

Transmit Diversity and Linear and Decision-Feedback Equalizations for Frequency-Selective Fading Channels

Ling Li, *Student Member, IEEE*, Yu-Dong Yao, *Senior Member, IEEE*, and Hongbin Li, *Member, IEEE*

Abstract—Rapid growth and increasing demands for near-ubiquitous high-quality high-data-rate services present the most challenges for wireless system design. As an effective method to provide such services, space-time (ST) coding is gaining more and more attention. This paper extends ST coding, originally designed for known frequency-nonsselective fading channels, to unknown frequency-selective channels. A novel scheme is presented to suppress intersymbol interference, coherently demodulate the information symbols with estimated channel state information in an ST transmit diversity wireless time-division multiple-access system that is equipped with multiple antennas at both transmit and receive sides. The proposed algorithm is powerful and computationally efficient. In addition to the discussion of system identifiability, both theoretical analysis and numerical simulation are presented to illustrate the performance of the proposed estimator and receiver in multipath fading channels.

Index Terms—Equalization, estimation, space-time (ST) coding, transmit diversity.

I. INTRODUCTION

RECENT advances in communications are driven by the requirements of next-generation wireless systems to provide high-data-rate and high-quality services anywhere at anytime. But the spectrum continues to be scarce and expensive, which creates new challenges in the development of telecommunication systems [1].

Wireless cellular systems are known to suffer from fading and intersymbol interference (ISI) caused by multipath propagation, which may degrade the system performance significantly. Therefore, the effective and efficient interference mitigation is required for a high-quality signal reception. While various work on such a topic has been done over recent years [2] in a single antenna system, spatial diversity (receive and/or transmit diversity [3], [4]) is now considered the enabling technique to realize the target of the wireless systems. It is proven that spatial receive diversity, realized by deploying multiple antennas at the receive side, can be used to effectively cancel interference without bandwidth expansion [5]–[7].

Manuscript received May 29, 2001; revised April 18, 2003. This work was supported in part by the New Jersey Center for Wireless Telecommunications (NJCWT) and the U.S. Army Research Office (ARO) under contract DAAD19-03-1-0184. This paper was presented in part at the IEEE Global Telecommunications Conference, San Antonio, TX, November 25-29, 2001.

The authors are with the Department of Electrical and Computer Engineering, Stevens Institute of Technology, Hoboken, NJ 07030 USA (e-mail: lli1@stevens-tech.edu; yyao@stevens-tech.edu; hli@stevens-tech.edu).

Digital Object Identifier 10.1109/TVT.2003.816613

While the algorithms using receive diversity are well documented, transmit diversity is a relatively new and attractive topic that is, in general, less straightforward to exploit when there is no feedback path to provide the transmitter with knowledge of the channel parameters. In the applications where receive diversity is applicable, transmit diversity can also be exploited together with receive diversity to further improve performance. Recent research in information theory has shown that large gains in capacity and reliability of communications over wireless channels could be achieved by exploiting the spatial diversity with multiple antennas at both the transmit and receive sides, which forms a multiple-input multiple-output (MIMO) system [8], [9].

Recently, various approaches for transmit diversity have been suggested. A delay diversity scheme was proposed by Wittneben [10], [11] for base-station simulcasting, in which more than one base station is utilized to send identical information. Later, a similar scheme was suggested by Winters [3] for a single base station in which copies of the same symbol are transmitted through multiple antennas at different time intervals. Another approach, known as BLAST (Bell Labs Layered Space-Time) [12], [13], features a layered architecture, which can achieve massive parallel transmission and very high data rates in a rich-scattering system by using a large number of antennas at the transmitter and at the receiver. The BLAST approach has reasonable complexity; however, its performance is not optimized for diversity and coding gain. It also suffers from error propagation.

Though the information capacity of wireless communication systems increases dramatically by employing multiple transmit and receive antennas, increasing the signal quality or reducing the effective error rate in a multipath fading channel is still a challenging issue. Consequently, space-time (ST) trellis coding relying on multiple antenna transmission was introduced in [14] as a blind transmit-coding scheme without requiring channel-state information (CSI) at the transmit side. Through grouping and encoding successive symbols, the coded symbols are sent simultaneously via an antenna array. By exploiting spatial and temporal diversity along with channel coding, it can attain diversity and coding gain with a minimal impact on existing spectral utilization. Besides the capacity improvement, it is very effective and able to provide good bit-error rate (BER) performance. However, due to the use of a maximum likelihood (ML) decoder implemented with the Viterbi algorithm, the decoding complexity (measured by the number of trellis states in the decoder) increases exponentially with transmission rate when the number of transmit antennas is fixed [4]. Thus, it has limited practical use, especially in high-data-rate applications in which

the complexity and capability of the transmitter and the receiver become an important issue.

ST block coding (STBC) [15], [16] appeared to be a simple alternative, which has the advantage of linear processing at the receive side. As an effective technique to combat fading in wireless communications, it guarantees good error performance over a broad range of channel realization while providing full diversity gain as the well-known maximal-ratio receiver combining (MRRC). The decreased sensitivity to fading may allow the use of higher level modulation schemes to increase the effective data rate or smaller reuse factors in a multicell environment to increase system capacity. As a breakthrough of wireless technologies, STBC substantially increases transmission rate and reliability, as well as the potential system capacity without any bandwidth expansion or any feedback from the receiver to the transmitter.

ST codes [14]–[16] were originally designed to provide a certain diversity order assuming known flat-fading channels. However, the assumption of flat fading is not always justified. Especially, wideband high-data-rate transmissions may lead to severe frequency-selective channel fading, which corrupts the received signal as a result of a loss of orthogonality among ST block coded symbols. It thus necessitates the use of channel equalization to compensate for ISI caused by multipath in order to improve the signal reception performance. In addition, the interference becomes more serious in a MIMO system than in a single input single output (SISO) system because of interantenna interference. Therefore, effective suppression of multipath interference and interantenna interference is critical in an STTD system.

Furthermore, the channel information, which is utilized to decode the received signals and to restore the initial transmitted values, is unknown in practice and has to be estimated. Many research efforts over recent years focus on the areas in which *a priori* knowledge is not available to the receivers [17] and the desired information is estimated and detected blindly. However, in some applications, especially in a mobile communication system, *a priori* knowledge is known to the receivers, although the actual transmitted symbol stream is unknown. In such a system, a known preamble is added to the message for training purpose. Such extra information may be exploited to enhance the accuracy of the estimates and may be used to simplify the computational complexity. Channel estimation in an STTD system, however, is more challenging than that in a single-antenna system since the number of unknown channel coefficients increases proportionally to the number of transmit antennas. Therefore, effective and efficient channel estimation schemes are critically important. Although the recently proposed differential ST coding algorithms for frequency-nonselective channels (e.g., [18]–[21] and references therein) obviate the requirement for channel estimation and, therefore, are particularly attractive in fast fading environments when channel estimation becomes very difficult or even infeasible, differential decoding of ST codes suffers approximately a 3-dB penalty in signal-to-noise ratio (SNR) as compared to coherent decoding, which requires channel information. Hence, channel estimation is well motivated, especially in cases when the channel experiences relatively slow fading and channel estimation is more reliable.

This paper differs from previous work [14]–[16] in that it investigates a combined transmit diversity, STBC, and equal-

ization strategy. It presents effective channel estimation and ISI suppression techniques by taking advantages of the structure of STBC in a system with two transmit antennas and M receive antennas. With the proposed efficient computation methods, it is shown that a practical bandwidth-efficient transmit diversity strategy in conjunction with interference suppression techniques can improve system performance dramatically. The results presented in this paper can be easily extended to a system with more than two transmit antennas using general STBC discussed in [16].

In particular, the system model is formulated in Section II. Efficient ISI cancellation techniques are presented in Section III by exploiting the structure of STBC codes and considering different criteria, zero forcing (ZF) and minimum mean-squared error (MMSE) with or without decision feedback. The corresponding simplified computation methods of matrix multiplication and inversion are introduced in Appendices I and II, respectively. Conditions for effective symbol recovery are investigated as well in this section. In Section IV, theoretical analysis for system BER performance and its upper bound are explored to assess our proposed algorithms. Section V considers a method of CSI estimation in signal reception. Finally, numerical results are presented in Section VI to illustrate the performance of the proposed channel estimator and the receiver in multipath fading channels. Concluding remarks are contained in Section VII.

Notation: Vectors (matrices) are denoted by boldface lower (upper) case letters; all vectors are column vectors; superscripts $(\cdot)^*$, $(\cdot)^T$, and $(\cdot)^H$ denote the complex conjugate, the transpose, and the conjugate transpose, respectively; \mathbf{I}_N denotes the $N \times N$ identity matrix; the subscript N is omitted when the dimension of the matrix \mathbf{I} is obvious; $\mathbf{0}$ denotes an all-zero vector (matrix); $\lceil \cdot \rceil$ denotes the smallest integer no less than the argument; $\|\cdot\|$ denotes the vector (matrix) 2-norm; $E[\cdot]$ denotes the statistical expectation; \otimes denotes the matrix Kronecker product; $\text{tr}(\cdot)$ denotes the trace of a matrix; finally, $[\cdot]_{kk}$ denotes the kk th component of a matrix.

II. PROBLEM FORMULATION

A. System Model

We focus on a discrete-time baseband model. Consider a wireless cellular system equipped with K ($K \geq 2$) transmit antennas and M ($M > 1$) receive antennas over frequency-selective fading channels perturbed by additive white Gaussian noise (AWGN). In particular, we assume Alamouti's ST coding [15] using $K = 2$ transmit antennas. Extensions to other STBC systems with more than two transmit antennas are straightforward. Fig. 1 depicts a diagram of the base-band ST-coded system. At the transmit side, the ST encoder (detailed in Section II-B) maps the incoming symbol stream $\{s(n)\}$ drawn from a certain constellation \mathcal{S} into two ST-coded symbol streams $\{c_1(n)\}$ and $\{c_2(n)\}$. Then, the two-coded symbol streams are sent out through transmit antenna 1 and transmit antenna 2, simultaneously. At the receive side, the channel estimator produces a channel estimate, which is then utilized by the receiver for interference cancellation and symbol detection.

Let $r_m(n)$ be the received symbol collected by receive antenna m at the n th symbol; $c_1(n)$ and $c_2(n)$ be the n th symbol

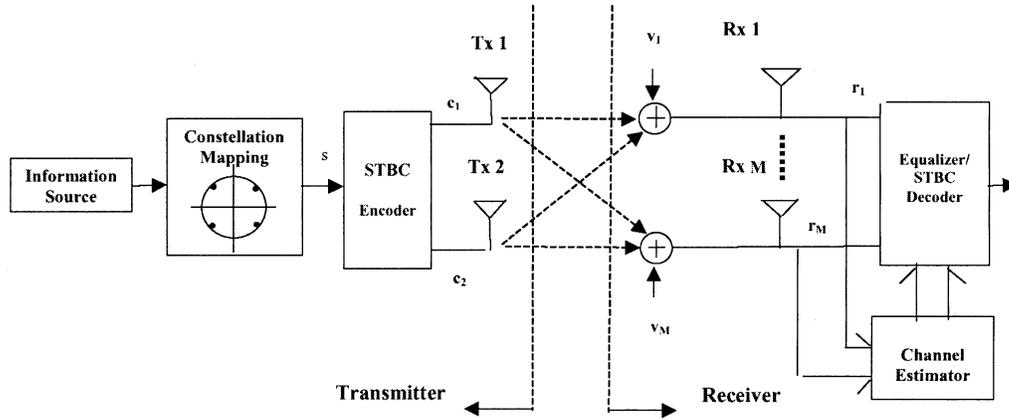


Fig. 1. Base-band ST-coded system.

sent from transmit antennas 1 and 2, respectively; $\{h_{1m}(l)\}_{l=0}^L$ and $\{h_{2m}(l)\}_{l=0}^L$ describe the channel impulse responses from transmit antennas 1 and 2 to receive antenna m , respectively, with L denoting the maximum channel order. The multipath channels include the physical channels and the transmit/receive filters and are modeled as finite impulse response (FIR) filters [22]. It is assumed that the transmitting antennas at the transmit side are placed far apart; similarly, the receiving antennas at the receive side are sufficiently far apart as well. This ensures that the transmitted symbols from the antennas undergo effectively independent fading, i.e., $h_{k_1 m_1}$ and $h_{k_2 m_2}$ are independent for $k_1 \neq k_2$ and/or $m_1 \neq m_2$. Another assumption is that the channels are invariant within a data frame though they may vary from frame to frame independently. Finally, let $v_m(n)$ describe the independent identically distributed (i.i.d.) AWGN of symbol n at the receive antenna m with zero mean and variance σ_v^2 , which is assumed to be uncorrelated with the transmitted signal.

Therefore, the complex baseband received signal at time n from the receive antenna m can be expressed as

$$r_m(n) = \sum_{k=1}^2 \sum_{l=0}^L c_k(n-l)h_{km}(l) + v_m(n), \quad m=1, \dots, M \quad (1)$$

which is the superposition of the transmitted symbols from two transmitters, each convoluted with the channel experienced by the transmitter and receiver pair. We collect P samples being sampled at the symbol rate from the output of each receive antenna: $\mathbf{r}_m(n) = [r_m(n), r_m(n+1), \dots, r_m(n+P-1)]^T \in \mathcal{C}^{P \times 1}$. And then let $\mathbf{r}(n) = [\mathbf{r}_1^T(n), \dots, \mathbf{r}_M^T(n)]^T \in \mathcal{C}^{MP \times 1}$ consist of the samples collected from the M receive antennas. Let $\mathbf{c}_1(n) = [c_1(n-L), c_1(n-L+1), \dots, c_1(n+P-1)]^T \in$

$\mathcal{C}^{q \times 1}$ describe the symbols sent from transmit antenna 1. Let $\mathbf{c}_2(n)$ be similarly formed from $c_2(n-L)$ to $c_2(n+P-1)$. Here, we assume that the total number of the transmitted symbols $q = L + P$ is even for an easy processing in an STTD system. Then, the input and output relation of the system can be expressed in a vector/matrix form

$$\mathbf{r}(n) = \mathcal{H}\mathbf{c}(n) + \mathbf{v}(n) \quad (2)$$

where

$$\mathcal{H} = \begin{bmatrix} \mathbf{H}_{11} & \mathbf{H}_{21} \\ \vdots & \vdots \\ \mathbf{H}_{1M} & \mathbf{H}_{2M} \end{bmatrix} \in \mathcal{C}^{MP \times 2q},$$

$$\mathbf{c}(n) = \begin{bmatrix} \mathbf{c}_1(n) \\ \mathbf{c}_2(n) \end{bmatrix} \in \mathcal{C}^{2q \times 1}. \quad (3)$$

The $P \times q$ Toeplitz channel matrix \mathbf{H}_{km} , $k = 1, 2$, $m = 1, \dots, M$ is given by (4), shown at the bottom of the page, and $\mathbf{v}(n)$ is the $MP \times 1$ channel noise vector, similarly constructed to $\mathbf{r}(n)$.

The problem of interest is, thus, to suppress ISI existing in the system and to coherently demodulate $\mathbf{c}_1(n)$ and $\mathbf{c}_2(n)$.

B. STBC Encoder

The STBC encoder in Fig. 1 exploits the STBC scheme introduced in [15]. The symbols $\{s(n)\} \in \mathcal{S}$ are divided into groups of two symbols each: $s(2n)$ and $s(2n+1)$, $n = 0, 1, 2, \dots$. These two adjacent symbols are input into the STBC encoder. The outputs of the ST block encoder may be consequently expressed in the following matrix form:

$$\mathbf{F} \triangleq \begin{bmatrix} c_1(2n) & c_1(2n+1) \\ c_2(2n) & c_2(2n+1) \end{bmatrix}$$

$$\mathbf{H}_{km} = \begin{bmatrix} h_{km}(L) & h_{km}(L-1) & \dots & h_{km}(0) & 0 & \dots & 0 \\ 0 & h_{km}(L) & \dots & h_{km}(1) & h_{km}(0) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & h_{km}(L) & h_{km}(L-1) & \dots & h_{km}(0) \end{bmatrix} \in \mathcal{C}^{P \times q} \quad (4)$$

where

$$\begin{aligned} c_1(2n) &= s(2n), & c_1(2n+1) &= -s^*(2n+1) \\ c_2(2n) &= s(2n+1), & c_2(2n+1) &= s^*(2n). \end{aligned} \quad (5)$$

The columns of \mathbf{F} are then transmitted over two successive symbol intervals with the elements of each column sent from two transmit antennas simultaneously.

C. Modified System Model

Therefore, the use of STBC with two transmit antennas implies the following structure in the transmitted vector $\mathbf{c}(n)$:

$$\begin{aligned} \mathbf{c}(n) &= \begin{bmatrix} \mathbf{c}_1(n) \\ \dots \\ \mathbf{c}_2(n) \end{bmatrix} = \begin{bmatrix} s(n-L) \\ -s^*(n-L+1) \\ \vdots \\ s(n+P-2) \\ -s^*(n+P-1) \\ \dots \\ s(n-L+1) \\ s^*(n-L) \\ \vdots \\ s(n+P-1) \\ s^*(n+P-2) \end{bmatrix} \\ &= \begin{bmatrix} s_r(n-L) + js_i(n-L) \\ -s_r(n-L+1) + js_i(n-L+1) \\ \vdots \\ s_r(n+P-2) + js_i(n+P-2) \\ -s_r(n+P-1) + js_i(n+P-1) \\ \dots \\ s_r(n-L+1) + js_i(n-L+1) \\ s_r(n-L) - js_i(n-L) \\ \vdots \\ s_r(n+P-1) + js_i(n+P-1) \\ s_r(n+P-2) - js_i(n+P-2) \end{bmatrix} \end{aligned} \quad (6)$$

with j denoting the imaginary unit.

Considering the spatial and temporal characteristic of STBC and using a transition matrix \mathcal{T} , defined as

$$\mathcal{T} = \begin{bmatrix} \mathbf{T}_1 \\ \mathbf{T}_2 \end{bmatrix} \in \mathcal{C}^{2q \times 2q} \quad (7)$$

the relation in (6) can be reformulated to

$$\mathbf{c}(n) = \mathcal{T} \mathbf{s}(n) \quad (8)$$

where the equivalent input vector $\mathbf{s}(n)$ is defined as $\mathbf{s}(n) = [s_r(n-L), s_i(n-L), \dots, s_r(n+P-1), s_i(n+P-1)]^T \in \mathcal{C}^{2q \times 1}$. Here, $s_r(n) = \Re\{s(n)\}$ represents the real part of symbol $s(n)$ and $s_i(n) = \Im\{s(n)\}$ represents the imaginary part. In \mathcal{T} , $\mathbf{T}_1 = \mathbf{I}_{q/2} \otimes \mathbf{A} \in \mathcal{C}^{q \times 2q}$ with $\mathbf{A} = \begin{bmatrix} 1 & j & 0 & 0 \\ 0 & 0 & -1 & j \end{bmatrix}$ and $\mathbf{T}_2 = \mathbf{I}_{q/2} \otimes \mathbf{B} \in \mathcal{C}^{q \times 2q}$ with $\mathbf{B} = \begin{bmatrix} 0 & 0 & 1 & j \\ 1 & -j & 0 & 0 \end{bmatrix}$. It is obvious that the transition matrix \mathcal{T} fully indicates the structure of STBC.

It follows from (2) and (8) that the relation between the input and the output is

$$\mathbf{r}(n) = \mathcal{H}\mathcal{T} \mathbf{s}(n) + \mathbf{v}(n) \quad (9)$$

in which the equivalent channel matrix $\mathcal{H}\mathcal{T}$ still keeps a block Toeplitz structure.

It is clear from (5) that STBC imposes a nonwhite spatial-temporal autocorrelation profile on the input sequences of $\mathbf{c}_1(n)$ and $\mathbf{c}_2(n)$. By introducing the transition matrix \mathcal{T} , however, the resulting correlation matrix $\mathbf{R}_{ss} = E[\mathbf{s}(n)\mathbf{s}^H(n)]$ is diagonal, since the equivalent transmitted symbols $\mathbf{s}(n)$ are uncorrelated [see (9)]. From now on, $\mathbf{R}_{ss} = \mathbf{I}_{2q}$ is assumed.

Therefore, with the discussion in Section II, the task is to estimate the unknown channel coefficients $\{h_{1m}\}_{l=0}^L$ and $\{h_{2m}\}_{l=0}^L$ and then to recover the transmitted information symbols $\{s(n)\}$ from the observations $\{r_m(n)\}$ corrupted by ISI and noise.

III. PROPOSED SCHEMES

Since the received signal at the receiver is the contaminated output over a frequency-selective fading channel driven by the desired signal, equalizers are necessary to compensate for channel distortion in order to decrease the error rate in signal detection. Various equalizers such as sequence and symbol-by-symbol can be employed to decode the desired signal while suppressing the interference. Although a sequence algorithm such as a maximum likelihood detector can be straightforwardly formulated, it will incur an exponential complexity with respect to the frame length [23]; hence, it has restricted availability. Accordingly, in this section, symbol-by-symbol linear equalizers (LE) and decision feedback (DFE) equalizers are considered using different criteria of zero-forcing (ZF) and minimum mean-squared error (MMSE).

Now we describe how to suppress ISI and detect desired signals. We restrict our attention first to the case in which the transmitter has no knowledge of the channel information, but the receiver has perfect knowledge of the channel. Specifically, we study a detector in which we assume the data within a subframe are correlated, but uncorrelated between subframes.

We want to detect a subframe of $P+L$ symbols. The data received within a frame are described by $\{r_m(n_f)\}_{n_f=0}^{N_f-1}$ with N_f denoting a frame length, from which non-overlapping received data vectors of length P are formed. Correspondingly, the transmitted data vectors with length $2(P+L) = 2q$ are constructed as well. That is, $\mathbf{r}_m^P(n) = [r_m(nP), r_m(nP+1), \dots, r_m(nP+P-1)]^T$, and $\mathbf{s}^P(n) = [s_r(nP-L), s_i(nP-L), \dots, s_r(nP+P-1), s_i(nP+P-1)]^T$, $n = 0, \dots, N_n-1$, where $N_n = \lceil N_f/P \rceil$. However, for clarity, from now on we still use the previous notation assuming $n \triangleq nP$, $\mathbf{r}_m(n) \triangleq \mathbf{r}_m^P(n)$, $\mathbf{r}(n) \triangleq \mathbf{r}^P(n)$, and $\mathbf{s}(n) \triangleq \mathbf{s}^P(n)$. The choice of the subframe length P is made by a tradeoff between performance and complexity: the larger the received data length P , the better the performance, whereas the more complex the receiver [24].

A. ZF Linear Equalization

First we study how to detect the symbols using a ZF equalizer. The receiver/equalizer $\mathbf{M}_{ZF} \in \mathcal{C}^{2q \times MP}$ is well known and can be obtained by minimizing the quadratic form [22], [23]

$$\mathbf{M}_{ZF} = \arg \min_{\mathbf{M}} (\mathbf{r}(n) - \mathcal{H}\mathcal{T} \hat{\mathbf{s}}_{ZF}(n))^H (\mathbf{r}(n) - \mathcal{H}\mathcal{T} \hat{\mathbf{s}}_{ZF}(n)) \quad (10)$$

which leads to $\hat{\mathbf{s}}_{\text{ZF}}(n) = \mathbf{M}_{\text{ZF}}\mathbf{r}(n) \in \mathcal{C}^{2q \times 1}$, the *soft* estimate of $\mathbf{s}(n)$. The solution to (10) is given by [22], [23]

$$\mathbf{M}_{\text{ZF}} = (\mathcal{T}^H \mathcal{H}^H \mathbf{R}_{vv}^{-1} \mathcal{H} \mathcal{T})^{-1} (\mathcal{H} \mathcal{T})^H \mathbf{R}_{vv}^{-1} \quad (11)$$

in which $\mathbf{R}_{vv} = \sigma_v^2 \mathbf{I}_{MP}$. It is noted from (11) that the ZF detector satisfies the constraint

$$\mathbf{M}_{\text{ZF}} \mathcal{H} \mathcal{T} = \mathbf{I}_{2q}. \quad (12)$$

As a result, the following is apparent:

$$\hat{\mathbf{s}}_{\text{ZF}}(n) = \mathbf{s}(n) + (\mathcal{T}^H \mathcal{H}^H \mathbf{R}_{vv}^{-1} \mathcal{H} \mathcal{T})^{-1} (\mathcal{H} \mathcal{T})^H \mathbf{R}_{vv}^{-1} \mathbf{v}(n). \quad (13)$$

Hence, from (13) ZF equalization yields symbol estimation by complete removing the ISI at the expense of increasing the additive noise level [22].

Here we suppose that CSI matrix \mathcal{H} as well as the signal and noise covariance matrices \mathbf{R}_{ss} and \mathbf{R}_{vv} are given (known or estimated beforehand using techniques such as training). Estimation of CSI is investigated in Section V of this paper.

B. MMSE Linear Equalization

Because a ZF equalizer may cause an increase in the noise level, a different linear equalizer is proposed with minimum mean-squared error criterion to trade off ISI for noise suppression.

In MMSE linear equalization, the receiver/equalizer $\mathbf{M}_{\text{MMSE}} \in \mathcal{C}^{MP \times 2q}$ is obtained with the cost function

$$\mathbf{M}_{\text{MMSE}} = \arg \min_{\mathbf{M}} E[\|\mathbf{s}(n) - \hat{\mathbf{s}}_{\text{MMSE}}(n)\|^2] \quad (14)$$

which minimizes the mean-squared error (MSE) between $\mathbf{s}(n)$ and its estimate $\hat{\mathbf{s}}_{\text{MMSE}}(n)$. It can be seen from the above equation that this equalizer is selected to maximize the associated signal-to-interference-and-noise ratio (SINR). The solution to (14) is known with the form [22], [23]

$$\mathbf{M}_{\text{MMSE}} = \mathbf{R}_{vv}^{-1} \mathcal{H} \mathcal{T} (\mathcal{T}^H \mathcal{H}^H \mathbf{R}_{vv}^{-1} \mathcal{H} \mathcal{T} + \mathbf{R}_{ss}^{-1})^{-1}. \quad (15)$$

Therefore, the *soft* estimate of $\mathbf{s}(n)$ is given by $\hat{\mathbf{s}}_{\text{MMSE}}(n) = \mathbf{M}_{\text{MMSE}}^H \mathbf{r}(n)$, which contains desired symbols, residual ISI, and noise.

Comparing (11) and (15), it can be demonstrated that the ZF detector and the MMSE detector are related with

$$\hat{\mathbf{s}}_{\text{MMSE}}(n) = \left(\mathbf{I}_{2q} + (\mathbf{R}_{ss} \mathcal{T}^H \mathcal{H}^H \mathbf{R}_{vv}^{-1} \mathcal{H} \mathcal{T})^{-1} \right)^{-1} \hat{\mathbf{s}}_{\text{ZF}}(n) \quad (16)$$

which states that the MMSE-LE can be interpreted as an extension of the ZF-LE by a Wiener estimator [25]. It is easily seen that the MMSE detection lessens the performance degradation of the ZF-LE caused by not taking into account the noise correlations existing in the decision variables.

When the noise level is very low in comparison with the signal, the ZF filter is approximately equivalent to the MMSE filter. That is, in the limit as $\sigma_v^2 \rightarrow 0$, the two equalizers yield the same solution [22], which could be easily demonstrated from (16).

C. ZF Decision Feedback Equalization

Linear equalizers, both ZF and MMSE, do not perform well when the underlying channels have deep spectral nulls in the

passband. Therefore, nonlinear equalizers have been developed as vital options to deal with such channels.

A DFE equalizer is a nonlinear equalizer that employs previously detected signals to eliminate the ISI caused by the previously detected signals on the currently detected signal. The use of the previously detected signals makes the equalizer output a nonlinear function of the data. Similarly, two different criteria, ZF and MMSE, can be used. First, the ZF case is described.

It is easily shown that the correlation of the noise term in (13) is given by its covariance matrix $(\mathcal{T}^H \mathcal{H}^H \mathbf{R}_{vv}^{-1} \mathcal{H} \mathcal{T})^{-1}$. Thus, the Cholesky decomposition [26] of the inverted covariance matrix has the form

$$\mathcal{T}^H \mathcal{H}^H \mathbf{R}_{vv}^{-1} \mathcal{H} \mathcal{T} = (\mathbf{D}\mathbf{U})^H (\mathbf{D}\mathbf{U}) \quad (17)$$

in which $\mathbf{U} \in \mathcal{C}^{2q \times 2q}$ is an upper triangular matrix with the diagonal having the value one and $\mathbf{D} \in \mathcal{C}^{2q \times 2q}$ is a diagonal matrix with real entries.

To describe the ZF-DFE, starting from (13) and (17), the modified received sequence is introduced by premultiplying the triangular matrix \mathbf{U}

$$\begin{aligned} \mathbf{r}'_{\text{ZF}}(n) &= \mathbf{U} \hat{\mathbf{s}}_{\text{ZF}}(n) \\ &= \mathbf{U} \mathbf{s}(n) + \mathbf{D}^{-1} \mathbf{D}^{-1} (\mathbf{U}^H)^{-1} \mathcal{T}^H \mathcal{H}^H \mathbf{R}_{vv}^{-1} \mathbf{v}(n) \\ &= \mathbf{s}(n) + (\mathbf{U} - \mathbf{I}) \mathbf{s}(n) \\ &\quad + \mathbf{D}^{-1} \mathbf{D}^{-1} (\mathbf{U}^H)^{-1} \mathcal{T}^H \mathcal{H}^H \mathbf{R}_{vv}^{-1} \mathbf{v}(n). \end{aligned} \quad (18)$$

Taking into account the fact that the matrix \mathbf{U} is upper triangular, the decision on a signal $s(n)$ can be made from (18) recursively in reverse order of the components of signal vector $\mathbf{s}(n)$ by using past decisions (high-indexed components) on previous signals (low-indexed components). Note that the term $(\mathbf{U} - \mathbf{I})\mathbf{s}(n)$ in (18) consists of the signals obtained already with reverse order decision. Decisions are made recursively as

$$\begin{aligned} \hat{\mathbf{s}}_{\text{ZF-DFE}}(n) &= \text{Qt}[r'_{\text{ZF}}(n)], \quad n \triangleq nP + P - 1 \\ \hat{\mathbf{s}}_{\text{ZF-DFE}}(n-k) &= \text{Qt}\left[r'_{\text{ZF}}(n-k) - \sum_{i=1}^k [\mathbf{U} - \mathbf{I}]_{n-k, n-k+i} \hat{\mathbf{s}}_{\text{ZF-DFE}}(n-k+i)\right] \end{aligned} \quad (19)$$

in which $k = 1, \dots, 2q - 1$ and $\text{Qt}[\cdot]$ indicates a quantization operation performed with a threshold detector. Subtracting the decisions $\hat{\mathbf{s}}_{\text{ZF-DFE}}$ from (18), the new statistic

$$\begin{aligned} \hat{\mathbf{t}}_{\text{ZF}}(n) &= \mathbf{r}'_{\text{ZF}}(n) - (\mathbf{U} - \mathbf{I}) \hat{\mathbf{s}}_{\text{ZF-DFE}}(n) \\ &= \mathbf{s}(n) + (\mathbf{U} - \mathbf{I})(\mathbf{s}(n) - \hat{\mathbf{s}}_{\text{ZF-DFE}}(n)) \\ &\quad + \mathbf{D}^{-1} \mathbf{D}^{-1} (\mathbf{U}^H)^{-1} \mathcal{T}^H \mathcal{H}^H \mathbf{R}_{vv}^{-1} \mathbf{v}(n) \end{aligned} \quad (20)$$

is obtained, which turns into

$$\hat{\mathbf{t}}_{\text{ZF}}(n) = \mathbf{s}(n) + \mathbf{D}^{-1} \mathbf{D}^{-1} (\mathbf{U}^H)^{-1} \mathcal{T}^H \mathcal{H}^H \mathbf{R}_{vv}^{-1} \mathbf{v}(n) \quad (21)$$

if all past decisions are correct. Therefore, the *soft* estimate of the desired component is obtained. The operation

¹For notation simplicity, we use the following definition here: $\hat{\mathbf{s}}_{\text{ZF-DFE}}(n) = [\hat{s}_{\text{ZF-DFE}}(nP + P - 2q), \dots, \hat{s}_{\text{ZF-DFE}}(nP + P - 1)]^T$. However, it should be kept in mind that $\hat{\mathbf{s}}_{\text{ZF-DFE}}(n) = [\hat{s}_{r, \text{ZF-DFE}}(nP - L), \hat{s}_{i, \text{ZF-DFE}}(nP - L), \dots, \hat{s}_{r, \text{ZF-DFE}}(nP + P - 1), \hat{s}_{i, \text{ZF-DFE}}(nP + P - 1)]^T$.

$(\mathbf{U} - \mathbf{I})\hat{\mathbf{s}}_{\text{ZF-DFE}}(n)$ in (20) constitutes the feedback operator. In [27], it is shown that the ZF-DFE is equivalent to a noise-cancelling detector derived from $\mathbf{U}^{-1}\mathbf{r}'_{\text{ZF}}(n)$.

D. MMSE Decision Feedback Equalization

An MMSE decision feedback equalizer (MMSE-DFE) is now similarly formed. The Cholesky decomposition of the matrix $\mathcal{T}^H \mathcal{H}^H \mathbf{R}_{vv}^{-1} \mathcal{H} \mathcal{T} + \mathbf{R}_{ss}^{-1}$ is

$$\mathcal{T}^H \mathcal{H}^H \mathbf{R}_{vv}^{-1} \mathcal{H} \mathcal{T} + \mathbf{R}_{ss}^{-1} = (\mathbf{D}' \mathbf{U}')^H \mathbf{D}' \mathbf{U}' \quad (22)$$

where $\mathbf{U}' \in \mathcal{C}^{2q \times 2q}$ and $\mathbf{D}' \in \mathcal{C}^{2q \times 2q}$ are similarly defined an upper triangular matrix and a diagonal matrix, respectively. The modified received sequence $\mathbf{r}''_{\text{MMSE}}$ can be introduced using the same approach with matrix \mathbf{U}' premultiplied

$$\begin{aligned} \mathbf{r}''_{\text{MMSE}}(n) &= \mathbf{U}' \hat{\mathbf{s}}_{\text{MMSE}}(n) \\ &= \mathbf{U}' (\mathcal{T}^H \mathcal{H}^H \mathbf{R}_{vv}^{-1} \mathcal{H} \mathcal{T} + \mathbf{R}_{ss}^{-1})^{-1} \\ &\quad \times \mathcal{T}^H \mathcal{H}^H \mathbf{R}_{vv}^{-1} \mathbf{r}(n). \end{aligned} \quad (23)$$

By exploiting Cholesky decomposition in (22) and its transformation $(\mathcal{H} \mathcal{T})^H \mathcal{H} \mathcal{T} = \sigma_v^2 ((\mathbf{D}' \mathbf{U}')^H (\mathbf{D}' \mathbf{U}') - \mathbf{I}_{2q})$ with $\mathbf{R}_{vv} = \sigma_v^2 \mathbf{I}_{MP}$ and $\mathbf{R}_{ss} = \mathbf{I}_{2q}$, the above equation can be expressed as

$$\begin{aligned} \mathbf{r}''_{\text{MMSE}}(n) &= \mathbf{U}' \mathbf{s}(n) - \mathbf{D}'^{-1} \mathbf{D}'^{-1} \mathbf{U}'^{-H} \mathbf{s}(n) \\ &\quad + \sigma_v^{-2} \mathbf{D}'^{-1} \mathbf{D}'^{-1} \mathbf{U}'^{-H} (\mathcal{H} \mathcal{T})^H \mathbf{v}(n) \\ &= \mathbf{s}(n) + (\mathbf{U}' - \mathbf{I}) \mathbf{s}(n) - \mathbf{D}'^{-1} \mathbf{D}'^{-1} \mathbf{U}'^{-H} \mathbf{s}(n) \\ &\quad + \sigma_v^{-2} \mathbf{D}'^{-1} \mathbf{D}'^{-1} \mathbf{U}'^{-H} (\mathcal{H} \mathcal{T})^H \mathbf{v}(n). \end{aligned} \quad (24)$$

The new statistic $\hat{\mathbf{t}}'_{\text{MMSE}}$ is then introduced similarly to (20). The process of deciding and subtracting past decisions is the same as described in Section III-C. If there is no error propagation, we obtain the *soft* estimate

$$\begin{aligned} \hat{\mathbf{t}}'_{\text{MMSE}}(n) &= (\mathbf{I} - \mathbf{D}'^{-1} \mathbf{D}'^{-1} \mathbf{U}'^{-H}) \mathbf{s}(n) \\ &\quad + \sigma_v^{-2} \mathbf{D}'^{-1} \mathbf{D}'^{-1} \mathbf{U}'^{-H} \\ &\quad \times (\mathcal{H} \mathcal{T})^H \mathbf{v}(n). \end{aligned} \quad (25)$$

From the previous discussion, it is noted that in order to fully recover the information signals, no separate ST decoding needs to be performed due to the introduction of the coding matrix \mathcal{T} , in which the spatial and temporal structure of STBC are inherent. It thus simplifies the total receiving processing. The overlapping signals of two successive subframes, because of block detection, are obtained by an arithmetic average of the corresponding points in these two subframes. The *hard* estimate of each desired symbol is finally obtained by comparing the *soft* estimate formed from $\hat{\mathbf{s}}_{\text{ZF}}(n)$, $\hat{\mathbf{s}}_{\text{MMSE}}(n)$, $\hat{\mathbf{t}}_{\text{ZF}}(n)$, $\hat{\mathbf{t}}'_{\text{MMSE}}(n)$, respectively, with every constellation point

$$\hat{\hat{s}}(n) = \arg \min_{s \in \mathcal{S}} |\hat{s}(n) - s| \quad (26)$$

where $|\cdot|$ denotes the Euclidean distance. For binary phase shift keying (BPSK), this reduces to $\hat{\hat{s}}(n) = \text{sign}(\text{Re}(\hat{s}(n)))$. Thus, the estimates of the transmitted symbols can be acquired.

From the previous discussion, it is obvious that the realization of equalizers is computationally demanding and that the calculation of matrix multiplication and inversion is critically important when implementing an equalizer. Direct evaluation of the matrix multiplication and inversion is computationally inefficient and, thus, not recommended, particularly when the subframe size P

is large. Simplified computation methods are consequently developed in Appendices I and II, which are able to reduce the computational load drastically.

E. Identifiability Condition

The formulation of the STTD system in Section II and the detection of the information symbols introduced in Section III imply the identifiability of a system.

It has been seen in Section III-A that $\mathbf{M}_{\text{ZF}} \mathcal{H} \mathcal{T} = \mathbf{I}_{2q}$. Given \mathcal{H} and \mathcal{T} , \mathbf{M}_{ZF} exists if and only if it is possible to invert $\mathcal{H} \mathcal{T}$; i.e., it has to be nonsingular. Invertibility of the $MP \times 2q$ matrix $\mathcal{H} \mathcal{T}$ requires it to be tall and of full column rank, which is, however, not necessarily the case in all conceivable situations. Since \mathcal{T} is square and of full rank, $\text{rank}(\mathcal{H} \mathcal{T}) = \text{rank}(\mathcal{H})$, which implies that the multichannel filtering matrix \mathcal{H} plays a central role in the equalization process. In particular, the equalizer exists on condition that the matrix \mathcal{H} is of full column rank, which necessitates a careful choice of the factor P to make it ‘‘tall.’’

Now we adopt the following assumption that the size P of the received signal vector $\mathbf{r}_m(n)$ for each receive antenna is greater than L , the maximum order of the channels. Therefore, for the existence of an FIR ZF-equalizing filter, we must have

- 1) $MP \geq 2q = 2(L + P)$;
- 2) $\text{rank}(\mathcal{H} \mathcal{T}) = 2q$.

In order to satisfy 1), we need to meet the following requirement: for a given $L \geq 1$, select M to satisfy $M \geq 3$ in a two-transmit antennas system. Then, given the number of receive antenna M and the channel order L , 1) is met easily by selecting P appropriately. The full column-rank requirement 2) is readily met when the channels have no common zeros (coprime). Detailed investigation on the characteristic of the channel matrix and the system identifiability is referred to in [28] and [29] and references herein. Therefore, when the above conditions of receive antenna number M and rank of channel matrix \mathcal{H} are satisfied, the ZF equalizer exists and has a unique solution. This also applies to the case of an MMSE equalizer, though, theoretically, it always exists.

It is derived from the discussion that equalization can be used to restore orthogonality of the transmitted symbols and to enhance the system performance, though equalization with a single receive antenna is not straightforward in an STTD system.

IV. PERFORMANCE ANALYSIS AND BER UPPER BOUND

A. BER Performance

In order to obtain an estimate of the system performance in terms of BER and to compare the different receivers, we resort to the SINR.

Let $\gamma(k)$ denote SINR relating to a data bit at the output of the equalizer. Then it can be calculated as

$$\gamma(k) = \frac{E[|\hat{s}(k)|^2]}{E[|\text{ISI}_{\text{res}} + \text{noise}|^2]} \quad (27)$$

which takes into account both noise and residual interference. Let $J(n)$ denote the MSE of the detection, defined by

$$\begin{aligned} J(n) &= E[|\mathbf{e}(n)|^2] = E[|\mathbf{s}(n) - \hat{\mathbf{s}}(n)|^2] \\ &= \text{tr} \left(E[(\mathbf{s}(n) - \hat{\mathbf{s}}(n))(\mathbf{s}(n) - \hat{\mathbf{s}}(n))^H] \right). \end{aligned} \quad (28)$$

Correspondingly, $J_s(k)$ is defined as $J_s(k) = E[\|s(k) - \hat{s}(k)\|^2]$ describing the MSE of the k th data bit of the detected vector. It can be seen from the following analysis that SINR and MSE are closely related.

1) *SINR of ZF-LE*: For a ZF-LE equalizer, it is easy to obtain from (13) that the MSE has the expression

$$J_{ZF}(n) = \text{tr} \left(\sigma_v^2 ((\mathcal{H}T)^H \mathcal{H}T)^{-1} \right) \quad (29)$$

and the SINR $\gamma(k)$ per data bit at the input to the nonlinear decision device is

$$\gamma_{ZF}(k) = \frac{1}{[\sigma_v^2 ((\mathcal{H}T)^H \mathcal{H}T)^{-1}]_{kk}} = \frac{1}{J_{s,ZF}(k)} \quad (30)$$

where $J_{s,ZF}(k)$ is the noise power of the estimate of bit k .

2) *SINR of MMSE-LE*: For an MMSE-LE equalizer, we start from (28) and get

$$J_{MMSE}(n) = \text{tr} (E[\mathbf{s}\mathbf{s}^H - \mathbf{s}\hat{\mathbf{s}}^H - \hat{\mathbf{s}}\mathbf{s}^H + \hat{\mathbf{s}}\hat{\mathbf{s}}^H]) \quad (31)$$

which leads to

$$\begin{aligned} J_{MMSE}(n) &= \text{tr} \left(\mathbf{I}_{2q} - (\mathcal{H}T)^H (\mathcal{H}T(\mathcal{H}T)^H + \sigma_v^2 \mathbf{I}_{MP})^{-1} \right. \\ &\quad \left. \times \mathcal{H}T \right), \\ &= \text{tr} \left((\mathbf{I}_{2q} + \sigma_v^{-2} (\mathcal{H}T)^H \mathcal{H}T)^{-1} \right). \end{aligned} \quad (32)$$

The proof of equation $\mathbf{I}_{2q} - (\mathcal{H}T)^H (\mathcal{H}T(\mathcal{H}T)^H + \sigma_v^2 \mathbf{I}_{MP})^{-1} \mathcal{H}T = (\mathbf{I}_{2q} + \sigma_v^{-2} (\mathcal{H}T)^H \mathcal{H}T)^{-1}$ is acquired immediately by exploiting the matrix inversion lemma [25].²

Because an MMSE equalizer minimizes the mean-squared value of the estimation error, it is easily demonstrated from the principle of orthogonality [25] that

$$E[\|\hat{\mathbf{s}}(n)\|^2] = \text{tr} \left(\mathbf{I}_{2q} - (\mathbf{I}_{2q} + \sigma_v^{-2} (\mathcal{H}T)^H \mathcal{H}T)^{-1} \right). \quad (33)$$

Consequently, it can be clearly shown that the SINR $\gamma(k)$ per data bit of the MMSE-LE equalizer is

$$\begin{aligned} \gamma_{MMSE}(k) &= \frac{1 - [\mathbf{I}_{2q} + \sigma_v^{-2} (\mathcal{H}T)^H \mathcal{H}T]_{kk}^{-1}}{[\mathbf{I}_{2q} + \sigma_v^{-2} (\mathcal{H}T)^H \mathcal{H}T]_{kk}^{-1}} \\ &= \frac{1 - J_{s,MMSE}(k)}{J_{s,MMSE}(k)} \end{aligned} \quad (34)$$

which is greater than $\gamma_{ZF}(k)$ in general [30]. It can be demonstrated that $J_{s,ZF}(k)$ and $J_{s,MMSE}(k)$ will be the normalized MSE, which is normalized to the variance of the input data bit, if \mathbf{R}_{ss} is not an identity matrix.

3) *SINR of ZF-DFE*: It is quite involved if not possible to evaluate the SINRs of the DFE receivers with error propagation. Therefore, we focus on the SINRs under the assumption that all past decisions are correct, i.e., have no error propagation effect.

Accordingly, the MSE and SINR per bit at the output of the ZF-DFE are

$$J_{ZF-DFE}(n) = \text{tr} (\mathbf{D}^{-2}) \quad (35)$$

$$\gamma_{ZF-DFE}(k) = [\mathbf{D}^2]_{kk} = J_{s,ZF-DFE}(k) \quad (36)$$

which are straightforwardly obtained from (21) and (17).

4) *SINR of MMSE-DFE*: Similar to the obtaining of SINR of the MMSE-LE receiver, we first evaluate the MSE of MMSE-DFE receivers without error propagation. From (25)

and (22), along with its transformation, the mean-squared error can be expressed as

$$\begin{aligned} J_{MMSE-DFE}(n) &= \text{tr} \left(\mathbf{D}'^{-1} \mathbf{D}'^{-1} \mathbf{U}'^{-H} \mathbf{U}'^{-1} \mathbf{D}'^{-1} \mathbf{D}'^{-1} \right. \\ &\quad \left. + \sigma_v^{-2} \mathbf{D}'^{-1} \mathbf{D}'^{-1} \mathbf{U}'^{-H} (\mathcal{H}T)^H \right. \\ &\quad \left. \times \mathcal{H}T \mathbf{U}'^{-1} \mathbf{D}'^{-1} \mathbf{D}'^{-1} \right) \\ &= \text{tr} \left(\mathbf{D}'^{-1} \mathbf{D}'^{-1} \mathbf{U}'^{-H} \mathbf{U}'^{-1} \mathbf{D}'^{-1} \mathbf{D}'^{-1} \right. \\ &\quad \left. + \mathbf{D}'^{-1} \mathbf{D}'^{-1} \mathbf{U}'^{-H} (\mathbf{U}'^H \mathbf{D}' \mathbf{D}' \mathbf{U}' - \mathbf{I}) \right. \\ &\quad \left. \times \mathbf{U}'^{-1} \mathbf{D}'^{-1} \mathbf{D}'^{-1} \right) \\ &= \text{tr} (\mathbf{D}'^{-2}). \end{aligned} \quad (37)$$

Similarly, using the orthogonality principle and the process to obtain $\gamma_{MMSE}(k)$, the SINR per bit becomes

$$\gamma_{MMSE-DFE}(k) = [\mathbf{D}'^2]_{kk} - 1 = J_{s,MMSE-DFE}(k) - 1. \quad (38)$$

Now it has been proven that the relationships between SINR $\gamma(k)$ and the normalized MSE hold for our STTD system with multiple transmit and multiple receive antennas as in the SISO case, which are given in (30), (34), (36), and (38), indicating different receiving schemes.

5) *BER Calculation*: From the given SINR $\gamma(k)$, an estimate of the resulting BER $P_e(\gamma(k))$ can be obtained over an AWGN channel according to different modulations [22]. $P_e(\gamma(k))$ is only an estimate of BER because the AWGN channel assumption is an approximation³ and is increased by counting the effect of multipath fading. An average error rate of the estimated BER $\bar{P}_e(\gamma(k))$ for frequency-selective channels can be obtained by taking average of $P_e(\gamma(k))$ with respect to the fading channels, which leads to

$$\bar{P}_e(\gamma(k)) = \int_0^\infty P_e(\gamma(k)) \cdot p_\gamma(\gamma(k)) d(\gamma(k)). \quad (39)$$

Here $p_\gamma(\gamma(k))$ is the probability density function of SINR $\gamma(k)$ determined by the frequency-selective fading channels. Finally, the estimate value \bar{P}_e of $\bar{P}_e(\gamma(k))$ is evaluated by an arithmetic average over all estimated bits.

B. BER Upper Bound

It is apparent that (39), when evaluated, would cause a rather formidable and awkward expression whose individual terms are not easily interpretable. Furthermore, it is not hard to conceive of a case in which an analytic result is not possible and a numerical calculation must be made. However, both of these circumstances, complicated expressions or numerical simulations, have a disadvantage in that they limit insight into the merits and limitations of a system and also prevent an easy comparison of one system with another. Accordingly, we now consider a scheme that leads to an approximate result, but has the feature to circumvent the difficulties mentioned above. This method gives an upper-bound approximation [31], an efficient asymptotic measure of system performance over frequency-selective channels. For uniformity, BER upper bounds are given for both ZF and MMSE cases, although exact $P_e(\gamma(k))$ exists for ZF equalizers.

²Let $\mathbf{A} = \mathbf{B}^{-1} + \mathbf{C}\mathbf{D}^{-1}\mathbf{C}^H$ where \mathbf{A} , \mathbf{B} , and \mathbf{D} are positive-definite matrices. Then $\mathbf{A}^{-1} = \mathbf{B} - \mathbf{B}\mathbf{C}(\mathbf{D} + \mathbf{C}^H\mathbf{B}\mathbf{C})^{-1}\mathbf{C}^H\mathbf{B}$.

³It should be noted that, for a linear ZF equalizer, it is possible to express the BER exactly with a Q -function, because the channel is perfectly equalized so that ISI is completely eliminated.

An upper bound was derived in [32] for a SISO system with coherent detection. It is easy to show that the exponent of the bound given in [32] and [33] is identical to MSE divided by the variance of the input signal and, therefore, the BER $\bar{P}_e(\gamma(k))$ in our STBC system is upper bounded by

$$\bar{P}_e(\gamma(k)) \leq \exp\left(\frac{1}{E[|s(k)|^2]}\right) E\left[\exp\left(\frac{-1}{J_s(k)}\right)\right] \quad (40)$$

which is subject to the effects of different equalizers. Then, upper bound of \bar{P}_e can be evaluated with an average operation. It is readily deduced that, when the noise is getting weak, the upper bound is asymptotically tight.

V. TRAINING-AIDED CHANNEL ESTIMATION

We have observed that channel parameters are needed for diversity combining, coherent detection, and decoding in an STTD system to achieve gains from both antenna and fading diversity. In mobile communication applications, *a priori* knowledge is known to the receivers, although the actual transmitted symbol stream is unknown. In such a system, a known preamble is added to the message for training purpose. Such extra information may be exploited to enhance the accuracy of the estimates and may be used to simplify the computational complexity of the estimation algorithm.

It has been shown that the input–output relation can be expressed in a form of (2). By denoting the input matrix $\tilde{\mathbf{C}}$ and the CSI vector $\tilde{\mathbf{h}}$, a new form of input–output relation is obtained with $\mathbf{r}_t(n)$ and $\mathbf{v}_t(n)$, defined similarly to $\mathbf{r}(n)$ and $\mathbf{v}(n)$, respectively

$$\mathbf{r}_t(n) = \tilde{\mathbf{C}}(n)\tilde{\mathbf{h}} + \mathbf{v}_t(n) \quad (41)$$

where $\tilde{\mathbf{C}}$ is expressed in (42), shown at the bottom of the page, with N_t denoting the length of the received symbols during the training period. Equation (42) has a block Toeplitz structure if we consider $[\mathbf{C}_1(n), \mathbf{C}_2(n)]$ as one block. Now the input signal vector $\mathbf{c}_k(\mathbf{n})$ in (2) has been reformatted to construct the training signal matrix $\mathbf{C}_k(n) \in \mathcal{C}^{N_t \times (L+1)}$, $k = 1, 2$ shown in (43) at the bottom of the page, and the CSI matrix \mathcal{H} is instead changed to a vector form

$$\tilde{\mathbf{h}} = [\mathbf{h}_{11}^T \quad \mathbf{h}_{21}^T \quad \cdots \quad \mathbf{h}_{1M}^T \quad \mathbf{h}_{2M}^T]^T \in \mathcal{C}^{2M(L+1) \times 1} \quad (44)$$

in which $\mathbf{h}_{km} = [h_{km}(L), h_{km}(L-1), \dots, h_{km}(0)]^T$, $k = 1, 2$, $m = 1, \dots, M$. Note that (43) is composed of known training

symbols and that the total number of these training symbols $N_t + L$ is assumed to be even.

By minimizing a cost function that composes the variance of error [25]

$$\hat{\mathbf{h}} = \arg \min_{\tilde{\mathbf{h}}} \|\mathbf{r}_t(n) - \tilde{\mathbf{C}}(n)\tilde{\mathbf{h}}\|^2 \quad (45)$$

we obtain the linear least-squares (LS) estimate $\hat{\mathbf{h}}$ of the CSI vector $\tilde{\mathbf{h}}$

$$\hat{\mathbf{h}} = \left(\tilde{\mathbf{C}}^H(n)\tilde{\mathbf{C}}(n)\right)^{-1} \tilde{\mathbf{C}}^H(n)\mathbf{r}_t(n). \quad (46)$$

This LS estimator turns out to be the best (in terms of having minimum mean-squared error) among all unbiased estimators and is the most efficient in the sense that it achieves the Cramér–Rao lower bound [34]. Its implementation can be completed recursively [25].

With training symbols, channel parameters are estimated sufficiently close to the desired settings such that much of ISI can be removed. Because of the channel variation and nonlinearity within one frame, the channel parameters should be tracked adaptively.

We expect the existence of the inverse in (46) and a unique solution to the linear LS problem exists only when the input matrix $\tilde{\mathbf{C}}$ has linearly independent columns; that is, when the input matrix $\tilde{\mathbf{C}}$ is of full column rank. This implies that the matrix $\tilde{\mathbf{C}}$ has at least as many rows as columns, i.e., $MN_t \geq 2M(L+1)$, which means $N_t \geq 2(L+1)$. Therefore, the training symbols should be long enough and chosen as a tradeoff between performance and spectral efficiency. Thus, provided that the input matrix $\tilde{\mathbf{C}}$ is of full column rank, the $2M(L+1)$ matrix $\tilde{\mathbf{C}}^H(n)\tilde{\mathbf{C}}(n)$ is nonsingular and the least-squares estimate has the unique value given in (46), which gives an efficient estimation of CSI.

As a method to estimate CSI, an LS algorithm is introduced so as to effectively remove ISI and to detect the desired signals. Another novel algorithm of estimating CSI and the noise level (plus cochannel interference) can be found in [35].

VI. SIMULATION RESULTS

In this section, we present our simulation results for the proposed techniques. The performances of the data detection and CSI estimation algorithms introduced in Sections III and V are illustrated. The theoretical analysis in Section IV is verified through numerical simulations.

$$\tilde{\mathbf{C}}(n) = \begin{bmatrix} \mathbf{C}_1(n) & \mathbf{C}_2(n) & \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{C}_1(n) & \mathbf{C}_2(n) & \mathbf{0} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{C}_1(n) & \mathbf{C}_2(n) \end{bmatrix} \in \mathcal{C}^{MN_t \times 2M(L+1)} \quad (42)$$

$$\mathbf{C}_k(n) = \begin{bmatrix} c_k(n-L) & c_k(n+1-L) & \cdots & c_k(n) \\ c_k(n+1-L) & c_k(n+2-L) & \cdots & c_k(n+1) \\ \vdots & \vdots & \vdots & \vdots \\ c_k(n+N_t-L-1) & c_k(n+N_t-L) & \cdots & c_k(n+N_t-1) \end{bmatrix} \quad (43)$$

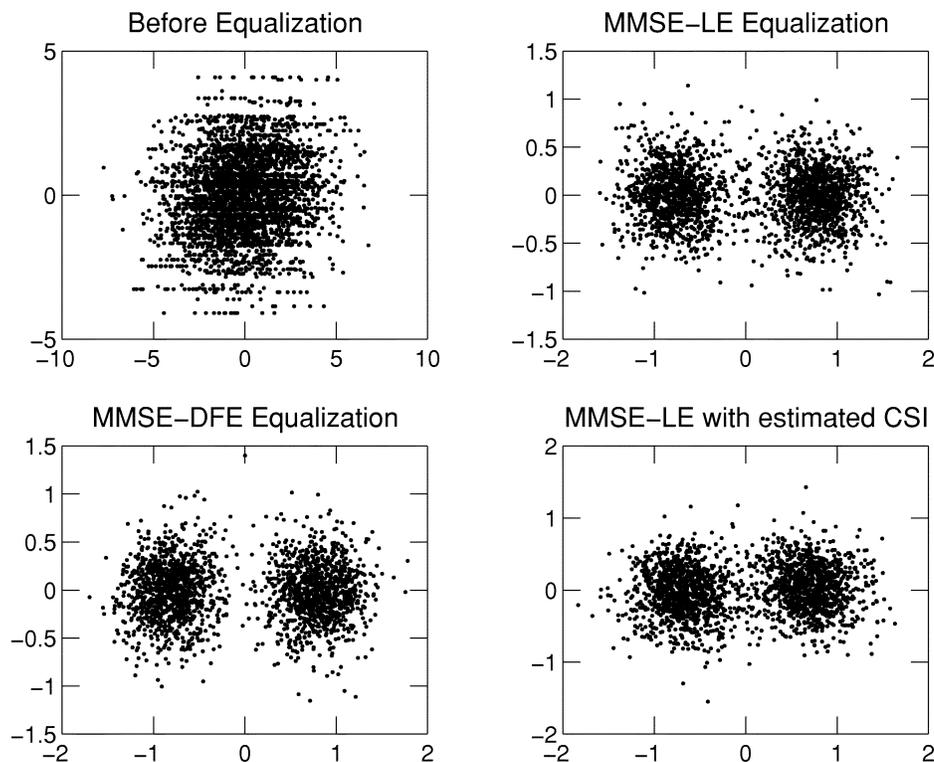


Fig. 2. Scattering diagrams ($M = 3$ Rx, SNR = 2 dB).

We consider an STTD system using a BPSK modulation scheme with $K = 2$ transmit antennas and $M = 3$ (or 5) receive antennas. Following the Rayleigh-fading channel assumption, the channel coefficients between each transmit antenna and each receive antenna $\{h_{km}\}$ are generated as complex Gaussian random variables with zero mean and equal variance that are independent for different k and/or m ; also, the channels stay unchanged within a frame and vary independently from frame to frame. Thus, the identical average signal power is received at each receive antenna from each transmit antenna. Here, the channel variance is chosen to guarantee the total received power obtained at one antenna from two transmit antennas is one. In the following examples, we set channel order $L = 1$ (i.e., a two-ray channel), the frame length $N_f = 162$, the received data length $P = 17$, and the training symbol length $N_t + L = 18$. The SNR is defined as $\text{SNR} = 10\log_{10}1/\sigma_v^2$ dB.

A. ISI Suppression and Diversity Advantage

Now we examine the performances of the proposed equalizers. First, we assume that the perfect knowledge of CSI is available at the receive side.

Fig. 2 depicts the scattering diagrams obtained with LE and DFE MMSE equalizers. All are under the condition that the received SNR = 2 dB. For comparison, a scattering diagram acquired with estimated CSI introduced in Section V is shown as well. It is evident that both LE and DFE equalizers can remove ISI effectively and that a nonlinear MMSE receiver performs better than does a linear one.

In Fig. 3, we examine the performances of the receivers discussed in Section III using BER curves. The BER is depicted as

the function of SNR. The BERs versus SNR with and without equalization in the case of $M = 3$ receive antennas are given. As a comparison, the BER without ISI (flat-fading case) is shown as well. Because of the assumed quasi-static characteristic of the fading channels, the BER results presented are averaged over large number of frames with different fading channels.

This figure demonstrates that all equalizers (ZF/MMSE with LE/DFE) are able to effectively suppress ISI and to detect the desired symbols compared with the case of no equalization applied. And at high SNR (e.g, SNR = 10 dB), MMSE equalizers result in the ZF solutions. It also demonstrates that an MMSE equalizer is superior to a ZF equalizer and that a nonlinear equalizer is superior to a linear equalizer (i.e., the MMSE-DFE performs best, followed by the ZF-DFE, the MMSE-LE, and the ZF-LE). Note that, in the cases of DFE equalizers, the effect of error propagation is included in our simulations.

Then we consider the transmit diversity advantage provided by STBC. We compare the proposed STTD system with the conventional system using receive diversity only (i.e., $K = 1$) without ST coding. The BER results of the conventional system are obtained with LE ZF and MMSE equalizers, both of which are implemented similarly to the equalizers in the STTD system. It is clearly seen that, compared with the cases with only one transmit antenna, better BER performances are achieved due to the higher diversity gain provided by STBC and multiple transmit antennas. It clearly motivates the use of ST coding in conventional systems.

Fig. 4 gives the results obtained similarly to the previous scenario with more receive antennas ($M = 5$). The BER curves validate that higher diversity gain can be achieved when more receive antennas are possible.

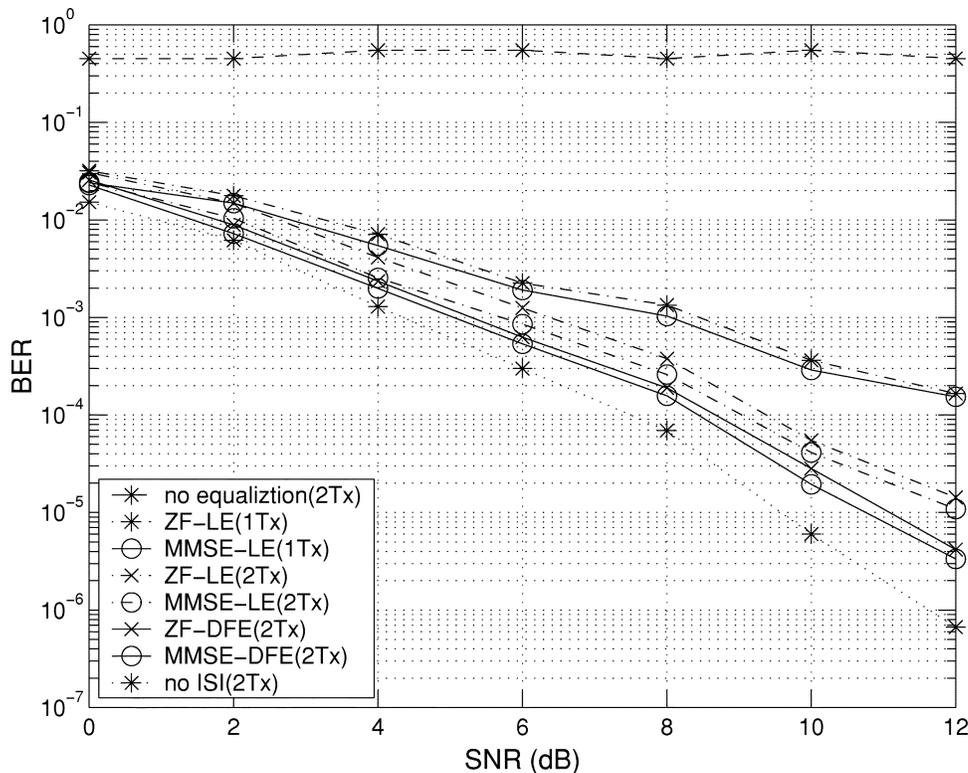


Fig. 3. BER versus SNR with $M = 3$ Rx (2Tx versus 1Tx).

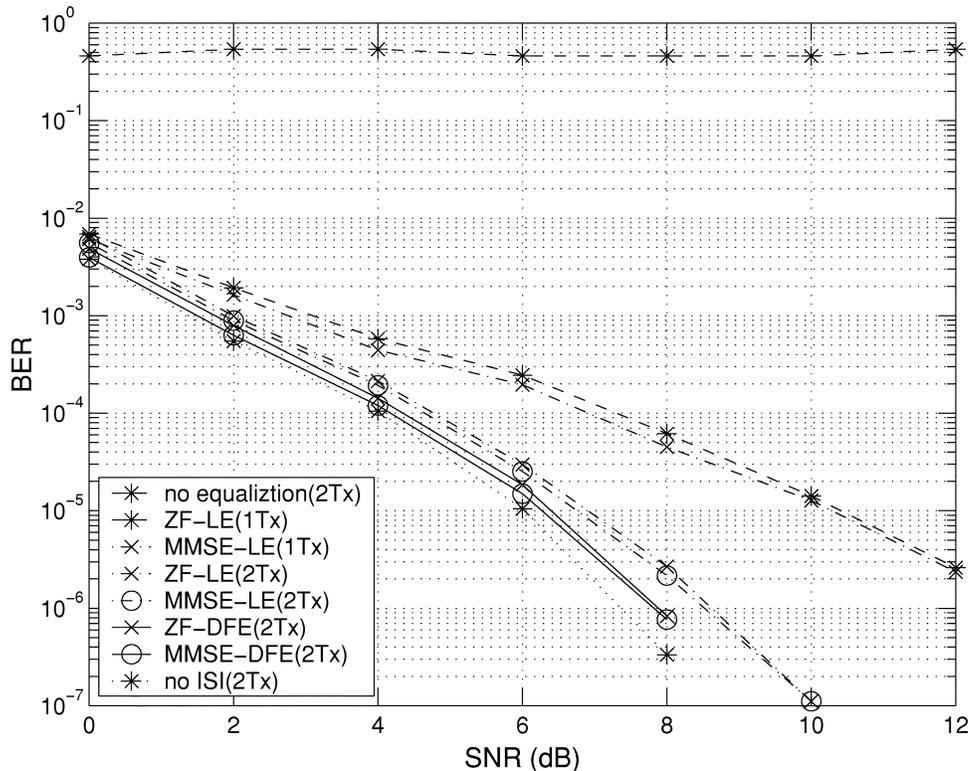


Fig. 4. BER versus SNR with $M = 5$ Rx (2Tx versus 1Tx).

B. BER Performance Analysis

While the simulated BER curves in Figs. 3 and 4 are obtained by simulating the whole process of the transmitted signal, fading channels with noise, and the receiver, the corresponding analytical BER curves in Figs. 5 and 6 are obtained with a

semi-analytical method by calculating $P_e(\gamma(k))$, which results from substituting different SINRs expressed in Section IV, i.e., (30), (34), (36), and (38) into the BER and SNR relation of a BPSK-modulated system [31]. The channel coefficients $\{h_{km}\}$ are created identically with those in the simulated cases given in

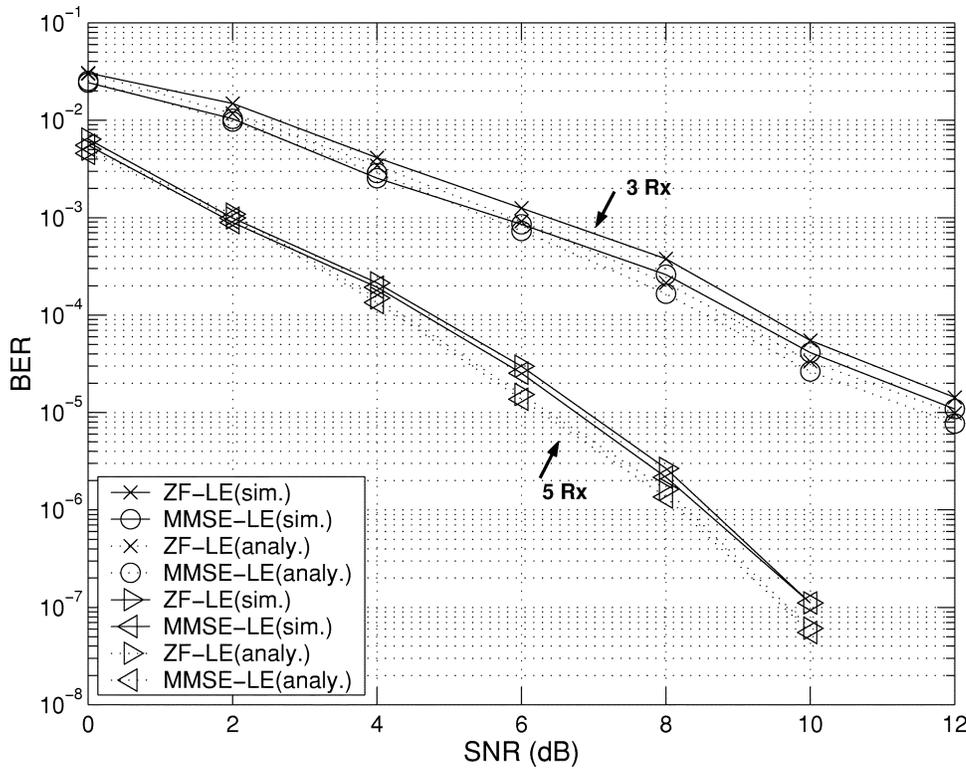


Fig. 5. Simulated and analytical BER versus SNR with LE.

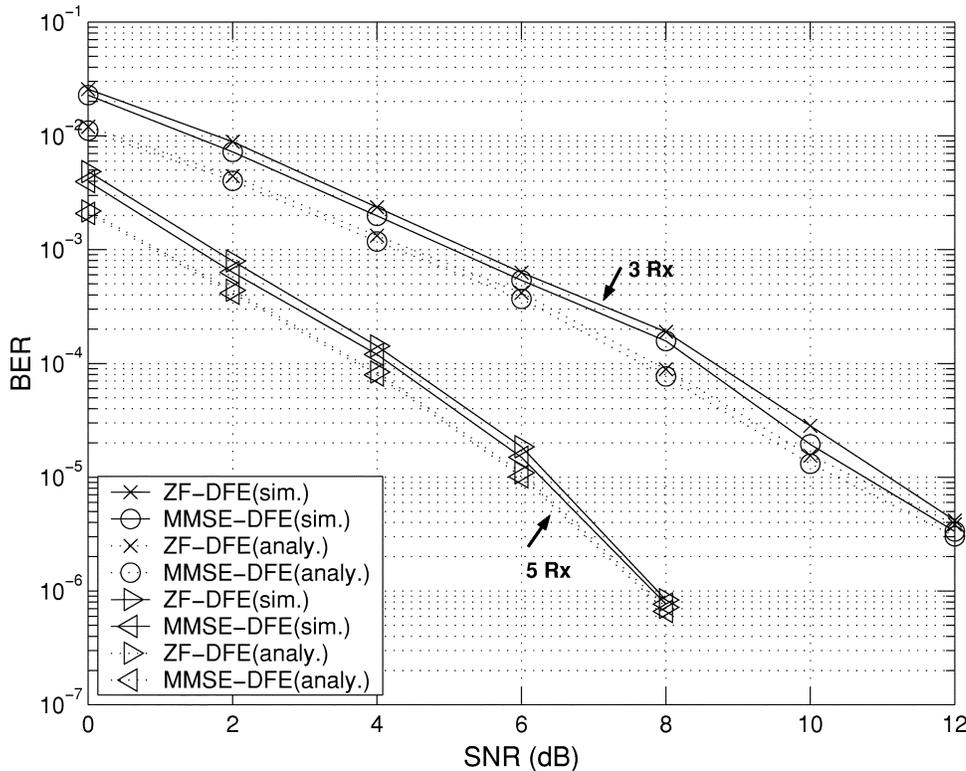


Fig. 6. Simulated and analytical BER versus SNR with DFE.

Figs. 3 and 4 to allow the calculation of $\bar{P}_e(\gamma(k))$ and \bar{P}_e . The BER results presented in Figs. 5 and 6 are also averaged over large number of frames with different fading channels.

The simulated and semi-analytical BER curves depicted in Figs. 5 and 6 with different equalization methods and antenna

numbers match very well, which represent the accuracy of theoretical expressions. Considering the data-estimation schemes with DFE, when the SNR becomes higher, the simulated and semi-analytical curves match better as a result of the fact that the semi-analytical results do not consider the error propagation

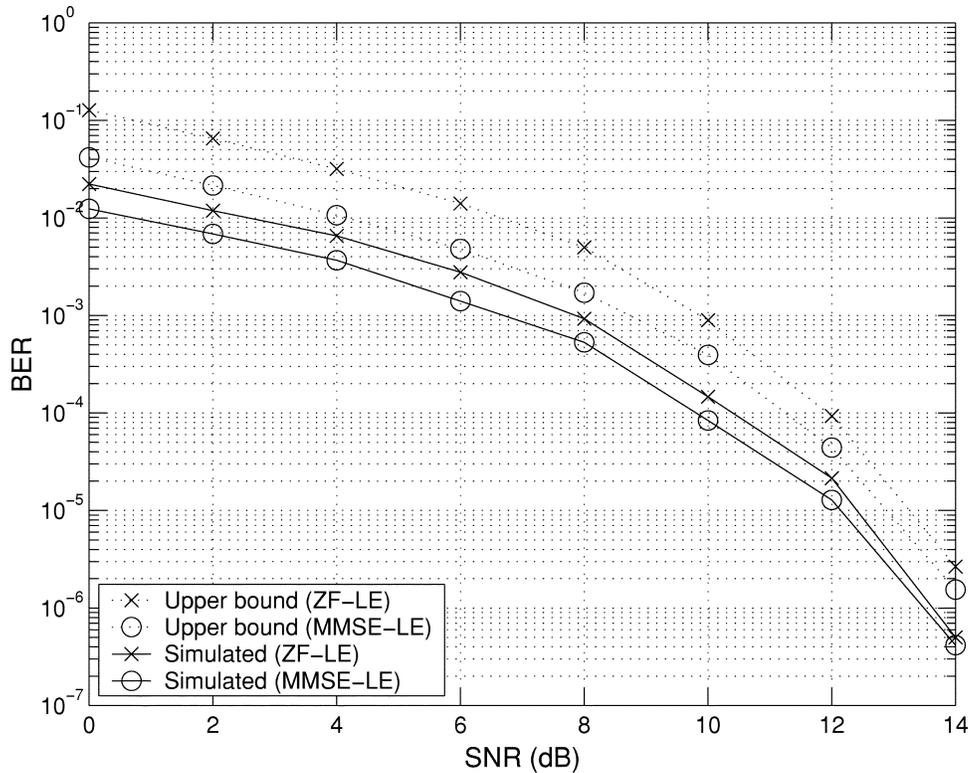


Fig. 7. BER upper bound versus SNR with $M = 3$ Rx.

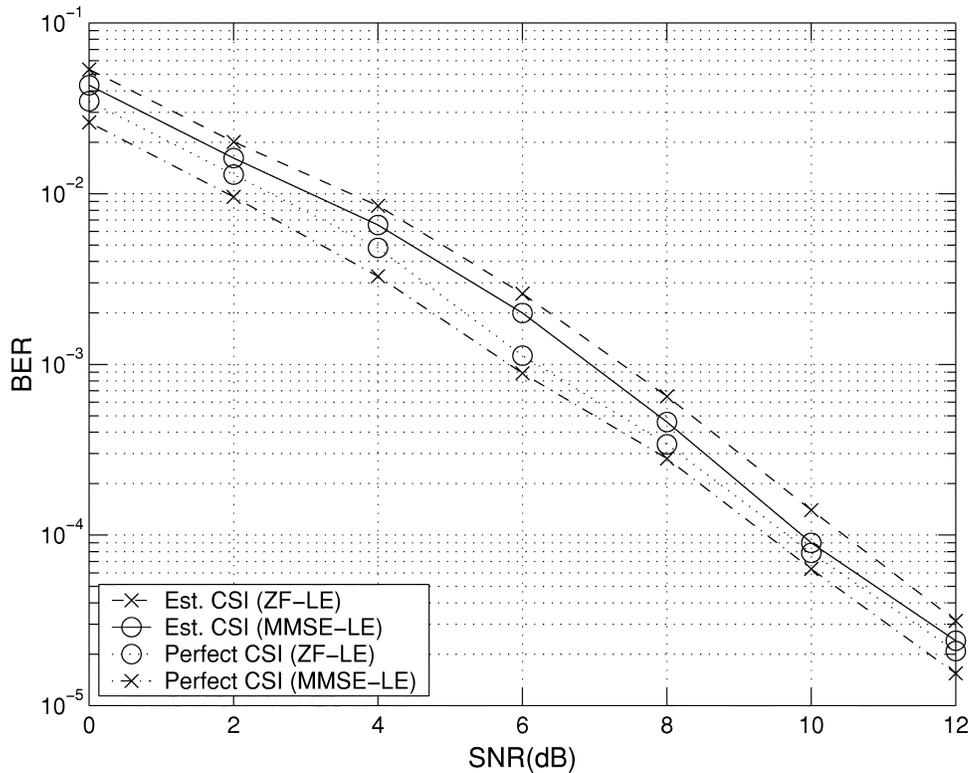


Fig. 8. BER versus SNR with/without channel estimation ($M = 3$ Rx).

effect while the simulated ones do; when the SNR goes higher, the impairing error-propagation effect becomes weaker.

Similarly, in Fig. 7, the simulation results of BER upper bounds of linear equalizers are also obtained with a semi-an-

alytical approach by taking an average over large number of frames. Based on the statistical model of the input symbol, the upper bounds for MMSE-LE and ZF-LE equalizers are calculated as a function of MSE per bit. Fig. 7 shows the

asymptotic tightness of the proposed BER upper bounds of the discussed equalizers and demonstrates the effectiveness of BER bound as a performance measurement. These bounds give an overview of performance of an STTD system using different equalization schemes, but can be evaluated with a reasonable computational load.

C. Channel Estimation

To illustrate the training-aided channel estimation method, comparisons are made in Fig. 8 between the cases with exact and estimated CSI. It is obviously seen that the estimates of CSI are accurate enough in terms of the resulting BERs, which illustrates that the proposed STTD system is tolerant to estimation error due to the existence of diversity and further inspires the utilization of STBC. With an increasing SNR, the performance curves with and without exact CSI are approaching, because the estimate of CSI is approaching the true value.

Simulation results shown in the aforementioned figures demonstrate the efficiency and effectiveness of our proposed methods to estimate channels and to mitigate interference. The results also state that the proposed BER analysis technique is good enough to evaluate the system performance. Compared to the case without ST coding, significant performance improvements can be obtained, which inspires the introduction of an STTD system.

VII. CONCLUSION

We have studied the problem of channel equalization and symbol detection in an STTD system operating in frequency-selective fading channels. The interference cancellation algorithm and the channel estimation method are explored for wireless cellular systems that utilize the spatial and temporal dimensions to combat interference and to detect signals. The proposed schemes are simple to implement, but have good performance. The performance of the resulting receiver in multipath Rayleigh-fading channels has been analyzed theoretically and shown numerically. The performance gain achieved by ST coding over the conventional system using receive diversity only has been demonstrated as well. It is concluded that, in a wireless system, equalization, diversity, and STBC coding can be used together to boost the received signal quality and link performance. Although the system with only two transmit antennas is investigated here, it can be straightforwardly extended to a system with more transmit antennas.

APPENDIX I MATRIX MULTIPLICATION

The discussion in Section III shows the significance of a computationally efficient algorithm to implement matrix multiplication and inversion, which may give rise to a performance enhancement with less computation. Because calculation of $\mathcal{H}\mathcal{T}$ and its conjugate can be easily completed employing the structure of \mathcal{T} , we first consider the evaluation of matrix multiplication $\mathcal{H}^H\mathcal{H}$.

It is easy to show that $\mathcal{H}^H\mathcal{H} = [\mathbf{A}_{11}, \mathbf{A}_{12}; \mathbf{A}_{21}, \mathbf{A}_{22}]$, where

$$\begin{aligned} \mathbf{A}_{11} &= \sum_{i=1}^M \mathbf{H}_{1i}^H \mathbf{H}_{1i}, & \mathbf{A}_{12} &= \sum_{i=1}^M \mathbf{H}_{1i}^H \mathbf{H}_{2i} \\ \mathbf{A}_{21} &= \sum_{i=1}^M \mathbf{H}_{2i}^H \mathbf{H}_{1i}, & \mathbf{A}_{22} &= \sum_{i=1}^M \mathbf{H}_{2i}^H \mathbf{H}_{2i}. \end{aligned} \quad (47)$$

Due to the Toeplitz structure of \mathbf{H}_{ki} , $k = 1, 2$, $i = 1, \dots, M$, a simple method can be found to fulfill the multiplication operation of $\mathbf{H}_{k1i}^H \mathbf{H}_{k2i}$.

As an example, let $\Phi = \mathbf{H}_{1i}^H \mathbf{H}_{1i} \in \mathcal{C}^{q \times q}$ and ϕ_{mn} denote its m th element. Since Φ is Hermitian symmetric, only the elements on and above the diagonal, i.e., ϕ_{mn} for $n \geq m$, need to be evaluated. Furthermore, it can be verified by direct calculation that ϕ_{mn} for $n \geq m$ is given by (48), where it was assumed that $P \geq L+1$ and $h(w) = 0$ for $w < 0$. For simplicity, we omit the subscript $1i$, i.e., $h(\cdot) = h_{1i}(\cdot)$. We see from (48) that the calculation of Φ reduces to a total of $(L+1)(L+2)/2$ quadratic terms $\{h(w_1)^* h(w_2)\}_{w_1, w_2=0}^L$ and their combinations (note that $h^*(w_1)h(w_2) = (h^*(w_2)h(w_1))^*$) and thus can be carried out easily. The similar procedure is applied to computing $\mathbf{H}_{2i}^H \mathbf{H}_{2i}$ [see (48), shown at the bottom of the page].

It is observed that

$$\mathbf{H}_{1i}^H \mathbf{H}_{2i} = (\mathbf{H}_{2i}^H \mathbf{H}_{1i})^H. \quad (49)$$

Hence, only the elements in the upper triangle of these two matrices need to be determined with the elements on the diagonal computed only once. This calculation can be performed by using the method introduced in (48). Therefore, when i is fixed, the total computational load of $\mathbf{H}_{1i}^H \mathbf{H}_{1i}$, $\mathbf{H}_{1i}^H \mathbf{H}_{2i}$, $\mathbf{H}_{2i}^H \mathbf{H}_{1i}$, and $\mathbf{H}_{2i}^H \mathbf{H}_{2i}$ is $(L+1)(2L+3)$ complex multiplication, their combinations and the corresponding complex transpose and complex conjugate operations. Finally, $\mathcal{H}^H\mathcal{H}$ is obtained by summing the related items over i s; then $\mathcal{T}^H\mathcal{H}^H\mathcal{H}\mathcal{T}$ is computed easily by taking advantage of the structure of \mathcal{T} and \mathcal{T}^H with the corresponding elements in $\mathcal{H}^H\mathcal{H}$ multiplied by $-1/+1$ or $j/-j$ and shifted accordingly.

$$\phi_{mn} = \begin{cases} \sum_{w=1}^m h^*(L-w+1)h(L-w+(m-2+1)), & 1 \leq m \leq L+1; m \leq n \leq m+L \\ \sum_{w=1}^{L+1} h^*(L-w+1)h(L-w+(m-n+1)), & L+2 \leq m \leq P; m \leq n \leq m+L \\ \sum_{w=1}^{L+P-m+1} h^*(w-1)h(w-1+(m-n)), & P+1 \leq m \leq P+L; m \leq n \leq P+L \\ 0, & \text{otherwise.} \end{cases} \quad (48)$$

APPENDIX II
MATRIX INVERSION

Regarding computation of matrix inverse $\mathcal{T}^H \mathcal{H}^H \mathcal{H} \mathcal{T}$, it is easy to verify that the inverse of \mathcal{T} has the form of $\mathcal{T}^{-1} = [\mathbf{I}_{q/2} \otimes \mathbf{C}, \mathbf{I}_{q/2} \otimes \mathbf{D}] \in \mathcal{C}^{2q \times 2q}$, where

$$\mathbf{C} = \frac{1}{2} \begin{bmatrix} 1 & 0 \\ -j & 0 \\ 0 & -1 \\ 0 & -j \end{bmatrix}, \quad \mathbf{D} = \frac{1}{2} \begin{bmatrix} 0 & 1 \\ 0 & j \\ 1 & 0 \\ -j & 0 \end{bmatrix}. \quad (50)$$

It is simple to get $(\mathcal{T}^H)^{-1}$ by noticing that $(\mathcal{T}^H)^{-1} = (\mathcal{T}^{-1})^H$.

Because $\mathcal{H}^H \mathcal{H}$ is composed of four blocks, each with reduced dimension $q \times q$, its inverse can be carried out by calculating inverse of the blocks with reduced dimension and performing corresponding multiplications and combinations exploiting the method introduced in [26].

By denoting $\mathbf{B} = (\mathcal{H}^H \mathcal{H})^{-1} = [\mathbf{B}_{11}, \mathbf{B}_{12}; \mathbf{B}_{21}, \mathbf{B}_{22}]$, we may have [26]

$$\begin{aligned} \mathbf{B}_{11} &= (\mathbf{A}_{11} - \mathbf{A}_{12} \mathbf{A}_{22}^{-1} \mathbf{A}_{21}) \\ &= \mathbf{A}_{11}^{-1} + \mathbf{A}_{11}^{-1} \mathbf{A}_{12} \mathbf{B}_{22} \mathbf{A}_{21} \mathbf{A}_{11}^{-1} \\ \mathbf{B}_{22} &= (\mathbf{A}_{22} - \mathbf{A}_{21} \mathbf{A}_{11}^{-1} \mathbf{A}_{12}) \\ &= \mathbf{A}_{22}^{-1} + \mathbf{A}_{22}^{-1} \mathbf{A}_{21} \mathbf{B}_{11} \mathbf{A}_{12} \mathbf{A}_{22}^{-1} \\ \mathbf{B}_{12} &= -\mathbf{A}_{11}^{-1} \mathbf{A}_{12} \mathbf{B}_{22} \\ \mathbf{B}_{21} &= -\mathbf{A}_{22}^{-1} \mathbf{A}_{21} \mathbf{B}_{11}. \end{aligned} \quad (51)$$

From (51), it can be observed that only the inversions of \mathbf{A}_{11} and \mathbf{A}_{22} have to be implemented. Considering that \mathbf{B} is Hermitian symmetric, only the evaluation of \mathbf{B}_{12} or \mathbf{B}_{21} is necessary. Therefore, instead of direct calculating the inverse of a matrix $\mathbf{B} \in \mathcal{C}^{2q \times 2q}$, it is carried out by taking inverse of the reduced dimension matrices $\mathbf{A}_{11} \in \mathcal{C}^{q \times q}$ and $\mathbf{A}_{22} \in \mathcal{C}^{q \times q}$ and then performing corresponding multiplications and combinations to obtain \mathbf{B}_{11} , \mathbf{B}_{22} , \mathbf{B}_{12} , and \mathbf{B}_{21} . The structure of \mathcal{T}^{-1} and $(\mathcal{T}^H)^{-1}$ is used to realize the inverse of $\mathcal{T}^H \mathcal{H}^H \mathcal{H} \mathcal{T}$ finally.

Therefore, $\mathcal{T}^H \mathcal{H}^H \mathcal{H} \mathcal{T}$ and its inversion can be achieved with considerably reduced computational complexity utilizing the procedures described in Appendices I and II, respectively.

REFERENCES

- [1] L. M. Correia and R. Prasad, "An overview of wireless broadband communications," *IEEE Commun. Mag.*, vol. 35, pp. 28–33, Jan. 1997.
- [2] J. K. Tugnait, L. Tong, and Z. Ding, "Single-user channel estimation and equalization," *IEEE Signal Processing Mag.*, vol. 17, no. 3, pp. 17–28, May 2000.
- [3] J. H. Winters, "The diversity gain of transmit diversity in wireless systems with Rayleigh fading," *IEEE Trans. Veh. Technol.*, vol. 47, pp. 119–123, Feb. 1998.
- [4] A. F. Naguib, N. Seshadri, and A. R. Calderbank, "Increasing data rate over wireless channels," *IEEE Signal Processing Mag.*, vol. 17, pp. 76–92, May 2000.
- [5] J. H. Winters, "Optimum combining in digital mobile radio with cochannel interference," in *IEEE J. Select. Areas Commun.*, vol. 2, July 1984, pp. 528–539.
- [6] J. H. Winters, J. Salz, and R. D. Gitlin, "The impact of antenna diversity on the capacity of wireless communication systems," *IEEE Trans. Commun.*, vol. 42, pp. 1740–1751, Feb./Apr. 1994.
- [7] Y. G. Li, "Spatial-temporal processing for wireless mobile systems with ISI and CCI," in *Proc. WCC-ICCT '00*, vol. 2, 2000, pp. 172–179.
- [8] E. Telatar, "Capacity of Multi-Antenna Gaussian Channels," AT&T-Bell Labs Internal Tech. Memo., 1995.
- [9] G. J. Foschini and M. J. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," *Wireless Personal Commun.*, vol. 6, no. 3, pp. 311–335, 1998.
- [10] A. Wittneben, "Basestation modulation diversity for digital SIMULCAST," *Proc. 41st IEEE Veh. Technol. Conf. (VTC '91)*, pp. 848–853, May 1991.
- [11] —, "A new bandwidth efficient transmit antenna modulation diversity scheme for linear digital modulation," in *Proc. IEEE Int. Conf. Commun. (ICC '93)*, Geneva, 1993, pp. 1630–1634.
- [12] G. J. Foschini, "Layered space-time architecture for wireless communication in a fading environment when using multi-element antennas," *The Bell Syst. Tech. J.*, vol. 1, no. 2, pp. 41–59, 1996.
- [13] G. J. Foschini, G. D. Golden, R. A. Valenzuela, and P. W. Wolniansky, "Simplified processing for high spectral efficiency wireless communication employing multi-element arrays," *IEEE J. Select. Areas Commun.*, vol. 17, pp. 1841–1852, Nov. 1999.
- [14] V. Tarokh, N. Seshadri, and A. R. Calderbank, "Space-time codes for high data rate wireless communications: Performance criterion and code construction," *IEEE Trans. Inform. Theory*, vol. 44, pp. 744–765, Mar. 1998.
- [15] S. M. Alamouti, "A simple transmit diversity technique for wireless communications," *IEEE J. Select. Areas Commun.*, vol. 16, pp. 1451–1458, Oct. 1998.
- [16] 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.
- [17] E. Moulines, P. Duhamel, J. F. Cardoso, and S. Mayrargue, "Subspace methods for the blind identification of multichannel FIR filters," *IEEE Trans. Signal Processing*, vol. 43, pp. 516–525, Feb. 1995.
- [18] V. T. Tarokh, S. M. Alamouti, and P. Poon, "New detection schemes for transmit diversity with no channel estimation," in *Proc. IEEE Int. Conf. Universal Pers. Commun. (ICUPC '98)*, vol. 2, Oct. 1998, pp. 917–920.
- [19] B. L. Hughes, "Differential space-time modulation," in *Proc. 1999 IEEE Wireless Commun. Networking Conf. (WCNC '99)*, New Orleans, LA, Sept. 22–29, 1999.
- [20] V. Tarokh and H. Jafarkhani, "A differential detection scheme for transmit diversity," *IEEE J. Select. Areas Commun.*, vol. 18, pp. 1169–1174, July 2000.
- [21] B. L. Hughes, "Differential space-time modulation," *IEEE Trans. Inform. Theory*, vol. 46, pp. 2567–2578, Nov. 2000.
- [22] J. G. Proakis, *Digital Communications*, 3rd ed. New York: McGraw-Hill, 1995.
- [23] S. Verdú, *Multuser Detection*, 1st ed. New York, NY: Cambridge, 1998.
- [24] T. Söderström and P. Stoica, *System Identification*. London, U.K.: Prentice-Hall, 1989.
- [25] S. Haykin, *Adaptive Filter Theory*, 3rd ed. Upper Saddle River, NJ: Prentice-Hall, 1996.
- [26] G. H. Golub and C. F. Van Loan, *Matrix Computations*, 3rd ed. Baltimore, MD: The Johns Hopkins Univ. Press, 1996.
- [27] A. Duel-Hallen, "Decorrelating decision-feedback multiuser detector for synchronous code-division multiple-access channel," *IEEE Trans. Commun.*, vol. 41, pp. 285–290, Feb. 1993.
- [28] Z. D. Wang and G. B. Giannakis, "Wireless multicarrier communications—Where Fourier meets Shannon," *IEEE Signal Processing Mag.*, vol. 17, pp. 29–48, May 2000.
- [29] P. Loubaton and E. Moulines, "On blind multiuser forward link channel estimation by the subspace method: Identifiability results," *IEEE Trans. Signal Processing*, vol. 48, pp. 2366–2376, Aug. 2000.
- [30] G. K. Kaleh, "Channel equalization for block transmission systems," *IEEE J. Select. Areas Commun.*, vol. 13, pp. 110–121, Jan. 1995.
- [31] H. Taub and D. L. Schilling, *Principles of Communication Systems*, 2nd ed. New York: McGraw-Hill, 1986.
- [32] G. J. Foschini and J. Salz, "Digital communications over fading radio channels," *The Bell Syst. Tech. J.*, vol. 62, no. 2, pp. 429–456, 1983.
- [33] J. H. Winters and J. Salz, "Upper bound on the bit-error rate of optimum combining in wireless systems," *IEEE Trans. Commun.*, vol. 46, pp. 1619–1624, Dec. 1998.
- [34] E. D. Carvalho and D. T. M. Slock, "Cramér-Rao bounds for semi-blind, blind and training sequence based channel estimation," in *Proc. 1st IEEE Signal Processing Workshop: Signal Processing Advances in Wireless Communications*, vol. 1, Apr. 1997, pp. 129–132.
- [35] L. Li, H. Li, and Y. Yao, "Channel estimation and interference suppression for space-time coded systems in frequency-selective fading channels," *Wireless Commun. Mobile Comput.*, vol. 2, no. 7, pp. 751–761, 2002.



Ling Li (S'98) received the B.S. and M.S. degrees from the University of Electronic Science and Technology of China (UESTC), Chengdu, in 1991 and 1994, respectively, both in electrical engineering. She is currently working toward the Ph.D. degree at the Department of Electrical and Computer Engineering, Stevens Institute of Technology, Hoboken, NJ.

Her current research interests include wireless communications, digital signal processing (DSP), networking, and multimedia communications.



Yu-Dong Yao (S'88–M'88–SM'94) received the B.Eng. and M.Eng. degrees from Nanjing University of Posts and Telecommunications, Nanjing, China, in 1982 and 1985, respectively, and the Ph.D. degree from Southeast University, Nanjing, China, in 1988, all in electrical engineering.

From 1989 to 1990, he was with Carleton University, Ottawa, ON, Canada, as a Research Associate working on mobile radio communications. He was with Spar Aerospace Ltd., Montreal, PQ, from 1990 to 1994, where he was involved in research on satellite communications.

He was with Qualcomm Inc., San Diego, CA, from 1994 to 2000, where he participated in research and development in wireless code-division multiple-access (CDMA) systems. He joined Stevens Institute of Technology, Hoboken, NJ, in 2000, where he is an Associate Professor in the Department of Electrical and Computer Engineering and a Director of the Wireless Information Systems Engineering Laboratory (WISELAB). His research interests include wireless communications and networks, spread spectrum and CDMA, and DSP for wireless systems.

Dr. Yao holds one Chinese patent and six United States patents. He is an Associate Editor of *IEEE COMMUNICATIONS LETTERS* and *IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY* and an Editor of *IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS*.



Hongbin Li (M'99) received the B.S. and M.S. degrees from the University of Electronic Science and Technology of China (UESTC), Chengdu, in 1991 and 1994, respectively, and the Ph.D. degree from the University of Florida, Gainesville, in 1999, all in electrical engineering.

From July 1996 to May 1999, he was a Research Assistant in the Department of Electrical and Computer Engineering, University of Florida. Since July 1999, he has been an Assistant Professor in the Department of Electrical and Computer Engineering, Stevens Institute of Technology, Hoboken, NJ. He was a Summer Visiting Faculty Member at the Air Force Research Laboratory, Rome, NY, in the summer of 2003. His current research interests include stochastic signal processing, sensor array processing, wireless communications, and radar imaging.

Dr. Li received the Jess H. Davis Memorial Award for excellence in research from Stevens Institute of Technology in 2001 and the Sigma Xi Graduate Research Award from the University of Florida in 1999. He is an Editor for the *IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS*. He is a member of Tau Beta Pi and Phi Kappa Phi.