

# Transmit-Antennae Space–Time Block Coding for Generalized OFDM in the Presence of Unknown Multipath

Zhiqiang Liu, Georgios B. Giannakis, *Fellow, IEEE*, Sergio Barbarossa, *Member, IEEE*, and Anna Scaglione, *Member, IEEE*

**Abstract**—Transmit antenna diversity has been exploited recently to develop high-performance space–time coders and simple maximum-likelihood decoders for transmissions over flat fading channels. Relying on block precoding, this paper develops generalized space–time coded multicarrier transceivers appropriate for wireless propagation over frequency-selective multipath channels. Multicarrier precoding maps the frequency-selective channel into a set of flat fading subchannels, whereas space–time encoding/decoding facilitates equalization and achieves performance gains by exploiting the diversity available with multiple transmit antennas. When channel state information is unknown at the receiver, it is acquired blindly based on a deterministic variant of the constant-modulus algorithm that exploits the structure of space–time block codes. To benchmark performance, the Cramér–Rao bound of channel estimates is also derived. System performance is evaluated both analytically and with simulations.

**Index Terms**—Blind channel estimation, multipath fading channels, space–time coding, transmit diversity.

## I. INTRODUCTION

**I**N RECENT YEARS, space–time (ST) coding has gained much attention as an effective transmit diversity technique to combat fading in wireless communications (see e.g., [8], [12], and references therein). ST coding relies on multi-antennae transmissions that are combined with appropriate signal processing at the receiver to provide diversity gain. ST trellis codes were first proposed in [17] to achieve maximum diversity and coding gains. However, for a fixed number of transmit antennas, their decoding complexity at the receiver increases exponentially with the transmission rate. To reduce decoding complexity, orthogonal ST block codes with two transmit antennas were first introduced in [2] and later generalized to an arbitrary number of transmit antennas in [16]. An attractive property of ST block codes is that maximum-likelihood (ML) decoding can be performed using only *linear* processing. For

complex constellations, ST block coding with two transmit antennas is the only block coding that provides full diversity without loss of transmission rate [16].

ST codes were originally designed for slow flat fading channels. Applications of ST codes to dispersive channels were dealt with in [1], [5], and [6] for orthogonal frequency division multiplexing (OFDM) systems<sup>1</sup>. Channel knowledge is assumed available at the receiver in [1] and is acquired through training [5], [6]. However, it is important to remark that ST decoding requires multichannel state information at the receiver. Thus, the achievable diversity gain comes at the price of proportional increase in the amount of training, which incurs efficiency loss especially in a rapidly varying environment. This motivates looking for receivers with blind channel estimation capabilities in the context of ST transmissions.

Toward this objective, we propose a novel ST-generalized OFDM (ST-GOFDM) transceiver. We consider a system with two transmit antennas and one receive antenna. Relying on symbol blocking, the ST block codes of [2] are incorporated into a generalized OFDM (GOFDM) transmitter [13], [14] to achieve transmit diversity in frequency-selective propagation. Different from ST-OFDM schemes in [1], [5], and [6] where performance suffers from consistent fading effects caused by common (or near common) channels zeros, ST-GOFDM offers additional robustness through what we term root hopping. In addition to performance improvement, the decoding simplicity of the ST block codes of [2] is retained in our system, even when communicating over multipath channels. By exploiting the special structure of ST block codes, we also develop a blind channel estimator based on a deterministic variant of the constant-modulus algorithm (CMA) [19]. Unlike [19], our ST-coded CMA is computationally simple. To benchmark the accuracy of our estimation algorithm, we also derive a closed form expression for the Cramér–Rao bound (CRB), assuming additive white Gaussian noise (AWGN) and modeling the unknown transmitted symbols as deterministic (nuisance) parameters. As a byproduct of our CRB derivation, we prove that imposing constraints on transmitted symbols is indispensable for blind channel estimation.

The paper is organized as follows. In Section II, we describe the system model; while in Section III, we present our transceiver design. The blind channel estimation algorithm is de-

Manuscript received October 1, 1999; revised October 5, 2000. This work was supported by the Army Research Office under Grant DAAG55-98-1-0336.

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

S. Barbarossa is with the Information and Communication Department, University of Roma “La Sapienza”, 00184 Rome, Italy (e-mail: sergio@infocom.ing.uniroma1.it).

A. Scaglione was with the Department of Electrical and Computer Engineering, University of Minnesota. She is now with the Electrical and Computer Engineering Department, University of New Mexico, Albuquerque, NM 87131 USA (e-mail: annas@ece.unm.edu).

Publisher Item Identifier S 0733-8716(01)04716-3.

<sup>1</sup>ST trellis codes were used in [1], while delay and permutation transmit diversity schemes were adopted in [5] and [6].

veloped in Section IV and its CRB based performance in Section V. Bit-error-rate (BER) performance is analyzed and extensive simulations are presented in Section VI. Section VII concludes this paper.

*Notation:* Column vectors (matrices) are denoted by boldface lower (upper) case letters. The superscripts  $\text{T}$ ,  $*$ ,  $\mathcal{H}$ , and  $\dagger$  stand for transpose, complex conjugate, complex conjugate transpose and matrix pseudoinverse, respectively.  $\mathbf{0}_{M \times N}$  denotes the  $M \times N$  matrix with all entries zero and  $\mathbf{I}_{M \times M}$  the  $M \times M$  identity matrix.

## II. SYSTEM MODELING

Consider a wireless communication system equipped with two transmit antennas and a single receive antenna. Fig. 1 represents the discrete-time equivalent *baseband* model of our ST-GOFDM transceiver. Similar to conventional OFDM, the information symbol sequence  $s(n)$  is parsed in blocks  $\mathbf{s}(n) := (s(nK), \dots, s(nK + K - 1))^\text{T}$  of size  $K \times 1$ . Our ST encoder maps every two consecutive symbol blocks  $\mathbf{s}(2n)$  and  $\mathbf{s}(2n + 1)$  to the following  $2K \times 2$  matrix:

$$\begin{pmatrix} \mathbf{s}(2n) & -\mathbf{s}^*(2n + 1) \\ \mathbf{s}(2n + 1) & \mathbf{s}^*(2n) \end{pmatrix} \begin{matrix} \rightarrow \text{time} \\ \downarrow \text{space} \end{matrix} \quad (1)$$

whose columns are transmitted in successive time intervals with the upper and lower blocks in a given column sent simultaneously through the first and second transmit antenna, respectively. Note that with  $K = 1$ , no blocking takes place, and (1) then reduces to the orthogonal ST block codes with two transmit antennas of [2]. However, as we will discuss in Section III, blocking is instrumental in developing *simple* ST coding/decoding algorithms for frequency selective channels. Prior to transmission, the  $K$ -long symbol block in each transmit antenna branch is mapped onto a  $P$ -long block with  $P > K$  through a redundant precoder described by the tall  $P \times K$  matrix  $\mathbf{C}$ . The precoder  $\mathbf{C}$  will convert frequency-selective fading channels to flat ones and will enable us to exploit the transmit diversity built by the ST mapping in (1). After parallel-to-serial (P/S) conversion and modulation (not shown in Fig. 1), the transmitted symbols from the  $i$ th transmit antenna propagate in successive time intervals  $T$  through a frequency-selective channel which is denoted by  $T$ -sampled impulse response vectors  $\mathbf{h}_i := (h_i(0), \dots, h_i(L))^\text{T}$ ,  $i = 1, 2$ . Each channel's impulse response includes transmit-receive (pulse shaping) filters, multipath, and relative delay between the two antennas. We will assume the following.

- a1) The two frequency-selective channels are finite impulse response (FIR) and an upper bound  $L$  on their orders is assumed available; i.e.,  $h_i(l) = 0$ ,  $\forall l \notin [0, L]$ ,  $i = 1, 2$ . Thus, each channel in Fig. 1 can be described by a  $P \times P$  Toeplitz convolution matrix  $\mathbf{H}_i$ , with  $(k, l)$ th entry  $h_i(k - l)$ ,  $i = 1, 2$ .
- a2) The last  $L = P - K$  rows of  $\mathbf{C}$  are set to zero; i.e., with  $c_k(p)$  denoting the  $(p, k)$ th entry of  $\mathbf{C}$ , we assume  $c_k(p) = 0$  for  $p = K + 1, \dots, P$ . This corresponds to padding our transmitted blocks with  $L$  trailing zeros.

Thanks to Assumption a2), the channel-induced interblock interference (IBI) is avoided [13], [14] and one can focus at each

received block separately. Alternately, we could have achieved IBI-free reception by inserting an  $L$ -long cyclic prefix to each transmitted block (similar to OFDM [1], [5]), and discarding it at the receiver. Either way, after S/P conversion (not shown in Fig. 1), each pair of two consecutive  $P$ -long received blocks  $\mathbf{y}(2n)$  and  $\mathbf{y}(2n + 1)$ , with  $\mathbf{y}(n) := (y(nP), \dots, y(nP + P - 1))^\text{T}$ , is given by

$$\begin{aligned} \mathbf{y}(2n) &= \mathbf{H}_1 \mathbf{C} \mathbf{s}(2n) + \mathbf{H}_2 \mathbf{C} \mathbf{s}(2n + 1) + \mathbf{w}(2n) \\ \mathbf{y}(2n + 1) &= -\mathbf{H}_1 \mathbf{C} \mathbf{s}^*(2n + 1) + \mathbf{H}_2 \mathbf{C} \mathbf{s}^*(2n) \\ &\quad + \mathbf{w}(2n + 1) \end{aligned} \quad (2)$$

where  $\mathbf{w}(n) := (w(nP), \dots, w(nP + P - 1))^\text{T}$  denotes AWGN. Based on the received data model (2), we will design next our transceivers to achieve transmit diversity gain in frequency selective channels.

## III. TRANSCIEVER DESIGN

Starting from the received data model (2), we wish to design the precoding matrix  $\mathbf{C}$  and the receiver matrices  $\mathbf{V}$  and  $\{\mathbf{G}_k\}_{k=1}^K$  (see Fig. 1) to recover the information block  $\mathbf{s}(n)$  from  $\mathbf{y}(n)$  with transmit diversity gain. We consider that the FIR channels are unknown to the transmitter which is always the case when feedback channels are not used. The two channels have to be estimated at the receiver, but only in this section, we will suppose that perfect channel state information (CSI) is available at the receiver. A blind channel estimator will be developed in the next section.

Orthogonal ST block codes with two transmit antennas were possible to design for flat fading channels [2]. However, multipath manifests itself in convolving ST-coded transmissions with FIR channels, and thereby destroys code orthogonality. Without an orthogonal code structure, ST decoding becomes extremely complex. Observe though that channel convolution becomes multiplication in the  $\mathcal{Z}$  domain where the code orthogonality can be retained. Our basic idea behind ST-GOFDM is to implement ST coding/decoding in the  $\mathcal{Z}$  domain by judicious design of the matrices  $\mathbf{C}$ ,  $\mathbf{V}$  and  $\{\mathbf{G}_k\}_{k=1}^K$ . Unlike ST-OFDM schemes in [1] and [5] which rely on FFT/IFFT operations and achieve diversity gain at fixed points in the frequency domain, ST-GOFDM resorts to  $\mathcal{Z}$ -transforms and offers some advantages over ST-OFDM as we will discuss later. It will also be shown that ST-OFDM is subsumed by ST-GOFDM if  $\mathbf{C}$  and  $\mathbf{V}$  are chosen appropriately.

Define the  $\mathcal{Z}$ -transforms:  $Y(n; z) = \sum_{p=0}^{P-1} y(nP + p)z^{-p}$ ;  $C_k(z) := \sum_{p=0}^{P-1} c_k(p + 1)z^{-p}$ ;  $H_i(z) = \sum_{l=0}^L h_i(l)z^{-l}$ ;  $W(n; z) = \sum_{p=0}^{P-1} w(nP + p)z^{-p}$ . Let us focus on the block  $\mathbf{y}(2n)$  in (2) first, and  $\mathcal{Z}$ -transform its entries to obtain

$$\begin{aligned} Y(2n; z) &= \sum_{\mu=1}^K [H_1(z)C_\mu(z)s(2nK + \mu - 1) \\ &\quad + H_2(z)C_\mu(z)s((2n + 1)K + \mu - 1)] \\ &\quad + W(2n; z). \end{aligned} \quad (3)$$

We choose  $K$  distinct points  $\{\rho_k\}_{k=1}^K$  on the complex plane. Our goal is to seek code polynomials  $C_\mu(z)$ ,  $\mu = 1, \dots, K$ , such that for each  $k$ ,  $Y(2n; \rho_k)$  contains the contribution from

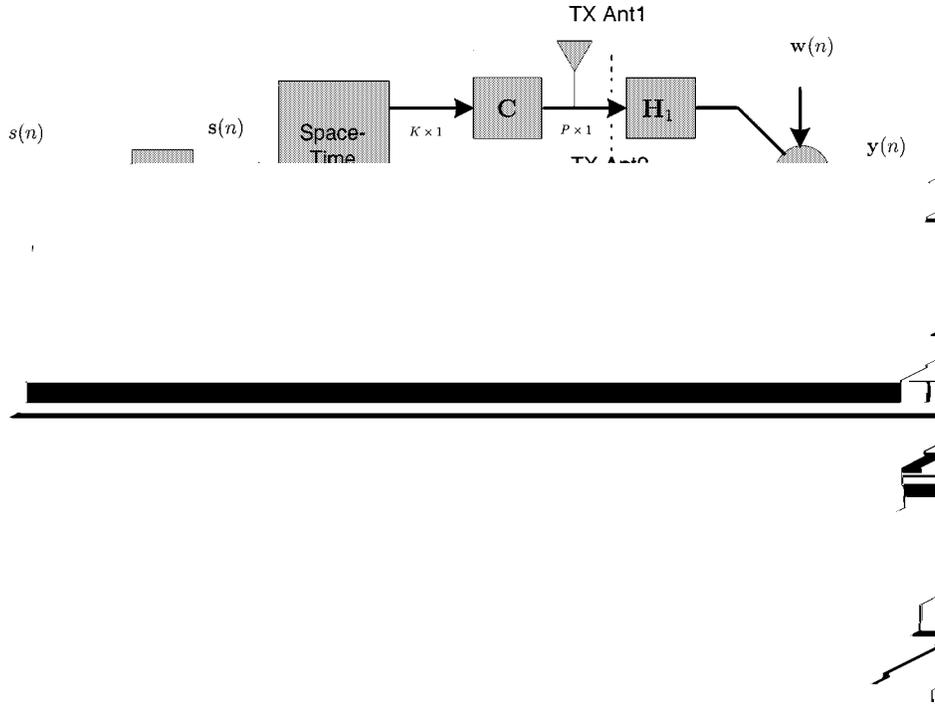


Fig. 1. Discrete-time equivalent baseband system model.

the  $k$ th summand in (3), namely from  $s(2nK + k - 1)$  and  $s((2n + 1)K + k - 1)$  only, regardless of  $H_1(z)$  and  $H_2(z)$ . To achieve this goal, we design  $C_\mu(\rho_k)$  to satisfy

$$C_\mu(\rho_k) = A\delta(k - \mu), \quad \forall \mu, k \in [1, K] \quad (4)$$

where  $A$  is a constant chosen to impose the transmission power constraint. For fixed  $\mu$ , (4) prescribes  $C_\mu(z)$  at  $K$  points  $\rho_k$ . Thus, polynomials  $C_\mu(z)$  satisfying (4) should have degree  $\deg[C_\mu(z)] \geq K - 1$ . When  $\deg[C_\mu(z)] = K - 1$ , the polynomial  $C_\mu(z)$  can be uniquely determined by Lagrange interpolation through the points  $\{\rho_k\}_{k=1}^K$  as

$$C_\mu(z) := \sum_{p=0}^{P-1} c_k(p+1)z^{-p} = A \frac{\prod_{k=1, k \neq \mu}^K (1 - \rho_k z^{-1})}{\prod_{k=1, k \neq \mu}^K (1 - \rho_k \rho_\mu^{-1})}. \quad (5)$$

The code construction as a Lagrange interpolating polynomial through the constraints in (4) offers a novel ST counterpart of [4]. Different from our symbol separating constraints (4), [4] utilizes constraints for multiuser and intersymbol interference elimination in a multiaccess/single-antenna setting (see also [13] and [14] where the Lagrange codes (5) offered GOFDM codes in a multiuser/single-antenna context).

Based on (5) and taking into account the  $L$  trailing zeros of Assumption a2), the code length is  $P = K + L$ . Since we deal with transmissions of  $K$ -long blocks, the bandwidth efficiency of our system is

$$\eta := \frac{K}{K + L}. \quad (6)$$

Note that for sufficiently large  $K$ , we have  $\eta \simeq 1$ ; hence, bandwidth is not overexpanded.

The design of the precoding matrix  $\mathbf{C}$  is specified in the  $\mathcal{Z}$  domain and depends on the selection of points  $\{\rho_k\}_{k=1}^K$  as in

(5). Let us consider a particular choice of the points  $\{\rho_k\}_{k=1}^K$ . Specifically, let  $\rho_k$ s be chosen regularly spaced around the unit circle on the complex plane on an FFT grid

$$\rho_k = e^{j(2\pi/K)(k-1)}, \quad \forall k \in [1, K]. \quad (7)$$

Plugging (7) into (5) reveals that  $c_\mu(p) = \sqrt{K} \exp(j2\pi\mu p/K)$ ,  $p = 1, \dots, K$ . Hence, the matrix  $\mathbf{C}$  corresponding to this selection becomes  $\mathbf{C} = \sqrt{K} \mathbf{T}_{zp} \mathbf{F}$  where  $\mathbf{F}$  is the  $K \times K$  inverse discrete Fourier transform (IDFT) matrix and  $\mathbf{T}_{zp} = [\mathbf{I}_K^T \mathbf{0}_{K \times L}^T]^T$  represents the matrix operator for padding the  $L$  trailing zeros.

Using the codes in (5) that satisfy (4), and evaluating (3) at  $z = \rho_k$ , we obtain  $\forall k \in [1, K]$

$$Y(2n; \rho_k) = AH_1(\rho_k)s(2nK + k - 1) + AH_2(\rho_k) \times s((2n + 1)K + k - 1) + W(2n; \rho_k). \quad (8)$$

Equation (8) confirms that only  $s(2nK + k - 1)$  and  $s((2n + 1)K + k - 1)$  contribute to  $Y(2n; \rho_k)$ . Similarly, following the steps to arrive at (8), we  $\mathcal{Z}$ -transform the block  $\mathbf{y}(2n + 1)$  in (2) and then use the condition (4) to obtain  $\forall k \in [1, K]$ ,

$$Y(2n + 1; \rho_k) = -AH_1(\rho_k)s^*((2n + 1)K + k - 1) + AH_2(\rho_k)s^*(2nK + k - 1) + W(2n + 1; \rho_k). \quad (9)$$

To express the  $\mathcal{Z}$ -transforms evaluated at  $z = \rho_k$  in (8) and (9), let the  $P \times 1$  Vandermonde vector  $\mathbf{v}(\rho, P)$  built from the complex constant  $\rho$  as  $\mathbf{v}(\rho, P) := (1, \rho^{-1}, \dots, \rho^{-(P-1)})^T$ . With  $z$  replacing  $\rho$ , the  $\mathcal{Z}$ -transform of any  $P \times 1$  vector  $\mathbf{x}(n)$  can be represented by  $X(n; z) = \mathbf{v}^T(z, P)\mathbf{x}(n)$ . We can thus express  $\{Y(n; \rho_k)\}_{k=1}^K$  as outputs of a filterbank composed of  $K$

parallel FIR filters each of length  $P$ , whose coefficients are the rows of the  $K \times P$  Vandermonde matrix (see also Fig. 1)

$$\mathbf{V} := (v(\rho_1, P), \dots, v(\rho_K, P))^T. \quad (10)$$

Again, considering the special choice of  $\{\rho_k\}_{k=1}^K$  in (7), we find  $\mathbf{V} = \mathbf{F}^H \mathbf{R}_{\text{oa}}$ , where the  $K \times P$  matrix  $\mathbf{R}_{\text{oa}} := [\mathbf{I}_K \mathbf{I}_{\text{oa}}]$  (with  $\mathbf{I}_{\text{oa}}$  denoting the first  $L$  columns of the  $K \times K$  identity matrix  $\mathbf{I}_K$ ) implements in matrix form the overlap-add operation. One remark is due at this point.

*Remark 1:* When  $\rho_k$ s are chosen as in (7), we have  $\mathbf{C} = \sqrt{K} \mathbf{T}_{\text{zp}} \mathbf{F}$  and  $\mathbf{V} = \mathbf{F}^H \mathbf{R}_{\text{oa}}$  so that matrix multiplication by  $\mathbf{C}$  and  $\mathbf{V}$  can be replaced by FFTs. Thus, we infer that (7) gives rise to an OFDM-like transmission with the cyclic prefix (CP) replaced by zero-padding (ZP) at the transmitter; and the CP remove by the overlap-add operation at the receiver. The pros and cons of CP and ZP in OFDM have been discussed in [11] and [20].

Equations (8) and (9) show that the information symbol  $s(n)$  is transmitted twice in two consecutive time intervals through two different channels. In order to decode  $s(n)$  with the embedded diversity gain through the repeated transmission, we define  $\tilde{\mathbf{y}}_k(n) := (Y(2n; \rho_k), Y^*(2n+1; \rho_k))^T$  and write (8) and (9) into a matrix/vector form

$$\tilde{\mathbf{y}}_k(n) = A \tilde{\mathbf{H}}_k \mathbf{s}_k(n) + \tilde{\mathbf{w}}_k(n), \quad \forall k \in [1, K] \quad (11)$$

where  $\mathbf{s}_k(n) := (s(2nK + k - 1), s((2n+1)K + k - 1))^T$ ,  $\tilde{\mathbf{w}}_k(n) := (W(2n; \rho_k), W^*(2n+1; \rho_k))^T$  and

$$\tilde{\mathbf{H}}_k := \begin{pmatrix} H_1(\rho_k) & H_2(\rho_k) \\ H_2^*(\rho_k) & -H_1^*(\rho_k) \end{pmatrix}. \quad (12)$$

Observe that  $\tilde{\mathbf{H}}_k$  in (12) is a scaled unitary matrix. As in [2], we can recover  $\mathbf{s}_k(n)$  by simply multiplying  $\tilde{\mathbf{y}}_k(n)$  by the  $2 \times 2$  matrix:

$$\mathbf{G}_k := \tilde{\mathbf{H}}_k^H, \quad \forall k \in [1, K] \quad (13)$$

and obtain from (11) the decision vector  $\mathbf{z}_k(n) := (z_{k,1}(n), z_{k,2}(n))^T = \mathbf{G}_k \tilde{\mathbf{y}}_k(n)$  as:

$$\mathbf{z}_k(n) = A(|H_1(\rho_k)|^2 + |H_2(\rho_k)|^2) \mathbf{s}_k(n) + \boldsymbol{\eta}_k(n) \quad (14)$$

where  $\boldsymbol{\eta}_k(n) := \tilde{\mathbf{H}}_k^H \tilde{\mathbf{w}}_k(n)$ . Equation (14) implies that transmit diversity gain of order two has been achieved for every  $\mathbf{s}_k(n)$  in our design. After detecting  $\hat{\mathbf{s}}_k(n)$  from  $\mathbf{z}_k(n)$ , the symbols  $s(n)$  can be retrieved by the P/S conversion of  $\hat{\mathbf{s}}_k(n)$  as shown in Fig. 1. Assuming that  $w(n)$  is white, it follows by definition that  $\tilde{\mathbf{w}}_k(n)$  is also white. Since  $\tilde{\mathbf{H}}_k$  is also unitary, we deduce that  $\boldsymbol{\eta}_k(n)$  is white as well. Four remarks are now in order.

*Remark 2:* Since  $\boldsymbol{\eta}_k(n)$  is white, detecting  $\mathbf{s}_k(n)$  from (14) amounts to solving two single symbol detection problems separately without loss of performance, as discussed in [2].

*Remark 3:* Detecting  $2K$ -long data blocks has been cast into  $2K$  single symbol detection problems, as implied by (11). Thus, the receiver complexity is very low and proportional to the block length  $K$ .

*Remark 4:* Since  $H_i(z)$  has degree  $\leq L$ , at most  $L$  of  $\{H_i(\rho_k)\}_{k=1}^K$  can be zero, and thus  $\mathbf{s}_k(n)$  cannot be recovered from (14) when  $H_1(\rho_k) = H_2(\rho_k) = 0$ . Hence, symbol

recovery can not be guaranteed when the two channels share common zeros at  $\{\rho_k\}_{k=1}^K$ , although it is unlikely to have  $H_1(\rho_k) = H_2(\rho_k) = 0$  when the two channels are uncorrelated.

*Remark 5:* In contrast to ST-OFDM in [1], [5], and [6], where  $\{\rho_k\}_{k=1}^K$  are fixed and equispaced around the unit circle, the flexibility to choose different  $\rho_k$ s in our system may offer some advantages. For example, as suggested in [13] and [14], we can periodically rotate (hop) the  $\rho_k$ s to ameliorate consistent fading effects caused by common (or close) channel zeros, as we will simulate in Section VI. Optimal design of  $\{\rho_k\}_{k=1}^K$  is an interesting future direction but goes beyond the scope of this paper.

Instead of assuming channel knowledge at the receiver, we will next equip our receiver with blind channel estimation capabilities.

#### IV. BLIND CHANNEL ESTIMATION

We pursue *blind* estimation of the two channels  $\mathbf{h}_i$ ,  $i = 1, 2$ , based on Assumptions a1), a2) and the following assumptions.

- a3) The modulated information symbols  $s(n)$  have constant modulus (CM);
- a4) The two channels  $\mathbf{h}_i$ ,  $i = 1, 2$  do not share common zeros, i.e., their transfer functions  $H_1(z)$  and  $H_2(z)$  are coprime polynomials;
- a5) The block size  $K$  is chosen to satisfy:  $K \geq 6L + 3$ .

Given  $\tilde{\mathbf{y}}_k(n)$  in (11), our blind channel estimation will be sought in two steps: First, we will exploit the structure of  $\tilde{\mathbf{H}}_k$  and the CM property of  $s(n)$  to develop a *deterministic* CMA that yields two estimates for two channel ratios  $H_1^*(\rho_k)/H_2(\rho_k)$  and  $-H_2^*(\rho_k)/H_1(\rho_k)$  for every  $k = 1, \dots, K$ , with an ambiguity of knowing which estimates correspond to what ratio. Second, we will exploit the FIR nature of the channel and develop an exhaustive search to resolve these  $K$  ambiguities and estimate jointly the two channels  $\mathbf{h}_1$  and  $\mathbf{h}_2$  with two remaining ambiguities: one is a scalar ambiguity; the other one is the ambiguity of distinguishing between  $\mathbf{h}_1$  and  $\mathbf{h}_2$ . These two ambiguities will be resolved by sending two training symbols.

##### A. Estimating Channel Ratios Using CMA

Consider the data model in (11) and simplify the notation by absorbing  $A$  into  $\tilde{\mathbf{H}}_k$  and  $\rho_k$  in the subscript  $k$ , to rewrite it as

$$\tilde{\mathbf{y}}_k(n) := (Y_{1k}(n), Y_{2k}^*(n))^T = \tilde{\mathbf{H}}_k \mathbf{s}_k(n) + \tilde{\mathbf{w}}_k(n), \quad (15)$$

where  $\mathbf{s}_k(n) := (s_{1k}(n), s_{2k}(n))^T$ ,  $\tilde{\mathbf{w}}_k(n) := (W_{1k}(n), W_{2k}^*(n))^T$  and

$$\tilde{\mathbf{H}}_k := \begin{pmatrix} H_{1k} & H_{2k} \\ H_{2k}^* & -H_{1k}^* \end{pmatrix}. \quad (16)$$

Without loss of generality <sup>2</sup>, we will fix the modulus  $|s_{1k}(n)|^2 = 1$  and  $|s_{2k}(n)|^2 = 1$  because all blind methods yield channel estimates up to a constant factor. Given only FFT processed data  $\tilde{\mathbf{y}}_k(n)$  as in (15), we seek an equalizer  $\mathbf{G}_k$

<sup>2</sup>In this paper, we deal with complex modulations only. Nonetheless, the extension to real modulations, e.g., binary phase-shift keying (BPSK), is straightforward and the constraints to be imposed are  $s_{1k}^2(n) = 1$  and  $s_{2k}^2(n) = 1$ .

such that  $\hat{\mathbf{s}}_k(n) := \mathbf{G}_k \tilde{\mathbf{y}}_k(n)$  has entries with unit constant modulus. Note that  $\mathbf{s}_k(n)$  need not be white. Our *deterministic* CM equalization is equivalent to a generalized eigenvalue problem which can be solved by the analytical CMA (ACMA) of [19]. However, we develop here a more simple algorithm which takes advantage of the specific structure of  $\tilde{\mathbf{H}}_k$  induced by the ST code design we described in Section II.

Following common (A)CMA practice, we will first consider blind channel estimation at sufficiently high signal-to-noise ratio (SNR), where the noise  $\tilde{\mathbf{w}}_k(n)$  can be neglected (noise effects will be tested in the simulations). Because  $\tilde{\mathbf{H}}_k$  is unitary and  $\mathbf{G}_k \tilde{\mathbf{H}}_k = \mathbf{I}$ , we look for  $\mathbf{G}_k$  which has a form

$$\mathbf{G}_k := \begin{pmatrix} g_{1k} & g_{2k} \\ g_{2k}^* & -g_{1k}^* \end{pmatrix}. \quad (17)$$

Writing  $\hat{\mathbf{s}}_k(n) := [\hat{s}_{1k}(n), \hat{s}_{2k}(n)]^T = \mathbf{G}_k (Y_{1k}(n), Y_{2k}^*(n))^T$  component-wise, and imposing the CM constraints  $|\hat{s}_{1k}(n)|^2 = 1$ ,  $|\hat{s}_{2k}(n)|^2 = 1$ , we arrive at  $\tilde{\mathbf{Y}}_k(n) \mathbf{g}_k = (1, 1)^T$ , where (18), shown at the bottom of the page. Next, we stack  $N$  blocks of data  $\{\tilde{\mathbf{y}}_k(n)\}_{n=0}^{N-1}$  and concatenate  $\tilde{\mathbf{Y}}_k(n) \mathbf{g}_k$  to arrive at

$$\tilde{\mathbf{Y}}_k \mathbf{g}_k = (1, \dots, 1)_{2N \times 1}^T \quad (19)$$

where  $\tilde{\mathbf{Y}}_k := (\tilde{\mathbf{Y}}_k^T(0), \dots, \tilde{\mathbf{Y}}_k^T(N-1))^T$ . If the  $2N \times 4$  matrix  $\tilde{\mathbf{Y}}_k$  had full column rank,  $\tilde{\mathbf{Y}}_k^\dagger$  would have yielded a unique solution of  $\mathbf{g}_k$  from (19). Unfortunately, the maximum column rank of  $\tilde{\mathbf{Y}}_k$  is only three [19]. Nevertheless, the solution of (19) can still be sought in the form  $\mathbf{g}_k = \mathbf{g}_{0k} + \lambda \mathbf{g}_{1k}$ , where  $\mathbf{g}_{0k}$  is a particular solution of (19) and  $\mathbf{g}_{1k}$  spans the one-dimensional kernel of  $\tilde{\mathbf{Y}}_k$ . With  $g_k(i)$  denoting the  $i$ th entry of  $\mathbf{g}_k$ ,  $\lambda$  can be determined by noting that in (18) we must have  $g_k(1)g_k(2) = g_k(3)g_k(4)$ . Because solving the latter leads to a second-order equation in  $\lambda$ , we end up with two possible solutions. In matrix/vector form, these solutions are given by

$$\begin{aligned} \tilde{\mathbf{G}}_k &:= \begin{pmatrix} \tilde{g}_{1k} & \tilde{g}_{2k} \\ \tilde{g}_{2k}^* & -\tilde{g}_{1k}^* \end{pmatrix} = \frac{\alpha_k}{|H_{1k}|^2 + |H_{2k}|^2} \begin{pmatrix} H_{1k}^* & H_{2k} \\ H_{2k}^* & -H_{1k} \end{pmatrix} \\ \tilde{\mathbf{g}}_k &:= (\tilde{g}_{1k} \tilde{g}_{1k}^* \quad \tilde{g}_{2k} \tilde{g}_{2k}^* \quad \tilde{g}_{1k} \tilde{g}_{2k}^* \quad \tilde{g}_{1k}^* \tilde{g}_{2k})^T \\ \tilde{\mathbf{G}}'_k &:= \begin{pmatrix} \tilde{g}'_{1k} & \tilde{g}'_{2k} \\ \tilde{g}'_{2k} & -\tilde{g}'_{1k} \end{pmatrix} = \frac{\alpha'_k}{|H_{1k}|^2 + |H_{2k}|^2} \begin{pmatrix} H_{2k}^* & -H_{1k} \\ H_{1k} & H_{2k} \end{pmatrix} \\ \tilde{\mathbf{g}}'_k &:= (\tilde{g}'_{1k} \tilde{g}'_{1k}^* \quad \tilde{g}'_{2k} \tilde{g}'_{2k}^* \quad \tilde{g}'_{1k} \tilde{g}'_{2k}^* \quad \tilde{g}'_{1k}^* \tilde{g}'_{2k})^T, \end{aligned} \quad (20)$$

where  $\alpha_k$  and  $\alpha'_k$  with  $|\alpha_k| = |\alpha'_k| = 1$  denote the phase ambiguities. One can readily verify that  $\tilde{\mathbf{G}}_k \tilde{\mathbf{H}}_k$  and  $\tilde{\mathbf{G}}'_k \tilde{\mathbf{H}}_k$  give rise to a diagonal and an anti-diagonal matrix, respectively (the CM property is satisfied in both cases, but in the second case the role of the two transmitted sequences is permuted, thus creating ambiguity). Since we cannot distinguish between the two solutions  $\tilde{\mathbf{g}}_k$  and  $\tilde{\mathbf{g}}'_k$ , it follows that estimating the equalizer by solving (19) leads to an ambiguity in choosing between the two possible equalizers  $\tilde{\mathbf{G}}_k$  and  $\tilde{\mathbf{G}}'_k$  (and thus the two channels) in addition

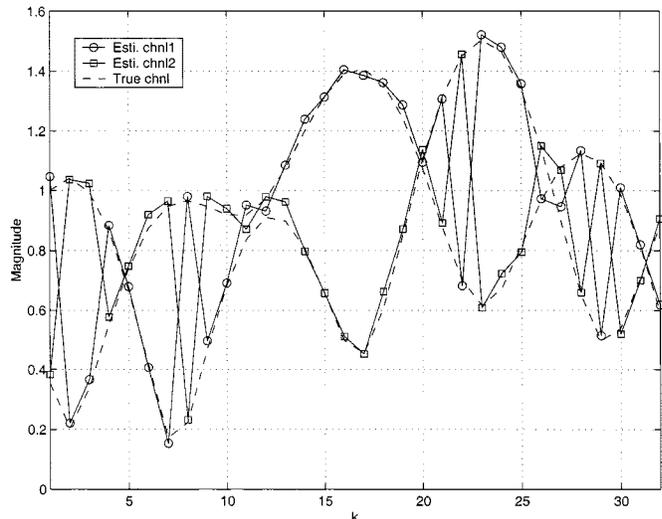


Fig. 2. Estimates of channel magnitude using "deterministic CMA".

to the phase ambiguities  $\alpha_k$  and  $\alpha'_k$  for every  $k$ . As an example, we obtain  $\mathbf{g}_k$  by randomly choosing one of two solutions of (19) at  $SNR = 15$  dB, and compute the two channel magnitude estimates as:  $|\hat{H}_{ik}| := \sqrt{g_k(i)/(g_k(1) + g_k(2))}$ ,  $i = 1, 2$ , that we plot in Fig. 2. We observe that indeed this simple method yields good estimates of the channel magnitude but we have ambiguity between the two channels for every  $k$ .

According to Assumption a5), we can have at least  $6L + 3$  solution pairs  $\{\tilde{\mathbf{g}}_k, \tilde{\mathbf{g}}'_k\}_{k=1}^{6L+3}$ , with  $\tilde{\mathbf{g}}_k$  and  $\tilde{\mathbf{g}}'_k$  consisting of products between  $H_{1k}(H_{1k}^*)$  and  $H_{2k}(H_{2k}^*)$  as seen from (20). However, Assumption a1) implies that  $H_{1k} = 0$  on at most  $L$  values of  $k$ , and likewise for  $H_{2k}$ . Hence, we can always find at least  $6L + 3 - 2L = 4L + 3$  solution pairs  $\{\tilde{\mathbf{g}}_k, \tilde{\mathbf{g}}'_k\}_{k=1}^{4L+3}$  for which both  $H_{1k}$  and  $H_{2k}$  are nonzero. For these  $k$ s, we can define the ratio:  $r(k) = g_k(3)/g_k(2)$ , and use (20) to infer that  $\forall k \in [1, 4L + 3]$

$$\begin{aligned} r(k) &= \begin{cases} \tilde{r}(k) := \frac{H_{1k}^*}{H_{2k}} = \frac{H_{1k}^*(\rho_k)}{H_{2k}(\rho_k)}, & \text{if } \mathbf{g}_k = \tilde{\mathbf{g}}_k \\ \tilde{r}'(k) := \frac{-H_{2k}^*}{H_{1k}} = \frac{-H_{2k}^*(\rho_k)}{H_{1k}(\rho_k)}, & \text{if } \mathbf{g}_k = \tilde{\mathbf{g}}'_k \end{cases} \end{aligned} \quad (21)$$

where we observe that the phase ambiguities  $\alpha_k$  and  $\alpha'_k$  have been eliminated by taking channel ratios. Now the ambiguity between  $\tilde{\mathbf{g}}_k$  and  $\tilde{\mathbf{g}}'_k$  in solution pairs  $\{\tilde{\mathbf{g}}_k, \tilde{\mathbf{g}}'_k\}_{k=1}^{4L+3}$  translates to the ambiguity between  $\tilde{r}(k)$  and  $\tilde{r}'(k)$  for every  $r(k)$ . To resolve this ambiguity, one approach is to estimate one channel by training, and retrieve the other channel by inspection from (21). Because only training for one channel is needed, such a "partially trained" CMA saves 50% overhead. Since  $L$  channel ratios are sufficient to estimate one channel, Assumption a5) can be relaxed to  $K \geq 3L$ . Note also that so far, we do not require Assumption a4) to obtain channel ratios and further estimate the

$$\begin{aligned} \tilde{\mathbf{Y}}_k(n) &:= \begin{pmatrix} |Y_{1k}(n)|^2 & |Y_{2k}(n)|^2 & Y_{1k}(n)Y_{2k}(n) & Y_{1k}^*(n)Y_{2k}^*(n) \\ |Y_{2k}(n)|^2 & |Y_{1k}(n)|^2 & -Y_{1k}(n)Y_{2k}(n) & -Y_{1k}^*(n)Y_{2k}^*(n) \end{pmatrix} \\ \mathbf{g}_k &:= (g_{1k}g_{1k}^* \quad g_{2k}g_{2k}^* \quad g_{1k}g_{2k}^* \quad g_{1k}^*g_{2k})^T. \end{aligned} \quad (18)$$

two channels if partial training is applied. However, Assumption a4) and a5) will be used next for deriving a fully blind channel estimator, that relies on an exhaustive search to resolve the ambiguity between  $r(k)$  and  $r'(k)$ .

### B. Resolving Channel Ambiguity

Exploiting the specific structure of ST codes, the CMA of the previous subsection yields two possible ratios for every  $k$ . Starting from the ratio pairs,  $\{\tilde{r}(k), \tilde{r}'(k)\}_{k=1}^{4L+3}$ , we exploit here the finite channel support and resort to an exhaustive search to estimate the two channels blindly.

Recalling that for each  $k \in [1, 4L+3]$ ,  $r(k)$  can be either  $\tilde{r}(k)$  or  $\tilde{r}'(k)$ , we infer that there are  $2^{4L+3}$  possible collections of  $r(k)$ s. With  $i_c$  denoting collection-index, we represent each of them with the  $(4L+3) \times 1$  vector  $\mathbf{r}_{i_c} := (r_{i_c}(1), \dots, r_{i_c}(4L+3))^T$ , and the entire collection with the set  $\mathcal{R} = \{\mathbf{r}_{i_c}, i_c = 1, \dots, 2^{4L+3}\}$ . Because each entry of  $\mathbf{r}_{i_c}$  comes either from  $\tilde{r}(k)$  or from  $\tilde{r}'(k)$ , the  $4L+3$  entries of every  $\mathbf{r}_{i_c}$  can be divided into two groups, namely the  $\tilde{r}$  group that contains  $\tilde{r}(k)$ s and the  $\tilde{r}'$  group that consists of  $\tilde{r}'(k)$ s. Thinking in terms of a coin-flipping experiment, we term the entries in each group as “same-side” entries. To resolve channel ambiguity, we will use the following lemma (see Appendix I for the proof):

*Lemma 1:* Under Assumptions a1)–a5),  $2L+2$  same-side entries of any  $\mathbf{r}_{i_c}$ , denoted by  $\{r_{i_c}(k_j), 1 \leq k_j \leq 4L+3\}_{j=1}^{2L+2}$ , enable identifiability of the two channels (within a scalar ambiguity) either as  $\mathbf{h}_{12} := (\mathbf{h}_1^T, \mathbf{h}_2^T)^T$ , or, as  $\mathbf{h}_{21} := (-\mathbf{h}_2^T, \mathbf{h}_1^T)^T$ . Using a common notation  $(\mathbf{b}_{i_c}^T, \mathbf{a}_{i_c}^T)^T$ , either  $\mathbf{h}_{12}$  or  $\mathbf{h}_{21}$  can be found by solving for the eigenvector corresponding to the minimum eigenvalue of the  $(2L+2) \times (2L+2)$  matrix  $\Theta_{i_c}$

$$\Theta_{i_c} := \begin{pmatrix} v^T(\rho_{k_1}^*, L+1) & -r_{i_c}(k_1)v^T(\rho_{k_1}, L+1) \\ \vdots & \vdots \\ v^T(\rho_{k_{2L+2}}^*, L+1) & -r_{i_c}(k_{2L+2})v^T(\rho_{k_{2L+2}}, L+1) \end{pmatrix} \quad (22)$$

Because each  $\mathbf{r}_{i_c}$  has  $4L+3$  entries, we infer that  $2L+2$  same-side entries can always be found. However, because an ambiguity appears for every  $k$ , we do not know where the  $2L+2$  same-side entries are. To locate these  $2L+2$  same-side entries, we resort to an exhaustive search following the following steps.

- s1) For each  $i_c \in [1, 2^{4L+3}]$ , use  $\{r_{i_c}(k)\}_{k=1}^{2L+2}$  and obtain  $(\mathbf{b}_{i_c}^T, \mathbf{a}_{i_c}^T)^T$  (with  $k_j$  replaced by  $j$  for  $j = 1, \dots, 2L+2$  in  $\Theta_{i_c}$ ), as in Lemma 1. We prove in Appendix I that if  $\{r_{i_c}(k)\}_{k=1}^{2L+2}$  are same-side entries, the matrix  $\Theta_{i_c}$  in (22) has nullity one and the solution  $(\mathbf{b}_{i_c}^T, \mathbf{a}_{i_c}^T)^T$  is either  $\mathbf{h}_{12}$  or  $\mathbf{h}_{21}$  as described in Lemma 1.
- s2) Use  $(\mathbf{b}_{i_c}, \mathbf{a}_{i_c})$  from s1) to form  $\forall k \in [1, 4L+3]$  the difference:

$$\Delta_{i_c}(k) := v^T(\rho_k^*, L+1)\mathbf{b}_{i_c}^* - r_{i_c}(k)v^T(\rho_k, L+1)\mathbf{a}_{i_c}. \quad (23)$$

- s3) Select the index:  $\mathbf{i}_c := \arg \min_{i_c} \sum_{k=1}^{4L+3} |\Delta_{i_c}(k)|^2$ .

The two possible solutions for the pair of our channels are either

$$\begin{pmatrix} \mathbf{h}_1 \\ \mathbf{h}_2 \end{pmatrix} = \alpha \begin{pmatrix} \mathbf{b}_{\mathbf{i}_c} \\ \mathbf{a}_{\mathbf{i}_c} \end{pmatrix} \quad \text{or} \quad \begin{pmatrix} \mathbf{h}_1 \\ \mathbf{h}_2 \end{pmatrix} = \alpha \begin{pmatrix} \mathbf{a}_{\mathbf{i}_c} \\ -\mathbf{b}_{\mathbf{i}_c} \end{pmatrix} \quad (24)$$

where  $\alpha$  denotes a scalar ambiguity which will be resolved by one training symbol.

We next prove that the channel estimates in (24) are unique. Observe that, out of all the  $2^{4L+3}$  vectors in  $\mathcal{R}$ , two vectors yield  $\sum_{k=1}^{4L+3} |\Delta_{i_c}(k)|^2 = 0$  from s1)–s2); namely, the one with entries  $r_{i_c}(k) = \tilde{r}(k)$ ,  $\forall k \in [1, 4L+3]$ , and the one with entries  $r_{i_c}(k) = \tilde{r}'(k)$ ,  $\forall k \in [1, 4L+3]$ . Due to the fact  $\sum_{k=1}^{4L+3} |\Delta_{i_c}(k)|^2 \geq 0$ , we deduce that  $\sum_{k=1}^{4L+3} |\Delta_{\mathbf{i}_c}(k)|^2 = 0$  and thereafter  $\Delta_{\mathbf{i}_c}(k) = 0$  for every  $k$ . As we mentioned before, there exist at least  $2L+2$  same-side entries in the resulting  $\mathbf{r}_{\mathbf{i}_c}$ . These entries must satisfy  $\Delta_{\mathbf{i}_c}(k) = 0$ . Based on Lemma 1, we infer that our exhaustive search from s1)–s3) yields two channel estimates as in (24).

Having reduced the ambiguity to choosing between these two solutions and resolving the scalar ambiguity, enables (almost blind) channel estimation using the received data and two training symbols only. Specifically, with two training symbols, we can estimate  $H_{11}$  and  $H_{21}$  (i.e.,  $H_{1k}$  and  $H_{2k}$  for  $k=1$ ) from which the two channels can be identified uniquely.

The computational complexity of our fully blind channel estimation algorithm is relatively high. The CMA requires a single value decomposition (SVD) of size  $2N \times 4$  for every  $k$ , which is still much simpler than the ACMA in [19], while the exhaustive search involves  $2^{2L+2}$  eigen value decompositions (EVDs) of size  $(2L+2) \times (2L+2)$ . Fortunately,  $L$  is small in typical applications [e.g.,  $L \approx 4$  in global system for mobile communication (GSM)] and the smallest null vector can be computed online [3]. Note also that, unlike existing statistical CMA variants that require long data sets, our ST-coded deterministic CMA is data-efficient and does not impose any input whiteness assumption.

To reduce complexity, we have also derived a preweighting scheme with different weights for  $s_{1k}(n)$  and  $s_{2k}(n)$  which avoids the exhaustive search [9], [10]. However, the price paid for the simplicity of [9] and [10] is reduced BER performance as we will illustrate in Section VI.

As a performance benchmark, the CRB of channel estimates will be derived next. Interestingly, the CRB derivation will provide additional insights to channel estimation issues arising with transmit antenna diversity.

### V. CRAMÉR–RAO BOUND

In this section, we derive the CRB of the channel estimates for ST-GOFDM. Our derivations will follow the general steps of [18] and start with the derivation of the Fisher’s information matrix (FIM) by treating the transmitted symbols as nuisance parameters. Interestingly, checking the existence of CRB under various constraints reveals that imposing constraints on transmitted symbols is indispensable for blind channel estimation for ST-GOFDM. For simplicity, the specific choice of  $\rho_k$ s in (7) is adopted in our derivations.

Denoting by  $\mathcal{D}(\mathbf{v}, D) := \text{diag}(v_1, \dots, v_D)$  a  $D \times D$  diagonal matrix built from the  $D \times 1$  vector  $\mathbf{v} := (v_1, \dots, v_D)^T$ , we consider (8) and (9) and absorb  $A$  into  $H_i(\rho_k)$  to cast them into matrix forms

$$\begin{aligned} \tilde{\mathbf{y}}(2n) &= \mathcal{D}(\tilde{\mathbf{h}}_1, K)\mathbf{s}(2n) + \mathcal{D}(\tilde{\mathbf{h}}_2, K)\mathbf{s}(2n+1) + \tilde{\mathbf{w}}(2n) \\ \tilde{\mathbf{y}}(2n+1) &= -\mathcal{D}(\tilde{\mathbf{h}}_1, K)\mathbf{s}^*(2n+1) + \mathcal{D}(\tilde{\mathbf{h}}_2, K)\mathbf{s}^*(2n) \\ &\quad + \tilde{\mathbf{w}}(2n+1) \end{aligned} \quad (25)$$

where  $\tilde{\mathbf{y}}(n) := (Y(n; \rho_1), \dots, Y(n; \rho_K))^T$ ,  $\tilde{\mathbf{h}}_i := (H_i(\rho_1), \dots, H_i(\rho_K))^T$ ,  $i = 1, 2$  and  $\tilde{\mathbf{w}}(n) := (W(n; \rho_1), \dots, W(n; \rho_K))^T = \mathbf{V}\mathbf{w}(n)$ . Clearly,  $\tilde{\mathbf{w}}(n)$  is zero mean, Gaussian and uncorrelated from  $s(n)$ . When  $\rho_k$ s are chosen as in (7), we have  $\mathbf{V} = \mathbf{F}^H \mathbf{R}_{\text{oa}} = \mathbf{F}^H [\mathbf{I}_K \mathbf{I}_{\text{oa}}]$  and the covariance matrix of  $\tilde{\mathbf{w}}(n)$  is, thus, given by

$$\mathbf{R}_{\tilde{\mathbf{w}}} = E \{ \mathbf{V}\mathbf{w}(n)\mathbf{w}^H(n)\mathbf{V}^H \} = \sigma_w^2 (\mathbf{K}\mathbf{I} + \mathbf{F}^H \mathbf{I}_{\text{oa}} \mathbf{I}_{\text{oa}}^H \mathbf{F}) \quad (26)$$

where  $\sigma_w^2$  is the variance of  $w(n)$ . It is easy to verify that the maximum entry of the matrix  $\mathbf{F}^H \mathbf{I}_{\text{oa}} \mathbf{I}_{\text{oa}}^H \mathbf{F}$  is  $L$ . Since the block length  $K \gg L$ , the matrix  $\mathbf{R}_{\tilde{\mathbf{w}}}$  can be well approximated as  $\mathbf{R}_{\tilde{\mathbf{w}}} \approx P\sigma_w^2 \mathbf{I}$ ; i.e.,  $W(n; \rho_k)$  is approximately white with variance  $\sigma^2 := P\sigma_w^2$ .

Let  $\boldsymbol{\theta} := (\mathbf{h}_1^T, \mathbf{h}_2^T, \mathbf{h}_1^H, \mathbf{h}_2^H, \mathbf{s}_N^T, \mathbf{s}_N^H)^T$  denote the  $4(L + KN + 1) \times 1$  complex vector containing coefficients of the two channels' impulse responses and the input vector  $\mathbf{s}_N := (\mathbf{s}^T(0), \mathbf{s}^T(2), \dots, \mathbf{s}^T(2N-2), \mathbf{s}^T(1), \mathbf{s}^T(3), \dots, \mathbf{s}^T(2N-1))^T$ . Because  $\tilde{\mathbf{w}}(n)$ 's in (25) are Gaussian, we can write the log-likelihood function conditioned on  $\mathbf{h}_1, \mathbf{h}_2$  and  $\mathbf{s}_N$ , as

$$\begin{aligned} f(\boldsymbol{\theta}) &:= -\log(2\pi\sigma)^{2KN} \\ &\quad - \frac{1}{\sigma^2} \sum_{n=0}^{N-1} \|\tilde{\mathbf{y}}(2n) - \mathcal{D}(\tilde{\mathbf{h}}_1, K)\mathbf{s}(2n) \\ &\quad \quad - \mathcal{D}(\tilde{\mathbf{h}}_2, K)\mathbf{s}(2n+1)\|^2 \\ &\quad - \frac{1}{\sigma^2} \sum_{n=0}^{N-1} (\|\tilde{\mathbf{y}}(2n+1) - \mathcal{D}(\tilde{\mathbf{h}}_1, K)\mathbf{s}^*(2n+1) \\ &\quad \quad - \mathcal{D}(\tilde{\mathbf{h}}_2, K)\mathbf{s}^*(2n)\|^2). \end{aligned} \quad (27)$$

The FIM for unbiased  $\boldsymbol{\theta}$  estimates is given by

$$\mathcal{F}(\boldsymbol{\theta}) := E_{\mathbf{h}_1, \mathbf{h}_2, \mathbf{s}_N} \{ \nabla_{\boldsymbol{\theta}}^H f(\boldsymbol{\theta}) \nabla_{\boldsymbol{\theta}} f(\boldsymbol{\theta}) \} = \begin{pmatrix} \mathcal{F}_{11} & \mathcal{F}_{12} \\ \mathcal{F}_{12}^H & \mathcal{F}_{22} \end{pmatrix} \quad (28)$$

where the submatrices  $\mathcal{F}_{11}$ ,  $\mathcal{F}_{12}$ , and  $\mathcal{F}_{22}$  correspond to partitioning channel and input parameters and have dimensionalities  $(4L+4) \times (4L+4)$ ,  $(4L+4) \times 4KN$ ,  $4KN \times 4KN$ , respectively. Specifically, with  $\mathcal{B}_1, \mathcal{B}_2, \mathcal{B}_3, \mathcal{B}_4$  and  $\mathcal{B}_5$  specified in Appendix II, we have

$$\begin{aligned} \mathcal{F}_{11} &:= \frac{1}{\sigma^2} \text{diag}(\mathcal{B}_1, \mathcal{B}_1, \mathcal{B}_1, \mathcal{B}_1) \\ \mathcal{F}_{22} &:= \frac{1}{\sigma^2} \mathbf{I}_{4N \times 4N} \otimes \mathcal{B}_3 \\ \mathcal{F}_{12} &:= \frac{1}{\sigma^2} (\mathcal{B}_4 \quad \mathcal{B}_5) \end{aligned}$$

where  $\otimes$  denotes the Kronecker product. Using the block matrix inversion formula, we find the CRB  $\mathcal{C}(\mathbf{h})$  for the channel estimates based on the top-left submatrix of  $\mathcal{F}^{-1}(\boldsymbol{\theta})$  as

$$\mathcal{C}(\mathbf{h}) = (\mathcal{F}_{11} - \mathcal{F}_{12} \mathcal{F}_{22}^{-1} \mathcal{F}_{12}^H)^{-1} \quad (29)$$

provided that the inverse in (29) exists. Unfortunately, direct substitution reveals that  $\mathcal{F}_{11} - \mathcal{F}_{12} \mathcal{F}_{22}^{-1} \mathcal{F}_{12}^H \equiv \mathbf{0}$ , when  $\rho_k$ 's are chosen as in (7). However, zero FIM (or infinite CRB) implies lack of identifiability. Hence, our blind channel estimation setup that treats transmitted (input) symbols as nuisance parameters can not guarantee channel identifiability, unless extra constraints are imposed.

Suppose we impose  $M$  continuously differentiable constraints on the two channels

$$\zeta_m(\mathbf{h}_1^T, \mathbf{h}_2^T, \mathbf{h}_1^H, \mathbf{h}_2^H) = 0, \quad m = 1, \dots, M \quad (30)$$

where the number of constraints (equations)  $M$  must be less than  $4L+4$  channel unknowns, because otherwise the channels can be obtained directly from the constraints. Along the lines of [15], we obtain the  $M \times 4(KN+L+1)$  gradient matrix of the constraints in (30) as:

$$\begin{aligned} \mathcal{G}_h &:= (\mathcal{E}_h, \mathbf{0}_{M \times 4KN}) \\ &:= \begin{pmatrix} \nabla_{\mathbf{h}_1^T} \zeta_1 & \nabla_{\mathbf{h}_2^T} \zeta_1 & \nabla_{\mathbf{h}_1^H} \zeta_1 & \nabla_{\mathbf{h}_2^H} \zeta_1 \\ \vdots & \vdots & \vdots & \vdots \\ \nabla_{\mathbf{h}_1^T} \zeta_M & \nabla_{\mathbf{h}_2^T} \zeta_M & \nabla_{\mathbf{h}_1^H} \zeta_M & \nabla_{\mathbf{h}_2^H} \zeta_M \end{pmatrix} \Bigg| \mathbf{0}_{M \times 4KN} \end{pmatrix}. \end{aligned} \quad (31)$$

Following [15], the constrained CRB requires the orthonormal basis of the nullspace  $\mathcal{N}(\mathcal{G}_h)$ . In our specific case, we let  $\mathcal{U}_h$  be a  $(4KN+4L+4) \times (4KN+4L+4-M)$  matrix with columns the basis vectors of  $\mathcal{N}(\mathcal{G}_h)$ , and compute first the constrained CRB for all the unknown parameter estimates

$$\mathcal{C}(\boldsymbol{\theta}) = \mathcal{U}_h (\mathcal{U}_h^H \mathcal{F}(\boldsymbol{\theta}) \mathcal{U}_h)^{-1} \mathcal{U}_h^H. \quad (32)$$

Our constrained CRB of the channel estimates would then be given by the  $(4L+4-M) \times (4L+4-M)$  top-left submatrix of  $\mathcal{C}(\boldsymbol{\theta})$ . Unfortunately, as we prove in the Appendix the matrix inverse in (32) does not exist *irrespective* of the channel constraints imposed. This reveals another important feature of our blind setup: channel constraints alone are not sufficient to guarantee channel identifiability. Hence, constraints (like CM) on our transmitted symbols are a *must* when it comes to guaranteeing blind channel identifiability.

With our constraints,  $|s(n)|^2 = 1$ ,  $n \in [0, 2NK-1]$ , the  $(2KN+M) \times 4(KN+L+1)$  gradient matrix of the CM and channel constraints in (30) turns out to be

$$\begin{aligned} \mathcal{G}_{CM,h} &:= \begin{pmatrix} \mathcal{E}_h & \mathbf{0}_{M \times 2KN} & \mathbf{0}_{M \times 2KN} \\ \mathbf{0}_{2KN \times (4L+4)} & \mathcal{D}^*(\mathbf{s}_N, 2KN) & \mathcal{D}(\mathbf{s}_N, 2KN) \end{pmatrix}. \end{aligned} \quad (33)$$

The corresponding  $4(KN + L + 1) \times (2KN + 4L + 4 - M)$  matrix  $\mathbf{U}_{CM,h}$  built with columns the orthonormal basis vector of  $\mathcal{N}(\mathcal{G}_{CM,h})$  is given by

$$\mathbf{U}_{CM,h} = \begin{pmatrix} \mathbf{v}_h & \mathbf{0}_{(4L+4) \times 2KN} \\ \mathbf{0}_{2KN \times (4L+4-M)} & \frac{1}{\sqrt{2}} \mathcal{D}^{*-1}(\mathbf{s}_N, 2KN) \\ \mathbf{0}_{2KN \times (4L+4-M)} & -\frac{1}{\sqrt{2}} \mathcal{D}^{-1}(\mathbf{s}_N, 2KN) \end{pmatrix} \quad (34)$$

where  $\mathbf{v}_h$  is the  $(4L + 4) \times (4L + 4 - M)$  matrix with columns the basis vectors of  $\mathcal{N}(\mathcal{E}_h)$ . The CRB of  $\boldsymbol{\theta}$  parameter estimates under CM constraints and  $M$  channel constraints is

$$\mathcal{C}_{CM,h}(\boldsymbol{\theta}) = \mathbf{U}_{CM,h} (\mathbf{U}_{CM,h}^H \mathcal{F}(\boldsymbol{\theta}) \mathbf{U}_{CM,h})^{-1} \mathbf{U}_{CM,h}^H. \quad (35)$$

Extracting the  $(4L + 4 - M) \times (4L + 4 - M)$  top-left submatrix of  $\mathcal{C}_{CM,h}(\boldsymbol{\theta})$  yields the constrained CRB for our channel parameter estimates as  $\mathcal{C}_h = \mathcal{C}_{CM,h}(1:4L + 4 - M, 1:4L + 4 - M)$ , where we used Matlab's notation  $\mathbf{X}(i_1:i_2, j_1:j_2)$  to denote the submatrix of  $\mathbf{X}$  formed by its  $i_1(j_1)$  through  $i_2(j_2)$  rows(columns).

As discussed in Section IV, to resolve the ambiguity between our two possible CMA-based channel estimates in (24), we utilize the pilot-based acquisition of  $H_{11}$  and  $H_{21}$ . Viewing the latter as channel constraints complementing our CM input constraints, enables evaluation of such a "fully constrained" CRB on our channel estimates. Specifically, we can rely on (35), but with  $\mathbf{v}_h$  in (33) being formed by replacing  $\mathcal{E}_h$  with

$$\mathcal{E}_h = \begin{pmatrix} \mathbf{I}_{1 \times (L+1)} & \mathbf{0}_{1 \times (L+1)} & \mathbf{0}_{1 \times (L+1)} & \mathbf{0}_{1 \times (L+1)} \\ \mathbf{0}_{1 \times (L+1)} & \mathbf{I}_{1 \times (L+1)} & \mathbf{0}_{1 \times (L+1)} & \mathbf{0}_{1 \times (L+1)} \end{pmatrix}. \quad (36)$$

In the ensuing section, we will test how close our ST-coded CMA-based channel estimates come to the constrained CRB derived in this section which benchmarks *algorithm independent* performance of blind channel estimates that rely on the same constraints.

## VI. ANALYTICAL AND SIMULATED PERFORMANCE

When the noise  $w(n)$  is AWGN, theoretical BER evaluation is possible for a given constellation. Starting from (14), we can compute the covariance matrix of  $\boldsymbol{\eta}_k(n)$  as

$$\mathbf{R}_{\boldsymbol{\eta}_k} = (|H_1(\rho_k)|^2 + |H_2(\rho_k)|^2) \mathbf{v}^H(\rho_k, P) \mathbf{v}(\rho_k, P) \sigma_w^2 \mathbf{I} \quad (37)$$

and derive the BER assuming, e.g., a quaternary phase-shift keying (QPSK) modulation scheme. Our figure of merit is the average BER, defined as  $\bar{P}_e = (2K)^{-1} \sum_{k=1}^K \sum_{i=1}^2 \bar{P}_{i,k}$ , where  $\bar{P}_{1,k}$  and  $\bar{P}_{2,k}$  denote the BERs for the sequences  $s(2nK + k - 1)$  and  $s((2n + 1)K + k - 1)$ , respectively. It follows from (14) and (37) that

$$\bar{P}_e = \frac{1}{K} \sum_{k=1}^K \mathcal{Q} \left( \sqrt{\frac{A^2 E_b (|H_1(\rho_k)|^2 + |H_2(\rho_k)|^2)}{N_0 v^H(\rho_k, P) v(\rho_k, P)}} \right) \quad (38)$$

where  $\mathcal{Q}(\cdot)$  denotes the  $\mathcal{Q}$ -function  $2E_b/N_0$  denotes bit SNR. Note that because each symbol is transmitted twice, we divide the transmit power by two for each transmit antenna to obtain (38). When points  $\{\rho_k\}_{k=1}^K$  are equispaced around the unit circle as in (7), (38) can be further simplified to

$$\bar{P}_e = \frac{1}{K} \sum_{k=1}^K \mathcal{Q} \left( \sqrt{\frac{K(|H_1(\rho_k)|^2 + |H_2(\rho_k)|^2) E_b}{(K + L) N_0}} \right). \quad (39)$$

Apart from the closed-form BER expressions in (38) and (39), we will resort to simulations in order to test performance and reveal additional salient features of our design.

In all simulations, QPSK modulation is employed and  $\rho_k$ s are chosen equispaced around the unit circle as in (7). All curves are averaged over 200 random channels which are generated based on the following two models.

### A. Multiray Time-Invariant Channel

An  $L$ th-order multiray (MR) channel consists of  $L + 1$  equal-power channel taps with each tap modeled as a complex Gaussian random variable with zero mean and variance  $1/(L + 1)$ . MR channels are used in Examples 1–4.

### B. Typical-Urban GSM Time-Varying Channel

The delay profile for typical-urban (TU) channels is tabulated in [5, Table I]. With the system parameters described in [5, Section III.A], a TU channel corresponds to an FIR channel of order  $L = 4$  whose  $L + 1 = 5$  channel taps are characterized by the Jakes' Doppler spectrum with a Doppler frequency of 40 Hz. TU channels are used in Example 5.

*Example 1 (Performance Gains with ST Coding):* ST-GOFDM is compared to conventional (single transmit antenna) OFDM, assuming that channels are flat ( $L = 0$ ) and known to the receiver. In OFDM, 16 subcarriers are used, and correspondingly in ST-GOFDM, we choose the block length  $K = 16$ . The results are depicted in Fig. 3 where ST-GOFDM is significantly better than OFDM. For MR channels of order  $L = 4$ , we choose 32 subcarriers for OFDM and  $K = 32$  for ST-GOFDM. Fig. 4 shows that ST-GOFDM outperforms OFDM considerably. The conventional OFDM transmits each symbol through a single fading channel. In contrast, ST-GOFDM transmits each symbol twice through two different fading channels. Equation (14) shows that the equivalent channel gain for each symbol is the sum of the squares of two different channels which is more reliable than a single channel. This explains why ST-GOFDM outperforms the OFDM at the expense of an extra antenna.

*Example 2 (Comparison with Pre-Weighting):* As mentioned in Section IV, blind channel estimators proposed in [9], [10] avoid costly exhaustive searching by employing preweighting. This example compares our CMA channel estimator to that based on preweighting. MR channels of order  $L = 4$  are used and the block length  $K = 32$ . The preweighting matrix in [9], [10] is chosen as  $\text{diag}(1.28, 0.6)$ , so that the transmit power for both schemes is identical. The BER curves depicted in Fig. 5 illustrate that exhaustive search outperforms

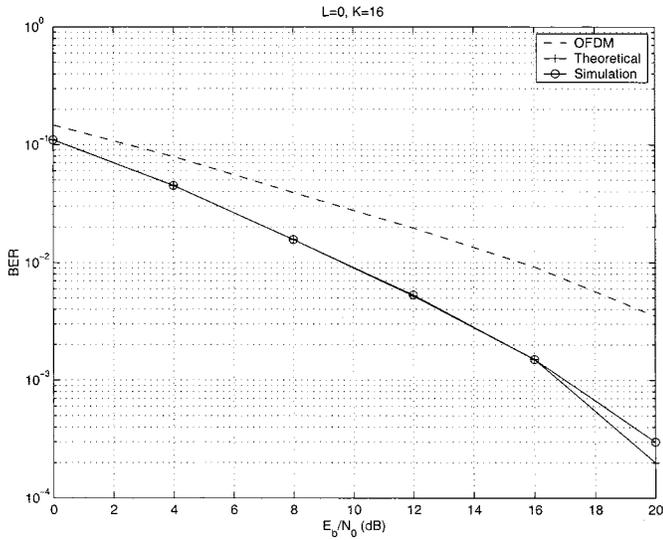


Fig. 3. Flat fading channels.

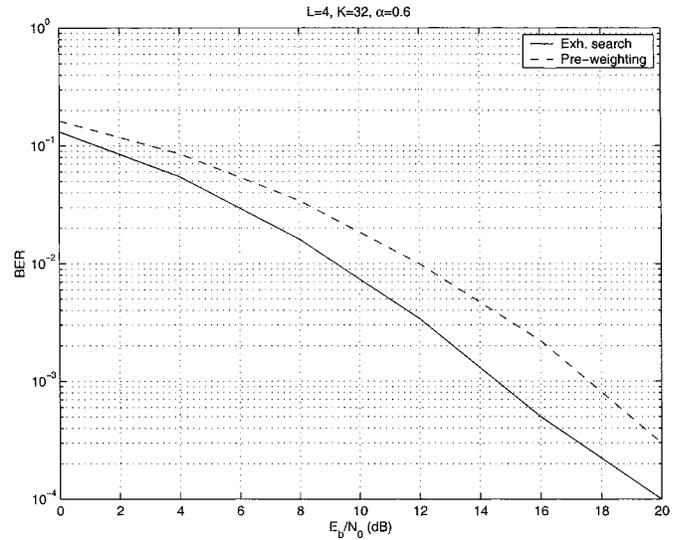


Fig. 5. Prewighting versus exhaustive search.

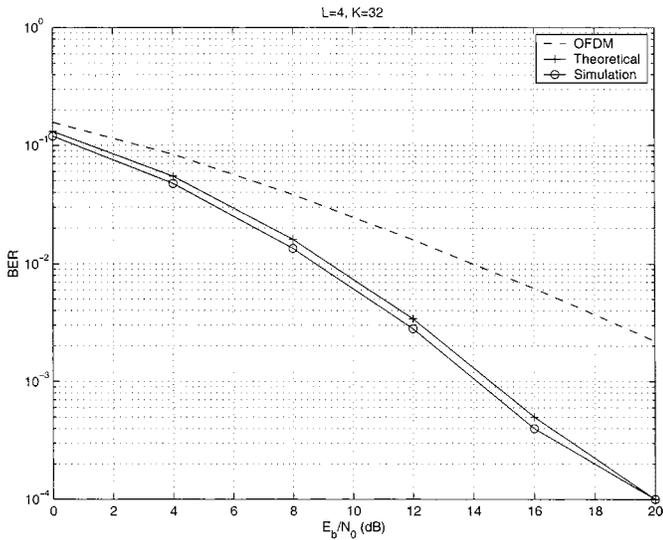


Fig. 4. Frequency selective channels.

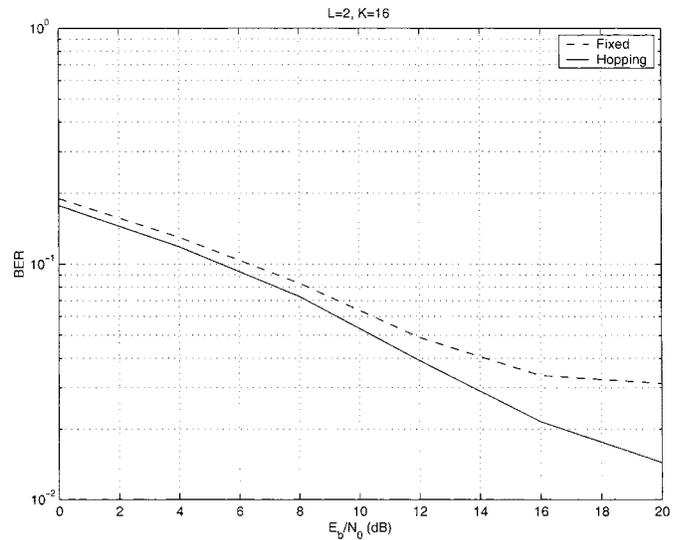


Fig. 6. Fixed versus hopping implementations.

preweighting at the expense of higher computational complexity at the receiver.

*Example 3 (Performance Improvements by Hopping  $\rho_k$ 's):* When the two channels are deeply faded at some  $\rho_k$ , i.e.,  $H_1(\rho_k) \approx 0$  and  $H_2(\rho_k) \approx 0$ , the decision vector in (14) may have low SNR and is not reliable. In the extreme case when  $H_1(\rho_k) = H_2(\rho_k) = 0$ , symbol recovery becomes impossible. Unlike ST-OFDM in [1], [5], and [6] where  $\rho_k$ s are fixed as in (7), ST-GOFDM is flexible to choose different  $\{\rho_k\}_{k=1}^K$ . Without CSI available at the transmitter, we test here a “root-rotating” approach, where  $\rho_k$ s are still equispaced around the unit circle but they rotate clockwise by a small angle  $\Delta_\theta$  every two blocks. By hopping the  $\rho_k$ s,  $H_i(\rho_k)$ s change every two blocks so that consistent deep fading for a certain  $\mathbf{s}_k(n)$  is avoided. We choose  $K = 16$ ,  $\Delta_\theta = \pi/64$ , and channels  $H_1(z) = (1 + jz^{-1})(1 + 0.5z^{-1})$  and  $H_2(z) = (1 + jz^{-1})(1 + 0.8z^{-1})$ , which share

a common zero at  $z = -j$ . Fig. 6 verifies that hopping the  $\rho_k$ 's exhibits better performance than fixed-root ST-OFDM. Note that when  $\rho_k$ s are rotated by  $\Delta_\theta$ , we have  $\mathbf{C} = \sqrt{(K)}\mathbf{T}_{zp}\mathbf{F}\mathbf{D}_\theta$  and  $\mathbf{V} = \mathbf{D}_\theta^H\mathbf{F}^H\mathbf{R}_{oa}$ , where  $\mathbf{D}_\theta := \text{diag}(1, \exp(j\Delta_\theta), \dots, \exp(j\Delta_\theta(K-1)))$ . Thus, low-complexity FFT operations can still be applied in this frequency-hopping ST-GOFDM system.

*Example 4 (Blind Channel Estimation and Its CRB):* To simulate the performance of our blind channel estimation algorithm in Section IV, we choose MR channels of order  $L = 2$  and  $K = 16$ . For estimating channel ratios,  $N = 64$  received blocks are used for each  $k$ . The performance of our system with estimated channels is shown against that with perfect CSI. The results are shown in Fig. 7 where we observe that the blind method entails a small penalty ( $< 2$  dB) in the overall system performance. Furthermore, we obtain the MSE of channel estimates and compute their CRB when  $H_1(\rho_1)$  and  $H_2(\rho_1)$  are known.

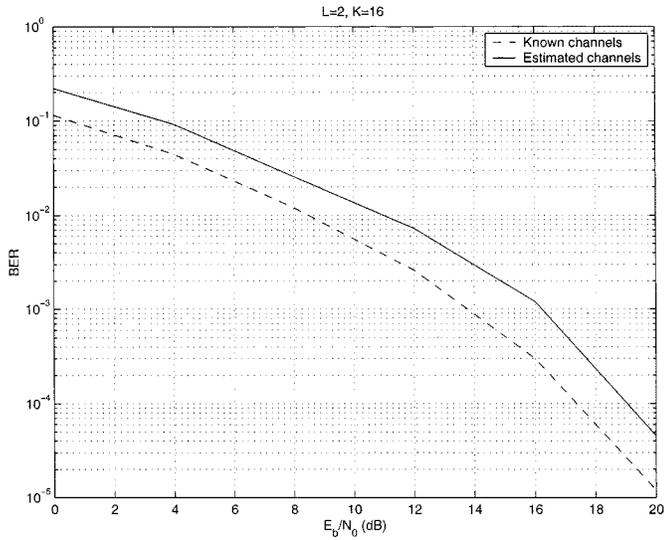


Fig. 7. Known versus estimated channels.

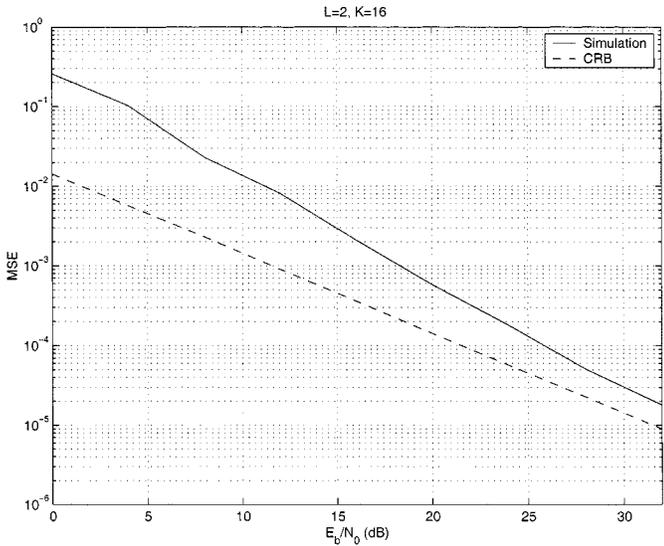
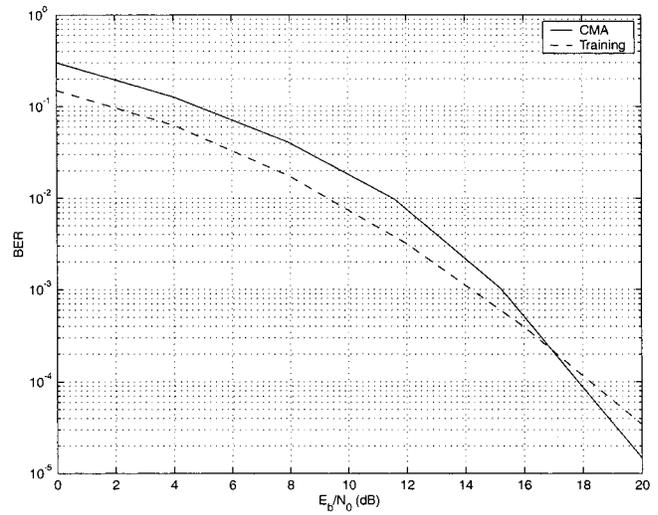


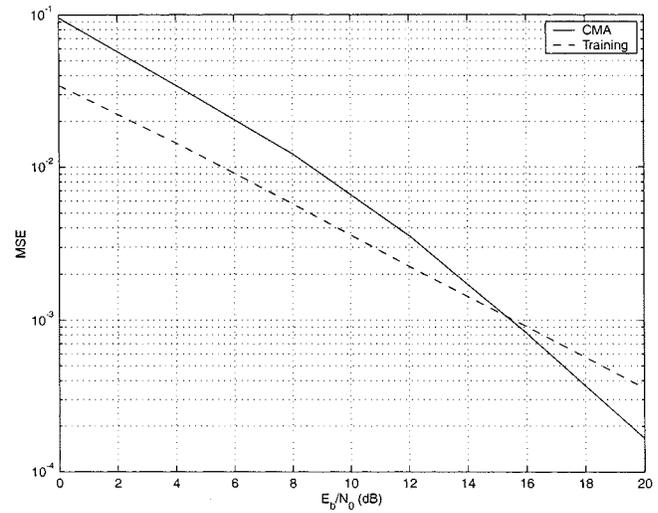
Fig. 8. Channel MSE and its CRB.

The curves depicted in Fig. 8 illustrate that the MSE of our channel estimates is very close to the CRB at high SNR.

*Example 5 (Comparison with Training-Based Channel Estimator):* We implement the training-based channel estimator [6] in our ST-GOFDM and compare its performance to our CMA blind estimator. TU channels ( $L = 4$ ) are used and the block length  $K = 32$  is chosen. To estimate the two channels, two training symbol blocks of size  $K \times 1$  (two OFDM symbols in [6]) are transmitted every 20 blocks. Thus, there is 10% overhead introduced in the training-based approach. For our CMA channel estimator,  $N = 20$  blocks of data are used to estimate the two channels with two training symbols sent every 20 blocks. The comparison in terms of BER and MSE of the channel estimates are depicted in Fig. 9(a) and 9(b), respectively. It is observed that the training-based channel estimator slightly outperforms the CMA channel estimator below



(a) BER



(b) Channel MSE

Fig. 9. Blind versus training-based GSM channel estimation.

SNR  $\approx 16$  dB. However, our blind channel estimator is better in the moderate to high SNR range. It is worthwhile to remark that, since  $K \gg 1$ , the training-based channel estimator induces a considerable loss in transmission rate.

## VII. CONCLUSION

In this paper, we proposed a novel ST-GOFDM transceiver suitable for frequency-selective multipath channels. Relying on symbol blocking, ST block codes designed for flat fading channels were extended to frequency-selective channels. By exploiting the specific structure of ST block codes, a blind channel estimation algorithm was also developed and compared with the CRB. In addition to the simplicity of the proposed transceiver, numerical simulations demonstrated superior performance over competing alternatives in simulated Rayleigh and typical urban GSM channels.

Because symbol recovery and blind channel estimation in this paper depend on channel zero locations, we currently investigate ST-coded transceiver designs irrespective of the channel zeros. Preliminary results relying on symbol blocking and long codes are reported in [7]. Other ongoing research topics include ST trellis coding for GOFDM systems and time-selective propagation of ST coded transmissions.

APPENDIX I  
PROOF OF LEMMA 1

Without loss of generality, we suppose that the  $2L + 2$  same-side entries are in the  $\tilde{r}$  group; namely, we have  $\{r(k_j) = \tilde{r}(k_j) = H_1^*(\rho_{k_j})/H_2(\rho_{k_j}), 1 \leq k_j \leq 4L + 3\}_{j=1}^{2L+2}$ . Enforcing the finite-channel support in Assumption a1), we find from the definition of  $H_i(\rho_{k_j})$ ,  $i = 1, 2$ , that  $\forall j \in [1, 2L + 2]$

$$\mathbf{v}^T(\rho_{k_j}^*, L + 1)\mathbf{h}_1^* - r(k_j)\mathbf{v}^T(\rho_{k_j}, L + 1)\mathbf{h}_2 = 0. \quad (40)$$

The homogeneous equations in (40) can be cast in the following matrix form

$$\begin{pmatrix} \mathbf{v}^T(\rho_{k_1}^*, L + 1) & -r(k_1)\mathbf{v}^T(\rho_{k_1}, L + 1) \\ \vdots & \vdots \\ \mathbf{v}^T(\rho_{k_{2L+2}}^*, L + 1) & -r(k_{2L+2})\mathbf{v}^T(\rho_{k_{2L+2}}, L + 1) \end{pmatrix} \times \begin{pmatrix} \mathbf{h}_1^* \\ \mathbf{h}_2 \end{pmatrix} = \boldsymbol{\Theta} \begin{pmatrix} \mathbf{h}_1^* \\ \mathbf{h}_2 \end{pmatrix} = 0. \quad (41)$$

Under Assumption a4), we now prove that  $\boldsymbol{\Theta}$  has nullity one by contradiction. Suppose  $\boldsymbol{\Theta}$  has nullity at least two, namely, we can find two different  $(2L + 2) \times 1$  vectors  $\mathbf{h} = (\mathbf{h}_1^T, \mathbf{h}_2^T)^T$  and  $\boldsymbol{\xi} = (\mathbf{b}^T, \mathbf{a}^T)^T$  which satisfy:  $\boldsymbol{\Theta}\mathbf{h} = \mathbf{0}$  and  $\boldsymbol{\Theta}\boldsymbol{\xi} = \mathbf{0}$ . Writing the latter component-wise and cancelling  $r(k_j)$ , we arrive at

$$H_1^*(\rho_{k_j})A(\rho_{k_j}) = H_2(\rho_{k_j})B^*(\rho_{k_j}), \quad \forall j \in [1, 2L + 2] \quad (42)$$

where  $B(\rho_{k_j}) := \mathbf{v}^T(\rho_{k_j}, L + 1)\mathbf{b}$  and  $A(\rho_{k_j}) := \mathbf{v}^T(\rho_{k_j}, L + 1)\mathbf{a}$ . Denote by  $B(z)$  and  $A(z)$  the  $\mathcal{Z}$ -transforms of  $\mathbf{b}$  and  $\mathbf{a}$ , respectively. Recall that  $\mathbf{b}$  and  $\mathbf{a}$  have length  $L + 1$  so that  $B(z)$  and  $A(z)$  have degree  $L$ . Because  $\rho_{k_j}$ ,  $j = 1, \dots, 2L + 2$  are chosen to be distinct, (42) prescribes  $H_1^*(z)A(z) = H_2(z)B^*(z)$  at  $2L + 2$  different points in the  $\mathcal{Z}$  domain. Thus, we can obtain from (42) that  $H_1^*(z)A(z) = H_2(z)B^*(z)$ . Under Assumption a4), the latter implies that  $B(z) = \alpha H_1(z)$  and  $A(z) = \alpha H_2(z)$  where  $\alpha$  stands for a scalar ambiguity. Hence,  $\mathbf{h} = \alpha \boldsymbol{\xi}$ , and  $\boldsymbol{\Theta}$  has nullity one so that  $\mathbf{h}_{12} = (\mathbf{h}_1^T, \mathbf{h}_2^T)^T$  (within a scale) can be identified as the unique nullvector of  $\boldsymbol{\Theta}$ . Likewise, when the  $2L + 2$  same-side entries come from the  $\tilde{r}'$  group, the unique nullvector of  $\boldsymbol{\Theta}$  will yield  $\mathbf{h}_{21} = (-\mathbf{h}_2^T, \mathbf{h}_1^T)^T$ . ■

APPENDIX II  
DERIVATION OF  $\mathcal{F}(\boldsymbol{\theta})$

In this appendix, we compute the entries of  $\mathcal{F}(\boldsymbol{\theta})$ . Defining  $\boldsymbol{\rho} := (\rho_1, \dots, \rho_K)^T$ , it follows by definition that  $H_i(\rho_k) := \mathbf{v}^T(\rho_k, L + 1)\mathbf{h}_i$ ,  $i = 1, 2$ , and

$$\begin{aligned} \frac{\partial \mathcal{D}(\tilde{\mathbf{h}}_i, K)}{\partial h_j(l)} &= \mathcal{D}^{\mathcal{H}}(\boldsymbol{\rho}, \mathbf{K})\delta(i - j) \\ \frac{\partial \mathcal{D}^{\mathcal{H}}(\tilde{\mathbf{h}}_i, K)}{\partial h_j^*(l)} &= \mathcal{D}(\boldsymbol{\rho}, \mathbf{K})\delta(i - j) \\ \frac{\partial \mathcal{D}(\tilde{\mathbf{h}}_i, K)}{\partial h_j^*(l)} &= 0 \\ \frac{\partial \mathcal{D}^{\mathcal{H}}(\tilde{\mathbf{h}}_i, K)}{\partial h_j(l)} &= 0. \end{aligned}$$

Taking partial derivatives of (27) with respect to  $\boldsymbol{\theta}$ , we obtain

$$\begin{aligned} -\frac{\partial f(\boldsymbol{\theta})}{\partial h_1(l)} &= \frac{1}{\sigma^2} \sum_{n=0}^{N-1} (-\tilde{\mathbf{w}}^{\mathcal{H}}(2n)\mathcal{D}^{\mathcal{H}}(\boldsymbol{\rho}, K)\mathbf{s}(2n) \\ &\quad + \tilde{\mathbf{w}}^{\mathcal{H}}(2n+1)\mathcal{D}^{\mathcal{H}}(\boldsymbol{\rho}, K)\mathbf{s}^*(2n+1)) \\ -\frac{\partial f(\boldsymbol{\theta})}{\partial h_2(l)} &= \frac{1}{\sigma^2} \sum_{i=1}^N (-\tilde{\mathbf{w}}^{\mathcal{H}}(2n)\mathcal{D}^{\mathcal{H}}(\boldsymbol{\rho}, K)\mathbf{s}(2n+1) \\ &\quad - \tilde{\mathbf{w}}^{\mathcal{H}}(2n+1)\mathcal{D}^{\mathcal{H}}(\boldsymbol{\rho}, K)\mathbf{s}^*(2n)) \\ -\frac{\partial f(\boldsymbol{\theta})}{\partial s_{1k}(n)} &= \frac{1}{\sigma^2} (-\tilde{\mathbf{w}}^{\mathcal{H}}(2n)\mathcal{D}(\tilde{\mathbf{h}}_1, K)\mathbf{e}_k \\ &\quad - \mathbf{e}_k^T \mathcal{D}^{\mathcal{H}}(\tilde{\mathbf{h}}_2, K)\tilde{\mathbf{w}}(2n+1)) \\ -\frac{\partial f(\boldsymbol{\theta})}{\partial s_{2k}(n)} &= \frac{1}{\sigma^2} (-\tilde{\mathbf{w}}^{\mathcal{H}}(2n)\mathcal{D}(\tilde{\mathbf{h}}_2, K)\mathbf{e}_k \\ &\quad + \mathbf{e}_k^T \mathcal{D}^{\mathcal{H}}(\tilde{\mathbf{h}}_1, K)\tilde{\mathbf{w}}(2n+1)) \\ \frac{\partial f(\boldsymbol{\theta})}{\partial h_i^*(l)} &= \left( \frac{\partial f(\boldsymbol{\theta})}{\partial h_i(l)} \right)^* \\ \frac{\partial f(\boldsymbol{\theta})}{\partial s_{ik}^*(n)} &= \left( \frac{\partial f(\boldsymbol{\theta})}{\partial s_{ik}(l)} \right)^* \end{aligned}$$

where  $s_{ik}(n) := s((2n + i - 1)K + k - 1)$ , and  $\mathbf{e}_k$  is the  $i$ th  $K \times 1$  canonical basis vector having all entries equal to zero except the  $k$ th one. Because the noise is white and circularly symmetric, we arrive at

$$\begin{aligned} E \left\{ \frac{\partial f(\boldsymbol{\theta})}{\partial h_i^*(l)} \frac{\partial f(\boldsymbol{\theta})}{\partial h_{i'}(l')} \right\} &= \delta(i - i') \frac{1}{\sigma^2} \sum_{k=1}^K \sum_{n=0}^{N-1} \rho_k^{l-l'} \\ &\quad \times (|s_{1k}(n)|^2 + |s_{2k}(n)|^2) \\ E \left\{ \frac{\partial f(\boldsymbol{\theta})}{\partial h_i^*(l)} \frac{\partial f(\boldsymbol{\theta})}{\partial h_{i'}^*(l')} \right\} &= E \left\{ \frac{\partial f(\boldsymbol{\theta})}{\partial h_i(l)} \frac{\partial f(\boldsymbol{\theta})}{\partial h_{i'}(l')} \right\} = 0 \\ E \left\{ \frac{\partial f(\boldsymbol{\theta})}{\partial s_{ik}^*(n)} \frac{\partial f(\boldsymbol{\theta})}{\partial s_{i'k'}(n')} \right\} &= \delta(n - n') \delta(i - i') \delta(k - k') \frac{1}{\sigma^2} \\ &\quad \times (|H_1(\rho_k)|^2 + |H_2(\rho_k)|^2) \\ E \left\{ \frac{\partial f(\boldsymbol{\theta})}{\partial s_{ik}^*(n)} \frac{\partial f(\boldsymbol{\theta})}{\partial s_{i'k'}(n')} \right\} &= E \left\{ \frac{\partial f(\boldsymbol{\theta})}{\partial s_{ik}(n)} \frac{\partial f(\boldsymbol{\theta})}{\partial s_{i'k'}(n')} \right\} = 0 \\ E \left\{ \frac{\partial f(\boldsymbol{\theta})}{\partial h_1^*(l)} \frac{\partial f(\boldsymbol{\theta})}{\partial s_{ik}(n)} \right\} &= \frac{1}{\sigma^2} s_{1k}^*(n) \rho_k^l H_i(\rho_k) \end{aligned}$$

$$\begin{aligned}
E \left\{ \frac{\partial f(\boldsymbol{\theta})}{\partial h_1^*(l)} \frac{\partial f(\boldsymbol{\theta})}{\partial s_{1k}^*(n)} \right\} &= -\frac{1}{\sigma^2} s_{2k}(n) \rho_k^l H_2(\rho_k) \\
E \left\{ \frac{\partial f(\boldsymbol{\theta})}{\partial h_1^*(l)} \frac{\partial f(\boldsymbol{\theta})}{\partial s_{2k}^*(n)} \right\} &= \frac{1}{\sigma^2} s_{2k}(n) \rho_k^l H_1(\rho_k) \\
E \left\{ \frac{\partial f(\boldsymbol{\theta})}{\partial h_2^*(l)} \frac{\partial f(\boldsymbol{\theta})}{\partial s_{ik}^*(n)} \right\} &= \frac{1}{\sigma^2} s_{2k}^*(n) \rho_k^l H_i(\rho_k) \\
E \left\{ \frac{\partial f(\boldsymbol{\theta})}{\partial h_2^*(l)} \frac{\partial f(\boldsymbol{\theta})}{\partial s_{1k}^*(n)} \right\} &= \frac{1}{\sigma^2} s_{1k}(n) \rho_k^l H_2(\rho_k) \\
E \left\{ \frac{\partial f(\boldsymbol{\theta})}{\partial h_2^*(l)} \frac{\partial f(\boldsymbol{\theta})}{\partial s_{2k}^*(n)} \right\} &= -\frac{1}{\sigma^2} s_{1k}(n) \rho_k^l H_1(\rho_k).
\end{aligned}$$

Let us define

$$\begin{aligned}
\mathbf{B}_1 &:= \sum_{n=0}^{N-1} \sum_{k=1}^K \sum_{i=1}^2 |s((2n+i-1)K+k-1)|^2 \\
&\quad \times v^T(\rho_k, L+1) \otimes v^*(\rho_k, L+1) \\
\mathbf{B}_2 &:= (v^*(\rho_1, L+1), \dots, v^*(\rho_K, L+1)) \\
\mathbf{B}_3 &:= \text{diag}(|H_1(\rho_1)|^2 + |H_2(\rho_1)|^2, \dots, \\
&\quad |H_1(\rho_K)|^2 + |H_2(\rho_K)|^2) \\
\mathcal{D}_{s_1} &:= (\mathcal{D}(s(0), K) \mathcal{D}(s(2), K), \dots, \mathcal{D}(s(2N-2), K)) \\
\mathcal{D}_{s_2} &:= (\mathcal{D}(s(1), K) \mathcal{D}(s(3), K), \dots, \mathcal{D}(s(2N-1), K)) \\
\mathbf{B}_4 &:= \begin{pmatrix} \mathbf{B}_2 \mathcal{D}(\tilde{\mathbf{h}}_1, K) \mathcal{D}_{s_1}^* & \mathbf{B}_2 \mathcal{D}(\tilde{\mathbf{h}}_2, K) \mathcal{D}_{s_1}^* \\ \mathbf{B}_2 \mathcal{D}(\tilde{\mathbf{h}}_1, K) \mathcal{D}_{s_2}^* & \mathbf{B}_2 \mathcal{D}(\tilde{\mathbf{h}}_2, K) \mathcal{D}_{s_2}^* \\ -\mathbf{B}_2^* \mathcal{D}^*(\tilde{\mathbf{h}}_2, K) \mathcal{D}_{s_2}^* & \mathbf{B}_2^* \mathcal{D}^*(\tilde{\mathbf{h}}_1, K) \mathcal{D}_{s_2}^* \\ \mathbf{B}_2^* \mathcal{D}^*(\tilde{\mathbf{h}}_2, K) \mathcal{D}_{s_1}^* & -\mathbf{B}_2^* \mathcal{D}^*(\tilde{\mathbf{h}}_1, K) \mathcal{D}_{s_1}^* \end{pmatrix} \\
\mathbf{B}_5 &:= \begin{pmatrix} -\mathbf{B}_2 \mathcal{D}(\tilde{\mathbf{h}}_2, K) \mathcal{D}_{s_2} & \mathbf{B}_2 \mathcal{D}(\tilde{\mathbf{h}}_1, K) \mathcal{D}_{s_2} \\ \mathbf{B}_2 \mathcal{D}(\tilde{\mathbf{h}}_2, K) \mathcal{D}_{s_1} & -\mathbf{B}_2 \mathcal{D}(\tilde{\mathbf{h}}_1, K) \mathcal{D}_{s_1} \\ \mathbf{B}_2^* \mathcal{D}^*(\tilde{\mathbf{h}}_1, K) \mathcal{D}_{s_1} & \mathbf{B}_2^* \mathcal{D}^*(\tilde{\mathbf{h}}_2, K) \mathcal{D}_{s_1} \\ \mathbf{B}_2^* \mathcal{D}^*(\tilde{\mathbf{h}}_1, K) \mathcal{D}_{s_2} & \mathbf{B}_2^* \mathcal{D}^*(\tilde{\mathbf{h}}_2, K) \mathcal{D}_{s_2} \end{pmatrix} \quad (43)
\end{aligned}$$

It follows from (28) that the FIM  $\mathcal{F}(\boldsymbol{\theta})$  is given by (28).

We next prove that the matrix inverse in (32) does not exist. Due to the structure of  $\mathcal{G}_h$  in (31), the matrix  $\mathcal{U}_h$  can be obtained as

$$\mathcal{U}_h = \begin{pmatrix} \mathcal{V}_h & \mathbf{0}_{(4L+4) \times 4KN} \\ \mathbf{0}_{4KN \times (4L+4)} & \mathbf{I}_{4KN \times 4KN} \end{pmatrix} \quad (44)$$

where  $\mathcal{V}_h$  is the nonnull matrix whose columns lie in the right null space of  $\mathcal{E}_h$ , which has dimensionality  $4L+4-M > 0$ . Using (28), (44), it can be readily shown that

$$\begin{aligned}
\det(\mathcal{U}_h^H \mathcal{F}(\boldsymbol{\theta}) \mathcal{U}_h) &= \det(\mathcal{F}_{22}) \det(\mathcal{V}_h^H (\mathcal{F}_{11} - \mathcal{F}_{12} \mathcal{F}_{22}^{-1} \mathcal{F}_{12}^H) \mathcal{V}_h) \\
&\equiv 0. \quad (45)
\end{aligned}$$

Thus, the matrix inverse in (32) does not exist. ■

## REFERENCES

- [1] D. Agrawal, V. Tarokh, A. Naguib, and N. Seshadri, "Space-time coded OFDM for high data-rate wireless communication over wideband channels," in *Proc. Vehicular Technology Conf.*, Ottawa, ON, Canada, May 18–21, 1998, pp. 2232–2236.
- [2] S. M. Alamouti, "A simple transmit diversity technique for wireless communications," *IEEE J. Select. Areas Commun.*, vol. 16, pp. 1451–1458, Oct. 1998.

- [3] P. Comon and G. Golub, "Tracking a few extreme singular values and vectors in signal processing," *Proc. IEEE*, vol. 78, pp. 1327–1343, Aug. 1990.
- [4] G. B. Giannakis, Z. Wang, A. Scaglione, and S. Barbarossa, "AMOUR—generalized multicarrier transceivers for blind CDMA regardless of multipath," *IEEE Trans. Commun.*, vol. 48, pp. 2064–2076, Dec. 2000.
- [5] Y. Li, J. C. Chung, and N. R. Sollenberger, "Transmitter diversity for OFDM systems and its impact on high-rate data wireless networks," *IEEE J. Select. Areas Commun.*, vol. 17, pp. 1233–1243, July 1999.
- [6] Y. Li, N. Seshadri, and S. Ariyavisitakul, "Channel estimation for OFDM systems with transmitter diversity in mobile wireless channels," *IEEE J. Select. Areas Commun.*, vol. 17, pp. 461–471, Mar. 1999.
- [7] Z. Liu and G. B. Giannakis, "Space-time block coded multiple access through frequency-selective fading Channels," *IEEE Trans. Commun.*, June 2001, to be published.
- [8] Z. Liu, G. B. Giannakis, B. Muquet, and S. Zhou, "Space-time coding for broadband wireless communications," *Wireless Syst. Mobile Comput.*, vol. 1, no. 1, pp. 33–53, Jan. 2001.
- [9] Z. Liu, G. B. Giannakis, A. Scaglione, and S. Barbarossa, "Block precoding and transmit-antenna diversity for decoding and equalization of unknown multipath channels," in *Proc. Asilomar Conf. Signals, Systems and Computers*, Pacific Grove, CA, Nov. 1–4, 1999, pp. 1557–1561.
- [10] B. Muquet, Z. Wang, G. B. Giannakis, M. de Courville, and P. Duhamel, "Cyclic prefix or zero-padding for multi-carrier transmissions?," in *Proc. Int. Conf. Communications*, New Orleans, LA, June 18–22, 2000, pp. 1049–1053.
- [11] A. F. Naguib, N. Seshadri, and R. Calderbank, "Space-time coding and signal processing for high data rate wireless communications," *IEEE Signal Processing Mag.*, vol. 17, pp. 76–92, May 2000.
- [12] A. Scaglione and G. B. Giannakis, "Design of user codes in QS-CDMA systems for MUI elimination in unknown multipath," *IEEE Commun. Lett.*, vol. 3, pp. 25–27, Feb. 1999.
- [13] A. Scaglione, G. B. Giannakis, and S. Barbarossa, "Lagrange/Vandermonde MUI eliminating user codes for quasynchronous CDMA in unknown multipath," *IEEE Trans. Signal Processing*, vol. 48, pp. 2057–2073, July 2000.
- [14] P. Stoica and B. C. Ng, "On the Cramér–Rao bound under parametric constraints," *IEEE Signal Processing Lett.*, vol. 5, pp. 177–179, July 1998.
- [15] V. Tarokh, H. Jafarkhani, and A. R. Calderbank, "Space-time block codes from orthogonal designs," *IEEE Trans. Inform. Theory*, vol. 45, pp. 1456–1467, July 1999.
- [16] V. Tarokh, N. Seshadri, and A. R. Calderbank, "Space-time codes for high data rate wireless communication: performance criterion and code construction," *IEEE Trans. Inform. Theory*, vol. 44, pp. 744–765, Mar. 1998.
- [17] A. van den Bos, "A Cramér–Rao lower bound for complex parameters," *IEEE Trans. Signal Processing*, vol. 42, p. 2859, Oct. 1994.
- [18] A. J. van der Veen and A. Paulraj, "An analytical constant modulus algorithm," *IEEE Trans. Signal Processing*, vol. 44, pp. 1136–1155, May 1996.
- [19] Z. Wang and G. B. Giannakis, "Wireless multicarrier communications: Where Fourier meets Shannon," *IEEE Signal Processing Mag.*, vol. 17, pp. 29–48, May 2000.



**Zhiqiang Liu** received the B.S. degree from the Department of Radio and Electronics, Peking University, Beijing, China, in 1991 and the M.E. degree from the Institute of Electronics, Chinese Academy of Science, Beijing, China, in 1994. He received the Ph.D. degree from the Department of Electrical and Computer Engineering, University of Minnesota, Minneapolis–St. Paul.

From 1995 to 1997, he was with the Department of Electrical Engineering, National University of Singapore as a Research Scholar. From 1997 to 1998, he was a Research Assistant with the University of Virginia, Charlottesville. His research interests include space–time coding, wideband wireless communication systems, multiuser detection, multicarrier, and blind channel estimation algorithms.



**Georgios B. Giannakis** (S'84–M'86–SM'91–F'97) received the Dipl. degree in electrical engineering from the National Technical University of Athens, Greece, in 1981. He received the MSc. degree in electrical engineering, the MSc. degree in mathematics, and the Ph.D. degree in electrical engineering from the University of Southern California (USC), Los Angeles, in 1983, 1986, and 1986, respectively.

After lecturing for one year at USC, he joined the University of Virginia (UVA) in 1987, where he became a Professor of electrical engineering in 1997,

Graduate Committee Chair, and Director of the Communications, Controls, and Signal Processing Laboratory in 1998. Since January 1999, he has been with the University of Minnesota, Minneapolis-St. Paul, as a Professor of electrical and computer engineering. His research interests include 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 120 journal papers and about 250 conference papers. Specific areas of expertise have included (poly)spectral analysis, wavelets, cyclostationary, and non-Gaussian signal processing with applications to SAR, array, and image processing. Current research topics focus on transmitter and receiver diversity techniques for single- and multiuser fading communication channels, compensation of nonlinear amplifier effects, redundant precoding and space-time coding for block transmissions, multicarrier, and wideband wireless communication systems.

Dr. Giannakis has received three best paper awards from the IEEE Signal Processing Society in 1992, 1998, and 2000, respectively. He also received the Society's Technical Achievement Award in 2000. He co-organized three IEEE-SP Workshops (Higher Order Statistics in 1993, Statistical Signal and Array Processing in 1996, and SP Advances in Wireless Communications in 1997) and guest (co-)edited four special issues. He has served as an Associate Editor for the IEEE TRANSACTIONS ON SIGNAL PROCESSING and the IEEE SIGNAL PROCESSING LETTERS, a secretary of the SP Conference Board, a member of the SP Publications Board, and a member and vice-chair of the SSAP Technical Committee. He is a member of the Editorial Board for the PROCEEDINGS OF THE IEEE, he chairs the SP for Communications Technical Committee, and serves as the Editor in Chief for the IEEE SIGNAL PROCESSING LETTERS. He is a member of the IEEE Fellows Election Committee, the IEEE-SP Society's Board of Governors, and a frequent consultant for the telecommunications industry.



**Sergio Barbarossa** (M'88) received the M.Sc. and Ph.D. degrees in electrical engineering from the University of Rome "La Sapienza," Rome, Italy, in 1984 and 1988, respectively.

From 1984 to 1986, he was a Radar System Engineer with Selenia, Rome, Italy. Since 1986, he has been with the Information and Communication Department, University of Rome "La Sapienza," where he is a Professor. In 1988, on leave from the University of Rome, he was a Research Engineer with the Environmental Research Institute of Michigan, Ann Arbor. During the spring of 1995 and the summer of 1997, he was a Visiting Faculty Member with the Department of Electrical Engineering, University of Virginia, Charlottesville, and from March 1999 to August 1999, he was a Visiting Professor with the University of Minnesota, Minneapolis. He has been involved in several international projects concerning phased array radars, synthetic aperture radars, and mobile communication systems. His current research interests are in the area of broadband mobile communication systems with specific attention to multiple access systems, optimal coding, synchronization, channel estimation, tracking, and equalization.

Dr. Barbarossa is a member of the IEEE Technical Committee on Signal Processing for Communications and is currently serving as an Associate Editor for the IEEE TRANSACTIONS ON SIGNAL PROCESSING.



**Anna Scaglione** (M'99) received the M.Sc. and Ph.D. degrees in 1995 and 1999, respectively, from the University of Rome "La Sapienza," Italy.

During 1997, she was with the University of Virginia as Visiting Research Assistant. In 1999, she was with the University of Minnesota as Postdoctoral Researcher. She recently joined the University of New Mexico, Albuquerque, as an Assistant Professor. Her research interests include signal processing for communications and her work has focused on optimal design of modulation and multiplexing techniques for

frequency selective, time varying channels.