Minimum BER Block Precoders for Zero-Forcing Equalization

Yanwu Ding, Timothy N. Davidson, *Member, IEEE*, Zhi-Quan Luo, *Senior Member, IEEE*, and Kon Max Wong, *Fellow, IEEE*

***Abstract*—In this paper, we determine the linear precoder that minimizes the bit error rate (BER) at moderate-to-high signal-to-noise ratios (SNRs) for block transmission systems with zero-forcing (ZF) equalization and threshold detection. The design is developed for the two standard schemes for eliminating inter-block interference, viz, zero padding (ZP) and cyclic prefix (CP). We show that both the ZP minimum BER precoder and the CP minimum BER precoder provide substantially lower error rates than standard block transmission schemes, such as orthogonal frequency division multiplexing (OFDM). The corresponding SNR gains can be on the order of several decibels. We also show that the CP minimum BER precoder can be obtained by a two-stage modification of the water-filling discrete multitone modulation (DMT) scheme in which the diagonal water-filling power loading is replaced by a full matrix consisting of a diagonal minimum mean square error power loading matrix post multiplied by a discrete Fourier transform (DFT) matrix.**

***Index Terms*—Block precoding, cyclic prefix, discrete multitone modulation (DMT), minimum bit error rate, orthogonal frequency division multiplexing (OFDM).**

I. Introduction

In THE transmission of digital data over dispersive media, channel induced inter-symbol interference (ISI) is a major performance limiting factor [14]. To mitigate such an effect, it is often helpful to transmit information-bearing data in equal-size blocks [7]. Examples of block based communication systems include important multicarrier (MC) systems such as orthogonal frequency division multiplexing (OFDM) [1], [6], which has been selected as the standard modulation scheme for terrestrial digital audio and video broadcasting in Europe, as well as discrete multitone (DMT) modulation [3], [15], which has been adopted for high-bit-rate digital subscriber line (HDSL) and asymmetric digital subscriber line (ADSL) systems. Recently, a broad class of linear block-by-block transmission schemes, which includes DMT and OFDM as special cases, has been studied in detail in [18], [19], and [21]. The block-based linear transmitter that maximizes the information rate was derived in [18]. However, to obtain the performance predicted by this type of design, we may need to employ sophisticated encoding and decoding structures, which may lead to unacceptable receiver complexity or communication latency. The design of block based linear transmitters for block-by-block linear receivers was studied in [19]. In particular, the transmitters that minimize the mean square error of the equalized symbols were derived for both zero-forcing (ZF) and minimum mean square error (MMSE) equalizers, under the assumption that the channel state information is known. For the case of ZF equalization, this idea has been extended to scenarios in which only the second-order statistics of the channel are known [12]. While the design of transmitters based on the minimum MSE criteria is mathematically tractable and results in simple realizations of the optimal precoder, and although such transmitters perform reasonably well in practice, the MMSE criterion does not guarantee minimum bit error rate (MBER). In this paper, we consider the design of an MBER linear precoder for a system with a block-by-block linear ZF receiver and threshold detection. We examine the bit error probability function and show that it is convex in the appropriate design parameters at moderate-to-high signal-to-noise ratios (SNRs). By exploiting the convexity, we obtain a closed-form expression for the MBER precoder. A simple test that determines whether the SNR is sufficiently high for the design to be optimal is provided, along with a natural scheme for dropping the low-gain subchannels to ensure that the SNR over the remaining channels is sufficiently high. Our simulation studies demonstrate that the performance improvement provided by the MBER precoder can be substantial and can be in excess of several decibels in SNR gain at a BER of around 10^{-4}.

Manuscript received April 12, 2002; revised February 21, 2003. This work was supported in part by the National Science and Engineering Research Council of Canada. The work of the third author is also supported by the Canada Research Chairs Program. The associate editor coordinating the review of this paper and approving it for publication was Dr. Helmut Boelcskei.

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

Digital Object Identifier 10.1109/TSP.2003.815387

II. Block-by-Block Transmission

In this paper, we employ the generalized block-by-block transceiver model developed in [19]. This model encompasses many modern communication systems, including OFDM and DMT. The system model is shown in Fig. 1. Related schemes that involve overlapped block-based transmission are also available [16], [19], [23]–[25], but we will focus on the block-by-block case here.

A. System Description

To describe our system model, we let

$$\boldsymbol{s}(n) \triangleq [s(nM),\ s(nM+1),\ \ldots,\ s(nM+M-1)]^T$$

denote the nth block of data to be transmitted. For each block of M data symbols, $P > M$ symbols are transmitted across the

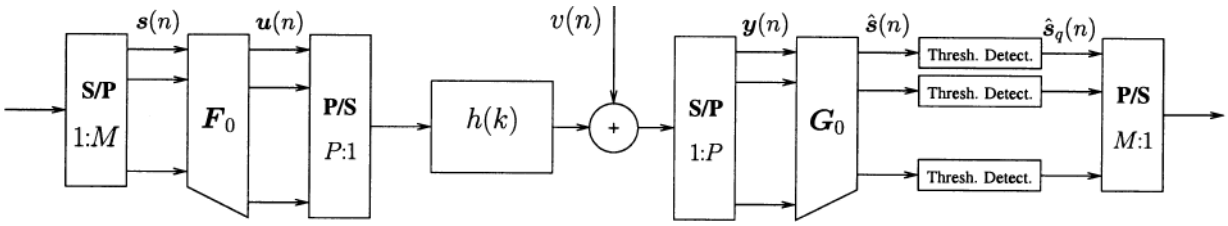


Fig. 1. Discrete-time baseband equivalent model for the block-by-block transceiver model developed in [19]. The $\mathbf{S/P}$ and $\mathbf{P/S}$ boxes denote serial-to-parallel and parallel-to-serial conversion, respectively. (The corresponding dimensions are noted.) The precoder is denoted by $\boldsymbol{F}_0$, the channel by $h(k)$, the additive noise by $v(n)$, and the equalizer by $\boldsymbol{G}_0$.

channel. This redundancy is the key to avoiding interblock interference (IBI) at the receiver, as we will see below. The vector of transmitted symbols

$$\boldsymbol{u}(n) \triangleq [u(nP),\ u(nP+1),\ \ldots,\ u(nP+P-1)]^T$$

is given by

$$\boldsymbol{u}(n) = \boldsymbol{F}_0 \boldsymbol{s}(n) \tag{1}$$

where $\boldsymbol{F}_0$ is $P \times M$. If

$$\boldsymbol{y}(n) \triangleq [y(nP),\ y(nP+1),\ \ldots,\ y(nP+P-1)]^T \tag{2}$$

denotes the nth block of receiver inputs, then the vector of equalized data symbols $\hat{\boldsymbol{s}}(n)$ is given by

$$\hat{\boldsymbol{s}}(n) = \boldsymbol{G}_0 \boldsymbol{y}(n) \tag{3}$$

where $\boldsymbol{G}_0$ is $M \times P$. The receiver inputs are the result of convolution with the channel impulse response $h(k)$ and the corruption by additive noise. Hence

$$y(n) = \sum_{k=-\infty}^{\infty} h(k)u(n-k) + v(n).$$

This can be written in vectorized form as

$$\boldsymbol{y}(n) = \sum_{k=-\infty}^{\infty} \boldsymbol{H}_k \boldsymbol{u}(n-k) + \boldsymbol{v}(n) \tag{4}$$

where the $P \times P$ matrices $\boldsymbol{H}_k$ are determined as

$$\boldsymbol{H}_k = \begin{bmatrix} h(kP) & h(kP-1) & \cdots & h(kP-P+1) \\ h(kP+1) & h(kP) & \ddots & h(kP-P+2) \\ \vdots & \ddots & \ddots & \vdots \\ h(kP+P-1) & h(kP+P-2) & \cdots & h(kP) \end{bmatrix}. \tag{5}$$

Hence, the equalized symbols can be written as

$$\hat{\boldsymbol{s}}(n) = \boldsymbol{G}_0 \sum_{k=-\infty}^{\infty} \boldsymbol{H}_k \boldsymbol{u}(n-k) + \boldsymbol{G}_0 \boldsymbol{v}(n). \tag{6}$$

B. Zero-Padded and Cyclic-Prefix Transmission

For finite impulse response (FIR) channels, (6) can be simplified by judicious choice of the block size and redundancy, as is well known in the special case of cyclic prefix-based transceivers (see, e.g., [1], [21]). To state this formally, we make the following assumptions.

- **A1)** The channel is an (at most) Lth-order FIR channel, with impulse response $h(k)$ satisfying $h(k) = 0$, for $k < 0$ and $k > L$.
- **A2)** The length of the block of transmitted symbols P is chosen so that $P \geq M + L$ and $P > 2L$.

Using (5) and these assumptions, (6) simplifies to

$$\hat{\boldsymbol{s}}(n) = \boldsymbol{G}_0 \boldsymbol{H}_0 \boldsymbol{F}_0 \boldsymbol{s}(n) + \boldsymbol{G}_0 \boldsymbol{H}_1 \boldsymbol{F}_0 \boldsymbol{s}(n-1) + \boldsymbol{G}_0 \boldsymbol{v}(n). \tag{7}$$

The IBI on the nth block of equalized symbols now comes only from the previous block, but this IBI will still limit the performance of the system at high SNR. To eliminate this IBI, we observe that since $h(k)$ is zero outside $0 \leq k \leq L$, $\boldsymbol{H}_1$ has nonzero elements only in its $L \times L$ top right submatrix. Therefore, IBI can be eliminated, irrespective of the actual impulse response of the channel, if we impose structure on $\boldsymbol{F}_0$ and $\boldsymbol{G}_0$ so that $\boldsymbol{G}_0 \boldsymbol{H}_1 \boldsymbol{F}_0 = \boldsymbol{0}$. In this paper, that structure is imposed by choosing $\boldsymbol{F}_0$ and $\boldsymbol{G}_0$ so that $\boldsymbol{F}_0 = \boldsymbol{T}\boldsymbol{F}$ and $\boldsymbol{G}_0 = \boldsymbol{G}\boldsymbol{R}$, where $\boldsymbol{F}$ and $\boldsymbol{G}$ are free parameters, and $\boldsymbol{T}$ and $\boldsymbol{R}$ are chosen such that $\boldsymbol{R}\boldsymbol{H}_1\boldsymbol{T} = \boldsymbol{0}$. There are two standard choices for $\boldsymbol{T}$ and $\boldsymbol{R}$, namely (see [21] and references therein) the following.

- **ZP)** *Zero-padded transmission:* Choose

$$\boldsymbol{T} = \boldsymbol{T}_{\text{zp}} = \begin{bmatrix} \boldsymbol{I}_{(P-L)} \\ \boldsymbol{0}_{L\times(P-L)} \end{bmatrix}$$

and

$$\boldsymbol{R} = \boldsymbol{R}_{\text{zp}} = \boldsymbol{I}_P$$

which results in $\boldsymbol{F} = \boldsymbol{F}_{\text{zp}}$ being a $(P-L) \times M$ "tall" matrix and $\boldsymbol{G} = \boldsymbol{G}_{\text{zp}}$ being an $M \times P$ "fat" matrix.

- **CP)** *Cyclic prefix transmission with removal of interfered samples at receiver:* Choose

$$\boldsymbol{T} = \boldsymbol{T}_{\text{cp}} = \begin{bmatrix} \boldsymbol{0}_{L\times(P-2L)} & \boldsymbol{I}_L \\ & \boldsymbol{I}_{(P-L)} \end{bmatrix}$$

and

$$\boldsymbol{R} = \boldsymbol{R}_{\text{cp}} = [\,\boldsymbol{0}_{(P-L)\times L} \boldsymbol{I}_{P-L}\,]$$

which results in $\boldsymbol{F} = \boldsymbol{F}_{\text{cp}}$ being a $(P-L) \times M$ "tall" matrix and $\boldsymbol{G} = \boldsymbol{G}_{\text{cp}}$ being an $M \times (P-L)$ "fat" matrix.

The vector of equalized symbols for both the zero-padding (ZP) and cyclic prefix (CP) options can be written in a common form

$$\hat{\boldsymbol{s}} = \boldsymbol{GHFs} + \boldsymbol{GRv} \tag{8}$$

where $\boldsymbol{H} = \boldsymbol{RH}_0\boldsymbol{T}$. For the ZP case, we have

$$\boldsymbol{H} = \boldsymbol{H}_{\rm zp} = \begin{bmatrix} h(0) & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ h(L) & \ddots & \ddots & 0 \\ 0 & \ddots & \ddots & h(0) \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \ddots & 0 & h(L) \end{bmatrix} \tag{9}$$

which is a $P \times (P-L)$ "tall" Toeplitz matrix. For the CP case, we have

$$\boldsymbol{H} = \boldsymbol{H}_{\rm cp} = \begin{bmatrix} h(0) & 0 & \cdots & 0 & h(L) & h(L-1) & \cdots & h(1) \\ h(1) & h(0) & 0 & \cdots & 0 & h(L) & \cdots & h(2) \\ \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots \\ 0 & \cdots & 0 & h(L) & h(L-1) & \cdots & h(1) & h(0) \end{bmatrix} \tag{10}$$

which is a $(P-L)\times(P-L)$ circulant matrix. The eigenvector matrix of such a circulant matrix is the (normalized) discrete Fourier transform (DFT) matrix [10], and advantage of this fact has been taken in the design of CP systems [1], [21]. In simple scenarios, CP schemes typically employ precoders and equalizers of the form $\boldsymbol{F}_{\rm cp} = \boldsymbol{D}_{P-L}^H \boldsymbol{\Delta}_T$ and $\boldsymbol{G}_{\rm cp} = \boldsymbol{\Delta}_R \boldsymbol{D}_{P-L}$, where $\boldsymbol{D}_N$ is the $N \times N$ (normalized) DFT matrix

$$[\boldsymbol{D}_N]_{k,\,n} = N^{-1/2} \exp(-j2\pi kn/N) \tag{11}$$

for $0 \le k,\, n \le N-1$, and $\boldsymbol{\Delta}_T$ and $\boldsymbol{\Delta}_R$ are the power loading and equalization matrices, respectively. When $M = P-L$, $\boldsymbol{\Delta}_T$ and $\boldsymbol{\Delta}_R$ are diagonal, and in that case, (8) simplifies to

$$\hat{\boldsymbol{s}}_{cp} = \boldsymbol{\Delta}_R \boldsymbol{\Delta}_H \boldsymbol{\Delta}_T \boldsymbol{s} + \boldsymbol{\Delta}_R \boldsymbol{D}_{P-L} \boldsymbol{v}_{\rm cp}. \tag{12}$$

Here, $\boldsymbol{v}_{\rm cp} = \boldsymbol{R}_{\rm cp}\boldsymbol{v}$, and $\boldsymbol{\Delta}_H$ is a diagonal matrix with ith diagonal element $H(e^{j2\pi i/(P-L)})$, $0 \le i \le P-L-1$, where $H(e^{j\omega}) = \sum_k h(k)\exp(-j\omega k)$ is the frequency response of the channel. The standard CP-OFDM scheme, in which $M = P-L$ and $\boldsymbol{\Delta}_T = \boldsymbol{I}_M$, is an example of this simple scheme. In more general scenarios, the power loading algorithm may result in $Q \le P-L$ subcarriers being allocated power, and we may transmit $M \le Q$ symbols over these subcarriers [3], [15], [21], [22]. For these scenarios, the precoder can be written in the form $\boldsymbol{F}_{\rm cp} = \boldsymbol{D}_{P-L}^H \boldsymbol{Q} \bar{\boldsymbol{\Delta}}_T \boldsymbol{\Theta}$, where $\bar{\boldsymbol{\Delta}}_T = \boldsymbol{Q}^T \boldsymbol{\Delta}_T \boldsymbol{Q}$ is a $Q \times Q$ diagonal matrix containing the nonzero (diagonal) elements of $\boldsymbol{\Delta}_T$, and $\boldsymbol{Q}$ is the corresponding $(P-L)\times Q$ selection matrix consisting of the columns of $\boldsymbol{I}_Q$ corresponding to the nonzero elements of $\boldsymbol{\Delta}_T$. The matrix $\boldsymbol{\Theta}$ is a $Q \times M$ matrix that satisfies $\boldsymbol{\Theta}^H\boldsymbol{\Theta} = \boldsymbol{I}_M$. In standard DMT schemes [3], [15], we typically have $M = Q$ and $\boldsymbol{\Theta} = \boldsymbol{I}_Q$, but in certain wireless applications, diversity gains can be obtained by choosing $Q = P-L$ and $M \le Q$ with $\boldsymbol{\Theta}$ chosen so that a certain rank condition holds [21], [22]. Although we will not impose any structure on our MBER CP precoders, our designs can be written in an analogous form, as we will show in Section V.

C. Derivation of BER

In our study of the block transmission systems described above, we will make the following assumptions in addition to Assumptions A1) and A2).

A3) During transmission, IBI has been eliminated by adopting either zero padded or cyclic-prefix-based transmission.

A4) The channel is completely known and ZF equalization is employed at the receiver. That is

$$\boldsymbol{G} = (\boldsymbol{HF})^{\dagger} \tag{13}$$

and hence $\hat{\boldsymbol{s}} = \boldsymbol{GHFs} + \boldsymbol{GRv} = \boldsymbol{s} + (\boldsymbol{HF})^{\dagger}\boldsymbol{Rv}$, where $\boldsymbol{X}^{\dagger} = \left(\boldsymbol{X}^H\boldsymbol{X}\right)^{-1}\boldsymbol{X}^H$ denotes the left pseudo-inverse of a "tall" matrix $\boldsymbol{X}$ of full column rank. A necessary condition for the left pseudo-inverse of $\boldsymbol{HF}$ to exist is that M, which is the number of columns of $\boldsymbol{F}$, is no more than rank$(\boldsymbol{H})$, [10]. (We will show below that this condition is also sufficient by constructing such an $\boldsymbol{F}$.)

A5) The transmitted symbols are equiprobable antipodal symbols (i.e., ± 1) uncorrelated with each other, i.e.,

$$E\{\boldsymbol{ss}^H\} = \boldsymbol{I}. \tag{14}$$

The antipodal assumption is made for simplicity, and our results can be extended to other constellations, such as 4-QAM/QPSK [5].

A6) The noise vector $\boldsymbol{v}$ is zero-mean, white and Gaussian, with covariance matrix

$$E\{\boldsymbol{vv}^H\} = \sigma^2\boldsymbol{I}. \tag{15}$$

We now derive an expression for the average BER of the system. For the communication system shown in Fig. 1, when $\boldsymbol{s}$ is transmitted, $\hat{\boldsymbol{s}}$, as given by (8), will be the received signal vector. The elements of this vector are then quantized by a threshold detector to obtain $\hat{\boldsymbol{s}}_q$, whose elements will be ± 1. The average BER of the detected signal is the average of the probability of error of each element of the block, i.e.,

$$P_e = \frac{1}{M}\sum_{m=1}^{M} P_{em} \tag{16}$$

where P_{em} denotes the BER of the mth symbol. Since the signal power of each data symbol is unity and the covariance matrix of the received noise is $\sigma^2\boldsymbol{GG}^H$, by following standard steps (e.g., [14]), it can be shown that the probability of the mth symbol in $\hat{\boldsymbol{s}}_q$ being in error can be written as

$$P_{em} = \frac{1}{2}\,\mathrm{erfc}\left(\frac{1}{\sqrt{2\sigma^2[\boldsymbol{GG}^H]_{mm}}}\right) \tag{17}$$

where $\mathrm{erfc}(\zeta) \triangleq (2/\sqrt{\pi})\int_{\zeta}^{\infty} e^{-z^2}\,dz$, and $[\boldsymbol{GG}^H]_{mm}$ denotes the (m, m)th element of the matrix $\boldsymbol{GG}^H$. The term $\sigma^2[\boldsymbol{GG}^H]_{mm}$ represents the noise variance in the receiver's estimate of the mth symbol of the transmitted signal vector. Substituting (17) into (16), we have

$$P_e = \frac{1}{2M}\sum_{m=1}^{M} \mathrm{erfc}\left(\frac{1}{\sqrt{2\sigma^2[\boldsymbol{GG}^H]_{mm}}}\right). \tag{18}$$

If we define $\phi(x) = \mathrm{erfc}(1/\sqrt{2\sigma^2 x})$ for $x > 0$, then

$$\frac{d^2\phi}{dx^2} = \frac{1}{\sqrt{\pi}}(2\sigma^2)^{-(1/2)}\exp\left(-\frac{1}{2\sigma^2 x}\right) \cdot\left(-\frac{3}{2}+\frac{1}{2\sigma^2 x}\right)x^{-(5/2)}. \tag{19}$$

Therefore, if $x < 1/3\sigma^2$, then $\partial^2\phi(x)/\partial x^2 > 0$. Applying this fact to (18), we determine that $\phi([\boldsymbol{GG}^H]_{mm})$ is a convex function if the noise power σ^2 is less than $1/3[\boldsymbol{GG}^H]_{mm}$. If this condition is satisfied for all m (i.e., if there is sufficiently large SNR at the receiver), then the average block BER P_e is also convex. The property of convexity is desirable in the development of a design algorithm for the precoder that minimizes the BER because any locally optimal precoder is also globally optimal. The development of such an algorithm is presented in the next section.

III. Design of the Minimum BER Precoder

We now design an optimum precoder $\boldsymbol{F}$ in (8) such that the minimum BER is achieved, subject to a bound on the transmission power. We will not impose any structure on $\boldsymbol{F}$, but in the CP case, we will show that the MBER precoder retains some of the structural features of the OFDM and DMT schemes. The transmission power is given by $\mathrm{tr}(E\{\boldsymbol{F}_0\boldsymbol{s}(\boldsymbol{F}_0\boldsymbol{s})^H\}) = \mathrm{tr}(\boldsymbol{F}_0\boldsymbol{F}_0^H)$. For ZP systems, $\mathrm{tr}(\boldsymbol{F}_0\boldsymbol{F}_0^H) = \mathrm{tr}(\boldsymbol{FF}^H)$, i.e., the transmission power is equal to the power used to transmit the data. For CP systems, the total transmission power is the power used to transmit the data plus the power used to transmit the cyclic prefix. However, the proportion of the total power used to transmit the cyclic prefix is typically small, and it is standard practice to define the transmission power as simply the power required to transmit the data [3], [9], [21]. (See [5] for further discussion.) Hence, $\mathrm{tr}(\boldsymbol{FF}^H)$ will also be used to represent the transmission power for CP systems. We can formulate the design problem as follows:

$$\begin{aligned}\min_{\boldsymbol{F}} \quad & P_e \\ \text{subject to} \quad & \mathrm{tr}(\boldsymbol{FF}^H) \le p_0\end{aligned} \tag{20}$$

where P_e is given by (18), the matrix $\boldsymbol{G}$ in (18) is the ZF equalizer in (13), which depends on both $\boldsymbol{F}$ and $\boldsymbol{H}$, and p_0 is a constant limiting the transmission power. Since P_e is convex at moderate-to-high SNRs, we can apply Jensen's inequality [4] to obtain the following lower bound on P_e:

$$\begin{aligned} P_e &= \frac{1}{2M}\sum_{m=1}^{M}\mathrm{erfc}\left(\frac{1}{\sqrt{2\sigma^2[\boldsymbol{GG}^H]_{mm}}}\right) \\ &\ge \frac{1}{2}\mathrm{erfc}\left(\frac{1}{\sqrt{\frac{2\sigma^2}{M}\sum_{m=1}^{M}[\boldsymbol{GG}^H]_{mm}}}\right) \qquad (21)\\ &= \frac{1}{2}\mathrm{erfc}\left(\sqrt{\frac{M}{2\sigma^2\mathrm{tr}(\boldsymbol{GG}^H)}}\right) \triangleq P_{e,\mathrm{LB}}. \qquad (22)\end{aligned}$$

Equality in (21) holds if and only if $[\boldsymbol{GG}^H]_{mm}$ are equal $\forall\, m \in [1, M]$. The inequality of (21) is valid only when P_e is convex, i.e., when

$$[\boldsymbol{GG}^H]_{mm} < \frac{1}{3\sigma^2}, \qquad \forall\, m \in [1, M]. \tag{23}$$

The quantity $P_{e,\mathrm{LB}}$ in (22) defines a lower bound on the BER P_e. To obtain an optimum precoder design, we will first minimize $P_{e,\mathrm{LB}}$. In doing so, we will show that a particular choice of precoder achieves this minimized lower bound.

Now, from Section II, we have learned that $\boldsymbol{F}$ is a "tall" $(P - L) \times M$ rectangular matrix. If we parameterize $\boldsymbol{F}$ in terms of its singular value decomposition [8], we have

$$\boldsymbol{F} = \boldsymbol{U}\begin{bmatrix}\boldsymbol{\Phi}\\ \boldsymbol{0}\end{bmatrix}\boldsymbol{V} \tag{24}$$

where $\boldsymbol{\Phi}$ is a $M \times M$ positive diagonal matrix, and $\boldsymbol{U}$ and $\boldsymbol{V}$ are square unitary matrices of dimension $(P - L)$ and M, respectively. From (9) and (10), the nature of the channel matrix $\boldsymbol{H}$ depends on whether ZP or CP is employed, but in either case, we can write

$$(\boldsymbol{H}^H\boldsymbol{H})^{-1} = \boldsymbol{W}\boldsymbol{\Lambda}\boldsymbol{W}^H \tag{25}$$

where $\boldsymbol{\Lambda}$ and $\boldsymbol{W}$ are, respectively, the (square) eigenvalue and eigenvector matrices of $(\boldsymbol{H}^H\boldsymbol{H})^{-1}$ and are both of dimension $(P - L)$. The eigenvalues $\{\lambda_i\}$ in $\boldsymbol{\Lambda}$, which are non-negative, are arranged in descending order. In the ZP case, $\boldsymbol{H}^H\boldsymbol{H}$ is guaranteed to be nonsingular as long as $h(k)$ is not identically zero. However, in the CP case, $\boldsymbol{H}^H\boldsymbol{H}$ will drop rank if $H(e^{j\omega})$ happens to have a zero at $\omega = 2\pi k/(P - L)$ for some integer k. For ease of exposition, we will explicitly exclude such channels in the CP case so that $\boldsymbol{H}^H\boldsymbol{H}$ will be assumed to be of full rank in both the ZP and CP cases. However, our minimum BER precoder remains valid for CP systems for which $H(e^{j2\pi k/(P-L)}) \to 0$ for some integer k because the corresponding subchannels will automatically be dropped by the subchannel dropping algorithm described in Section IV.

Given (24), for a ZF equalizer $\boldsymbol{G} = (\boldsymbol{HF})^{\dagger}$, we can write

$$\boldsymbol{GG}^H = \boldsymbol{V}^H\boldsymbol{\Phi}^{-1}\left([\,\boldsymbol{I}_M \quad \boldsymbol{0}\,]\boldsymbol{U}^H\boldsymbol{H}^H\boldsymbol{H}\boldsymbol{U}\begin{bmatrix}\boldsymbol{I}_M\\ \boldsymbol{0}\end{bmatrix}\right)^{-1}\boldsymbol{\Phi}^{-1}\boldsymbol{V}. \tag{26}$$

Now, since erfc$(\cdot)$ is a monotonically decreasing function, to minimize $P_{e,\mathrm{LB}}$ in (22), we need only minimize tr$(\boldsymbol{G}\boldsymbol{G}^H)$. For notational simplicity, we define

$$\boldsymbol{Z}(\boldsymbol{U}) = \left([\boldsymbol{I}_M \quad \boldsymbol{0}]\boldsymbol{U}^H\boldsymbol{H}^H\boldsymbol{H}\boldsymbol{U}\begin{bmatrix}\boldsymbol{I}_M \\ \boldsymbol{0}\end{bmatrix}\right)^{-1}. \tag{27}$$

As one would expect from (24), $\boldsymbol{Z}(\boldsymbol{U})$ is a function of only the first M columns of $\boldsymbol{U}$. However, our notation enhances the connections with related transmission schemes, as discussed in Section II-B. Now, the problem of minimizing $P_{e,\mathrm{LB}}$ over $\boldsymbol{F}$, subject to tr$(\boldsymbol{F}\boldsymbol{F}^H) \leq p_0$, the "convexity constraint" in (23) and $\boldsymbol{G}$ being a ZF equalizer, is equivalent to

$$\min_{\boldsymbol{U},\boldsymbol{\Phi},\boldsymbol{V}} \quad \mathrm{tr}(\boldsymbol{\Phi}^{-2}\boldsymbol{Z}(\boldsymbol{U})) \tag{28a}$$

$$\text{subject to} \quad \mathrm{tr}(\boldsymbol{\Phi}^2) \leq p_0 \tag{28b}$$

$$\text{and} \quad [\boldsymbol{V}^H\boldsymbol{\Phi}^{-1}\boldsymbol{Z}(\boldsymbol{U})\boldsymbol{\Phi}^{-1}\boldsymbol{V}]_{mm} \leq \frac{1}{3\sigma^2} \tag{28c}$$

where we have used the unitary property of $\boldsymbol{V}$ and the relation tr$(\boldsymbol{A}\boldsymbol{B}) = \mathrm{tr}(\boldsymbol{B}\boldsymbol{A})$ for dimensionally compatible matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ to simplify the objective in (28a).

The problem in (28) is quite awkward to solve directly, due to the presence of (28c), but, as we will show below, the minimization of (28a) over $\boldsymbol{U}$ and $\boldsymbol{\Phi}$ subject to (28b) has an analytic solution. Therefore, we will solve the problem in (28) by first dealing with (28c) for any $\boldsymbol{U}$ and $\boldsymbol{\Phi}$ and then solving the remaining problem. To do so, we observe that the design parameter $\boldsymbol{V}$ does not enter the objective in (28a) nor the power constraint in (28b). Rather, its role is to try to satisfy the constraint in (28c), which ensures the convexity of P_e in (18) and, hence, the validity of the lower bound in (22). For any given $\boldsymbol{U}$ and $\boldsymbol{\Phi}$, an optimal choice for $\boldsymbol{V}$ is one that maximizes the minimum constraint satisfaction (or minimizes the maximum constraint violation) in (28c), i.e., the $\boldsymbol{V}$ that minimizes the largest diagonal element on the left-hand side of (28c). Such a $\boldsymbol{V}$ is provided by the following lemma.

Lemma 1: Given an $M \times M$ positive semi-definite (symmetric) matrix $\boldsymbol{A} = \boldsymbol{\Psi}\boldsymbol{\Gamma}\boldsymbol{\Psi}^{\mathbf{H}}$, where $\boldsymbol{\Gamma}$ is a diagonal matrix whose diagonal elements are the eigenvalues of $\boldsymbol{A}$ and $\boldsymbol{\Psi}\boldsymbol{\Psi}^H = \boldsymbol{I}$, then we have the following.

i)

$$\min_{\boldsymbol{V}\boldsymbol{V}^H=\boldsymbol{I}} \max_m [\boldsymbol{V}^H\boldsymbol{A}\boldsymbol{V}]_{mm} = \frac{\mathrm{tr}(\boldsymbol{A})}{M}. \tag{29}$$

ii) The minimum value of (29) can be achieved by choosing

$$\boldsymbol{V} = \boldsymbol{\Psi}\boldsymbol{D}_M \tag{30}$$

where $\boldsymbol{D}_M$ denotes the $M \times M$ (normalized) DFT matrix.

Proof: We first prove that the right-hand side of (29) is a lower bound for the left-hand side. Then, we show that the $\boldsymbol{V}$ in (30) achieves this lower bound.

i) Since $\boldsymbol{V}$ is unitary, tr$(\boldsymbol{V}^H\boldsymbol{A}\boldsymbol{V}) = \mathrm{tr}(\boldsymbol{A})$. Furthermore, since $\boldsymbol{A}$ is positive semidefinite, the diagonal elements of $\boldsymbol{A}$ and $\boldsymbol{V}^H\boldsymbol{A}\boldsymbol{V}$ are non-negative. Given the set of all length N sequences of non-negative numbers $\{x_i\}_{i=1}^N$ that sum to y, the sequence minimizing the maximum value of x_i is $x_i = y/N$. Applying that result to the left-hand side of (29) and observing that the constraint on $\boldsymbol{V}$ may restrict the values that the diagonal elements of $\boldsymbol{V}^H\boldsymbol{A}\boldsymbol{V}$ can take on, we have that $\min_{\boldsymbol{V}\boldsymbol{V}^H=\boldsymbol{I}} \max_m [\boldsymbol{V}^H\boldsymbol{A}\boldsymbol{V}]_{mm} \geq \mathrm{tr}(\boldsymbol{A})/M$.

ii) Let $\boldsymbol{V} = \boldsymbol{\Psi}\boldsymbol{\Upsilon}$, where $\boldsymbol{\Upsilon}$ is some unitary matrix. Then, we have

$$[\boldsymbol{V}^H\boldsymbol{A}\boldsymbol{V}]_{mm} = [\boldsymbol{\Upsilon}^H\boldsymbol{\Gamma}\boldsymbol{\Upsilon}]_{mm} = \sum_{\ell=1}^{M} \gamma_\ell |\upsilon_{m\ell}|^2$$

where γ_ℓ is the ℓth diagonal element of $\boldsymbol{\Gamma}$, and $\upsilon_{m\ell}$ is the (m, ℓ)th element of $\boldsymbol{\Upsilon}$. If $\boldsymbol{\Upsilon}$ is chosen to be the (normalized) DFT matrix, i.e., if $\boldsymbol{V} = \boldsymbol{\Psi}\boldsymbol{D}_M$, then since the magnitude of each element of the DFT matrix is equal to $|d_{m\ell}|^2 = 1/M$, we have that

$$[\boldsymbol{V}^H\boldsymbol{A}\boldsymbol{V}]_{mm} = \sum_{\ell=1}^{M} \gamma_\ell \frac{1}{M} = \frac{\mathrm{tr}(\boldsymbol{A})}{M}, \quad \text{for all } 1 \leq m \leq M.$$

This completes the proof. ■

Without compromising Lemma 1, we point out that there are other matrices $\boldsymbol{V}$ that will achieve the minimum in (29), including the (normalized) inverse DFT matrix $\boldsymbol{D}_M^H$ and, if M is a power of two, the $M \times M$ normalized Hadamard matrix. All that we require for $\boldsymbol{V}$ to be optimal is for it to be unitary and for all its elements to have the same magnitude.

Lemma 1 provides the key to obtaining an analytic solution to the problem of minimizing $P_{e,\mathrm{LB}}$ in (28), as we show below. It also shows us that there is a precoder that minimizes $P_{e,\mathrm{LB}}$ and actually achieves this lower bound, as we point out later on. To derive an analytic solution to the problem in (28), we observe that by applying Lemma 1, there exists a $\boldsymbol{V}$ satisfying (28c) if and only if the optimal value of the objective is less than $M/(3\sigma^2)$. Hence, the problem in (28) can be solved in the following two stages. First, minimize (28a) over $\boldsymbol{\Phi}$ and $\boldsymbol{U}$, subject to (28b), i.e.,

$$\min_{\boldsymbol{U},\boldsymbol{\Phi}} \quad \mathrm{tr}\left(\boldsymbol{\Phi}^{-2}\boldsymbol{Z}(\boldsymbol{U})\right) \tag{31a}$$

$$\text{subject to} \quad \mathrm{tr}(\boldsymbol{\Phi}^2) \leq p_0. \tag{31b}$$

Let $\boldsymbol{\Phi}_{\mathrm{opt}}$ and $\boldsymbol{U}_{\mathrm{opt}}$ denote the optimal solution to the problem in (31). Then, there exists a complementary $\boldsymbol{V}_{\mathrm{opt}}$ that satisfies (28c) if and only if

$$\mathrm{tr}\left(\boldsymbol{\Phi}_{\mathrm{opt}}^{-2}\boldsymbol{Z}(\boldsymbol{U}_{\mathrm{opt}})\right) \leq \frac{M}{3\sigma^2}. \tag{32}$$

Furthermore, if this condition holds, one such $\boldsymbol{V}_{\mathrm{opt}}$ is $\boldsymbol{\Psi}\boldsymbol{D}_M$, where $\boldsymbol{Z}(\boldsymbol{U}_{\mathrm{opt}}) = \boldsymbol{\Psi}\boldsymbol{\Gamma}\boldsymbol{\Psi}^H$ is an eigen decomposition of $\boldsymbol{Z}(\boldsymbol{U}_{\mathrm{opt}})$. We will show below that $\boldsymbol{Z}(\boldsymbol{U}_{\mathrm{opt}})$ is diagonal, and hence, $\boldsymbol{\Psi} = \boldsymbol{I}$, and the normalized DFT matrix is an optimal choice for $\boldsymbol{V}$. With the above analysis, we have reduced the awkward design problem in (28) to the solution of the simpler optimization problem in (31) followed by the feasibility test in (32) and construction of $\boldsymbol{V}_{\mathrm{opt}}$ via Lemma 1, should (32) show that the solution to the simplified problem in (31) is feasible for the problem in (28).

All that remains is to solve the simpler optimization problem in (31). This problem is equivalent to the design of the $(P-L)\times$

M minimum MSE precoder for ZF equalization. To show that equivalence, we note that since we are using a ZF equalizer, the mean square error in $\hat{\boldsymbol{s}}$ in (8), namely $\mathrm{tr}(E\{(\hat{\boldsymbol{s}}-\boldsymbol{s})(\hat{\boldsymbol{s}}-\boldsymbol{s})^H\})$, can be written as $\sigma^2\mathrm{tr}(\boldsymbol{G}\boldsymbol{G}^H)$. Using (26) and the unitary invariance of the trace function, we have that $\mathrm{tr}(\boldsymbol{G}\boldsymbol{G}^H) = \mathrm{tr}(\boldsymbol{\Phi}^{-2}\boldsymbol{Z}(\boldsymbol{U}))$ and, hence, the equivalence. As shown in the Appendix, the precoder that minimizes the squared error in $\hat{\boldsymbol{s}}$ for ZF equalization has

$$\boldsymbol{\Phi}_{\mathrm{opt}} = \sqrt{\frac{p_0}{\mathrm{tr}\left(\boldsymbol{\Lambda}_M^{1/2}\right)}}\,\boldsymbol{\Lambda}_M^{1/4} \tag{33}$$

$$\boldsymbol{U}_{\mathrm{opt}} = [\,\boldsymbol{W}_M \quad \boldsymbol{W}_M^{\perp}\,] \tag{34}$$

where $\boldsymbol{\Lambda}_M = \mathrm{diag}(\lambda_{P-L-M+1}, \ldots, \lambda_{P-L})$ consists of the *smallest* M eigenvalues of $(\boldsymbol{H}^H\boldsymbol{H})^{-1}$ arranged in descending order, $\boldsymbol{W}_M$ is of size $(P-L)\times M$ and consists of the last M columns of $\boldsymbol{W}$ (which are the eigenvectors corresponding to those M smallest eigenvalues), and $\boldsymbol{W}_M^{\perp}$ consists of the first $P-L-M$ columns of $\boldsymbol{W}$. (For the case where $M = P-L$, this result was stated, but not proved, in [19, Part I fn 5].)

By substituting (33) and (34) into (32), we find that the optimal solution to the problem in (31) is feasible for the problem in (28) if and only if

$$\frac{p_0}{\mathrm{tr}\left(\boldsymbol{\Lambda}_M^{1/2}\right)^2} \geq \frac{3\sigma^2}{M}. \tag{35}$$

Furthermore, $\boldsymbol{Z}(\boldsymbol{U}_{\mathrm{opt}})$ is diagonal, and hence, we can simply choose $\boldsymbol{V}_{\mathrm{opt}} = \boldsymbol{D}_M$. Therefore, if (35) is satisfied, then a precoder that solves the problem in (28), i.e., minimizes $P_{e,\mathrm{LB}}$ subject to the lower bound being valid and to a bound on the transmitted power, is given by

$$\boldsymbol{F}_{\mathrm{opt}} = \sqrt{\frac{p_0}{\mathrm{tr}\left(\boldsymbol{\Lambda}_M^{1/2}\right)}}\,\boldsymbol{W}_M\boldsymbol{\Lambda}_M^{1/4}\boldsymbol{D}_M. \tag{36}$$

Moreover, this $\boldsymbol{F}_{\mathrm{opt}}$ results in the diagonal elements of $\boldsymbol{G}\boldsymbol{G}^H$ being the same [see the part ii) of the proof of Lemma 1], and hence, the actual BER of this precoder achieves the minimized lower bound. Therefore, $\boldsymbol{F}_{\mathrm{opt}}$ in (36) is a minimum BER precoder. The minimized BER is

$$P_{e,\min} = \frac{1}{2}\,\mathrm{erfc}\left(\sqrt{\frac{Mp_0}{2\sigma^2\left(\mathrm{tr}(\boldsymbol{\Lambda}_M^{1/2})\right)^2}}\right). \tag{37}$$

IV. Remarks

The following remarks on the minimum BER precoder design are in order.

1) The above design for a minimum BER precoder regards the precoder $\boldsymbol{F}$ as a general "tall" matrix of full column rank and is therefore applicable to both the ZP and CP options for the elimination of IBI.
2) The MBER precoder design is obtained from further exploitation of the MMSE precoder for ZF equalization given by (33) and (34). The set of all MMSE precoders for ZF equalization is given by

$$\boldsymbol{F}_{\text{MMSE-ZF}} = \boldsymbol{U}_{\mathrm{opt}}\begin{bmatrix}\boldsymbol{\Phi}_{\mathrm{opt}}\\ \boldsymbol{0}\end{bmatrix}\boldsymbol{V} = \boldsymbol{W}_M\boldsymbol{\Phi}_{\mathrm{opt}}\boldsymbol{V} \tag{38}$$

where $\boldsymbol{V}$ is an arbitrary $M\times M$ unitary matrix. Therefore, the minimum BER precoder is also an MMSE-ZF precoder, but the reverse is not necessarily true. In particular, only those $\boldsymbol{V}$ in (38) with elements of equal magnitude achieve minimum BER. The role of this special class of $\boldsymbol{V}$ matrices is to ensure that the diagonal elements of $\boldsymbol{G}\boldsymbol{G}^H$ are equal and, hence, that we achieve the minimized lower bound on the BER generated by Jensen's inequality and the minimization of the MSE. In practice, the role of this class of $\boldsymbol{V}$ matrices is to distribute the noise power across the subchannels so that the SNR at the input to each detector in Fig. 1 is the same. The effectiveness of this class of $\boldsymbol{V}$ matrices will be illustrated in Section VI.
3) If we define the ratio of the transmitted signal power to the receiver noise power (SNR) to be $\rho \triangleq p_0/(P\sigma^2)$, then from (35), the above MBER design is valid if

$$\rho > \frac{3(\mathrm{tr}(\boldsymbol{\Lambda}_M^{1/2}))^2}{MP} \triangleq \rho_c. \tag{39}$$

This condition is a function of the channel (through $\boldsymbol{\Lambda}$) and the block sizes M and P and can be evaluated before one attempts to solve the problem in (31). In order to satisfy this condition, we can either increase the transmitting power to raise the SNR or drop the subchannels corresponding to the largest elements of $\boldsymbol{\Lambda}_M$ to lower the value of ρ_c. Dropping subchannels corresponds to avoiding transmission on the "low-gain" subchannels and reallocating transmission power among the surviving ones. Although the denominator in (39) is also decreased, the numerator decreases more rapidly, and therefore, the value of ρ_c diminishes. The benefit of dropping subchannels is that the MBER precoder is guaranteed without violating the transmission power budget, but the block size, and hence transmission rate, is lower. [Recall from Assumption A5) that the bit loading is uniform.] The value of ρ_c determines the SNR below which subchannel dropping is required and thus determines how the tradeoff is made between the BER and transmission rate. The process of dropping subchannels and the corresponding MBER precoder design can be implemented by the following steps:
— First, determine the new block size $\bar{M}$: Let $\bar{M} = M$, and while $\rho < 3(\mathrm{tr}(\boldsymbol{\Lambda}_{\bar{M}}^{1/2}))^2/(P\bar{M})$, set $\bar{M} = \bar{M}-1$.
— Then, construct $\boldsymbol{F}_{\text{MBER-DROP}}$, which is the MBER precoder after dropping subchannels, via (36) with M replaced by $\bar{M}$, i.e.,

$$\boldsymbol{F}_{\mathrm{MBER-DROP}} = \sqrt{\frac{p_0}{\mathrm{tr}\left(\boldsymbol{\Lambda}_{\bar{M}}^{1/2}\right)}}\,\boldsymbol{W}_{\bar{M}}\boldsymbol{\Lambda}_{\bar{M}}^{1/4}\boldsymbol{D}_{\bar{M}}. \tag{40}$$

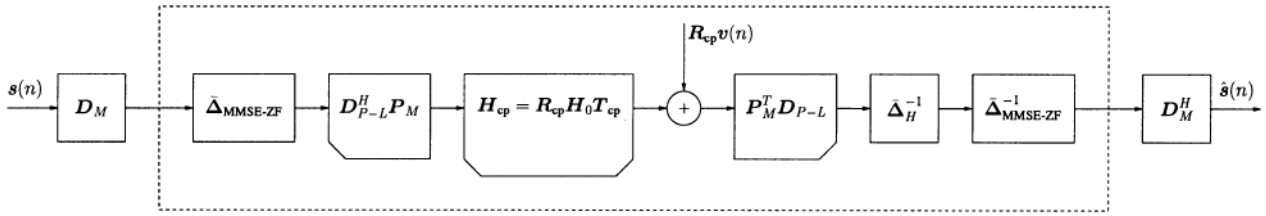


Fig. 2. MBER precoder for the CP transmission scheme. The system inside the dashed box can be viewed as a DMT system with MMSE-ZF power loading (and ZF equalization).

V. MBER Precoders for CP Systems

The MBER precoder design scheme in Section III is applicable to both ZP and CP transmission schemes. In both these schemes, the precoder is a $(P-L)\times M$ full column rank matrix. However, in the case of CP transmission, we can use the special structure of $\boldsymbol{H}_{\text{cp}}$ in (10) to simplify the expression for $\boldsymbol{F}_{\text{MBER}}$. In that case, $\boldsymbol{H}=\boldsymbol{H}_{\text{cp}}=\boldsymbol{D}_{P-L}^H\boldsymbol{\Delta}_H\boldsymbol{D}_{P-L}$, where the $(i,\,i)$th element of the diagonal matrix $\boldsymbol{\Delta}_H$ is $H(e^{j2\pi i/(P-L)})$, $0\le i\le P-L-1$. Let us define a square permutation matrix $\boldsymbol{P}$ of dimension $(P-L)$ such that the diagonal elements of $\boldsymbol{P}^T\boldsymbol{\Delta}_H\boldsymbol{P}$ are ordered so that their magnitudes are in increasing order. Furthermore, let

$$\bar{\boldsymbol{\Delta}}_H=[\,\boldsymbol{0}\quad \boldsymbol{I}_M\,]\,\boldsymbol{P}^T\boldsymbol{\Delta}_H\boldsymbol{P}\begin{bmatrix}\boldsymbol{0}\\ \boldsymbol{I}_M\end{bmatrix} \tag{41}$$

denote the $M\times M$ diagonal matrix containing the M subcarrier gains with the largest magnitudes. Then, $\boldsymbol{\Lambda}_M=(\bar{\boldsymbol{\Delta}}_H^H\bar{\boldsymbol{\Delta}}_H)^{-1}$, $\boldsymbol{W}=\boldsymbol{D}_{P-L}^H\boldsymbol{P}$, and $\boldsymbol{W}_M=\boldsymbol{D}_{P-L}^H\boldsymbol{P}_M$, where $\boldsymbol{P}_M=\boldsymbol{P}[\begin{smallmatrix}\boldsymbol{0}\\ \boldsymbol{I}_M\end{smallmatrix}]$ contains the last M columns of $\boldsymbol{P}$. Therefore, for CP systems, the condition for the validity of our MBER precoder can be rewritten as

$$\rho>\rho_{c,\,\text{cp}}=\frac{3\left(\text{tr}\left((\bar{\boldsymbol{\Delta}}_H^H\bar{\boldsymbol{\Delta}}_H)^{-1/2}\right)\right)^2}{MP} \tag{42}$$

and the CP-MBER precoder can be rewritten as

$$\boldsymbol{F}_{\text{CP-MBER}}=\sqrt{\frac{p_0}{\text{tr}\left((\bar{\boldsymbol{\Delta}}_H^H\bar{\boldsymbol{\Delta}}_H)^{-1/2}\right)}}\cdot\boldsymbol{D}_{P-L}^H\boldsymbol{P}_M\left(\bar{\boldsymbol{\Delta}}_H^H\bar{\boldsymbol{\Delta}}_H\right)^{-1/4}\boldsymbol{D}_M. \tag{43}$$

Using (38), the set of all MMSE precoders for CP systems with ZF equalization can be expressed as

$$\boldsymbol{F}_{\text{CP-MMSE-ZF}}=\boldsymbol{D}_{P-L}^H\boldsymbol{P}_M\bar{\boldsymbol{\Delta}}_{\text{MMSE-ZF}}\boldsymbol{V} \tag{44}$$

where $\boldsymbol{V}$ is an arbitrary unitary matrix, and

$$\bar{\boldsymbol{\Delta}}_{\text{MMSE-ZF}}=\sqrt{\frac{p_0}{\text{tr}\left(\left(\bar{\boldsymbol{\Delta}}_H^H\bar{\boldsymbol{\Delta}}_H\right)^{-1/2}\right)}}\left(\bar{\boldsymbol{\Delta}}_H^H\bar{\boldsymbol{\Delta}}_H\right)^{-1/4} \tag{45}$$

is the $M\times M$ diagonal MMSE power loading matrix for ZF equalization. When $\boldsymbol{V}=\boldsymbol{I}$, the MMSE-ZF precoder in (44) takes a similar form to that of the standard DMT scheme discussed at the end of Section II-B; see also (53) and (54) later. The difference is that in standard DMT, the diagonal power loading is calculated via water filling, whereas (45) is the diagonal power loading that minimizes the MSE (for ZF equalization). An advantage of the MMSE-ZF power loading in (45) is that it has a closed form and, hence, is rather simple to implement. In contrast, the water-filling power loading requires the solution of a nonlinear optimization problem via an iterative algorithm. The CP-MBER precoder in (43) also takes a similar form to that of a standard DMT system, but in this case, the diagonal power loading is replaced by a full matrix consisting of a diagonal MMSE-ZF power loading matrix post-multiplied by a DFT matrix. The block diagram of a CP transmission system with MBER precoding is shown in Fig. 2, where the receiver is a ZF equalizer, i.e., $\boldsymbol{G}_{\text{cp}}=(\boldsymbol{H}_{\text{cp}}\boldsymbol{F}_{\text{cp}})^{\dagger}=\boldsymbol{D}_M^H\bar{\boldsymbol{\Delta}}_{\text{MMSE-ZF}}^{-1}\bar{\boldsymbol{\Delta}}_H^{-1}\boldsymbol{P}_M^T\boldsymbol{D}_{P-L}$. The system inside the dashed box can be regarded as a DMT system with MMSE-ZF power loading.

VI. Performance Evaluation

We consider the performance of several ZP and CP precoders in binary transmission over a well-conditioned channel, and an ill-conditioned channel, and the average performance of these precoders over a class of randomly generated channels. As mentioned in Section IV, the SNR is defined as $\rho=p_0/(P\sigma^2)$. In all of our designs, the transmission power will be normalized, i.e., $p_0=1$.

For ZP precoders, comparisons are given among the precoders designed by the criteria of MBER, MMSE (for ZF equalization) [19], maximum SNR [19], and zero-padded OFDM [13], [19]. From the set of MMSE precoders in (38), we choose the one in which the unitary matrix degree of freedom is given by $\boldsymbol{V}=\boldsymbol{I}$. If M symbols are to be transmitted in each block, the ZP precoders are expressed, respectively, as follows:

$$\boldsymbol{F}_{\text{ZP-MBER}}=\sqrt{\frac{p_0}{\text{tr}(\boldsymbol{\Lambda}_M^{1/2})}}\,\boldsymbol{W}_M\boldsymbol{\Lambda}_M^{1/4}\boldsymbol{D}_M \tag{46}$$

$$\boldsymbol{F}_{\text{ZP-MMSE-ZF-I}}=\sqrt{\frac{p_0}{\text{tr}(\boldsymbol{\Lambda}_M^{1/2})}}\,\boldsymbol{W}_M\boldsymbol{\Lambda}_M^{1/4} \tag{47}$$

$$\boldsymbol{F}_{\text{ZP-MSNR}}=\sqrt{\frac{p_0}{\text{tr}(\boldsymbol{\Lambda}_M)}}\,\boldsymbol{W}_M\boldsymbol{\Lambda}_M^{1/2} \tag{48}$$

$$\boldsymbol{F}_{\text{ZP-OFDM}}=\sqrt{\frac{p_0}{M}}\,\boldsymbol{D}_{P-L}^H\boldsymbol{Q}_M \tag{49}$$

where $\boldsymbol{Q}_M$ consists of M columns of $\boldsymbol{I}_{P-L}$. [For the ZP-OFDM precoder in (49) and the CP-OFDM precoder in (52) below, we will usually choose $M=P-L$, in which case, $\boldsymbol{Q}_M=\boldsymbol{I}_{P-L}$.] For CP transmission schemes, we examine the following precoders: MBER, MMSE (for ZF equalization), OFDM, and water-filling DMT. The unitary matrix in the MMSE design is

again chosen to be the identity matrix. The precoders have the following forms:

$$\boldsymbol{F}_{\text{CP-MBER}} = \sqrt{\frac{p_0}{\text{tr}\left((\bar{\boldsymbol{\Delta}}_H^H \bar{\boldsymbol{\Delta}}_H)^{-1/2}\right)}} \cdot \boldsymbol{D}_{P-L}^H \boldsymbol{P}_M \left(\bar{\boldsymbol{\Delta}}_H^H \bar{\boldsymbol{\Delta}}_H\right)^{-1/4} \boldsymbol{D}_M \quad (50)$$

$$\boldsymbol{F}_{\text{CP-MMSE-ZF-I}} = \boldsymbol{D}_{P-L}^H \boldsymbol{P}_M \bar{\boldsymbol{\Delta}}_{\text{MMSE-ZF}} \quad (51)$$

$$\boldsymbol{F}_{\text{CP-OFDM}} = \sqrt{\frac{p_0}{M}}\, \boldsymbol{D}_{P-L}^H \boldsymbol{Q}_M \quad (52)$$

$$\boldsymbol{F}_{\text{WF-DMT}} = \boldsymbol{D}_{P-L}^H \boldsymbol{Q}_{\text{WF}} \bar{\boldsymbol{\Delta}}_{\text{WF}} \quad (53)$$

where $\bar{\boldsymbol{\Delta}}_{\text{WF}} = \boldsymbol{Q}_{\text{WF}}^T \boldsymbol{\Delta}_{\text{WF}} \boldsymbol{Q}_{\text{WF}}$. Here, $\boldsymbol{\Delta}_{\text{WF}}$ is the water-filling power loading matrix that maximizes the achievable information rate, and $\boldsymbol{Q}_{\text{WF}}$ consists of the columns of $\boldsymbol{I}_{P-L}$ corresponding to the nonzero values of $\boldsymbol{\Delta}_{\text{WF}}$. The ith diagonal element of $\boldsymbol{\Delta}_{\text{WF}}$ is [3], [15]

$$[\boldsymbol{\Delta}_{\text{WF}}]_{ii} = \begin{cases} \sqrt{\eta - \dfrac{\sigma^2}{[\boldsymbol{\Delta}_H^H \boldsymbol{\Delta}_H]_{ii}}}, & \text{if } \eta > \dfrac{\sigma^2}{[\boldsymbol{\Delta}_H^H \boldsymbol{\Delta}_H]_{ii}} \\ 0, & \text{otherwise} \end{cases} \quad (54)$$

where $[\boldsymbol{\Delta}_H^H \boldsymbol{\Delta}_H]_{ii} = |H(e^{j2\pi i/(P-L)})|^2, 0 \le i \le P-L$, and η is a constant chosen such that power $\text{tr}(\boldsymbol{\Delta}_{\text{WF}} \boldsymbol{\Delta}_{\text{WF}}^H) = p_0$. From (54), it is clear that subchannels with low gain might not be allocated any power, i.e., they might be dropped. However, the channel dropping criterion in (54) is different from that in Section IV (see Example 2). Before proceeding with the examples, we emphasize the fact that in order to provide a consistent comparison between the various transmission schemes, we will employ a uniform bit loading (one bit per subchannel) to each of the $\bar{M}$ subchannels that are allocated power. As such, our results for the water-filling scheme are best considered as a reference or benchmark because in practice, a nonuniform bit loading would normally be used for water-filling power loaded systems.

Example 1: In this example, we examine the performance of the various linear transceivers when the data blocks are sent over a channel with tap coefficients $\{0.3038 + 0.2554j, 0.5056 + 0.5587j, 0.2855 + 0.0035j, 0.2834 + 0.1843j, 0.2793 + 0.0305j\}$. The data blocks are of length $M = 32$, the channel has order $L = 4$, and we choose the transmitted block size to be $P = 36$. In this case, $M = P - L$, and hence, $\boldsymbol{\Lambda}_M = \boldsymbol{\Lambda}$ and $\boldsymbol{W}_M = \boldsymbol{W}$. The magnitude of a $P - L = 32$-point DFT of the impulse response of the channel is shown in Fig. 3. It can be observed that the ratio of the largest to the smallest subchannel gains is 4. For ZP schemes, the critical value of SNR in (39) for this channel is $\rho_{c,\text{zp}} = 7.46$ dB, whereas that for CP schemes is $\rho_{c,\text{cp}} = 7.61$ dB. (For a given channel $h(k)$ and block sizes M and P, $\boldsymbol{H}_{\text{cp}} = [\boldsymbol{I}_{(P-L)} \quad \begin{matrix}\boldsymbol{I}_L \\ \boldsymbol{0}_{L\times(P-2L)}\end{matrix}]\boldsymbol{H}_{\text{zp}}$, [21], and hence, it can be shown that $\rho_{c,\text{cp}} \ge \rho_{c,\text{zp}}$.) We compare the BER performance of the various linear transceivers at SNRs between 0 and 18 dB. Fig. 4 shows the performance of the various ZP schemes, and Fig. 5 shows the performance of the various CP schemes. We have not used the subchannel dropping scheme here but will do so in Fig. 6. It is observed that in both cases, the MBER design provides significantly improved performance over the other schemes when $\rho \ge \rho_c$. At lower SNRs, the differences between the performance of the various schemes become small. We note from Fig. 4 that even though the ZP-MSNR scheme is not an MMSE design, at high SNR, it outperforms the ZP-MMSE-ZF

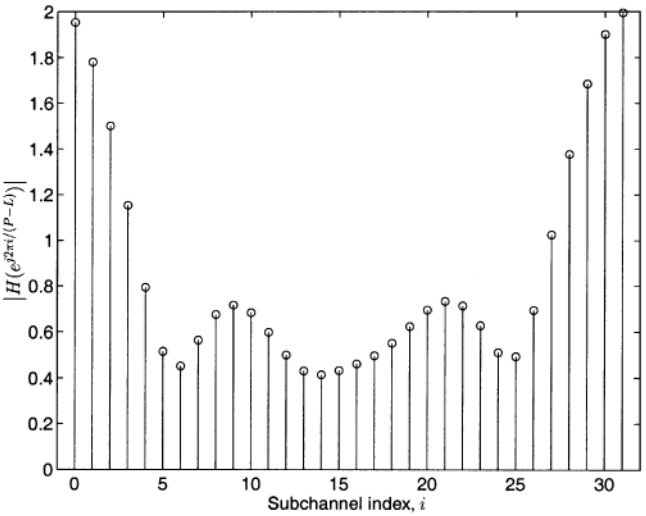


Fig. 3. Frequency response of the channel in Example 1.

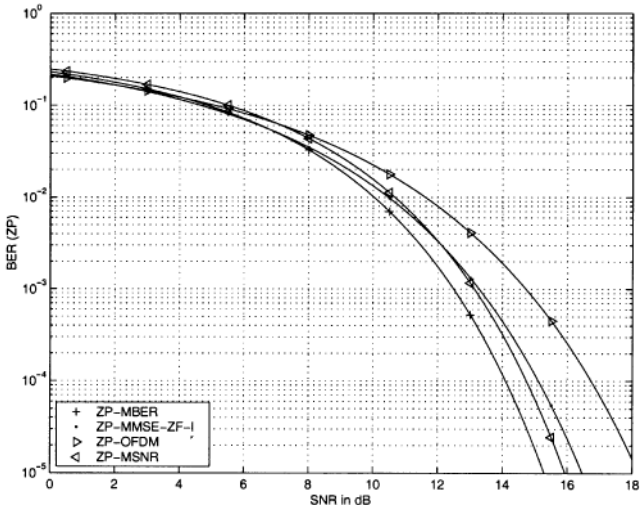


Fig. 4. BER performance of ZP precoders in Example 1.

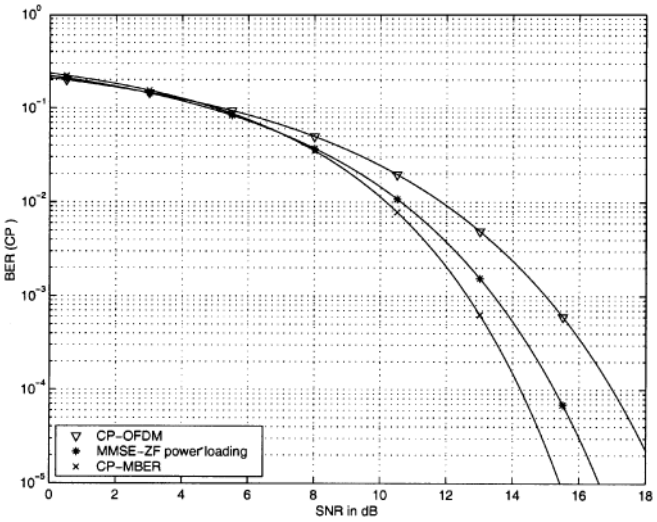
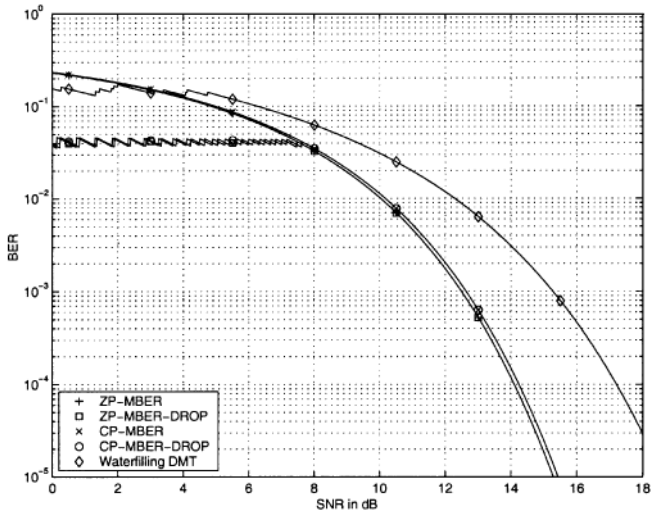
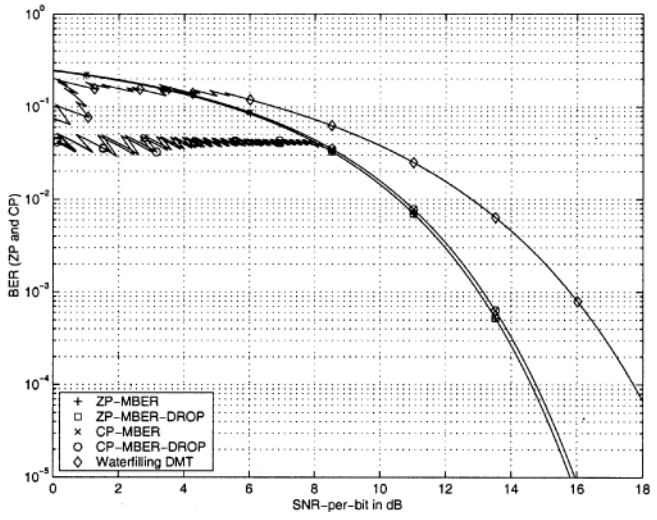


Fig. 5. BER performance of CP precoders in Example 1.



(a)



(b)

Fig. 6. BER performance of CP-MBER and ZP-MBER precoders with sub-channel dropping and the water-filling DMT precoder in Example 1. (a) BER versus (block) SNR. (b) BER versus average SNR-per-bit.

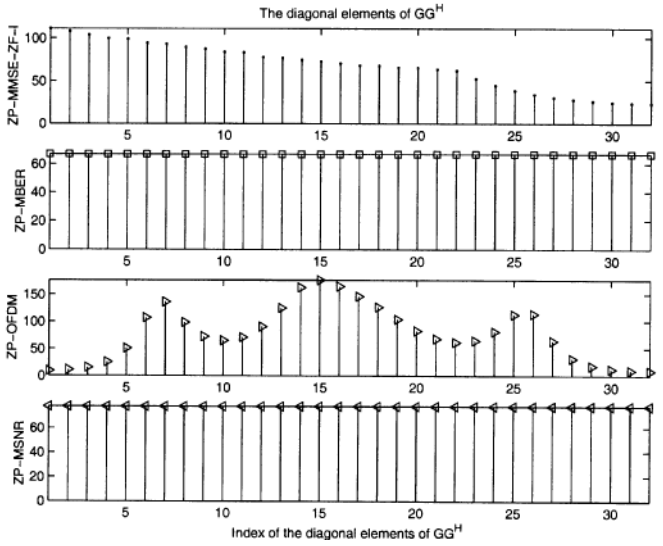


Fig. 7. Diagonal elements of $\boldsymbol{GG}^H$ for the ZP schemes in Example 1.

design in which the unitary matrix degree of freedom is chosen as the identity matrix. However, our minimum BER scheme, which is an MMSE design with a special unitary matrix degree of freedom, outperforms both the ZP-MMSE-ZF-I design and the ZP-MSNR design.

Fig. 6(a) presents the performance of the ZP and CP minimum BER schemes in which the system is allowed to drop the subchannels with poor frequency responses. We compare the performance of these two schemes with that of the DMT scheme equipped with the water-filling algorithm of (54) (and uniform bit loading). As expected, it is observed that at high SNR, the performance of the MBER schemes with subchannel dropping (designated as ZP-MBER-DROP and CP-MBER-DROP in Fig. 6) is the same as that of the ZP-MBER and CP-MBER schemes, respectively. At SNRs below ρ_c, the ZP-MBER-DROP and CP-MBER-DROP schemes begin to drop subchannels, whereas for the water-filling DMT scheme, subchannels are dropped below about 5 dB. Over the whole SNR range, both the ZP-MBER-DROP and CP-MBER-DROP schemes provide better BER performance than uniformly bit-loaded water-filling DMT. The gain of these new optimum schemes over DMT at a BER of 10^{-4} is about 3.5 dB. Recall that when a subchannel is dropped, the power that was allocated to that subchannel is distributed over the remaining subchannels, and therefore, the BER improves in a discontinuous fashion. This leads to the "jagged" nature of the curves in a Fig. 6(a) at low SNR. Given our uniform bit allocation policy, subchannel dropping results in a reduction of the number of bits per block and, hence, a drop in the bit rate. To include the effects of this rate change in our analysis, in Fig. 6(b), we plot the BER performance data from Fig. 6(a) against the average SNR-per-bit $p_0/(\bar{M}\sigma^2)$, rather than the block SNR $p_0/(P\sigma^2)$, which is used in Fig. 6(a). Since the subchannel dropping criteria of both our MBER schemes and the water-filling DMT scheme are based on a constraint on the transmitted power per block, not per subchannel, the BER versus average SNR-per-bit curves can be multivalued at low SNR. However, the basic trends of both graphs in Fig. 6 are the same.

To explain the performance advantage of our designs, Figs. 7 and 8 show the diagonal elements of the matrices $\boldsymbol{GG}^H$ for the different ZP and CP schemes. The comparison here is evaluated at $\rho = 18$ dB so that for all designs, none of the subchannels are dropped. Recall from Section III that in order to minimize the BER, we first need to minimize $\mathrm{tr}(\boldsymbol{GG}^H)$. For the ZP schemes, we note that neither the ZP-OFDM nor ZP-MSNR are MMSE designs. Therefore, the values of $\mathrm{tr}(\boldsymbol{GG}^H)$ for these schemes are higher than that for the MMSE designs. However, we also note that for the ZP-MSNR scheme, the diagonal elements of $\boldsymbol{GG}^H$ are equal. On the other hand, among the MMSE designs [which minimize $\mathrm{tr}(\boldsymbol{GG}^H)$], the ZP-MBER scheme is the only design for which $\boldsymbol{GG}^H$ has equal diagonal elements. Similar observations can be obtained from Fig. 8. Both CP-OFDM and water-filling DMT are not MMSE designs and do not result in the product $\boldsymbol{GG}^H$ having equal diagonal elements. The CP-MBER scheme is the only CP design that results in $\boldsymbol{GG}^H$ having equal diagonal elements while minimizing $\mathrm{tr}(\boldsymbol{GG}^H)$.

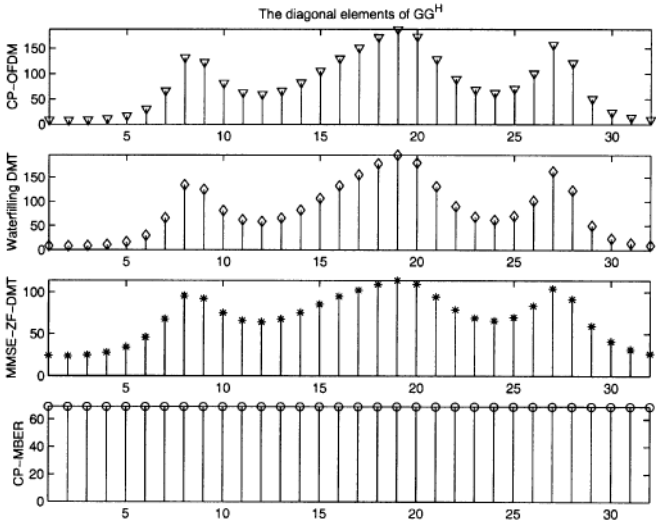


Fig. 8. Diagonal elements in GG^H for the CP schemes in Example 1.

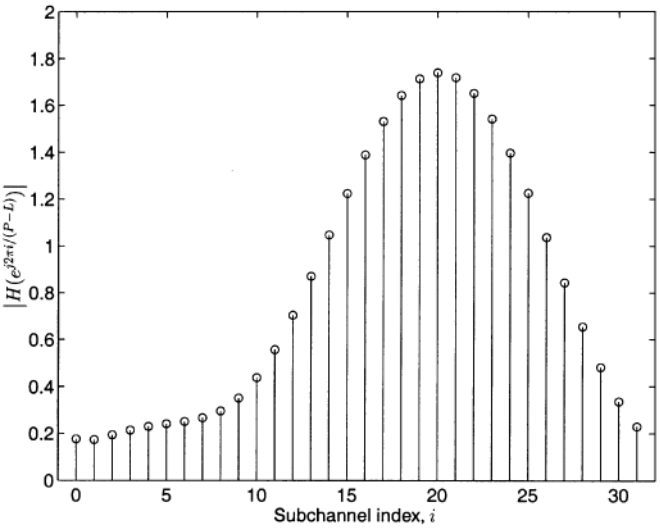


Fig. 9. Frequency response of the channel in Example 2.

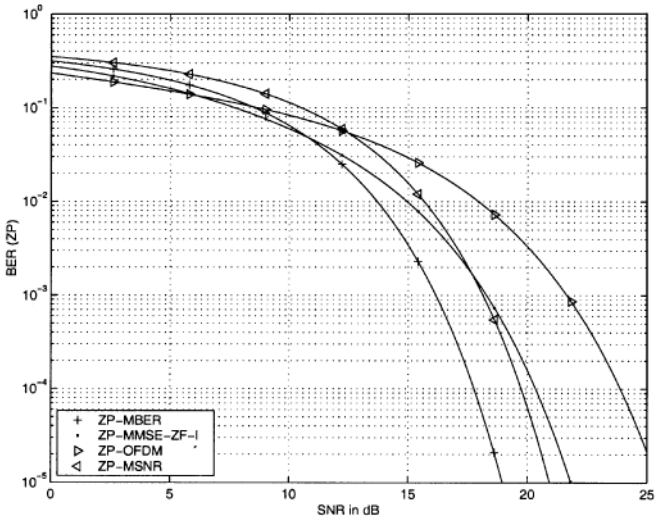


Fig. 10. BER performance of the ZP precoders in Example 2.

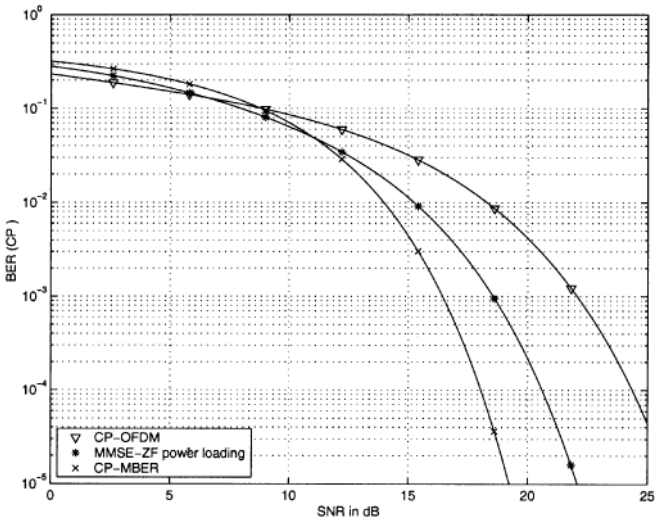
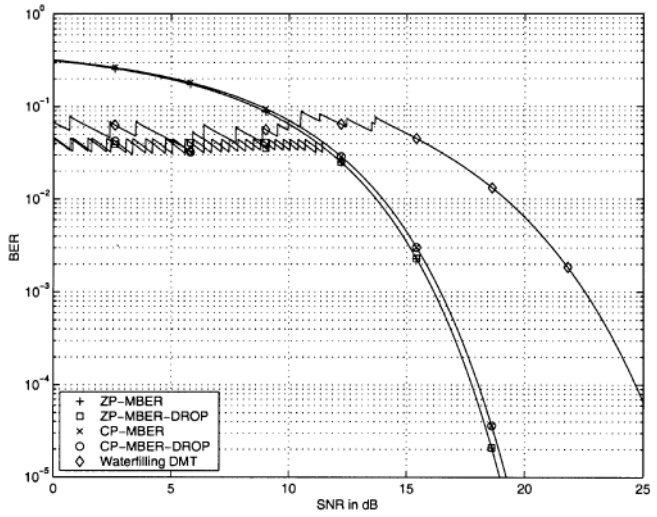


Fig. 11. BER performance of the CP precoders in Example 2.

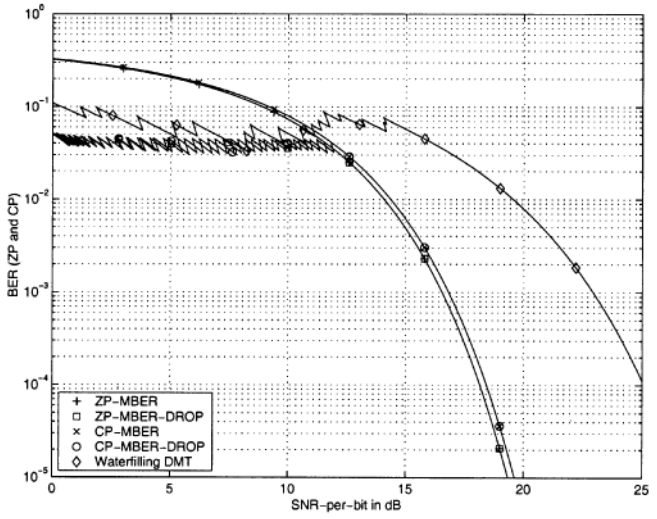
Example 2: In this example, we examine the performance of the various block transmission schemes when the data are sent over a channel with tap coefficients $\{0.6121, -0.5331 - 0.4481j, 0.369j, 0.0513 - 0.0388j\}$. The data blocks are of length $M = 32$ and the channel is of order $L = 3$, and we choose the transmitted block size to be $P = 35$. The magnitude of a $P - L = 32$ point DFT of the impulse response of the channel is shown in Fig. 9. The ratio of the largest to the smallest subchannel gain is 8.75, which means that this channel is more frequency selective than the one in Example 1. For ZP schemes, the SNR threshold of (39) for this channel is $\rho_{c,\mathrm{zp}} = 11.13$ dB, whereas that for CP schemes is $\rho_{c,\mathrm{cp}} = 11.41$ dB. Figs. 10 and 11 show the BER performance of the ZP and CP schemes, respectively. Similar to Example 1, the MBER designs show clearly superior performance when $\rho \geq \rho_c$. When $\rho < \rho_c$, the differences between the BERs of each scheme become small. The SNR gain of the MBER precoders over OFDM at a BER of 10^{-4} is about 7 dB in this case.

Fig. 12(a) compares the BER performance of schemes in which the subchannels with low gain are allowed to be dropped at low SNR. At SNRs below the corresponding ρ_c, both the ZP-MBER-DROP and CP-MBER-DROP schemes start dropping subchannels. These dropping schemes provide substantial improvement over the performance of the corresponding ZP-MBER and CP-MBER schemes. The water-filling DMT system carries out its own subchannel dropping scheme. Since the subchannel dropping schemes are different, the water-filling DMT scheme starts dropping subchannels at a higher SNR of around 14 dB. Despite the fact that the water-filling DMT scheme employs a smaller block size than our MBER precoders for SNRs below about 14 dB (see Fig. 13), the MBER scheme provides better BER performance. This trend is also clear from Fig. 12(b), in which we have plotted the same BER performance data against average SNR-per-bit rather than the block SNR. As was mentioned in the discussion in Example 1, the BER versus average SNR-per-bit curves can be multivalued at low SNR.

Example 3: In Examples 1 and 2, we examined the BER performance of various precoders in a "good-quality" and a "moderate-quality" channel, respectively. In this example, we



(a)



(b)

Fig. 12. BER performance of the CP-MBER and ZP-MBER precoders and the water-filling DMT precoder in Example 2. (a) BER versus (block) SNR. (b) BER versus average SNR-per-bit.

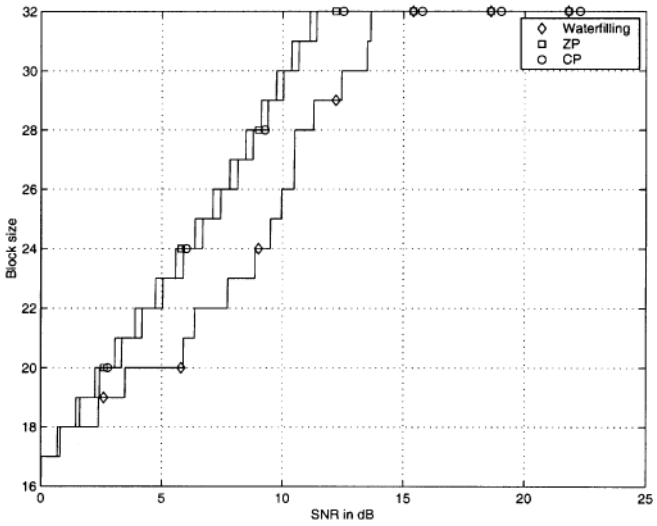


Fig. 13. Block sizes generated by the ZP-MBER-DROP, CP-MBER-DROP, and water-filling DMT schemes in Example 2.

examine the average BER performance of some precoders over a class of randomly generated channels. The (complex-valued) taps of the channels were generated independently from a zero-mean circular Gaussian distribution with unit variance per dimension. Each channel realization was then normalized so that the impulse response had unit two-norm. The channels were of order $L = 4$, and the block sizes were chosen as $M = 16$ and $P = M + L = 20$. The average BER performance curves for various ZP precoders over 2000 channel realizations from this class are shown in Fig. 14(a). The average BER performance curves for various CP schemes are shown in Fig. 14(b), and those for the MBER schemes with subchannel dropping and the water-filling DMT precoders are shown in Fig. 14(c). These curves indicate that the trends established in the previous two examples extend to the average case. Furthermore, the SNR gains of our MBER designs remain significant. For example, from Fig. 14(a), it can be seen that at a BER of 10^{-3} the SNR gain of our ZP-MBER precoder (without sub-channel dropping) over the ZP-MMSE-ZF-I precoder is around 2 dB, and from Fig. 14(b), the corresponding SNR gain of our CP-MBER precoder (without sub-channel dropping) over CP-OFDM is around 7.5 dB.

VII. Conclusion

In this paper, we have designed a linear block-by-block precoder that minimizes the BER achieved with ZF equalization and threshold detection. The design was obtained by observing that at high SNR the expression for the BER is a convex function of the magnitude of the diagonal elements of the equalizer. A lower bound for the BER was derived, and its minimum value was obtained by minimizing the trace of the matrix product $\boldsymbol{G}\boldsymbol{G}^H$, where $\boldsymbol{G}$ is the equalization matrix. We then showed that this (minimized) lower bound on the BER can be attained by certain members of the set of precoders that minimize the lower bound. The set of precoders that minimize the lower bound was shown to be the set of MMSE precoders for ZF equalization—a set which is parameterized by a unitary matrix degree of freedom. The MBER precoders were obtained from special choices of this unitary matrix. One such choice is the (normalized) DFT matrix. The design scheme is flexible because it applies directly to both the zero-padding and cyclic-prefix schemes for avoiding inter-block interference and allows the block sizes to be chosen freely [up to Assumption A2)]. We also provided a natural modification of the design method for low SNR scenarios. In the modified design, the low gain subchannels are systematically dropped (thus reducing the data block size) in order to maintain a sufficiently high SNR on the remaining subchannels.

Performance comparisons have been provided to demonstrate that the MBER design can provide substantially lower BER than all current designs. The improvement in BER performance over some commonly used systems may be as much as several decibels in SNR at reasonable BER levels. Another advantage of our design is that the MBER precoder has a simple analytic form, whereas the water-filling-based designs require the solution of a nonlinear optimization problem using an iterative

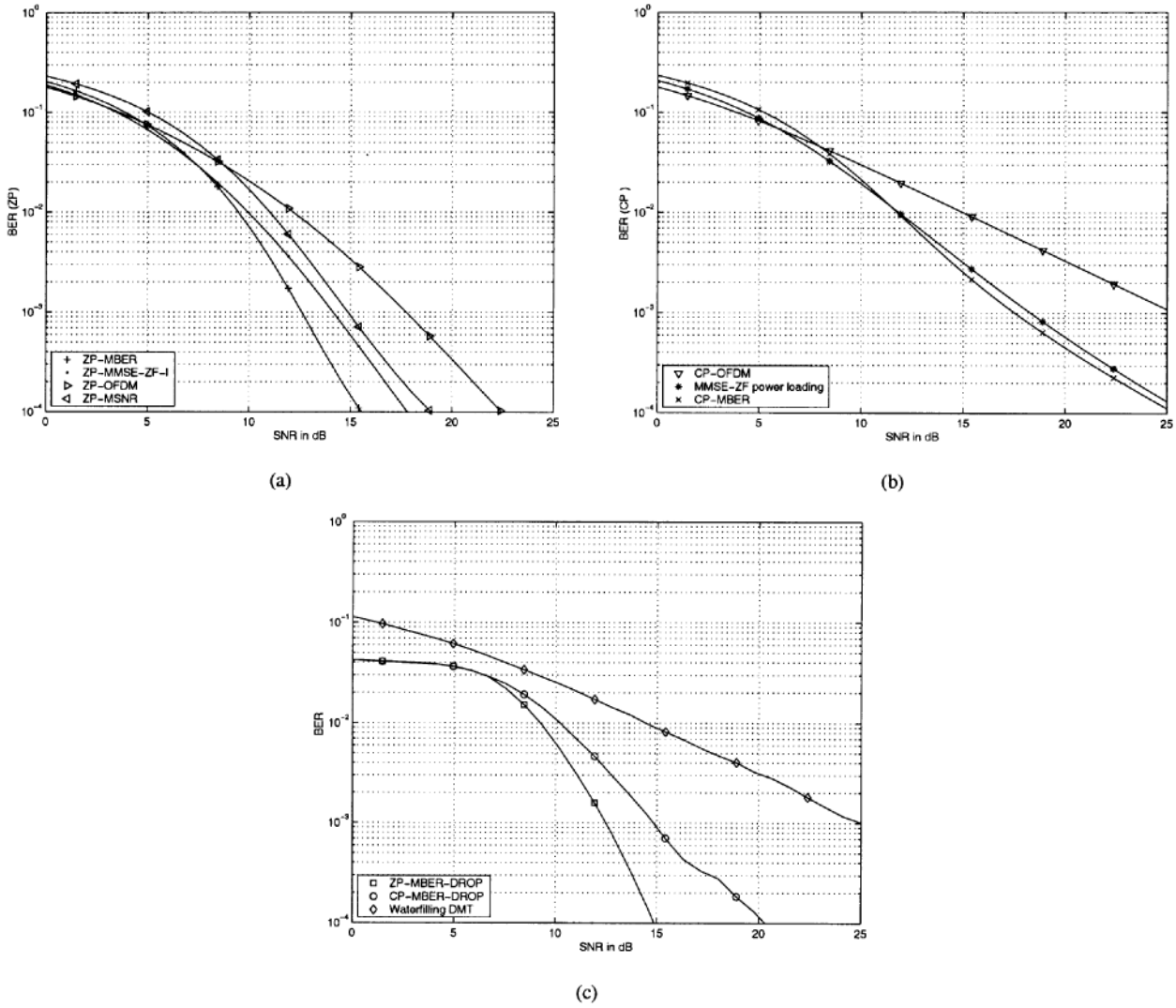


Fig. 14. Average BER performance of various precoders over the class of channels in Example 3. (a) BER versus (block) SNR for ZP precoders. (b) BER versus (block) SNR for CP precoders. (c) BER versus (block) SNR for subchannel dropping precoders.

algorithm. The extra computational load incurred when operating an MBER design in place of one of the currently used systems is minimal, being only the additional DFTs at the transmitter and the receiver. Thus, the MBER design proposed in this paper is an attractive alternative for realizing linear precoders for block-by-block data transmission with zero-forcing equalization.

The optimal design obtained in the paper is for a single-user system, with ZF equalization and threshold detection, white uncoded data, white noise, uniform bit loading, and a known channel. We are currently working on extending the ideas in this paper to various other schemes such as multiuser systems, different equalization techniques (MMSE [2], decision feedback, etc.), maximum likelihood detection, colored noise, and alternative "bit-loading" schemes, as well as other scenarios in which the channel is imprecisely known at the transmitter. To keep the exposition of our minimum BER design method reasonably simple, we have restricted our attention to single-input, single-output block transmission systems in this paper. However, the algebraic model of multiple-input, multiple-output (MIMO) block transmission systems, such as those considered in [17], [20], and [26] is essentially the same as that in (8), and hence, our design approach extends naturally to the MIMO case.

APPENDIX
OPTIMUM VALUES OF $\boldsymbol{\Phi}$ AND $\boldsymbol{U}$ IN (31)

If $M = P - L$, an optimal product $\boldsymbol{U\Phi}$ can be found directly using a straightforward Lagrange multiplier argument [11]. In the more general case where $M \leq P - L$, it is more convenient to first find the optimal $\boldsymbol{\Phi}$ for a given choice of $\boldsymbol{U}$ and then find an optimal $\boldsymbol{U}$. For a fixed $\boldsymbol{U}$, the optimization problem in (31) can be reparameterized in terms of $\boldsymbol{\Gamma} = \boldsymbol{\Phi}^2$ as

$$\min_{\boldsymbol{\Gamma}} \operatorname{tr}\left(\boldsymbol{\Gamma}^{-1}\boldsymbol{Z}\right), \qquad \text{subject to } \operatorname{tr}(\boldsymbol{\Gamma}) \leq p_0. \tag{55}$$

Here, as in (27), $\boldsymbol{Z} = \boldsymbol{Z}(\boldsymbol{U}) = \left([\,\boldsymbol{I}_M \quad \boldsymbol{0}\,]\boldsymbol{U}^H\boldsymbol{H}^H\boldsymbol{H}\boldsymbol{U}\left[\begin{smallmatrix}\boldsymbol{I}_M\\ \boldsymbol{0}\end{smallmatrix}\right]\right)^{-1}$. The Lagrangian function for this problem is

$$L(\boldsymbol{\Gamma}, \mu) = \operatorname{tr}\left(\boldsymbol{\Gamma}^{-1}\boldsymbol{Z}\right) + \mu\left(\operatorname{tr}(\boldsymbol{\Gamma}) - p_0\right) \tag{56}$$

where μ is the Lagrange multiplier. Equating the derivatives of (56) with respect to $\mathbf{\Gamma}$ and μ to zero, we obtain the following necessary conditions for optimality:

$$-\left(\mathbf{\Gamma}^{-1}\boldsymbol{Z}\mathbf{\Gamma}^{-1}\right)^T + \mu\boldsymbol{I}_M = \mathbf{0}$$
$$\operatorname{tr}(\mathbf{\Gamma}) = p_0. \tag{57}$$

To satisfy these necessary conditions, we require $\mathbf{\Gamma}^2 = (1/\mu)\boldsymbol{Z}^T$, and hence

$$\mathbf{\Gamma} = \frac{1}{\sqrt{\mu}}\left([\boldsymbol{I}_M \quad \mathbf{0}]\boldsymbol{U}^H\boldsymbol{W}\mathbf{\Lambda}^{-1}\boldsymbol{W}^H\boldsymbol{U}\begin{bmatrix}\boldsymbol{I}_M\\ \mathbf{0}\end{bmatrix}\right)^{-1/2} \tag{58}$$

where, as in (25), $\boldsymbol{W}\mathbf{\Lambda}\boldsymbol{W}^H$ is an eigenvalue decomposition of $(\boldsymbol{H}^H\boldsymbol{H})^{-1}$ with eigenvalues arranged in descending order. However, since $\mathbf{\Gamma} = \mathbf{\Phi}^2$, and $\mathbf{\Phi}$ is diagonal, the optimal $\mathbf{\Gamma}$ must be diagonal and positive semidefinite. Therefore, since $\boldsymbol{U}$ is a unitary matrix, the optimal product $\boldsymbol{U}^H\boldsymbol{W}$ must be a permutation matrix, say, $\boldsymbol{P}$. Hence, $\boldsymbol{U}_{\text{opt}} = \boldsymbol{W}\boldsymbol{P}$. Making the dependence of $\boldsymbol{Z}$ on $\boldsymbol{U}$ explicit, the resulting objective in (55) is

$$\operatorname{tr}\left(\mathbf{\Gamma}_{\text{opt}}^{-1}\boldsymbol{Z}(\boldsymbol{W}\boldsymbol{P})\right) = \sqrt{\mu}\sum_{i=1}^{M}\left[\boldsymbol{P}^T\mathbf{\Lambda}^{1/2}\boldsymbol{P}\right]_{ii}.$$

An optimal $\boldsymbol{P}$ places the M smallest elements of $\mathbf{\Lambda}^{1/2}$ in the top left corner of $\boldsymbol{P}^T\mathbf{\Lambda}^{1/2}\boldsymbol{P}$. Since the elements of $\mathbf{\Lambda}$ are arranged in descending order, an optimal $\boldsymbol{P}$ is

$$\boldsymbol{P}_{\text{opt}} = \begin{bmatrix}\mathbf{0} & \boldsymbol{I}_{(P-L-M)}\\ \boldsymbol{I}_M & \mathbf{0}\end{bmatrix}$$

and hence, an optimal $\boldsymbol{U}$ is $\boldsymbol{U}_{\text{opt}} = \boldsymbol{W}\boldsymbol{P}_{\text{opt}} = [\boldsymbol{W}_M \ \boldsymbol{W}_M^{\perp}]$, as given in (34). Hence, $\mathbf{\Gamma}_{\text{opt}} = \mathbf{\Lambda}_M^{1/2}/\sqrt{\mu}$, where $\mathbf{\Lambda}_M = \operatorname{diag}(\lambda_{P-L-M+1}, \ldots, \lambda_{P-L})$. By choosing $\mu = \operatorname{tr}(\mathbf{\Lambda}_M^{1/2})/p_0$ so that (57) is satisfied, we have that

$$\mathbf{\Phi}_{\text{opt}} = \mathbf{\Gamma}_{\text{opt}}^{1/2} = \sqrt{\frac{p_0}{\operatorname{tr}\left(\mathbf{\Lambda}_M^{1/2}\right)}}\,\mathbf{\Lambda}_M^{1/4}.$$

Since $\boldsymbol{Z}(\boldsymbol{U})$ is positive semidefinite (by construction), $\lambda_i \geq 0$ and therefore, the diagonal elements of $\mathbf{\Phi}$ are real and non-negative, as required. Hence, (33) holds.

REFERENCES

[1] J. A. C. Bingham, "Multicarrier modulation for data transmission: An idea whose time has come," *IEEE Commun. Mag.*, vol. 28, pp. 5–14, May 1990.

[2] S. S. Chan, T. N. Davidson, and K. M. Wong, "Asymptotically minimum bit error rate block precoders for minimum mean square error equalization," in *Proc. Second IEEE Sensor Array Multichannel Signal Processing Workshop*, Rosslyn, VA, Aug. 2002.

[3] J. S. Chow, J. C. Tu, and J. M. Cioffi, "Performance evaluation of a multichannel transceiver system for ADSL and VHDSL receivers," *IEEE J. Select. Areas Commun.*, vol. 9, pp. 909–919, Aug. 1991.

[4] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York: Wiley, 1991.

[5] Y. W. Ding, "Minimum BER block precoders for zero-forcing equalization," M.Eng. Thesis, McMaster Univ., Hamilton, ON, Oct. 2001.

[6] N. J. Fliege, "Orthogonal multiple carrier data transmission," *Eur. Trans. Telecommun.*, vol. 3, pp. 255–264, May 1992.

[7] G. D. Forney and M. V. Eyuboğlu, "Combined equalization and coding using precoding," *IEEE Commun. Mag.*, pp. 25–34, Dec. 1991.

[8] G. H. Golub and C. F. Van Loan, *Matrix Computations*, Third ed. Baltimore, MD: Johns Hopkins Univ. Press, 1996.

[9] B. S. Krongold, K. Ramchandran, and D. L. Jones, "Computationally efficient optimal power allocation algorithms for multicarrier communication systems," *IEEE Trans. Commun.*, vol. 48, pp. 23–27, Jan. 2000.

[10] P. Lancaster and M. Tismenetsky, *The Theory of Matrices*. New York: Academic, 1985.

[11] J. P. Milanović, "Design of robust redundant precoding filter banks for uncertain frequency-selective channels," M. Eng. Thesis, McMaster Univ., Hamilton, ON, Canada, Sept. 2000.

[12] J. P. Milanović, T. N. Davidson, Z.-Q. Luo, and K. M. Wong, "Design of robust redundant precoding filterbanks with zero-forcing equalizers for unknown frequency-selective channels," in *Proc. Int. Conf. Acoust., Speech, Signal Processing*, Istanbul, Turkey, June 2000.

[13] B. Muquet, M. de Courville, P. Duhamel, and G. Giannakis, "OFDM with trailing zeros versus OFDM with cyclic prefix: Links, comparisons and application to the HiperLAN/2 system," in *Proc. Int. Conf. Commun.*, New Orleans, LA, June 2000.

[14] J. G. Proakis, *Digital Communications*, Fourth ed. New York: McGraw-Hill, 2001.

[15] A. Ruiz, J. M. Cioffi, and S. Kastuia, "Discrete multiple tone modulation with coset coding for the spectrally shaped channel," *IEEE Trans. Commun.*, vol. 40, pp. 1012–1029, June 1992.

[16] S. D. Sandberg and M. A. Tzannes, "Overlapped discrete multiple tone modulation for high speed copper wire communications," *IEEE J. Select. Areas Commun.*, vol. 13, pp. 1571–1585, Dec. 1995.

[17] H. Sampath, P. Stoica, and A. Paulraj, "Generalized linear precoder and decoder design for MIMO channels using the weighted MMSE criterion," *IEEE Trans. Commun.*, vol. 49, pp. 2198–2206, Dec. 2001.

[18] A. Scaglione, S. Barbarossa, and G. B. Giannakis, "Filterbank transceivers optimizing information rate in block transmissions over dispersive channels," *IEEE Trans. Inform. Theory*, vol. 45, pp. 1988–2006, Apr. 1999.

[19] A. Scaglione, G. B. Giannakis, and S. Barbarossa, "Redundant filterbank precoders and equalizers, Parts I and II," *IEEE Trans. Signal Processing*, vol. 47, pp. 1988–2022, July 1999.

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

[21] Z. Wang and G. B. Giannakis, "Wireless multicarrier communications," *IEEE Signal Processing Mag.*, pp. 29–48, May 2000.

[22] ——, "Linearly precoded or coded OFDM against wireless channel fades?," in *Proc. Third IEEE Workshop Signal Processing Advances Wireless Commun.*, Taoyuan, Taiwan, R.O.C., Mar. 2001.

[23] K. M. Wong, J. Wu, T. N. Davidson, and Q. Jin, "Wavelet packet division multiplexing and wavelet packet design under timing error effects," *IEEE Trans. Signal Processing*, vol. 45, pp. 2877–2890, Dec. 1997.

[24] G. Wornell, "Emerging application of multirate signal processing and wavelet in digital communications," *Proc. IEEE*, vol. 84, pp. 586–603, Apr. 1996.

[25] X.-G. Xia, "New precoding for intersymbol interference cancellation using nonmaximally decimated multirate filterbanks with ideal FIR equalizers," *IEEE Trans. Signal Processing*, vol. 45, pp. 2431–2441, Oct. 1997.

[26] J. Yang and S. Roy, "On joint transmitter and receiver optimization for multiple-input-multiple-output (MIMO) transmission systems," *IEEE Trans. Commun.*, vol. 42, pp. 3221–3231, Dec. 1994.

Yanwu Ding received the B.Eng. degree from Southwest Jiaotong University, Chengdu, Sichuan Province, China in 1985 and the M.Eng. degree from McMaster University, Hamilton, ON, Canada, in 2002.

She is currently a research associate with the Department of Electrical and Computer Engineering, McMaster University. She has previously held positions with the Department of Communication and Control Engineering, Northern Jiaotong University, Beijing, China, and the Information Center of State Bureau of Quality Technical Supervision, Beijing, China. Her research interests are in signal processing and communications.

Ms. Ding was a recipient of the Outstanding Thesis Research Award from McMaster University in 2002 and the Outstanding Graduate Award from Southwest Jiaotong University in 1985.

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, Oxford, U.K., in 1995.

He is currently an assistant professor with the Department of Electrical and Computer Engineering, McMaster University, Hamilton, ON, Canada. 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 at the Communications Research Laboratory, McMaster University, the Adaptive Signal Processing Laboratory, UWA, and the Australian Telecommunications Research Institute, Curtin University of Technology, Perth.

Dr. Davidson was awarded the 1991 J. A. Wood Memorial Prize (for "the most outstanding [UWA] graduand" in the pure and applied sciences) and the 1991 Rhodes Scholarship for Western Australia.

Zhi-Quan Luo (SM'03) was born in Nanchang, Jiangxi province, China. He received the B.Sc. degree in applied mathematics in 1984 from Peking University, Beijing, China. From 1984 to 1985, he studied at the Nankai Institute of Mathematics, Tianjin, China. He received the Ph.D. degree in operations research from the Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge, in 1989.

In 1989, he joined the Department of Electrical and Computer Engineering, McMaster University, Hamilton, ON, Canada, where he is now a Professor and holds the Canada Research Chair in Information Processing. His research interests lie in the union of large-scale optimization, information theory and coding, data communications, and signal processing. He is presently serving as an associate editor for *Journal of Optimization Theory and Applications*, *SIAM Journal on Optimization*, *Mathematics of Computation*, *Mathematics of Operations Research*, and *Optimization and Engineering*.

Prof. Luo is a member of SIAM and MPS and is an Associate Editor for the IEEE TRANSACTIONS ON SIGNAL PROCESSING.

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

He started working at the Transmission Division of Plessey Telecommunications Research Ltd., London, in 1969. In October 1970, he was on leave from Plessey, pursuing postgraduate studies and research at the 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, Halifax, NS, Canada, and in 1981, he 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 from 1986 to 1987 and again from 1988 to 1994. He was on leave as Visiting Professor with the Department of Electronic Engineering, Chinese University of Hong Kong, from 1997 to 1999. At present, he holds the title of NSERC-Mitel Professor of Signal Processing and is the Director of the Communication Technology Research Centre at McMaster University. His research interest is in signal processing and communication theory, and he has published over 170 papers in this area.

Prof. Wong was the recipient of the IEE Overseas Premium for the best paper in 1989, and 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 also served as an Associate Editor of the IEEE TRANSACTIONS ON SIGNAL PROCESSING from 1996 to 1998 and has been the chairman of the Sensor Array and Multichannel Signal Processing Technical Committee of the Signal Processing Society since 1998. 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 honored 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.