Diversity Analysis and Design of Space–Time Multiblock Codes for MIMO Systems Equipped With Linear MMSE Receivers

Jing Liu, Timothy N. Davidson, *Member, IEEE*, and K. Max Wong, *Fellow, IEEE*

***Abstract*—This paper addresses the problem of designing optimum full-symbol-rate linear space–time block codes (STBC) for a multi-input multi-output (MIMO) communication system with M transmitter and $N \geq M$ receiver antennas and a linear minimum mean square error (MMSE) receiver. By analyzing the detection error probability expression for the optimized STBC, it is shown that for QAM signaling, the maximum diversity gain for such a system is $N-M+1$. The minimum probability of error STBC design is then extended to systems in which the transmission spans L independent realizations from a block fading channel model, and a (multiblock) linear MMSE receiver is employed. Necessary and sufficient conditions for the optimality of the code are obtained, and a systematic design method for generating codes that satisfy these conditions is presented. The detection error probability and diversity gain of this optimized linear multiblock transceiver are analyzed. It is proved that the error probability decreases with L, and it is shown numerically that the diversity gain increases with L. Thus, if the corresponding latency can be accommodated, for sufficiently large L an optimally designed multiblock system with a linear receiver can exploit the temporal diversity provided by the block-fading channel and achieve higher diversity gain than that of any single-block system of the same symbol rate with a maximum likelihood (ML) receiver. The optimized multiblock linear system achieves this diversity at a substantially lower computational cost. In fact, the structure of the optimal codes can be exploited to significantly reduce the cost of the multiblock linear receiver.**

***Index Terms*—Detection complexity, diversity gain, linear space–time block codes, multiblock transmission, multi-input multi-output (MIMO) systems.**

I. INTRODUCTION

THE design of space–time block codes (STBC) for high rate transmission with low detection error probability is one of the core problems in multi-input multi-output (MIMO) communication systems, e.g., [1]–[10]. A key performance metric for STBCs is the diversity gain. For an arbitrary STBC and an arbitrary receiver, the diversity gain, d, can be defined as [11]

$$d = -\lim_{\rho \to \infty} \frac{\log \mathcal{P}_e}{\log \rho} \tag{1}$$

where $\mathcal{P}_e$ is the chosen measure of the error probability and ρ is the signal-to-noise ratio (SNR). Although it is asymptotic in nature, the diversity gain plays a key role in space–time code design because it is an indication of the rate of decay of the error probability with SNR in the high SNR region; e.g., [12]. For the narrowband richly scattered block fading MIMO channel, the maximum achievable diversity gain is $d = MN$, where M and N are the number of transmitter and receiver antennas, respectively, and several full-symbol-rate STBCs that enable the maximum likelihood (ML) receiver to extract this "full" diversity provided by the channel have been proposed; e.g., [13]–[15]. However, when higher order constellations are employed, the computational requirements of the ML detector are often beyond the capability of the envisioned communication device.

A linear minimum mean square error (MMSE) receiver,[1] on the other hand, is relatively simple to implement. The full-symbol-rate linear STBC that minimizes the bit error rate (BER) for a MIMO system employing a linear MMSE receiver has been presented in [16], [17]. Those codes were shown to provide significantly better BER performance with linear receivers than existing STBCs, but, due to the simplicity of the linear receivers, the diversity provided by the MIMO channel cannot be fully exploited. Using the expression for the minimum detection error probability, we will show in Section II of this paper that the *maximum* achievable diversity gain for a general MIMO system with an arbitrary full-symbol-rate linear STBC and a linear MMSE receiver is $N-M+1$, which is, in general, lower than the full diversity, MN.

The achievable diversity of a STBC for a block-fading MIMO channel can be increased by allowing the codewords to span multiple blocks and employing a multiblock ML receiver [18], [19]. This enables the transceiver to take advantage of the temporal diversity provided by the channel, but the computational cost of the multiblock ML receiver is often prohibitive. As in the single-block case, the use of a linear MMSE receiver reduces this cost, but in the multiblock case the design of appropriate STBC for linear receivers has yet to be resolved. Some experiments performed in [20] have shown that by interchanging

[1]That is, a linear MMSE spatial equalizer followed by symbol-by-symbol detection.

Manuscript received November 24, 2006; revised October 02, 2008 and July 29, 2009. Date of current version September 15, 2010.

J. Liu was with the Department of Electrical and Computer Engineering, McMaster University, Hamilton, ON L8S 4K1, Canada. She is now with Siemens Canada, Peterborough, ON K9J 6X7, Canada.

T. N. Davidson and K. M. Wong are with the Department of Electrical and Computer Engineering, McMaster University, Hamilton, ON L8S 4K1, Canada (e-mail: davidson@mcmaster.ca; wongkm@mcmaster.ca).

Communicated by E. Viterbo, Associate Editor for Coding Techniques.

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TIT.2010.2059639

elements of a (single-block) space–time coded symbol stream between different blocks, the system error performance may be improved.

In Section III of this paper, we will consider the design of an optimum linear multiblock STBC for a MIMO system equipped with a (multiblock) linear MMSE receiver. As in [18], we adopt the conventional independent block fading model in which the channel coefficients remain constant for one block and vary independently from block to block; e.g., [21]. For such a system, we present necessary and sufficient conditions on the structure of the code for the detection error probability to be minimized, and we show that this minimum error probability is a decreasing function of the number of blocks, L. We also show that for codes with this optimal structure, the computational cost of the generic multiblock linear receiver can be substantially reduced. Furthermore, by evaluating the analytic expression for the detection error probability numerically, we show that the diversity order grows with L. Therefore, L can be chosen so that the combination of the optimal full-symbol-rate multiblock STBC and a linear MMSE receiver achieves a higher diversity gain than that of any single-block STBC with ML detection. For a system in which latency can be accommodated, such a scheme would provide high diversity without the high computational complexity associated with ML detection. This is a particular advantage when higher order constellations are employed.

II. Single-Block Transmission

In Section II-A, we will review the optimality conditions in [16] on linear STBCs for single-block transmission and linear reception. Then, in Section II-B, we will determine the diversity gain of codes that satisfy those optimality conditions.

A. Optimal Linear STBC for Single-Block Transmission

Consider a MIMO communication system equipped with M transmitter antennas and N $(N \geq M)$ receiver antennas, in which each of the $N \times M$ channels linking these antennas is assumed to be *block flat-fading*, i.e., the fading channel coefficient h_{nm} that links the mth transmitter and the nth receiver antennas remains constant during a period of T time slots (designated a *block*) and may change to another state independently after the duration elapses. We assume h_{nm} to be independent and identically distributed (i.i.d.) circularly symmetric complex Gaussian with zero mean and unit variance. The channel state information is assumed to be unknown to the transmitter but available at the receiver. Let us form the channel gain matrix $\boldsymbol{H}$ with elements, h_{nm}. If we denote $[x_{1t} \cdots x_{Mt}]^T$ as the vector of signals transmitted by the M antennas at the tth time slot, then, over T time slots, the block of signals to be transmitted through $\boldsymbol{H}$ can be represented by the $M \times T$ signal matrix $\boldsymbol{X}$. For single block transmission, the received signals, denoted by the $N \times T$ matrix $\boldsymbol{Y}$, of the system can then be written as

$$\boldsymbol{Y} = \sqrt{\frac{\rho}{M}}\boldsymbol{H}\boldsymbol{X} + \boldsymbol{W} \tag{2}$$

where ρ is the SNR at each receiver antenna and $\boldsymbol{W}$ is the additive noise, which has a circularly symmetric complex white Gaussian distribution of unit variance, denoted $\mathcal{CN}(\mathbf{0}, \boldsymbol{I})$.

We will focus on systems that employ a full-symbol-rate linear STBC. Since we are considering systems in which $N \geq M$, that means that MT symbols are transmitted per block, and the code word $\boldsymbol{X}$ in (2) is a linear combination of code matrices $\boldsymbol{C}_i$

$$\boldsymbol{X} = \sum_{i=1}^{MT} s_i \boldsymbol{C}_i \tag{3}$$

where s_i are the symbols to be transmitted and are selected from a particular constellation. These symbols s_i are assumed to be independent with zero mean and unit variance. For reasons of implementation simplicity, we will focus on systems with a linear MMSE receiver, and a natural goal is to design the set $\{\boldsymbol{C}_i\}$ so that the detection error probability for that receiver is minimized, while satisfying the power constraint

$$\sum_{i=1}^{MT} \operatorname{tr}\left(\boldsymbol{C}_i^H \boldsymbol{C}_i\right) = MT \tag{4}$$

where $\operatorname{tr}(\cdot)$ denotes trace of a matrix. Several researchers have studied this problem in the past few years [4], [5], [16], [17], [20], [22]. In particular, the following result from [16] provides necessary and sufficient conditions on the structure of the set $\{\boldsymbol{C}_i\}$ to achieve the minimum BER for a system with 4-QAM signaling (and a linear MMSE receiver), under the standard approximation [23] that the residual (spatial) intersymbol interference (ISI) is Gaussian.[2]

Lemma 1 ([16, Theorem 1]): Consider a MIMO system transmitting 4-QAM signals using a full-symbol-rate linear STBC and a linear MMSE receiver. Under the standard Gaussian approximation of the interference at the input to each symbol detector, the average BER, $\mathcal{P}_{\mathrm{es}}$, has a lower bound given by

$$\mathcal{P}_{\mathrm{es}} \geq \mathrm{E}_{\boldsymbol{H}}\left\{Q\left(\sqrt{\frac{M}{\operatorname{tr}\left((\boldsymbol{I} + \frac{\rho}{M}\boldsymbol{H}^H\boldsymbol{H})^{-1}\right)} - 1}\right)\right\} \triangleq \mathcal{P}_{\mathrm{smin}} \tag{5}$$

where $\mathrm{E}_{\boldsymbol{H}}\{\cdot\}$ denotes expectation over the random channels. Equality in (5) holds if and only if the code matrices $\boldsymbol{C}_i$ satisfy the following two conditions:

i) the set $\{\boldsymbol{C}_i\}$ is trace-orthogonal, i.e., $\operatorname{tr}(\boldsymbol{C}_i\boldsymbol{C}_j^H) = \delta_{ij}$ for $i, j = 1, 2, \ldots, MT$, where δ_{ij} is the Kronecker delta;

ii) each $\boldsymbol{C}_i$ is unitary up to a constant, i.e., $\boldsymbol{C}_i\boldsymbol{C}_i^H = \frac{1}{M}\boldsymbol{I}$, $i = 1, 2, \ldots, MT$. □

The Q-function in (5) is defined as $Q(z) = \frac{1}{\sqrt{2\pi}}\int_z^\infty \exp(-x^2/2)dx$. Lemma 1 states that the jointly unitary and trace-orthogonal code structure is necessary and sufficient for minimizing the average BER for single-block transmission of 4-QAM signals, under the Gaussian ISI approximation. These optimality conditions also apply to any square QAM constellation (with Gray coding) at high SNR [16]. They are also necessary and sufficient for the minimization of the average mean square error (MSE) of the detector inputs, irrespective of the transmitted constellation [25]. A code that satisfies the

[2]The distribution of the ISI converges almost surely to a proper (circular) Gaussian distribution as M increases [24].

conditions of Lemma 1 can be systematically generated from a normalized DFT matrix [16].

B. Performance Analysis

For codes designed for single-block transmission, the minimum BER given by Lemma 1 is in the form of the expected value of a Q-function over the random channels. To gain insight into the role of each parameter in the performance of a system employing such codes, we would like to capture the optimal performance in a form that is independent of the random channels. To do so, we examine the minimum BER at high SNRs and determine the diversity gain, d, cf. (1). This can be done by concentrating on the dominant term in the expression in (5) at high SNR, i.e.,

$$\mathcal{P}_{\rm smin} = \mathcal{K}^{-1}\rho^{-d} + (\text{terms involving higher orders of } \rho^{-1})$$

where the coefficient $\mathcal{K}$ is often referred to as the *coding gain* [2].

Since the channels h_{nm} are assumed to be i.i.d. zero-mean Gaussian random variables, the channel matrix $\boldsymbol{H}$ is of distribution $\mathcal{CN}(\boldsymbol{0},\boldsymbol{I})$ and the matrix product $\boldsymbol{H}^H\boldsymbol{H}$ is of *Wishart Distribution* [26], [27]. Let λ_m, $m = 1,2,\ldots,M$, denote the eigenvalues of $\boldsymbol{H}^H\boldsymbol{H}$ ordered as $\lambda_1 > \lambda_2 > \cdots > \lambda_M > 0$. (Inequality between adjacent eigenvalues occurs almost surely.) Then [26], [27], the joint probability density function (pdf) of the eigenvalues is given by

$$p(\lambda_1,\ldots,\lambda_M) = \alpha(M,N)\exp\left(-\sum_{i=1}^{M}\lambda_i\right) \times \prod_{i=1}^{M}\lambda_i^{N-M}\prod_{i<j}^{M}(\lambda_i-\lambda_j)^2 \quad (6)$$

where $\alpha(M,N) = \pi^{M(M-1)}/(2^{MN/2}\,\Gamma_{cM}(M)\,\Gamma_{cM}(N))$, with $\Gamma_c(\,\cdot\,)$ being the Gamma function for complex multivariate, which is defined as $\Gamma_{cM}(\nu) = \pi^{M(M-1)/2}\prod_{i=1}^{M}\Gamma(\nu-i+1)$.

Observe that in (5) the random channel appears in the form $\boldsymbol{H}^H\boldsymbol{H}$ only, and the trace term in the denominator can be expressed as

$$\mathrm{tr}\Big(\boldsymbol{I}+\frac{\rho}{M}\boldsymbol{H}^H\boldsymbol{H}\Big)^{-1} = \sum_{m=1}^{M}\Big(1+\frac{\rho}{M}\lambda_m\Big)^{-1}. \quad (7)$$

Using the standard bound on the Q-function [28]

$$Q(z) \le \frac{1}{2}\exp\left(-\frac{z^2}{2}\right), \quad z \ge 0 \quad (8)$$

and combining (7) and (8) we have the following upper bound on the minimum BER:

$$\mathcal{P}_{\rm smin} \le \frac{\sqrt{e}}{2}\int\cdots\int \exp\left\{-\frac{M}{2\sum_{m=1}^{M}\left(1+\frac{\rho}{M}\lambda_m\right)^{-1}}\right\} \times p(\lambda_1,\ldots,\lambda_M)\,d\lambda_1\cdots d\lambda_M \quad (9)$$

where the joint pdf of the eigenvalues $p(\lambda_1,\ldots,\lambda_M)$ is given by (6). For finite number of transmission antennas, the following result can be obtained by evaluating (9) and focusing on the term containing the lowest order of ρ^{-1}.

Theorem 1: For a MIMO system that transmits 4-QAM signals using the optimal linear STBC described in Lemma 1, the diversity gain $d(N,M)$ and coding gain $\mathcal{K}(N,M)$ obtained by a linear MMSE receiver are given by

$$d(N,M) = N-M+1 \quad (10a)$$

$$\mathcal{K}^{-1}(N,M) = \frac{\alpha(M,N)2^{N-M}e^{-(M-1)/2}(N-M)!}{\alpha(M-1,N+1)}. \quad (10b)$$

□

Proof: See Appendix A. ■

Remarks:

1) Although the derivation provided here is for 4-QAM signals only, Theorem 1 can be extended to a general (Gray coded) square QAM constellation with $2b$ bits per symbol, where b is a positive integer. In particular, using the BER approximation in [29] (and the Gaussian ISI approximation), the minimum BER achieved by a unitary trace-orthogonal code at high SNR is given by [16]

$$\tilde{\mathcal{P}}_{\rm smin} \simeq \mathrm{E}_{\boldsymbol{H}}\left\{\zeta Q\left(\sqrt{\vartheta\left(\frac{M}{\mathrm{tr}\left(\left(\boldsymbol{I}+\frac{\rho}{M}\boldsymbol{H}^H\boldsymbol{H}\right)^{-1}\right)}-1\right)}\right) + \eta Q\left(3\sqrt{\vartheta\left(\frac{M}{\mathrm{tr}\left(\left(\boldsymbol{I}+\frac{\rho}{M}\boldsymbol{H}^H\boldsymbol{H}\right)^{-1}\right)}-1\right)}\right)\right\} \quad (11)$$

where $\zeta = \frac{2^b-1}{b2^{b-1}}, \eta = \frac{2^b-2}{b2^{b-1}}$ and $\vartheta = \frac{3b}{4^b-1}$. We observe that the constants ζ, η and ϑ are positive and independent of ρ, and hence (11) and (5) have the same dominant order of ρ^{-1}, i.e., the same diversity order. However, the coding gain does depend on the signal constellation.

2) The dependence of the diversity gain in (10a) on $(N-M)$ implies that increasing the degrees of freedom (i.e., M and N) of a full-symbol-rate system does not guarantee improved diversity gain for a linear receiver.

3) It is interesting to compare the diversity gain in (10a) with that of a MIMO system in which the data symbols are transmitted directly, without a space–time code (i.e., spatial multiplexing), and a linear zero-forcing (ZF) receiver is employed. In [30], [31], [32, p. 546], and [33], it was shown that the diversity of that scheme is also $N-M+1$. This shows that when a linear receiver is employed, the use of a BER-optimal full-symbol-rate STBC does not increase the diversity gain.[3] (The lower BER obtained by the optimal STBC is due to an increase in the coding gain.) This is in contrast to the case of ML detection. In that case, the diversity of the direct transmission scheme is N, whereas a system that employs a well-designed full-symbol-rate STBC can obtain the maximum diversity gain, which is MN.

Since the diversity achieved in (10a) is less than the maximum diversity obtainable for single-block coded systems employing

[3]Analogous insight is obtained from the recently derived diversity-versus-multiplexing trade-off for linear receivers [34].

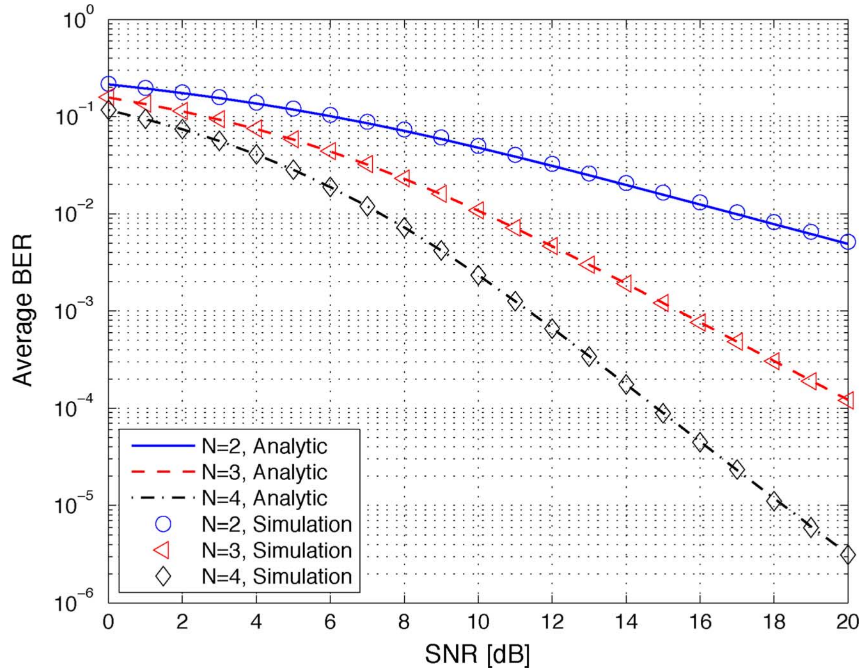

Fig. 1. $\mathcal{P}_{\rm smin}$ (curve) and simulated average BER (symbols) for optimized single-block transmission. $M = 2, N = 2, 3, 4$.

an ML detector, in the next section we will develop a multiblock transmission scheme that increases the diversity of MIMO systems with a linear MMSE receiver. Before we do so, we provide a numerical example that verifies the diversity analysis above.

Example 1: In this example, we verify the diversity analysis in Theorem 1 by evaluating the BER performance of an optimal single-block transmission scheme with 4-QAM signaling in an i.i.d. Rayleigh fading environment. (Performance comparisons with other STBCs were provided in [16].) The performance is evaluated using the expression in (5) and via simulation. Two of the specific space–time codes used in the simulations are provided in Appendix C.

1) In the first experiment, we consider a system with $M = T = 2$, and hence the transmission data rate is $R = 4$ bits per channel use (pcu). In Fig. 1 we consider the cases of $N = 2, 3$ and 4 receiver antennas and we show that the negative slopes of the (analytic and simulated) average BER curves increase with N and are equal to $(N - M + 1)$, verifying the results in Theorem 1. (These curves also illustrate that the Gaussian approximation of the ISI leads to an accurate expression for the minimum average BER, even when M is only 2.)
2) In the second experiment, we fix $M = N = T = n$, and we examine the BER performance for $n = 2, 4$ and 8. The resulting averaged BER curves are plotted in Fig. 2. We observe that at high SNR, all the curves have the same slope, and hence, the same diversity gain. However, the larger the value of M, the lower the BER that the system achieves, even though the data rate is $2M$ bits pcu. These observations verify the analysis (cf. Theorem 1), which showed that diversity gain is only related to $(N - M)$, and that systems with larger M have larger coding gains. □

III. Optimal Linear STBC for Multiblock Transmission With Linear MMSE Receiver

In the optimal STBC for single-block transmission presented in Section II-A, each data symbol is distributed among the MT elements of the matrix $\boldsymbol{X}$, but is constrained to the given block of T channel uses. In multiblock transmission [18], each symbol may be distributed over all the elements of the L block signal matrices $\boldsymbol{X}_\ell, \ell = 1, 2, \ldots, L$, and this offers the opportunity to take advantage of the time diversity offered by the block-fading channel. In this section, we examine the design of a multiblock linear STBC for use with a (multiblock) linear MMSE receiver, and show that a substantial increase in diversity can be achieved.

Consider L blocks of the independent block fading channel model, and let $\boldsymbol{H}_\ell$ denote the channel matrix for the ℓth block. For full-symbol-rate transmission, we require LMT symbols to be transmitted during these L blocks, i.e., $\{s_i\}, i = 1, 2, \ldots, LMT$. We consider a linear multiblock space time code in which each $\boldsymbol{X}_\ell$ can be written as

$$\boldsymbol{X}_\ell = \sum_{i=1}^{LMT} s_i \boldsymbol{C}_{i\ell}, \quad \ell = 1, 2, \ldots, L \tag{12}$$

where $\boldsymbol{C}_{i\ell}$ are the $M \times T$ matrices to be designed and are constrained to have the power evenly allocated over all blocks,

$$\sum_{i=1}^{LMT} \operatorname{tr}\left(\boldsymbol{C}_{i\ell}^H \boldsymbol{C}_{i\ell}\right) = MT, \quad \forall \ell. \tag{13}$$

Using (12), the received signal for each transmitted block is given by

$$\boldsymbol{Y}_\ell = \sqrt{\frac{\rho}{M}} \boldsymbol{H}_\ell \left(\sum_{i=1}^{LMT} s_i \boldsymbol{C}_{i\ell} \right) + \boldsymbol{W}_\ell, \quad \ell = 1, 2, \ldots, L. \tag{14}$$

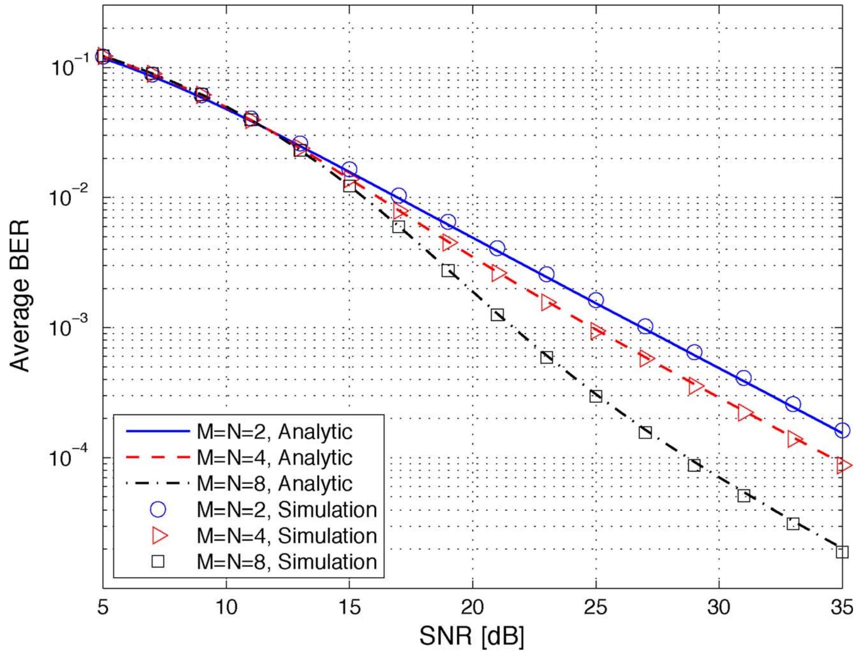

Fig. 2. $\mathcal{P}_{\rm smin}$ (curve) and simulated average BER (symbols) for optimized single-block transmission. $M = N = 2, 4, 8$.

Since the received signals from each block contain information from all LMT symbols, we need to jointly detect the signals from all L blocks at the receiver. If we collect the LMT symbols transmitted in blocks $1 \leq \ell \leq L$ in the vector $\boldsymbol{s} = [s_1, \ldots, s_{LMT}]^T$, and if we vectorize each block of received data in (14) with $\boldsymbol{y}_\ell = \mathrm{vec}(\boldsymbol{Y}_\ell)$ and $\boldsymbol{w}_\ell = \mathrm{vec}(\boldsymbol{W}_\ell)$, then we can write the L blocks of received signals as

$$\underbrace{\begin{bmatrix} \boldsymbol{y}_1 \\ \boldsymbol{y}_2 \\ \vdots \\ \boldsymbol{y}_L \end{bmatrix}}_{\boldsymbol{y}} = \sqrt{\frac{\rho}{M}} \underbrace{\begin{bmatrix} \boldsymbol{I}_T \otimes \boldsymbol{H}_1 & & & \\ & \boldsymbol{I}_T \otimes \boldsymbol{H}_2 & & \\ & & \ddots & \\ & & & \boldsymbol{I}_T \otimes \boldsymbol{H}_L \end{bmatrix}}_{\boldsymbol{\mathcal{H}}} \times \underbrace{\begin{bmatrix} \mathrm{vec}\left(\boldsymbol{C}_{11}\right) \cdots \mathrm{vec}\left(\boldsymbol{C}_{(LMT)1}\right) \\ \mathrm{vec}\left(\boldsymbol{C}_{12}\right) \cdots \mathrm{vec}\left(\boldsymbol{C}_{(LMT)2}\right) \\ \vdots \\ \mathrm{vec}\left(\boldsymbol{C}_{1L}\right) \cdots \mathrm{vec}\left(\boldsymbol{C}_{(LMT)L}\right) \end{bmatrix}}_{\boldsymbol{F}} \boldsymbol{s} + \underbrace{\begin{bmatrix} \boldsymbol{w}_1 \\ \boldsymbol{w}_2 \\ \vdots \\ \boldsymbol{w}_L \end{bmatrix}}_{\boldsymbol{w}}. \quad (15)$$

Using the underbraced terms, we can rewrite (15) as

$$\boldsymbol{y} = \sqrt{\frac{\rho}{M}} \boldsymbol{\mathcal{H}} \boldsymbol{F} \boldsymbol{s} + \boldsymbol{w} \quad (16)$$

where the equivalent channel matrix $\boldsymbol{\mathcal{H}}$ is of dimension $LNT \times LMT$, and $\boldsymbol{F}$ is the $LMT \times LMT$ square coding matrix that is to be designed. We note that the ith column of $\boldsymbol{F}$, denoted by $\boldsymbol{f}_i$, is an LMT-dimensional vector containing elements of the code matrices for the ith symbol s_i in all L blocks, i.e.,

$$\boldsymbol{f}_i = \left[\, \mathrm{vec}^T(\boldsymbol{C}_{i1}) \quad \mathrm{vec}^T(\boldsymbol{C}_{i2}) \quad \cdots \quad \mathrm{vec}^T(\boldsymbol{C}_{iL}) \,\right]^T. \quad (17)$$

The (multiblock) linear MMSE equalizer for the system model in (16) is (e.g., [35])

$$\boldsymbol{G} = \sqrt{\frac{\rho}{M}} \left(\boldsymbol{I} + \frac{\rho}{M} \boldsymbol{F}^H \boldsymbol{\mathcal{H}}^H \boldsymbol{\mathcal{H}} \boldsymbol{F} \right)^{-1} \boldsymbol{F}^H \boldsymbol{\mathcal{H}}^H \quad (18)$$

and the equalized signal $\hat{\boldsymbol{s}}$ is

$$\hat{\boldsymbol{s}} = \boldsymbol{G}\boldsymbol{y} = \sqrt{\frac{\rho}{M}} \boldsymbol{G} \boldsymbol{\mathcal{H}} \boldsymbol{F} \boldsymbol{s} + \boldsymbol{G}\boldsymbol{w}. \quad (19)$$

Each of the elements of $\hat{\boldsymbol{s}}$ is then detected by a symbol-by-symbol detector. The symbol error covariance matrix of the detector inputs is given by

$$\boldsymbol{V}_{\rm se} \triangleq \mathrm{E}\{(\boldsymbol{s} - \hat{\boldsymbol{s}})(\boldsymbol{s} - \hat{\boldsymbol{s}})^H\} = \left(\boldsymbol{I} + \frac{\rho}{M} \boldsymbol{F}^H \boldsymbol{\mathcal{H}}^H \boldsymbol{\mathcal{H}} \boldsymbol{F} \right)^{-1}. \quad (20)$$

For later convenience, we define $\boldsymbol{\sigma} = [\boldsymbol{s}_{\rm re}^T, \boldsymbol{s}_{\rm im}^T]^T$, where $\boldsymbol{s}_{\rm re}$ and $\boldsymbol{s}_{\rm im}$ are the real and imaginary parts of $\boldsymbol{s}$, respectively, and $\hat{\boldsymbol{\sigma}}$ to be the corresponding equalized signal. The error covariance of $\boldsymbol{\sigma}$ can be written as

$$\boldsymbol{V} \triangleq \mathrm{E}\{(\boldsymbol{\sigma} - \hat{\boldsymbol{\sigma}})(\boldsymbol{\sigma} - \hat{\boldsymbol{\sigma}})^T\} = \left(\boldsymbol{I} + \frac{\rho}{M} \boldsymbol{J}^H \hat{\boldsymbol{F}}^H \hat{\boldsymbol{\mathcal{H}}}^H \hat{\boldsymbol{\mathcal{H}}} \hat{\boldsymbol{F}} \boldsymbol{J} \right)^{-1} \quad (21)$$

where $\hat{\boldsymbol{\mathcal{H}}}$ is of dimension $2LNT \times 2LMT$, and $\hat{\boldsymbol{F}}$ and $\boldsymbol{J}$ are both $2LMT \times 2LMT$ and are respectively defined as

$$\hat{\boldsymbol{\mathcal{H}}} \triangleq \begin{bmatrix} \boldsymbol{\mathcal{H}} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{\mathcal{H}}^* \end{bmatrix}, \quad \hat{\boldsymbol{F}} \triangleq \begin{bmatrix} \boldsymbol{F} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{F}^* \end{bmatrix}$$

$$\boldsymbol{J} \triangleq \frac{1}{\sqrt{2}} \begin{bmatrix} \boldsymbol{I}_{LMT} & j\boldsymbol{I}_{LMT} \\ \boldsymbol{I}_{LMT} & -j\boldsymbol{I}_{LMT} \end{bmatrix}$$

where the superscript "$*$" denotes complex conjugate.

As is in the case of single-block STBC design, we will first consider the bit error probability for 4-QAM symbols. If we approximate the residual ISI at the output of the linear MMSE equalizer by zero-mean Gaussian noise [23], [24], the averaged error probability over the $2LMT$ bits can be written as

$$\mathcal{P}_{\rm em} = {\rm E}_{\boldsymbol{\mathcal{H}}}\left\{\frac{1}{2LMT}\sum_{k=1}^{2LMT} Q\left(\sqrt{[\boldsymbol{V}]_{kk}^{-1}-1}\right)\right\}. \tag{22}$$

The objective now is to obtain the structure of codes that minimize this error probability; i.e., the structure of $\arg\min_{\boldsymbol{F}} \mathcal{P}_{\rm em}$ subject to the power constraints in (13). The solution to this problem is provided by the following theorem:

Theorem 2: Consider a MIMO system transmitting 4-QAM signals using a full-symbol-rate multiblock linear STBC and a linear MMSE receiver. Under the standard Gaussian approximation of the interference at the input to each symbol detector, the average BER $\mathcal{P}_{\rm em}$ has a lower bound given by

$$\mathcal{P}_{\rm em} \geq {\rm E}_{\{\boldsymbol{H}_\ell\}}\left\{Q\left(\sqrt{\frac{LM}{\sum_{\ell=1}^{L}{\rm tr}\left(\left(\boldsymbol{I}+\frac{\rho}{M}\boldsymbol{H}_\ell^H\boldsymbol{H}_\ell\right)^{-1}\right)}-1}\right)\right\}$$
$$\triangleq \mathcal{P}_{\rm mmin}. \tag{23}$$

Equality in (23) holds if and only if the code matrices $\boldsymbol{C}_{i\ell}$ satisfy the following two conditions:

i) the set $\{\boldsymbol{C}_{i\ell}\}$ is trace-orthogonal in an L-block sense, i.e.,

$$\sum_{\ell=1}^{L}{\rm tr}\left(\boldsymbol{C}_{i\ell}^H\boldsymbol{C}_{j\ell}\right) = \delta_{ij} \quad \text{for } i,j = 1,2,\ldots,MT \tag{24}$$

ii) each $\boldsymbol{C}_{i\ell}$ is unitary up to a constant, i.e., $\boldsymbol{C}_{i\ell}\boldsymbol{C}_{i\ell}^H = \frac{1}{ML}\boldsymbol{I}, i=1,2,\ldots,MT, \ell=1,2,\ldots,L$. □

Proof: See Appendix B. ■

Remarks on Theorem 2:

1) Lemma 1 is a special case of Theorem 2; i.e., for $L=1$, the statements are identical.
2) Following similar arguments to those in [16], [25], and [36] and those that follow Theorem 1, the two conditions in Theorem 2 can be proved to be necessary and sufficient in minimizing BER for any square QAM signal constellation at high SNR, and necessary and sufficient for the minimization of the average MSE of the detector inputs, irrespective of the transmitted constellation.
3) Since $\boldsymbol{C}_{i\ell}$ is the code for the ith symbol transmitted through the ℓth realization of channel, the first optimality condition enforces an orthogonality relation between different symbols while maintaining equal power allocation for each symbol. The second optimality condition ensures that each symbol, s_i, is evenly distributed over all the channel fading coefficients for each of the L channel realizations (via the use of a unitary coding matrix $\boldsymbol{C}_{i\ell}$), and the power allocated to transmit each symbol through each of the channel realizations is equal.
4) Equation (23) shows that $\mathcal{P}_{\rm mmin}$ is a function of L; i.e., $\mathcal{P}_{\rm mmin}(L)$. In the following, we will prove that $\mathcal{P}_{\rm mmin}(L) > \mathcal{P}_{\rm mmin}(L+1)$; i.e., the detection error probability decreases with the increase of L. For notational simplicity, we write $\mathcal{P}_{\rm mmin}(L) = {\rm E}\{\Upsilon(\frac{1}{L}\sum_{\ell=1}^{L}\xi_\ell)\}$, where $\Upsilon(x) \triangleq Q(\sqrt{x^{-1}-1})$, and $\xi_\ell = \frac{1}{M}{\rm tr}(\boldsymbol{I}+\frac{\rho}{M}\boldsymbol{H}_\ell^H\boldsymbol{H}_\ell)^{-1}$. First, we note that $\Upsilon(x)$ is a convex function of x when $0 < x \leq 1$; e.g., [16]. Now, consider having $(L+1)$ independent blocks, among which we choose any combination of L different blocks for transmission. In total, there are $(L+1)$ different combinations. For each combination of L blocks, we jointly transmit signals that have been space–time coded in the manner described in Theorem 2, and the expected value of the error probability over the random channels is $P_{\rm mmin}(L)$. Averaging over all the $L+1$ combinations, this value remains unchanged; i.e.,

$$\begin{aligned}\mathcal{P}_{\rm mmin}(L) &= {\rm E}\left\{\Upsilon\left(\frac{1}{L}\sum_{\ell=1}^{L}\xi_\ell\right)\right\}\\ &= \frac{1}{L+1}\sum_{i=1}^{L+1}{\rm E}\left\{\Upsilon\left(\frac{1}{L}\sum_{\ell=1,\ell\neq i}^{L+1}\xi_\ell\right)\right\}.\end{aligned} \tag{25}$$

Applying Jensen's inequality [37] to the right side of (25), which contains the convex function $\Upsilon(x)$, we arrive at

$$\begin{aligned}\mathcal{P}_{\rm mmin}(L) &\geq {\rm E}\left\{\Upsilon\left(\frac{1}{L+1}\sum_{i=1}^{L+1}\left(\frac{1}{L}\sum_{\ell=1,\ell\neq i}^{L+1}\xi_\ell\right)\right)\right\}\\ &= {\rm E}\left\{\Upsilon\left(\frac{1}{L+1}\sum_{\ell=1}^{L+1}\xi_\ell\right)\right\}\\ &= \mathcal{P}_{\rm mmin}(L+1).\end{aligned} \tag{26}$$

Equality in (26) holds iff the channel realization in each of the $(L+1)$ blocks is identical. Therefore, under our independent block fading model, we have strict inequality in (26) with probability one.

IV. Performance Analysis: Multiblock Transmission Using Optimum Code

In the previous section, we derived an expression whose expected value over the random channels yields the minimum BER $\mathcal{P}_{\rm mmin}$ for multiblock transmission; cf. (23). We also proved that $\mathcal{P}_{\rm mmin}$ decreases as L increases. However, to understand the roles that each parameter plays, we would like to evaluate the expectation in (23). First, by writing $\boldsymbol{\lambda}_\ell = [\lambda_{\ell 1},\ldots,\lambda_{\ell M}]$, with $\lambda_{\ell m}$ being the mth eigenvalue for $\boldsymbol{H}_\ell^H\boldsymbol{H}_\ell, \mathcal{P}_{\rm mmin}$ can be written as

$$\mathcal{P}_{\rm mmin} = \int\cdots\int Q\left(\sqrt{\frac{ML}{\sum_{\ell=1}^{L}\sum_{m=1}^{M}(1+\frac{\rho}{M}\lambda_{\ell m})^{-1}}-1}\right) \times p(\boldsymbol{\lambda}_1,\ldots,\boldsymbol{\lambda}_L)\,d\boldsymbol{\lambda}_1\cdots d\boldsymbol{\lambda}_L \tag{27}$$

where $p(\boldsymbol{\lambda}_1,\ldots,\boldsymbol{\lambda}_L)$ denotes the joint pdf of all $\lambda_{\ell m}, \ell = 1,\ldots,L,\ m = 1,\ldots,M$. Since the channel matrices $\boldsymbol{H}_\ell$ are statistically independent for different ℓ, we have $p(\boldsymbol{\lambda}_1,\ldots,\boldsymbol{\lambda}_L) = \prod_{\ell=1}^{L}p(\boldsymbol{\lambda}_\ell)$, where each $p(\boldsymbol{\lambda}_\ell) = p(\lambda_{\ell 1},\ldots,\lambda_{\ell M})$ satisfies (6). Due to the complexity of the Q-function and $p(\boldsymbol{\lambda}_\ell)$, the integration in (27) appears to be intractable. However, we observe from (27) that $\mathcal{P}_{\rm mmin}$

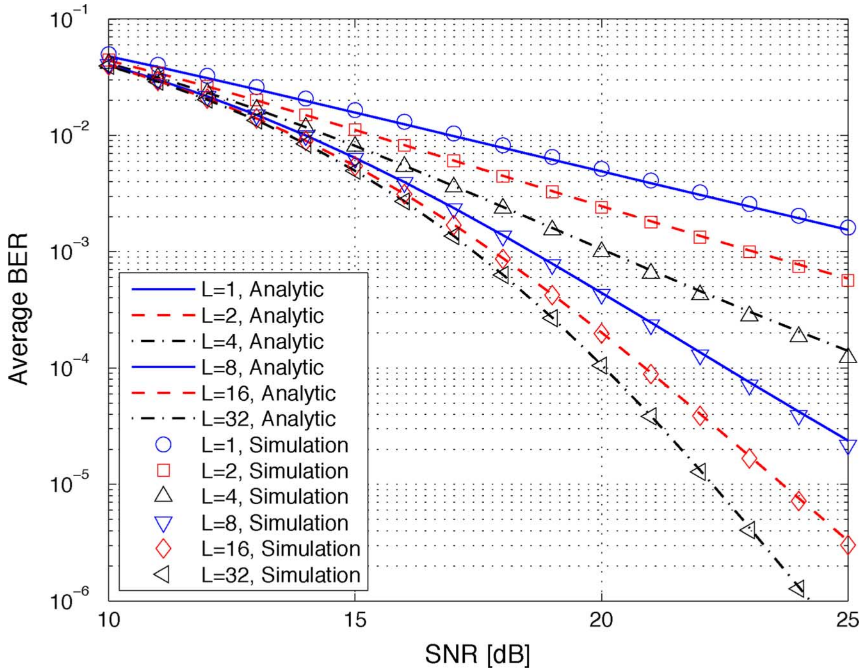

Fig. 3. $\mathcal{P}_{\rm mmin}$ (curve) and simulated average BER (symbols) for optimized multiblock transmission for different L for the case of $M = N = 2$.

depends on the parameters $M, N, L,$ and ρ. To examine the properties of $\mathcal{P}_{\rm mmin}$ in relation to multiblock design, in the following we evaluate numerically $\mathcal{P}_{\rm mmin}$ in terms of $\rho, L, M,$ and N, keeping $M = N$ so that in the case of $L = 1$, the diversity $d = 1$.

A. Finite L

Since the channel eigenvalues are independent from one block to another, the evaluation of (27) with respect to L is relatively simple. However, the channel eigenvalues are not independent within each block, hence this evaluation could be quite tedious for large M. The task can be made simpler following the general procedure outlined below.

i) We let $\xi_{\ell m} = (1 + \frac{\rho}{M}\lambda_{\ell m})^{-1}$, so that $p_{\xi_\ell}(\xi_{\ell 1}, \ldots, \xi_{\ell M}) = p(\lambda_{\ell 1}, \ldots, \lambda_{\ell M})/|\frac{\partial(\xi_{\ell 1}, \ldots, \xi_{\ell M})}{\partial(\lambda_{\ell 1}, \ldots, \lambda_{\ell M})}|$.

ii) Let $\zeta_\ell = \sum_{m=1}^{M} \xi_{\ell m}$. Then we can obtain $p_{\zeta_\ell}(\zeta_\ell)$ by successive integration such that, for $\ell = 1, 2, \ldots, L$

$$p_{\zeta_\ell}(\zeta_\ell) = \int \cdots \int p(\zeta_\ell - u_{M-1}, u_{M-1} - u_{M-2}, \ldots, u_2 - u_1, u_1)\, du_1 \cdots du_{M-1} \quad (28)$$

where $u_1 = \xi_{\ell M}, u_2 = \xi_{\ell(M-1)} + \xi_{\ell M}, \ldots, u_{M-1} = \xi_{\ell 2} + \cdots + \xi_{\ell M}$.

iii) Let $\bar{\zeta} = \sum_{\ell=1}^{L} \zeta_\ell$. Since the random variables ζ_ℓ are independent from block to block, we have

$$p_{\bar{\zeta}}(\bar{\zeta}) = p_{\zeta_1}(\zeta_1) \star p_{\zeta_2}(\zeta_2) \star \cdots \star p_{\zeta_L}(\zeta_L) \quad (29)$$

where "$\star$" denotes convolution, which can be efficiently carried out using the FFT.

iv) Let $z = 1/\bar{\zeta}$, then $p_z(z) = p(\bar{\zeta})/|\frac{dz}{d\bar{\zeta}}|$ and (27) becomes

$$\mathcal{P}_{\rm mmin} = \int Q(\sqrt{LMz - 1}) p_z(z)\, dz. \quad (30)$$

For finite L and M, the above procedure yields, numerically, the value of $\mathcal{P}_{\rm mmin}$, from which we can appraise the gain obtained from the multiblock design, as we will demonstrate in the following simple example.

Example 2: In this example we evaluate the performance of an optimized multiblock linear transceiver via numerical evaluation of $\mathcal{P}_{\rm mmin}$ and BER simulations, under the i.i.d. Rayleigh channel model that we have considered. (Some of the specific space–time codes used in the simulations are provided in Appendix C.) In Fig. 3 we have plotted the bit error rate against the SNR for different values of L for the case of $M = T = N = 2$, for which $R = 4$ bits pcu. This figure confirms the statement in Remark 4 following Theorem 2 that $\mathcal{P}_{\rm mmin}$ decreases as L increases, and it illustrates the manner in which $\mathcal{P}_{\rm mmin}$ decreases with increasing L.

In order to assess the impact of increasing L on the achieved diversity, we numerically evaluated (and scaled) the negative slope of the error rate curves at high SNR; cf. (1). The (scaled) slopes at an SNR of 24 dB are plotted in Fig. 4 for the cases of $M = T = N = \{1, 2, 4\}$ and a number of values of L. We observe that, as predicted by Theorem 1, for $L = 1$ all three systems have a slope of $N - M + 1 = 1$. We also observe that the slopes become larger (i.e., the error rate curves become steeper) as L increases. As one might intuitively expect, for $L > 1$ the high SNR slope of the systems with more antennas is steeper than those with fewer antennas, even through the data rate is $2M$ bits pcu. □

B. For Asymptotically Large L

When the code is jointly designed for a large number of blocks, the effect of L can be analyzed in the following way. Let the fractional part in the argument of the Q-function in (27) be written as $1/(\frac{1}{L}\sum_{\ell=1}^{L} \frac{1}{M} \sum_{m=1}^{M} (1 + \frac{\rho}{M}\lambda_{\ell m})^{-1})$. As

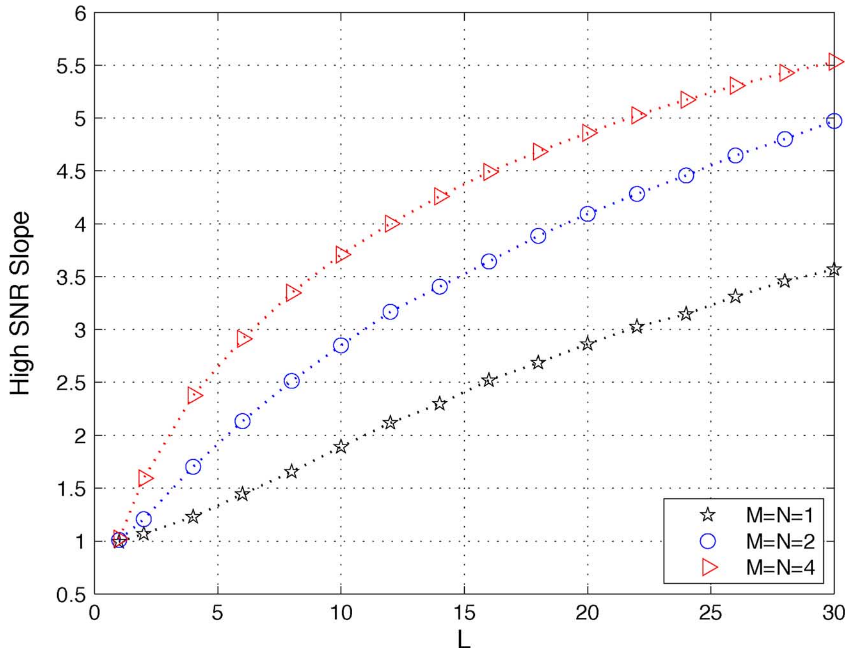

Fig. 4. Numerical approximation of $d\log \mathcal{P}_{\mathrm{mmin}}/d\log\rho$ at $\rho = 24$ dB for different values of L for the cases of $M = N = 1$, $M = N = 2$, and $M = N = 4$.

$L \to \infty$, we can apply the law of large numbers and write the denominator as

$$\begin{aligned}\bar{\xi} &= \mathrm{E}\left\{\frac{1}{M}\sum_{m=1}^{M}\frac{1}{1+\frac{\rho}{M}\lambda_m}\right\}\\ &= \frac{1}{M}\sum_{m=1}^{M}\mathrm{E}\left\{\frac{1}{1+\frac{\rho}{M}\lambda_m}\right\} = \mathrm{E}\left\{\frac{1}{1+\frac{\rho}{M}\lambda}\right\} \end{aligned}\tag{31}$$

where the last step is valid for any positive number M since the summand is a constant. The pdf of λ is the marginal density, which can be obtainable from the joint pdf of the *ordered* eigenvalues $\lambda_1 > \cdots > \lambda_M$ in (6) by first deriving the joint pdf of the *unordered* eigenvalues and then integrating, resulting in [3]

$$p(\lambda) = \frac{1}{M}\lambda^{N-M}e^{-\lambda}\sum_{k=0}^{M-1}\frac{k!}{(k+N-M)!}\left(L_k^{N-M}(\lambda)\right)^2 \tag{32}$$

where $L_k^{N-M}(\lambda) = \sum_{i=0}^{k}\binom{k+N-M}{k-i}\frac{(-\lambda)^i}{i!}$ is the associated Laguerre polynomial of order $(N-M)$ and degree k, with $\binom{n}{k}$ being the binomial coefficient. Substituting (32) into (31), we obtain

$$\bar{\xi} = \frac{1}{M}\int_0^{\infty}\frac{\lambda^{N-M}\sum_{p=0}^{2(M-1)}a_p\lambda^p}{1+\frac{\rho}{M}\lambda}e^{-\lambda}\,d\lambda \tag{33}$$

where, for notational simplicity, we have rewritten the square of $L_k^{N-M}(\lambda)$ as a polynomial in ascending powers of λ with coefficients a_p. If we let $x = \lambda + \frac{M}{\rho}$, (33) can be written as

$$\begin{aligned}\bar{\xi} &= \rho^{-1}e^{M/\rho}\sum_{p=N-M}^{N+M-2}a_p\int_{\frac{M}{\rho}}^{\infty}\left(x-\frac{M}{\rho}\right)^p\frac{e^{-x}}{x}dx\\ &= \rho^{-1}e^{M/\rho}\sum_{p=N-M}^{N+M-2}\left\{a_p\sum_{q=0}^{p}\binom{p}{q}\left(\frac{M}{\rho}\right)^{p-q}\right.\\ &\qquad\qquad\left.\times\int_{\frac{M}{\rho}}^{\infty}x^{q-1}e^{-x}dx\right\}\end{aligned}\tag{34}$$

In the following, we analyze the diversity gain for two cases.

1) $N = M = 1$: In this case, (34) equals

$$\bar{\xi} = \rho^{-1}e^{(\rho^{-1})}\int_{\rho^{-1}}^{\infty}\frac{e^{-x}}{x}\,dx = -\rho^{-1}e^{(\rho^{-1})}\mathrm{Ei}(-\rho^{-1}) \tag{35}$$

where the exponential integral is defined as [38] $\mathrm{Ei}(\varepsilon) \triangleq \int_{-\infty}^{\varepsilon}\frac{e^t}{t}dt = \gamma + \ln(-\varepsilon) + \sum_{k=1}^{\infty}\frac{\varepsilon^k}{k!k}$, with γ being Euler's constant $\gamma = 0.5772157\ldots$. Therefore, (35) can be written as

$$\bar{\xi} = e^{(\rho^{-1})}\rho^{-1}\left(\ln\rho - \gamma - \sum_{k=1}^{\infty}\frac{(-\rho^{-1})^k}{k!k}\right)$$

where the terms inside parentheses are ordered in descending order of ρ. At high SNR, we have

$$\bar{\xi}|_{\rho\to\infty} = \rho^{-1}\ln\rho. \tag{36}$$

Since $z = 1/(LM\bar{\xi})$, applying the upper bound on the Q-function in (8) to (30), and using (36), the upper bound on the minimum error probability is given by

$$\mathcal{P}_{\mathrm{mmin}}|_{\substack{L\to\infty\\ \rho\to\infty}} \leq \frac{1}{2e}\exp\left(-\frac{\rho}{2\ln\rho}\right) \tag{37}$$

It can be shown that $\lim_{\rho\to\infty}\frac{\exp(-\frac{\rho}{2\ln\rho})}{\rho^{-K}} = 0$ for any finite positive integer K, and hence we conclude that as L increases, the maximum diversity gain achievable by linear MMSE receiver grows without bound.

2) $\max(N-M, M-1) \geq 1$: Let $B_{pq} = a_pM^{p-q}\binom{p}{q}$. Then the term inside the braces in (34) can be integrated. Evaluating it at high SNR, we obtain the expression in (38), shown at the bottom of the page. Comparing (38) with (36), we observe that the terms (38) are of higher order of ρ^{-1} than that in (36). Therefore, the maximum diversity gain grows unboundedly with L in this case, too.

V. Generation and Detection of the Optimum Multiblock Code

A. Code Generation

For the case of single-block transmission (i.e., $L=1$), there are various ways of generating space–time code matrices that satisfy the optimality conditions in Lemma 1; cf. [16]. In this section we generalize the DFT-based method described therein in order to generate matrices that satisfy the optimality conditions for the multiblock case (cf. Theorem 2) for an arbitrary number of blocks, L.

The basic building block of a multiblock code is the $M \times T$ matrix $\boldsymbol{C}_{i\ell}$ given by (17), $i=1,2,\ldots,LMT, \ell=1,2,\ldots,L$. To construct these building blocks, the following simple procedure can be taken.

1. We first generate MT matrices $\boldsymbol{\Omega}_{mt}, m=1,2,\ldots,M, t=1,2,\ldots,T, T \geq M$, using the following steps.
   a) Form a $T \times T$ row permutation matrix such that
   $$\boldsymbol{P}=\begin{bmatrix} \mathbf{0} & 1 \\ \boldsymbol{I}_{T-1} & \mathbf{0} \end{bmatrix}$$
   where $\boldsymbol{I}_{T-1}$ is the $(T-1)\times(T-1)$ identity matrix.
   b) Form the $M \times M$ normalized DFT matrix $\boldsymbol{D}_M = [\boldsymbol{d}_M(1), \boldsymbol{d}_M(2), \cdots, \boldsymbol{d}_M(M)]$, with $\boldsymbol{d}_M(m)$ being the mth column.
   c) Generate the matrices $\boldsymbol{\Omega}_{mt}$ according to the following equation:
   $$\boldsymbol{\Omega}_{mt} = [\,\mathrm{Diag}(\boldsymbol{d}_M(m)) \quad \mathbf{0}\,]\,\boldsymbol{P}^{t-1}, \qquad m=1,2,\ldots,M, \quad t=1,2,\ldots,T$$
   where $\mathrm{Diag}(\boldsymbol{x})$ is the matrix obtained by putting the M elements of the vector $\boldsymbol{x}$ on the diagonal of an $M \times M$ diagonal matrix, and $\mathbf{0}$ is an $M \times (T-M)$ zero matrix. It can be easily verified that all the MT matrices $\boldsymbol{\Omega}_{mt}$ satisfy the unitarity and trace-orthogonality conditions stated in Theorem 1, and hence $\{\boldsymbol{\Omega}_{mt}\}$ forms an optimal single-block code.
2. Now, generate an $L \times L$ normalized DFT matrix and denote it by $\boldsymbol{D}_L$. The set of optimal building block code matrices can now be obtained by
   $$[\boldsymbol{C}_{i1}^T \quad \boldsymbol{C}_{i2}^T \quad \cdots \quad \boldsymbol{C}_{iL}^T]^T = \boldsymbol{d}_L(\ell) \otimes \boldsymbol{\Omega}_{mt}, \quad m=1,2,\ldots,M, t=1,2,\ldots,T, \ell=1,2,\ldots,L \tag{39}$$
   where $i=(M-1)TL+(t-1)L+\ell$ and $\boldsymbol{d}_L(\ell)$ denotes the ℓth column vector of the matrix $\boldsymbol{D}_L$. □

We now verify that the code matrix generated by the proposed algorithm satisfies the optimality conditions for a multiblock code in Theorem 2: Since $\boldsymbol{\Omega}_{mt}$ is a scaled unitary matrix and since each element of $\boldsymbol{D}_M$ has equal nonzero magnitude, it follows that $\boldsymbol{C}_{i\ell}$ is a scaled unitary matrix, and hence the second condition in Theorem 2 is satisfied. To simplify our verification of the first condition, we first map the double index mt in an arbitrary one-to-one fashion to the single index, r, where $r=1,2,\ldots,MT$. Now, for $i=(r_1-1)L+\ell_1$ and $j=(r_2-1)L+\ell_2$, we have

$$\begin{aligned}\sum_{\ell=1}^{L} \mathrm{tr}\left(\boldsymbol{C}_{i\ell}^H \boldsymbol{C}_{j\ell}\right) &= \mathrm{tr}\left(\left(\boldsymbol{d}_L(\ell_1)\otimes\boldsymbol{\Omega}_{r_1}\right)^H\left(\boldsymbol{d}_L(\ell_2)\otimes\boldsymbol{\Omega}_{r_2}\right)\right) \\ &= \mathrm{tr}\left(\left(\boldsymbol{d}_L(\ell_1)^H\boldsymbol{d}_L(\ell_2)\right)\otimes\left(\boldsymbol{\Omega}_{r_1}^H\boldsymbol{\Omega}_{r_2}\right)\right) \\ &= \delta_{\ell_1\ell_2}\cdot\mathrm{tr}(\boldsymbol{\Omega}_{r_1}^H\boldsymbol{\Omega}_{r_2}) = \delta_{\ell_1\ell_2}\delta_{r_1r_2} = \delta_{ij}\end{aligned}$$

which is the first condition in Theorem 2.

B. Detection of the Optimum Multiblock Code

For an L-block STBC, a total of LMT symbols have to be jointly processed at the receiver. Thus, the complexity of a given receiver is expected to be larger than that of the corresponding receiver for the single-block case. However, as we will show below, for codes with the optimal structure given in Theorem 2, the per-symbol computational cost of the linear MMSE receiver can be greatly reduced. Furthermore, for optimal codes constructed using the above technique, the computational cost can be further reduced.

The computational cost of a linear (multiblock) MMSE receiver is dominated by the cost of obtaining the equalized signal vector

$$\hat{\boldsymbol{s}}=\boldsymbol{G}y=\sqrt{\frac{\rho}{M}}\left(\boldsymbol{I}+\frac{\rho}{M}\boldsymbol{F}^H\boldsymbol{\mathcal{H}}^H\boldsymbol{\mathcal{H}}\boldsymbol{F}\right)^{-1}\boldsymbol{F}^H\boldsymbol{\mathcal{H}}^H\boldsymbol{\mathcal{H}}\boldsymbol{y}. \tag{40}$$

If, in accordance with Theorem 2, unitary trace-orthogonal signaling is employed, the special structure of the code matrix $\boldsymbol{F}$ can be used to reduce the cost of evaluating (40). First, the unitarity of $\boldsymbol{F}$ can be used to rewrite $\boldsymbol{G}$ as

$$\boldsymbol{G}=\sqrt{\frac{\rho}{M}}\boldsymbol{F}^H\mathrm{Diag}(\boldsymbol{\Psi}_1,\ldots,\boldsymbol{\Psi}_L) \tag{41}$$

$$B_{pq}\rho^{-(1+p-q)}e^{M/\rho}\int_{\frac{M}{\rho}}^{\infty} x^{q-1}e^{-x}dx = \begin{cases} B_{pq}\rho^{-(1+p-q)}\sum_{\eta=0}^{q-1}\left(\frac{M}{\rho}\right)^{\eta}\frac{(q-1)!}{\eta!}, & q \geq 1 \\ B_{pq}\rho^{-(1+p-q)}e^{M/\rho}\left(\ln\frac{\rho}{M}-\gamma-\sum_{k=1}^{\infty}\frac{(-M\rho^{-1})^k}{k!k}\right)\Big|_{\rho\to\infty} \approx B_{pq}\rho^{-(1+p-q)}\ln\rho, & q=0 \end{cases}. \tag{38}$$

TABLE I
NUMBER OF MULTIPLICATIONS REQUIRED TO DETECT EACH BLOCK OF MT SYMBOLS; $M = N = T = 2$; QPSK SIGNALING

Detector	Multiplications
LMMSE, $L = 1$	49
LMMSE, $L = 4$	69
LMMSE, $L = 32$	193
ML, single-block	5120

where our definition of $\mathrm{Diag}(\cdot)$ has been generalized to allow for matrix arguments, and

$$\boldsymbol{\Psi}_\ell = \boldsymbol{I}_T \otimes \left(\boldsymbol{I}_M + \frac{\rho}{M}\boldsymbol{H}_\ell^H \boldsymbol{H}_\ell\right)^{-1} \boldsymbol{H}_\ell^H, \quad \ell = 1, 2, \ldots, L. \tag{42}$$

Equation (41) is obtained by using the definition of $\boldsymbol{\mathcal{H}}$ in (15) together with the property of Kronecker product that $(\boldsymbol{C} \otimes \boldsymbol{D})^{-1} = \boldsymbol{C}^{-1} \otimes \boldsymbol{D}^{-1}$. We observe that the expression in (41) contains the inverse of L matrices of size $M \times M$, whereas the expression in (40) contains the inverse of a matrix of size $LMT \times LMT$. If we substitute (41) into (40), the equalized signal vector can be written as

$$\hat{\boldsymbol{s}} = \sqrt{\frac{\rho}{M}} \boldsymbol{F}^H \boldsymbol{\gamma} \tag{43}$$

where $\boldsymbol{\gamma}$ is the $LMT \times 1$ vector defined by $\boldsymbol{\gamma} = [(\boldsymbol{\Psi}_1 \boldsymbol{y}_1)^T, \ldots, (\boldsymbol{\Psi}_L \boldsymbol{y}_L)^T]^T$.

In our evaluation of the computational cost of computing $\hat{\boldsymbol{s}}$, we will focus on the number of multiplications required. Computing $\boldsymbol{\gamma}$ from $\boldsymbol{y}$ using the above expression, and expression for Ψ_ℓ in (42), requires LN_Π multiplications, where $N_\Pi = 8M^3/3 + 2NM^2 + NTM$ is the number of multiplications required to compute each $\boldsymbol{\Psi}_i \boldsymbol{y}_i$. Computing $\hat{\boldsymbol{s}}$ from $\boldsymbol{\gamma}$ using (43) requires a further $(LMT)^2$ multiplications. Therefore, the number of multiplications required to detect each of the L blocks of MT symbols is $N_\Pi + L(MT)^2$, which grows only linearly with L, with coefficient $(MT)^2$. That said, if one adopts the particular code structure proposed in Section V-A, the cost of computing $\hat{\boldsymbol{s}}$ can be substantially reduced. From the code construction algorithm in Section V-A, we observe that each column of matrix $\boldsymbol{F}$ has ML nonzero elements, each of which is generated from two DFT matrices, $\boldsymbol{D}_M$ and $\boldsymbol{D}_L$. As a result, each element of the equalized signal $\hat{\boldsymbol{s}}$ can be written as

$$\sum_{\ell=0}^{L-1} \sum_{m=0}^{M-1} W_M^{K_1 m} W_L^{K_2 \ell} \gamma_{m+M\ell} = \sum_{\ell=0}^{L-1} W_L^{K_2 \ell} \left(\sum_{m=0}^{M-1} W_M^{K_1 m} \gamma_{m+M\ell}\right) \tag{44}$$

where $K_1 \leq M-1$ and $K_2 \leq L-1$ are integers, and $W_K = \exp(-j2\pi/K)$. Equation (44) can be computed efficiently by a standard FFT algorithm [39], [40]. For example, if M and L are self-composite with base 2, then computing $\hat{\boldsymbol{s}}$ from $\boldsymbol{\gamma}$ requires $LMT(L\log_2 M + \log_2 L)$ multiplications. By exploiting the FFT algorithm in the computation of $\hat{\boldsymbol{s}}$ from $\boldsymbol{\gamma}$ using (43), the number of multiplications required to detect each of the L blocks of MT symbols is reduced to $N_\Pi + MT(L\log_2 M + \log_2 L)$. As L increases, the growth of this expression with L is dominated by the linear term, which has the coefficient $MT\log_2 M$. (Recall that the corresponding coefficient for a generic optimally structured code is $(MT)^2$.) For comparison, for a generic full-symbol-rate single-block transmission scheme with a constellation of cardinality μ, the ML detector requires (at most) $MT(MT+1)\mu^{MT}$ multiplications to detect each block (of MT symbols).[4]

To provide a concrete comparison between the computational cost of the multiblock linear MMSE receiver with the signaling scheme provided in Section V-A and the ML receiver for single block communication, we consider a MIMO system with $M = N = T = 2$ and a signal constellation of size $\mu = 4$; e.g., QPSK. The number of multiplications required to detect each block of MT symbols is provided in Table I. From this table, it can be seen that with the signaling scheme proposed in Section V-A, the computational cost per-block of multiblock signaling grows slowly with the number of blocks, L, and is much less than that of the ML detector for single-block transmission. A performance comparison of the systems in Table I will be provided in Example 3 in the ensuing section.

VI. PERFORMANCE COMPARISON

In this section, we compare the BER performance of the proposed multiblock code design with that of some conventional single-block STBCs.

Example 3: We consider the case of $M = N = 2$ antennas and 4-QAM signaling, which results in a transmission data rate of $R = 4$ bits pcu. In Fig. 5 we compare the average BER performance of the proposed multiblock scheme with linear MMSE reception to that of some conventional single-block MIMO transmission systems with ML detection, namely, the direct transmission scheme, in which no STBC is used (i.e., spatial multiplexing), and the scheme that transmits using the "Golden Code" [15], which possesses a constant minimum determinant and hence high coding gain. It is well known that the combination of the Golden code transmission and ML detection extracts the full diversity offered by the channel (i.e., $MN = 4$ in this example), and that direct transmission does not enable full diversity. We observe from Fig. 5 that at high SNR (above 19 dB), the proposed 8-block code with the linear MMSE receiver performs better than single-block direct transmission with ML detection. Now, from Fig. 4, we can predict that for $L \geq 20$, the diversity gain for the proposed multiblock scheme with linear reception will be higher than 4, which is the full diversity of the single-block system. This is indeed the case in Fig. 5. In particular, the performance curve for the multiblock scheme with $L = 22$ shows a slightly higher diversity gain, while the scheme with $L = 32$ shows a significantly higher diversity gain than that of the Golden code with ML detection. It is worth noting that these gains are achieved

[4]While this paper was under review, some specific single-block transmission schemes for which ML detection can be achieved using fewer operations have been developed (e.g., [41]), but the cost of such schemes remains exponential in μ.

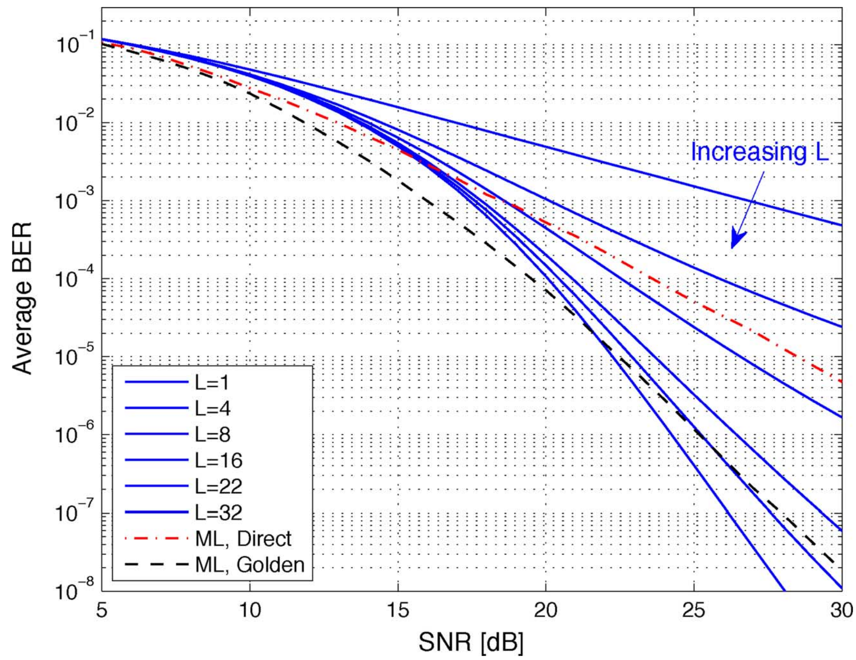

Fig. 5. Performance of the proposed multiblock transmission scheme (with a linear MMSE receiver) and the single-block direct transmission and Golden code schemes with ML detection for the case of $M = N = 2$.

at low computational cost, as shown in Table I in Section V-B. However, we should also point out that Fig. 5 demonstrates that at low SNRs, systems with ML detection can provide better performance than those with linear MMSE receivers. □

VII. Conclusion

In this paper, we have analyzed the maximum diversity gain associated with a MIMO system equipped with a linear MMSE receiver, for which an STBC of minimum detection error probability has been designed. The calculated diversity gain of $(N-M+1)$ has been verified by simulations. Indeed, the same diversity gain has previously been shown to be achievable in the simpler cases of direct transmission (i.e., spatial multiplexing) with linear ZF reception. Here, we showed that employing the optimum unitary trace-orthogonal code together with the MMSE receiver will not increase the diversity gain, but we also showed that this scheme achieves a lower BER due to its higher coding gain. To improve the performance of a MIMO system with linear reception, we considered multiblock transmission, in which the space–time code spans L realizations of a block fading channel. Such a scheme provides additional time diversity, and we designed an optimum STBC that minimizes the probability of error under (multiblock) linear MMSE reception. Our analysis showed that the diversity order of the proposed system grows with L, while the per-symbol detection complexity is little more than that for a single-block design with a linear receiver. Thus, if the latency of jointly detecting L signal blocks can be accommodated, the multiblock code design is an attractive alternative for increasing the diversity gain of a MIMO system at low computational cost.

Appendix A
Proof of Theorem 1

Substituting the pdf of the eigenvalues in (6) into the RHS of (9) and denoting it by $\eta(\rho)$, we have

$$\begin{aligned}\eta(\rho) = &\frac{\alpha(M,N)\sqrt{e}}{2}\int_0^\infty d\lambda_1 \int_0^{\lambda_1} d\lambda_2 \cdots \\ &\times \int_0^{\lambda_{M-2}} \exp\left(-\sum_{i=1}^{M-1}\lambda_i\right) \\ &\times \prod_{i=1}^{M-1}\lambda_i^{(N-M)}\prod_{i<j}^{M-1}(\lambda_i-\lambda_j)^2\, d\lambda_{M-1} \\ &\times \int_0^{\lambda_{M-1}} \exp\left[-\left(\frac{M/2}{\sum_{i=1}^{M}\left(1+\frac{\rho}{M}\lambda_i\right)^{-1}}+\lambda_M\right)\right] \\ &\qquad \times \lambda_M^{(N-M)}\prod_{i=1}^{M-1}(\lambda_i-\lambda_M)^2\, d\lambda_M. \qquad (45)\end{aligned}$$

The proof of the theorem requires successively integrating (45) by parts. In the following, we examine the result of each stage of integration by parts and select the nonzero term that has the lowest order of ρ^{-1}. To proceed, we first introduce the following function sequences to simplify the notation. For $1 \le m \le M$, we define

$$\mu_m = \sum_{i=1}^{m}\left(1+\frac{\rho}{M}\lambda_i\right)^{-1}, \quad \mu_0 = 0 \qquad (46a)$$

$$\begin{aligned}\phi_m = &\frac{M/2}{\mu_{m-1}+(M-m+1)\left(1+\frac{\rho}{M}\lambda_m\right)^{-1}} \\ &+ (M-m+1)\lambda_m \qquad (46b)\end{aligned}$$

$$\theta_m = \lambda_m^{(N-M)(M-m+1)}\prod_{i=1}^{m-1}(\lambda_i-\lambda_m)^{2(M-m+1)} \qquad (46c)$$

with $\theta_1 = \lambda_1^{(N-M)M}$. From (46b), we have $\phi'_m(\lambda_m) = \frac{\rho(M-m+1)/2}{(\mu_{m-1}(1+\frac{\rho}{M}\lambda_m)+(M-m+1))^2} + (M-m+1)$, where $(\cdot)'$ denotes the first derivative w.r.t. λ_m. Since ϕ'_m is also a function of other eigenvalues $\lambda_i, i \le m$, we define

$$\psi_{mi}(\lambda_i) = \frac{\rho(M-m+1)/2}{\left(\mu_{i-1}\left(1+\frac{\rho}{M}\lambda_i\right)+(M-i+1)\right)^2} + (M-m+1), \quad \text{for } i \le m. \tag{47}$$

Since the lower limit of integration is $\lambda_m = 0$, we examine the values of ϕ_m and ϕ'_m at this limit. From (46b), we have $\phi_m|_0 = (\frac{M/2}{\mu_{m-1}+(M-m+1)})$. Now, the eigenvalues are ordered by strict inequality such that $\lambda_1 > \lambda_2 > \cdots > \lambda_M$. Therefore, for $i < m$, λ_i is finite and at high SNR, $(1+\frac{\rho}{M}\lambda_i)^{-1}$ is negligible with respect to 1. Hence, $\mu_{m-1}|_{\lambda_m=0} \approx 0$ and therefore we have $\phi_m|_0 = (\frac{M/2}{M-m+1})$. Similarly, $\phi'_m|_{\lambda_m=0} \approx (\frac{\rho/2}{M-m+1} + (M-m+1)) \approx \frac{\rho}{2(M-m+1)}$ for a finite number of transmitter antennas.

To evaluate (45), we use "integration by parts" such that for general functions Φ and Θ of λ

$$\int_{\Lambda_1}^{\Lambda_2} e^{-\Phi}\Theta d\lambda \triangleq A + I \tag{48}$$

where $A \triangleq -\frac{e^{-\Phi}}{\Phi'}\Theta|_{\Lambda_1}^{\Lambda_2}$ and $I \triangleq \int_{\Lambda_1}^{\Lambda_2} e^{-\Phi}\frac{\partial}{\partial\lambda}(\frac{1}{\Phi'}\Theta)\, d\lambda$. The proof of the theorem now follows the procedure as outlined below.

a) $N-M=0$: Let I_{M1} denote the innermost integral in (45). Integrating I_{M1} by parts we have

$$\begin{aligned} I_{M1} &= \int_0^{\lambda_{M-1}} \exp\left(-\phi_M(\lambda_M)\right)\theta_M(\lambda_M)\, d\lambda_M \\ &= -\left.\frac{\exp(-\phi_M)}{\phi'_M(\lambda_M)\theta_M}\right|_0^{\lambda_{M-1}} \\ &\quad + \int_0^{\lambda_{M-1}} \exp\left(-\phi_M(\lambda_M)\right) \\ &\quad\quad \times \frac{\partial}{\partial\lambda_M}\left(\frac{1}{\phi'_M(\lambda_M)}\theta_M(\lambda_M)\right) d\lambda_M \\ &\triangleq A_{M1} + I_{M2} \end{aligned} \tag{49}$$

where $A_{M1} = -\frac{\exp(-\phi_M)}{\phi'_M(\lambda_M)}\theta_M|_0^{\lambda_{M-1}}$ and $I_{M2} = \int_0^{\lambda_{M-1}} \exp(-\phi_M(\lambda_M))\frac{\partial}{\partial\lambda_M}(\frac{1}{\phi'_M(\lambda_M)}\theta_M(\lambda_M))\, d\lambda_M$. Now, the value of A_{M1} is the difference of the two terms obtained by substituting the upper and lower limits. For the upper limit $\lambda_M = \lambda_{M-1}, \theta_M$ always contains the factor $(\lambda_{M-1}-\lambda_{M-1})$ and is therefore zero. For the lower limit $\lambda_M = 0, \theta_M$ contains the factor $\lambda_M^{(N-M)}|_{\lambda_M=0}$. If $N-M=0$, since $0^0 = 1$, then we have $\theta_M|_{\lambda_M=0} = \prod_{i=1}^{M-1}\lambda_i^2$. Substituting $e^{-\phi_M}|_{\lambda_M=0}$ and $\phi'_M(\lambda_M)|_{\lambda_M=0}$, we have $A_{M1} = \frac{2}{\rho}\exp(-\frac{M}{2})\prod_{i=1}^{M-1}\lambda_i^2$. Further integration w.r.t. $\lambda_{M-1},\ldots,\lambda_1$ will not increase the order of ρ^{-1} in this term and thus this contains the lowest order of ρ^{-1}. If $(N-M) \neq 0$, then $A_{M1} = 0$ and we have to examine I_{M2}.

b) $(N-M) > 0$: In this case, for (49), $A_{M1} = 0$ and we examine I_{M2}. Integrating I_{M2} by parts

$$\begin{aligned} I_{M2} &= \int_0^{\lambda_{M-1}} \exp(-\phi_M)\frac{\partial}{\partial\lambda_M}\left(\frac{1}{\phi'_M(\lambda_M)}\theta_M\right) d\lambda_M \\ &= A_{M2} + I_{M3} \end{aligned} \tag{50}$$

where

$$\begin{aligned} A_{M2} &= \frac{-e^{-\phi_M}}{\phi'_M}\frac{\partial}{\partial\lambda_M}\left.\left(\frac{1}{\phi'_M}\theta_M\right)\right|_0^{\lambda_{M-1}} \\ &= \frac{-e^{-\phi_M}}{\phi'_M}\left[-\frac{\phi''_M}{(\phi'_M)^2}\theta_M + \frac{1}{\phi'_M}\theta'_M\right]_0^{\lambda_{M-1}} \\ &= \frac{-e^{-\phi_M}}{(\phi'_M)^2}\left.\theta'_M\right|_0^{\lambda_{M-1}}, \end{aligned} \tag{51a}$$

$$I_{M3} = \int_0^{\lambda_{M-1}} e^{-\phi_M}\frac{\partial}{\partial\lambda_M}\left(\frac{1}{\phi'_M}\frac{\partial}{\partial\lambda_M}\left(\frac{1}{\phi'_M}\theta_M\right)\right) d\lambda_M. \tag{51b}$$

The last step of (51a) is obtained by observing that the first term is zero for both limits, since $\theta_M|_0^{\lambda_M-1} = 0$. Now

$$\theta'_M|_0^{\lambda_{M-1}} = -(N-M)\left.\lambda_M^{(N-M-1)}\right|_0 \prod_{i=1}^{M-1}\lambda_i^2 \tag{52}$$

where this result is due to the fact that the other terms of the differential all vanish after the limits are substituted, i.e., $\theta'_M|_0^{\lambda_{M-1}}$ is only nonzero in the limit $\lambda_M \to 0$. Thus, using (52) in (51a), we have

$$A_{M2} = \left(\frac{2}{\rho}\right)^2 \exp\left(-\frac{M}{2}\right) \times (N-M)\left(\left.\lambda_M^{(N-M-1)}\right|_0 \prod_{i=1}^{M-1}\lambda_i^2\right) \tag{53}$$

which contains $(\rho^{-1})^2$. Similar to the case of A_{M1}, if $N-M-1=0$, then A_{M2} contains the lowest order of ρ^{-1}. Further integration w.r.t. $\lambda_1,\ldots,\lambda_{M-1}$ will not change the order of ρ^{-1}. However, if $N-M-1>0$, then $A_{M2} = 0$ and we have to integrate I_{M3} in (51b) by parts, i.e.,

$$I_{M3} = A_{M3} + I_{M4}$$

where

$$\begin{aligned} A_{M3} &= -\frac{e^{-\phi_M}}{\phi'_M}\frac{\partial}{\partial\lambda_M}\left.\left[\frac{1}{\phi'_M}\frac{\partial}{\partial\lambda_M}\left(\frac{1}{\phi'_M}\theta_M\right)\right]\right|_0^{\lambda_{M-1}} \\ &= \frac{e^{-\phi_M}}{(\phi'_M)^3}\theta''_M|^0_{\lambda_{M-1}}, \end{aligned} \tag{54a}$$

$$I_{M4} = \int_0^{\lambda_{M-1}} e^{-\phi_M} \times \frac{\partial}{\partial\lambda_M}\left(\frac{1}{\phi'_M}\frac{\partial}{\partial\lambda_M}\left(\frac{1}{\phi'_M}\frac{\partial}{\partial\lambda_M}\left(\frac{1}{\phi'_M}\theta_M\right)\right)\right) d\lambda_M \tag{54b}$$

where (54a) is obtained by observing that both θ_M and θ'_M are equal to zero at the upper and lower limits under the condition $N-M-1>0$. Now

$$\theta''_M|^0_{\lambda_{M-1}} = (N-M)(N-M-1)\,\lambda_M^{(N-M-2)}\Big|_0 \prod_{i=1}^{M-1}\lambda_i^2 + 2\lambda_{M-1}^{(N-M)}\prod_{i=1}^{M-2}(\lambda_i-\lambda_{M-1})^2. \quad (55)$$

Equation (55) is the result after the limits have been substituted into the second derivative. Here, $\theta''_M|^0_{\lambda_{M-1}}$ consists of two terms. If $N-M-2=0$, then the first term is independent of ρ. From (54a), the factor $(e^{-\phi_M(0)}/(\phi'_M(0))^3)$, which contains the factor $(\rho^{-1})^3$, is independent of λ_{M-1} or any other eigenvalues and the order of ρ^{-1} will not increase upon further integration w.r.t. $\lambda_{M-1},\ldots,\lambda_1$. Hence, this is the lowest order of ρ^{-1}. On the other hand, if $N-M-2>0$, the first term in (55) is zero. Now, we can obtain the lowest order term in two different ways.

i) Consider I_{M4} of (54b) and perform integration by parts. Here, we focus on reducing the power index of λ_M^{N-M-2} to zero so that the integration has nonzero result at $\lambda_M=0$. The order of ρ^{-1} will not increase upon further integration w.r.t. the other eigenvalues.

ii) When the upper limit $\lambda_M=\lambda_{M-1}$ is substituted, the second term in (55) is nonzero. Putting this term back in (54a), together with the factor $[e^{-\phi_M}/(\phi'_M)^3]|_{\lambda_M=\lambda_{M-1}}$ which is a function of ρ and λ_{M-1}, this whole quantity of A_{M3} has to be further integrated w.r.t. $\lambda_{M-1},\ldots,\lambda_1$, and therefore increasing the order of ρ^{-1}.

Thus, the problem of seeking the lowest order of ρ^{-1} has been reduced to the following questions. Is it the terms in Step i) for which, in integrating I_{M4} by parts, the index of $\lambda_M|_0$ is reduced to zero after differentiation w.r.t. λ_M? Or is it the term in Step ii) which yields the term involving the lowest order of ρ^{-1} after all the integrations w.r.t. $\lambda_{M-1},\ldots,\lambda_1$? We will examine the questions in parts c) and d) below.

c) The order of ρ^{-1} by Step i): The first integration of I_{M1} in (49) yields $(\phi'_M)^{-1}$ together with θ_M which involves no reduction in the index. The second integration I_{M2} yields $(\phi'_M)^{-2}$ together with θ'_M and reduces the index by 1. Thus, to reduce the index of $\lambda_M|_0$ from $(N-M)$ to zero, we need to differentiate the function θ_M $(N-M)$ times. This yields $(\phi'_M)^{-(N-M+1)}$, which contains the SNR term $\rho^{-(N-M+1)}$. The nonzero component of the integration is $((N-M)!\,(2/\rho)^{N-M+1}e^{-\frac{M}{2}}\prod_{i=1}^{M-1}\lambda_i^2)$, which is independent of λ_M. Substituting this nonzero term into (45), we obtain the term in $\eta(\rho)$ containing the lowest order of ρ^{-1}, namely

$$A_0=(N-M)!\frac{\alpha(M,N)\exp\left(-\frac{M-1}{2}\right)}{2}\left(\frac{2}{\rho}\right)^{N-M+1} \times\int_0^\infty d\lambda_1\int_0^{\lambda_1}d\lambda_2\cdots\int_0^{\lambda_{M-2}}\exp\left(-\sum_{i=1}^{M-1}\lambda_i\right) \times\prod_{i=1}^{M-1}\lambda_i^{N-M+2}\prod_{i<j}(\lambda_i-\lambda_j)^2\,d\lambda_{M-1} \quad (56)$$

where $\alpha(M,N)$ was defined in (6). Observe that the integration in (56) will not increase the order of ρ^{-1}. Now, multiplying both the numerator and denominator of (56) by $\alpha(M-1,N+1)$, we can write

$$A_0=(N-M)!\frac{\alpha(M,N)\exp\left(-\frac{M-1}{2}\right)}{2\alpha(M-1,N+1)}\left(\frac{2}{\rho}\right)^{(N-M+1)} \times\left\{\alpha(M-1,N+1)\int_0^\infty d\lambda_1\int_0^{\lambda_1}d\lambda_2\cdots \times\int_0^{\lambda_{M-2}}\exp\left(-\sum_{i=1}^{M-1}\lambda_i\right)\prod_{i=1}^{M-1}\lambda_i^{N-M+2} \times\prod_{i<j}(\lambda_i-\lambda_j)^2\,d\lambda_{M-1}\right\} \quad (57a)$$

$$=(N-M)!\frac{\alpha(M,N)\exp\left(-\frac{M-1}{2}\right)2^{N-M}}{\alpha(M-1,N+1)}\rho^{-(N-M+1)}. \quad (57b)$$

The last step in (57b) is obtained by observing that the integrand inside the braces of (57a) is the same pdf as (6), and thus its integration results in unity. From (57b), we observe that for high SNR the order of ρ^{-1} is given by

$$d_1=N-M+1 \quad (58)$$

This is thus the lowest order of ρ^{-1} contained in the term obtained when the lower limit $\lambda_M=0$ is substituted.

d) The order of ρ^{-1} obtained by Step ii): Putting the second term of (55) into (54a) and then the result into (45), and taking the terms involving $\lambda_1,\ldots,\lambda_{M-2}$ outside, the integral w.r.t. λ_{M-1} is

$$I_{(M-1)1}=2\int_0^{\lambda_{M-2}}\frac{e^{-\phi_{M-1}}}{(\psi_{M(M-1)}(\lambda_{M-1}))^3}\,\theta_{M-1}\,d\lambda_{M-1} \quad (59)$$

where ϕ_{M-1} and θ_{M-1} have been defined in a general form in (46b) and (46c), respectively. We note that the indices of the factors λ_{M-1} and $(\lambda_i-\lambda_{M-1})$ in θ_{M-1} have increased by one extra fold and by 2, respectively, in comparison to the corresponding indices in θ_M. Equation (59) has a similar structure to (49), and hence integrating (59) by parts follows a similar pattern, i.e.,

$$I_{(M-1)1}=2(A_{(M-1)1}+I_{(M-1)2}) \quad (60)$$

where $A_{(M-1)1}=\frac{-\exp(-\phi_{M-1})}{(\psi_{M(M-1)}(\lambda_{M-1}))^3\phi'_{M-1}}\,\theta_{M-1}|_0^{\lambda_{M-1}}$ and $I_{(M-1)2}=\int_0^{\lambda_{M-2}}e^{-\phi_{M-1}}\left(\frac{\theta_{M-1}}{(\psi_{M(M-1)}(\lambda_{M-1}))^3\phi'_{M-1}}\right)'d\lambda_{M-1}$. Due to the property of $\theta_{M-1}|_0^{\lambda_{M-2}}=0$, we only have to examine the effect of $I_{(M-1)2}$ on the index of ρ^{-1}. Again, applying integration by parts of (48) to $I_{(M-1)2}$, we have

$$I_{(M-1)2}=A_{(M-1)2}+I_{(M-1)3} \quad (61)$$

where

$$A_{(M-1)2} = -\frac{e^{-\phi_{M-1}}}{\phi'_{M-1}}\left(\frac{1}{(\psi_{M(M-1)}(\lambda_{M-1}))^3\phi'_{M-1}}\theta_{M-1}\right)'\Bigg|_0^{\lambda_{M-2}}$$

and

$$I_{(M-1)3} = \int_0^{\lambda_{M-2}} e^{-\phi_{M-1}}\left(\frac{1}{\phi'_{M-1}}\left(\frac{1}{(\psi_{M(M-1)}(\lambda_{M-1}))^3\phi'_{M-1}} \times \theta_{M-1}\right)'\right)' d\lambda_{M-1}.$$

The term $A_{(M-1)2}$ is zero since θ'_{M-1} contains the term $4\lambda_{M-1}^{2(N-M)}(\lambda_{M-2}-\lambda_{M-1})^3$ which yields zero at either limit of 0 or λ_{M-2}. A nonzero result is obtained from $I_{(M-1)3}$ at the limit of $\lambda_{M-1}=\lambda_{M-2}$ by having three more steps of integration by parts on $I_{(M-1)3}$, each involving the differentiation of $(\lambda_{M-2}-\lambda_{M-1})^3$ and resulting in the power index being reduced to zero. This nonzero term is given by

$$\frac{2\cdot 4!\, e^{-\phi_{M-2}}\theta_{M-2}}{(\psi_{M(M-2)}(\lambda_{M-2}))^3(\psi_{(M-1)(M-2)}(\lambda_{M-2}))^5} \tag{62}$$

Further integration of (62) w.r.t. $\lambda_{M-2},\ldots,\lambda_1$ repeats the same procedure as in (59) and reveals similar structures, i.e., each time the factors in the denominator increases by $(\phi'_m)^{2(M-m+1)+1}$ and the subscript of θ reduces by 1. Continuing the procedure until the integration w.r.t. λ_1, we have (63)

$$I_1 = \int_0^{\infty}\frac{\left(\prod_{i=1}^{M-1}(2i)!\right)e^{-\phi_1(\lambda_1)}}{(\psi_{M1}(\lambda_1))^3(\psi_{(M-1)1}(\lambda_1))^5\cdots\psi_{21}(\lambda_1)^{2(M-1)+1}}\theta_1\, d\lambda_1 \tag{63}$$

where, from (46b) and (46c), $\phi_1(\lambda_1) = (\frac{1}{2} + (\frac{\rho}{2M}+M)\lambda_1)$ and $\theta_1 = \lambda_1^{(N-M)M}$. We note that $e^{-\phi_1(\lambda_1)}|_{\lambda_1=\infty}=0$, and the nonzero result of integrating I_1 comes only from the lower limit $\lambda_1=0$. Again, this term is arrived at by repeated integration by parts until the index of the term $\lambda_1^{(N-M)M}$ in θ_1 reaches 0. The resulting nonzero term with minimum steps of integration by parts is given in (64), at the bottom of the page. Now, we can evaluate the power index of ρ^{-1} in (64). We observe that each ψ or ϕ' contains ρ (unity power index). Thus, ignoring the constant coefficients, the power index of ρ^{-1} in A_1 is the sum of the exponents of the functions ψ and ϕ' in the denominator of (64), i.e.,

$$\sum_{i=1}^{M-1}(2i+1)+(N-M)M+1 = MN \triangleq d_2. \tag{65}$$

Equation (65) provides the order of ρ^{-1} obtained by continuing the integration of $I_{(M-1)1}$ w.r.t. $\lambda_{M-1},\ldots,\lambda_1$.

e) We now compare the two lowest orders d_1 and d_2 established respectively in (58) and (65). We note that $d_2-d_1 = MN-(N-M+1)=(N+1)(M-1)\geq 0$ since $N,M>0$. Therefore, d_1 is the lowest order of ρ^{-1} in the result of evaluating $\eta(\rho)$ of (45), and is therefore the diversity order for the single-block MIMO system. ■

Appendix B
Proof of Theorem 2

We can prove Theorem 2 by first deriving several consecutively achievable lower bounds on the error probability given in (22) until we reach a minimum independent of the design parameters. This is followed by examining the conditions required on the optimal $\boldsymbol{F}$ to obtain each of the lower bounds.

1) First Lower Bound: Using the convexity property of the Q-function [16] in (22), we obtain the first lower bound on $\mathcal{P}_{\text{em}}$ by employing Jensen's inequality [37], i.e.,

$$\mathcal{P}_{\text{em}} = \mathrm{E}_{\boldsymbol{\mathcal{H}}}\left\{\frac{1}{2LMT}\sum_{k=1}^{2LMT}Q\left(\sqrt{\frac{1}{[\boldsymbol{V}]_{kk}}-1}\right)\right\}$$
$$\geq \mathrm{E}\left\{Q\left(\sqrt{\frac{2LMT}{\sum_{k=1}^{2LMT}[\boldsymbol{V}]_{kk}}-1}\right)\right\} \tag{66a}$$
$$= \mathrm{E}\left\{Q\left(\sqrt{\frac{LMT}{\mathrm{tr}(\boldsymbol{V}_{\text{se}})}-1}\right)\right\} \tag{66b}$$

where $\boldsymbol{V}_{\text{se}}$ is the symbol error covariance matrix given by (20). Equation (66b) holds due to the fact that $\mathrm{tr}(\boldsymbol{V}) = \mathrm{tr}(2\boldsymbol{V}_{\text{se}})$. Equality in (66a) holds iff the diagonal elements of $\boldsymbol{V}$ are all equal, i.e., if the MSEs of all the bits are equal. Since we are transmitting 4-QAM symbols having symmetry in both real and imaginary parts, this is equivalent to

$$[\boldsymbol{V}_{\text{se}}]_{ii} = [\boldsymbol{V}_{\text{se}}]_{jj}, \quad \forall\, i,j = 1,2,\ldots,LMT. \tag{67}$$

2) Second Lower Bound: The following two lemmas [42] are provided here to facilitate the development of the second lower bound.

$$A_1 = \frac{((N-M)M)!\prod_{i=1}^{M-1}(2i)!\, e^{-1/2}}{(\psi_{M1}(\lambda_1))^3(\psi_{(M-1)1}(\lambda_1))^5\cdots\psi_{21}(\lambda_1)^{(2(M-1)+1)}\phi'_1(\lambda_1)^{((N-M)M+1)}}\Bigg|_{\lambda_1=0} \tag{64}$$

Lemma 2: For any square matrix $\mathbf{\Gamma}$ we have $\operatorname{tr}((\mathbf{I}+\mathbf{\Gamma}\mathbf{\Gamma}^H)^{-1}) = \operatorname{tr}((\mathbf{I}+\mathbf{\Gamma}^H\mathbf{\Gamma})^{-1})$. □

Lemma 3: For any nonsingular Hermitian symmetric positive semidefinite (PSD) matrix $\boldsymbol{Z} = \begin{bmatrix} \boldsymbol{Z}_{11} & \boldsymbol{Z}_{12} \\ \boldsymbol{Z}_{12}^H & \boldsymbol{Z}_{22} \end{bmatrix}$ we have $\operatorname{tr}(\boldsymbol{Z}^{-1}) \geq \operatorname{tr}(\boldsymbol{Z}_{11}^{-1}) + \operatorname{tr}(\boldsymbol{Z}_{22}^{-1})$, where equality holds iff $\boldsymbol{Z}_{12} = \mathbf{0}$, i.e., iff $\boldsymbol{Z}$ is block diagonal. □

Lemma 3 can be repeatedly applied to generalize the result to a nonsingular Hermitian symmetric PSD matrix partitioned into multiple submatrices.

Applying Lemma 2 to the error covariance matrix $\boldsymbol{V}_{\mathrm{se}}$ in (20) yields

$$\operatorname{tr}(\boldsymbol{V}_{\mathrm{se}}) = \operatorname{tr}\left(\left(\boldsymbol{I} + \frac{\rho}{M}\mathbf{\Xi}\boldsymbol{F}\boldsymbol{F}^H\mathbf{\Xi}^H\right)^{-1}\right) \tag{68}$$

where we define $\mathbf{\Xi} \triangleq (\boldsymbol{\mathcal{H}}^H\boldsymbol{\mathcal{H}})^{\frac{1}{2}}$. From the definition of $\boldsymbol{\mathcal{H}}$ in (15), we can see that $\mathbf{\Xi}$ is an $LMT \times LMT$ block diagonal matrix. Each diagonal block $\tilde{\mathbf{\Xi}}_\ell \triangleq \mathbf{I}_T \otimes (\boldsymbol{H}_\ell^H\boldsymbol{H}_\ell)^{\frac{1}{2}}$, $\ell = 1, \ldots, L$, is also block diagonal of dimension $MT \times MT$, containing T identical $M \times M$ subblocks of $\mathbf{\Xi}_\ell \triangleq (\boldsymbol{H}_\ell^H\boldsymbol{H}_\ell)^{\frac{1}{2}}$.

Now, writing $\boldsymbol{B} = \boldsymbol{F}\boldsymbol{F}^H$ in (68), we note that this $LMT \times LMT$ nonsingular Hermitian symmetric PSD matrix can be partitioned into L^2 blocks of $MT \times MT$ submatrices $\boldsymbol{B}_{ij}, i, j = 1, \ldots, MT$. Denoting by $\boldsymbol{B}_{\ell\ell}$ the ℓth $MT \times MT$ submatrix on the diagonal of $\boldsymbol{B}$ and applying Lemma 3, a lower bound for $\operatorname{tr}(\boldsymbol{V}_{\mathrm{se}})$ in (68) can be obtained, namely

$$\operatorname{tr}(\boldsymbol{V}_{\mathrm{se}}) \geq \sum_{\ell=1}^{L} \operatorname{tr}\left(\left(\boldsymbol{I}_{MT} + \frac{\rho}{M}\tilde{\mathbf{\Xi}}_\ell\boldsymbol{B}_{\ell\ell}\tilde{\mathbf{\Xi}}_\ell^H\right)^{-1}\right). \tag{69}$$

Equality in (69) holds iff $\boldsymbol{B}_{ij} = \mathbf{0}$, for $i, j = 1, \ldots, L, i \neq j$. But each $\boldsymbol{B}_{\ell\ell}$ can be further partitioned into T^2 submatrices $\boldsymbol{B}_{\ell\ell}^{ij}, i, j = 1, \ldots, T$, of dimension $M \times M$. Denoting by $\boldsymbol{B}_{\ell\ell}^{tt}$ the tth $M \times M$ submatrix on the diagonal of $\boldsymbol{B}_{\ell\ell}$ and applying Lemma 3 to each of the bracketed term in (69) results in

$$\operatorname{tr}(\boldsymbol{V}_{\mathrm{se}}) \geq \sum_{\ell=1}^{L}\sum_{t=1}^{T} \operatorname{tr}\left(\left(\boldsymbol{I}_M + \frac{\rho}{M}\mathbf{\Xi}_\ell\boldsymbol{B}_{\ell\ell}^{tt}\mathbf{\Xi}_\ell^H\right)^{-1}\right) \tag{70}$$

where equality in (70) holds iff

$$\boldsymbol{B}_{\ell\ell}^{ij} = \mathbf{0} \quad \text{for } i, j = 1, \ldots, T, \quad i \neq j \tag{71}$$

i.e., $\boldsymbol{B} = \boldsymbol{F}\boldsymbol{F}^H$ is block diagonal having nonzero $M \times M$ diagonal submatrices.

Furthermore, for any positive definite matrix $\boldsymbol{Z}$, $\operatorname{tr}(\boldsymbol{Z}^{-1})$ is convex with respect to $\boldsymbol{Z}$ [37]. Thus, applying Jensen's inequality to this convex function in the inner sum of (70) and employing Lemma 2 leads to

$$\frac{1}{T}\sum_{t=1}^{T} \operatorname{tr}\left(\left(\boldsymbol{I}_M + \frac{\rho}{M}\mathbf{\Xi}_\ell\boldsymbol{B}_{\ell\ell}^{tt}\mathbf{\Xi}_\ell^H\right)^{-1}\right) \geq \operatorname{tr}\left(\left(\boldsymbol{I}_M + \frac{\rho}{M}\mathbf{\Xi}_\ell\bar{\boldsymbol{B}}_{\ell\ell}\mathbf{\Xi}_\ell^H\right)^{-1}\right) \tag{72}$$

where $\bar{\boldsymbol{B}}_{\ell\ell} \triangleq \frac{1}{T}\sum_{t=1}^{T}\boldsymbol{B}_{\ell\ell}^{tt}$. Equality in (72) holds iff all $\boldsymbol{B}_{\ell\ell}^{tt}$ are equal, i.e.,

$$\boldsymbol{B}_{\ell\ell}^{tt} = \bar{\boldsymbol{B}}_{\ell\ell} \quad \text{for } t = 1, \ldots, T. \tag{73}$$

Combining (70) with (72) results in

$$\operatorname{tr}(\boldsymbol{V}_{\mathrm{se}}) \geq T\sum_{\ell=1}^{L} \operatorname{tr}\left(\left(\boldsymbol{I}_M + \frac{\rho}{M}\mathbf{\Xi}_\ell\bar{\boldsymbol{B}}_{\ell\ell}\mathbf{\Xi}_\ell^H\right)^{-1}\right) \tag{74}$$

with equality holding iff (71) and (73) are satisfied. Since the function $f(x) = Q(\sqrt{x^{-1} - 1})$ is monotonically increasing with x, substituting (74) in (66b) establishes the second lower bound for the average asymptotic BER $\mathcal{P}_{\mathrm{em}}$ such that

$$\mathcal{P}_{\mathrm{em}} \geq \mathrm{E}\left\{Q\left(\sqrt{\frac{ML}{\sum_{\ell=1}^{L}\operatorname{tr}\left((\boldsymbol{I}_M + \frac{\rho}{M}\mathbf{\Xi}_\ell\bar{\boldsymbol{B}}_{\ell\ell}\mathbf{\Xi}_\ell^H)^{-1}\right)} - 1}\right)\right\}. \tag{75}$$

Again, equality in (75) holds iff the conditions in (71) and (73) are met simultaneously.

3) Final Lower Bound: Equation (75) still depends on the design variable $\boldsymbol{F}$ and therefore needs to be further minimized to obtain a constant achievable lower bound. We observe that the elements of the channel matrix $\boldsymbol{H}_\ell$ are of i.i.d. distribution, hence the stochastic properties of $\boldsymbol{H}_\ell$ remain unchanged by the multiplication of a unitary matrix. By pre- and postmultiplying permutation matrices to the eigenvalue matrix of $\bar{\boldsymbol{B}}_{\ell\ell}$ in (75) and following similar procedures of averaging over all $M!$ permutations, as in [3], [16], we arrive at

$$\mathcal{P}_{\mathrm{em}} \geq \mathrm{E}\left\{Q\left(\sqrt{\frac{ML}{\sum_{\ell=1}^{L}\operatorname{tr}\left((\boldsymbol{I}_M + \frac{\rho}{M}\boldsymbol{H}_\ell^H\boldsymbol{H}_\ell)^{-1}\right)} - 1}\right)\right\} \tag{76}$$

with equality holding iff

$$\bar{\boldsymbol{B}}_{\ell\ell} = \boldsymbol{I}, \quad \ell = 1, \ldots, L. \tag{77}$$

The lower bound in (76) is independent of the design matrix $\boldsymbol{F}$ and is therefore a genuine lower bound on the value that can be achieved by a designer, in the sense that it can be computed prior to the design.

4) Optimal Linear STBC: The minimum BER given in (76) can be achieved iff the four conditions in (67), (71), (73), and (77) are satisfied simultaneously. Equations (71), (73), and (77) jointly imply that $\boldsymbol{B} = \boldsymbol{F}\boldsymbol{F}^H = \boldsymbol{I}$, i.e., $\boldsymbol{F}$ must be unitary. Equation (67) requires all the diagonal elements of $\boldsymbol{V}_{\mathrm{se}}$ to be equal, i.e.,

$$[\boldsymbol{V}_{\mathrm{se}}]_{ii} = \frac{1}{LMT}\sum_{i=1}^{LMT}[\boldsymbol{V}_{\mathrm{se}}]_{ii} = \frac{1}{LMT}\operatorname{tr}(\boldsymbol{V}_{\mathrm{se}}) \tag{78}$$

and this must hold for any channel realization.

Now, (20) can be rewritten as $\boldsymbol{V}_{\mathrm{se}} = (\boldsymbol{F}^H\boldsymbol{F} + \frac{\rho}{M}\boldsymbol{F}^H\boldsymbol{\mathcal{H}}^H\boldsymbol{\mathcal{H}}\boldsymbol{F})^{-1} = \boldsymbol{F}^H(\boldsymbol{I} + \frac{\rho}{M}\boldsymbol{\mathcal{H}}^H\boldsymbol{\mathcal{H}})^{-1}\boldsymbol{F}$, where we have used the condition that $\boldsymbol{F}$ must be unitary. Thus, the ith diagonal element of $\boldsymbol{V}_{\rm se}$ is

$$\begin{aligned}[\boldsymbol{V}_{\rm se}]_{ii} &= \boldsymbol{f}_i^H \left(\boldsymbol{I}_{LMT} + \frac{\rho}{M}\boldsymbol{\mathcal{H}}^H\boldsymbol{\mathcal{H}}\right)^{-1} \boldsymbol{f}_i \\ &= \operatorname{tr}\left(\sum_{\ell=1}^{L} \boldsymbol{C}_{i\ell}^H \left(\boldsymbol{I}_M + \frac{\rho}{M}\boldsymbol{H}_\ell^H\boldsymbol{H}_\ell\right)^{-1} \boldsymbol{C}_{i\ell}\right) \\ &= \operatorname{tr}\left(\sum_{\ell=1}^{L} \left(\boldsymbol{I}_M + \frac{\rho}{M}\boldsymbol{H}_\ell^H\boldsymbol{H}_\ell\right)^{-1} \boldsymbol{C}_{i\ell}\boldsymbol{C}_{i\ell}^H\right) \qquad (79)\end{aligned}$$

where $\boldsymbol{f}_i$ is the ith column of the coding matrix $\boldsymbol{F}$ and $\boldsymbol{C}_{i\ell}$ is the coding submatrix in the ℓth block for the ith symbol. The second step in (79) is possible because of the relationship between $\boldsymbol{f}_i$ and $\boldsymbol{C}_{i\ell}$ in (17). On the other hand, the right hand side of (78) can also be written as

$$\begin{aligned}\operatorname{tr}(\boldsymbol{V}_{\rm se}) &= \operatorname{tr}\left(\boldsymbol{I}_{LMT} + \frac{\rho}{M}\boldsymbol{\mathcal{H}}^H\boldsymbol{\mathcal{H}}\right)^{-1} \\ &= T\operatorname{tr}\left(\sum_{\ell=1}^{L} \boldsymbol{I}_M + \frac{\rho}{M}\boldsymbol{H}_\ell^H\boldsymbol{H}_\ell\right)^{-1}. \qquad (80)\end{aligned}$$

Substituting (79) and (80) respectively into the left and right sides of (78), the condition becomes

$$\begin{aligned}&\operatorname{tr}\left(\sum_{\ell=1}^{L} \left(\boldsymbol{I}_M + \frac{\rho}{M}\boldsymbol{H}_\ell^H\boldsymbol{H}_\ell\right)^{-1} \boldsymbol{C}_{i\ell}\boldsymbol{C}_{i\ell}^H\right) \\ &\qquad = \frac{1}{ML}\operatorname{tr}\left(\sum_{\ell=1}^{L} \left(\boldsymbol{I}_M + \frac{\rho}{M}\boldsymbol{H}_\ell^H\boldsymbol{H}_\ell\right)^{-1}\right). \qquad (81)\end{aligned}$$

Thus, for any channel realization $\boldsymbol{H}_\ell$, the condition in (67) holds iff $\boldsymbol{C}_{i\ell}\boldsymbol{C}_{i\ell}^H = \frac{1}{ML}\boldsymbol{I}_M$. ■

APPENDIX C
MULTIBLOCK STBC FOR SIMULATIONS

In this section, we provide two of the multiblock codes that were used in the simulations. The codes were designed using the method in Section V-A are represented by the $LMT \times LMT$ matrix $\boldsymbol{F}$.

1) Code for $M = N = T = 2, L = 2$:

$$\boldsymbol{F} = \frac{1}{2}\begin{bmatrix} 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & -1 & -1 \\ 1 & 1 & -1 & -1 & 0 & 0 & 0 & 0 \\ 1 & -1 & 1 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & -1 & 1 & -1 \\ 0 & 0 & 0 & 0 & 1 & -1 & -1 & 1 \\ 1 & -1 & -1 & 1 & 0 & 0 & 0 & 0 \end{bmatrix}$$

2) Code for $M = N = T = 2, L = 4$:

$$\boldsymbol{F} = [\boldsymbol{F}_A, \boldsymbol{F}_B]$$

where

$$\boldsymbol{F}_A = \frac{1}{2\sqrt{2}}\begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 \\ 1 & -i & -1 & i & 1 & -i & -1 & i \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & -i & -1 & i & -1 & i & 1 & -i \\ 1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & -1 & 1 & -1 & -1 & 1 & -1 & 1 \\ 1 & i & -1 & -i & 1 & i & -1 & -i \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & i & -1 & -i & -1 & -i & 1 & i \end{bmatrix}$$

$$\boldsymbol{F}_B = \frac{1}{2\sqrt{2}}\begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & -i & -1 & i & 1 & -i & -1 & i \\ 1 & -i & -1 & i & -1 & i & 1 & -i \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 \\ 1 & -1 & 1 & -1 & -1 & 1 & -1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & i & -1 & -i & 1 & i & -1 & -i \\ 1 & i & -1 & -i & -1 & -i & 1 & i \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

REFERENCES

[1] S. M. Alamouti, "A simple transmit diversity technique for wireless communications," *IEEE J. Select. Areas Commun.*, vol. 16, pp. 1451–1458, Oct. 1998.

[2] V. Tarokh, N. Seshadri, and A. R. Calderbank, "Space-time codes for high data rate wireless communication: Performance criterion and code construction," *IEEE Trans. Inf. Theory*, vol. 44, pp. 744–765, Mar. 1998.

[3] I. E. Telatar, "Capacity of multi-antenna Gaussian channels," *Eur. Trans. Telecommun.*, vol. 10, pp. 585–595, Nov./Dec. 1999.

[4] B. Hassibi and B. M. Hochwald, "High-rate codes that are linear in space and time," *IEEE Trans. Inf. Theory*, vol. 48, pp. 1804–1824, Jul. 2002.

[5] R. W. Heath and A. J. Paulraj, "Linear dispersion codes for MIMO systems based on frame theory," *IEEE Trans. Signal Process.*, vol. 50, pp. 2429–2441, Oct. 2002.

[6] O. Tirkkonen and A. Hottinen, "Square-matrix embeddable space-time codes for complex signal constellations," *IEEE Trans. Inf. Theory*, vol. 48, pp. 384–395, Feb. 2002.

[7] W. Su and X.-G. Xia, "Two generalized complex orthogonal space-time block codes of rates 7/11 and 3/5 for 5 and 6 transmit antennas," *IEEE Trans. Inf. Theory*, vol. 49, pp. 313–316, Jan. 2003.

[8] X.-B. Liang, "Orthogonal designs with maximal rates," *IEEE Trans. Inf. Theory*, vol. 49, pp. 2468–2503, Oct. 2003.

[9] X.-B. Liang and X.-G. Xia, "On the nonexistence of rate-one generalized complex orthogonal designs," *IEEE Trans. Inf. Theory*, vol. 49, pp. 2984–2988, Nov. 2003.

[10] H. El Gamal and M. O. Damen, "Universal space-time coding," *IEEE Trans. Inf. Theory*, vol. 49, pp. 1097–1119, May 2003.

[11] L. Zheng and D. N. C. Tse, "Diversity gain and multiplexing: A fundamental tradeoff in multiple antenna channels," *IEEE Trans. Inf. Theory*, vol. 49, pp. 1073–1096, May 2003.

[12] S. N. Diggavi, N. Al-Dhahir, A. Stamoulis, and A. R. Calderbank, "Great expectations: The value of spatial diversity in wireless networks," *Proc. IEEE*, vol. 92, pp. 219–270, Feb. 2004.

[13] X. Ma and G. B. Giannakis, "Full-diversity full rate complex-field space-time coding," *IEEE Trans. Signal Process.*, vol. 51, pp. 2917–2930, Nov. 2003.

[14] M. O. Damen, H. El Gamal, and N. C. Beaulieu, "Linear threaded algebraic space-time constellations," *IEEE Trans. Inf. Theory*, vol. 49, pp. 2372–2388, Oct. 2003.

[15] F. Oggier, G. Rekaya, J.-C. Belfiore, and E. Viterbo, "Perfect space time block codes," *IEEE Trans. Inf. Theory*, vol. 52, pp. 3885–3902, Sep. 2006.

[16] J. Liu, J.-K. Zhang, and K. M. Wong, "On the design of minimum BER linear space-time block codes for MIMO systems with MMSE receivers," *IEEE Trans. Signal Process.*, vol. 54, pp. 3147–3158, Aug. 2006.

[17] A. Fasano and S. Barbarossa, "Trace-orthogonal space-time coding," *IEEE Trans. Signal Process.*, vol. 56, pp. 2017–2034, May 2008.

[18] H. El Gamal and A. R. Hammons, "On the design of algebraic space-time codes for MIMO block fading channels," *IEEE Trans. Inf. Theory*, vol. 49, pp. 151–163, Jan. 2003.

[19] H.-F. Lu, "Constructions of multiblock space-time coding schemes that achieve the diversity-multiplexing tradeoff," *IEEE Trans. Inf. Theory*, vol. 54, pp. 3790–3796, Aug. 2008.

[20] R. H. Gohary and T. N. Davidson, "Design of linear dispersion codes: Some asymptotic guidelines and their implementation," *IEEE Trans. Wireless Commun.*, vol. 4, pp. 2892–2906, Nov. 2005.

[21] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge, U.K.: Cambridge Univ. Press, 2005.

[22] S. Sandhu and A. Paulraj, "Space-time block coding: A capacity perspective," *IEEE Commun. Letters*, vol. 4, pp. 384–386, Dec. 2000.

[23] H. V. Poor and S. Verdú, "Probability of error in MMSE multiuser detection," *IEEE Trans. Inf. Theory*, vol. 43, pp. 858–871, May 1997.

[24] J. Zhang and E. K. P. Chong, "Linear MMSE multiuser receivers: MAI conditional weak convergence and network capacity," *IEEE Trans. Inf. Theory*, vol. 48, pp. 2114–2122, Jul. 2002.

[25] J. Liu, "Linear STBC design for MIMO systems with MMSE receivers," Master's thesis, McMaster Univ., Hamilton, ON, Canada, Jun. 2004.

[26] R. J. Muirhead, *Aspects of Multivariate Statistical Theory*. New York: Wiley, 1982.

[27] T. Ratnarajah, R. Vaillancourt, and M. Alvo, "Eigenvalues and condition numbers of complex random matrices," *SIAM J. Matrix Anal. Appl.*, vol. 26, pp. 441–456, 2004.

[28] M. K. Simon and M.-S. Alouini, "A unified approach to the performance analysis of digital communication over generalized fading channels," *Proc. IEEE*, vol. 86, pp. 1860–1877, Sep. 1998.

[29] K. Cho and D. Yoon, "On the general BER expression of one and two dimensional amplitude modulations," *IEEE Trans. Commun.*, vol. 50, pp. 1074–1080, Jul. 2002.

[30] 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. 1994.

[31] D. Gore, R. W. Heath Jr., and A. Paulraj, "On performance of the zero forcing receivers in presence of transmit correlation," in *Proc. Int. Symp. Inf. Theory*, Lausanne, Switzerland, Jun.–Jul. 2002, p. 159.

[32] J. R. Barry, E. A. Lee, and D. G. Messerschmitt, *Digital Communication*, 3rd ed. Norwell, MA: Kluwer, 2004.

[33] X. Ma and W. Zhang, "Performance analysis for MIMO systems with lattice-reduction aided linear equalization," *IEEE Trans. Commun.*, vol. 56, pp. 309–918, Feb. 2008.

[34] K. R. Kumar, G. Caire, and A. L. Moustakas, "Asymptotic performance of linear receivers in MIMO fading channels," *IEEE Trans. Inf. Theory*, vol. 55, pp. 4398–4418, Oct. 2009.

[35] A. Scaglione, P. Stoica, S. Barbarossa, G. Giannakis, and H. Sampath, "Optimal designs for space-time linear precoders and decoders," *IEEE Trans. Signal Process.*, vol. 50, pp. 1051–1064, May 2002.

[36] J. Liu, J.-K. Zhang, and K. M. Wong, "Design of optimal orthogonal linear codes in MIMO systems for MMSE receivers," in *Proc. IEEE Conf. Acoust., Speech, Signal Process.*, Montreal, QC, Canada, May 2004, vol. 4, pp. 725–728.

[37] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.

[38] N. N. Lebedev, *Special Functions and Their Applications*. Englewood Cliffs, NJ: Prentice-Hall, 1965.

[39] P. Duhamel and M. Vetterli, "Fast Fourier transforms: A tutorial review and a state of the art," *Signal Process.*, vol. 19, pp. 259–299, 1990.

[40] M. Frigo and S. Johnson, "The design and implementation of FFTW3," *Proc. IEEE*, vol. 93, pp. 216–231, Feb. 2005.

[41] E. Biglieri, Y. Hong, and E. Viterbo, "On fast-decodable space-time block codes," *IEEE Trans. Inf. Theory*, vol. 55, pp. 524–530, Feb. 2009.

[42] R. A. Horn and C. R. Johnson, *Topics in Matrix Analysis*. Cambridge, U.K.: Cambridge Univ. Press, 1994.

Jing Liu received the B.S. degree in electronic science from Nan Kai University, Tianjin, China, in 1997, and the M.A.Sc. and Ph.D. degrees in electrical engineering from McMaster University, Hamilton, ON, Canada, in 2004 and 2008, respectively, where she conducted research in wireless communications and signal processing.

She is currently a Senior Designer at Siemens Canada, Peterborough, ON.

Dr. Liu received the Outstanding Thesis Award from McMaster University in 2004, the Outstanding Teaching Assistant Award from McMaster University in 2006, and the 2007 Young Author Best Paper Award from IEEE Signal Processing Society. She was the holder of an NSERC postgraduate scholarship for 2006–2008.

Timothy N. Davidson (M'96) received the B.Eng. (Hons. I) degree in electronic engineering from the University of Western Australia (UWA), Perth, in 1991 and the D.Phil. degree in engineering science from the University of Oxford, U.K., in 1995.

He has held research positions with the Communications Research Laboratory, McMaster University; the Adaptive Signal Processing Laboratory, UWA; and the Australian Telecommunications Research Institute, Curtin University of Technology, Perth. He is currently a Professor in the Department of Electrical and Computer Engineering, McMaster University, Hamilton, ON, Canada, where he holds the (Tier II) Canada Research Chair in Communication Systems and is currently serving as Associate Director of the School of Computational Engineering and Science. His research interests lie in the general areas of communications, signal processing, and control.

Dr. Davidson received the 1991 J. A. Wood Memorial Prize from UWA, and the 1991 Rhodes Scholarship for Western Australia. He is serving, or has served, as an Associate Editor of the IEEE TRANSACTIONS ON SIGNAL PROCESSING, the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS—PART II: EXPRESS BRIEFS, and *Optimization and Engineering*. He has also served as a Guest Coeditor of issues of the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS, the IEEE JOURNAL OF SELECTED TOPICS IN SIGNAL PROCESSING, and the *EURASIP Journal on Advances in Signal Processing*. He is a member of the IEEE Signal Processing Society's Technical Committee on Signal Processing for Communications and Networking. He is a Registered Professional Engineer in the Province of Ontario.

K. Max Wong (SM'81–F'02) received the B.Sc.(Eng.), DIC, Ph.D., and D.Sc.(Eng.) degrees, all in electrical engineering, from the University of London, U.K., in 1969, 1972, 1974 and 1995, respectively.

He started working at the Transmission Division of Plessey Telecommunications Research Ltd., England, in 1969. In October 1970 he was on leave from Plessey pursuing postgraduate studies and research at Imperial College of Science and Technology, London. In 1972, he rejoined Plessey as a Research Engineer and worked on digital signal processing and signal transmission. In 1976, he joined the Department of Electrical Engineering, Technical University of Nova Scotia, Canada, and in 1981, moved to McMaster University, Hamilton, ON, Canada, where he has been a Professor since 1985 and served as Chairman of the Department of Electrical and Computer Engineering in 1986–1987, 1988–1994 and 2003–2008. He was on leave as Visiting Professor at the Department of Electronic Engineering of the Chinese University of Hong Kong from 1997 to 1999. At present, he holds the Canada Research Chair in Signal Processing at McMaster University. His research interest is in signal processing and communication theory, and he has published over 240 papers in the area.

Prof. Wong is a Fellow of the Institution of Electrical Engineers, a Fellow of the Royal Statistical Society, and a Fellow of the Institute of Physics. He was the recipient of a number of prestigious recognitions and awards nationally and internationally including the IEE Overseas Premium for the best paper in 1989, the ITG Best Paper Award in 2003, and is also the coauthor of the papers that received the IEEE Signal Processing Society "Best Young Author" awards of 2006 and 2008. In 1996, the degree of D.Sc. from the University of London, U.K., was conferred upon him by Her Royal Highness Princess Anne in recognition of his standing as an international authority in signal processing. More recently, he has also been elected as Fellow of the Canadian Academy of Engineering as well as Fellow of the Royal Society of Canada. In addition to his outstanding scientific achievements, he has also assumed a number of influential technical positions including the membership of various provincial and federal technical advisory committees in and outside Canada. He was an Associate Editor of the IEEE TRANSACTIONS ON SIGNAL PROCESSING from 1996–1998 and served as Chair of the Sensor Array and Multi-channel Signal Processing Technical Committee of the IEEE Signal Processing Society in 2002–2004. He was the recipient of a medal presented by the International Biographical Centre, Cambridge, U.K., for his outstanding contributions to the research and education in signal processing in May 2000, and was honoured with the inclusion of his biography in the two books: *Outstanding People of the 20th Century* and *2000 Outstanding Intellectuals of the 20th Century* published by IBC to celebrate the arrival of the new millennium. In 2009, he was named a Distinguished Visiting Fellow of the Royal Academy of Engineering, U.K. In 2010, he was named a recipient of the prestigious Research Award of the Alexander von Humboldt Foundation, and was received in Berlin by the President of the Federal Republic of Germany.