= -\,\mathbf{g}^T(n)\mathbf{A}^T(n,\omega_0,\mathbf{e}_l)\mathbf{A}(n,\omega_0,\mathbf{e}_m)\mathbf{g}(n)\end{aligned} \tag{48}$$

where (40) has been used. Therefore, we obtain that for $l \neq m$

$$\mathbf{g}^T(n)\mathbf{A}^T(n,\omega_0,\mathbf{e}_l)\mathbf{A}(n,\omega_0,\mathbf{e}_m)\mathbf{g}(n) = 0 \tag{49}$$

holds true and the proof of (39) is complete. Using (39), we can write

$$\begin{aligned}\mathbf{Q}^T\mathbf{Q} &= \sum_{n=1}^{n_T}\mathbf{U}^T\mathbf{B}^T(n,\hat{\omega}_0,\mathbf{g}(n))\mathbf{B}(n,\hat{\omega}_0,\mathbf{g}(n))\mathbf{U}\\ &= \sum_{n=1}^{n_T}\|\mathbf{g}(n)\|^2\mathbf{U}^T\mathbf{U}\\ &= \sum_{n=1}^{n_T}\|\mathbf{g}(n)\|^2\mathbf{I}_{n_o}.\end{aligned} \tag{50}$$

With (50), Lemma 1 is proven. ■

References

[1] V. Tarokh, N. Seshadri, and A. R. Calderbank, "Space-time codes for high data rates wireless communication: Performance criterion and code construction," *IEEE Trans. Inf. Theory*, vol. 44, no. 3, pp. 744–765, Mar. 1998.

[2] S. M. Alamouti, "A simple transmit diversity technique for wireless communications," *IEEE J. Sel. Areas Commun.*, vol. 45, no. 8, pp. 1451–1458, Oct. 1998.

[3] V. Tarokh, H. Jafarkhani, and A. R. Calderbank, "Space-time block codes from orthogonal designs," *IEEE Trans. Inf. Theory*, vol. 45, no. 7, pp. 1456–1467, Jul. 1999.

[4] N. Ammar and Z. Ding, "On blind channel identifiability under space-time coded transmission," in *Proc. Asilomar Conf. Signals, Syst. Comp.*, Pacific Grove, CA, Nov. 2002, vol. 1, pp. 664–668.

[5] S. Shahbazpanahi, A. B. Gershman, and J. H. Manton, "Closed-form blind MIMO channel estimation for orthogonal space-time block codes," *IEEE Trans. Signal Process.*, vol. 53, no. 12, pp. 4506–4517, Dec. 2005.

[6] M. Gharavi-Alkhansari and A. B. Gershman, "Constellation space invariance of orthogonal space-time block codes," *IEEE Trans. Inf. Theory*, vol. 51, no. 1, pp. 331–334, Jan. 2005.

[7] O. Besson and P. Stoica, "On parameter estimation of MIMO flat-fading channels with frequency offsets," *IEEE Trans. Signal Process.*, vol. 51, no. 3, pp. 602–613, Mar. 2003.

[8] Y. Yao and G. Giannakis, "Blind carrier frequency offset estimation in SISO, MIMO and multiuser OFDM systems," *IEEE Trans. Commun.*, vol. 53, no. 1, pp. 173–183, Jan. 2005.

[9] X. Ma, M.-K. Oh, G. B. Giannakis, and D.-J. Park, "Hopping pilots for estimation of frequency-offsets and multi-antenna channels in MIMO-OFDM," *IEEE Trans. Commun.*, vol. 53, no. 1, pp. 162–172, Jan. 2005.

[10] F. Gini and G. B. Giannakis, "Frequency offset and symbol timing recovery in flat fading channels: A cyclostationary approach," *IEEE Trans. Commun.*, vol. 46, no. 3, pp. 400–411, Mar. 1998.

[11] S. Shahbazpanahi, A. B. Gershman, and G. B. Giannakis, "Joint blind channel and carrier frequency offset estimation in orthogonally space-time block coded MIMO systems," in *Proc. IEEE Workshop Signal Process. Adv. Wireless Commun.*, New York, Jun. 2005, pp. 363–367.

[12] S. Shahbazpanahi, A. B. Gershman, and G. B. Giannakis, "Semiblind channel and carrier frequency-offset parameter estimation for orthogonally space-time block coded MIMO systems," in *Proc. ICASSP'07*, Honolulu, HI, Apr. 2007.

[13] E. G. Larsson and P. Stoica, *Space-Time Block Coding for Wireless Communications*. Cambridge, U.K.: Cambridge Univ. Press, 2003.

[14] B. Hassibi and B. M. Hochwald, "High-rate codes that are linear in space and time," *IEEE Trans. Inf. Theory*, vol. 48, no. 7, pp. 1804–1824, Jul. 2002.

[15] S. Shahbazpanahi, M. Beheshti, A. B. Gershman, M. Gharavi-Alkhansari, and K. M. Wong, "Minimum variance linear receivers for multiaccess MIMO wireless systems with space-time block coding," *IEEE Trans. Signal Process.*, vol. 52, no. 12, pp. 3306–3313, Dec. 2004.

[16] S. Talwar, M. Viberg, and A. Paulraj, "Blind separation of synchronous co-channel digital signals using an antenna array — Part I: Algorithms," *IEEE Trans. Signal Process.*, vol. 44, no. 5, pp. 1184–1197, May 1996.

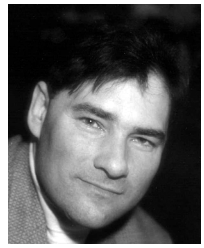
[17] X. Ma, C. Tepedelenlioglu, G. B. Giannakis, and S. Barbarossa, "Non-data-aided carrier offset estimators for OFDM with null subcarriers: Identifiability, algorithms, and performance," *IEEE J. Sel. Areas Commun.*, vol. 19, pp. 2504–2515, Dec. 2001.

[18] S. Shahbazpanahi, A. B. Gershman, and G. B. Giannakis, "Semiblind multi-user MIMO channel estimation using Capon and MUSIC techniques," *IEEE Trans. Signal Process.*, vol. 54, pp. 3581–3591, Sep. 2006.

Shahram Shahbazpanahi (M'02) was born in Sanandaj, Kurdistan, Iran. He received the B.Sc., M.Sc., and Ph.D. degrees in electrical engineering from Sharif University of Technology, Tehran, Iran, in 1992, 1994, and 2001, respectively.

From September 1994 to September 1996, he was a faculty member with the Department of Electrical Engineering, Razi University, Kermanshah, Iran. From July 2001 to March 2003, he was a Postdoctoral Fellow with the Department of Electrical and Computer Engineering, McMaster University, Hamilton, Ontario, Canada. From April 2003 to September 2004, he was a Visiting Researcher with the Department of Communication Systems, University of Duisburg-Essen, Duisburg, Germany. From September 2004 to April 2005, he was a Lecturer and Adjunct Professor with the Department of Electrical and Computer Engineering, McMaster University. Since July 2005, he has been with the Faculty of Engineering and Applied Science, University of Ontario Institute of Technology, Oshawa, Canada, where he holds an Assistant Professor position. His research interests include statistical and array signal processing, space-time adaptive processing, detection and estimation, smart antennas, spread spectrum techniques, MIMO communications, DSP programming, and hardware/real-time software design for telecommunication systems.

Dr. Shahbazpanahi is currently serving as Associate Editor for the IEEE TRANSACTIONS ON SIGNAL PROCESSING and IEEE SIGNAL PROCESSING LETTERS.



Alex B. Gershman (M'97–SM'98–F'06) received the Diploma and Ph.D. degrees in radiophysics from the Nizhny Novgorod State University, Russia, in 1984 and 1990, respectively.

From 1984 to 1999, he held several full-time and visiting research appointments in Russia, Switzerland, and Germany. In 1999, he joined the Department of Electrical and Computer Engineering, McMaster University, Hamilton, Ontario, Canada, where he became a professor in 2002. From April 2005, he has been with the Darmstadt University of Technology, Darmstadt, Germany, as a professor of Communication Systems. His research interests are in the area of signal processing and communications with the primary emphasis on array processing, beamforming, MIMO and multiuser communications, and estimation and detection theory.

Dr. Gershman is the recipient of several awards including the 2004 IEEE Signal Processing Society (SPS) Best Paper Award; the IEEE Aerospace and Electronic Systems Society (AESS) Barry Carlton Award for the best paper published in 2004, the 2002 Young Explorers Prize from the Canadian Institute for Advanced Research (CIAR); the 2001 Wolfgang Paul Award from the Alexander von Humboldt Foundation, Germany; and the 2000 Premier's Research Excellence Award, Ontario, Canada. He also coauthored the paper that received the 2005 IEEE SPS Young Author Best Paper Award. He is currently Editor-in-Chief for the IEEE SIGNAL PROCESSING LETTERS, Chair of the Sensor Array and Multichannel (SAM) Technical Committee of the IEEE SPS, and a member of the IEEE SPS Publications and Technical Directions Boards. He was an Associate Editor of the IEEE TRANSACTIONS ON SIGNAL PROCESSING (1999–2005); Technical Co-Chair of the IEEE International Symposium on Signal Processing and Information Technology (ISSPIT), Darmstadt, Germany, December 2003; Technical Co-Chair of the 4th IEEE Sensor Array and Multichannel Signal Processing Workshop, Waltham, MA, June 2006; General Co-Chair of the 1st IEEE Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP), Puerto Vallarta, Mexico, December 2005, and Tutorial Chair of EUSIPCO, Florence, Italy, September 2006. He is General Co-Chair of the 5th IEEE Sensor Array and Multichannel Signal Processing Workshop, Darmstadt, Germany, July 2008.

Georgios B. Giannakis (F'97) received the Diploma in electrical engineering from the National Technical University of Athens, Greece, in 1981. From September 1982 to July 1986 he was with the University of Southern California (USC), where he received the M.Sc. degree in electrical engineering in 1983, the M.Sc. degree in mathematics in 1986, and the Ph.D. degree in electrical engineering in 1986.

After lecturing for one year at USC, he joined the University of Virginia in 1987, where he became a professor of Electrical Engineering in 1997. Since 1999, he has been a professor with the Department of Electrical and Computer Engineering, University of Minnesota, where he now holds an ADC Chair in Wireless Telecommunications. His general interests span the areas of communications and signal processing, estimation and detection theory, time-series analysis, and system identification—subjects on which he has published more than 220 journal papers, 380 conference papers, and two edited books. Current research focuses on transmitter and receiver diversity techniques for single- and multiuser fading communication channels, complex-field and space-time coding, multicarrier, ultra-wideband wireless communication systems, cross-layer designs, and sensor networks.

Dr. Giannakis is the (co-)recipient of six paper awards from the IEEE Signal Processing and Communications Societies (1992, 1998, 2000, 2001, 2003, 2004). He also received Technical Achievement Awards from the Signal Processing Society in 2000 and from EURASIP in 2005. He served as Editor in Chief for the IEEE SIGNAL PROCESSING LETTERS, as Associate Editor for the IEEE TRANSACTIONS ON SIGNAL PROCESSING and the IEEE SIGNAL PROCESSING LETTERS, as secretary of the Signal Processing Conference Board, as member of the Signal Processing Publications Board, as a member and vice-chair of the Statistical Signal and Array Processing (SSAP) Technical Committee (TC), as chair of the Signal Processing for Communications (SPCOM) TC, and as a member of the IEEE Fellows Election Committee. He has also served as a member of the IEEE Signal Processing Society's Board of Governors, the Editorial Board for the PROCEEDINGS OF THE IEEE and the steering committee of the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS.