Optimal ISI-Free DMT Transceivers for Distorted Channels with Colored Noise

Yuan-Pei Lin, *Member, IEEE,* and See-May Phoong, *Member, IEEE*

***Abstract*—The design of optimal DMT transceivers for distorted channel with colored noise has been of great interest. Of particular interest is the class of block based DMT, where the transmitter and the receiver consist of constant matrices. Two types of block-based DMT transceivers will be considered: the DMT system with zero padding (ZP-DMT) and the DMT system with general prefix (GP-DMT). We will derive the bit allocation formula. For a given channel and channel noise spectrum, we will design the ISI-free optimal transceiver that minimizes the transmission power for a given transmission rate and probability of error. For both ZP-DMT and GP-DMT systems, the optimal ISI-free transceiver can be given in closed from. We will see that for both classes, the optimal transceiver has an orthogonal transmitter. Simulation shows that the optimal DMT system can achieve the same transmission rate and the same probability of error with a much lower transmission power compared with other existing DMT systems.**

I. INTRODUCTION

THERE has been great interest in the design of DMT systems recently. Fig. 1 shows an example of an M-band DMT transceiver over channel $C(z)$ with additive noise $\nu(n)$. The example is the so-called *block-based* DMT, where the transmitter and the receiver consist of constant matrices. The encoding at the transmitter end and the decoding at the receiver end are done blockwise. When the receiver outputs are identical to the transmitter inputs in the absence of channel noise, the transceiver is said to be *ISI free*. The introduction of redundancy is typical of DMT transmitters so that the receiver can cancel ISI due to the channel. The cyclic prefix and zero padding (or trailing zeros) are commonly used forms of redundancy. For example, the cyclic prefix is used in DFT-based DMT systems [1], and zero padding is considered in [2]–[5].

Block-based DMT transceivers have been studied extensively. In the commonly used DFT-based DMT, the transmitter and the receiver are DFT matrices [1]. In [2], more general orthogonal matrices are proposed. It is shown therein that for additive white Gaussian noise (AWGN) frequency-selective channels, the optimal orthogonal transmitter consists of eigenvectors associated with the channel. In [3], optimal transceivers that minimize the total output noise power are developed.

Manuscript received August 3, 2000; revised July 18, 2001. This work was supported in part by National Science Council of Taiwan under Contracts 89-2213-E-009-118 and NSC 89-2213-E-002-063, Ministry of Education, under Contract 89-E-FA06-2-4, Taiwan, R.O.C., and the Lee and MTI Center for Networking Research. The associate editor coordinating the review of this paper and approving it for publication was Prof. Arnab K. Shaw.

Y.-P. Lin is with the Department of Electrical and Control Engineering, National Chiao-Tung University, Hsinchu, Taiwan, R.O.C. (e-mail: ypl@cc.nctu.edu.tw).

S.-M. Phoong is with the Graduate Institute of Communication Engineering and the Department of Electrical Engineering, National Taiwan University, Taipei, Taiwan, R.O.C.

Publisher Item Identifier S 1053-587X(01)09230-3.

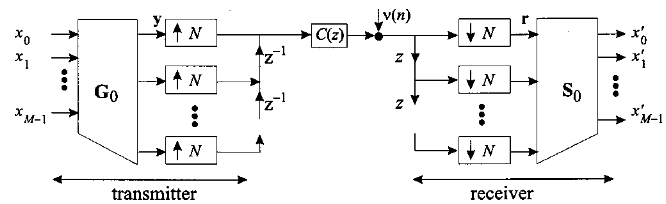


Fig. 1. M-band block-based DMT transceiver over channel $C(z)$ with noise $\nu(n)$.

Information rate optimized DMT systems are considered in [4]. In [5], the authors start from an orthogonal transmitter with zero padding. The optimal orthogonal transmitter that minimizes the transmission power for a fixed probability of error and transmission bit rate is derived.

In this paper, we consider general block-based DMT transceivers: DMT systems with zero padding (ZP-DMT) and DMT systems with general prefix (GP-DMT). The transmitters are general constant matrices that are not limited to the orthogonal class. The widely used DMT systems with the cyclic prefix fall into the category of the GP-DMT systems. For these two types of systems, optimal bit allocation will be derived. Under optimal bit allocation, we will develop optimal transceivers that minimize the transmission power for any given transmission rate and probability of error. It turns out that for both ZP-DMT and GP-DMT systems, the optimal transceiver that minimizes the transmission power has an orthogonal transmitter. That is, there is no loss of generality in using orthogonal transmitter for designing optimal block-based DMT systems. Furthermore, for both ZP-DMT and GP-DMT systems, the optimal transceivers can be given in closed form in terms of the given channel and channel noise. We will see that in most cases, the ZP-DMT system outperforms the GP-DMT system. However, the superiority of the ZP-DMT system is not guaranteed as these are two different classes of system. We will also see one toy example where the GP-DMT system performs better.

The sections are organized as follows. Sections II–V are devoted to the class of ZP-DMT systems. Section II gives a review of the ISI-free transceiver solutions. The class of ZP-DMT transceivers with the ISI-free property will be parameterized. Section III gives the measure of optimality in this paper and states the problem. For a given transmission bit rate, probability of error and a given ISI-free transceiver, we derive the optimal bit allocation in Section IV. Based on the optimal bit allocation, the optimal ZP-DMT transceiver is presented in Section V. In Section VI, the GP-DMT systems are studied in steps paralleling those of ZP-DMT systems in Sections II–V. Comparisons of the proposed optimal system and other existing DMT systems are presented in Section VI. Examples are given in Section VIII. Concluding remarks are given in Section IX.

A. Connection with Subband Coding

The connection between an M-band filter bank and an M-band DMT transceiver (or transmultiplexer) has been well-known [6]–[9]. When the analysis and synthesis banks of a perfect reconstruction filterbank are interchanged, the new structure becomes a DMT transceiver. The M-band DMT system is often considered to be the dual of an M-band subband coder; the block-based DMT system can be considered the dual of the transform coder. In subband coding, we know that optimal orthogonal transform coder is as good as any biorthogonal transform coder [10]. However, the design problem of subband coders is different from that of DMT transceivers. The analysis for subband coding does not carry through for DMT systems in general.

The DMT systems obtained from interchanging the analysis and synthesis banks have interpolation ratio $N = M$, and there is no redundancy in the transmitter output. When the channel $P(z)$ is a delay, the M-band DMT transceiver is ISI free if the corresponding filterbank has perfect reconstruction [6]. However, when the channel is not a delay, the perfect reconstruction property of the filterbank no longer translates to the ISI-free property of DMT transceivers. It has been shown that with $N = M$, ISI-free solutions can only be obtained with IIR transceivers. The IIR filters are unstable if the channel does not have minimum phase [11].

In practical designs of DMT systems, unlike in transform coding, the interpolation ratio N is larger than M to enable ISI cancellation. For the two classes of block based transceivers considered in this paper (the GP-DMT and ZP-DMT systems), we cannot draw an analogy to the design problem of subband coders. Despite the differences in the designs of optimal subband coders and DMT transceivers, it turns out that the results we will show in this paper are similar to those of optimal transform coding. For both GP-DMT and ZP-DMT systems, orthogonal transmitters are as good as any transmitter.

B. Notations

1) Boldfaced lower-case letters are used to represent vectors, and boldfaced upper-case letters are reserved for matrices. The notation $\mathbf{A}^T$ represents the transpose of $\mathbf{A}$.
2) The function $\mathcal{E}[y]$ denotes the expected value of the random variable y.
3) The notation $\mathbf{I}_M$ is used to represent the $M \times M$ identity matrix. The subscript is omitted whenever the size is clear from the context.
4) The notation $\mathbf{J}_M$ is used to represent the $M \times M$ reversal matrix. For example, for $M = 3$

$$\mathbf{J}_3 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}.$$

II. ISI-FREE BLOCK-BASED DMT WITH ZERO PADDING [3], [4]

Consider Fig. 1, where an M-band block-based DMT system is shown. Usually, the channel is modeled as an LTI filter $C(z)$ with additive noise $\nu(n)$. Assume that $\nu(n)$ is a zero-mean wide sense stationary random process and that $C(z)$ is an FIR filter of order L. In some applications, e.g., DSL, where the channel has a very long impulse response, typically, an equalizer is used to shorten the impulse response to an FIR filter with a small order.

For block-based DMT systems, the length of redundant samples is chosen to be L so that the receiver can remove ISI due to $C(z)$, and decoding can be performed blockwise. Therefore, the interpolation ratio N is $N = M + L$. In the case of zero padding (ZP-DMT), the transmitter and receiver are of the form

$$\mathbf{G}_0 = \begin{pmatrix} \mathbf{G} \\ \mathbf{0} \end{pmatrix}, \quad \text{and} \quad \mathbf{S}_0 = \mathbf{S}$$

where $\mathbf{G}$ is of dimensions M by M, the bottom matrix $\mathbf{0}$ is of dimensions L by M, and $\mathbf{S}$ is an M by N constant matrix. The system is ISI free if $x'_k(n) = x_k(n)$, for $k = 0, 1, \ldots M - 1$, in the absence of channel noise. Due to the padded L zeros, there is no overlapping between adjacent blocks. Processing is performed on a block-by-block basis; the dependence on n will be omitted.

With the assumption that the channel $C(z)$ is an FIR filter of order L, we can write

$$C(z) = c_0 + c_1 z^{-1} + \cdots + c_L z^{-L}$$

with nonzero c_0 and c_L. The received vector $\mathbf{r}$ shown in Fig. 1 can be given in terms of the transmitted vector as $\mathbf{r} = \mathbf{C}(y_0 \; y_1 \; \cdots \; y_{M-1})^T$, where y_k denotes the kth element of $\mathbf{y}$, and $\mathbf{C}$ is an N by M lower triangular Toeplitz matrix given by

$$\mathbf{C} = \begin{pmatrix} c_0 & 0 & \cdots & 0 & & 0 \\ c_1 & c_0 & & \vdots & & 0 \\ \vdots & & \ddots & & & \vdots \\ c_L & c_{L-1} & & & & 0 \\ 0 & c_L & & & & 0 \\ \vdots & \vdots & \ddots & & & \vdots \\ 0 & 0 & & c_L & \cdots & c_0 \\ \vdots & \vdots & & & \ddots & \vdots \\ 0 & 0 & & 0 & \cdots & c_L \end{pmatrix}.$$

Fig. 1 can be redrawn as Fig. 2 using the matrix representation of the channel. The $N \times 1$ vector $\boldsymbol{\nu}$ shown in Fig. 2 is the blocked version of the scalar noise process $\nu(n)$.

The ZP-DMT transceivers with an ISI-free property can be parameterized using the matrix representation. Applying the singular value decomposition (SVD), we can express $\mathbf{C}$ as

$$\mathbf{C} = \underbrace{(\mathbf{U}_0 \quad \mathbf{U}_1)}_{\mathbf{U}} \begin{pmatrix} \mathbf{\Lambda} \\ \mathbf{0} \end{pmatrix} \mathbf{V} = \mathbf{U} \begin{pmatrix} \mathbf{\Lambda} \\ \mathbf{0} \end{pmatrix} \mathbf{V} \qquad (1)$$

where $\mathbf{U}$ and $\mathbf{V}$ are orthonormal matrices of dimensions, respectively, $N \times N$ and $M \times M$, and $\mathbf{\Lambda}$ is an $M \times M$ diagonal matrix. The column vectors of $\mathbf{U}$ are the eigenvectors of the $N \times N$ matrix $\mathbf{C}\mathbf{C}^T$, and the column vectors of $\mathbf{V}^T$ are the eigenvectors of the $M \times M$ matrix $\mathbf{C}^T\mathbf{C}$. The diagonal elements of $\mathbf{\Lambda}$ are the eigenvalues of $\mathbf{C}^T\mathbf{C}$. As $\mathbf{C}$ has full rank M, the matrix $\mathbf{\Lambda}$ is nonsingular. From Fig. 2, we see that the

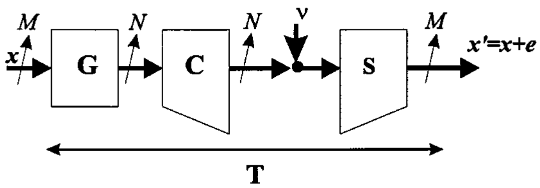


Fig. 2. Matrix representation of DMT systems with zero padding.

overall DMT system is $\mathbf{T} = \mathbf{SCG}$. The ISI-free condition becomes $\mathbf{SCG} = \mathbf{I}_M$. Since $\mathbf{G}$ is M by M, the ISI-free condition implies that $\mathbf{G}$ is nonsingular. Premultiplying the above equation by $\mathbf{G}$ and post-multiplying by $\mathbf{G}^{-1}$, we get $(\mathbf{GS})\mathbf{C} = \mathbf{I}_M$. This means that $\mathbf{GS}$ is a left inverse of $\mathbf{C}$, which is of dimensions $N \times M$, and its left inverses are not unique. Therefore, we can write $\mathbf{S} = \mathbf{G}^{-1}\mathbf{B}$, where $\mathbf{B}$ is any left inverse of $\mathbf{C}$ (i.e., $\mathbf{BC} = \mathbf{I}_M$). Using the SVD of $\mathbf{C}$ as in (1), its left inverse $\mathbf{B}$ can be expressed as

$$\mathbf{B} = \mathbf{V}^T \mathbf{\Lambda}^{-1} [\mathbf{I}\ \mathbf{A}] \mathbf{U}^T$$

where $\mathbf{A}$ is an arbitrary $M \times L$ matrix. Summarizing the results, we have the following theorem.

Theorem 1—Parameterization of ZP-DMT Systems[3], [4]: The block-based DMT system with zero padding in Fig. 1 is ISI free if and only if the zero-padded transceiver satisfies the following.

i) $\mathbf{G}$ is an $M \times M$ nonsingular matrix.
ii) $\mathbf{S} = \mathbf{G}^{-1}\mathbf{B}$, where $\mathbf{B} = \mathbf{V}^T \mathbf{\Lambda}^{-1} [\mathbf{I}\ \mathbf{A}] \mathbf{U}^T$, for an arbitrary $M \times L$ matrix $\mathbf{A}$. The matrices $\mathbf{U}$, $\mathbf{V}$, and $\mathbf{\Lambda}$ are as in the SVD of $\mathbf{C}$ in (1).

In the above formulation, $\mathbf{U}$, $\mathbf{V}$, and $\mathbf{\Lambda}$ are determined by the channel matrix $\mathbf{C}$. After the ISI-free condition is imposed, the parameters left in the transceiver design are the nonsingular matrix $\mathbf{G}$ and the $M \times L$ matrix $\mathbf{A}$, which is completely arbitrary. The development in Sections III–V will be based on the ISI-free transceiver described in Theorem 1.

III. Problem Formulation

In this paper, we assume that the inputs x_k are PAM symbols of b_k bits. Without much loss of generality, we further assume that x_k have zero mean and that they are uncorrelated with each other. That is

$$\mathcal{E}[x_k x_m] = \sigma_{x_k}^2 \delta(k-m).$$

This can always be obtained with proper interleaving. The average bit rate per symbol in this case becomes

$$b = \frac{1}{M} \sum_{k=0}^{M-1} b_k.$$

The actual transmission rate is $R_b = (M/N)\, b$ bits per sample due to zero padding.

The transmission power P is the average energy of the vector $\mathbf{y} = (y_0\ y_1\ \cdots\ y_{N-1})^T$, as shown in Fig. 1

$$P = \frac{1}{N} \sum_{k=0}^{M-1} \sigma_{y_k}^2.$$

The summation has only M terms as the last L elements of $\mathbf{y}$ are the padded zeros. As the inputs x_k are uncorrelated and have zero mean, $\sigma_{y_k}^2$ is given by

$$\sigma_{y_k}^2 = \sum_{n=0}^{M-1} [\mathbf{G}]_{kn}^2 \sigma_{x_n}^2.$$

Using this expression, we can write the transmission power as

$$P = \frac{1}{N} \sum_{n=0}^{M-1} \sigma_{x_n}^2 \left[\sum_{k=0}^{M-1} [\mathbf{G}]_{kn}^2 \right].$$

Define $||\mathbf{g}_k||_2^2 = \sum_{\ell=0}^{M-1} [\mathbf{G}]_{\ell k}^2$ as the energy of the kth column of $\mathbf{G}$. The transmission power can be written as

$$P = \frac{1}{N} \sum_{k=0}^{M-1} \sigma_{x_k}^2 ||\mathbf{g}_k||_2^2. \tag{2}$$

Note that the quantity $\sigma_{x_k}^2 ||\mathbf{g}_k||_2^2$ is the transmission power due to the kth band. Under the ISI-free condition, for a fixed bit rate and a fixed probability of error P_e, we will find the ZP-DMT transceiver that minimizes the transmission power. The optimization process involves two steps. We will show in Section IV that the bits b_k can be optimally allocated to minimize the transmission power for any given transceiver. Under the optimal bit allocation, the optimal ZP-DMT transceiver will then be derived in Section V.

IV. Optimal Bit Allocation

For a given transceiver, a fixed probability of error P_e, and the average bit rate per input symbol b, we present the optimal bit allocation $\{b_k\}_{k=0}^{M-1}$ with $b = (1/M) \sum_{k=0}^{M-1} b_k$ such that the transmission power in (2) is minimized.

At the receiver end, the output of the kth band is $x'_k = x_k + e_k$, where e_k comes entirely from channel noise as the transceiver achieves zero ISI. Define the $M \times 1$ output noise vector as $\mathbf{e} = (e_0\ e_1\ \cdots\ e_{M-1})^T$; then

$$\mathbf{e} = \mathbf{S}\boldsymbol{\nu} = \mathbf{G}^{-1}\mathbf{B}\boldsymbol{\nu}.$$

Assuming the PAM symbols of the kth band carry b_k bits, the probability of error for the kth band is given by [12]

$$P_e(k) = 2(1 - 2^{-b_k}) Q\left(\sqrt{\frac{3\sigma_{x_k}^2}{(2^{2b_k} - 1)\sigma_{e_k}^2}} \right)$$

where

$$Q(\tau) = \frac{1}{\sqrt{2\pi}} \int_\tau^\infty e^{-t^2/2}\, dt, \qquad \tau \ge 0.$$

For a fixed probability of error P_e across all bands, we need to have $P_e(0) = P_e(1) = \cdots = P_e(M-1) = P_e$. Under the high bit rate assumption $2^{b_k} - 1 \approx 2^{b_k}$, we can see that $\sigma_{x_k}^2$ and $\sigma_{e_k}^2$ satisfy

$$2Q\left(\sqrt{\frac{3\sigma_{x_k}^2}{2^{2b_k} \sigma_{e_k}^2}} \right) = P_e.$$

We can rearrange the above equation as

$$\sigma_{x_k}^2 = c2^{2b_k}\sigma_{e_k}^2, \quad \text{where} \quad c = \frac{1}{3}\left(Q^{-1}\{P_e/2\}\right)^2. \tag{3}$$

For the fixed probability of error, the transmission power in (2) becomes

$$P = \frac{c}{N}\sum_{k=0}^{M-1} 2^{2b_k}\,\sigma_{e_k}^2\|\mathbf{g}_k\|_2^2.$$

Applying the arithmetic mean-geometric mean inequality, we obtain

$$P \geq \frac{cM}{N}\,2^{2b}\prod_{k=0}^{M-1}\left(\sigma_{e_k}^2\|\mathbf{g}_k\|_2^2\right)^{1/M} \triangleq P_{opt,\,bit}. \tag{4}$$

The equality holds if and only if $2^{2b_k}\sigma_{e_k}^2\|\mathbf{g}_k\|_2^2$ are the same for $k = 0, 1, \ldots, M-1$. Notice that $P_{opt,\,bit}$ depends only on b and $\sigma_{e_k}^2\|\mathbf{g}_k\|_2^2$, where $\sigma_{e_k}^2$ is determined once the receiver is known, and $\|\mathbf{g}_k\|_2^2$ is determined once the transmitter is given. Therefore, when the transceiver is given and average bit rate per symbol b is fixed, $P_{opt,\,bit}$ is the lower bound of the transmission power independent of the bit allocation $\{b_k\}_{k=0}^{M-1}$. The lower bound $P_{opt,\,bit}$ is achieved if and only if bits are allocated such that $2^{2b_k}\sigma_{e_k}^2\|\mathbf{g}_k\|_2^2$ are equalized. Solving for the optimal b_k, we have

$$b_k = b - \log_2\left(\sigma_{e_k}\|\mathbf{g}_k\|_2\right) + \frac{1}{M}\,\log_2\left(\prod_{\ell=0}^{M-1}\sigma_{e_\ell}\|\mathbf{g}_\ell\|_2\right). \tag{5}$$

Remarks:

1) The optimal bit allocation equalizes the terms $2^{2b_k}\sigma_{e_k}^2\|\mathbf{g}_k\|_2^2$. From (3), this means that the transmission power contributed from each band is the same.
2) If the transmitting vectors $\mathbf{g}_k$ have equal energy (i.e., $\|\mathbf{g}_0\|_2 = \|\mathbf{g}_1\|_2 = \cdots = \|\mathbf{g}_{M-1}\|_2$), then the symbol variances $\sigma_{x_k}^2$ are also equalized. In this case, we can see from (3) that more bits are assigned to bands where $\sigma_{e_k}^2$ is small and fewer bits where $\sigma_{e_k}^2$ is large. This is similar to the bit allocation in water filling scheme. More bits are transmitted in less noisy bands.

V. Optimal ZP-DMT Transceivers

From the previous derivation, we know that the transceiver in Theorem 1 renders the received signal ISI free for all possible choices of nonsingular $\mathbf{G}$ and an arbitrary $M \times L$ matrix $\mathbf{A}$. Under optimal bit allocation, the power lower bound $P_{opt,\,bit}$ depends on the choice of the transceiver. Next, we will show how to choose $\mathbf{G}$ and $\mathbf{A}$ so that $P_{opt,\,bit}$ is minimized. For a given $\mathbf{A}$, we will derive the optimal $\mathbf{G}$, based on which $\mathbf{A}$ is optimized.

Optimal $\mathbf{G}$*:* We first express the quantity $P_{opt,\,bit}$ in (4) in terms of $\mathbf{G}$. The energy of the kth column of $\mathbf{G}$ is $\|\mathbf{g}_k\|_2^2 = [\mathbf{G}^T\mathbf{G}]_{kk}$. Letting $\mathbf{q} = \mathbf{B}\boldsymbol{\nu}$, then $\mathbf{e} = \mathbf{G}^{-1}\mathbf{q}$. The $M \times M$ autocorrelation matrix $\mathbf{R}_e$ of the noise vector $\mathbf{e}$ is given by $\mathbf{R}_e = \mathbf{S}\mathbf{R}_\nu\mathbf{S}^T = \mathbf{G}^{-1}\mathbf{R}_q\mathbf{G}^{-T}$, where $\mathbf{R}_q$ is the autocorrelation matrix of the vector $\mathbf{q}$. The output noise $\sigma_{e_k}^2$ of the kth band is equal to $[\mathbf{R}_e]_{kk}$ or $[\mathbf{G}^{-1}\mathbf{R}_q\mathbf{G}^{-T}]_{kk}$. Therefore, (4) can be rewritten as

$$P_{opt,\,bit} = c\frac{M}{N}\,2^{2b}\left(\prod_{k=0}^{M-1}\left[\mathbf{G}^T\mathbf{G}\right]_{kk}\left[\mathbf{G}^{-1}\mathbf{R}_q\mathbf{G}^{-T}\right]_{kk}\right)^{1/M}.$$

The Hadamard inequality for an $M \times M$ positive definite matrix $\mathbf{P}$ states that [13]

$$\prod_{k=0}^{M-1}[\mathbf{P}]_{kk} \geq \det(\mathbf{P})$$

with equality if and only if $\mathbf{P}$ is a diagonal matrix. Using this inequality, we have

$$\begin{aligned} P_{opt,\,bit} &\geq c\frac{M}{N}\,2^{2b}\left(\det\left(\mathbf{G}^T\mathbf{G}\right)\det\left(\mathbf{G}^{-1}\mathbf{R}_q\mathbf{G}^{-T}\right)\right)^{1/M} \\ &= c\frac{M}{N}\,2^{2b}(\det\mathbf{R}_q)^{1/M} \triangleq P_{opt,\,bit,\,\mathbf{G}}. \end{aligned} \tag{6}$$

The equality holds if and only if i) $\mathbf{G}^T\mathbf{G}$ is diagonal and ii) $\mathbf{G}^{-1}\mathbf{R}_q\mathbf{G}^{-T}$ is diagonal. The lower bound $P_{opt,\,bit,\,G}$ does not depend on the transmitter $\mathbf{G}$, and it is achieved if and only if $\mathbf{G}$ satisfy both conditions i) and ii). The first condition means that $\mathbf{G}$ is an orthogonal matrix. The second condition means that $\mathbf{G}^{-1}$ decorrelates the noise vector $\mathbf{q}$. Let the Schur decomposition of $\mathbf{R}_q$ be

$$\mathbf{R}_q = \mathbf{Q}\boldsymbol{\Sigma}\mathbf{Q}^T.$$

Combining conditions i) and ii), we see that $\mathbf{G}$ is an orthogonal matrix of the form $\mathbf{G} = \mathbf{Q}\mathbf{D}$, where $\mathbf{D}$ is a diagonal nonsingular matrix. Since $\mathbf{G} = \mathbf{Q}\mathbf{D}$ achieves the lower bound for any nonsingular $\mathbf{D}$, without loss of generality, we can choose $\mathbf{D} = \mathbf{I}$, and

$$\mathbf{G} = \mathbf{Q}, \quad \text{and} \quad \mathbf{S} = \mathbf{Q}^T\mathbf{B}$$

where

$$\mathbf{B} = \mathbf{V}^T\boldsymbol{\Lambda}^{-1}[\mathbf{I}\ \mathbf{A}]\mathbf{U}^T. \tag{7}$$

Remark: To understand the effect of nonidentity $\mathbf{D}$, we let $\mathbf{D}' = \mathrm{diag}(d_0\ \ 1\ \ \cdots\ \ 1)$. At the transmitting side, $\|\mathbf{g}_0'\|_2^2 = d_0^2\|\mathbf{g}_0\|_2^2$, and at the receiving end, $\sigma_{e_0'}^2 = \sigma_{e_0}^2/d_0^2$. As the product $\sigma_{e_0'}^2\|\mathbf{g}_0'\|_2^2 = \sigma_{e_0}^2\|\mathbf{g}_0\|_2^2$ remains the same, bit allocation remains the same. It follows from (3) that $\sigma_{x_0'}^2 = \sigma_{x_0}^2/d_0^2$; the transmission power $\|\mathbf{g}_0'\|_2^2\sigma_{x_0'}^2$ due to the 0th band remains the same.

Optimal $\mathbf{A}$*:* From (6), we see that given any $\mathbf{A}$, the achievable lower bound is $P_{opt,\,bit,\,\mathbf{G}} = c(M/N)\,2^{2b}(\det\mathbf{R}_q)^{1/M}$. The matrix $\mathbf{A}$ should be chosen such that $\det(\mathbf{R}_q)$ is minimized. Using the facts that $\mathbf{R}_q = \mathbf{B}\mathbf{R}_\nu\mathbf{B}^T$ and $\mathbf{B} = \mathbf{V}^T\boldsymbol{\Lambda}^{-1}[\mathbf{I}\ \mathbf{A}]\mathbf{U}^T$, we get

$$\det(\mathbf{R}_q) = \det\left(\mathbf{V}^T\boldsymbol{\Lambda}^{-1}[\mathbf{I}\ \mathbf{A}]\mathbf{U}^T\mathbf{R}_\nu\mathbf{U}\begin{pmatrix}\mathbf{I}\\ \mathbf{A}^T\end{pmatrix}\boldsymbol{\Lambda}^{-1}\mathbf{V}\right).$$

Since $\mathbf{V}$ is orthonormal and $\det(\mathbf{V}) = 1$, we can simplify the above expression to

$$\det(\mathbf{R}_q) = \det(\boldsymbol{\Lambda}^{-2})\det\left([\mathbf{I}\ \mathbf{A}]\mathbf{U}^T\mathbf{R}_\nu\mathbf{U}\begin{pmatrix}\mathbf{I}\\ \mathbf{A}^T\end{pmatrix}\right). \tag{8}$$

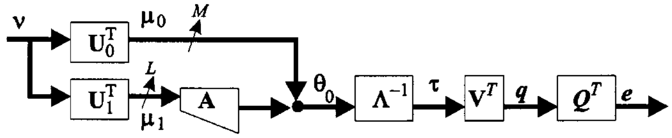


Fig. 3. Receiver in the form of $\mathbf{S} = \mathbf{Q}^T\mathbf{V}^T\mathbf{\Lambda}^{-1}[\mathbf{I}\ \mathbf{A}]\mathbf{U}^T$ for the derivation of an optimal $\mathbf{A}$.

Note that $\mathbf{\Lambda}$, $\mathbf{U}$, and $\mathbf{R}_\nu$ are fixed once the channel and input noise are given. Thus, the optimal $\mathbf{A}$ is such that $\det([\mathbf{I}\ \mathbf{A}]\mathbf{U}^T\mathbf{R}_\nu\mathbf{U}[\mathbf{I}\ \mathbf{A}]^T)$ is minimized. In the following, we will derive a closed-form expression for such an optimal $\mathbf{A}$.

For the convenience of explanation, we draw the receiver in the form of (7) as in Fig. 3, where $\mathbf{U}_0$ and $\mathbf{U}_1$ are submatrices of $\mathbf{U}$, as in (1). The vectors $\boldsymbol{\mu}_0$ and $\boldsymbol{\mu}_1$ in Fig. 3 are given by

$$\boldsymbol{\mu} = \begin{pmatrix} \boldsymbol{\mu}_0 \\ \boldsymbol{\mu}_1 \end{pmatrix} = \begin{pmatrix} \mathbf{U}_0^T\boldsymbol{\nu} \\ \mathbf{U}_1^T\boldsymbol{\nu} \end{pmatrix} = \mathbf{U}^T\boldsymbol{\nu}. \tag{9}$$

Define the vectors $\boldsymbol{\theta}$, $\boldsymbol{\theta}_0$, and $\boldsymbol{\theta}_1$ as

$$\boldsymbol{\theta} = \begin{pmatrix} \boldsymbol{\theta}_0 \\ \boldsymbol{\theta}_1 \end{pmatrix} = \begin{pmatrix} \mathbf{I}_M & \mathbf{A} \\ \mathbf{0} & \mathbf{I}_L \end{pmatrix} \begin{pmatrix} \boldsymbol{\mu}_0 \\ \boldsymbol{\mu}_1 \end{pmatrix} = \begin{pmatrix} \boldsymbol{\mu}_0 + \mathbf{A}\boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_1 \end{pmatrix}. \tag{10}$$

The vector $\boldsymbol{\theta}_0$ is as illustrated in Fig. 3. Let $\mathbf{R}_\theta$, $\mathbf{R}_{\theta_0}$, and $\mathbf{R}_{\theta_1}$ be the autocorrelation matrices of $\boldsymbol{\theta}$, $\boldsymbol{\theta}_0$, and $\boldsymbol{\theta}_1$ respectively. These matrices are related as

$$\mathbf{R}_\theta = \begin{pmatrix} \mathbf{R}_{\theta_0} & \mathbf{R}_{\theta_{01}} \\ \mathbf{R}_{\theta_{01}}^T & \mathbf{R}_{\theta_1} \end{pmatrix}$$

where $\mathbf{R}_{\theta_{01}} = \mathcal{E}[\boldsymbol{\theta}_0\boldsymbol{\theta}_1^T]$. Using (9) and (10), it is not difficult to see that

$$\mathbf{R}_{\theta_0} = [\mathbf{I}\ \mathbf{A}]\mathbf{U}^T\mathbf{R}_\nu\mathbf{U}[\mathbf{I}\ \mathbf{A}]^T.$$

Therefore, the optimal $\mathbf{A}$ should be such that $\det(\mathbf{R}_{\theta_0})$ is minimized.

Using the Fischer's inequality for positive definite matrices and its extension given in the Appendix, we have

$$\det(\mathbf{R}_{\theta_0}) \geq \frac{\det(\mathbf{R}_\theta)}{\det(\mathbf{R}_{\theta_1})}$$

with equality if and only if $\mathbf{R}_{\theta_{01}} = \mathbf{0}$. Using (9) and (10), we have $\det(\mathbf{R}_\theta) = \det(\mathbf{R}_\mu) = \det(\mathbf{R}_\nu)$ and $\det(\mathbf{R}_{\theta_1}) = \det(\mathbf{R}_{\mu_1})$. Using these relations, the above equation can be expressed as

$$\det(\mathbf{R}_{\theta_0}) \geq \frac{\det(\mathbf{R}_\nu)}{\det(\mathbf{R}_{\mu_1})}.$$

Note that the lower bound is independent of $\mathbf{A}$. Moreover, this lower bound can be obtained if and only if $\mathbf{A}$ is chosen such that

$$\mathbf{R}_{\theta_{01}} = \mathcal{E}\left[(\boldsymbol{\mu}_0 + \mathbf{A}\boldsymbol{\mu}_1)\boldsymbol{\mu}_1^T\right] = \mathbf{0}.$$

Solving the above equation, the optimal $\mathbf{A}$ is uniquely given by

$$\mathbf{A} = -\mathbf{U}_0^T\mathbf{R}_\nu\mathbf{U}_1\left(\mathbf{U}_1^T\mathbf{R}_\nu\mathbf{U}_1\right)^{-1}. \tag{11}$$

Summarizing the results, we have theorem.

Theorem 2: The optimal $\mathbf{A}$ that minimizes $\det([\mathbf{IA}]\mathbf{U}^T\mathbf{R}_\nu\mathbf{U}[\mathbf{I}\ \mathbf{A}]^T)$ is such that $\mathcal{E}[(\boldsymbol{\mu}_0 + \mathbf{A}\boldsymbol{\mu}_1)\boldsymbol{\mu}_1^T] = \mathbf{0}$, where $\boldsymbol{\mu}_0$ and $\boldsymbol{\mu}_1$ are given in (9). Equivalently, $\mathbf{A}$ is given uniquely by (11). Furthermore, the minimum is given by

$$\det(\mathbf{R}_\nu)/\det(\mathbf{U}_1^T\mathbf{R}_\nu\mathbf{U}_1).$$

Remarks:

- Theorem 2 states that $\mathbf{A}$ should be chosen such that $\mathcal{E}[(\boldsymbol{\mu}_0 + \mathbf{A}\boldsymbol{\mu}_1)\boldsymbol{\mu}_1^T] = \mathbf{0}$. This implies that the optimal solution of $\mathbf{A}$ is the optimal estimator of $\boldsymbol{\mu}_0$, given the observation of $-\boldsymbol{\mu}_1$.
- The solution of $\mathbf{A}$ in (11) minimizes not only $\det(\mathbf{R}_q) = \det(\mathbf{R}_e)$ but the total output noise power given by $\sum_{k=0}^{M-1}\sigma_{e_k}^2$ or $trace(\mathbf{R}_e)$ as well. To see this, we consider Fig. 3. Since the vectors $\mathbf{e}$ and $\boldsymbol{\tau}$ are related through the orthonormal matrix $\mathbf{Q}^T\mathbf{V}^T$, which preserves traces, we have $trace(\mathbf{R}_\tau) = trace(\mathbf{R}_e)$. On the other hand, $[\mathbf{R}_\tau]_{kk} = [\mathbf{R}_{\theta_0}]_{kk}/[\mathbf{\Lambda}]_{kk}^2$, and minimization of $trace(\mathbf{R}_\tau)$ can be achieved if each term $[\mathbf{R}_{\theta_0}]_{kk}$ can be individually minimized. Notice that $\boldsymbol{\theta} = \boldsymbol{\mu}_0 + \mathbf{A}\boldsymbol{\mu}_1$. Upon invoking the orthogonality principle, $\mathbf{A}$ should be chosen as the optimal estimator of $\boldsymbol{\mu}_0$ based on the observation $\boldsymbol{\mu}_1$. Therefore, the solution of $\mathbf{A}$ given in (11) is also optimal for minimizing the total output noise $trace(\mathbf{R}_e)$.

We can summarize our results as follows:

Theorem 3 : Consider the M-Band ZP-DMT System in Fig. 2. Assume that the inputs are PAM symbols of b_k bits. For any fixed probability of error P_e and any fixed transmission bit rate per symbol b, the transceiver is ISI free and minimizes the transmission power P in (2) if and only if the following are true.

i) The matrix $\mathbf{A}$ is given by $\mathbf{A} = -\mathbf{U}_0^T\mathbf{R}_\nu\mathbf{U}_1(\mathbf{U}_1^T\mathbf{R}_\nu\mathbf{U}_1)^{-1}$, where $\mathbf{R}_\nu$ is the autocorrelation matrix of the noise vector $\boldsymbol{\nu}$, and $\mathbf{U}_0$ and $\mathbf{U}_1$ are as defined in (1).

ii) The transmitter $\mathbf{G} = \mathbf{Q}$, where $\mathbf{Q}$ is the orthonormal matrix such that $\mathbf{Q}^T\mathbf{R}_q\mathbf{Q}$ is diagonal. The matrix $\mathbf{R}_q$ is given by $\mathbf{R}_q = \mathbf{B}\mathbf{R}_\nu\mathbf{B}^T$ and $\mathbf{B} = \mathbf{V}^T\mathbf{\Lambda}^{-1}[\mathbf{I}\ \mathbf{A}]\mathbf{U}^T$, where $\mathbf{U}$, $\mathbf{V}$, and $\mathbf{\Lambda}$ are given in (1).

iii) The receiver is given by $\mathbf{S} = \mathbf{Q}^T\mathbf{B}$.

iv) The bits b_k are allocated as $b_k = b - \log_2(\sigma_{e_k}\|\mathbf{g}_k\|_2) + (1/M)\log_2(\prod_{\ell=0}^{M-1}\sigma_{e_\ell}\|\mathbf{g}_\ell\|_2)$.

The minimum transmission power is

$$P_{ZP-DMT} = c\,2^{2b}\frac{M}{N}\left[\det(\mathbf{\Lambda}^{-2})\frac{\det(\mathbf{R}_\nu)}{\det\left(\mathbf{U}_1^T\mathbf{R}_\nu\mathbf{U}_1\right)}\right]^{1/M}.$$

The theorem indicates that the optimal ZP-DMT transceiver has an orthogonal transmitter. There is no loss of generality in using orthogonal transmitters.

VI. Optimal ISI-Free DMT Transceivers with General Prefix

In this section, we will consider the case of DMT transceiver with general prefix (GP-DMT). In this case, the transmitter and receiver of the DMT system in Fig. 1 are given by

$$\mathbf{G}_0 = \tilde{\mathbf{G}}, \quad \mathbf{S}_0 = [\mathbf{0} \quad \tilde{\mathbf{S}}]$$

where $\tilde{\mathbf{G}}$, $\mathbf{0}$, and $\tilde{\mathbf{S}}$ are of dimensions $(M+L)$ by M, M by L, and M by M, respectively. Note that the ZP-DMT system is not a special case of the GP-DMT system. The first L columns of the receiver $\mathbf{S}_0$ in GP-DMT systems are constrained to be zeros, whereas those of $\mathbf{S}_0$ in ZP-DMT systems are not constrained. Due to the zero columns, the GP-DMT system is also known as the *leading-zero* DMT [3]. In the special case of cyclic prefix, the first L rows of $\tilde{\mathbf{G}}$ will be the same as its last L rows. Using a technique similar to that given in Section II, Fig. 1 can be redrawn as Fig. 4, where the matrix $\tilde{\mathbf{C}}$ is an M by N upper triangular Toeplitz matrix given by

$$\tilde{\mathbf{C}} = \begin{pmatrix} c_L & c_{L-1} & \cdots & c_0 & 0 & \cdots & 0 \\ 0 & c_L & c_{L-1} & \cdots & c_0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & & \ddots & \vdots \\ 0 & \cdots & 0 & c_L & c_{L-1} & \cdots & c_0 \end{pmatrix}.$$

Comparing the above equation with the expression of $\mathbf{C}$ in (2), one realizes that the two channel matrices are related by $\tilde{\mathbf{C}} = \mathbf{J}_M \mathbf{C}^T \mathbf{J}_N$, where $\mathbf{J}_k$ is the k by k reversal matrix. Therefore, the SVD of $\tilde{\mathbf{C}}$ can be expressed as

$$\tilde{\mathbf{C}} = \tilde{\mathbf{U}}\,[\mathbf{\Lambda} \quad \mathbf{0}]\underbrace{\begin{pmatrix} \tilde{\mathbf{V}}_0 \\ \tilde{\mathbf{V}}_1 \end{pmatrix}}_{\tilde{\mathbf{V}}} \tag{12}$$

where $\tilde{\mathbf{U}} = \mathbf{J}_M \mathbf{V}^T$ and $\tilde{\mathbf{V}} = \mathbf{U}^T \mathbf{J}_N$ are unitary matrices of dimensions M by M and N by N, respectively. Notice that the noise vector $\tilde{\boldsymbol{\nu}}$, which was obtained by blocking the channel noise $\nu(n)$, is of dimensions $M \times 1$. Using an approach similar to that in Section II, we can get the following theorem for ISI-free GP-DMT systems.

Theorem 4—Parameterization of ISI-Free GP-DMT Systems: The block-based DMT system with a general prefix in Fig. 4 is ISI free if and only if the transmitter and receiver satisfy the following.

i) $\tilde{\mathbf{S}}$ is an $M \times M$ nonsingular matrix.
ii) $\tilde{\mathbf{G}} = \tilde{\mathbf{B}}\tilde{\mathbf{S}}^{-1}$, where $\tilde{\mathbf{B}} = \tilde{\mathbf{V}}^T \binom{\mathbf{I}}{\tilde{\mathbf{A}}} \mathbf{\Lambda}^{-1} \tilde{\mathbf{U}}^T$, for an arbitrary $L \times M$ matrix $\tilde{\mathbf{A}}$.

Using a derivation that is similar to that in Section III, one will find that the transmission power can be expressed as

$$\tilde{P} = \frac{1}{N} \sum_{k=0}^{M-1} \sigma_{x_k}^2 \|\tilde{\mathbf{g}}_k\|_2^2 \tag{13}$$

where $\|\tilde{\mathbf{g}}_k\|_2^2 = \sum_{\ell=0}^{M-1} [\tilde{\mathbf{G}}]_{\ell k}^2$ is the energy of the kth column of $\tilde{\mathbf{G}}$. For a given probability of error and average bit rate $b = 1/M \sum_{k=0}^{M-1} b_k$, the bits b_k can be optimally allocated to minimize the above arithmetic mean. Under the optimal bit allocation, the transmission power becomes

$$\tilde{P}_{opt,\,bit} = \frac{cM}{N}\, 2^{2b} \prod_{k=0}^{M-1} \left(\tilde{\sigma}_{e_k}^2 \|\tilde{\mathbf{g}}_k\|_2^2\right)^{1/M} \tag{14}$$

where $\tilde{\sigma}_{e_k}^2$ is the variance of the output noise at the kth band. From the above expression, we see that under the optimal bit allocation, the transmission power $\tilde{P}_{opt,\,bit}$ depends on the transceiver. In the ISI-free solutions of Theorem 4, we are free to choose any $L \times M$ matrix $\tilde{\mathbf{A}}$ and $M \times M$ nonsingular matrix $\tilde{\mathbf{S}}$. In the following, these two matrices are optimized to minimize the transmission power $\tilde{P}_{opt,\,bit}$.

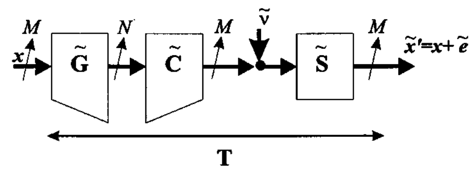


Fig. 4. Matrix representation of DMT systems with a general prefix.

Optimal $\tilde{\mathbf{S}}$*:* The transmission power $\tilde{P}_{opt,\,bit}$ in (14) can be rewritten as

$$\tilde{P}_{opt,\,bit} = c\frac{M}{N}\, 2^{2b} \left(\prod_{k=0}^{M-1} \left[\tilde{\mathbf{S}}^{-T}\tilde{\mathbf{B}}^T\tilde{\mathbf{B}}\tilde{\mathbf{S}}^{-1}\right]_{kk} \left[\tilde{\mathbf{S}}\tilde{\mathbf{R}}_\nu\tilde{\mathbf{S}}^T\right]_{kk} \right)^{1/M}$$

where $\tilde{\mathbf{R}}_\nu$ is the $M \times M$ autocorrelation matrix of the channel noise $\nu(n)$ in Fig. 1. As both $\tilde{\mathbf{R}}_\nu$ and $\tilde{\mathbf{B}}^T\tilde{\mathbf{B}}$ are positive semidefinite, we can apply the Hadamard inequality to obtain

$$\begin{aligned} \tilde{P}_{opt,bit} &\geq c2^{2b}\frac{M}{N} \left[\det\left[\tilde{\mathbf{S}}^{-T}\tilde{\mathbf{B}}^T\tilde{\mathbf{B}}\tilde{\mathbf{S}}^{-1}\right] \det[\tilde{\mathbf{S}}\tilde{\mathbf{R}}_\nu\tilde{\mathbf{S}}^T]\right]^{1/M} \\ &= c2^{2b}\frac{M}{N} \left[\det\left[\tilde{\mathbf{B}}^T\tilde{\mathbf{B}}\right] \det\left[\tilde{\mathbf{R}}_\nu\right]\right]^{1/M} \\ &\triangleq \tilde{P}_{opt,\,bit,\,\tilde{\mathbf{S}}}. \end{aligned} \tag{15}$$

Note that the lower bound $\tilde{P}_{opt,\,bit,\,\tilde{\mathbf{S}}}$ is independent of the receiver $\tilde{\mathbf{S}}$, and it is achieved if and only if the receiver $\tilde{\mathbf{S}}$ simultaneously satisfies the following two conditions.

i) $\tilde{\mathbf{S}}\tilde{\mathbf{R}}_\nu\tilde{\mathbf{S}}^T$ is diagonal.
ii) $\tilde{\mathbf{S}}^{-T}\tilde{\mathbf{B}}^T\tilde{\mathbf{B}}\tilde{\mathbf{S}}^{-1}$ is diagonal; that is, $\tilde{\mathbf{G}}^T\tilde{\mathbf{G}}$ is diagonal.

To derive such an optimal $\tilde{\mathbf{S}}$, we first decompose $\tilde{\mathbf{R}}_\nu$ as

$$\tilde{\mathbf{R}}_\nu = \mathbf{T}\mathbf{D}\mathbf{T}^T = \left(\mathbf{T}\mathbf{D}^{1/2}\mathbf{T}^T\right)\left(\mathbf{T}\mathbf{D}^{1/2}\mathbf{T}^T\right) = \tilde{\mathbf{R}}_\nu^{1/2}\tilde{\mathbf{R}}_\nu^{1/2}$$

for some unitary matrix $\mathbf{T}$ and some diagonal matrix $\mathbf{D}$ with non-negative entries. To avoid degenerate cases, we assume $\tilde{\mathbf{R}}_\nu$ is positive definite so that $\tilde{\mathbf{R}}_\nu^{1/2}$ is invertible. Using the above decomposition, it is not difficult to see that Condition i) is satisfied if $\tilde{\mathbf{S}}$ has the following form:

$$\tilde{\mathbf{S}} = \tilde{\mathbf{Q}}\tilde{\mathbf{R}}_\nu^{-1/2}$$

where $\tilde{\mathbf{Q}}$ is any M by M unitary matrix. Substituting the above expression into Condition ii), one can show that Condition ii) is satisfied if the unitary matrix $\tilde{\mathbf{Q}}$ diagonalizes the positive semidefinite matrix

$$\tilde{\mathbf{R}}_\nu^{1/2}\tilde{\mathbf{B}}^T\tilde{\mathbf{B}}\tilde{\mathbf{R}}_\nu^{1/2}.$$

Therefore, we conclude that the lower bound $\tilde{P}_{opt,\,bit,\,\tilde{\mathbf{S}}}$ given in (15) is achievable. Note that once the channel and noise are given, this bound depends only on $\tilde{\mathbf{B}}$, whose only free parameter is the matrix $\tilde{\mathbf{A}}$. In what follows, we will optimize $\tilde{\mathbf{A}}$ so that this lower bound is minimized.

Optimal $\tilde{\mathbf{A}}$*:* From the expression of $\tilde{P}_{opt,\,bit,\,\tilde{\mathbf{S}}}$ in (15), we see that the matrix $\tilde{\mathbf{A}}$ should be chosen such that $\det[\tilde{\mathbf{B}}^T\tilde{\mathbf{B}}]$ is minimized. Using the fact that $\tilde{\mathbf{B}} = \tilde{\mathbf{V}}^T\binom{\mathbf{I}}{\tilde{\mathbf{A}}}\mathbf{\Lambda}^{-1}\tilde{\mathbf{U}}^T$, where $\tilde{\mathbf{U}}$ and $\tilde{\mathbf{V}}$ are square unitary matrices defined in (12), we have

$$\det\left[\tilde{\mathbf{B}}^T\tilde{\mathbf{B}}\right] = \det[\mathbf{\Lambda}^{-2}]\det\left[\mathbf{I}+\tilde{\mathbf{A}}^T\tilde{\mathbf{A}}\right].$$

As $\tilde{\mathbf{A}}^T\tilde{\mathbf{A}}$ is positive semidefinite, it can be verified that

$$\det\left[\tilde{\mathbf{B}}^T\tilde{\mathbf{B}}\right] \geq \det[\mathbf{\Lambda}^{-2}].$$

Equality holds if and only if $\tilde{\mathbf{A}} = \mathbf{0}$. In this case, the optimal transmitter is given by

$$\tilde{\mathbf{G}} = \tilde{\mathbf{V}}_0^T\mathbf{\Lambda}^{-1}\tilde{\mathbf{U}}^T\tilde{\mathbf{S}}^{-1}$$

where $\tilde{\mathbf{S}}$ is the optimal receiver derived in the previous subsection. Summarizing all the results in this section, we have the following theorem.

Theorem 5: Consider the M-band GP-DMT system in Fig. 4. Assume that the inputs are PAM symbols of b_k bits. For any fixed probability of error P_e and any fixed transmission bit rate per symbol b, the transceiver is ISI free and minimizes the transmission power $\tilde{P}$ in (13) if and only if the following are true.

i) The matrix $\tilde{\mathbf{B}} = \tilde{\mathbf{V}}_0^T\mathbf{\Lambda}^{-1}\tilde{\mathbf{U}}^T$, where $\tilde{\mathbf{U}}$ and $\tilde{\mathbf{V}}_0$ are defined in (12).

ii) The receiver is given by $\tilde{\mathbf{S}} = \tilde{\mathbf{Q}}\tilde{\mathbf{R}}_\nu^{-1/2}$, where $\tilde{\mathbf{R}}_\nu$ is the $M \times M$ autocorrelation matrix of the noise vector $\boldsymbol{\nu}$, and $\tilde{\mathbf{Q}}$ is the unitary matrix that diagonalizes $(\tilde{\mathbf{R}}_\nu^{1/2}\tilde{\mathbf{B}}^T\tilde{\mathbf{B}}\tilde{\mathbf{R}}_\nu^{1/2})$.

iii) The transmitter $\tilde{\mathbf{G}} = \tilde{\mathbf{B}}\tilde{\mathbf{S}}^{-1}$.

iv) The bits b_k are allocated as $b_k = b - \log_2(\tilde{\sigma}_{e_k}\|\tilde{\mathbf{g}}_k\|_2) + (1/M)\log_2(\prod_{\ell=0}^{M-1}\tilde{\sigma}_{e_\ell}\|\tilde{\mathbf{g}}_\ell\|_2)$.

The minimum transmission power is

$$P_{GP-DMT} = c2^{2b}\frac{M}{N}\left[\det(\mathbf{\Lambda}^{-2})\det\left(\tilde{\mathbf{R}}_\nu\right)\right]^{1/M}.$$

Remarks:

1) For the inequality to hold in (15), Conditions i) and ii) should hold simultaneously. Condition i) means that the receiver $\tilde{\mathbf{S}}$ decorrelates the noise vector. Condition ii) implies that the transmitter $\tilde{\mathbf{G}}$ is orthogonal. Therefore, in GP-DMT systems, there is no loss of generality in using orthogonal transmitters.

2) In GP-DMT systems, nonzero prefix samples are padded at the beginning of every M samples. There is inter-block-interference (IBI) in the received signal due to the channel. To remove IBI, the receiver retains only M samples of each block. Although the transmitter sends out N samples for each input M symbols, the receiver use only M samples for decoding. On the other hand, in ZP-DMT systems, zeros are padded at the end of every M samples. After passing through the channel, samples are spread to nonoverlapping blocks of length N. As there is no IBI, all the N samples can be used for decoding. There are more observations than unknowns; the dimension of the signal subspace is M, whereas the received signal has dimension $N = M + L$. The eigenstructure of the signal subspace can be exploited to our advantage in ZP-DMT systems. Therefore, the performance of the ZP-DMT systems is generally better than that of the GP-DMT system, as we will see later. However, as ZP-DMT and GP-DMT systems are two different classes of DMT systems, it is possible to construct toy examples (see Section VIII) such that the GP-DMT system is better.

Special Case—Cyclic Prefix: In the commonly used case of cyclic prefix, the transmitter has the form

$$\tilde{\mathbf{G}} = \begin{pmatrix}\mathbf{0} & \mathbf{I}_L\\ \mathbf{I}_M & \end{pmatrix}\tilde{\mathbf{G}}_M \tag{16}$$

where $\tilde{\mathbf{G}}_M$ is some M by M nonsingular matrix. Using (12), the ISI-free condition becomes

$$\tilde{\mathbf{S}}\underbrace{\tilde{\mathbf{C}}\begin{pmatrix}\mathbf{0} & \mathbf{I}_L\\ \mathbf{I}_M & \end{pmatrix}}_{\tilde{\mathbf{C}}_{cyc}}\tilde{\mathbf{G}}_M = \mathbf{I}.$$

Observe that the matrix $\tilde{\mathbf{C}}_{cyc}$ is circulant, with the first column given by $(c_0\ c_1\ \cdots\ c_L\ 0\ \cdots\ 0)^T$. It is known that it can be diagonalized using a DFT matrix, and the eigenvalues are the M-point DFT of the channel impulse response. It is not guaranteed to be nonsingular; it is singular if the channel has zeros at $e^{j2k\pi/M}$ for some integer k. Let us assume that $\tilde{\mathbf{C}}_{cyc}$ is nonsingular, and let

$$\tilde{\mathbf{B}}_{cyc} = \begin{pmatrix}\mathbf{0} & \mathbf{I}_L\\ \mathbf{I}_M & \end{pmatrix}\tilde{\mathbf{C}}_{cyc}^{-1}.$$

For any nonsingular $\tilde{\mathbf{S}}$, the transmitter that achieves the ISI-free condition is given by $\tilde{\mathbf{G}} = \tilde{\mathbf{B}}_{cyc}\tilde{\mathbf{S}}^{-1}$. Note that in the cyclic prefix case, once the channel $\tilde{\mathbf{C}}$ is given, the matrix $\tilde{\mathbf{B}}_{cyc}$ is fixed. The only free parameter in an ISI-free DMT system with cyclic prefix is the nonsingular matrix $\tilde{\mathbf{S}}$. By repeating the earlier optimization process, we can obtain the optimal ISI-free DMT system with cyclic prefix as follows:

Theorem 6: Consider Fig. 4. Suppose that the ISI-free DMT system with cyclic prefix exists. In this case, its transmitter of the cyclic prefix DMT system has the form of (16). Assume that the inputs are PAM symbols of b_k bits. For any fixed probability of error P_e and any fixed transmission bit rate per symbol b, the transceiver is ISI free and minimizes the transmission power $\tilde{P}$ in (13) if and only if the following are true.

i) The receiver is given by $\tilde{\mathbf{S}} = \tilde{\mathbf{Q}}\tilde{\mathbf{R}}_\nu^{-1/2}$, where $\tilde{\mathbf{R}}_\nu$ is the autocorrelation matrix of the noise vector $\tilde{\boldsymbol{\nu}}$, and $\tilde{\mathbf{Q}}$ is the unitary matrix that diagonalizes $(\tilde{\mathbf{R}}_\nu^{1/2}\tilde{\mathbf{B}}_{cyc}^T\tilde{\mathbf{B}}_{cyc}\tilde{\mathbf{R}}_\nu^{1/2})$.

ii) The transmitter is $\tilde{\mathbf{G}} = \tilde{\mathbf{B}}_{cyc}\tilde{\mathbf{S}}^{-1}$.

iii) The bits b_k are allocated as $b_k = b - \log_2(\tilde{\sigma}_{e_k}\|\tilde{\mathbf{g}}_k\|_2) + (1/M)\log_2(\prod_{\ell=0}^{M-1}\tilde{\sigma}_{e_\ell}\|\tilde{\mathbf{g}}_\ell\|_2)$.

The minimum transmission power is

$$P_{CYC\text{-}DMT} = c2^{2b}\frac{M}{N}\left[\det\left[\tilde{\mathbf{B}}_{cyc}^T\tilde{\mathbf{B}}_{cyc}\right]\det\left[\tilde{\mathbf{R}}_\nu\right]\right]^{1/M}.$$

Remarks:

1) As cyclic prefix is a special case of a general one, we can immediately conclude that the performance $P_{GP\text{-}DMT} \leq P_{CYC\text{-}DMT}$. The gain can be substantial, as will be demonstrated by the example in Section VIII.
2) The conventional DFT-based DMT systems belong to the class of GP-DMT with cyclic prefix. The receiver $\tilde{\mathbf{S}}$ is $\tilde{\mathbf{S}} = \mathbf{W}\tilde{\mathbf{C}}_{cyc}^{-1}$, where $\mathbf{W}$ is the $M \times M$ DFT matrix. Because $\tilde{\mathbf{C}}_{cyc}^{-1}$ is also circulant, one can verify that $\tilde{\mathbf{S}} = \mathbf{\Sigma}^{-1}\mathbf{W}$, where $\mathbf{\Sigma}$ is a diagonal matrix. The diagonal entries of $\mathbf{\Sigma}$ are the M-point DFT of the channel impulse response. As $P_{CYC\text{-}DMT}$ achieves the minimum transmission power among the class of DMT systems with a cyclic prefix, we have $P_{GP\text{-}DMT} \leq P_{CYC\text{-}DMT} \leq P_{DFT}$. In fact, $P_{CYC\text{-}DMT}$ can be much smaller than P_{DFT}, as we will see later.

VII. Comparison of Different DMT Systems with Zero Padding

The design of optimal transceiver for a given channel with additive colored noise is considered in [3]. The measure of optimality and the problem addressed therein are different from here. For a given input autocorrelation matrix and transmission power, the authors derive the optimal transceiver that minimizes the total output noise power, but bit allocation is not considered, and the solution is not ISI free. In [4], information rate as a measure of optimality is considered. The DMT transceiver that maximizes the mutual information between the transmitted block and the received blocked is developed. However, the resulting transceiver may not be optimal for all encoding schemes. This is demonstrated in the example shown in [4]; the DFT-based DMT system with QAM symbols compares favorably with the information rate maximized DMT system in some cases.

In this section, we will show that the optimal ZP-DMT has a very nice connection to the systems proposed in [5] and [2]. All three systems belong to the class of ISI-free DMT systems with zero padding. We will show that the optimal ZP-DMT transceiver always outperforms these two systems, except in degenerated cases when they have the same performance. We will give the conditions when the optimal ZP-DMT transceiver reduces to the system in [5] and the vector coding system in [2].

The DMT systems with orthogonal transmitters in [5] and the vector coding systems in [2] are both ISI free. According to Theorem 1, they can be described in terms of the matrices $\mathbf{A}$ and $\mathbf{G}$. The DMT transceiver proposed in [5] has an orthogonal transmitter with

$$\mathbf{A} = \mathbf{0}, \quad \text{and} \quad \mathbf{G} = \mathbf{Q}$$

where $\mathbf{Q}$ is the orthonormal matrix such that $\mathbf{Q}^T\mathbf{R}_q\mathbf{Q}$ is diagonal. Under optimal bit allocation, the transmission power $P_{[5]}$ in this case is [5],

$$P_{[5]} = c2^{2b}\frac{M}{N}\left[\det\left(\mathbf{\Lambda}^{-2}\right)\det(\mathbf{R}_{\mu_0})\right]^{1/M}.$$

The vector coding transceiver proposed in [2] has

$$\mathbf{A} = \mathbf{0}, \quad \text{and} \quad \mathbf{G} = \mathbf{V}^T.$$

Under optimal bit allocation, one can verify that the transmission power $P_{VC,[2]}$ of the vector coding transceiver is [2]

$$P_{VC,[2]} = c2^{2b}\frac{M}{N}\left[\det(\mathbf{\Lambda}^{-2})\prod_{k=0}^{M-1}[\mathbf{R}_{\mu_0}]_{kk}\right]^{1/M}.$$

Let us compute the ratio $P_{ZP\text{-}DMT}/P_{VC,[2]}$

$$\frac{P_{ZP\text{-}DMT}}{P_{VC,[2]}} = \left(\frac{\det(\mathbf{R}_\nu)}{\det\left(\mathbf{R}_{\mu_1}\right)\prod_{k=0}^{M-1}[\mathbf{R}_{\mu_0}]_{kk}}\right)^{1/M}$$
$$= \underbrace{\left(\frac{\det(\mathbf{R}_\nu)}{\det\left(\mathbf{R}_{\mu_0}\right)\det\left(\mathbf{R}_{\mu_1}\right)}\right)^{1/M}}_{P_{ZP\text{-}DMT}/P_{[5]}} \cdot \underbrace{\left(\frac{\det\left(\mathbf{R}_{\mu_0}\right)}{\prod_{k=0}^{M-1}[\mathbf{R}_{\mu_0}]_{kk}}\right)^{1/M}}_{P_{[5]}/P_{VC,[2]}}.$$

The above equation gives the precise connection between $P_{ZP\text{-}DMT}$ and $P_{[5]}$, as well as between $P_{[5]}$ and $P_{VC,[2]}$.

i) Comparison of $P_{ZP\text{-}DMT}$ and $P_{[5]}$: The ratio $P_{ZP\text{-}DMT}/P_{[5]} \leq 1$ with equality if and only if $\boldsymbol{\mu}_0$ and $\boldsymbol{\mu}_1$ are uncorrelated, $\mathcal{E}[\boldsymbol{\mu}_0\boldsymbol{\mu}_1^T] = \mathbf{0}$. To see this, let us consider the autocorrelation matrix of the vector $\boldsymbol{\mu}$ defined in (9)

$$\mathbf{R}_\mu = \begin{pmatrix} \mathbf{R}_{\mu_0} & \mathcal{E}[\boldsymbol{\mu}_0\boldsymbol{\mu}_1^T] \\ \mathcal{E}\left[\boldsymbol{\mu}_0\boldsymbol{\mu}_1^T\right]^T & \mathbf{R}_{\mu_1} \end{pmatrix}.$$

Using Fischer's inequality (see the Appendix), one can see that $\det(\mathbf{R}_\nu) = \det(\mathbf{R}_\mu) \leq \det(\mathbf{R}_{\mu_0})\det(\mathbf{R}_{\mu_1})$. Equality holds if and only if $\mathcal{E}[\boldsymbol{\mu}_0\boldsymbol{\mu}_1^T] = \mathbf{0}$. The improvement of $P_{ZP\text{-}DMT}$ over $P_{[5]}$ comes from the exploitation of the signal subspace or the flexibility in the design of $\mathbf{A}$ for the left inverse of $\mathbf{C}$.

ii) Comparison of $P_{[5]}$ and $P_{VC,[2]}$: Using the Hadamard inequality, the ratio $P_{[5]}/P_{VC,[2]} \leq 1$ with equality if and only if $\mathbf{R}_{\mu_0}$ is a diagonal matrix, i.e., elements of $\boldsymbol{\mu}_0$ are uncorrelated. The gain is from the decorrelation of the noise vector $\boldsymbol{\mu}_0$. There is no improvement if elements of $\boldsymbol{\mu}_0$ are not correlated.

iii) Comparison of $P_{ZP\text{-}DMT}$ and $P_{VC,[2]}$: The ratio $P_{ZP\text{-}DMT}/P_{VC,[2]} \leq 1$ with equality if and only if $\boldsymbol{\mu}_0$ and $\boldsymbol{\mu}_1$ are uncorrelated, and $\mathbf{R}_{\mu_0}$ is a diagonal matrix.

When the channel noise $\nu(n)$ is additive white Gaussian noise (AWGN), the elements of the noise vector $\boldsymbol{\nu}$ are uncorrelated. The autocorrelation matrix $\mathbf{R}_\nu = \sigma_\nu^2\mathbf{I}_M$ is the identity matrix, except for a scalar. The vector $\boldsymbol{\mu}_0$ and $\boldsymbol{\mu}_1$ are obtained from $\boldsymbol{\nu}$ through the orthonormal transformation $\mathbf{U}^T$. Therefore, $\boldsymbol{\mu}_0$

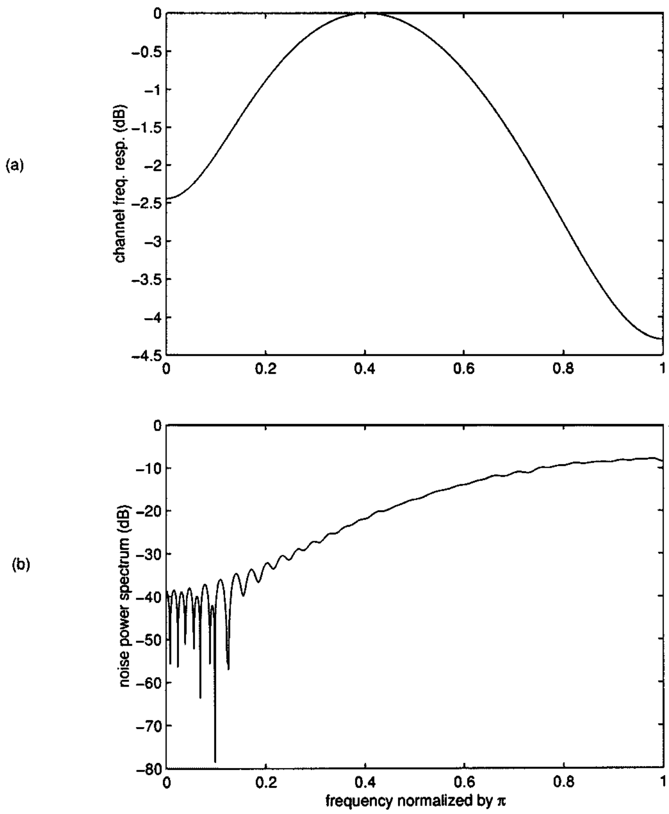


Fig. 5. (a) Frequency response of the channel $|C(e^{j\omega})|$. (b) Power spectrum of the channel noise $\nu(n)$.

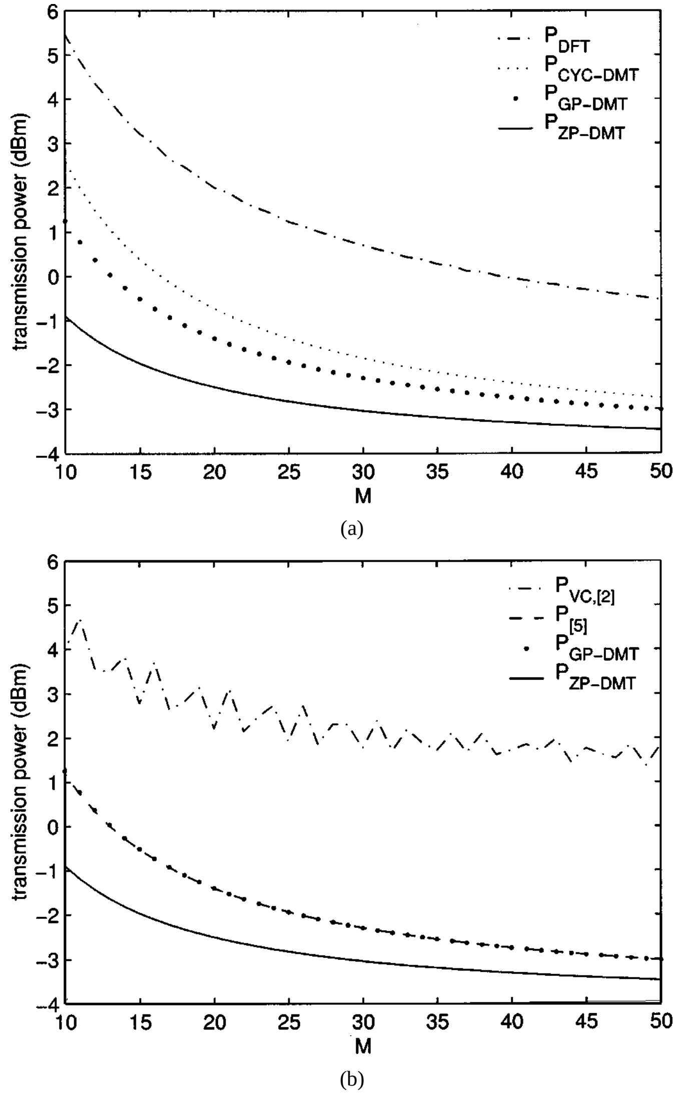


Fig. 6. (a) Performance comparison of DMT systems with a prefix. (b) Performance comparison of DMT systems with zero padding.

and $\boldsymbol{\mu}_1$ are uncorrelated, and moreover, $\mathbf{R}_{\mu_0}$ and $\mathbf{R}_{\mu_1}$ are diagonal. In this case, $P_{ZP\text{-}DMT} = P_{[5]} = P_{VC,[2]}$, and the optimal transceiver reduces to the vector coding solution given in [2]. Notice that in applications such as DSL, the channel model of LTI filter $C(z)$ plus additive noise $\nu(n)$ in Fig. 1 is usually obtained after time domain equalization; $C(z)$ is the equalized channel, and $\nu(n)$ is the noise after equalization. Even if the original channel noise is white, the equalized noise $\nu(n)$ can be colored.

VIII. Examples

For the first two examples, the channel $C(z)$ and power spectrum for the colored noise $\nu(n)$ used are shown in Fig. 5(a) and (b). These parameters are obtained from a typical ADSL environment. The channel $C(z)$ in this case is the equivalent channel after time domain equalization, and $C(z)$ has order $L = 4$. The bit error rate $P_e = 10^{-6}$ and average bit rate per sample is $R_b = (M/N)b = 2$.

Example 1—DMT Systems with a Nonzero Prefix: Fig. 6(a) shows the results of the three prefix DMT systems considered in Section VI:

1) optimal prefix system $P_{GP\text{-}DMT}$;
2) system with cyclic prefix $P_{CYC\text{-}DMT}$;
3) DFT-based DMT system with cyclic prefix P_{DFT}.

The results are shown for $M = 10$ to 50. As a comparison, the result of $P_{ZP\text{-}DMT}$ is also shown in the plot. The optimal GP-DMT systems has approximately 2.5 to 5 dB gain over the conventional DFT-based systems; the gain of $P_{GP\text{-}DMT}$ over $P_{CYC\text{-}DMT}$ is approximately 1 dB for $M = 10$. In both cases, the gain is larger for small M. In addition, we can see that the zero-padding ZP-DMT system performs better than the general-prefix system.

Example 2—DMT Systems with Zero Padding: Fig. 6(b) compares the three zero-padding DMT systems:

1) $P_{ZP\text{-}DMT}$;
2) $P_{[5]}$;
3) $P_{VC,[2]}$.

$P_{GP\text{-}DMT}$ is also shown in the plot for comparison. Note that $P_{ZP\text{-}DMT}$ is approximately 5 dB smaller than $P_{VC,[2]}$, and $P_{[5]}$ is 4–5 dB smaller than $P_{VC,[2]}$. The transmission power $P_{ZP\text{-}DMT}$ is 0.5–2 dB less than $P_{[5]}$, depending on M. The gain of $P_{[5]}$ over $P_{VC,[2]}$ comes from the decorrelation of the noise vector. The gain of $P_{ZP\text{-}DMT}$ over $P_{[5]}$ comes from the exploitation of the eigenstructure of the signal subspace. The improvement is more significant for small number of bands M. When M is large, the gap becomes smaller. This is because, for large M, the dimension M of signal subspace is almost as large as the dimension N of the received signal.

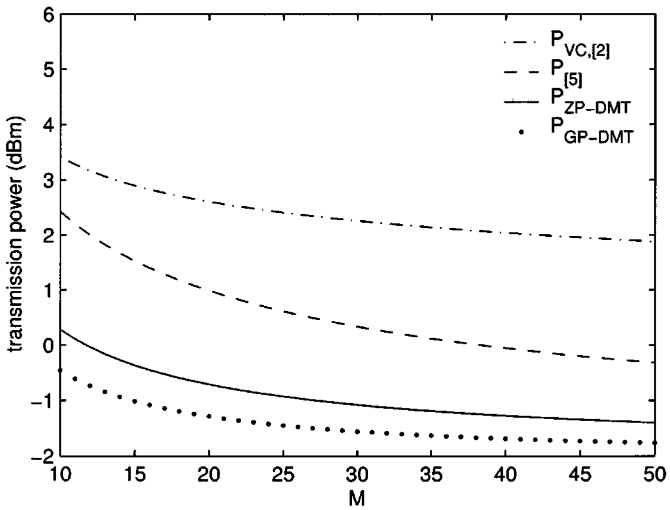


Fig. 7. Toy example. Performance comparison of $P_{GP\text{-}DMT}$, $P_{ZP\text{-}DMT}$, $P_{[5]}$, and $P_{VC,[2]}$.

In most cases, the ZP-DMT system outperforms the *GP-DMT* system as in the above example, and the curve $P_{GP\text{-}DMT}$ is very close to the curve $P_{[5]}$. However, the ZP-DMT system is not always better than the GP-DMT system; the two curves $P_{GP\text{-}DMT}$ and $P_{[5]}$ are not necessarily close. Consider the following toy example.

Example 3—Toy Example: Consider the channel $C(z) = 1 + 2z^{-1} + z^{-2}$ and the channel noise $\nu(n)$ with power spectrum $S_\nu(e^{j\omega}) = |C(e^{j\omega})|^2$. The results of $P_{GP\text{-}DMT}$ and the three zero-padding systems $P_{ZP\text{-}DMT}$, $P_{[5]}$, and $P_{VC,[2]}$ are shown in Fig. 7. In this toy example, $P_{GP\text{-}DMT}$ is better than $P_{ZP\text{-}DMT}$ and significantly better than $P_{[5]}$ and $P_{VC,[2]}$. The three zero-padding system have $P_{ZP\text{-}DMT} < P_{[5]} < P_{VC,[2]}$, as expected from the results shown in Section VII.

IX. Concluding Remarks

In this paper, we consider two classes of block-based DMT transceivers: the DMT system with a general prefix (GP-DMT) and the system with zero padding (ZP-DMT). We have shown that for these two classes, the optimal transceivers that minimize the transmission power for a given bit rate and probability of error can be given in closed form. Furthermore, we demonstrate that for both classes, the optimal transceiver has an orthogonal transmitter; there is no loss of generality in using orthogonal transmitters.

Appendix
Fischer's Inequality for Matrices and its Extension

Lemma 1 (Fischer's Inequality [13]): Suppose that

$$\mathbf{P} = \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B}^T & \mathbf{C} \end{pmatrix}$$

is a positive definite matrix that is partitioned so that $\mathbf{A}$ and $\mathbf{C}$ are square and nonempty. Then

$$\det(\mathbf{P}) \le \det(\mathbf{A})\det(\mathbf{C}). \tag{17}$$

Proof: The proof can be found in [13]. We will repeat it here as the proof will be used in later part of the paper. Letting $\mathbf{X} = -\mathbf{A}^{-1}\mathbf{B}$, we have

$$\begin{aligned}\det\mathbf{P} &= \det\begin{pmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{X}^T & \mathbf{I} \end{pmatrix}\begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B}^T & \mathbf{C} \end{pmatrix}\begin{pmatrix} \mathbf{I} & \mathbf{X} \\ \mathbf{0} & \mathbf{I} \end{pmatrix} \\ &= \det\begin{pmatrix} \mathbf{A} & \mathbf{0} \\ \mathbf{0} & \mathbf{C} - \mathbf{B}^T\mathbf{A}^{-1}\mathbf{B} \end{pmatrix} \\ &= \det(\mathbf{A})\det\left(\mathbf{C} - \mathbf{B}^T\mathbf{A}^{-1}\mathbf{B}\right).\end{aligned}$$

Note that the matrix $\mathbf{C} - \mathbf{B}^T\mathbf{A}^{-1}\mathbf{B}$ is positive definite as $\mathbf{P}$ is positive definite. Using this and the facts that both $\mathbf{A}$ and $\mathbf{C}$ are positive definite, it is not difficult to show that

$$\det\left(\mathbf{C} - \mathbf{B}^T\mathbf{A}^{-1}\mathbf{B}\right) \le \det(\mathbf{C}).$$

Lemma 2: Let $\mathbf{P}$ be the positive definite matrix defined in Lemma 1. Then, $\det(\mathbf{P}) = \det(\mathbf{A})\det(\mathbf{C})$ if and only if $\mathbf{B} = \mathbf{0}$.

Proof: From the proof of Lemma 1, we know that proving Lemma 2 is equivalent to proving the following statement:

$$\det\left(\mathbf{C} - \mathbf{B}^T\mathbf{A}^{-1}\mathbf{B}\right) = \det(\mathbf{C}) \tag{18}$$

if and only if $\mathbf{B} = \mathbf{0}$. The *if* part is clear. To show the *only if* part, we first decompose $\mathbf{C} = \mathbf{E}^{-T}\mathbf{E}^{-1}$ and $\mathbf{A}^{-1} = \mathbf{F}^T\mathbf{F}$ for some positive definite matrices $\mathbf{E}$ and $\mathbf{F}$. This can always be done as both $\mathbf{C}$ and $\mathbf{A}^{-1}$ are positive definite. Using these decompositions, we can rewrite (18) as

$$\det\left(\mathbf{I} - (\mathbf{FBE})^T(\mathbf{FBE})\right) = \det(\mathbf{I}) = 1. \tag{19}$$

Note that the matrix $(\mathbf{FBE})^T(\mathbf{FBE})$ can be diagonalized by some unitary matrix. Consider its diagonalized form, and let the corresponding diagonal matrix that consists of all the eigenvalues be $\mathbf{\Lambda} = \mathrm{diag}(\lambda_0, \lambda_1, \ldots, \lambda_{n-1})$. Then, we can rewrite (19) as

$$\det(\mathbf{I} - \mathbf{\Lambda}) = \prod_{i=0}^{n-1}(1 - \lambda_i) = 1. \tag{20}$$

As the matrix $\mathbf{I} - (\mathbf{FBE})^T(\mathbf{FBE})$ is positive definite and $(\mathbf{FBE})^T(\mathbf{FBE})$ is positive semi definite, the eigenvalues λ_i must satisfy $0 \le \lambda_i < 1$. Using this, we conclude that (20) holds if and only if $\lambda_i = 0$ for all i. In other words, (19) holds if and only if $(\mathbf{FBE}) = \mathbf{0}$. Since $\mathbf{F}$ and $\mathbf{E}$ are positive definite matrices, we have $(\mathbf{FBE}) = \mathbf{0}$ if and only if $\mathbf{B} = \mathbf{0}$.

References

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

[2] S. Kasturia, J. T. Aslanis, and J. M. Cioffi, "Vector coding for partial response channels," *IEEE Trans. Inform. Theory*, pp. 741–762, July 1990.

[3] A. Scaglione, G. B. Giannakis, and S. Barbarossa, "Redundant filterbank precoders and equalizers Part I: Unification and optimal designs," *IEEE Trans. Signal Processing*, vol. 47, pp. 1988–2006, July 1999.

[4] 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. 1019–1032, Apr. 1999.

[5] Y.-P. Lin and S.-M. Phoong, “Perfect discrete multitone modulation with optimal transceivers,” *IEEE Trans. Signal Processing*, vol. 48, pp. 1702–1711, June 2000.

[6] M. Vetterli, “Perfect transmultiplexers,” in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, Tokyo, Japan, Apr. 1986, pp. 2567–2570.

[7] A. N. Akansu *et al.*, “Orthogonal transmultiplexers in communication: A review,” *IEEE Trans. Signal Processing*, vol. 46, pp. 979–995, Apr. 1998.

[8] I. Kalet, “Multitone modulation,” in *Subband and Wavelet Transforms: Design and Applications*, A. N. Akansu and M. J. T. Smith, Eds. Boston, MA: Kluwer, 1995.

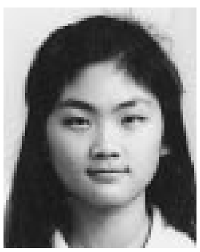
[9] Y.-P. Lin, P. P. Vaidyanathan, S. Akkarakaran, and S.-M. Phoong, “On the duality of optimal DMT systems and biorthogonal subband coders,” in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, June 2000.

[10] P. P. Vaidyanathan, “Theory of optimal orthonormal coder,” *IEEE Trans. Signal Processing*, pp. 1528–1543, June 1998.

[11] Y.-P. Lin and S.-M. Phoong, “Minimum redundancy for ISI-free FIR DMT transceivers,” IEEE Trans. Signal Processing, submitted for publication.

[12] J. G. Proakis, *Digital Communications*. New York: McGraw-Hill, 1995.

[13] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge, U.K.: Cambridge Univ. Press, 1985.



Yuan-Pei Lin (S’93–M’97) was born in Taipei, Taiwan, R.O.C., in 1970. She received the B.S. degree in control engineering from the National Chiao-Tung University (NCTU), Hsinchu, Taiwan, in 1992 and the M.S. and the Ph.D. degrees in electrical engineering from the California Institute of Technology, Pasadena, in 1993 and 1997, respectively.

She joined the Department of Electrical and Control Engineering of NCTU in 1997. Her research interests include multirate filterbanks, wavelets, and applications to communication systems. She is currently an Associate Editor for *Multidimensional Systems and Signal Processing* with Academic Press.

See-May Phoong (M’96) was born in Johor, Malaysia, in 1968. He received the B.S. degree in electrical engineering from the National Taiwan University (NTU), Taipei, Taiwan, R.O.C., in 1991 and the M.S. and Ph.D. degrees in electrical engineering from the California Institute of Technology (Caltech), Pasadena, in 1992 and 1996, respectively.

He joined the faculty of the Department of Electronic and Electrical Engineering, Nanyang Technological University, Singapore, from September 1996 to September 1997. Since September 1997, he has been an Assistant Professor with the Institute of Communication Engineering and Electrical Engineering, NTU. His interests include signal compression, transform coding, and filterbanks and their applications to communication.

Dr. Phoong received the 1997 Wilts Prize at Caltech for outstanding independent research in electrical engineering.