

Perfect Discrete Multitone Modulation with Optimal Transceivers

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

Abstract—Recently, discrete Fourier transform (DFT)-based discrete multitone modulation (DMT) systems have been widely applied to various applications. In this paper, we study a broader class of DMT systems using more general unitary matrices instead of DFT matrices. For this class, we will show how to design the optimal DMT systems over frequency-selective channels with colored noise. In addition, asymptotical performance of DFT-based and optimal DMT systems will be studied and shown to be equivalent. However, for a moderate number of bands, the optimal DMT system offers significant gain over the DFT-based DMT system, as will be demonstrated by examples.

Index Terms—DMT, optimal DMT, perfect transceiver, zero ISI.

I. INTRODUCTION

RECENTLY, there has been considerable interest in applying the discrete multitone modulation (DMT) technique to high-speed data transmission over frequency selective channels such as asymmetrical digital subscriber loops (ADSL's) and high-speed digital subscriber loops (HDSL's) [1]–[4]. Fig. 1 shows an M -band DMT system over a frequency-selective channel $C(z)$ with additive noise $e(n)$. The channel is divided into M bands using the transmitting filters $F_k(z)$ and receiving filters $H_k(z)$. The input bit stream is parsed and coded as modulation symbols, e.g., PAM or QAM. In [5] and [6], Kalet shows that the DMT system with ideal filters can achieve a signal-to noise-ratio (SNR) within 8–9 dB of the channel capacity.

In practice, to cancel intersymbol interference (ISI), usually, some degree of redundancy is introduced, and the interpolation ratio $N > M$ [1], [2]. The length of the transmitting and receiving filters is usually also N . In the widely used discrete Fourier transform (DFT) based DMT system, the transmitting and receiving filters are DFT filters. Redundancy takes the form of cyclic prefix. For a given probability of error and transmission power, bits can be allocated among the bands to achieve maximum total bit rate $R_{b, \max}$. Very high speed data transmission can be achieved using a DFT-based DMT system at a relatively low cost [1]. This technique is currently playing an important role in high speed modems for ADSL and HDSL [3]. Canceling

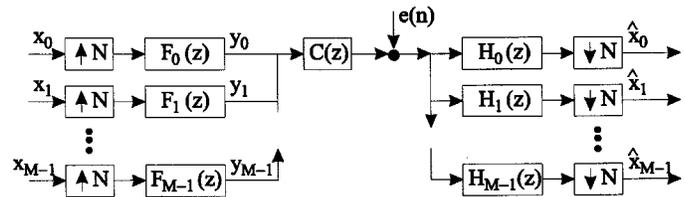


Fig. 1. System model for an M -band DMT system over a frequency-selective channel.

channel ISI by introducing redundancy using a multirate precoding technique has been studied by Xia in [7].

In the DMT system, the maximum bit rate $R_{b, \max}$ depends on the choice of the transmitting and receiving filters. The use of more general orthogonal transmitting filters instead of DFT filters is proposed in [8]. From the viewpoint of multidimensional signal constellations, it is shown that for additive white Gaussian noise (AWGN) frequency-selective channels, the optimal transmitting and receiving filters are eigenvectors associated with the channel. However, in HDSL applications, the dominating noise is often colored noise known as near end cross talk (NEXT) [1].

In this paper, we will use the polyphase approach that has enjoyed great success in filter bank theory to study the DMT system [2], [9]. We will derive a modified DFT-based DMT system that has a better noise rejection property but the same cost as the traditional DFT-based system. Moreover, optimal transceiver for colored noise will be studied in detail. In particular, we will show how to assign bits among the bands so that the total transmitting power can be minimized for a given bit rate. Based on the optimal bit allocation, the optimal transceiver is derived. In [6], the DFT-based DMT system is proposed as a practical DMT implementation, but its optimality has not been discussed. We will show that the DFT-based DMT systems are asymptotically optimal, although they are not optimal for a finite number of bands. Furthermore, the asymptotical performance of the DFT-based DMT system is the same as that of the DMT system with ideal filters in [5], [6]. Although the DFT-based DMT system is asymptotically optimal, the optimal transceiver provides significant gain over the DFT-based system for a modest number of bands. Examples will be given to demonstrate this.

A. Outline

In Section II, we derive the polyphase representation of the system model that characterizes the channel, the transmitter, and receiver in the DMT system. Based on the polyphase representation, a modified DFT-based DMT system is proposed in Section III. In Section IV, we develop a more general class of perfect

Manuscript received November 30, 1998; revised December 20, 1999. This work was supported by NSC under Grants 89-2213-E-009-118 and 89-2213-E-002-063, Taiwan, R.O.C. The associate editor coordinating the review of this paper and approving it for publication was Dr. Hitoshi Kiya.

Y.-P. Lin is with the Department 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 Department of Electrical Engineering and Institute of Communications Engineering, National Taiwan University, Taipei, Taiwan, R.O.C.

Publisher Item Identifier S 1053-587X(00)04064-2.

DMT systems. For this class, we will derive the optimal DMT system for a given channel in Section V. The asymptotical performance of the optimal DMT and the DFT-based DMT system will be studied in Section VI.

B. Notations

- 1) Boldfaced lower-case letters are used to represent vectors, and boldfaced upper-case letters are reserved for matrices. The notations \mathbf{A}^T and \mathbf{A}^\dagger represent the transpose and transpose-conjugate of \mathbf{A} .
- 2) The function $\delta(n)$ is defined as

$$\delta(n) = \begin{cases} 1, & n = 0 \\ 0, & \text{otherwise.} \end{cases}$$

- 3) The notation $\text{diag}(a_0, a_1, \dots, a_{M-1})$ denotes an $M \times M$ diagonal matrix with diagonal elements a_k .

II. SYSTEM MODEL AND POLYPHASE REPRESENTATION

Consider Fig. 1, where an M -band DMT system is shown. Usually, the channel is modeled as an LTI filter $C(z)$ with additive noise $e(n)$. Assume that $C(z)$ is an FIR filter of order L (a reasonable assumption after time domain equalization) and $e(n)$ is a zero-mean wide sense stationary random process. For a given number of bands M , the interpolation ratio N is chosen as $N = M + L$. Redundancy is introduced so that the receiver can remove ISI due to $C(z)$ and decoding can be performed blockwise. As the interpolation ratio $N \geq M$, we say that the system is over interpolated. The filters $F_k(z)$ and $H_k(z)$ are called transmitting and receiving filters, respectively. In the DMT system, $F_k(z)$ and $H_k(z)$ have length $=N$. When the outputs \hat{x}_k , $k = 0, 1, \dots, M-1$ are identical to the inputs x_k , $k = 0, 1, \dots, M-1$ in the absence of channel noise, we say that the system is ISI-free or perfect reconstruction (PR).

The transmitting filters $F_k(z)$ have N coefficients

$$F_k(z) = \sum_{n=0}^{N-1} f_k(n)z^{-n}.$$

We can write the $1 \times M$ transmitting bank as

$$\begin{pmatrix} F_0(z) & F_1(z) & \dots & F_{M-1}(z) \end{pmatrix} \\ = (1 \quad z^{-1} \quad \dots \quad z^{-(N-1)}) \mathbf{G}$$

where the $N \times M$ matrix \mathbf{G} has $[\mathbf{G}]_{n,k} = f_k(n)$. The implementation of the transmitter is as shown in Fig. 2. Let the receiving filters be $H_k(z) = \sum_{n=0}^{N-1} h_k(n)z^n$. (Noncausal filters are used here for notational convenience; $N-1$ delays can be added to obtain causal filters.) In a similar manner, we can write the $M \times 1$ receiving bank as

$$\begin{pmatrix} H_0(z) \\ H_1(z) \\ \vdots \\ H_{M-1}(z) \end{pmatrix} = \mathbf{S} \begin{pmatrix} 1 \\ z \\ \vdots \\ z^{N-1} \end{pmatrix}$$

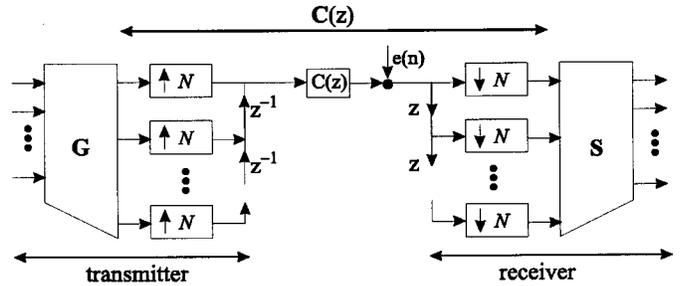


Fig. 2. Polyphase representation of the transmitter and the receiver.

where the $M \times N$ matrix \mathbf{S} has $[\mathbf{S}]_{k,n} = h_k(n)$. The implementation of the receiver is as shown in Fig. 2. Using polyphase representation, we can decompose the channel as

$$C(z) = C_0(z^N) + C_1(z^N)z^{-1} + \dots + C_{N-1}(z^N)z^{-N+1}.$$

It can be shown [9] that the $N \times N$ system incorporating the delay chain, the channel $C(z)$, and the advance chain in Fig. 2 can be redrawn as in Fig. 3. The $N \times N$ channel matrix $\mathbf{C}(z)$ is defined as

$$\mathbf{C}(z) = \begin{pmatrix} C_0(z) & z^{-1}C_{N-1}(z) & \dots & z^{-1}C_1(z) \\ C_1(z) & C_0(z) & & \\ \vdots & & \ddots & \vdots \\ C_{N-1}(z) & C_{N-2}(z) & \dots & C_0(z) \end{pmatrix}. \quad (1)$$

Matrices in the above form are known as pseudo-circulant matrices, and their properties can be found in [9]. With the assumption that the channel $C(z)$ is an FIR filter of order L , we can write $C(z) = c_0 + c_1z^{-1} + \dots + c_Lz^{-L}$. The channel matrix $\mathbf{C}(z)$ is pseudo-circulant with the first column given by $(c_0 \quad c_1 \quad \dots \quad c_L \quad 0 \quad \dots \quad 0)^T$. It can be partitioned as a constant matrix and a transfer matrix with z

$$\mathbf{C}(z) = \begin{pmatrix} \mathbf{C}_0 & \mathbf{C}_1(z) \\ \mathbf{C}_1(z) & \mathbf{C}_2(z) \\ \vdots & \vdots \end{pmatrix} \quad (2)$$

where \mathbf{C}_0 is a lower triangular Toeplitz matrix given by

$$\mathbf{C}_0 = \begin{pmatrix} c_0 & 0 & \dots & 0 & 0 \\ c_1 & c_0 & & \vdots & 0 \\ \vdots & & \ddots & & \vdots \\ c_L & c_{L-1} & & & 0 \\ 0 & c_L & & & 0 \\ \vdots & \vdots & \ddots & & \vdots \\ 0 & 0 & & c_L & \dots & c_0 \\ \vdots & \vdots & & \ddots & \vdots & \\ 0 & 0 & & 0 & \dots & c_L \end{pmatrix}. \quad (3)$$

A. Perfect Reconstruction Condition

From Fig. 3, we see that the overall transfer function $\mathbf{T}(z)$ of the DMT system is

$$\mathbf{T}(z) = \mathbf{S}\mathbf{C}(z)\mathbf{G}. \quad (4)$$

When $\mathbf{T}(z) = \mathbf{I}$, the DMT system has the ISI-free property.

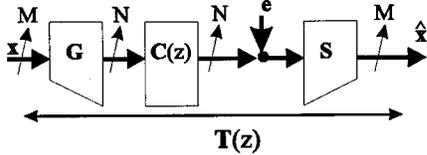


Fig. 3. Polyphase representation of the DMT system.

Fig. 4. Equivalent multirate system for the transfer function $T_{kn}(z)$.

B. Interchange of Transmitter and Receiver

One immediate advantage of the polyphase approach is that it tells us that we can interchange the transmitting and receiving filters and still preserve the PR property. To see this, observe that the matrix $C(z)$ is Toeplitz. It satisfies

$$C(z) = J_N C^T(z) J_N \quad (5)$$

where J_N is the $N \times N$ reversal matrix. For example

$$J_2 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Note that $J_N J_N = I$. Using this fact and (4) and (5), we can get

$$T^T(z) = (G^T J_N) C(z) (J_N S^T).$$

Therefore, if a DMT system with transmitter and receiver pair (G, S) is perfect, then the DMT system $(J_N S^T, G^T J_N)$ is also perfect. This implies that we can exchange the transmitting filters $F_k(z)$ and receiving filters $H_k(z)$, and the system is still perfect, even when the channel is a frequency selective one. This result will be used later in Section III.

C. PR Condition on the Transmitting and Receiving Filters

The polyphase representation allows us to redraw the multirate filter bank in Fig. 1 as the $M \times M$ LTI system in Fig. 3. Note that $T_{kn}(z)$ is the transfer function from the n th input to the k th output; therefore, $T_{kn}(z)$ represents the multirate system in Fig. 4. By applying the so-called polyphase identity [9], such an interconnection yields an LTI system, and the transfer function is the zeroth polyphase of $H_k(z)C(z)F_n(z)$. Therefore, we have

$$T_{kn}(z) = (H_k(z)C(z)F_n(z))_{\downarrow N}.$$

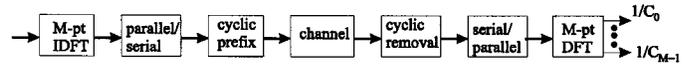
The PR condition can be also be expressed as

$$(H_k(z)C(z)F_n(z))_{\downarrow N} = \delta(k-n), \quad \text{for } 0 \leq k, n \leq M-1.$$

Since $(H_k(z)C(z)F_n(z))_{\downarrow N} = (F_n(z)C(z)H_k(z))_{\downarrow N}$, we can interchange the transmitting and receiving filters without affecting the PR property.

III. MODIFIED DFT-BASED DMT SYSTEMS

The block diagram of the DFT-based DMT system is shown in Fig. 5. The transmitter performs two operations: computing

Fig. 5. M -band DFT-based DMT system.

the M -point inverse DFT of each input block and adding cyclic prefix of length L [10]. This is equivalent to choosing the interpolation ratio N as

$$N = M + L.$$

In general, the length of the prefix L is smaller than M . The redundancy allows the receiver to remove ISI, and the overall system is perfect. The receiver consists of an M -point DFT matrix and M scalars $1/C_k$ for $k = 0, 1, \dots, M-1$, where C_k are the M -point DFT of the channel impulse response [10]. It has the great advantage that the whole system is almost channel independent, except for the M scalars $1/C_k$.

With cyclic prefix added, the transmitter is given by

$$G = \begin{pmatrix} W_1^\dagger \\ W^\dagger \end{pmatrix}$$

where W is the $M \times M$ unitary DFT matrix with

$$[W]_{kn} = \frac{1}{\sqrt{M}} e^{-j(2\pi/M)kn} \quad \text{for } 0 \leq k, n \leq M-1 \quad (6)$$

and W_1 is the $M \times L$ submatrix of W that contains the last L columns of W (assuming $L < M$). The receiver is

$$S = \Gamma^{-1} \begin{pmatrix} 0 & W \end{pmatrix} \quad (7)$$

where Γ is the $M \times M$ diagonal matrix $\text{diag}(C_0, C_1, \dots, C_{M-1})$. The transmitting and receiving filters are, respectively

$$F_k(z) = W^{(L-M)k} \sum_{n=0}^{N-1} W^{-nk} z^{-n},$$

$$H_k(z) = \frac{1}{C_k} z^L \sum_{n=0}^{M-1} W^{nk} z^n, \quad \text{where } W = e^{-j(2\pi/M)}.$$

The receiving filters are DFT filters of length M , and hence, the frequency responses will have a main lobe of width $2\pi/M$.

Now, if we exchange the transmitter and the receiver (with slight modifications), we get the modified DFT-based DMT system [15],

$$G = \begin{pmatrix} W_0^\dagger \\ 0 \end{pmatrix}, \quad S = \Gamma^{-1} [W W_0]$$

where W_0 is the $M \times L$ submatrix of W that contains the first L columns of W . The transmitting and receiving filters are now

$$F_k(z) = \sum_{n=0}^{M-1} W^{-nk} z^{-n}, \quad H_k(z) = \frac{1}{C_k} \sum_{n=0}^{N-1} W^{nk} z^n.$$

The modified system has the same complexity as the conventional case. However, the new receiving filters are DFT filters with length N instead of M in the conventional case. This allows the new system to enjoy additional advantages. First, the new receiving filters have a narrower bandwidth $2\pi/N$. Fig. 6

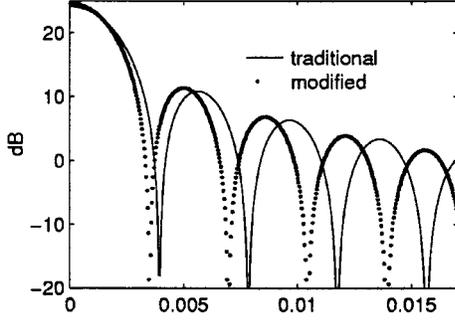


Fig. 6. Magnitude responses of the first receiving filters in the conventional DFT-based DMT system and the modified system for $L = 32$ and $M = 256$. (frequency normalized by 2π).

gives a comparison of the conventional and new receiving filters for the same transmitting power. Only the first receiving filters of these two systems are shown as the other receiving filters are shifted versions of the first filter. The narrower main lobe in the modified case gives a better performance in rejecting out-of-band noise [4]. Moreover, the new receiving filters have longer length. The channel noise will be averaged over a longer block, and the effect of impulsive noise will be reduced.

Note that although the proposed scheme has potential advantage when the channel noise is narrowband or impulsive, its performance is not necessarily better in all types of channel environments. We can always find a channel environment where the conventional system performs as well or better.

IV. GENERALIZED PERFECT DMT SYSTEMS

The transmitter of the modified DMT system can be viewed as the coding of the input block of size M using DFT vectors plus the padding of L zeros. The signal in the k th band is transmitted using the k th DFT vector. We can generalize the system by using the more general orthogonal vectors for transmission instead of DFT vectors. The transmitter becomes a general unitary matrix followed by the padding of L zeros, i.e.,

$$\mathbf{G} = \begin{pmatrix} \mathbf{G}_0 \\ \mathbf{0} \end{pmatrix} \quad (8)$$

where \mathbf{G}_0 is an arbitrary $M \times M$ unitary matrix. As the channel has order L , there will not be any interblock interference (IBI) due to a nonideal channel. With the partition of the channel matrix $\mathbf{C}(z)$ in (2), it follows that

$$\mathbf{C}(z)\mathbf{G} = \mathbf{C}_0\mathbf{G}_0$$

where \mathbf{C}_0 is as defined in (3). Now, the condition for perfect reconstruction becomes

$$\mathbf{S}\mathbf{C}_0\mathbf{G}_0 = \mathbf{I}. \quad (9)$$

This means that \mathbf{S} should be a left inverse of the constant matrix $\mathbf{C}_0\mathbf{G}_0$. Using singular value decomposition (SVD), we can decompose \mathbf{C}_0 as

$$\mathbf{C}_0 = \underbrace{[\mathbf{U}_0 \mathbf{U}_1]}_{\mathbf{U}} \begin{pmatrix} \Lambda \\ \mathbf{0} \end{pmatrix}_{N \times M} \mathbf{V}^T = \mathbf{U}_0 \mathbf{A} \mathbf{V}^T \quad (10)$$

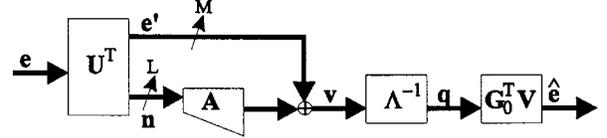


Fig. 7. Block diagram of the general receiver solution.

where \mathbf{U} and \mathbf{V} are, respectively, $N \times N$ and $M \times M$ unitary matrices. The column vectors of \mathbf{U} are the eigenvectors of $\mathbf{C}_0\mathbf{C}_0^T$, and the column vectors of \mathbf{V} are the eigenvectors of $\mathbf{C}_0^T\mathbf{C}_0$. The matrix \mathbf{A} is diagonal

$$\mathbf{A} = \text{diag}(\lambda_0, \lambda_1, \dots, \lambda_{M-1}). \quad (11)$$

The diagonal elements λ_k are the singular values of \mathbf{C}_0 , which are positive as \mathbf{C}_0 has full rank. Therefore, \mathbf{A}^{-1} exists. The SVD of \mathbf{C}_0 immediately gives us one possible choice of \mathbf{S} such that the PR condition in (9) is satisfied

$$\mathbf{S} = \mathbf{G}_0^T \mathbf{V} \mathbf{A}^{-1} \mathbf{U}_0^T. \quad (12)$$

However, the above equation gives only one possible solution. To obtain all solutions, we note that the PR condition in (9) only requires that \mathbf{S} be a left inverse of $\mathbf{C}_0\mathbf{G}_0$. As $\mathbf{C}_0\mathbf{G}_0$ is of dimension $N \times M$, the receiver \mathbf{S} is not unique. In fact, we can choose

$$\mathbf{S} = \mathbf{G}_0^T \mathbf{V} \mathbf{A}^{-1} (\mathbf{I} \ \mathbf{A}) \begin{pmatrix} \mathbf{U}_0^T \\ \mathbf{U}_1^T \end{pmatrix} \quad (13)$$

where \mathbf{A} is an arbitrary $M \times L$ matrix. The flexibility of \mathbf{A} can be exploited to improve the frequency selectivity of the receiving filters [11] or to minimize the total output noise power. The discussion of the later is given next.

A. MMSE Receiver

When the DMT system is perfect, the output noise $\hat{\mathbf{e}} = \hat{\mathbf{x}} - \mathbf{x}$ comes entirely from the channel noise \mathbf{e} . We define the average output noise power \mathcal{E} as $\mathcal{E} = (1/M)E(\hat{\mathbf{e}}^T \hat{\mathbf{e}})$. To analyze the output noise, we draw in Fig. 7 the general receiver \mathbf{S} in (13)

$$\mathbf{e}' = \mathbf{U}_0^T \mathbf{e}, \quad \text{and} \quad \mathbf{n} = \mathbf{U}_1^T \mathbf{e} \quad (14)$$

$$\mathbf{v} = \mathbf{e}' + \mathbf{A}\mathbf{n} \quad (15)$$

$$\mathbf{q} = \mathbf{A}^{-1}\mathbf{v} \quad (16)$$

$$\hat{\mathbf{e}} = \mathbf{G}_0^T \mathbf{V} \mathbf{q}. \quad (17)$$

Observe that the last part $\mathbf{G}_0^T \mathbf{V}$ is a unitary matrix and that unitary matrices preserve input energy; therefore, $E(\hat{\mathbf{e}}^T \hat{\mathbf{e}}) = E(\mathbf{q}^T \mathbf{q})$. As \mathbf{A}^{-1} is a diagonal matrix with positive diagonal elements, $E(\mathbf{q}^T \mathbf{q})$ is minimized if $\sigma_{v_k}^2$ is minimized for each k . However, v_k is related to the vectors \mathbf{e}' and \mathbf{n} by

$$v_k = [\mathbf{e}']_k + \mathbf{a}_k^T \mathbf{n}$$

where $[\mathbf{e}']_k$ denotes the k th element of \mathbf{e}' , and \mathbf{a}_k^T is the k th row of \mathbf{A} . Now, the minimization of $\sigma_{v_k}^2$ for each k can be considered as a linear estimation problem; \mathbf{a}_k should be the optimal predictor of \mathbf{e}'_k given the observation vector \mathbf{n} . By the

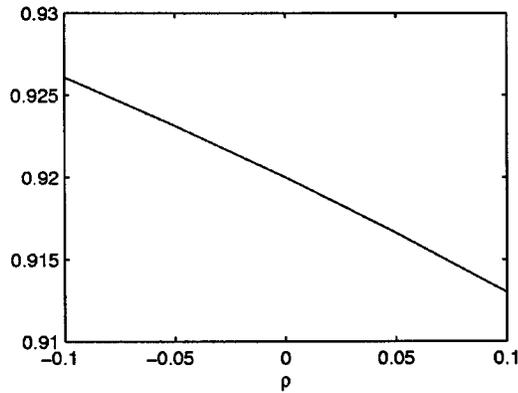


Fig. 8. Plot of the ratio of the average noise power for the MMSE receiver over the average noise power for the receiver in (12) for the channel $1 + \rho z^{-1}$.

orthogonality principle, the optimal choice of \mathbf{A} is given by $\mathbf{A} = -E(\mathbf{e}'\mathbf{n}^T)(E(\mathbf{n}\mathbf{n}^T))^{-1}$. Using (14), we have

$$\mathbf{A} = -\mathbf{U}_0^T \mathbf{R}_N \mathbf{U}_1 (\mathbf{U}_1^T \mathbf{R}_N \mathbf{U}_1)^{-1}$$

where $\mathbf{R}_N = E(\mathbf{e}\mathbf{e}^T)$ is the $N \times N$ autocorrelation matrix of \mathbf{e} . The optimal solution of \mathbf{A} is zero if \mathbf{e}' and \mathbf{n} are uncorrelated. One case where this happens is when the channel noise $e(n)$ is white. The noise vector \mathbf{e} has autocorrelation matrix $\mathbf{R}_N = \sigma_e^2 \mathbf{I}$ and after the unitary transformation \mathbf{U}^T , the vectors \mathbf{e}' and \mathbf{n} are uncorrelated.

Example 1: Let the channel be $C(z) = 1 + \rho z^{-1}$ with a NEXT noise source [1]. As ρ changes from -1 to 1 , the channel changes from a lowpass filter to a highpass filter. Using ρ as a parameter, we can observe the gain of using the MMSE receiver in different channel environments. For the same transmitter $\mathbf{G}_0 = \mathbf{V}$, we compute the average output noise power \mathcal{E} when the receiver is as in (13) and the average output noise power \mathcal{E}_{mmse} when MMSE receiver is used. Fig. 8 shows the ratio $\mathcal{E}_{mmse}/\mathcal{E}$ as a function of the parameter ρ .

B. Minimum-Redundancy DMT Systems

In the generalization of (8), each input block of size M is passed through a unitary transformation, and then, L zeros are appended to each block. As the channel has order L , there is no block overlapping. The receiver can easily perform blockwise decoding as there is no interblock interference (IBI). However zero padding introduces redundancy and equalization is done at the expense of a data rate loss of $L/(M+L)$. The question that arises is the following: Is it possible to reduce the amount of zero padding without introducing IBI? The answer is in the affirmative. With minor conditions on the channel $c(n)$, the number of padding zeros can be halved. In particular, the minimum number of padding zeros is $\lceil L/2 \rceil$, where the function $\lceil a \rceil$ denotes the smallest integer that is greater than or equal to the number a .

Observe that to make successive block decoding independent, the receiver requires at least M samples that are not contaminated by IBI from adjacent blocks. Suppose the number of padding zeros is $\ell < L$ and that the output block of the transmitter has size $M + \ell$. After passing through the channel, the information of each input block spreads over $M + L$ samples, and the information spills over to the next block by $L - \ell$ samples.

Therefore, the first $L - \ell$ samples of each size $M + \ell$ block contain information from the previous block. The number of samples that contains no IBI available to the receiver is $(M + \ell) - (L - \ell) = M + (2\ell - L)$. Therefore, we need $2\ell - L \geq 0$ or $\ell_{\min} = \lceil L/2 \rceil$. Suppose L is even and $\ell = L/2$. Then, the IBI-free samples are the last M samples of each block. For the transmitter input block \mathbf{x} , the receiver gets $\mathbf{C}_0 \mathbf{G}_0 \mathbf{x}$, where \mathbf{C}_0 is as defined in (3). Therefore, the clean samples are $\mathbf{C}_{10} \mathbf{G}_0 \mathbf{x}$, where \mathbf{C}_{10} is the bottom $M \times M$ submatrix of \mathbf{C}_0 . That is

$$\mathbf{C}_{10} = \begin{pmatrix} c_{L/2} & \cdots & c_0 & 0 & \cdots & 0 \\ \vdots & \ddots & & \ddots & & \\ c_L & & c_{L/2} & & & \\ 0 & \ddots & & \ddots & \ddots & 0 \\ \vdots & & & & & c_0 \\ 0 & & 0 & c_L & \cdots & c_{L/2} \end{pmatrix}_{M \times M}$$

As long as \mathbf{C}_{10} is nonsingular, we can invert $\mathbf{C}_{10} \mathbf{G}_0$ to recover \mathbf{x} .

V. OPTIMAL TRANSCEIVERS

We first derive the bit allocation formula for the generalized DMT system such that the transmitting power can be minimized for a given bit rate and a probability of symbol error. Then, we show how to design the optimal transceiver for arbitrary colored noise.

A. Bit Allocation

Let the number of bits allocated for the k th band be b_k ; then, the average number of bits per symbol is $b = (1/M) \sum_{k=0}^{M-1} b_k$. To account for the bit rate reduction due to zero padding, the average bit rate R_b is

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

The input power of the k th band is $\sigma_{x_k}^2$, which is also the output signal power of the k th band at the receiver end due to the PR property. Suppose the output noise power of the k th band is $\hat{\sigma}_{e_k}^2$. For a given probability of error P_e , most modulation systems under high bit rate assumption satisfy

$$\sigma_{x_k}^2 = c 2^{2b_k} \hat{\sigma}_{e_k}^2 \quad (18)$$

where the constant c depends on P_e . For example, in the case of PAM, the probability of error P_e is related to signal power $\sigma_{x_k}^2$ and noise power $\hat{\sigma}_{e_k}^2$ by

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

where

$$Q(y) = \frac{1}{\sqrt{2\pi}} \int_y^\infty e^{-t^2/2} dt, \quad y \geq 0.$$

Under high bit rate assumption, we have $2^{b_k} - 1 \approx 2^{b_k}$. This approximation leads to (18) where c is given by $c = 1/3(Q^{-1}(P_e/2))^2$.

Define $\mathcal{P}(R_b, P_e, M)$ as the transmission power needed in an M -band DMT system for a given bit rate R_b and probability of error P_e . As the unitary transformation \mathbf{G}_0 preserves energy, the average transmission power is

$$\mathcal{P}(R_b, P_e, M) = \frac{1}{M} \sum_{k=0}^{M-1} \sigma_{x_k}^2.$$

Applying the arithmetic mean (AM) over geometric mean (GM) inequality to the above equation and using (18), we have

$$\begin{aligned} \mathcal{P}(R_b, P_e, M) &\geq \left(\prod_{k=0}^{M-1} \sigma_{x_k}^2 \right)^{1/M} \\ &= c 2^{2R_b(N/M)} \left(\prod_{k=0}^{M-1} \hat{\sigma}_{e_k}^2 \right)^{1/M}. \end{aligned} \quad (19)$$

The equality holds if and only if the bits are optimally allocated according to

$$b_k = b - \log_2 \hat{\sigma}_{e_k} + \frac{1}{M} \log_2 \prod_{k=0}^{M-1} \hat{\sigma}_{e_k}. \quad (20)$$

Let us define the coding gain of bit allocation \mathcal{CG} as

$$\mathcal{CG} = \frac{\mathcal{P}_{direct}(R_b, P_e, M)}{\mathcal{P}(R_b, P_e, M)}$$

where $\mathcal{P}_{direct}(R_b, P_e, M)$ is the power needed when there is no bit allocation. Without bit allocation, we have $b_k = b$, for $k = 0, 1, \dots, M-1$; therefore

$$\mathcal{P}_{direct}(R_b, P_e, M) = c 2^{2R_b(N/M)} \frac{1}{M} \sum_{k=0}^{M-1} \hat{\sigma}_{e_k}^2. \quad (21)$$

Using (19), the coding gain of bit allocation is

$$\mathcal{CG} = \frac{\frac{1}{M} \sum_{k=0}^{M-1} \hat{\sigma}_{e_k}^2}{\left(\prod_{k=0}^{M-1} \hat{\sigma}_{e_k}^2 \right)^{1/M}} \geq 1. \quad (22)$$

The above inequality follows from the *AM* over *GM* inequality. The coding gain is the *AM* over the *GM* of the output noise variances $\hat{\sigma}_{e_k}^2$. Note that this ratio depends on the choice of unitary transformation \mathbf{G}_0 of the transmitter. In the next subsection, we show how to design \mathbf{G}_0 so that the coding gain can be maximized.

B. Design of Optimal Transceivers

Fig. 9 shows the individual parts of the receiver \mathbf{S} in (12). We see the last part of the receiver is the unitary matrix $\mathbf{G}_0^T \mathbf{V}$. Let us call it \mathbf{Q} . As \mathbf{Q} preserves input energy, we have $\sum_{k=0}^{M-1} \hat{\sigma}_{e_k}^2 = \sum_{k=0}^{M-1} \tilde{\sigma}_{e_k}^2$, which is a quantity independent of \mathbf{G} . The maximization of coding gain in (22) becomes the

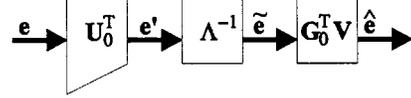


Fig. 9. Receiver block diagram for the derivation of the optimal DMT system.

problem of minimizing the product $\prod_{k=0}^{M-1} \hat{\sigma}_{e_k}^2$. Let $\hat{\mathbf{R}}$ be the autocorrelation matrix of $\hat{\mathbf{e}}$; then, $\hat{\sigma}_{e_k}^2$ are the diagonal entries of $\hat{\mathbf{R}}$, and $\prod_{k=0}^{M-1} \hat{\sigma}_{e_k}^2 = \prod_{k=0}^{M-1} [\hat{\mathbf{R}}]_{kk}$. The matrix $\hat{\mathbf{R}}$ is related to $\tilde{\mathbf{R}}$, which is the autocorrelation function of $\tilde{\mathbf{e}}$, by

$$\hat{\mathbf{R}} = \mathbf{Q} \tilde{\mathbf{R}} \mathbf{Q}^T. \quad (23)$$

Using the Hadamard inequality for positive definite matrices [17], we have

$$\prod_{k=0}^{M-1} \hat{\sigma}_{e_k}^2 = \prod_{k=0}^{M-1} [\hat{\mathbf{R}}]_{kk} \geq \det \hat{\mathbf{R}} = \det \tilde{\mathbf{R}}$$

which is a fixed quantity independent of \mathbf{Q} . The equality holds if and only if the matrix $\hat{\mathbf{R}}$ is diagonal (see, e.g., [9]). Therefore, the optimal \mathbf{Q} is the unitary matrix \mathbf{Q}_0 that decorrelates $\tilde{\mathbf{e}}$. In this case, $\prod_{k=0}^{M-1} \hat{\sigma}_{e_k}^2$ is minimized and is equal to $\det \tilde{\mathbf{R}}$. The minimum power required in the optimal DMT system is

$$\mathcal{P}_{opt}(R_b, P_e, M) = c 2^{2R_b N/M} (\det \tilde{\mathbf{R}})^{1/M}. \quad (24)$$

Comparing $\mathcal{P}_{opt}(R_b, P_e, M)$ and $\mathcal{P}_{direct}(R_b, P_e, M)$ in (21), we obtain the coding gain

$$\mathcal{CG} = \frac{\frac{1}{M} \text{trace}(\tilde{\mathbf{R}})}{(\det \tilde{\mathbf{R}})^{1/M}}.$$

Note that this is the coding gain formula for the optimal transmitter when the input random vector has autocorrelation matrix $\tilde{\mathbf{R}}$. The optimal DMT is given by

$$\mathbf{G} = \begin{pmatrix} \mathbf{V} \mathbf{Q}_0^T \\ \mathbf{0} \end{pmatrix}, \quad \mathbf{S} = \mathbf{Q}_0 \Lambda^{-1} \mathbf{U}_0^T. \quad (25)$$

Note that the receiver can be replaced by the MMSE receiver corresponding to the above choice of \mathbf{G} as derived in the previous section to further minimize the average output noise power.

AWGN Channels: When the channel noise is a white process, the autocorrelation matrix $\mathbf{R}_N = \sigma_e^2 \mathbf{I}$ is a diagonal matrix, and so is $\tilde{\mathbf{R}}$. The noise vector $\tilde{\mathbf{e}}$ is already uncorrelated; therefore, $\mathbf{Q} = \mathbf{I}$. The optimal transmitter \mathbf{G}_0 is simply $\mathbf{G}_0 = \mathbf{V}$, and the receiver is $\mathbf{S} = \mathbf{U}_0^T$. (This optimal solution in this case is consistent with what Kasturia *et al.* have obtained for AWGN channels from the viewpoint of multidimensional signal constellations.) We call this design of DMT system AWGN-optimal as it is optimal for AWGN noise. The coding gain is

$$\mathcal{CG}_{\max} = \frac{\frac{1}{M} \sum_{k=0}^{M-1} 1/\lambda_k^2}{\left(\prod_{k=0}^{M-1} 1/\lambda_k^2 \right)^{1/M}}$$

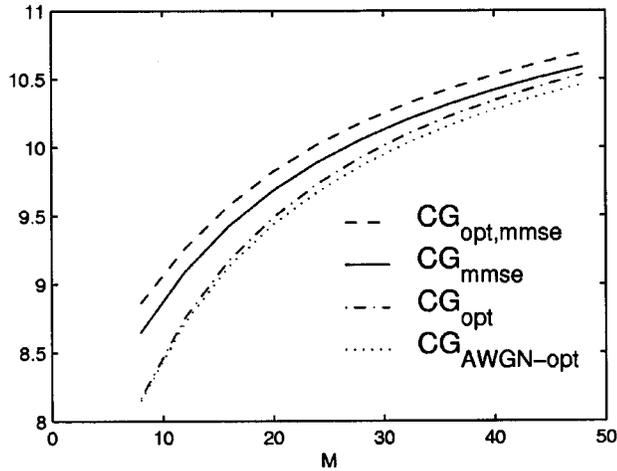


Fig. 10. Comparison of coding gains for different DMT systems.

where λ_k are the diagonal entries of the diagonal matrix $\mathbf{\Lambda}$ given in (11).

Example 2: Let the channel be $C(z) = 1 + 0.5z^{-1}$ with a NEXT noise source. For the same probability of error and same bit rate, Fig. 10 shows the coding gain for different DMT systems:

- 1) coding gain $\mathcal{CG}_{\text{AWGN-opt}}$ for AWGN-optimal DMT, i.e., the system with transmitter $\mathbf{G}_0 = \mathbf{V}$ and receiver $\mathbf{S} = \mathbf{U}_0 \mathbf{\Lambda}^{-1}$;
- 2) coding gain $\mathcal{CG}_{\text{opt}}$ for the optimal DMT in (25);
- 3) coding gain $\mathcal{CG}_{\text{mmse}}$ for the DMT with AWGN-optimal transmitter $\mathbf{G}_0 = \mathbf{V}$ and a corresponding MMSE receiver;
- 4) coding gain $\mathcal{CG}_{\text{opt,mmse}}$ for the DMT with the optimal transmitter design in (25) and a corresponding MMSE receiver.

Optimal DMT Systems with Ideal Filters: In the DMT systems that we have discussed so far, the transmitting and receiving filters have length N . The transmitter \mathbf{G} and receiver \mathbf{S} are constant matrices. If we allow the DMT system to have longer filters, we gain extra design flexibility. For example, it can be shown that it is possible to obtain perfect reconstruction DMT transceiver without introducing redundancy to the system if ideal filters can be used [11]. In this case the transmitter and the receiver are $M \times M$ transfer matrices $\mathbf{G}(e^{j\omega})$ and $\mathbf{S}(e^{j\omega})$. The channel matrix $\mathbf{C}(e^{j\omega})$ is $M \times M$ and pseudo-circulant. Suppose the transmitter $\mathbf{G}(e^{j\omega})$ is orthonormal, i.e., $\mathbf{G}(e^{j\omega})$ is unitary for all ω ([9, ch. 6]). When the channel $C(z)$ has no zeros on the unit circle, it can be verified that the channel matrix $\mathbf{C}(e^{j\omega})$ is nonsingular. In this case, if we choose

$$\mathbf{S}(e^{j\omega}) = \mathbf{G}^\dagger(e^{j\omega}) \mathbf{C}^{-1}(e^{j\omega})$$

then the DMT system is perfect. We can verify that the receiving filters have the form $H_k(e^{j\omega}) = F_k^*(e^{j\omega})/C(e^{j\omega})$. Using a procedure similar to that in Section V-A, it can be further verified that under optimal bit allocation, the transmitting power is

$$P_{\text{ideal}} = c2^{2R_b(N/M)} E_0$$

where

$$E_0 = \left(\prod_{k=0}^{M-1} \int_{-\pi}^{\pi} |F_k(e^{j\omega})|^2 \frac{S_{ee}(e^{j\omega})}{|C(e^{j\omega})|^2} \frac{d\omega}{2\pi} \right)^{1/M}$$

where $S_{ee}(e^{j\omega})$ is the power spectral density of the channel noise $e(n)$ in Fig. 1. The design of the optimal DMT system for minimizing transmission power becomes the problem of designing orthonormal filters $F_k(e^{j\omega})$ such that E_0 is minimized. This problem is the same as designing optimal orthonormal filters for maximizing coding gain in filter bank theory [12]. The filters $F_k(e^{j\omega})$ should be the optimal orthonormal filters for the power spectrum $S_{ee}(e^{j\omega})/|C(e^{j\omega})|^2$.

VI. ASYMPTOTICAL OPTIMALITY OF DFT-BASED DMT SYSTEMS

Although the DFT-based DMT systems are not optimal in general, they are asymptotically optimal, regardless of the type of channel noise. The performance of the DFT-based DMT systems approaches that of optimal DMT systems as the number of bands M increases. In particular, for a given error probability and bit rate, the power required in DFT-based DMT system approaches that required in the optimal system when M is sufficiently large.

Let \mathbf{R}_k be the $k \times k$ autocorrelation function of the noise process $e(n)$. Using $\mathbf{R}' = \mathbf{U}_0^T \mathbf{R}_N \mathbf{U}_0$ and $\tilde{\mathbf{R}} = \mathbf{\Lambda}^{-1} \mathbf{R}' \mathbf{\Lambda}^{-1}$, we can rewrite the transmission power in (24) as

$$\mathcal{P}_{\text{opt}}(R_b, P_e, M) = c2^{2R_b N/M} \frac{(\det(\mathbf{U}_0^T \mathbf{R}_N \mathbf{U}_0))^{1/M}}{\left(\prod_{k=0}^{M-1} \lambda_k^2 \right)^{1/M}} \quad (26)$$

under optimal bit allocation. For the DFT-based system, the receiver is as given in (7) so that the minimum power under optimal bit allocation is

$$\mathcal{P}_{\text{DFT}}(R_b, P_e, M) = c2^{2R_b N/M} \frac{\left(\prod_{k=0}^{M-1} [\mathbf{W} \mathbf{R}_M \mathbf{W}^\dagger]_{kk} \right)^{1/M}}{\left(\prod_{k=0}^{M-1} |C_k|^2 \right)^{1/M}} \quad (27)$$

where \mathbf{W} is the $M \times M$ DFT matrix, and C_k is the M -point DFT of the channel impulse response.

Using the distribution of eigenvalues for Toeplitz matrices [13], we can show that (see Appendix A)

$$\begin{aligned} \lim_{M \rightarrow \infty} \left(\prod_{k=0}^{M-1} \lambda_k^2 \right)^{1/M} &= \lim_{M \rightarrow \infty} \left(\prod_{k=0}^{M-1} |C_k|^2 \right)^{1/M} \\ &= \exp \left(\int_{-\pi}^{\pi} \ln |C(e^{j\omega})|^2 \frac{d\omega}{2\pi} \right). \end{aligned} \quad (28)$$

In addition, using properties of positive definite matrices, we can show that (see Appendix B)

$$\lim_{M \rightarrow \infty} (\det(\mathbf{U}_0^T \mathbf{R}_N \mathbf{U}_0))^{1/M} = \exp\left(\int_{-\pi}^{\pi} \ln S_{ee}(e^{j\omega}) \frac{d\omega}{2\pi}\right). \quad (29)$$

On the other hand, it can be shown that [14]

$$\begin{aligned} \lim_{M \rightarrow \infty} \left(\prod_{k=0}^{M-1} [\mathbf{W} \mathbf{R}_M \mathbf{W}^\dagger]_{kk} \right)^{1/M} \\ = \exp\left(\int_{-\pi}^{\pi} \ln S_{ee}(e^{j\omega}) \frac{d\omega}{2\pi}\right). \end{aligned} \quad (30)$$

With the equalities in (28)–(30), we can establish that

$$\begin{aligned} \lim_{M \rightarrow \infty} \mathcal{P}_{opt}(R_b, P_e, M) \\ = \lim_{M \rightarrow \infty} \mathcal{P}_{DFT}(R_b, P_e, M) \\ = c 2^{2R_b} \exp\left(\int_{-\pi}^{\pi} \ln \frac{S_{ee}(e^{j\omega})}{|C(e^{j\omega})|^2} \frac{d\omega}{2\pi}\right). \end{aligned} \quad (31)$$

This is the same bound achieved in [5] by studying the asymptotical performance of M -band DMT systems with ideal brick-wall filters.

Note that the DMT system developed in [8] does not achieve this bound asymptotically. To see this, let $C(z) = 1$; then, the transmitter and receiver are identity matrices. The coding gain of the system in [8] is one, regardless of the number of bands. On the other hand, the coding gain corresponding to the asymptotic bound in (31) is always greater than one if the channel noise is not white.

Example 3: Let the channel be $C(z) = 1 + 0.5z^{-1}$. For the same probability of error and same bit rate, we can obtain the ratio of power needed in optimal system over the power needed in the DFT-based system ($\mathcal{P}_{opt}(R_b, P_e, M)/(\mathcal{P}_{DFT}(R_b, P_e, M))$) using (26) and (27)

$$\begin{aligned} \frac{\mathcal{P}_{opt}(R_b, P_e, M)}{\mathcal{P}_{DFT}(R_b, P_e, M)} \\ = \frac{\left(\prod_{k=0}^{M-1} |C_k|^2\right)^{1/M} (\det(\mathbf{U}_0^T \mathbf{R}_N \mathbf{U}_0))^{1/M}}{\left(\prod_{k=0}^{M-1} \lambda_k^2\right)^{1/M} \left(\prod_{k=0}^{M-1} [\mathbf{W} \mathbf{R}_M \mathbf{W}^\dagger]_{kk}\right)^{1/M}}. \end{aligned}$$

Note that the ratio is a quantity independent of the given bit rate and probability of error. Fig. 11 shows the ratio as a function of M for two different noise sources: the AWGN and NEXT noise source. From Fig. 11, we see that the ratio ($\mathcal{P}_{opt}(R_b, P_e, M)/(\mathcal{P}_{DFT}(R_b, P_e, M))$) approaches 1 as the number of bands M increases. However, in the colored NEXT case, the ratio approach 1 only for very large M ; the optimal system provides significant gain for a moderate number of bands.

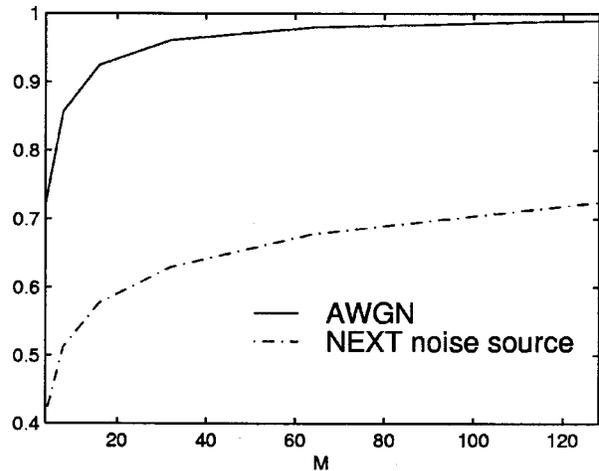


Fig. 11. Asymptotical performances of the DFT-based and the optimal DMT systems. The ratio of the power needed in the optimal DMT system over the power needed in DFT-based system for the same probability of error and the same bit rate for two types of channel noise.

APPENDIX A PROOF OF (28)

Equation (28) is a result for sequences of asymptotically equivalent matrices. Define the strong norm $\|\cdot\|$ and the weak norm $|\cdot|$ of an $n \times n$ matrix \mathbf{A} , respectively, as

$$\begin{aligned} \|\mathbf{A}\| &= \max_{\mathbf{v}} \left(\frac{\mathbf{v}^\dagger \mathbf{A} \mathbf{v}}{\mathbf{v}^\dagger \mathbf{v}} \right)^{1/2} \\ |\mathbf{A}| &= \left(\frac{1}{n} \text{trace}(\mathbf{A}^\dagger \mathbf{A}) \right)^{1/2}. \end{aligned}$$

Let \mathbf{A}_n and \mathbf{B}_n be two sequences of $n \times n$ Hermitian matrices. The subscript n of \mathbf{A}_n denotes the sequence index and indicates that the size of \mathbf{A}_n is $n \times n$. The size of the matrices \mathbf{A}_n grows with the sequence index n . The sequences \mathbf{A}_n and \mathbf{B}_n are said to be asymptotically equivalent [13] if

$$\|\mathbf{A}_n\|, \|\mathbf{B}_n\| \leq M_1 < \infty, \quad \text{and} \quad \lim_{n \rightarrow \infty} |\mathbf{A}_n - \mathbf{B}_n| = 0.$$

Suppose \mathbf{A}_n and \mathbf{B}_n have eigenvalues $\alpha_{n,k}$ and $\beta_{n,k}$, and $M_0 \leq \alpha_{n,k}, \beta_{n,k} \leq M_1$. In [13], Gray shows that

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=0}^{n-1} F(\alpha_{n,k}) = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=0}^{n-1} F(\beta_{n,k})$$

where $F(\cdot)$ is an arbitrary function that is continuous on $[M_0, M_1]$.

To show (28), we observe that $|\lambda_k|^2$, for $k = 0, 1, \dots, M-1$, are the eigenvalues of $\mathbf{A}_M = \mathbf{C}_0^T \mathbf{C}_0$, where the subscript M indicates that \mathbf{A}_M is an $M \times M$ matrix. Now, we construct a sequence of matrices that is asymptotically equivalent to \mathbf{A}_M , and their eigenvalues are $|C_k|^2$. Partition \mathbf{C}_0 as

$$\mathbf{C}_0 = \begin{pmatrix} \mathbf{C}_{00} \\ \mathbf{C}_{01} \end{pmatrix}$$

where \mathbf{C}_{00} is an $M \times M$ matrix. Define

$$\hat{\mathbf{C}}_0 = \mathbf{C}_{00} + \begin{pmatrix} \mathbf{C}_{01} \\ \mathbf{0} \end{pmatrix}.$$

Then, it can be verified that $\hat{\mathbf{C}}_0$ is an $M \times M$ circulant matrix with the zeroth column given by

$$(c_0 \ c_1 \ \cdots \ c_L \ 0 \ \cdots \ 0)^T.$$

It is known that circulant matrices can be diagonalized by DFT matrices

$$\hat{\mathbf{C}}_0 = \mathbf{W}\mathbf{T}\mathbf{W}^\dagger$$

where \mathbf{W} is the $M \times M$ unitary DFT matrix as defined in (6). The matrix \mathbf{T} is diagonal, and the diagonal elements are C_k , which is the M -point DFT of the channel impulse response $c(n)$. Letting $\mathbf{B}_M = \hat{\mathbf{C}}_0^\dagger \hat{\mathbf{C}}_0$, we then have

$$\mathbf{B}_M = \mathbf{W}\mathbf{T}^\dagger \mathbf{T}\mathbf{W}.$$

The diagonal terms of $\mathbf{T}^\dagger \mathbf{T}$ are $|C_k|^2$ so that the eigenvalues of \mathbf{B}_M are $|C_k|^2$. It can be verified that \mathbf{A}_M and \mathbf{B}_M are asymptotically equivalent; therefore

$$\lim_{M \rightarrow \infty} \frac{1}{M} \sum_{k=0}^{M-1} \ln \lambda_k^2 = \lim_{M \rightarrow \infty} \frac{1}{M} \sum_{k=0}^{M-1} \ln |C_k|^2.$$

Note that

$$\exp\left(\frac{1}{M} \sum_{k=0}^{M-1} \ln \lambda_k^2\right) = \left(\prod_{k=0}^{M-1} \lambda_k^2\right)^{1/M}$$

and

$$\exp\left(\frac{1}{M} \sum_{k=0}^{M-1} \ln |C_k|^2\right) = \left(\prod_{k=0}^{M-1} |C_k|^2\right)^{1/M}.$$

Letting M go to infinity, we obtain

$$\lim_{M \rightarrow \infty} \exp\left(\frac{1}{M} \sum_{k=0}^{M-1} \ln \lambda_k^2\right) = \lim_{M \rightarrow \infty} \left(\prod_{k=0}^{M-1} \lambda_k^2\right)^{1/M}$$

and

$$\lim_{M \rightarrow \infty} \exp\left(\frac{1}{M} \sum_{k=0}^{M-1} \ln |C_k|^2\right) = \lim_{M \rightarrow \infty} \left(\prod_{k=0}^{M-1} |C_k|^2\right)^{1/M}.$$

Observe that C_k for $k = 0, 1, \dots, M-1$ are the M -point DFT of $c(n)$; they are samples of $C(e^{j\omega})$, i.e., $C_k = C(e^{j2k\pi/M})$. Therefore

$$\begin{aligned} \lim_{M \rightarrow \infty} \exp\left(\frac{1}{M} \sum_{k=0}^{M-1} \ln |C_k|^2\right) \\ = \exp\left(\int_{-\pi}^{\pi} \ln |C(e^{j\omega})|^2 \frac{d\omega}{2\pi}\right). \end{aligned}$$

Equation (28) follows.

APPENDIX B PROOF OF (29)

Note that the matrix $\mathbf{U}_0^T \mathbf{R}_N \mathbf{U}_0$ is the $M \times M$ leading principle submatrix of $\mathbf{P} = \mathbf{U}^T \mathbf{R}_N \mathbf{U}$, where \mathbf{U} is as defined in (10). Let the eigenvalues of \mathbf{P} be ordered as $\gamma_0 \leq \gamma_1 \leq \dots \leq$

γ_{N-1} . Using the interlacing property of eigenvalues for positive definite matrices [17], it can be shown that $\det(\mathbf{U}_0^T \mathbf{R}_N \mathbf{U}_0)$ is bounded between the product of the M largest eigenvalues and the product of the M smallest eigenvalues

$$\gamma_0 \gamma_1 \cdots \gamma_{M-1} \leq \det(\mathbf{U}_0^T \mathbf{R}_N \mathbf{U}_0) \leq \gamma_L \gamma_{L+1} \cdots \gamma_{N-1}.$$

Suppose the power spectral density $S_{ee}(e^{j\omega})$ of the channel noise has minimum $S_{\min} > 0$ and maximum $S_{\max} < \infty$. Then, these eigenvalues are bounded between S_{\min} and S_{\max} , in particular, $S_{\min} \leq \gamma_0 \leq \gamma_1 \leq \dots \leq \gamma_{N-1} \leq S_{\max}$. It follows that

$$\begin{aligned} \det(\mathbf{U}_0^T \mathbf{R}_N \mathbf{U}_0) &\leq \gamma_L \gamma_{L+1} \cdots \gamma_{N-1} \\ &= \frac{\det \mathbf{P}}{\gamma_0 \gamma_1 \cdots \gamma_{L-1}} \\ &\leq \frac{\det \mathbf{P}}{\gamma_0^L} \\ &\leq \frac{\det \mathbf{P}}{S_{\min}^L} \end{aligned} \quad (32)$$

$$\begin{aligned} \det(\mathbf{U}_0^T \mathbf{R}_N \mathbf{U}_0) &\geq \gamma_0 \gamma_1 \cdots \gamma_{M-1} \\ &= \frac{\det \mathbf{P}}{\gamma_M \gamma_{M+1} \cdots \gamma_{N-1}} \\ &\geq \frac{\det \mathbf{P}}{\gamma_{N-1}^L} \\ &\geq \frac{\det \mathbf{P}}{S_{\max}^L} \end{aligned} \quad (33)$$

Combining (32) and (33), we have

$$\frac{\det \mathbf{P}}{S_{\max}^L} \leq \det(\mathbf{U}_0^T \mathbf{R}_N \mathbf{U}_0) \leq \frac{\det \mathbf{P}}{S_{\min}^L}. \quad (34)$$

In addition, observe that $\det \mathbf{P} = \det \mathbf{R}_N$. The matrix \mathbf{R}_N is Toeplitz, and it is the $N \times N$ autocorrelation matrix of $S_{ee}(e^{j\omega})$. It is known that [16]

$$\lim_{N \rightarrow \infty} (\det \mathbf{R}_N)^{1/N} = \exp\left(\int_{-\pi}^{\pi} \ln S_{ee}(e^{j\omega}) \frac{d\omega}{2\pi}\right).$$

Letting M go to ∞ in (34), we arrive at (29).

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] A. N. Akansu, P. Duhamel, X. Lin, and M. de Courville, "Orthogonal transmultiplexers in communication: A review," *IEEE Trans. Signal Processing*, vol. 46, pp. 979-995, Apr. 1998.
- [3] J. W. Lechleider, "High bit rate digital subscriber lines: A review of HDSL progress," *IEEE J. Select. Areas Commun.*, vol. 9, pp. 769-784, Aug. 1991.
- [4] G. W. Wornell, "Emerging applications of multirate signal processing and wavelets in digital communications," *Proc. IEEE*, vol. 84, Apr. 1996.
- [5] I. Kalet, "The multitone channel," *IEEE Trans. Commun.*, vol. 37, pp. 119-124, Feb. 1989.
- [6] 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.
- [7] X. G. Gia, "A new precoding for ISI cancellation using multirate filterbanks," in *Proc. Int. Symp. Circuits Syst.*, Hong Kong, 1997.
- [8] S. Kasturia, J. T. Aslanis, and J. M. Cioffi, "Vector coding for partial response channels," *IEEE Trans. Inform. Theory*, vol. 36, July 1990.
- [9] P. P. Vaidyanathan, *Multirate Systems and Filter Banks*. Englewood Cliffs, NJ: Prentice-Hall, 1993.
- [10] J. G. Proakis, *Digital Communications*. New York: McGraw-Hill, 1995.

- [11] Y.-P. Lin and S.-M. Phoong, "Wavelet based DMT systems," in preparation.
- [12] P. P. Vaidyanathan, "Theory of optimal orthonormal coder," *IEEE Trans. Signal Processing*, vol. 46, pp. 1528–1543, June 1998.
- [13] R. M. Gray, "On the asymptotic eigenvalue distribution of toeplitz matrices," *IEEE Trans. Information Theory*, vol. IT-16, Nov. 1972.
- [14] A. Gersho and R. M. Gray, *Vector Quantization and Signal Compression*. Boston, MA: Kluwer, 1991.
- [15] Y.-P. Lin and S.-M. Phoong, "Perfect discrete wavelet multitone modulation for fading channels," in *Proc. IEEE Int. Workshop Intelligent Signal Process. Commun. Syst.*, Melbourne, Australia, Nov. 1998.
- [16] N. S. Jayant and P. Noll, *Digital Coding of Waveforms*. Englewood Cliffs, NJ: Prentice-Hall, 1984.
- [17] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge, U.K.: Cambridge Univ. Press, 1985.



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 filter banks and their applications to communication.

Dr. Phoong is the recipient of the 1997 Wilts Prize from Caltech for outstanding independent research in electrical engineering.



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 Ph.D. degrees, both in electrical engineering, from California Institute of Technology, Pasadena, in 1993 and 1997, respectively.

She joined the Department of Electrical and Control Engineering, NCTU, in 1997. Her research interests include multirate filter banks, wavelets, and applications to communication systems. She is currently

an Associate Editor of *Multidimensional Systems and Signal Processing*.