ASYMPTOTICAL OPTIMALITY OF DFT BASED DMT TRANSCEIVERS

Yuan-Pei Lin
Dept. Electrical and Control Engr.
National Chiao Tung Univ.
Hsinchu, Taiwan, R.O.C.

See-May Phoong
Dept. of EE & Inst. of Comm. Engr.
National Taiwan Univ.
Taipei, Taiwan, R.O.C.

ABSTRACT

The DMT (discrete multitone modulation) technique has been widely applied to data transmission over fading channels of twisted pairs. It has been shown that the DMT system with ideal filters can achieve within 8 to 9 dB of the channel capacity of ADSL. The DFT based DMT system is proposed as a practical DMT implementation but its optimality is never asserted. In this paper we will show that the DFT based DMT systems are asymptotically optimal although they are not optimal for finite number of channels. The DFT based DMT system and the DMT system with ideal filters achieve the same bound. However, for a modest number of channels the optimal transceiver can provide substantial gain over the DFT based system as will be demonstrated by examples.

1. INTRODUCTION

Recently there has been great interest in applying the discrete multitone modulation (DMT) technique to high speed data transmission over fading channels such as ADSL and HDSL [1][2]. Fig. 1 shows an M-channel DMT system over a fading channel $C(z)$ with additive noise $e(n)$. The channel is divided into M subchannels using the transmitting filters $F_k(z)$ and receiving filters $H_k(z)$. The input is parsed and coded as modulation symbols, e.g., QAM (quadrature amplitude modulation). With judicious power and bit allocation, DMT can provide significant gain over fading channels. In [3], Kalet shows that the DMT system with ideal filters can achieve within 8 to 9 dB of the channel capacity of ADSL.

In the widely used DFT based DMT system, the transmitting and receiving filters are DFT filters. For a given probability of error and transmission power, bits can be allocated among the subchannels to achieve maximum total bit rate $R_{b,max}$. Very high speed data transmission can be achieved using 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.

THE WORK WAS SUPPORTED BY NSC 88-2218-E-009-016 AND BY NSC 88-2213-E-002-080, TAIWAN, R.O.C.

In the DMT system the 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 [4]. From the view point of multidimentional signal constellations it is shown that, for AWGN fading channels the optimal transmitting and receiving filters are eigen vectors associated with the channel. However in ADSL or HDSL applications, the channel noise is often the colored NEXT noise due to cross talk [3]. For fading channels with general colored noise source, the optimal transceiver is derived in [5]. The optimal transceiver decomposes the channel into eigen channels by incorporating the channel frequency response and the noise power spectrum.

The DFT based DMT system is proposed as a practical DMT implementation but its optimality is not asserted. In this paper we will show that the DFT based DMT systems are asymptotically optimal. The performance of the DFT based DMT systems becomes close to that of optimal DMT systems when the channel number M is sufficiently large. Furthermore the asymptotical performance of these two systems is the same as that of the DMT system with ideal filters in [3]. Although the DFT based DMT system is asymptotical optimal, the optimal transceiver provides significant gain over the DFT based system for a modest number of channels. An example with NEXT noise source will be given to demonstrate this.

2. ZERO ISI DMT TRANSCEIVER

Consider the system model of an M-channel DMT transceiver over a fading channel $C(z)$ with additive noise $e(n)$ in Fig. 1. Suppose the channel $C(z)$ is an FIR filter with order L, which is a reasonable assumption after channel equalization. In practice, to cancel ISI (inter-symbol interference) some degree of redundancy is introduced and the interpolation ratio $N > M$. Usually we have $N = M + L$. The length of the transmitting and receiving filters is also N.

Polyphase representation. The DMT system can be redrawn as in Fig. 2 using polyphase decomposition [2]. The transmitter $\mathbf{G}$ is an $N \times M$ constant matrix; the kth column of $\mathbf{G}$ contains the coefficients of the transmitting filter

0-7803-5471-0/99/$10.00©1999 IEEE

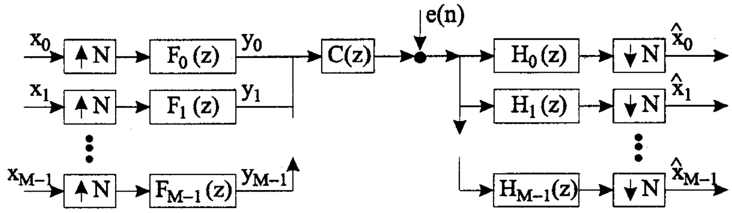

Figure 1: An M-channel DMT system over a fading channel.

$F_k(z)$. The receiver $\mathbf{S}$ is an $M \times N$ constant matrix; the kth row of $\mathbf{S}$ contains the coefficients of the receiving filter $H_k(z)$. The matrix $\mathbf{C}(z)$ is an $N \times N$ pseudo circulant matrix [6] with the first column given by

$$\begin{pmatrix} c_0 & c_1 & \cdots & c_L & 0 & \cdots & 0 \end{pmatrix}^T$$

where $\{c_n\}_{n=0}^{L}$ is the channel impulse response. The condition for zero ISI becomes $\mathbf{SC}(z)\mathbf{G} = \mathbf{I}$.

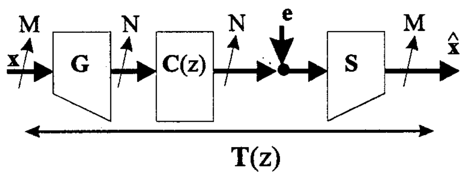

Figure 2: The polyphase representation of the DMT system.

DFT based DMT systems. In the DFT based DMT system (Fig. 3), the transmitting and receiving filters are DFT filters. Redundancy takes the form of cyclic prefix. Define $\mathbf{W}$ as the $M \times M$ DFT matrix with $[\mathbf{W}]_{mn} = 1/\sqrt{M}e^{-j2\pi mn/M}$, for $0 \leq m, n < M$. One can verify that DFT based DMT system has transmitter and receiver given by

$$\mathbf{G} = [\mathbf{W}\mathbf{W}_1]^\dagger, \quad \mathbf{S} = \mathbf{\Gamma}^{-1}[\mathbf{0}\ \mathbf{W}], \tag{1}$$

where $\mathbf{W}_1$ is a submatrix of $\mathbf{W}$ that contains the first L columns of $\mathbf{W}$ and $\mathbf{\Gamma}$ is the diagonal matrix $diag(C_0, C_1, \cdots, C_{M-1})$ with $\{C_k\}_{k=0}^{M-1}$ denoting the M point DFT of c_n.

Zero ISI DMT systems. Using singular value decomposition, we can decompose $\mathbf{C}_0$ as,

$$\mathbf{C}_0 = \underbrace{[\mathbf{U}_0\mathbf{U}_1]}_{\mathbf{U}} \begin{pmatrix} \mathbf{\Lambda} \\ \mathbf{0} \end{pmatrix}_{N \times M} \mathbf{V}^T = \mathbf{U}_0\mathbf{\Lambda}\mathbf{V}^T, \tag{2}$$

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

Consider the case that the transmitter is a unitary transformation followed by padding of L zeros, in particular

$$\mathbf{G} = \begin{pmatrix} \mathbf{G}_0 \\ \mathbf{0} \end{pmatrix}, \tag{3}$$

where $\mathbf{G}_0$ is an arbitrary $M \times M$ unitary matrix. For zero ISI, we can choose

$$\mathbf{S} = \mathbf{G}_0^T\mathbf{V}\mathbf{\Lambda}^{-1}\mathbf{U}_0^T. \tag{4}$$

When the transmitter is chosen as $\mathbf{G}_0 = \mathbf{V}^T$, the receiver is $\mathbf{S} = \mathbf{\Lambda}^{-1}\mathbf{U}_0^T$. This becomes the DMT system developed in [4].

3. TRANSMISSION POWER

For a given average bit rate R_b, the design of the transmitter and receiver affects the required transmission power. Let $\mathbf{R}_N$ be the $N \times N$ autocorrelation matrix of the channel noise process $e(n)$. The $M \times 1$ output noise vector of the receiver has autocorrelation function given by

$$\widehat{\mathbf{R}} = \mathbf{S}\mathbf{R}_N\mathbf{S}^T.$$

Let the number of bits allocated to the k-th channel be b_k, then the average bit rate is $R_b = \frac{1}{N}\sum_{k=0}^{M-1} b_k$. The actual bit rate is $\frac{1}{T}R_b$, where T is the sampling period of the system. Let $P(R_b, P_e, M)$ be the transmission power required for the M channel transceiver to achieve an average bit rate of R_b and probability of error P_e. With optimal bit allocation, the transmission power for the given transceiver is minimized and is equal to [5]

$$P(R_b, P_e, M) = c2^{2R_bN/M}(\Pi_{k=0}^{M-1}[\mathbf{S}\mathbf{R}_N\mathbf{S}^T]_{kk})^{1/M}, \tag{5}$$

where the constant c depends on the given probability of symbol error P_e and the modulation scheme.

In the DFT based DMT system, the receiver is $\mathbf{S} = \mathbf{\Gamma}^{-1}[\mathbf{0}\ \mathbf{W}]$ as given in (1). In this case we can verify that transmission power is

$$\begin{aligned} &P_{DFT}(R_b, P_e, M) \\ = \ & c2^{2R_bN/M}\frac{\left(\Pi_{k=0}^{M-1}[\mathbf{W}\mathbf{R}_M\mathbf{W}^\dagger]_{kk}\right)^{1/M}}{\det(\mathbf{\Gamma}^\dagger\mathbf{\Gamma})^{1/M}}. \end{aligned}$$

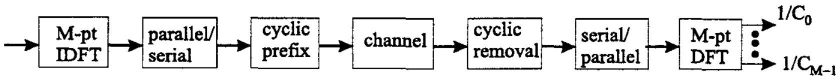

Figure 3: An M-channel DFT based DMT system.

From (5) we see that the transmission power can be further minimized by optimizing the transceiver. Using the optimal transceiver, the minimum transmission power is [5]

$$P_{opt}(R_b, P_e, M) = c2^{2R_b N/M} \frac{\left(\det(\mathbf{U}_0^T \mathbf{R}_N \mathbf{U}_0)\right)^{1/M}}{\det(\mathbf{\Lambda}^2)^{1/M}}.$$

4. ASYMPTOTICAL PERFORMANCE

In this section we will show that the DFT based DMT systems are asymptotically optimal although they are not optimal for finite number of channels. For a given error probability and bit rate, we will show that the power required in DFT based DMT system approaches that of the optimal system for large M. In particular,

$$\begin{aligned} & \lim_{M\to\infty} P_{opt}(R_b, P_e, M) \\ = & \lim_{M\to\infty} P_{DFT}(R_b, P_e, M) \\ = & c2^{2R_b} exp\left(\int_{-\pi}^{\pi} \ln \frac{S_{ee}(e^{j\omega})}{|C(e^{j\omega})|^2} \frac{d\omega}{2\pi}\right). \end{aligned} \quad (6)$$

Note that this is the same bound achieved by the DMT system with ideal filters as derived in [3]. The proof can be done in two steps.

Step 1: Using the distribution of eigenvalues for Toeplitz matrices [7], we are able to show that

$$\begin{aligned} & \lim_{M\to\infty} \det(\mathbf{\Lambda}^2)^{1/M} = \lim_{M\to\infty} \det(\mathbf{\Gamma}^\dagger\mathbf{\Gamma})^{1/M} \\ & = exp\left(\int_{-\pi}^{\pi} \ln |C(e^{j\omega})|^2 \frac{d\omega}{2\pi}\right), \end{aligned} \quad (7)$$

where $C(e^{j\omega})$ is the Fourier transform of c_n.

Step 2: Using properties of positive definite matrices, we can show that

$$\begin{aligned} & \lim_{M\to\infty} \left(\det(\mathbf{U}_0^T \mathbf{R}_N \mathbf{U}_0)\right)^{1/M} \\ = & exp\left(\int_{-\pi}^{\pi} \ln S_{ee}(e^{j\omega}) \frac{d\omega}{2\pi}\right). \end{aligned} \quad (8)$$

On the other hand, properties of Toeplitz matrices give us [8],

$$\begin{aligned} & \lim_{M\to\infty} \left(\Pi_{k=0}^{M-1} [\mathbf{W}\mathbf{R}_M\mathbf{W}^\dagger]_{kk}\right)^{1/M} \\ = & exp\left(\int_{-\pi}^{\pi} \ln S_{ee}(e^{j\omega}) \frac{d\omega}{2\pi}\right). \end{aligned} \quad (9)$$

With the equalities in (7)-(9), we can establish (6).

Proof of (7): Eq. (7) 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

$$||A|| = \max_{\mathbf{v}} \left(\frac{\mathbf{v}^\dagger \mathbf{A}^\dagger \mathbf{A} \mathbf{v}}{\mathbf{v}^\dagger \mathbf{v}}\right)^{1/2}, \ |\mathbf{A}| = (\frac{1}{n} trace(\mathbf{A}^\dagger \mathbf{A}))^{1/2}.$$

Two sequences of $n \times n$ matrices $\mathbf{A}_n$ and $\mathbf{B}_n$ are said to be asymptotically equivalent [7] if

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

Suppose $\mathbf{A}_n$ and $\mathbf{B}_n$ have eigenvalues $\alpha_{n,k}$ and $\beta_{n,k}$ and $M_0 \le \alpha_{n,k}, \beta_{n,k} \le M_1$. In [7], Gray shows that

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

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

To show (7), we observe that

$$\det(\mathbf{\Lambda}^2) = \Pi_{k=0}^{M-1} \lambda_k^2, \text{ and } \det(\mathbf{\Gamma}^\dagger \mathbf{\Gamma}) = \Pi_{k=0}^{M-1} |C_k|^2.$$

The values λ_k^2, for $k = 0, 1, \cdots, M-1$, 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 $|C_k|^2$. Define

$$\widehat{\mathbf{C}}_0 = \begin{pmatrix} \mathbf{I}_M & \begin{matrix} \mathbf{I}_L \\ \mathbf{0} \end{matrix} \end{pmatrix} \mathbf{C}_0.$$

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

$$\begin{pmatrix} c_0 & c_1 & \cdots & c_L & 0 & \cdots & 0 \end{pmatrix}^T.$$

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

$$\widehat{\mathbf{C}}_0 = \mathbf{W} \mathbf{\Gamma} \mathbf{W}^\dagger,$$

where $\mathbf{\Gamma}$ is the diagonal matrix in (1). Let $\mathbf{B}_M = \widehat{\mathbf{C}}_0^\dagger \widehat{\mathbf{C}}_0$, then 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, so

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

The second equality of (7) follows readily from the fact that C_k are the M points DFT of c_n. $\Delta\Delta$

Proof of (8): 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 (2). Let the eigenvalues of $\mathbf{P}$ be ordered as $\gamma_0 \le \gamma_1 \le \cdots \le \gamma_{N-1}$. Using the interlacing property of eigenvalues for positive definite matrices [9], 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, i.e.,

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

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,

$$S_{min} \le \gamma_0 \le \gamma_1 \le \cdots \le \gamma_{N-1} \le S_{max}.$$

It follows that

$$\det(\mathbf{U}_0^T\mathbf{R}_N\mathbf{U}_0) \le \Pi_{k=L}^{N-1}\gamma_k = \frac{\det \mathbf{P}}{\Pi_{k=0}^{L-1}\gamma_k} \le \frac{\det \mathbf{P}}{\gamma_0^L} \le \frac{\det \mathbf{P}}{S_{min}^L}$$
$$\det(\mathbf{U}_0^T\mathbf{R}_N\mathbf{U}_0) \ge \Pi_{k=0}^{M-1}\gamma_k = \frac{\det \mathbf{P}}{\Pi_{k=M}^{N-1}\gamma_k} \ge \frac{\det \mathbf{P}}{\gamma_{N-1}^L} \ge \frac{\det \mathbf{P}}{S_{max}^L}$$

Combining the above two equalities, we have

$$\frac{\det \mathbf{P}}{S_{max}^L} \le \det(\mathbf{U}_0^T\mathbf{R}_N\mathbf{U}_0) \le \frac{\det \mathbf{P}}{S_{min}^L}. \qquad (10)$$

Also 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 [10]

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

Letting M go to ∞ in (10), we arrive at (8). $\Delta\Delta$

Note that the DMT system developed in [4] 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 [4] is one regardless of the number of channels. On the other hand, the coding gain corresponding to the asymptotic bound in (6) is always greater than one if the channel noise is not white.

Example. Suppose the channel $C(z)$ is an FIR filter of order 1 and $C(z) = 1 + 0.5z^{-1}$. For the same probability of error and same bit rate, Fig. 4 shows $\frac{\mathcal{P}_{opt}(R_b,P_e,M)}{\mathcal{P}_{DFT}(R_b,P_e,M)}$, the ratio of power needed in optimal system over the power needed in the DFT-based system. We plot the ratio as a function of M for two different noise sources, the AWGN and NEXT noise source, which is colored channel noise due to cross talk [3].

From Fig. 4 we see that, for both noise sources the ratio $\frac{\mathcal{P}_{opt}(R_b,P_e,M)}{\mathcal{P}_{DFT}(R_b,P_e,M)}$ approaches unity as the channel number M increases. But for the NEXT noise channel, the ratio approaches unitary only for very large M. We can see that for a modest number of channel the optimal system provides substantial gain.

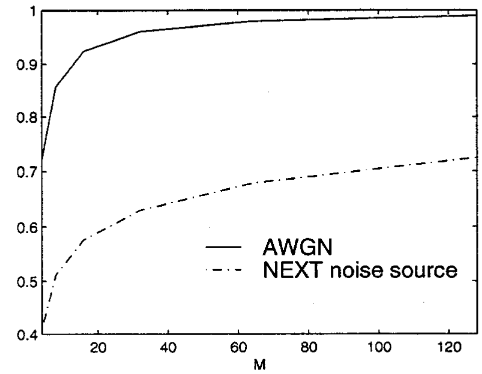

Figure 4: The ratio of the power needed in the optimal DMT system over the power needed in DFT based system for the same probability of error and the same bit rate.

5. 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., Aug. 1991.

[2] A. N. Akansu, et. al., "Orthogonal Transmultiplexers in Communication: A Review," *IEEE Trans. SP*, April 1998.

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

[4] S. Kasturia, J. T. Aslanis, and J. M. Cioffi, "Vector Coding for Partial Response Channels," *IEEE Trans. Inform. Theory*, July 1990.

[5] Yuan-Pei Lin and See-May Phoong, "Perfect Discrete Multitone Modulation with Optimal Transceivers," submitted to IEEE Trans. SP.

[6] P. P. Vaidyanathan, *Multirate Systems and Filter Banks*, Prentice-Hall, 1993.

[7] R. M. Gray, "On the Asymptotic Eigenvalue Distribution of Toeplitz Matrices," *IEEE Trans. Information Theory*, Nov. 1972.

[8] A. Gersho and R. M. Gray, *Vector Quantization and Signal Compression,* Kluwer Academic Publishers, Boston, 1991.

[9] R. A. Horn and C. R. Johnson, *Matrix Analysis*, Cambridge University Press, 1985.

[10] N. S. Jayant and P. Noll, *Digital Coding of Waveforms*, Prentice-Hall, 1984.