

Design of Block Transceivers With Decision Feedback Detection

Fang Xu, Timothy N. Davidson, *Member, IEEE*, Jian-Kang Zhang, and K. Max Wong, *Fellow, IEEE*

Abstract—This paper presents a method for jointly designing the transmitter–receiver pair in a block-by-block communication system that employs (intra)block decision feedback detection. We provide closed-form expressions for transmitter–receiver pairs that simultaneously minimize the arithmetic mean squared error (MSE) at the decision point (assuming perfect feedback), the geometric MSE, and the bit error rate of a uniformly bit-loaded system at moderate-to-high signal-to-noise ratios. Separate expressions apply for the “zero-forcing” and “minimum MSE” (MMSE) decision feedback structures. In the MMSE case, the proposed design also maximizes the Gaussian mutual information and suggests that one can approach the capacity of the block transmission system using (independent instances of) the same (Gaussian) code for each element of the block. Our simulation studies indicate that the proposed transceivers perform significantly better than standard transceivers and that they retain their performance advantages in the presence of error propagation.

Index Terms—Bit error rate, block precoding, channel capacity, decision feedback detection, minimum mean-square error, mutual information, zero-forcing.

I. INTRODUCTION

BLOCK-BY-BLOCK communication is an effective scheme for the transmission of data over dispersive media; e.g., [28]–[30], [41], [42]. In such “vector” communication schemes, blocks of data are transmitted in a manner that avoids interference between the received blocks, and hence the detector need only operate on a block-by-block basis. Two popular examples of block-by-block communication schemes are orthogonal frequency-division multiplexing (OFDM) [5] and discrete multitone modulation (DMT) [8]. In addition, certain multiple antenna systems operate in a block-by-block fashion (e.g., [18], [20], [26], [36], [43], and [45]), and block-by-block detection schemes appear in some multiuser detectors for synchronous code-division multiple-access systems [14], [15], [47]. In general, an optimal detector for a block transmission system must make a decision on the received data block as a whole, although in certain cases, such as OFDM and DMT, the elements of that block can be decoupled and simpler detection schemes obtained. Unfortunately, maximum likelihood detection of the transmitted vector can

be rather computationally expensive, and simpler detectors based on linear equalization and (disjoint) symbol-by-symbol detection may incur a significant performance loss. A useful compromise between performance and complexity can be obtained by employing intra)block decision feedback detection [4], [10], [14], [15], [18], [20], [22], [28], [44], [47], [51]. In an intra)block decision feedback detector the individual symbols which constitute a given block are detected sequentially, with the “intra)block interference” from previously detected symbols being subtracted before the decision on the current symbol is made. Such schemes fall into the class of generalized decision feedback equalizers [10]. In multiple antenna communication schemes intra)block decision feedback is sometimes referred to as “nulling and cancelling” [4], [18], [20], and in multiuser detection the corresponding concept is sometimes referred to as “successive interference cancellation” [14], [15], [22], [47].

The goal of this paper is to jointly design the linear transmitter matrix and the receiver feedforward and feedback matrices so as to optimize the performance of a block-by-block communication system with an intra)block decision feedback detector (BDFD). The design is based on knowledge of the channel, and hence is an appropriate choice for systems in which there is timely, reliable feedback from the receiver to the transmitter. The proposed approach provides closed-form expressions for transceivers that minimize the arithmetic mean (over the block) of the expected squared errors (MSE) at the input to the (scalar) decision device that is implicit in the BDFD, under the standard assumption [3], [9], [10], [17], [40], [52] that the previous decisions were correct.

The expressions depend on the nature of the BDFD, and separate expressions are provided for the zero-forcing (ZF) and minimum mean square error (MMSE) BDFDs. In order to help distinguish our designs from previous work, we point out that if one is given a transmitter matrix, the design of the feedforward and feedback matrices of a ZF or MMSE-BDFD that minimize the MSE is well known; e.g., [2], [4], [9], [10], [17], [20], and [40]. However, the joint minimum MSE design of the transmitter and receiver matrices has previously been deemed to be difficult (e.g., [52, p. 1338]), and hence several authors have suggested minimizing a particular lower bound on the MSE, namely, the geometric mean of the expected squared errors; e.g., [9], [10], [52]. We will minimize the geometric MSE as the first step in our approach, but we will also show how the unitary matrix that parameterizes the set of transceivers which minimize the geometric MSE can be chosen so that the (arithmetic) MSE attains its minimized lower bound.

Transceivers designed in the manner we propose have several additional desirable properties. In particular, the inputs to

Manuscript received July 10, 2004; revised February 25, 2005. This work was supported in part by the National Science and Engineering Research Council of Canada. The work of T. N. Davidson was also supported by the Canada Research Chairs program. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Martin Haardt.

The authors are with the Department of Electrical and Computer Engineering, McMaster University, Hamilton, ON L8S 4K1, Canada (e-mail: xufang042003@yahoo.ca; davidson@mcmaster.ca; jkzhang@mail.ece.mcmaster.ca; wongkm@mcmaster.ca).

Digital Object Identifier 10.1109/TSP.2005.861779

the (scalar) decision device are uncorrelated and have equal signal-to-interference-and-noise ratios (SINRs). In fact, the minimum SINR over the elements of the block is maximized. As a result, the average bit error rate (BER) is (essentially) minimized. More precisely, for systems with a ZF-BDFD, our design minimizes the average BER for (uncoded) uniform quaternary phase-shift keying (QPSK) signaling at moderate-to-high signal-to-noise ratios (SNRs) and also minimizes the dominant components of the BER for uniform M -ary quadrature amplitude modulation (QAM) signaling.¹ For systems with an MMSE-BDFD, our design minimizes the average BER under an assumption that the residual intrablock interference is Gaussian.

For the MMSE-BDFD, it is reasonably well known [9], [10], [17], [40], [52] that any transmitter that minimizes the geometric MSE (including the proposed design) also maximizes the mutual information between the transmitter and receiver for Gaussian signals. However, the standard choice from the set of transmitters that minimize the geometric MSE does not minimize the (arithmetic) MSE and produces inputs to the decision device that have potentially different SINRs for each element of the block. Therefore, in order to achieve reliable communication at rates which approach the capacity of the block transmission system, different codes (and constellations) may need to be applied for each element of the block [10]. An advantage of the proposed design is that from within the set of transmitters that minimize the geometric MSE (and maximize the Gaussian mutual information), we obtain a transceiver that also minimizes the arithmetic MSE, minimizes the BER, and provides uncorrelated inputs to the decision device that have identical (and maximized) SINRs. Since the MMSE-BDFD is a “canonical” receiver [9], [10], [23], this suggests that by using the proposed design, reliable communication at rates approaching the capacity of the block transmission system can be achieved by using independent instances of the same (Gaussian) code in each element of block.

As mentioned earlier, our designs are based on the standard assumption [3], [9], [10], [17], [40], [52] that the previous symbols were correctly detected. However, error propagation is not catastrophic in block-by-block communication schemes because errors can only propagate within a single block (e.g., [10] and Section II). Bounds for the conventional symbol-by-symbol decision feedback equalizer (DFE) [1], [16] also suggest that good performance should be maintained in the presence of error propagation, and our simulations confirm this prediction. Furthermore, our simulation studies indicate that the proposed transceivers perform significantly better than standard transceivers and that they retain their performance advantages in the presence of error propagation.

Notation: The notation adopted in this paper is fairly standard. We conform to the following conventions: scalars are denoted by lowercase letters; vectors by bold lowercase letters; and matrices by bold upper case letters. The symbol \mathbf{I}_N denotes the identity matrix of size N , and $\mathbf{0}_{N \times M}$ denotes the $N \times M$ matrix of zeros. The symbol $|\mathbf{A}|$ denotes the determinant

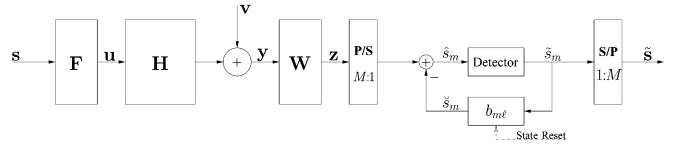


Fig. 1. A generic block-by-block communication system with intrablock decision feedback detection. The P/S block denotes parallel-to-serial conversion with the last element of the input block becoming the first output, and the S/P block denotes serial-to-parallel conversion with the first input becoming the last element of the output block.

of a matrix \mathbf{A} , and $\text{tr}(\mathbf{A})$ denotes its trace. The symbol $E[\cdot]$ denotes the expectation operator; $(\cdot)^H$ the complex-conjugate transpose operation; $(\cdot)^T$ the transpose operation; and $[\cdot]_{ij}$ the element at the intersection of the i th row and j th column of a matrix.

II. BLOCK-BY-BLOCK TRANSMISSION

We consider the generic block-by-block transmission system with intrablock decision feedback detection illustrated in Fig. 1. In this system, a block of M data symbols \mathbf{s} is linearly precoded to construct a block of $K \geq M$ channel symbols, $\mathbf{u} = \mathbf{F}\mathbf{s}$, which is transmitted over the channel. The receiver independently processes a block of $P \geq M$ received samples in order to detect the data vector \mathbf{s} . The received block \mathbf{y} can be written as

$$\mathbf{y} = \mathbf{H}\mathbf{F}\mathbf{s} + \mathbf{v} \quad (1)$$

where the $P \times K$ matrix \mathbf{H} captures the effects of the channel and \mathbf{v} is a length P vector of additive noise samples. We will assume that the noise is circularly symmetric [37] (or proper [35]) and Gaussian, with zero mean and positive definite correlation matrix $E[\mathbf{v}\mathbf{v}^H] = \mathbf{R}_{vv}$. We will also assume that the data symbols have zero mean and are white,² of unit energy, and not correlated with the noise, (i.e., $E[\mathbf{s}\mathbf{s}^H] = \mathbf{I}$ and $E[\mathbf{s}\mathbf{v}^H] = \mathbf{0}$). The model in (1) is applicable in many applications, including zero-padded or cyclic-prefixed block transmission over a scalar finite impulse response channel that is constant over the duration of the block; e.g., [6], [12], [28]–[30], [41], [42], and [44]. In the zero-padded case, \mathbf{H} is a tall, lower triangular, full column rank Toeplitz matrix whose columns contain the impulse response of the channel, and in the cyclic-prefixed case \mathbf{H} is a square circulant matrix whose columns contain the channel impulse response. The model in (1) is also applicable in vector transmission over a narrow-band multiple antenna channel (e.g., [18] and [20]), in which case \mathbf{H} has no deterministic structure; in space-time block transmission over a (quasi-static) narrow-band multiple antenna channel (e.g., [26] and [45]), in which case \mathbf{H} has a block diagonal structure; and in block transmission over a (quasi-static) frequency-selective multiple antenna channel (e.g., [36] and [43]), in which case \mathbf{H} is either block Toeplitz or block circulant.

The intrablock decision feedback detector first preprocesses the received block \mathbf{y} with an $M \times P$ feedforward matrix \mathbf{W} to form $\mathbf{z} = \mathbf{W}\mathbf{y}$. (The functional form of \mathbf{W} depends on

¹Our design for the ZF-BDFD coincides with the one that minimizes the block error rate [56], [57], but the design approach taken in this paper is substantially different from that taken in [56] and [57].

²In the case where $E[\mathbf{s}\mathbf{s}^H]$ is not a scaled identity matrix, a data whitening matrix can readily be absorbed into the precoder, so long as the data covariance matrix is known (and full rank).

whether the ZF- or MMSE-BDFD is implemented; see Section III.) The detection of the transmitted symbols $s_m = [\mathbf{s}]_m$ then proceeds sequentially, starting from $m = M$, by making a scalar decision on $\hat{s}_M = z_M$ and then $\hat{s}_m = z_m - \check{s}_m$, $m = M - 1, M - 2, \dots, 1$, where $\check{s}_m = \sum_{\ell=m+1}^M b_{m\ell} \hat{s}_\ell$ is the output of the feedback filter, with $b_{m\ell}$ being its coefficients. The states of that filter \check{s}_ℓ are the previously detected symbols in the block, and the filter coefficients are different for each element of the block (indexed by m). Once a given block has been detected, the states of the feedback filter are reset to zero. That is, the symbols are detected on a block-by-block basis and hence error propagation between blocks is avoided.

If the filter coefficients $b_{m\ell}$ are arranged in a strictly upper triangular $M \times M$ matrix

$$\mathbf{B} = \begin{bmatrix} 0 & b_{12} & b_{13} & \cdots & b_{1M} \\ 0 & 0 & b_{23} & \cdots & b_{2M} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & b_{(M-1)M} \\ 0 & 0 & 0 & \cdots & 0 \end{bmatrix}$$

the operation of the block transceiver in Fig. 1 is equivalent to successively making decisions on the elements of

$$\hat{\mathbf{s}} = \mathbf{WHF}\mathbf{s} + \mathbf{W}\mathbf{v} - \mathbf{B}\check{\mathbf{s}} \quad (2)$$

starting from the M th row. That interpretation leads to the convenient conceptual model in Fig. 2. We observe that when $\mathbf{B} = \mathbf{0}$, the system in Fig. 2 reduces to a block transmission system with linear equalization and disjoint detection; e.g., [6], [12], [36], and [41]–[43]. In fact, many of the results for the linear case can be obtained by setting $\mathbf{B} = \mathbf{0}$ in the expressions we will derive herein.

If we denote the error between the input to the detector and the transmitted data symbols by $\mathbf{e} = \hat{\mathbf{s}} - \mathbf{s}$, then

$$\mathbf{e} = (\mathbf{WHF} - \mathbf{I})\mathbf{s} - \mathbf{B}\check{\mathbf{s}} + \mathbf{W}\mathbf{v}. \quad (3)$$

Under the assumption of correct past decisions (i.e., when deciding s_m , $\check{s}_\ell = s_\ell$ for all $m + 1 \leq \ell \leq M$), \mathbf{e} simplifies to

$$\mathbf{e} = (\mathbf{WHF} - \mathbf{I} - \mathbf{B})\mathbf{s} + \mathbf{W}\mathbf{v}. \quad (4)$$

The covariance of this error will play a key role in our designs. Under our statistical models for \mathbf{s} and \mathbf{v} , the covariance matrix of the error is

$$\mathbf{R}_{ee} = E[\mathbf{e}\mathbf{e}^H] = (\mathbf{WHF} - \mathbf{B} - \mathbf{I})(\mathbf{WHF} - \mathbf{B} - \mathbf{I})^H + \mathbf{WR}_{vv}\mathbf{W}^H. \quad (5)$$

The (arithmetic) MSE of the detector input is simply $\bar{e}^2 = \text{tr}(E[(\hat{\mathbf{s}} - \mathbf{s})(\hat{\mathbf{s}} - \mathbf{s})^H])/M = \text{tr}(\mathbf{R}_{ee})/M$.

III. MINIMUM MSE TRANSCEIVERS

In this section, our goal is to jointly design the transceiver elements \mathbf{F} , \mathbf{B} , and \mathbf{W} so that the (arithmetic) MSE is minimized, subject to a bound p_0 on the average transmitted power and constraints which ensure that the receiver performs either

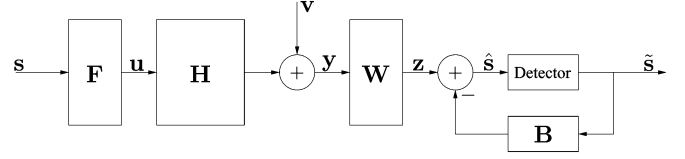


Fig. 2. A convenient conceptual model for Fig. 1.

ZF or MMSE decision-feedback detection. The average transmitted power is given by $E[\text{tr}(\mathbf{F}\mathbf{s}(\mathbf{F}\mathbf{s})^H)] = \text{tr}(\mathbf{F}\mathbf{F}^H)$, and hence the design problem can be stated as

$$\min_{\mathbf{F}, \mathbf{B}, \mathbf{W}} \text{tr}((\mathbf{WHF} - \mathbf{B} - \mathbf{I})(\mathbf{WHF} - \mathbf{B} - \mathbf{I})^H + \mathbf{WR}_{vv}\mathbf{W}^H) \quad (6a)$$

$$\text{subject to } \text{tr}(\mathbf{F}\mathbf{F}^H) \leq p_0 \quad (6b)$$

$$\text{and a functional relationship between } \mathbf{F}, \mathbf{B}, \text{ and } \mathbf{W}. \quad (6c)$$

The functional relationship between \mathbf{F} , \mathbf{B} , and \mathbf{W} determines whether the BDFD is of the ZF type or the MMSE type. This optimization problem is rather difficult to solve directly because it is not convex and hence is subject to the standard difficulties associated with the potential for multiple local minima. However, we will use the following stages to find a solution $(\mathbf{F}, \mathbf{B}, \mathbf{W})$ whose performance is optimal.

- 1) Obtain a (tight) lower bound on the MSE, and minimize that lower bound, subject to the constraint on transmission power.
- 2) Derive a triple $(\mathbf{F}, \mathbf{B}, \mathbf{W})$ whose performance achieves the minimized lower bound.

In the following sections, we will perform the above stages to obtain the minimized lower bounds on the MSE and optimal transceivers for the ZF and MMSE BDFDs, respectively. The matrix $\mathbf{H}^H\mathbf{R}_{vv}^{-1}\mathbf{H}$ will play a key role in our designs. For later convenience we let

$$\mathbf{V}\mathbf{\Lambda}\mathbf{V}^H = \mathbf{H}^H\mathbf{R}_{vv}^{-1}\mathbf{H} \quad (7)$$

represent the eigenvalue decomposition of $\mathbf{H}^H\mathbf{R}_{vv}^{-1}\mathbf{H}$, with eigenvalues λ_i arranged in nonincreasing order along the diagonal of $\mathbf{\Lambda}$. For an integer $1 \leq k \leq K$, we also define $\tilde{\mathbf{V}}_k$ to be the first k columns of \mathbf{V} and $\tilde{\mathbf{\Lambda}}_k$ to be the upper left $k \times k$ block of $\mathbf{\Lambda}$. In the development of our designs, we will find it convenient to parameterize the $K \times M$ precoder matrix \mathbf{F} of rank q in terms of its singular value decomposition

$$\mathbf{F} = \mathbf{\Theta} [\mathbf{\Phi} \quad \mathbf{0}_{q \times (M-q)}] \mathbf{\Psi} \quad (8)$$

where $\mathbf{\Theta}$ contains q columns of a $K \times K$ unitary matrix, $\mathbf{\Phi}$ is a diagonal positive definite $q \times q$ matrix, and $\mathbf{\Psi}$ is an $M \times M$ unitary matrix.

A. Zero-Forcing BDFD

The zero-forcing criterion imposes the following relationship between \mathbf{W} , \mathbf{F} , and \mathbf{B} [see (4)]:

$$\mathbf{WHF} = \mathbf{B} + \mathbf{I}. \quad (9)$$

Given a $P \times K$ matrix \mathbf{H} and an integer $M \leq \min\{P, K\}$, there exists a $K \times M$ matrix \mathbf{F} , an $M \times P$ matrix \mathbf{W} , and an $M \times M$ strictly upper triangular matrix \mathbf{B} such that (9) is satisfied if and only if $\text{rank}(\mathbf{H}) \geq M$, and we will make the assumption that this condition holds.³ In order to satisfy (9), \mathbf{F} must be chosen so that it has rank M and that $\text{rank}(\mathbf{H}\mathbf{F}) = M$.

By substituting (9) into (4) and (5), the covariance matrix of the error can be written as

$$\mathbf{R}_{ee, \text{ZF}} = \mathbf{W}\mathbf{R}_{vv}\mathbf{W}^H. \quad (10)$$

If we define $\check{\mathbf{W}} = \mathbf{W}\mathbf{R}_{vv}^{1/2}$, then the design problem (6) can be rewritten as

$$\min_{\check{\mathbf{W}}, \mathbf{B}, \mathbf{F}} \text{tr}(\check{\mathbf{W}}\check{\mathbf{W}}^H) \quad (11a)$$

$$\text{subject to } \text{tr}(\mathbf{F}\mathbf{F}^H) \leq p_0 \quad (11b)$$

$$\check{\mathbf{W}}\check{\mathbf{H}}\mathbf{F} = \mathbf{B} + \mathbf{I} \quad (11c)$$

where $\check{\mathbf{H}} = \mathbf{R}_{vv}^{-1/2}\mathbf{H}$. From (11), it is clear that for a given \mathbf{F} for which there exists a solution to (11c) and a given \mathbf{B} , the optimal $\check{\mathbf{W}}$ is $\check{\mathbf{W}} = (\mathbf{B} + \mathbf{I})(\check{\mathbf{H}}\mathbf{F})^+$, where $(\cdot)^+$ denotes the (minimum-norm) Moore–Penrose pseudoinverse. Therefore, the optimal receiver feedforward matrix can be written as

$$\mathbf{W}_{\text{ZF}} = (\mathbf{B} + \mathbf{I})(\check{\mathbf{H}}\mathbf{F})^+\mathbf{R}_{vv}^{-1/2}. \quad (12)$$

Since $\check{\mathbf{H}}\mathbf{F}$ has at least as many rows as it has columns and has full column rank

$$(\check{\mathbf{H}}\mathbf{F})^+ = (\mathbf{F}^H\check{\mathbf{H}}^H\check{\mathbf{H}}\mathbf{F})^{-1}\mathbf{F}^H\check{\mathbf{H}}^H. \quad (13)$$

If we let $\mathbf{U} = \mathbf{B} + \mathbf{I}$, the design problem in (11) has been reduced to

$$\min_{\mathbf{U}, \mathbf{F}} \text{tr}(\mathbf{U}(\check{\mathbf{H}}\mathbf{F})^+((\check{\mathbf{H}}\mathbf{F})^+)^H\mathbf{U}^H) \quad (14a)$$

$$\text{subject to } \text{tr}(\mathbf{F}\mathbf{F}^H) \leq p_0 \quad (14b)$$

$$\begin{aligned} &\mathbf{U} \text{ being a unit-diagonal} \\ &\text{upper triangular matrix.} \end{aligned} \quad (14c)$$

The first stage in our solution of (14) is to derive and minimize a lower bound on the objective function (14a). The lower bound that we will use is a simple consequence of the arithmetic-geometric mean inequality [27, p. 535]. In particular, for an $M \times M$ positive semidefinite matrix \mathbf{X}

$$\text{tr}(\mathbf{X})/M \geq |\mathbf{X}|^{1/M} \quad (15)$$

with equality holding if and only if $\mathbf{X} = \alpha\mathbf{I}$ for some $\alpha \geq 0$. For convenience, we will refer to (15) as the trace-determinant inequality.

Applying (15) to (14a), a lower bound on the mean-square error is

$$\begin{aligned} \tilde{e}_{\text{ZF}}^2 &= \text{tr}(\mathbf{W}_{\text{ZF}}\mathbf{R}_{vv}\mathbf{W}_{\text{ZF}}^H)/M \\ &\geq |\mathbf{U}(\check{\mathbf{H}}\mathbf{F})^+((\check{\mathbf{H}}\mathbf{F})^+)^H\mathbf{U}^H|^{1/M} \end{aligned} \quad (16a)$$

$$= |\mathbf{F}^H\check{\mathbf{H}}^H\check{\mathbf{H}}\mathbf{F}|^{-1/M} \quad (16b)$$

where we have used the fact that \mathbf{U} is a unit-diagonal upper-triangular matrix and thus $|\mathbf{U}| = 1$ and the expression for $(\check{\mathbf{H}}\mathbf{F})^+$ in (13). Observe that (16b) depends only on the transmitter \mathbf{F} and is independent of $\mathbf{U} = \mathbf{B} + \mathbf{I}$. It is also of interest to

³If M were a design variable, rather than a parameter of the problem, one could guarantee that this condition holds by simply choosing $M \leq \text{rank}(\mathbf{H})$.

point out that the bound in (16a) is equivalent to stating that the arithmetic MSE is bounded below by the geometric MSE; i.e., $\text{tr}(\mathbf{R}_{ee, \text{ZF}})/M \geq |\mathbf{R}_{ee, \text{ZF}}|^{1/M}$. Therefore, the problem of minimizing the lower bound in (16a) corresponds to minimizing the geometric MSE.

The lower bound in (16) can be minimized simply by maximizing $|\mathbf{F}^H\check{\mathbf{H}}^H\check{\mathbf{H}}\mathbf{F}|$; i.e., by solving

$$\max_{\mathbf{F}} |\mathbf{F}^H\mathbf{H}^H\mathbf{R}_{vv}^{-1}\mathbf{H}\mathbf{F}| \quad (17a)$$

$$\text{subject to } \text{tr}(\mathbf{F}\mathbf{F}^H) \leq p_0. \quad (17b)$$

Using the ordered eigendecomposition of $\mathbf{H}^H\mathbf{R}_{vv}^{-1}\mathbf{H}$ in (7) and applying the trace-determinant inequality (15), we have that

$$|\mathbf{F}^H\mathbf{H}^H\mathbf{R}_{vv}^{-1}\mathbf{H}\mathbf{F}| = |\Phi^2| |\Theta^H\mathbf{H}^H\mathbf{R}_{vv}^{-1}\mathbf{H}\Theta| \quad (18a)$$

$$\leq \left(\frac{\text{tr}(\Phi^2)}{M}\right)^M \prod_{i=1}^M \lambda_i \quad (18b)$$

$$\leq \left(\frac{p_0}{M}\right)^M \prod_{i=1}^M \lambda_i. \quad (18c)$$

Therefore, for any ZF-BDFD system, the (arithmetic) MSE is bounded below by

$$\tilde{e}_{\text{ZF}}^2 \geq \frac{M}{p_0} \left(\prod_{i=1}^M \lambda_i\right)^{-1/M}. \quad (19)$$

This bound depends only on the parameters M and p_0 and the M largest eigenvalues of $\mathbf{H}^H\mathbf{R}_{vv}^{-1}\mathbf{H}$.

The second stage of the derivation of the proposed design is to determine matrices \mathbf{F} and \mathbf{B} so that the minimized lower bound on the arithmetic MSE in (19) is achieved. To do so, we point out that according to the trace-determinant inequality (15) and the eigenvalue decomposition of $\mathbf{H}^H\mathbf{R}_{vv}^{-1}\mathbf{H}$ in (7), the bound in (18b) holds with equality if and only if $\Phi = \alpha\mathbf{I}$ for some $\alpha > 0$ and $\Theta = \tilde{\mathbf{V}}_M\mathbf{P}$, where $\tilde{\mathbf{V}}_M$ was defined after (7) and \mathbf{P} is an arbitrary permutation matrix. According to the power constraint in (17b), the bound in (18c) is achieved if and only if $\alpha = \sqrt{p_0/M}$. Therefore, precoders of the form $\mathbf{F} = \sqrt{p_0/M}\tilde{\mathbf{V}}_M\mathbf{\Psi}$, where $\mathbf{\Psi}$ is an arbitrary $M \times M$ unitary matrix, minimize the geometric MSE of a ZF-BDFD system. The remaining task is to determine matrices $\mathbf{\Psi}$ such that the bound in (16a) holds with equality. To do so, we observe that the trace-determinant inequality (15) holds with equality if and only if $\mathbf{X} = \alpha\mathbf{I}$ for some $\alpha \geq 0$. Therefore, (16a) holds with equality if and only if we can choose $\mathbf{\Psi}$ such that $\mathbf{R}_{ee, \text{ZF}} = \sigma_e^2\mathbf{I}$, where $\sigma_e^2 = (M/p_0)(\prod_{i=1}^M \lambda_i)^{-1/M}$. That is, we can achieve the minimized lower bound on the arithmetic MSE if and only if we can find a $\mathbf{\Psi}$ such that

$$\frac{M}{p_0}\mathbf{U}\mathbf{\Psi}^H\tilde{\Lambda}_M^{-1}\mathbf{\Psi}\mathbf{U}^H = \sigma_e^2\mathbf{I} \quad (20)$$

where $\tilde{\Lambda}_M$ was defined after (7). By taking the Cholesky factor, solving (20) is equivalent to solving

$$\sqrt{\frac{M}{p_0}}\mathbf{U}\mathbf{\Psi}^H\tilde{\Lambda}_M^{-1/2} = \sigma_e\mathbf{Q}^H \quad (21)$$

where \mathbf{Q} is an $M \times M$ unitary matrix. That is, we can reduce the search for a pair (\mathbf{F}, \mathbf{B}) such that the minimized lower bound on the MSE is achieved to the search for a unit-diagonal upper

triangular matrix \mathbf{U} and unitary matrices $\mathbf{\Psi}$ and \mathbf{Q} that satisfy (21). Substituting σ_e into (21), we get

$$\tilde{\mathbf{\Lambda}}_M^{1/2} \mathbf{\Psi} = \mathbf{Q} \tilde{\mathbf{U}} \quad (22)$$

where $\tilde{\mathbf{U}} = (\prod_{i=1}^M \lambda_i)^{1/(2M)} \mathbf{U}$. The following result, which is a special case of a more general result in [56] and [57], indicates that a solution to (22) exists.

Lemma 1: Let $\mathbf{\Gamma}$ be a diagonal nonsingular $M \times M$ matrix. There exists a unitary matrix \mathbf{S} such that $\mathbf{\Gamma}\mathbf{S}$ has an equal-diagonal ‘‘R-factor’’ in its (standard) QR decomposition; i.e., $\exists \mathbf{S}$ such that $\mathbf{\Gamma}\mathbf{S} = \mathbf{Q}\mathbf{R}$, where \mathbf{Q} is an $M \times M$ unitary matrix and \mathbf{R} is an upper triangular matrix with equal diagonal elements $[\mathbf{R}]_{ii} = (\prod_{k=1}^M \gamma_k)^{1/M}$ for $i = 1, 2, \dots, M$, where γ_k is the k th diagonal element of $\mathbf{\Gamma}$. \square

The matrix \mathbf{S} in Lemma 1 can be obtained by suitably modifying [57, Algorithm 5]. The modified algorithm is provided in Appendix I. Using that algorithm, we can obtain $\mathbf{\Psi}$ in (22). By performing the QR decomposition of $\tilde{\mathbf{\Lambda}}_M^{1/2} \mathbf{\Psi}$, we obtain an upper triangular matrix $\tilde{\mathbf{U}}$ whose diagonal elements are all equal to $(\prod_{i=1}^M \lambda_i)^{1/(2M)}$. Finally, we obtain \mathbf{U} using $\mathbf{U} = (\prod_{i=1}^M \lambda_i)^{-1/(2M)} \tilde{\mathbf{U}}$. Thus, we have established the following proposition.

Proposition 1: The (arithmetic) mean-square error $\text{tr}(\mathbf{R}_{ee})/M$ of a block-by-block transceiver with a ZF-BDFD achieves its minimized lower bound of $(M/p_0)(\prod_{i=1}^M \lambda_i)^{-1/M}$ when the precoder $\mathbf{F} = \sqrt{(p_0)/(M)} \tilde{\mathbf{V}}_M \mathbf{\Psi}_{\text{ZF}}$, where $\mathbf{\Psi}_{\text{ZF}}$ is obtained by applying the algorithm in Appendix I to $\tilde{\mathbf{\Lambda}}_M^{1/2}$. The corresponding feedback matrix $\mathbf{B} = \mathbf{U} - \mathbf{I}$, where \mathbf{U} is the unit-diagonal upper triangular matrix $\mathbf{U} = (\prod_{i=1}^M \lambda_i)^{-1/(2M)} \tilde{\mathbf{U}}$ and $\tilde{\mathbf{U}}$ is obtained from the QR decomposition in (22). Substituting such \mathbf{F} and \mathbf{B} into (12) yields the feedforward matrix \mathbf{W} . \square

From the above derivation, it is apparent that the precoder in Proposition 1, which minimizes the arithmetic MSE, also minimizes the geometric MSE. However, a precoder that minimizes the geometric MSE does not necessarily minimize the arithmetic MSE.

B. MMSE-BDFD

In this section, we consider joint transmitter–receiver design for a system based on the MMSE-BDFD. The approach is similar to that for the ZF-BDFD in the previous section, but the details are substantially different.

Recall from Section II and Fig. 2 that the received vector is $\mathbf{y} = \mathbf{H}\mathbf{F}\mathbf{s} + \mathbf{v}$. Hence, the error between $\hat{\mathbf{s}}$ and \mathbf{s} is $\mathbf{e} = \mathbf{W}\mathbf{y} - (\mathbf{B} + \mathbf{I})\mathbf{s}$. The covariance matrix of \mathbf{y} is $\mathbf{R}_{yy} = (\mathbf{H}\mathbf{F})(\mathbf{H}\mathbf{F})^H + \mathbf{R}_{vv}$, and cross-correlation matrix of \mathbf{s} and \mathbf{y} is $\mathbf{R}_{sy} = (\mathbf{H}\mathbf{F})^H = \mathbf{R}_{ys}^H$. In order to determine the minimum MSE feedforward matrix \mathbf{W}_{MMSE} , we exploit the standard first-order necessary condition for optimality known as the orthogonality principle [39], namely, $E[\mathbf{e}\mathbf{y}^H] = \mathbf{W}\mathbf{R}_{yy} - (\mathbf{B} + \mathbf{I})\mathbf{R}_{sy} = \mathbf{0}$. Therefore

$$\mathbf{W}_{\text{MMSE}} = (\mathbf{B} + \mathbf{I})\mathbf{R}_{sy}\mathbf{R}_{yy}^{-1}. \quad (23)$$

Substituting (23) into (5), and invoking the matrix inversion lemma $(\mathbf{A} + \mathbf{C}\mathbf{B}^{-1}\mathbf{D})^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{C}(\mathbf{B} + \mathbf{D}\mathbf{A}^{-1}\mathbf{C})^{-1}\mathbf{D}\mathbf{A}^{-1}$ [32], the covariance matrix of the error can be written as

$$\mathbf{R}_{ee, \text{MMSE}} = (\mathbf{B} + \mathbf{I}) (\mathbf{I} + \mathbf{F}^H \mathbf{H}^H \mathbf{R}_{vv}^{-1} \mathbf{H} \mathbf{F})^{-1} (\mathbf{B} + \mathbf{I})^H. \quad (24)$$

Our goal is to design the \mathbf{F} and \mathbf{B} to minimize the MSE subject to the power constraint. Letting $\mathbf{U} = \mathbf{B} + \mathbf{I}$, the design problem (6) can be rewritten as

$$\min_{\mathbf{F}, \mathbf{U}} \text{tr} \left(\mathbf{U} (\mathbf{I} + \mathbf{F}^H \mathbf{H}^H \mathbf{R}_{vv}^{-1} \mathbf{H} \mathbf{F})^{-1} \mathbf{U}^H \right) \quad (25a)$$

$$\text{subject to } \text{tr}(\mathbf{F}\mathbf{F}^H) \leq p_0, \text{ and } \mathbf{U} \text{ being a unit-diagonal upper triangular matrix.} \quad (25b)$$

Following the first stage outlined at the beginning of Section III, we now obtain and minimize a lower bound on the MSE. According to the trace-determinant inequality (15), we have that

$$\begin{aligned} & \text{tr} \left(\mathbf{U} (\mathbf{I} + \mathbf{F}^H \mathbf{H}^H \mathbf{R}_{vv}^{-1} \mathbf{H} \mathbf{F})^{-1} \mathbf{U}^H \right) \\ & \geq M \left| \mathbf{U} (\mathbf{I} + \mathbf{F}^H \mathbf{H}^H \mathbf{R}_{vv}^{-1} \mathbf{H} \mathbf{F})^{-1} \mathbf{U}^H \right|^{1/M} \\ & = M \left| \mathbf{I} + \mathbf{F}^H \mathbf{H}^H \mathbf{R}_{vv}^{-1} \mathbf{H} \mathbf{F} \right|^{-1/M}. \end{aligned} \quad (26)$$

Therefore, the lower bound on the MSE can be minimized by solving

$$\max_{\mathbf{F}} \left| \mathbf{I} + \mathbf{F}^H \mathbf{H}^H \mathbf{R}_{vv}^{-1} \mathbf{H} \mathbf{F} \right| \quad (27a)$$

$$\text{subject to } \text{tr}(\mathbf{F}\mathbf{F}^H) \leq p_0. \quad (27b)$$

As in the ZF case, the problem of minimizing the lower bound depends only on the transmitter. We point out that the objective in (27a) is equivalent to minimizing the geometric MSE implicit in (26). Furthermore, the logarithm of the objective in (27a) is the mutual information between the transmitter and receiver for Gaussian signals. (An analogous observation has been made in several similar contexts [9], [10], [17], [40], [52].) Hence, minimizing the lower bound on the arithmetic MSE in (26) is equivalent to maximizing the Gaussian mutual information.

Given that the problem in (27) is equivalent to maximizing the mutual information for Gaussian signals, the solution involves a ‘‘waterfilling’’ power allocation over the eigenvectors of $\mathbf{H}^H \mathbf{R}_{vv}^{-1} \mathbf{H}$, [50]. More formally, the solution depends on a parameter $r \leq K$ which is the largest integer satisfying $1/\lambda_r < (p_0 + \sum_{j=1}^r \lambda_j^{-1})/r$. If we define $q = \min\{r, M\}$, then the following set of precoders⁴ minimize the lower bound [50], $\mathbf{F} = \tilde{\mathbf{V}}_q [\mathbf{\Phi} \ \mathbf{0}_{q \times (M-q)}] \mathbf{\Psi}$, where $\mathbf{\Phi}$ is a $q \times q$ diagonal matrix with diagonal elements satisfying

$$|\phi_{ii}|^2 = \frac{1}{q} \left(p_0 + \sum_{j=1}^q \lambda_j^{-1} \right) - \lambda_i^{-1} \quad (28)$$

⁴If $M = K$ and $r = K$, or if $\lambda_q > \lambda_{q+1}$, this set is the set of all precoders that minimize the lower bound.

and Ψ is an arbitrary $M \times M$ unitary matrix.⁵ In that case, the minimal value of the lower bound on the MSE generated by (26) and (27) is

$$\bar{e}_{\text{MMSE}}^2 \geq q^{q/M} \left(p_0 + \sum_{j=1}^q \lambda_j^{-1} \right)^{-q/M} \prod_{j=1}^q \lambda_j^{-1/M} \quad (29)$$

which is independent of our design parameters \mathbf{F} and \mathbf{B} .

Moving to the second stage of our general approach, we now determine a transceiver that achieves the minimized lower bound in (29). For ease of exposition, we define $\check{\Phi} = [\Phi \ \mathbf{0}_{q \times (M-q)}]$. Substituting $\mathbf{F} = \check{\mathbf{V}}_q \check{\Phi} \Psi$ into (24) and (25a), the arithmetic MSE is $\text{tr}(\mathbf{R}_{ee, \text{MMSE}})/M$, where

$$\mathbf{R}_{ee, \text{MMSE}} = \mathbf{U} \Psi^H (\mathbf{I}_M + \check{\Phi}^T \check{\Lambda}_q \check{\Phi})^{-1} \Psi \mathbf{U}^H. \quad (30)$$

Using the trace-determinant inequality (15), for the MSE to achieve its minimized lower bound, we must choose \mathbf{U} and Ψ so that $\mathbf{R}_{ee, \text{MMSE}} = \check{\sigma}_e^2 \mathbf{I}$, where $\check{\sigma}_e^2 = q^{q/M} (p_0 + \sum_{j=1}^q \lambda_j^{-1})^{-q/M} \prod_{j=1}^q \lambda_j^{-1/M}$. That is, a system of the form in (28) achieves the minimized lower bound on the MSE in (30) if and only if we can find $\check{\mathbf{U}} = (1/\check{\sigma}_e) \mathbf{U}$ and unitary matrices Ψ and \mathbf{Q} so that

$$(\mathbf{I}_M + \check{\Phi}^T \check{\Lambda}_q \check{\Phi})^{1/2} \Psi = \mathbf{Q} \check{\mathbf{U}}. \quad (31)$$

According to Lemma 1, there exists a unitary matrix Ψ such that the QR decomposition of $(\mathbf{I}_M + \check{\Phi}^T \check{\Lambda}_q \check{\Phi})^{1/2} \Psi$ has an upper triangular “R-factor” with diagonal elements all equal to $|(\mathbf{I}_M + \check{\Phi}^T \check{\Lambda}_q \check{\Phi})^{1/2} \Psi|^{1/(2M)}$. This unitary matrix can be obtained by applying the algorithm in Appendix I to $(\mathbf{I}_M + \check{\Phi}^T \check{\Lambda}_q \check{\Phi})^{1/2}$. We summarize this result in the following proposition.

Proposition 2: The mean-square error $\text{tr}(\mathbf{R}_{ee})/M$ for a block-by-block transceiver with an MMSE-BDFD achieves its minimized lower bound (29) when the precoder $\mathbf{F} = \check{\mathbf{V}}_q [\Phi \ \mathbf{0}_{q \times (M-q)}] \Psi_{\text{MMSE}}$, where Φ satisfies (28), and Ψ_{MMSE} is obtained by applying the algorithm in Appendix I to $(\mathbf{I}_M + \check{\Phi}^T \check{\Lambda}_q \check{\Phi})^{1/2}$. The corresponding feedback matrix $\mathbf{B} = \mathbf{U} - \mathbf{I}$, where \mathbf{U} is the unit-diagonal upper triangular matrix $\mathbf{U} = \check{\sigma}_e \check{\mathbf{U}}$ and $\check{\mathbf{U}}$ is obtained from the QR decomposition in (31). Substituting such \mathbf{F} and \mathbf{B} into (23) yields the feedforward matrix \mathbf{W} . \square

As was the case for the ZF-BDFD in Section III-A, the precoder in Proposition 2, which minimizes the arithmetic MSE, lies within the set of precoders that minimize the geometric MSE, but a precoder chosen arbitrarily from the set of precoders that minimize the geometric MSE does not necessarily minimize the arithmetic MSE. This observation provides a connection between the proposed design and an earlier design for a more general overlapping block transmission system in which the transmitter was designed to minimize the geometric MSE [52]. In the context of the block-by-block transmission schemes that we have considered, the design in [52] corresponds to

choosing $\Psi = \mathbf{I}_M$ rather than choice of $\Psi = \Psi_{\text{MMSE}}$ in Proposition 2. While the choice of $\Psi = \mathbf{I}_M$ results in a system that minimizes the geometric MSE, it does not minimize the arithmetic MSE in the general case. In addition, the SINR for each element of the block may be different. In contrast, the choice of $\Psi = \Psi_{\text{MMSE}}$ minimizes the geometric MSE and the arithmetic MSE, and provides an equal SINR for each element of the block.

The choice of Ψ also has an impact on the nature of coding strategies for approaching the capacity of the block-by-block transmission system. From the discussion following (27), it is evident that the Gaussian mutual information is maximized by choosing $M = r$ and employing a transmitter matrix of the form $\mathbf{F} = \check{\mathbf{V}}_r \Phi \Psi$, where Φ satisfies (28) and Ψ is an arbitrary $r \times r$ unitary matrix. Since the MMSE-BDFD is a “canonical” receiver⁶ for Gaussian signals [9], [10], [23], this suggests that by using sufficiently powerful codes, reliable communication at rates approaching the capacity of the block transmission system can be achieved by employing any \mathbf{F} of this form and the MMSE-BDFD [9], [10], [23]. The choice $\Psi = \mathbf{I}_r$ results in a “vector coding” scheme [10], [29], [30], [36], [41] in which the feedback component of the MMSE-BDFD is inactive; i.e., $\mathbf{B} = \mathbf{0}$. Vector coding induces an equivalent system with r parallel Gaussian subchannels, each with a possibly different SNR ρ_i . (Standard DMT modulation schemes [5], [8] are a class of vector coding schemes.) Therefore, one can approach the capacity of the block transmission scheme by choosing the code for the i th element of the block to be one that approximates the ideal Gaussian code of rate $b_i = \log_2(1 + \rho_i)$ bits per channel use. (Such approximations will often involve the selection of a constellation for each element of the block.) The choice $\Psi = \Psi_{\text{MMSE}}$ results in a system in which the feedback component of the MMSE-BDFD is active, and the inputs to the decision device are uncorrelated and have identical SINRs ρ . Since the MMSE-BDFD is a canonical receiver, this suggests that one can also approach the capacity of the block transmission system by employing an independent instance of the same approximation of the ideal Gaussian code of rate $b = \log_2(1 + \rho)$ for each element of the block. The MMSE-BDFD used when $\Psi = \Psi_{\text{MMSE}}$ is more complicated to implement than the linear detector of the vector coding scheme because of the need to compute the feedback signal. However, the vector coding approach requires the design (and implementation) of (up to) r codes, one for each element of the block, whereas the proposed design requires the design of only one code.

IV. BIT ERROR RATE PERFORMANCE

In this section, we show that the (\mathbf{F}, \mathbf{B}) pairs designed in Section III to minimize the arithmetic MSE also minimize the (dominant components of the uncoded) BER of a block transmission system with uniform bit loading at moderate-to-high block SNRs. We define the average BER of the detected signal

⁵The rank of the resulting product $\mathbf{H}\mathbf{F}$ is q , and hence if M were a design variable rather than a parameter of the problem, a natural choice for M would be $M = r$.

⁶The term “canonical” is used to denote the fact that in the absence of error propagation, employing an MMSE-BDFD in place of the optimal detector does not reduce the achievable data rate [9], [10]. Methods for exploiting this property of the MMSE-BDFD were described in [24] and [48].

to be the average of the probability of error of each element of the block; i.e.,

$$P_e = \frac{1}{M} \sum_{i=1}^M P_{e,i} \quad (32)$$

where $P_{e,i}$ denotes the BER of the i th symbol s_i . For ease of exposition, we will deal with the ZF and MMSE-BDFDs separately. We will begin with the case of the ZF-BDFD.

A. ZF-BDFD

For the ZF-BDFD⁷ and for square⁸ QAM signaling with $2b_i$ bits per symbol, if all the previous decisions are correct, $P_{e,i}$ is closely approximated⁹ by [7]

$$P_{e,i} \approx \tilde{P}_{e,i} = \alpha_i \operatorname{erfc}(\sqrt{\beta_i \rho_{i,\text{ZF}}}) + \zeta_i \operatorname{erfc}(3\sqrt{\beta_i \rho_{i,\text{ZF}}}) \quad (33)$$

where $\operatorname{erfc}(x) = (2/\sqrt{\pi}) \int_x^\infty e^{-z^2} dz$ is the error function complement, $\rho_{i,\text{ZF}}$ is the decision point SNR for the i th symbol in the block, $\alpha_i = (\sqrt{4^{b_i}} - 1)/(b_i \sqrt{4^{b_i}})$, $\beta_i = 3b_i/(4^{b_i} - 1)$, and $\zeta_i = (\sqrt{4^{b_i}} - 2)/(b_i \sqrt{4^{b_i}})$. Hence

$$P_e \approx \tilde{P}_e = \frac{1}{M} \sum_{i=1}^M \tilde{P}_{e,i}.$$

Under the assumption that all the previous symbols were correctly detected, we have that

$$\rho_{i,\text{ZF}} = \frac{E[s_i^2]}{E[|\hat{s}_i - s_i|^2]} \quad (34)$$

and under our assumptions that $E[\mathbf{ss}^H] = \mathbf{I}$ and $E[\mathbf{sv}^H] = \mathbf{0}$, this expression simplifies to

$$\rho_{i,\text{ZF}} = \frac{1}{[\mathbf{R}_{ee,\text{ZF}}]_{ii}}. \quad (35)$$

Therefore, the average BER can be closely approximated by

$$P_e \approx \tilde{P}_e = \frac{1}{M} \sum_{i=1}^M \alpha_i \operatorname{erfc}\left(\sqrt{\beta_i/[\mathbf{R}_{ee,\text{ZF}}]_{ii}}\right) + \zeta_i \operatorname{erfc}\left(3\sqrt{\beta_i/[\mathbf{R}_{ee,\text{ZF}}]_{ii}}\right). \quad (36)$$

Since our precoders generate equal decision point SNRs for each element of the block, we will assume uniform bit-loading in the remainder of this section, and therefore we will drop the element index i in α_i, β_i and ζ_i . When $[\mathbf{R}_{ee,\text{ZF}}]_{ii} < 2\beta/3$,

⁷We implicitly assume that $\operatorname{rank}(\mathbf{H}) \geq M$ so that the ZF-BDFD exists.

⁸For notational simplicity, we have restricted our attention to square QAM constellations. The extension to rectangular QAM constellations can be derived in a straightforward manner using the BER expressions in [7] and [53].

⁹In the case of QPSK signaling, the expression in (33), in which $\zeta_i = 0$, is exact.

which corresponds to moderate-to-high SNRs, \tilde{P}_e is a convex function of $[\mathbf{R}_{ee}]_{ii}$, [12], [13], [36]. By applying Jensen's inequality [11] to (36), we obtain the following lower bound on the average BER:

$$\tilde{P}_e \geq \alpha \operatorname{erfc}\left(\sqrt{\beta M/\operatorname{tr}(\mathbf{R}_{ee,\text{ZF}})}\right) + \zeta \operatorname{erfc}\left(3\sqrt{\beta M/\operatorname{tr}(\mathbf{R}_{ee,\text{ZF}})}\right). \quad (37)$$

Equality in (37) holds if and only if the diagonal elements of $\mathbf{R}_{ee,\text{ZF}}$ are equal.

Equation (37) exposes an intriguing relationship between the (arithmetic) MSE and the BER. Since minimizing $\operatorname{tr}(\mathbf{R}_{ee,\text{ZF}})$ simultaneously minimizes both terms in the summation on the right-hand side of (37), minimizing the lower bound on \tilde{P}_e in (37) is equivalent to minimizing the MSE; i.e., it is equivalent to minimizing $\operatorname{tr}(\mathbf{R}_{ee,\text{ZF}})$. Therefore, the lower bound on \tilde{P}_e achieves its minimum value if the MSE is minimal. However, for the actual \tilde{P}_e to achieve its lower bound [i.e., for (37) to hold with equality], the diagonal elements of $\mathbf{R}_{ee,\text{ZF}}$ must be identical.¹⁰ Fortunately, the design proposed in Proposition 1 results in $\mathbf{R}_{ee,\text{ZF}} = \sigma_e^2 \mathbf{I}$, and hence the proposed design, which minimizes the (arithmetic) MSE of a ZF-BDFD, also minimizes the BER of the ZF-BDFD at moderate-to-high SNRs, in the sense that it minimizes \tilde{P}_e in (36).

B. MMSE-BDFD

The analysis of the previous section can be extended to the case of the MMSE-BDFD if the residual intrablock interference on each element of the block is approximated by a Gaussian random variable. For large block sizes, this approximation is (almost surely) sufficiently accurate for all but the last few elements of the block (see [25], [38], and [54]), and hence it is appropriate for our analysis. In order to account for the bias in the MMSE-BDFD (e.g., [9]), we can express the BER as a function of the decision point SINR of the i th element of the block [9], [10], [36]

$$\rho_{i,\text{MMSE}} = \frac{1}{[\mathbf{R}_{ee,\text{MMSE}}]_{ii}} - 1. \quad (38)$$

(Note that $0 \leq [\mathbf{R}_{ee,\text{MMSE}}]_{ii} < 1$.) By replacing $\rho_{i,\text{ZF}}$ in (33) by $\rho_{i,\text{MMSE}}$, the BER of the MMSE-BDFE can be approximated by

$$P_e \approx \tilde{P}_e = \frac{1}{M} \sum_{i=1}^M \alpha_i \operatorname{erfc}\left(\sqrt{\beta_i([\mathbf{R}_{ee,\text{MMSE}}]_{ii})^{-1} - 1}\right) + \zeta_i \operatorname{erfc}\left(3\sqrt{\beta_i([\mathbf{R}_{ee,\text{MMSE}}]_{ii})^{-1} - 1}\right). \quad (39)$$

As was the case for the ZF-BDFD, this function is convex in $[\mathbf{R}_{ee,\text{MMSE}}]_{ii}$ when $[\mathbf{R}_{ee,\text{MMSE}}]_{ii}$ is below a (reasonably large)

¹⁰The alternative analysis in [47] generates a related observation.

threshold [6], [36], and hence for a system in which uniform bit loading is applied, Jensen's inequality can be used to show that

$$\tilde{P}_e \geq \alpha \operatorname{erfc} \left(\sqrt{\beta(M/\operatorname{tr}(\mathbf{R}_{ee, \text{MMSE}}) - 1)} \right) + \zeta \operatorname{erfc} \left(3 \sqrt{\beta(M/\operatorname{tr}(\mathbf{R}_{ee, \text{MMSE}}) - 1)} \right) \quad (40)$$

with equality holding when the diagonal elements of $\mathbf{R}_{ee, \text{MMSE}}$ are equal. Hence, using similar arguments to those used in the case of the ZF-BDFD, the design proposed in Proposition 2, which minimizes the arithmetic MSE of the MMSE-BDFD and results in $\mathbf{R}_{ee, \text{MMSE}} = \check{\sigma}_e^2 \mathbf{I}$, also minimizes the BER of the MMSE-BDFD at moderate-to-high SNRs, in the sense that it minimizes \tilde{P}_e in (39).¹¹

V. PERFORMANCE ANALYSIS

In Section IV, it was shown that the precoders that we designed in Section III (essentially) minimize the BER of the BDFD, under the assumption that the decisions that are fed back in the receiver are correct. It can also be shown (see Appendix II) that under the same assumption, the optimized system for an MMSE-BDFD provides a lower BER than the optimized system for a ZF-BDFD, and that each optimized BDFD system provides a lower BER than the optimized system for the corresponding linear detector; see [6], [12], and [36]. That said, an incorrect decision in a BDFD can make it more likely that subsequent errors will occur by feeding back incorrect decisions. This may lead to error propagation across the block. (Recall that error propagation between blocks is explicitly avoided in block-by-block communication systems.) A standard bound on the probability of error of a conventional decision feedback equalizer in the presence of error propagation is a simple multiple of the probability of error in the absence of error propagation [16]. This suggests that the systems designed in Section III should perform well in the presence of error propagation. (A bound that is sometimes tighter [1] generates similar insight.) In this section, we seek to verify these suggestions by analyzing, via simulation, the (uncoded) BER performance of the system when error propagation may occur.

We will consider two communication scenarios: zero-padded block transmission [41], [42], [44] through a (quasi-static) scalar finite impulse response (FIR) frequency-selective fading channel that is constant over the length of the block; and transmission through a narrow-band (i.e., frequency-flat) multiple antenna fading channel with at least as many receive antennas as transmit antennas [18]. In the first scenario, the channel matrix \mathbf{H} is a tall, lower triangular, Toeplitz matrix, but in the second scenario \mathbf{H} does not possess any deterministic structure. We will evaluate the average BER performance of various transceivers for these channels in the presence of additive white Gaussian noise at the receiver; i.e., $\mathbf{R}_{vv} = \sigma^2 \mathbf{I}$. We will plot the BER performance curves as a function of the (system) SNR,

which we define as being the ratio of the transmitted energy per symbol to the noise variance; i.e., $(p_0/M)/\sigma^2$.

In addition to the transceivers we designed for the ZF-BDFD and MMSE-BDFD in Section III, for which the precoders are denoted by $\mathbf{F}_{\text{OPT-ZF-BDFD}}$ and $\mathbf{F}_{\text{OPT-MMSE-BDFD}}$, respectively, when $M = K$, we will also consider the direct transmission scheme, for which the precoder is

$$\mathbf{F}_{\text{I}} = \sqrt{p_0/M} \mathbf{I}_M \quad (41)$$

and the discrete Fourier transform (DFT) precoded scheme, for which the precoder is

$$\mathbf{F}_{\text{DFT}} = \sqrt{p_0/M} \mathbf{D}^H \quad (42)$$

where \mathbf{D} is the normalized $M \times M$ DFT matrix. For the precoders in (41) and (42), the receiver matrices \mathbf{B} and \mathbf{W} are chosen according to the (separate) design procedures for the ZF-BDFD and MMSE-BDFD in [44]. (Note that the precoders in the direct and DFT schemes are channel independent.) For all these precoders, we provide BER curves for the idealized detector, in which the decisions that are fed back are correct, and for the practical detector, in which the actual decisions are fed back (and hence error propagation may occur).

In order to assess the extent of the performance gains (derived in Appendix II) of the optimized BDFD systems over the optimized system for the corresponding linear detector, we will include the performance of systems with linear ZF and MMSE detection and precoders designed so that the BER at moderate-to-high block SNRs is minimized [6], [12], [36]. Using the notational conventions in Sections II and III, in particular the ordered eigendecomposition $\mathbf{H}^H \mathbf{R}_{vv}^{-1} \mathbf{H} = \mathbf{V} \mathbf{\Lambda} \mathbf{V}^H$, a minimum BER precoder for the linear ZF detector is [12]

$$\mathbf{F}_{\text{OPT-ZF-L}} = \sqrt{p_0/\operatorname{tr}(\tilde{\mathbf{\Lambda}}_M^{-1/2})} \tilde{\mathbf{V}}_M \tilde{\mathbf{\Lambda}}_M^{-1/4} \mathbf{D} \quad (43)$$

and one for the linear MMSE detector is [6], [36]

$$\mathbf{F}_{\text{OPT-MMSE-L}} = \tilde{\mathbf{V}}_k [\mathbf{\Upsilon} \quad \mathbf{0}_{k \times (M-k)}] \mathbf{D} \quad (44)$$

where the integer $k = \min\{\ell, M\}$, where ℓ is the largest integer such that

$$\lambda_\ell^{-1/2} \left(\sum_{j=1}^{\ell} \lambda_j^{-1/2} \right) - \sum_{j=1}^{\ell} \lambda_j^{-1} < p_0$$

and $\mathbf{\Upsilon}$ is a $k \times k$ diagonal matrix with diagonal elements satisfying

$$|v_{ii}|^2 = \left(\frac{p_0 + \sum_{j=1}^k \lambda_j^{-1}}{\sum_{j=1}^k \lambda_j^{-1/2}} \right) \lambda_i^{-1/2} - \lambda_i^{-1}.$$

A. Scalar Frequency-Selective Fading Channel

In this section, we consider the case of zero-padded block transmission through a (quasi-static) scalar FIR frequency-selective fading channel. In this case, the direct transmission

¹¹Note that if $M > r$, then $\operatorname{rank}(\mathbf{F}) < M$ and hence the lower bound on the BER in (40) will be quite high. If M were a design variable, rather than a parameter of the problem, reducing the symbol rate to $M = r$ would result in a substantial reduction in the error rate of the optimized system.

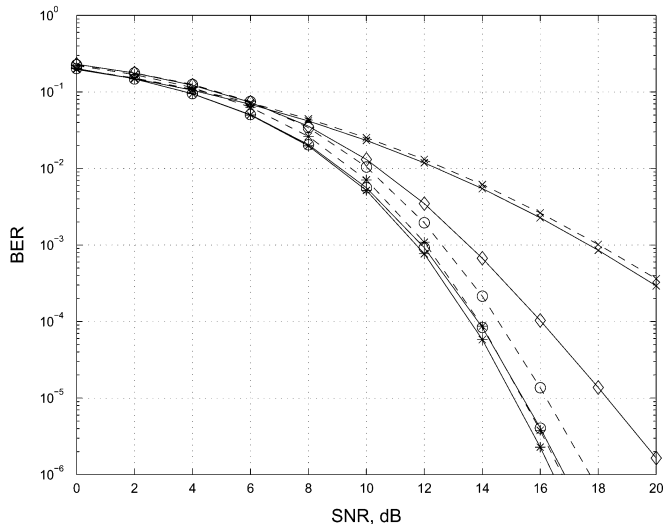


Fig. 3. Average BER performance of the ZF-BDFD for the various precoders and the linear ZF detector with its optimal precoder in the scalar frequency-selective fading channel scenario in Section V-A. The solid curves denote performance achieved in the absence of error propagation, and the dashed curves incorporate the effects of error propagation. *: optimized scheme $\mathbf{F}_{\text{OPT-ZF-BDFD}}$; o: direct (SCZP) \mathbf{F}_{I} ; x: DFT (ZP-OFDM), \mathbf{F}_{DFT} ; \diamond : optimized linear ZF scheme $\mathbf{F}_{\text{OPT-ZF-L}}$.

scheme in (41) is sometimes referred to as the single-carrier zero-padded (SCZP) scheme [49], and the DFT precoded scheme is sometimes called the zero-padded OFDM (ZP-OFDM) scheme [34]. We consider a scenario in which the channel is of length $L + 1 = 5$ and L zeros are appended to each block of channel symbols \mathbf{u} . The symbol block \mathbf{s} is of length $M = 16$, and we consider square precoders \mathbf{F} . (Hence, $K = 16$ and $P = K + L = 20$.) Each element of \mathbf{s} is an independently selected symbol from the 4-QAM constellation, with each constellation point being equally likely. In Fig. 3 we plot the BER for the ZF-BDFD transceivers, averaged over 10 000 channel realizations. (In the optimized designs, the transceiver was redesigned for each channel realization.) For each channel realization, the tap coefficients were generated independently from a zero-mean circular complex Gaussian distribution and then normalized so that the impulse response had unit energy. It is clear from the solid curves in Fig. 3 that in the absence of error propagation, the design proposed in Proposition 1 performs better than all the other transmission schemes,¹² although the SNR gain over the direct transmission (SCZP) scheme is rather small (around 0.5 dB at a BER of 10^{-4}). Furthermore, the dashed curves demonstrate that this performance advantage is maintained in the presence of error propagation. In particular, the performance of the proposed scheme in the presence of error propagation is as good as the performance of the SCZP scheme in the absence of error propagation. The combination of the DFT transmitter (ZP-OFDM) and the ZF-BDFD performs poorly at moderate-to-high block SNRs. In fact, it is apparent from Fig. 3 that the linear ZF detection scheme with its minimum BER precoder [12] performs better than the combination of the DFT transmitter

¹²As predicted by the derivation in Section IV-A, the proposed precoder performs better than all other transmission schemes for each realization of the channel.

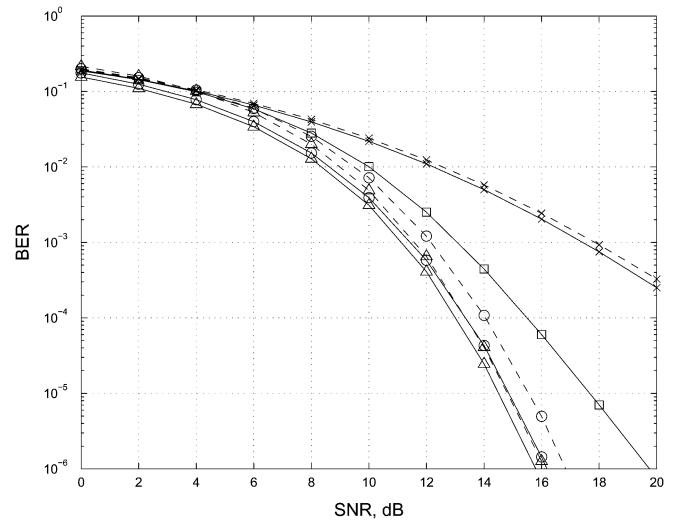


Fig. 4. Average BER performance of the MMSE-BDFD for the various precoders and the linear MMSE detector with its optimal precoder in the scalar frequency-selective fading channel scenario in Section V-A. The solid curves denote performance achieved in the absence of error propagation, and the dashed curves incorporate the effects of error propagation. \triangle : optimized scheme $\mathbf{F}_{\text{OPT-MMSE-BDFD}}$; o: direct (SCZP) \mathbf{F}_{I} ; x: DFT (ZP-OFDM) \mathbf{F}_{DFT} ; \square : optimized linear MMSE scheme $\mathbf{F}_{\text{OPT-MMSE-L}}$.

and the ZF-BDFD. However, as predicted by the analysis in Appendix II, the optimal precoder for the ZF-BDFD provides substantially better performance than the combination of the linear ZF detector and its minimum BER precoder.

The corresponding results for the MMSE-BDFD are provided in Fig. 4. The same trends are observed and the SNR gains are at least as large. Furthermore, the improved BER performance of the optimized MMSE-BDFD system over the optimized ZF-BDFD system predicted by the analysis in Appendix II can be clearly observed. In both Figs. 3 and 4, the performance of the optimized scheme in the absence of error propagation is indistinguishable from the corresponding bound on \hat{P}_e in Section IV; see (37) and (40), respectively.

An interesting by-product of the above performance evaluation is the good performance provided by the (channel-independent) direct transmission scheme (SCZP). In fact, the SCZP scheme is an optimal channel independent transmission scheme for systems that employ linear [31] or maximum likelihood [49], [55] detection, and it approaches the diversity-multiplexing tradeoff for a standard class of FIR channels as the block length grows [21]. These desirable characteristics are due, in part, to the fact that the SCZP scheme preserves the good conditioning properties implicit in the tall lower-triangular Toeplitz structure of the channel matrix.

B. Multiple Antenna Systems

In this example, we consider the case of narrow-band transmission over a multiple antenna channel with at least as many receiver antennas as transmitter antennas. In this scenario, the combination of the direct transmission scheme and a BDFD is sometimes referred to as (uncoded) V-BLAST with a (fixed-order) “nulling and cancelling” receiver [4], [18], [20]. We consider a standard Rayleigh model for the channel in which the

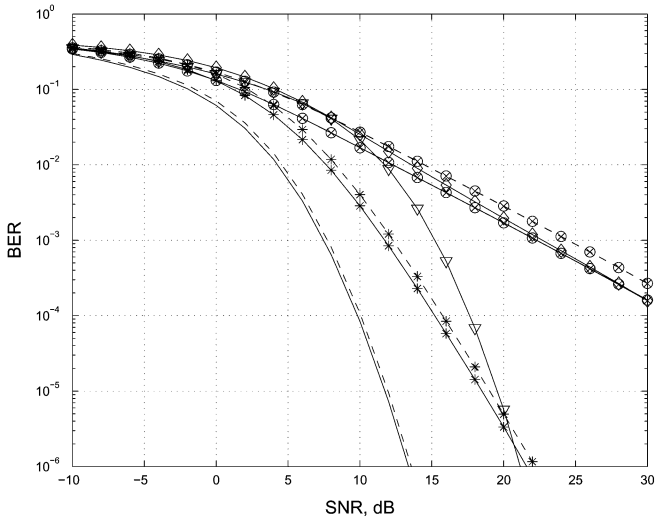


Fig. 5. Average BER performance of the ZF-BDFD for the various precoders and the linear ZF detector with its optimal precoder in the narrow-band multiple antenna scenario in Section V-B with three transmitter antennas, three receiver antennas, and $M = 3$ symbols per block. The solid curves denote performance achieved in the absence of error propagation, and the dashed curves incorporate the effects of error propagation. *: optimized scheme $\mathbf{F}_{\text{OPT-ZF-BDFD}}$; \circ : direct \mathbf{F}_i ; \times : DFT \mathbf{F}_{DFT} ; \diamond : optimized linear ZF scheme $\mathbf{F}_{\text{OPT-ZF-L}}$; ∇ : OSTBC. For later reference, the unmarked curves are for the optimized scheme with $M = 2$.

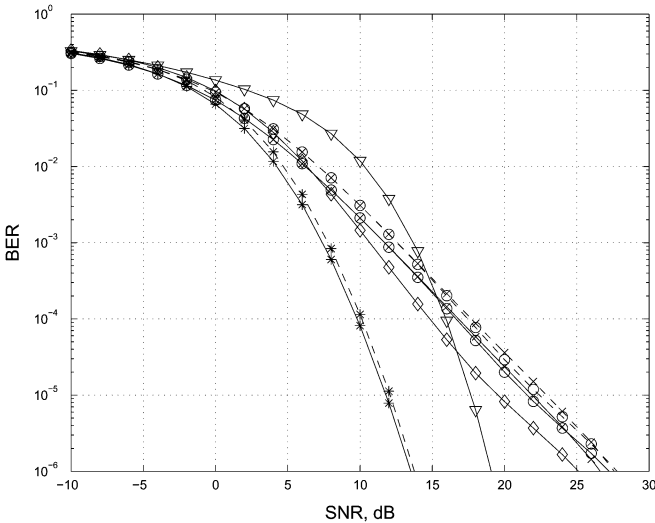


Fig. 6. Average BER performance of the ZF-BDFD for the various precoders and the linear ZF detector with its optimal precoder in the narrow-band multiple antenna scenario in Section V-B with three transmitter antennas and four receiver antennas. The legend is the same as that in Fig. 5.

paths between antennas are modeled as independent zero-mean circular Gaussian random variables of unit variance.

We will focus on scenarios with $K = 3$ transmitter antennas and $P = 3$ or 4 receiver antennas in which $M = K = 3$ symbols are transmitted per channel use. Each element of \mathbf{s} is an independent and equally likely 4-QAM symbol. Therefore, the bit rate of each scheme is 6 bits-per-channel-use (bpcu). In Figs. 5 and 6, we plot the average BER performance over 10 000 channel realizations of the various transmission schemes with the ZF receivers, and in Figs. 7 and 8 we plot the corresponding curves for the MMSE receivers. While most of the basic trends from the case of the scalar frequency-selective channels are

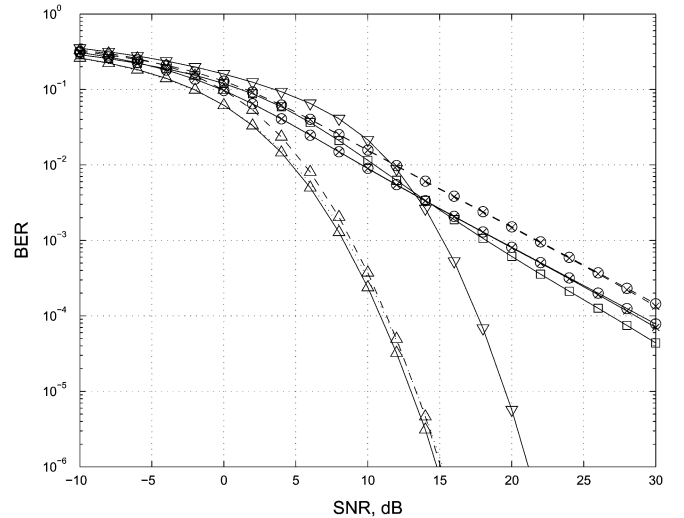


Fig. 7. Average BER performance of the MMSE-BDFD for the various precoders and the linear MMSE detector with its optimal precoder in the narrow-band multiple antenna scenario in Section V-B with three transmitter antennas and three receiver antennas. The solid curves denote performance achieved in the absence of error propagation, and the dashed curves incorporate the effects of error propagation. \triangle : optimized scheme $\mathbf{F}_{\text{OPT-MMSE-BDFD}}$; \circ : direct \mathbf{F}_i ; \times : DFT \mathbf{F}_{DFT} ; \square : optimized linear ZF scheme $\mathbf{F}_{\text{OPT-MMSE-L}}$; ∇ : OSTBC. The dotted curve denotes the lower bound on \hat{P}_e in (40).

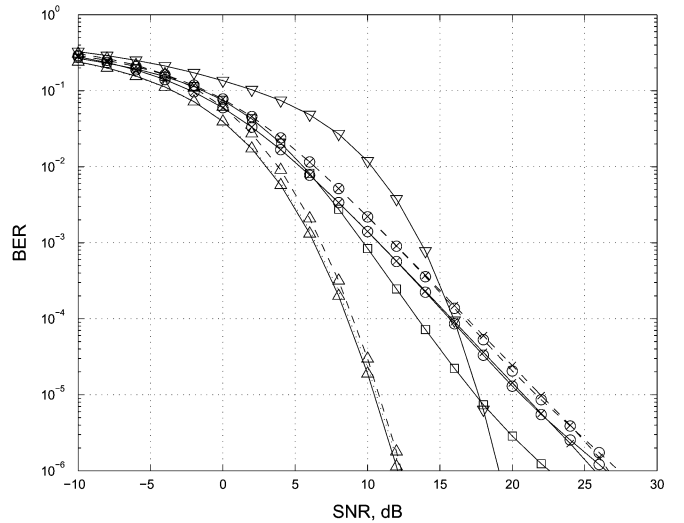


Fig. 8. Average BER performance of the MMSE-BDFD for the various precoders and the linear MMSE detector with its optimal precoder in the narrow-band multiple antenna scenario in Section V-B with three transmitter antennas, four receiver antennas. The legend is the same as that in Fig. 7.

maintained in the multiple antenna scenario, the performance advantages of the precoders designed in Section III are much greater. (The SNR gains are of the order of 6–8 dB at a BER of 10^{-4} .) This can be attributed to the fact that the channel matrix \mathbf{H} does not possess any deterministic structure. In particular, the probability of encountering a channel matrix that does not have M substantial singular values is not negligible. Since the proposed designs provide significantly better performance in those cases, the average performance is also substantially improved.

As expected, the performance of the optimized ZF-BDFD scheme in the absence of error propagation in Figs. 5 and 6 is equal to the lower bound on \hat{P}_e in (37). (Recall that we are using 4-QAM signaling.) However, in the MMSE-BDFD case,

the lower bound on \tilde{P}_e in (40) is distinguishable from the simulated BER in the absence of error propagation. This is due to the fact that the block size ($M = 3$) is small enough for the inaccuracy of the Gaussian approximation of the residual interference to result in a discernible difference between the BER and \tilde{P}_e . That said, even for this small block size, \tilde{P}_e is an accurate approximation of the BER in the absence of error propagation.

A few other features of Figs. 5–8 are worthy of note. First, the average performances of the direct and DFT transmission schemes are essentially the same. This is to be expected because the statistics of \mathbf{H} are unitarily invariant. Second, the increase in the diversity provided by the channel when using $P = 4$ receiver antennas rather than $P = 3$ is clear from the different slopes of the BER curves at high SNR. Finally, the performance advantage of the optimized MMSE-BDFD scheme over the optimized ZF-BDFD scheme is significant in the case of $P = 4$ receiver antennas and is substantial in the case of $P = 3$. The performance advantage of the optimized MMSE-BDFD scheme is due, in part, to the fact the power allocated to the first M eigenmodes of $\mathbf{H}^H \mathbf{R}_{vv}^{-1} \mathbf{H}$ depends on the corresponding eigenvalues. In particular, weak eigenmodes might not be allocated any power at all. In contrast, the optimized ZF-BDFD scheme allocates power uniformly over these eigenmodes. The larger performance advantage of the optimized MMSE-BDFD scheme in the case of $P = 3$ is due to the larger probability of encountering a channel matrix such that $\mathbf{H}^H \mathbf{R}_{vv}^{-1} \mathbf{H}$ does not have $M = 3$ significant eigenvalues.

For reference, we have included the performance of a standard orthogonal space–time block coding (OSTBC) scheme in Figs. 5–8. (Like the direct and DFT transmission schemes, OSTBC schemes were designed to be applied without knowledge of the channel at the transmitter.) We have used the (symbol) rate 3/4 code in [19] (which is a simplified version of that in [45]), and hence in order to achieve a bit rate of 6 bpcu, a natural choice for the underlying constellation is 256-QAM. (We assume that the channel is constant for the four channel uses that are required to transmit the codewords.) As expected, at high SNR, the OSTBC scheme provides better BER performance than that direct transmission (V-BLAST) scheme. However, the proposed precoder (which exploits knowledge of the channel) provides substantially better performance when $P = 4$ receiver antennas are employed, and when $P = 3$ and the MMSE-BDFD receiver is used.

When $P = 3$ receiver antennas are employed and the ZF-BDFD is used, the OSTBC scheme performs better than the optimized scheme at high SNRs. This does not contradict the optimality of the proposed transceiver design, because the values of m , k and P , and the structure of the channel matrix, are different for the OSTBC scheme.¹³ The good performance of the OSTBC scheme at high SNRs is simply a manifestation of the tradeoff between error rate (achievable diversity) and symbol rate in multiple antenna fading channels without outer codes [46]. (That tradeoff is related to the fundamental diversity-multiplexing tradeoff [58].) The symbol rate of the OSTBC scheme is significantly lower than that of the proposed

scheme.¹⁴ Hence, in the range of SNRs in which noise dominates the error performance, the proposed scheme provides better performance than the OSTBC scheme, but in the SNR range in which the channel condition dominates the error performance, the OSTBC scheme provides better performance. To illustrate that point, in Fig. 5 we plotted with unmarked curves the performance of the proposed ZF-BDFD scheme with a symbol rate of $M = 2$ (as distinct from the scheme with $M = 3$ described above). In order to maintain a bit rate of 6 bpcu, the elements of \mathbf{s} were taken, in an independent and equally likely fashion, from an 8-QAM constellations, and for consistency, the SNR was defined to be $(p_0/3)/\sigma^2$. Over the range of SNRs considered, the performance of the proposed ZF-BDFD scheme with $M = 2$ is substantially better than that of the OSTBC scheme, with SNR gains of over 7 dB.

VI. CONCLUSION

In this paper, we have jointly designed the precoder and the feedback matrix of a block-by-block transmission scheme equipped with a zero-forcing or MMSE intrablock decision feedback detector. The designs minimize the arithmetic mean of the expected squared errors at the decision point, under the standard assumption that the previous symbols were correctly detected. The covariance matrix of the minimized error is white, and hence the proposed designs also minimize the (dominant components of the) bit error rate of a uniformly bit-loaded transmission system. In our simulations, the proposed systems performed significantly better than standard precoding systems and retained their performance advantages in the presence of error propagation. In the case of the MMSE-BDFD, the proposed design also maximizes the Gaussian mutual information. Since the MMSE-BDFD is a “canonical” receiver [9], [10], [23], this suggests that by using the proposed transceiver design, one can approach the capacity of the block transmission system using (independent instances of) the same (Gaussian) code for each element of the block.

APPENDIX I

ALGORITHM FOR LEMMA 1

To state the algorithm succinctly, we make the following definitions: $g = (\prod_{k=1}^M \gamma_k^2)^{1/M}$; $[\mathbf{S}]_{\cdot k}$ denotes the k th column of \mathbf{S} and $s_{\ell k}$ denotes its elements; \mathbf{Z}_k denotes the first k columns of \mathbf{S} and \mathbf{Z}_k^\perp denotes its orthogonal complement; $\mathcal{P}_{\mathbf{A}} = \mathbf{I} - \mathbf{A}(\mathbf{A}^H \mathbf{A})^{-1} \mathbf{A}^H$. The recursion will be based on the $(M - k) \times (M - k)$ matrix

$$\mathbf{A}^{(k)} = (\mathbf{\Gamma} \mathbf{Z}_k^\perp)^H \mathcal{P}_{(\mathbf{\Gamma} \mathbf{Z}_k^\perp)} \mathbf{\Gamma} \mathbf{Z}_k^\perp. \quad (45)$$

For convenience, we assume that the elements of $\mathbf{\Gamma}$ are arranged in nonincreasing order. The algorithm proceeds as follows.

- 1) Initialization: Set $k = 1$. An explicit solution for the first column of \mathbf{S} is $s_{11} = \sqrt{(g - \gamma_M^2)/(\gamma_1^2 - \gamma_M^2)}$, $s_{M1} = \sqrt{(\gamma_1^2 - g)/(\gamma_1^2 - \gamma_M^2)}$, $s_{\ell 1} = 0$ for $\ell = 2, 3, \dots, M - 1$.

¹³In this example, the channel matrix for the OSTBC scheme is $\mathbf{I}_4 \otimes \mathbf{H}$, where \otimes denotes the Kronecker product and \mathbf{H} is the channel matrix for the other schemes. The corresponding block sizes are $P = 12$, $K = 12$, and $M = 3$.

¹⁴In particular, in four consecutive channel uses, the proposed scheme transmits $4M = 12$ symbols, whereas the OSTBC scheme transmits only three symbols.

- 2) Construct $\mathbf{A}^{(k)}$ in (45) and its eigendecomposition, $\mathbf{A}^{(k)} = \mathbf{V}^{(k)} \mathbf{\Lambda}^{(k)} (\mathbf{V}^{(k)})^H$.
- 3) Set the $(k+1)$ th column of \mathbf{S} to be

$$[\mathbf{S}]_{\cdot k+1} = \mathbf{Z}_k^\perp \mathbf{V}^{(k)} \mathbf{y}^{(k)},$$

where

$$y_1^{(k)} = \sqrt{(g - \lambda_{M-k}^{(k)}) / (\lambda_1^{(k)} - \lambda_{M-k}^{(k)})},$$

$$y_{M-k}^{(k)} = \sqrt{(\lambda_1^{(k)} - g) / (\lambda_1^{(k)} - \lambda_{M-k}^{(k)})}, y_\ell^{(k)} = 0$$

for $\ell = 2, 3, \dots, M - k - 1$.

- 4) Increment k . If $k \leq M - 2$ return to 2). Otherwise, set $[\mathbf{S}]_{\cdot M} = \mathbf{Z}_{M-2}^\perp \mathbf{V}^{(M-2)} \mathbf{y}^{(M-1)}$, where

$$y_1^{(M-1)} = -\sqrt{(g - \lambda_2^{(M-2)}) / (\lambda_1^{(M-2)} - \lambda_2^{(M-2)})},$$

$$y_2^{(M-1)} = \sqrt{(\lambda_1^{(M-2)} - g) / (\lambda_1^{(M-2)} - \lambda_2^{(M-2)})}.$$

APPENDIX II

ANALYTIC PERFORMANCE COMPARISONS

It was shown in Section IV that the precoders designed in Section III achieve the minimized value of the lower bound on \tilde{P}_e ; see (37) and (40). Therefore, the relative BER performance of the optimized ZF-BDFD and MMSE-BDFD systems in the absence of error propagation can be determined by simply comparing the optimal values of the MSE $\bar{e}^2 = \text{tr}(\mathbf{R}_{ee})/M$. (A preliminary version of this appendix appeared in [33], and related results on the MSEs of conventional decision feedback equalizers appear in [2, ch. 8].)

In order to ensure that the ZF systems exist, we will assume that $\text{rank}(\mathbf{H}) \geq M$, and to simplify the comparisons, we will also assume that the transmitted power p_0 is large enough that $q = M$ in (28) for the MMSE-BDFD and $\ell = M$ in (44) for the linear MMSE detector.¹⁵ Proposition 1 states that the minimum value of the MSE for a ZF-BDFD system is

$$\bar{e}_{\text{OPT-ZF-BDFD}}^2 = \frac{M}{p_0} |\tilde{\mathbf{\Lambda}}_M|^{-1/M}$$

and Proposition 2 states that the minimum value of the MSE for an MMSE-BDFD system is

$$\bar{e}_{\text{OPT-MMSE-BDFD}}^2 = \frac{M}{p_0 + \text{tr}(\tilde{\mathbf{\Lambda}}_M^{-1})} |\tilde{\mathbf{\Lambda}}|^{-1/M}.$$

Since $\tilde{\mathbf{\Lambda}}_M$ is positive definite, $\bar{e}_{\text{OPT-MMSE-BDFD}}^2 < \bar{e}_{\text{OPT-ZF-BDFD}}^2$, and hence, in the absence of error propagation, the optimized MMSE-BDFD system will provide a

¹⁵The assumption that $\text{rank}(\mathbf{H}) \geq M$ ensures that there is a threshold value for p_0 above which $q = M$ and $\ell = M$.

lower BER than the optimized ZF-BDFD system. While it is intuitively obvious that for a given precoder, the MMSE-BDFD will provide a lower MSE than the ZF-BDFD, in the case of optimized precoders, this lower MSE leads directly to a lower BER.

The analysis of Section IV remains valid for systems with linear detectors, so long as the constraint $\mathbf{B} = \mathbf{0}$ is enforced. Therefore, we can compare the BER performance of an optimized BDFD system with that of the system that is optimized for the corresponding linear detector by simply comparing their minimum MSEs. The minimum MSE of a system with a linear ZF detector is [12]

$$\begin{aligned} \bar{e}_{\text{OPT-ZF-L}}^2 &= \frac{1}{Mp_0} \left(\text{tr} \left(\tilde{\mathbf{\Lambda}}_M^{-1/2} \right) \right)^2 \\ &\geq \frac{M}{p_0} |\tilde{\mathbf{\Lambda}}_M|^{-1/M} = \bar{e}_{\text{OPT-ZF-BDFD}}^2 \end{aligned}$$

where we have used the trace-determinant inequality (15). Therefore, in the absence of error propagation, the optimized system for the ZF-BDFD will provide a lower BER than the optimized system for the linear ZF detector. Similarly, the minimum MSE of a system with a linear MMSE detector is [6], [36]

$$\begin{aligned} \bar{e}_{\text{OPT-MMSE-L}}^2 &= \frac{1}{M \left(p_0 + \text{tr} \left(\tilde{\mathbf{\Lambda}}_M^{-1} \right) \right) - \left(\text{tr} \left(\tilde{\mathbf{\Lambda}}_M^{-1/2} \right) \right)^2} \\ &\quad \times \left(\text{tr} \left(\tilde{\mathbf{\Lambda}}_M^{-1/2} \right) \right)^2 \\ &> \frac{M}{p_0 + \text{tr} \left(\tilde{\mathbf{\Lambda}}_M^{-1} \right)} |\tilde{\mathbf{\Lambda}}_M|^{-1/M} \\ &= \bar{e}_{\text{OPT-MMSE-BDFD}}^2 \end{aligned}$$

and hence the optimized system for the MMSE-BDFD provides a lower BER than the optimized system for the linear MMSE detector. As observed in [6], $\bar{e}_{\text{OPT-MMSE-L}}^2 \leq \bar{e}_{\text{OPT-ZF-L}}^2$, and hence the optimized system for the linear MMSE detector provides a lower BER than the optimized system for the linear ZF detector.

ACKNOWLEDGMENT

The authors gratefully acknowledge the assistance of S. S. Chan of McMaster University in the preparation of the simulation results reported in Section V-A and of Q. Meng of McMaster University in the preparation of Appendix II.

REFERENCES

- [1] S. A. Altekar and N. C. Beaulieu, "Upper bounds on the error probability of decision feedback equalization," *IEEE Trans. Inf. Theory*, vol. 39, pp. 145–156, Jan. 1993.
- [2] J. R. Barry, E. A. Lee, and D. G. Messerschmitt, *Digital Communication*, 3rd ed. Boston, MA: Kluwer, 2004.
- [3] C. A. Belfiore and J. H. Park Jr., "Decision feedback equalization," *Proc. IEEE*, vol. 67, pp. 1143–1156, Aug. 1979.
- [4] E. Biglieri, G. Taricco, and A. Tulino, "Decoding space-time codes with BLAST architectures," *IEEE Trans. Signal Process.*, vol. 50, pp. 2547–2552, Oct. 2002.

- [5] J. A. C. Bingham, "Multicarrier modulation for data transmission: An idea whose time has come," *IEEE Commun. Mag.*, pp. 5–14, May 1990.
- [6] S. S. Chan, T. N. Davidson, and K. M. Wong, "Asymptotically minimum BER linear block precoders for MMSE equalization," *Proc. Inst. Elect. Eng. Commun.*, vol. 151, pp. 297–304, Aug. 2004.
- [7] 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, July 2002.
- [8] J. S. Chow, J. C. Tu, and J. M. Cioffi, "Performance evaluation of a multichannel transceiver system for ADSL and VDSL receivers," *IEEE J. Sel. Areas Commun.*, vol. 9, pp. 909–919, Aug. 1991.
- [9] J. M. Cioffi, G. P. Dudevoir, M. V. Eyuboglu, and G. D. Forney, "MMSE decision-feedback equalizers and coding—Part I: Equalization results; Part II: Coding results," *IEEE Trans. Commun.*, vol. 43, pp. 2582–2604, Oct. 1995.
- [10] J. M. Cioffi and G. D. Forney, "Generalized decision-feedback equalization for packet transmission with ISI and Gaussian noise," in *Communications, Computation, Control and Signal Processing*, A. Paulraj, V. Roychowdhury, and C. Schaper, Eds. Boston, MA: Kluwer, 1997.
- [11] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York: Wiley, 1991.
- [12] Y. W. Ding, T. N. Davidson, Z.-Q. Luo, and K. M. Wong, "Minimum BER block precoders for zero-forcing equalization," *IEEE Trans. Signal Process.*, vol. 51, pp. 2410–2423, Sep. 2003.
- [13] Y. W. Ding, T. N. Davidson, and K. M. Wong, "On improving the BER performance of rate-adaptive block transceivers, with applications to DMT," in *Proc. IEEE Global Commun. Conf.*, San Francisco, CA, Dec. 2003.
- [14] A. Duel-Hallen, "Decorrelating decision-feedback multiuser detector for synchronous code-division multiple-access channels," *IEEE Trans. Commun.*, vol. 41, pp. 285–290, Feb. 1993.
- [15] —, "A family of multiuser decision-feedback detectors for asynchronous code-division multiple-access channels," *IEEE Trans. Commun.*, vol. 43, pp. 421–434, Feb./Mar./Apr. 1995.
- [16] D. L. Duttweiler, J. E. Mazo, and D. G. Messerschmitt, "An upper bound on the error probability in decision-feedback equalization," *IEEE Trans. Inf. Theory*, vol. IT-20, pp. 490–497, Jul. 1974.
- [17] D. Falconer and G. Foschini, "Theory of minimum mean-square-error QAM systems employing decision feedback equalization," *Bell Syst. Tech. J.*, vol. 52, pp. 1821–1849, Dec. 1973.
- [18] G. J. Foschini, G. D. Golden, R. A. Valenzuela, and P. W. Wolniansky, "Simplified processing for high spectral efficiency wireless communication employing multi-element arrays," *IEEE J. Sel. Areas Commun.*, vol. 17, pp. 1841–1852, Nov. 1999.
- [19] G. Ganesan and P. Stoica, "Space-time block codes: A maximum SNR approach," *IEEE Trans. Inf. Theory*, vol. 47, pp. 1650–1656, May 2001.
- [20] G. Gini and J. M. Cioffi, "On the relation between V-BLAST and GDFE," *IEEE Commun. Lett.*, vol. 5, pp. 364–366, Sept. 2001.
- [21] L. Gropop and D. N. C. Tse, "Diversity/multiplexing tradeoff in ISI channels," in *Proc. Int. Symp. Information Theory*, Chicago, June 2004, p. 97.
- [22] T. Guess, "Optimal sequences for CDMA with decision-feedback receivers," *IEEE Trans. Inf. Theory*, vol. 49, pp. 886–900, Apr. 2003.
- [23] T. Guess and M. K. Varanasi, "An information theoretic framework for deriving canonical decision-feedback receivers in Gaussian channels," *IEEE Trans. Inf. Theory*, vol. 51, pp. 173–187, Jan. 2005.
- [24] —, "A new successively decodable coding technique for intersymbol-interference channels," in *Proc. Int. Symp. Information Theory*, Sorrento, Italy, Jun. 2000.
- [25] D. Guo, S. Verdù, and L. K. Rasmussen, "Asymptotic normality of linear multiuser receiver outputs," *IEEE Trans. Inf. Theory*, vol. 48, pp. 3080–3095, Dec. 2002.
- [26] 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.
- [27] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge, UK: Cambridge University Press, 1985.
- [28] G. K. Kaleh, "Channel equalization for block transmission system," *IEEE J. Sel. Areas Commun.*, vol. 13, pp. 110–121, Jan. 1995.
- [29] S. Kasturia, J. T. Aslanis, and J. M. Cioffi, "Vector coding for partial response channels," *IEEE Trans. Inf. Theory*, vol. 36, pp. 741–762, Jul. 1990.
- [30] J. W. Lechleider, "The optimum combination of block codes and receivers for arbitrary channels," *IEEE Trans. Commun.*, vol. 38, pp. 615–621, May 1990.
- [31] Y. Lin and S. M. Phoong, "BER minimized OFDM systems with channel independent precoders," *IEEE Trans. Signal Process.*, vol. 51, pp. 2369–2380, Sep. 2003.
- [32] J. R. Magnus and H. Neudecker, *Matrix Differential Calculus with Applications in Statistics and Econometrics*. New York: Wiley, 1988.
- [33] Q. Meng, J.-K. Zhang, and K. M. Wong, "Block data transmission: A comparison of performance for the MBER precoder designs," presented at the IEEE Sensor Array Multichannel Signal Processing Workshop, Sitges, Spain, Jul. 2004.
- [34] B. Muquet, Z. Wang, G. B. Giannakis, M. de Courville, and P. Duhamel, "Cyclic prefixing or zero padding for wireless multicarrier transmissions?," *IEEE Trans. Commun.*, vol. 50, pp. 2136–2148, Dec. 2002.
- [35] F. D. Neesser and J. L. Massey, "Proper complex random processes with applications to information theory," *IEEE Trans. Inf. Theory*, vol. 39, pp. 1293–1302, Jul. 1993.
- [36] D. P. Palomar, J. M. Cioffi, and M. A. Lagunas, "Joint Tx-Rx beamforming design for multicarrier MIMO channels: A unified framework for convex optimization," *IEEE Trans. Signal Process.*, vol. 51, pp. 2381–2401, Sep. 2003.
- [37] B. Picinbono, "On circularity," *IEEE Trans. Signal Process.*, vol. 42, pp. 3474–3482, Dec. 1994.
- [38] H. V. Poor and S. Verdù, "Probability of error in MMSE multiuser detection," *IEEE Trans. Inf. Theory*, vol. 43, pp. 858–871, May 1997.
- [39] J. G. Proakis, *Digital Communications*, 4th ed. New York: McGraw Hill, 2001.
- [40] J. Salz, "Optimum mean-square decision feedback equalization," *Bell Syst. Tech. J.*, vol. 52, pp. 1341–1373, Oct. 1973.
- [41] A. Scaglione, S. Barbarossa, and G. B. Giannakis, "Filterbank transceivers optimizing information rate in block transmissions over dispersive channels," *IEEE Trans. Inf. Theory*, vol. 45, pp. 1019–1032, Apr. 1999.
- [42] A. Scaglione, G. B. Giannakis, and S. Barbarossa, "Redundant filterbank precoders and equalizers part I: Unification and optimal designs," *IEEE Trans. Signal Process.*, vol. 47, pp. 1988–2005, Jul. 1999.
- [43] A. Scaglione, P. Stoica, S. Barbarossa, G. B. Giannakis, and H. Sampath, "Optimal designs for space-time linear precoders and decoders," *IEEE Trans. Signal Process.*, vol. 50, pp. 1051–1064, May 2002.
- [44] A. Stamouli, G. B. Giannakis, and A. Scaglione, "Block FIR decision-feedback equalizers for filterbank precoded transmissions with blind channel estimation capabilities," *IEEE Trans. Commun.*, vol. 49, pp. 69–83, Jan. 2001.
- [45] V. Tarokh, H. Jafarkhani, and A. R. Calderbank, "Space-time block codes from orthogonal designs," *IEEE Trans. Inf. Theory*, vol. 45, pp. 1456–1467, Jul. 1999.
- [46] B. Varadarajan and J. R. Barry, "The rate-diversity trade-off for linear space-time codes," in *Proc. IEEE Veh. Tech. Conf.*, vol. 1, Vancouver, Canada, Sep. 2002, pp. 67–71.
- [47] M. K. Varanasi, "Decision feedback multiuser detection: A systematic approach," *IEEE Trans. Inf. Theory*, vol. 45, pp. 110–121, Jan. 1999.
- [48] M. K. Varanasi and T. Guess, "Optimum decision feedback multiuser equalization with successive decoding achieves the total capacity of the Gaussian multiple-access channel," in *Proc. Asilomar Conf. Signals, Systems, Computers*, Monterey, CA, Nov. 1997.
- [49] Z. Wang, X. Ma, and G. B. Giannakis, "OFDM or single-carrier block transmissions?," *IEEE Trans. Commun.*, vol. 52, pp. 380–394, Mar. 2004.
- [50] H. S. Witsenhausen, "A determinant maximization problem occurring in the theory of data communication," *SIAM J. Appl. Math.*, vol. 29, pp. 515–522, Nov. 1975.
- [51] F. Xu, T. N. Davidson, and K. M. Wong, "Design of block-by-block transceivers with successive detection," in *Proc. IEEE Workshop Signal Processing Adv. Wireless Commun.*, Rome, Italy, Jun. 2003.
- [52] J. Yang and S. Roy, "Joint transmitter-receiver optimization for multi-input multi-output systems with decision feedback," *IEEE Trans. Inf. Theory*, vol. 40, pp. 1334–1347, Sep. 1994.
- [53] D. Yoon and K. Cho, "General bit error probability of rectangular quadrature amplitude modulation," *Electron. Lett.*, vol. 38, pp. 131–133, Jan. 2002.
- [54] J. Zhang, E. K. P. Chong, and D. N. C. Tse, "Output MAI distributions of linear MMSE multiuser receivers in DS-SSMA systems," *IEEE Trans. Inf. Theory*, vol. 47, pp. 1128–1144, Mar. 2001.
- [55] J.-K. Zhang, T. N. Davidson, and K. M. Wong, "Optimal precoder for block transmission over frequency-selective fading channels," *Proc. Inst. Elect. Eng. Commun.*, vol. 154, pp. 463–468, Aug. 2005.
- [56] J.-K. Zhang, A. Kavčić, X. Ma, and K. M. Wong, "Design of unitary precoders for ISI channels," in *Proc. Int. Conf. Acoust., Speech, Signal Processing*, Orlando, FL, Apr. 2002, pp. 2265–2268.
- [57] J.-K. Zhang, A. Kavčić, and K. M. Wong, "Equal diagonal QR decomposition and its application to precoder design for successive-cancellation detection," *IEEE Trans. Inf. Theory*, vol. 51, pp. 154–172, Jan. 2005.

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



Fang Xu received the B.Eng. and M.Eng. degrees in electrical engineering from Huazhong University of Science and Technology, China and the M.A.Sc. degree in electrical engineering from McMaster University, Hamilton, ON, Canada, in 2003.

Currently she is a Technical Support Specialist with Raytheon Canada Limited.



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

He is currently an Associate 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. His research interests are in signal processing, communications, and control,

with current activity focused on signal processing for digital communication systems. He has held research positions with the Communications Research Laboratory at McMaster University, the Adaptive Signal Processing Laboratory at UWA, and the Australian Telecommunications Research Institute at Curtin University of Technology, Perth.

Dr. Davidson received the 1991 J. A. Wood Memorial Prize from the University of Western Australia and the 1991 Rhodes Scholarship for Western Australia.



Jian-Kang Zhang received the B.S. degree in mathematics and information science from Shaanxi Normal University, Xi'an, China, the M.S. degree in mathematics from Northwest University, Xi'an, and the Ph.D. degree in electrical engineering from Xidian University, Xi'an.

He is now with the Department of Electrical and Computer Engineering, McMaster University, Hamilton, ON, Canada. His research interests include multirate filterbanks, wavelet and multiwavelet transforms and their applications, number theory transform, and fast algorithms. His current research focuses on multiscale wavelet transform, random matrices, precoding and space-time coding for block transmissions, and multicarrier and wide-band wireless communication systems.



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

He joined the Transmission Division, Plessey Telecommunications Research Ltd., U.K., in 1969. In 1970 he was on leave from Plessey pursuing post-graduate 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 joined McMaster University, Hamilton, ON, Canada, where he has been a Professor since 1985 and was Chairman of the Department of Electrical and Computer Engineering in 1986–1987 and 1988–1994. He was a Visiting Professor at the Department of Electronic Engineering, Chinese University of Hong Kong, from 1997 to 1999. At present, he holds the NSERC-Mitel Professorship of Signal Processing and is Director of the Communication Technology Research Centre, McMaster University. He is also serving another five-year term as Chair of the Department. His research interest is in signal processing and communication theory and he has published over 200 papers in the area.

Prof. Wong is a Fellow of the Institution of Electrical Engineers, the Royal Statistical Society, and the Institute of Physics. He received the Institution of Electrical Engineers Overseas Premium for the best paper in 1989. He was an Associate Editor of the *IEEE TRANSACTIONS ON SIGNAL PROCESSING* from 1996 to 1998 and Chair of the Sensor Array and Multi-channel Signal Processing Technical Committee of the Signal Processing Society in 2002–2004. He received a medal presented by the International Biographical Centre, Cambridge, U.K., for his "outstanding contributions to the research and education in signal processing" in 2000. He has appeared in *Outstanding People of the 20th Century* and *2000 Outstanding Intellectuals of the 20th Century*.