BIT ERROR RATE MINIMIZING CHANNEL SHORTENING EQUALIZERS FOR SINGLE CARRIER CYCLIC PREFIXED SYSTEMS

*Richard K. Martin**

The Air Force Inst. of Tech.
Dept. of ECE
WPAFB, OH, 45433
richard.martin
@afit.edu

Koen Vanbleu

Broadcom Corporation
Mechelen
Belgium
koen.vanbleu
@broadcom.com

Geert Ysebaert

Alcatel Bell
Antwerpen
Belgium
(email available
upon request)

ABSTRACT

Single carrier cyclic prefixed (SCCP) communications are a close relative of multicarrier communications. Both types of systems are robust to multipath, provided that the channel delay spread is shorter than the guard interval between transmitted blocks. If this condition is not met, a channel shortening equalizer can be used to shorten the channel to the desired length. Previous work on channel shortening has largely been in the context of digital subscriber lines, a wireline system that allows bit allocation, thus it has focused on maximizing the bit rate for a given bit error rate (BER). We propose and evaluate a channel shortener that attempts to directly minimize the BER for an SCCP system. The problem is shown to be analytically distinct from the analogous problem in multicarrier systems.

***Index Terms*—** equalization

1. INTRODUCTION

There are two types of cyclic prefixed systems: multicarrier modulation and single-carrier cyclic prefixed (SCCP) modulation, a.k.a. single-carrier frequency domain equalization (SC-FDE) [1], [2], [3]. Examples of wireline multicarrier systems include power line communications (HomePlug) and digital subscriber lines (DSL). Examples of wireless multicarrier systems include wireless local area networks (IEEE 802.11a/g, HIPERLAN/2, MMAC), wireless metropolitan area networks (IEEE 802.16), digital video/audio broadcasting in Europe, satellite radio (Sirius and XM Radio), and multiband ultra wideband (IEEE 802.15.3a). SCCP modulation has not been widely implemented, but it is gaining support in the literature.

Cyclic prefixed systems are robust to multipath, provided that the delay spread of the channel is less than the length of the cyclic prefix (CP) between transmitted blocks. If the channel is short, then channel equalization can be done tone-wise in the frequency domain by a bank of complex scalars. This is called a frequency-domain equalizer (FEQ). However, if the channel is longer than the CP, additional equalization is required, often in the form of a channel shortening equalizer (CSE) [a.k.a. a time-domain equalizer (TEQ)], which is a filter at the receiver front end. A survey of CSE design for DSL can be found in [4].

*Funded in part by the Air Force Office of Scientific Research. The views expressed in this paper are those of the authors, and do not reflect the official policy or position of the United States Air Force, Department of Defense, or the U.S. Government. This document has been approved for public release; distribution unlimited.

Channel shortening was first applied to maximum likelihood sequence estimation [5]. More recently, it has been used to shorten the long wireline impulse responses encountered by DSL [6], [7]. While early designs were based on heuristic cost functions, recent designs have adressed maximizing the bit rate for a given bit error rate (BER) [8], [9], which is appropriate in wireline multicarrier systems that allow bit allocation.

Wireless systems generally have a fixed bit allocation, and receiver performance is measured in terms of BER. Moreover, in wireline systems, once bit allocation has been done, the CSE can be used to minimize the BER of that bit allocation. Previously, the authors investigated BER minimizing CSE design for multicarrier systems [10]. However, the additional IFFT at the end of SCCP systems causes coupling that drastically changes this design problem compared to multicarrier systems. Hence, the main goals of this paper are to model the BER for SCCP systems, and to develop and assess a CSE that minimizes this BER.

2. SYSTEM MODEL

We assume a multiple-input multiple-output (MIMO) channel model with L transmit antennas and P receive antennas. Throughout, $(\cdot)^*$, $(\cdot)^T$, $(\cdot)^H$, and $\mathcal{E}\{\cdot\}$ denote complex conjugate, matrix transpose, conjugate transpose, and statistical expectation, respectively.

The system model is shown in Fig. 1. The complex finite-alphabet data symbols (usually multi-level QAM data) are blocked into groups of length N. The k^{th} such block for transmitter l is denoted $\widetilde{\mathbf{x}}_l(k)$. A cyclic prefix is inserted by copying the last ν samples of the block to the beginning of the block, and then the $S = N + \nu$ samples are transmitted serially. The i^{th} transmitted data sample is denoted $x_l(i)$. Note that l is the index of the transmit antenna, p is the index of the receive antenna, k is the block index, n is the tone index, i is the sample index, and $\jmath = \sqrt{-1}$ is the unit imaginary number.

The redundancy in the transmitted signal due to the CP can be represented by

$$x_l\left(Sk+i\right) = x_l\left(Sk+i+N\right), \qquad (1)$$
$$1 \le l \le L, \quad 1 \le i \le \nu, \quad -\infty < k < \infty.$$

Let $\mathbf{h}_{l,p}$ be an FIR filter of length L_h, which models the channel from transmit antenna l to receive antenna p, and let $\mathbf{w}_p$ be an FIR filter of length L_w, which is the CSE to be designed for antenna p. Let $\mathbf{H}_{l,p}$ be the tall channel convolution matrix for $\mathbf{h}_{l,p}$, which is an $L_c \times L_w$ Toeplitz matrix, where $L_c = L_h + L_w - 1$.

1-4244-0728-1/07/$20.00 ©2007 IEEE

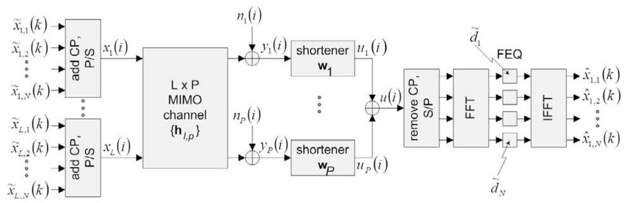

Fig. 1. Complex baseband SCCP system model.

Define the transmitted signal vectors

$$\mathbf{x}_l(i) = [x_l(i), \cdots, x_l(i - L_c + 1)]^T, \tag{2}$$

$$\mathbf{x}(i) = \left[\mathbf{x}_1^T(i), \cdots, \mathbf{x}_L^T(i)\right]^T, \tag{3}$$

and similarly for $\boldsymbol{\eta}_p(i)$, $\boldsymbol{\eta}(i)$, $\mathbf{y}_p(i)$, $\mathbf{y}(i)$. We can compactly write the CSE input vector as

$$\mathbf{y}(i) = \begin{bmatrix} \mathbf{H}_{1,1}^T, & \cdots, & \mathbf{H}_{L,1}^T \\ \vdots & \ddots & \vdots \\ \mathbf{H}_{1,P}^T, & \cdots, & \mathbf{H}_{L,P}^T \end{bmatrix} \mathbf{x}(i) + \boldsymbol{\eta}(i). \tag{4}$$

Passing the signal through the bank of CSEs yields

$$u(i) = \mathbf{w}^T \mathbf{y}(i). \tag{5}$$

After channel shortening, the cyclic prefix is discarded and a discrete Fourier transform (DFT), implemented by the fast Fourier transform (FFT), is used to convert the data to the frequency domain. We use $\mathcal{F}$ to denote the unitary DFT matrix, with element (m, n) given by $(1/\sqrt{N})e^{-\jmath 2\pi mn/N}$. The DFT requires an estimate of the transmission delay Δ, since the length N DFT input vector for block k is

$$\mathbf{u}(k) = [u(Sk + \nu + 1 + \Delta), \cdots, u(S(k+1) + \Delta)]^T. \tag{6}$$

The delay Δ is a design parameter whose choice affects the values of the optimal CSE as well as the performance that can be attained. To invert the channel in the frequency domain, the DFT is computed, the FEQ $\widetilde{\mathbf{d}}$ is applied, and an inverse DFT (IDFT) is computed,

$$\widetilde{\mathbf{u}}(k) = \mathcal{F}\, \mathbf{u}(k), \tag{7}$$

$$\widehat{\mathbf{u}}(k) = \widetilde{\mathbf{d}} \odot \widetilde{\mathbf{u}}(k), \tag{8}$$

$$\widehat{\mathbf{x}}_1(k) = \mathcal{F}^H \widehat{\mathbf{u}}(k), \tag{9}$$

where $\odot$ denotes element-by-element multiplication, and we assume that the receiver is attempting to recover the data from transmitter $l = 1$. In a multiuser scheme, the data for $l = 2, \cdots, L$ can be ignored, or a multi-user detection technique can be used to mitigate the interference. In a single user scheme, the data $\widetilde{x}_l$ may be the same on all transmitters l; or an Alamouti transmit diversity space-time code may be used [11]. Interleaving and forward error correction blocks can be included, although for conciseness, they are not depicted here.

3. BER MODEL

The goal of this section is to model the BER of an SCCP system, which we will attempt to optimize in Section 4.

The BER will be averaged over the N elements of the final IFFT output, $\widehat{\mathbf{x}}_1(k)$. We assume that the total residual interference and noise on each output sample is Gaussian, and that M-level QAM signalling is used on each tone. The probability of error on the PAM component of sample m is given by [12, pp. 225–226]

$$P_{\sqrt{M}}(m) = 2\left(1 - \frac{1}{\sqrt{M}}\right) Q\left(\sqrt{\frac{3}{M-1}\mathrm{SNR}_n}\right), \tag{10}$$

hence the SER of sample m is

$$P_M(m) = 2P_{\sqrt{M}}(m) - \left(P_{\sqrt{M}}(m)\right)^2, \tag{11}$$

where SNR_m is the effective signal-to-interference-and-noise ratio on sample m (which we will refer to as the output SNR); and $Q(x)$ is the Q-function, which is the integral of a unit Gaussian PDF from x to infinity. For the $M = 4$ case, (10) is the BER for sample m, and it reduces to $Q\left(\sqrt{\mathrm{SNR}_m}\right)$, which we use here for simplicity of notation. Averaging (10) and (11) over the N output samples, the BER and the SER for the output of an SCCP system with $M = 4$ are

$$BER_{sccp} = \frac{1}{N}\sum_{m=1}^{N} Q\left(\sqrt{\mathrm{SNR}_m}\right), \tag{12}$$

$$SER_{sccp} = \frac{1}{N}\sum_{m=1}^{N} \left(2Q\left(\sqrt{\mathrm{SNR}_m}\right) - Q^2\left(\sqrt{\mathrm{SNR}_m}\right)\right). \tag{13}$$

Either can be used as the objective function, although we focus on the BER. At this point, we need a model for the output SNR.

The DFT output can be written as

$$\mathcal{F}\underbrace{(\mathbf{Y}_k \mathbf{w})}_{\mathbf{u}(k)} = \underbrace{(\mathcal{F}\mathbf{Y}_k)}_{\widetilde{\mathbf{Y}}_{k,N}} \mathbf{w}, \tag{14}$$

where $\mathbf{Y}_k$ is a block Toeplitz matrix of size $N \times (PL_w)$, where each $N \times L_w$ sub-block contains the data that will be convolved with the p^{th} CSE, $\mathbf{w}_p$, and successive rows are vectors $\mathbf{y}^T(i)$ for successive values of i. Then $\widetilde{\mathbf{Y}}_{k,N}$ is an $N \times (PL_w)$ matrix as well, with n^{th} the row denoted by $\widetilde{\mathbf{y}}_{k,n}$. Define $\mathbf{Q}_{m,n}$ to be element (m, n) of the (unitary) IDFT matrix, with $\mathbf{Q} = \mathcal{F}^H$. Passing (14) through the FEQ and the IDFT, and taking output sample m for user $l = 1$ yields

$$\widehat{\mathbf{x}}_1(k)[m] = \sum_{n=1}^{N} \mathbf{Q}_{m,n} \widetilde{d}_n \widetilde{\mathbf{y}}_{k,n} \mathbf{w}. \tag{15}$$

Define the correlation terms

$$\sigma^2 \triangleq \mathcal{E}\left\{|\widetilde{\mathbf{x}}_1^*(k)[m]|^2\right\}, \tag{16}$$

$$\boldsymbol{\varphi}_{m,n} \triangleq \mathcal{E}\left\{\widetilde{\mathbf{x}}_1^*(k)[m]\widetilde{\mathbf{y}}_{k,n}\right\}, \tag{17}$$

$$\boldsymbol{\Sigma}_{m,n}^2 \triangleq \mathcal{E}\left\{\widetilde{\mathbf{y}}_{k,m}^H \widetilde{\mathbf{y}}_{k,n}\right\}, \tag{18}$$

which have dimensions 1×1, $1 \times PL_w$, and $PL_w \times PL_w$, respectively. For simplicity, we assume that σ^2 is independent of m. The output SNR of sample m is the ratio of the power of the desired signal, $\widetilde{\mathbf{x}}_1(k)[m]$, to the power of the total interference and noise,

$$\mathrm{SNR}_m = \frac{\sigma^2}{\mathcal{E}\left\{|\widetilde{e}_{k,m}|^2\right\}} = \frac{\sigma^2}{\mathcal{E}\left\{|\widehat{\mathbf{x}}_1(k)[m] - \widetilde{\mathbf{x}}_1(k)[m]|^2\right\}}. \tag{19}$$

Using (15), the denominator expands as

$$\mathcal{E}\left\{|\widetilde{e}_{k,m}|^2\right\} = \sum_{n_1,n_2} \mathbf{Q}_{m,n_1}^* \mathbf{Q}_{m,n_2} \widetilde{d}_{n_1}^* \widetilde{d}_{n_2} \mathbf{w}^H \boldsymbol{\Sigma}_{n_1,n_2}^2 \mathbf{w}$$
$$- \sum_n \mathbf{Q}_{m,n} \widetilde{d}_n \boldsymbol{\varphi}_{m,n} \mathbf{w} - \sum_n \mathbf{Q}_{m,n}^* \widetilde{d}_n^* \boldsymbol{\varphi}_{m,n}^* \mathbf{w}^* + \sigma^2. \tag{20}$$

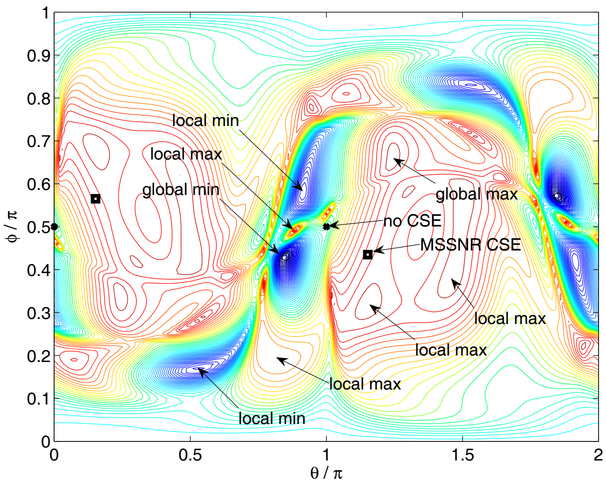

Fig. 2. Contours of the BER for a 3-tap CSE under a unit norm constraint. The CSE is parameterized in spherical coordinates by the angles θ and ϕ. The cost function is symmetric with respect to a sign change in the CSE ($\mathbf{w} \rightarrow -\mathbf{w}$), i.e. $(\theta, \phi) \rightarrow (\theta + \pi, \pi - \phi)$.

The unbiased MMSE FEQ for sample m is found by setting the correlation of the input and output for sample m equal to the transmitted power for sample m, or equivalently

$$\mathcal{E}\left\{\widetilde{\mathbf{x}}_1^*(k)[m] \sum_{n=1}^{N} \mathbf{Q}_{m,n}\widetilde{d}_n \widetilde{\mathbf{y}}_{k,n}\mathbf{w}\right\} = \mathcal{E}\left\{\widetilde{\mathbf{x}}_1^*(k)[m]\widetilde{\mathbf{x}}_1(k)[m]\right\},$$

$$\sum_{n=1}^{N} \mathbf{Q}_{m,n}\widetilde{d}_n \boldsymbol{\varphi}_{m,n}\mathbf{w} = \sigma^2. \tag{21}$$

With this value of the FEQ, the output SNR becomes

$$\text{SNR}_m = \frac{\sigma^2}{\sum_{n_1,n_2} \mathbf{Q}^*_{m,n_1}\mathbf{Q}_{m,n_2}\widetilde{d}^*_{n_1}\widetilde{d}_{n_2}\mathbf{w}^H \boldsymbol{\Sigma}^2_{n_1,n_2}\mathbf{w} - \sigma^2}. \tag{22}$$

Eq. (21) can be rewritten in matrix form as

$$\widetilde{\mathbf{d}}^T \mathbf{A}_m \mathbf{w} = \sigma^2, \quad m \in \{1, \cdots, N\}, \tag{23}$$

where row $\mathbf{A}_m[n,:] = \mathbf{Q}_{m,n}\boldsymbol{\varphi}_{m,n}$. Collecting these N equations into a vector and solving for $\widetilde{\mathbf{d}}$ yields

$$\widetilde{\mathbf{d}} = \sigma^2 \begin{bmatrix} \mathbf{w}^T\mathbf{A}_1^T \\ \vdots \\ \mathbf{w}^T\mathbf{A}_N^T \end{bmatrix}^{-1} \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}. \tag{24}$$

Together, (12), (22), and (23) allow us to evaluate the BER for a given CSE $\mathbf{w}$ and unbiased MMSE FEQ $\widetilde{\mathbf{d}}$ based on that CSE.

We can visualize the BER and SNR model by using a 3-tap unit norm CSE parameterized by the two angles of spherical coordinates. The channel is $\mathbf{h} = [1, -0.3, 0.7]$ with a desired length of $\nu+1 = 2$, FFT size $N = 8$, and 20 dB SNR. For each value of the CSE, the unbiased MMSE FEQ is given by (23). Fig. 2 shows log-spaced contours of the BER. The BER is extremely multimodal, and numerical optimization of this cost surface is an ambitious goal.

Estimate the correlation terms in (16), (17), and (18) from time averages. Choose a neighborhood size σ_{step} and a unit norm initial guess $\mathbf{w}_{best}$, then loop through:

1. Generate $\mathbf{w}_{step}$ as a circularly Gaussian random vector, i.i.d. with zero mean and variance σ^2_{step}
2. $\mathbf{w}_{trial} = \frac{\mathbf{w}_{best}+\mathbf{w}_{step}}{\|\mathbf{w}_{best}+\mathbf{w}_{step}\|}$
3. Use (12), (22), and (23) to compute BER_{trial}
4. If $BER_{trial} < BER_{best}$, then $\mathbf{w}_{best} = \mathbf{w}_{trial}$ and $BER_{best} = BER_{trial}$

Stop at a given number of iterations or target BER.

Fig. 3. Greedy minimum error rate (G-MER) CSE design algorithm.

4. BER MINIMIZATION PROCEDURE

Since the BER is intractable to direct minimization, a heuristic approach is needed. The approach we propose is a greedy search. At each iteration, we search in a neighborhood of the current best solution, and if the new CSE has a lower BER, we accept it. The algorithm is summarized in Fig. 3. Note that evaluation of the analytic BER model is required in order to evaluate each tentative update.

The BER model is invariant to scaling the CSE, hence the unit norm constraint is used for convenience. If the step size is $\sigma_{step} = \alpha L_w^{-1/2}$, then the average update size is $\left(\mathcal{E}\left\{\mathbf{w}^H_{step}\mathbf{w}_{step}\right\}\right)^{1/2} = \alpha$. For our simulations, we use $\alpha = 0.01$, i.e. each step has a magnitude of about 1% of the magnitude of the current CSE. The initialization for $\mathbf{w}_{best}$ should be a cheap-to-compute CSE that has the best performance of all such designs, so that there is a reasonable chance of starting in the valley of the global minimum of the BER.

There are two drawbacks to the greedy search. First, it requires computation of the BER model at each iteration, which is very expensive. If N_{corr} symbols are used to compute the correlation terms, then the greedy search requires $\frac{1}{2}(PL_w)^2 N^2 N_{corr}$ complex multiply-and-accumulate (MAC) operations to compute the correlation terms, and a further $(PL_w)^2 N^3$ complex MACs per iteration. Second, as with a gradient descent method, the global minimum is only achieved if the initialization lies somewhere in the valley of the global minimum. In order to not become trapped in a local minimum, the greedy search can be generalized to simulated annealing (SA) [14]. SA occasionally allows upwards steps, but the probability of allowing them decreases according to a user-defined "cooling schedule." Under certain conditions (including infinite run time and an infinitesimally slow cooling schedule), simulated annealing will find the global minimum of the cost function. However, this further adds to the complexity, since a large number of iterations is required.

5. SIMULATIONS

The algorithms to be compared are the MMSNR design [7], the MMSE design [5], the MDS design [15], the Min-IBI design [13], and the greedy search (G-MER). We also compare to the BER when no CSE is used. The FFT size is $N = 16$ and the CP length is $\nu = 4$, which are small since the output SNR model of (22) is so expensive to compute. The signal constellation is 4-QAM. The channels are Rayleigh fading with 10 significant taps and an approximately exponential delay profile as in [16]. There are $L = 1$ transmit antenna and $P = 2$ receive antennas. The CSE has $L_w = 16$ taps per an-

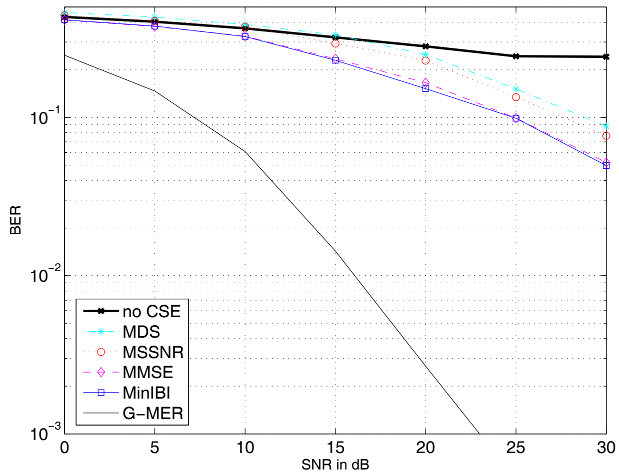

Fig. 4. BER vs. SNR for various CSE design algorithms.

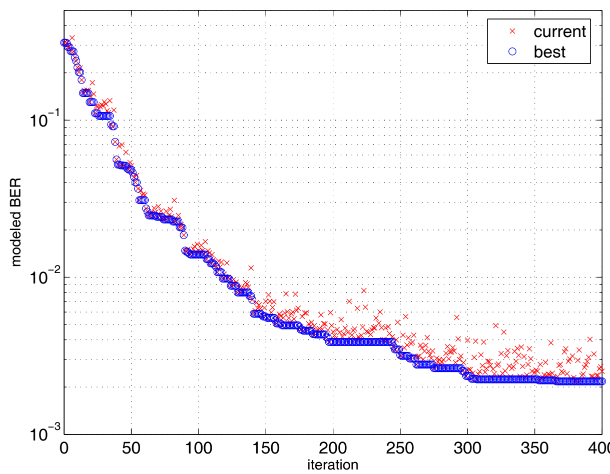

Fig. 5. History of the BER model for one run of the G-MER algorithm. The SNR was 15 dB.

tenna. The correlation parameters are estimated using 100 blocks of training, and the Min-IBI design is used as the initialization. The BER values will be measured over 200 independent trials (channel, input data, and noise), using 2000 blocks of data each. The desired delay will be chosen heuristically (similar to [16]) rather than performing a global search.

Fig. 4 shows the measured BER versus SNR in dB, for this SCCP system; and Fig. 5 shows the calculated BER versus iteration number, for the greedy search. Note that the BER model of (12), (22), and (23) is only used to perform the greedy search, and the actual BER assessment in Fig. 4 uses the actual measured BER, not the model. Aside from the greedy search, the channel shorteners considered (MDS, MSSNR, MMSE, and Min-IBI) are the only ones that the authors are aware of that do not explicitly take into account the multicarrier signal structure, hence they are the only ones that can be directly applied to the SCCP system for comparison. The design that performs the best by far is the greedy search (G-MER). However, it is too computationally intensive for real-time implementation; of the remaining designs, the Min-IBI design performs the best. Clearly, there is a need for a new design that is computationally cheaper than the greedy search but performs better than the Min-IBI design.

6. REFERENCES

[1] D. D. Falconer, S. L. Ariyavisitakul, A. Benyamin-Seeyar, and B. Eidson, "Frequency Domain Equalization for Single-carrier Broadband Wireless Systems," *IEEE Commun. Mag.*, vol. 40, pp. 58–66, Apr. 2002.

[2] H. Sari, G. Karam, and I. Jeanclaude, "Frequency-Domain Equalization of Mobile Radio and Terrestrial Broadcast Channels," in *Proc. IEEE Global Comm. Conf. (GLOBECOM) 94*, San Fransisco, CA, Nov. 1994, pp. 1–5.

[3] H. Sari, G. Karam, and I. Jeanclaude, "Transmission Techniques for Digital Terrestrial TV Broadcasting," *IEEE Commun. Mag.*, pp. 100–109, Feb. 1995.

[4] R. K. Martin, K. Vanbleu, M. Ding, G. Ysebaert, M. Milosevic, B. L. Evans, M. Moonen, and C. R. Johnson, Jr., "Unification and Evaluation of Equalization Structures and Design Algorithms for Discrete Multitone Modulation Systems," *IEEE Trans. Signal Proc.*, vol. 53, no. 10, pp. 3880–3894, Oct. 2005.

[5] D. D. Falconer and F. R. Magee, "Adaptive Channel Memory Truncation for Maximum Likelihood Sequence Estimation," *Bell Sys. Tech. Journal*, pp. 1541–1562, Nov. 1973.

[6] N. Al-Dhahir and J. M. Cioffi, "Optimum Finite-Length Equalization for Multicarrier Transceivers," *IEEE Trans. on Comm.*, vol. 44, no. 1, pp. 56–64, Jan. 1996.

[7] P. J. W. Melsa, R. C. Younce, and C. E. Rohrs, "Impulse Response Shortening for Discrete Multitone Transceivers," *IEEE Trans. on Comm.*, vol. 44, pp. 1662–1672, Dec. 1996.

[8] G. Arslan, B. L. Evans, and S. Kiaei, "Equalization for Discrete Multitone Receivers To Maximize Bit Rate," *IEEE Trans. Signal Proc.*, vol. 49, no. 12, pp. 3123–3135, Dec. 2001.

[9] K. Vanbleu, G. Ysebaert, G. Cuypers, M. Moonen, and K. Van Acker, "Bitrate Maximizing Time-Domain Equalizer Design for DMT-based Systems," *IEEE Trans. on Comm.*, vol. 52, no. 6, pp. 871–876, June 2004.

[10] R. K. Martin, G. Ysebaert, K. Vanbleu, and A. G. Klein, "Bit Error Rate Minimizing Channel Shortening Equalizers for Multicarrier Systems," in *Proc. IEEE Workshop on Sig. Proc. Advances in Wireless Comm.*, Cannes, France, July 2005.

[11] S. M. Alamouti, "A Simple Transmit Diversity Technique for Wireless Communications," *IEEE Journal on Selected Areas in Comm.*, vol. 16, pp. 1451–1458, Oct. 1998.

[12] G. L. Stuber, *Principles of Mobile Communication*, Kluwer Academic Publishers, Boston, 1996.

[13] S. Celebi, "Interblock Interference (IBI) Minimizing Time-Domain Equalizer (TEQ) for OFDM," *IEEE Signal Processing Letters*, vol. 10, no. 8, pp. 232–234, Aug. 2003.

[14] C. R. Reeves, *Modern Heuristic Techniques for Combinatorial Problems*, MMcGraw-Hill Book Company, London, 1995.

[15] R. Schur and J. Speidel, "An Efficient Equalization Method to Minimize Delay Spread in OFDM/DMT Systems," in *Proc. IEEE Int. Conf. on Comm.*, Helsinki, Finland, June 2001, vol. 5, pp. 1481–1485.

[16] R. K. Martin, J. M. Walsh, and C. R. Johnson, Jr., "Low Complexity MIMO Blind, Adaptive Channel Shortening," *IEEE Trans. Signal Proc.*, vol. 53, no. 4, pp. 1324–1334, Apr. 2005.