Blind, Adaptive Channel Shortening by Sum-squared Auto-correlation Minimization (SAM)

J. Balakrishnan,[*] R. K. Martin, and C. R. Johnson, Jr.[†]
School of Electrical and Computer Engineering
Cornell University
Ithaca, NY 14853
jai@ti.com, {frodo,johnson}@ece.cornell.edu

Abstract

We propose a new blind, adaptive channel shortening algorithm for updating a time-domain equalizer (TEQ) in a system employing multicarrier modulation. The technique attempts to minimize the sum-squared auto-correlation of the combined channel-TEQ impulse response outside a window of desired length. The proposed algorithm, "Sum-squared Auto-correlation Minimization" (SAM), assumes the source sequence to be white and wide-sense stationary, and it is implemented as a stochastic gradient descent algorithm. Simulation results demonstrating the success of the SAM algorithm are provided.

1 Introduction

Multicarrier modulation (MCM) has been gaining in popularity over recent years. One reason for this is the ease with which MCM can combat channel dispersion, provided the channel delay spread is not greater than the length of the cyclic prefix (CP). However, if the CP is not long enough, the orthogonality of the sub-carriers is lost and this causes both inter-carrier interference (ICI) and inter-symbol interference (ISI).

A well known technique to deal with an inadequate CP length is the use of a time-domain equalizer (TEQ) in the receiver. The TEQ is a filter that shortens the effective channel to the length of the CP. The TEQ design problem has been extensively studied in the literature. Falconer and Magee [1] proposed a minimum MSE method for channel shortening, which was designed for the maximum likelihood sequence estimation problem. Melsa, Younce, and Rohrs [2] proposed the maximum shortening SNR method, which attempts to maximize the ratio of energy inside a window of the channel to energy outside the window. In DSL, the true performance metric to optimize is the maximum bit allocation that can be acheived for a fixed error probability [3], [4]. Optimizing the MSE or SSNR does not necessarily optimize the bit rate.

The above techniques are non-adaptive (except [1]), and all require training. The MMSE solution [1] can be implemented adaptively, but it converges very slowly [5]. Chow's algorithm converges more quickly, but it usually converges to a suboptimal setting [5]. Lashkarian and Kiaei [6] proposed an iterative implementation of the maximum bit rate method of [3], but as cited in [4], the method in [3] makes some inaccurate assumptions and is not optimal. Furthermore, the method in [6] is iterative rather than adaptive, since it assumes knowledge of large matrices that depend upon the channel and are not adaptively updated.

De Courville, *et al.* have proposed a blind, adaptive TEQ [7] that relies on the presence of unused subcarriers within the transmission bandwidth. However, it shortens the channel to a single spike rather than to a window. This is an overly stringent criterion, so its performance is expected to be suboptimal. Martin, *et al.* [8], [9] have proposed a low-complexity adaptive TEQ algorithm known as MERRY, but it only updates once per symbol. The SAM algorithm proposed in this paper is blind and adaptive, and its performance is competitive with MERRY. SAM has higher complexity, but it converges much faster and does not require an estimate of the symbol placement within the data stream.

This paper is organized as follows. Section 2 presents the system model. Sections 3 and 4 discuss the SAM cost function and gradient descent algorithm. Section 5 studies properties of the cost function. Section 6 provides simulations, and Section 7 concludes.

---

[*]Jaiganesh Balakrishnan was with Cornell University when this work was performed. He is now with Texas Instruments, Dallas, TX 75243

[†]This work was supported in part by NSF grant ECS-9811297 and NxtWave Communications (now ATI).

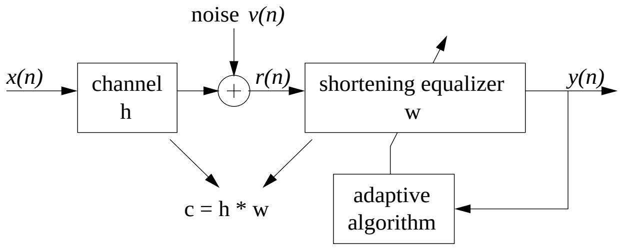

Figure 1: System model for an adaptive TEQ.

2 System Model

The system model is shown in Fig. 1. The received sequence $r(n)$ is generated by passing the transmitted data $x(n)$ through a length L_h+1 channel $\mathbf{h}$ and adding samples of the noise $v(n)$,

$$r(n) = \sum_{k=0}^{L_h} h(k)x(n-k) + v(n). \tag{1}$$

The equalized data $y(k)$ is obtained by filtering the received data with a length L_w+1 TEQ $\mathbf{w}$,

$$y(n) = \sum_{k=0}^{L_w} w(k)r(n-k) = \mathbf{w}^T\mathbf{r}_n, \tag{2}$$

where $\mathbf{r}_n = [r(n)\ r(n-1)\ \cdots\ r(n-L_w)]^T$. Let $\mathbf{c} = \mathbf{h} \star \mathbf{w}$ denote the effective channel of length L_c+1. We make the following assumptions.

1. The source sequence $x(n)$ is white, zero-mean and wide-sense stationary (W.S.S).
2. The relation $2L_c < N_{fft}$ holds, i.e. the effective channel has length less than half the FFT size[1].
3. The source sequence $x(n)$ is real with unit variance.
4. The noise sequence $v(n)$ is zero-mean, i.i.d., uncorrelated to the source sequence and has a variance σ_v^2.

The first assumption is critical for the proposed channel shortening algorithm. Assumption two is important for analytical reasons, but if it is modestly violated the performance degradation should be minor. The last two assumptions are for notational simplicity.

[1]Multicarrier systems employ a block structure. Modulation is performed by an FFT of the same size as the block.

3 Sum-squared Auto-correlation Minimization

This section motivates the use of the SAM cost function, and shows how to blindly measure it from the data. Consider the auto-correlation sequence of the impulse response of the effective channel,

$$R_{cc}(l) = \sum_{k=0}^{L_c} c(k)c(k-l). \tag{3}$$

For the effective channel $\mathbf{c}$ to be zero outside a window of size $\nu+1$, it is necessary for the auto-correlation values $R_{cc}(l)$ to satisfy

$$R_{cc}(l) = 0, \qquad \forall\ |l| > \nu. \tag{4}$$

Hence, one possible way of performing channel shortening is by ensuring that (4) is satisfied by the auto-correlation function of the effective channel. The trivial solution $\mathbf{c} = \mathbf{0}$ (or equivalently $\mathbf{w} = \mathbf{0}$) can be avoided by imposing a constraint, e.g. $\|\mathbf{c}\|_2^2 = 1$ or equivalently $R_{cc}(0) = 1$.

We define a cost function $J_{\nu+1}$ in an attempt to minimize the sum-squared auto-correlation terms, i.e.,

$$J_{\nu+1} = \sum_{l=\nu+1}^{L_c} |R_{cc}(l)|^2. \tag{5}$$

The TEQ optimization problem can then be stated as

$$\mathbf{w}^{opt} = \arg_{\mathbf{w}} \min_{\|\mathbf{h}\|_2^2=1} J_{\nu+1}. \tag{6}$$

The auto-correlation function of the sequence $y(n)$ can be written as

$$\begin{aligned} R_{yy}(l) &= \mathrm{E}[y(n)y(n-l)] \\ &= \mathrm{E}\left[\left(\mathbf{c}^T\mathbf{x}_n + \mathbf{w}^T\mathbf{v}_n\right)\left(\mathbf{x}_{n-l}^T\mathbf{c} + \mathbf{v}_{n-l}^T\mathbf{w}\right)\right] \\ &= \mathbf{c}^T\mathrm{E}\left[\mathbf{x}_n\mathbf{x}_{n-l}^T\right]\mathbf{c} + \mathbf{w}^T\mathrm{E}\left[\mathbf{v}_n\mathbf{v}_{n-l}^T\right]\mathbf{w} \end{aligned} \tag{7}$$

where $\mathbf{x}_n = [x(n)\ x(n-1)\ \cdots\ x(n-L_h-L_w)]^T$, and $\mathbf{v}_n = [v(n)\ v(n-1)\ \cdots\ v(n-L_w)]^T$. Since $v(n)$ is i.i.d., $\mathrm{E}\left[\mathbf{v}_n\mathbf{v}_{n-l}^T\right]$ will be Toeplitz, with only one diagonal of nonzero entries. It becomes a shifting matrix, i.e. its affect on a vector is to shift the elements of the vector up or down (depending on l). $\mathrm{E}\left[\mathbf{x}_n\mathbf{x}_{n-l}^T\right]$ becomes another shifting matrix, provided that the assumption $2L_c < N_{fft}$ holds. If this is violated, then the matrix is still Toeplitz, but for some values of l there will be another diagonal of nonzero entries, corresponding to the correlation between samples in the transmitted symbol end and samples in the transmitted cyclic prefix. Fortunately, assumption 2 is a reasonable one, as can be seen by considering the CSA test loop channels [10] for the case of DSL: $L_h \cong 200$, $L_w \cong 32$, and $N_{fft} = 512$, so $2\,(200+32) < 512$.

Now (7) can be simplified to

$$\begin{aligned} R_{yy}(l) &= \sum_{k=0}^{L_c} c(k)c(k-l) + \sigma_v^2 \sum_{k=0}^{L_w} w(k)w(k-l) \\ &= R_{cc}(l) + \sigma_v^2\, R_{ww}(l). \end{aligned} \tag{8}$$

Thus, we may approximate the cost function of (5) by

$$\begin{aligned} \hat{J}_{\nu+1} &= \sum_{l=\nu+1}^{L_c} |\mathrm{E}[y(n)y(n-l)]|^2 \\ &= \sum_{l=\nu+1}^{L_c} |R_{cc}(l)|^2 + 2\sigma_v^2 \sum_{l=\nu+1}^{L_w} R_{cc}(l)R_{ww}(l) \\ &\quad + \sigma_v^4 \sum_{l=\nu+1}^{L_w} |R_{ww}(l)|^2 . \end{aligned} \tag{9}$$

If the TEQ length L_w+1 is shorter than the CP length ν (as in [2], [4]), both noise terms in (9) vanish due to the empty summations. If L_w is significantly longer than ν, for typical SNR values σ_v^4 will be very small, so we can neglect the last term in (9). Furthermore, the summands in the second term will be both positive and negative, so they will often add to a small value. Combining this with the fact that the second summation is multiplied by the (small) noise variance, we are justified in ignoring the second term in (9) as well. This leaves us with $\hat{J}_{\nu+1} \cong J_{\nu+1}$ (and $\hat{J}_{\nu+1} = J_{\nu+1}$ exactly if $L_w < \nu+1$). Accordingly, we will henceforth drop the hat on $J_{\nu+1}$.

4 Adaptive Algorithm

The steepest gradient-descent algorithm over the cost surface $J_{\nu+1}$ is

$$\mathbf{w}^{new} = \mathbf{w}^{old} - \mu \nabla_{\mathbf{w}} \left(\sum_{l=\nu+1}^{L_c} \mathrm{E}[y(n)y(n-l)]^2 \right), \tag{10}$$

where μ denotes the step size and $\nabla_{\mathbf{w}}$ denotes the gradient with respect to $\mathbf{w}$. To implement this algorithm, an instantaneous cost function is defined, where the expectation operation is replaced by a moving average over a user-defined window of length N.

$$J_{\nu+1}^{inst}(k) = \sum_{l=\nu+1}^{L_c} \left\{ \sum_{n=kN}^{(k+1)N-1} \frac{y(n)y(n-l)}{N} \right\}^2 . \tag{11}$$

The value of N is a design parameter. It should be large enough to give a reliable estimate of the expectation, but no larger, as the algorithm complexity is proportional to N. The "stochastic" gradient-descent algorithm is then given by

$$\begin{aligned} \mathbf{w}(k+1) = \mathbf{w}(k) - \mu \sum_{l=\nu+1}^{L_c} \Bigg[\left\{ \sum_{n=kN}^{(k+1)N-1} \frac{y(n)y(n-l)}{N} \right\} \\ \cdot \left\{ \sum_{n=kN}^{(k+1)N-1} \left(\frac{y(n)\mathbf{r}_{n-l} + y(n-l)\mathbf{r}_n}{N} \right) \right\} \Bigg] \end{aligned} \tag{12}$$

The blind TEQ update algorithm described in (12) will be referred to as the *Sum-squared Auto-correlation Minimization* (SAM) algorithm.

An alternate implementation comes from using auto-regressive (AR) estimates instead of moving average (MA) estimates. Let

$$\mathbf{q}^1 = (1-\alpha)\mathbf{q}^1 + \alpha\, y(n) \begin{bmatrix} r(n-\nu-1) \\ \vdots \\ r(n-L_c-L_w) \end{bmatrix}$$

$$\mathbf{q}^2 = \mathbf{W}\mathbf{q}^1$$

$$\mathbf{Q} = (1-\alpha)\mathbf{Q} + \alpha \begin{bmatrix} r(n) \\ \vdots \\ r(n-L_w) \end{bmatrix} \begin{bmatrix} y(n-\nu-1) \\ \vdots \\ y(n-L_c) \end{bmatrix}^T$$

where $0 < \alpha < 1$ is a design parameter and $\mathbf{W}$ is the $(L_c-\nu) \times (L_c+L_w-\nu)$ convolution matrix of the equalizer. Then the update rule can be written as

$$\begin{aligned} \mathbf{w}(n+1) &= \mathbf{w}(n) - \mu \sum_{l=\nu+1}^{L_c} \{\mathrm{E}\,[y(n)y(n-l)]\} \cdot \\ &\qquad \{\mathrm{E}\,[y(n)\mathbf{r}_{n-l} + y(n-l)\mathbf{r}_n]\} \\ &\cong \mathbf{w}(n) - \mu \sum_{l=\nu+1}^{L_c} \left\{\mathbf{q}^2_{l-\nu}\right\} \cdot \\ &\quad \left\{ \begin{bmatrix} \mathbf{q}^1_{l-\nu} \\ \vdots \\ \mathbf{q}^1_{l-\nu+L_w} \end{bmatrix} + \begin{bmatrix} \mathbf{Q}_{1,l-\nu} \\ \vdots \\ \mathbf{Q}_{L_w+1,l-\nu} \end{bmatrix} \right\} \end{aligned} \tag{13}$$

With both implementations, $\mathbf{w}$ must be periodically renormalized to enforce the constraint $\|\mathbf{c}\|_2^2 = 1$. (The constraint may also be implemented by adding a penalty term onto the cost function.) As the source sequence is assumed to be white,

$$E[y(n)y(n)] = \|\mathbf{c}\|_2^2, \tag{14}$$

and the norm of $\mathbf{c}$ can be determined by monitoring the energy of the output sequence $y(n)$. A more easily implementable constraint is $\|\mathbf{w}\|_2^2 = 1$, since we have ready access to $\mathbf{w}$, but not to $\mathbf{c}$, so this is the constraint used in the simulations in Section 6.

The AR implementation requires $4L_w(L_c - \nu)$ multiply-adds and one division per update, whereas the MA implementation requires $3NL_w(L_c - \nu)$ multiply-adds and one division per update.

5 Properties of the Cost Function

As is typical of blind equalization algorithms, for instance the constant modulus algorithm (CMA) [11], SAM's cost surface is multi-modal. For SAM, this can be explained by the following theorem.

Theorem 1 *The SAM cost function is invariant to the operation* $\mathbf{w} \rightarrow \overline{\mathbf{w}}$*, where* $\overline{\mathbf{w}}$ *denotes* $\mathbf{w}$ *with the order of its elements reversed.*

Proof: Consider the autocorrelation sequences of the combined channels $\mathbf{c}_1 = \mathbf{h} \star \mathbf{w}$ and $\mathbf{c}_2 = \mathbf{h} \star \overline{\mathbf{w}}$.

$$\begin{aligned} R_{c_1c_1} &= \mathbf{c}_1 \star \overline{\mathbf{c}}_1 = (\mathbf{h} \star \mathbf{w}) \star \overline{(\mathbf{h} \star \mathbf{w})} \\ &= \mathbf{h} \star \mathbf{w} \star \overline{\mathbf{h}} \star \overline{\mathbf{w}} \\ &= (\mathbf{h} \star \overline{\mathbf{w}}) \star (\overline{\mathbf{h}} \star \mathbf{w}) \\ &= \mathbf{c}_2 \star \overline{\mathbf{c}}_2 = R_{c_2c_2}. \end{aligned} \tag{15}$$

Thus the auto-correlation sequence (and hence the SAM cost) is invariant to time-reversing $\mathbf{w}$. ■

Whenever there is a minimum of the SAM cost surface, say at $\mathbf{w}_o$, there will also be another minimum at $\overline{\mathbf{w}}_o$. Even though the SAM cost is the same, the achievable bit rate will not be the same for the two settings, so one of each pair of minima may be in an undesirable location. Also, since the SAM cost surface is symmetric with respect to $\mathbf{w} \Leftrightarrow \overline{\mathbf{w}}$, there will be minima, maxima, or saddle points along the subspace $\mathbf{w} = \overline{\mathbf{w}}$.

Consider the following example. The channel is $\mathbf{h} = [1, 0.3, 0.2]$, the cyclic prefix is 1 (so we want a 2-tap channel), there is no noise, and the 3-tap TEQ $\mathbf{w}$ satisfies $\|\mathbf{w}\|_2 = 1$. We can represent the TEQ in spherical coordinates: $w_0 \stackrel{\triangle}{=} w_x = \cos(\theta)\sin(\phi)$, $w_1 \stackrel{\triangle}{=} w_z = \cos(\phi)$, $w_2 \stackrel{\triangle}{=} w_y = \sin(\theta)\sin(\phi)$. Then $\mathbf{w} \rightarrow \overline{\mathbf{w}}$ is equivalent to reflecting θ over $\frac{\pi}{4}$ or $\frac{5\pi}{4}$, and $\mathbf{w} \rightarrow -\mathbf{w}$ is equivalent to the combination of reflecting ϕ over $\frac{\pi}{2}$ and adding π to θ (mod 2π).

Fig. 2 shows a contour plot of the SAM cost function. The contours are logarithmically spaced to show detail near the minima. There are four minima, but they all have equivalent values of the SAM cost, due to the equivalencies $\mathbf{w} \Leftrightarrow -\mathbf{w}$ and $\mathbf{w} \Leftrightarrow \overline{\mathbf{w}}$.

We compare the locations of these minima to the commonly used shortening SNR (SSNR) TEQ design [2]. The two global maxima of the SSNR are shown as circles in Fig. 2. The global maxima of the SSNR match up nicely with two of the global minima of the SAM cost. When we find a global minimum of the SAM cost, we may switch to the other global minimum of SAM by reversing the order of taps in $\mathbf{w}$ if that yields a higher shortening SNR.

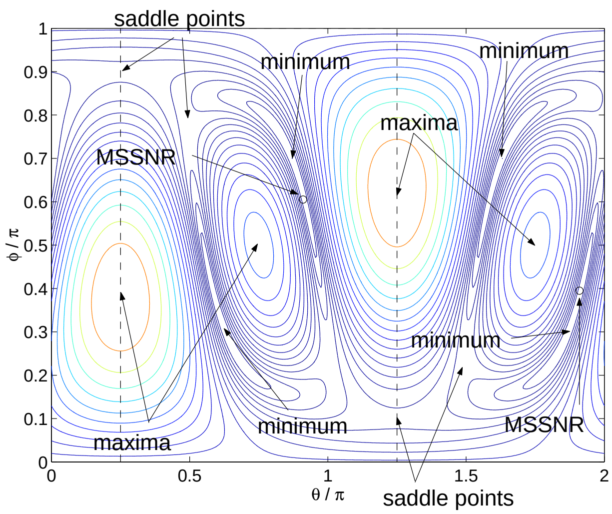

Figure 2: Contours of the SAM cost function. The two circles are the global maxima of the SSNR. The contours are logarithmically spaced to show detail.

6 Simulations

This section provides a numerical performance assessment of SAM in an ADSL environment. The Matlab code is available at [12]. The cyclic prefix was $\nu = 32$, $N_{fft} = 512$, the TEQ had 16 taps, and the channel was CSA test loop 1 [10], available at [13]. The noise was AWGN, with $\sigma_x^2\|\mathbf{c}\|^2/\sigma_v^2 = 40$ dB. No crosstalk was present. 75 symbols were used (of 544 samples each), and SAM used the AR implementation of (13) with $\alpha = 1/100$ and with the constraint $\|\mathbf{w}\|_2 = 1$. The initialization was a single spike, and the step size was 10 (such a large step size can be used because the SAM cost is very small, so the update size is still small).

Fig. 3 shows the SAM cost and achievable bit rate versus the iteration number, compared to the bit rate of the maximum SSNR TEQ and the matched filter bound (MFB). The bit rate is determined based on

$$R = \sum_{i \in \text{used tones}} \log_2\left(1 + \frac{SNR_i}{\Gamma}\right). \tag{16}$$

The SNR was computed using a 6 dB margin and a 4.2 dB coding gain. For more details, see [4] or [13].

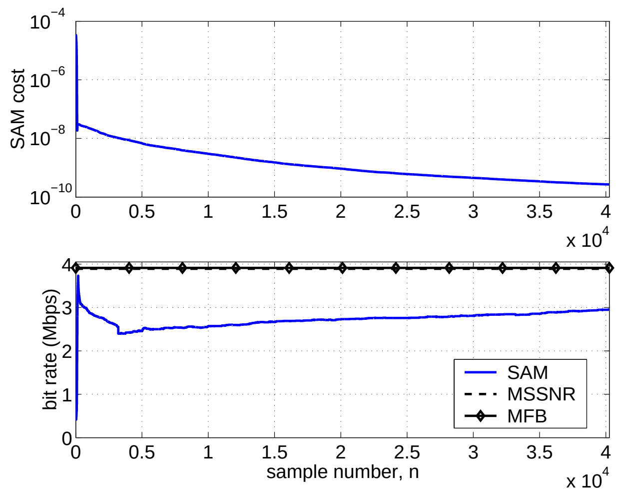

Figure 3: SAM cost (top) and achievable bit rate (bottom) vs. sample number, compared to the maximum shortening SNR TEQ and the matched filter bound.

The fact that the SAM cost is not monotonically decreasing in the first few hundred samples is because of the renormalization, which causes the algorithm to be only approximately a gradient descent algorithm. The bit rate is not monotonically increasing because the SAM cost bears no direct relation to the bit rate. At 340 iterations, SAM achieves 96% of the MFB, but then drops, and eventually rises again to 74% of the bound. The fact that the SAM cost is steadily decreasing when the bit rate decreases and then increases again indicates that the SAM minima and the bit rate maxima are not in the same location.

7 Conclusions and Future Work

A new blind, adaptive channel shortening algorithm based on a windowed sum-squared auto-correlation minimization has been proposed. The effectiveness of the algorithm to blindly shorten the channel has been demonstrated numerically. Proper initialization of the TEQ is necessary to ensure the convergence of the SAM algorithm to a good minima. Further studies are needed to characterize the cost function and formulate suitable design rules to ensure good performance. Robustness of the algorithm to violation of the assumption of source whiteness needs to be investigated further as well.

References

[1] 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.

[2] 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.

[3] 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.

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

[5] J. S. Chow, J. M. Cioffi, and J. A. C. Bingham, "Equalizer Training Algorithms for Multicarrier Modulation Systems," in *Proc. IEEE Int. Conf. on Comm.*, May 1993, pp. 761–765.

[6] N. Lashkarian and S. Kiaei, "Optimum Equalization of Multicarrier Systems: A Unified Geometric Approach," *IEEE Trans. on Comm.*, vol. 49, pp. 1762–1769, Oct. 2001.

[7] M. de Courville, P. Duhamel, P. Madec, and J. Palicot, "Blind equalization of OFDM systems based on the minimization of a quadratic criterion," in *Proceedings of the Int. Conf. on Communications*, Dallas, TX, June 1996, pp. 1318–1321.

[8] R. K. Martin, W. A. Sethares, J. Balakrishnan, and C. R. Johnson, Jr., "Blind, Adaptive Channel Shortening for Multicarrier Systems," Invited paper, to appear in *Proc. 36th Asilomar Conf. on Signals, Systems, and Computers*, Pacific Grove, CA, Nov. 2002.

[9] R. K. Martin, J. Balakrishnan, W. A. Sethares, and C. R. Johnson, Jr., "A Blind, Adaptive TEQ for Multicarrier Systems," *IEEE Signal Processing Letters*, Nov. 2002.

[10] K. Sistanizadeh, "Loss characteristics of the proposed canonical ADSL loops with 100-Ohm termination at 70, 90, and 120 F," ANSI T1E1.4 Committee Contribution, no. 161, Nov. 1991.

[11] C. R. Johnson, Jr., P. Schniter, T. J. Endres, J. D. Behm, D. R. Brown, and R. A. Casas, "Blind Equalization Using the Constant Modulus Criterion: A Review," *Proceedings of the IEEE*, vol. 86, pp. 1927–1950, Oct. 1998.

[12] R. K. Martin, "Matlab code for papers by R. K. Martin," available at http://bard.ece.cornell.edu/matlab/martin/index.html .

[13] G. Arslan, M. Ding, B. Lu, Z. Shen, and B. L. Evans, "TEQ design toolbox," The University of Texas at Austin, http://www.ece.utexas.edu/∼bevans/projects/adsl/dmtteq/dmtteq.html.