Blind, Adaptive Channel Shortening for Multicarrier Systems

R. K. Martin
Cornell University
Ithaca, NY 14853
frodo@ece.cornell.edu

J. Balakrishnan*
Texas Instruments
Dallas, TX 75243
jai@ti.com

W. A. Sethares
University of Wisconsin
Madison, WI 53706
sethares@ece.wisc.edu

C. R. Johnson, Jr.†
Cornell University
Ithaca, NY 14853
johnson@ece.cornell.edu

Abstract

This paper exploits the cyclic prefix to create a blind, adaptive, globally convergent channel shortening algorithm, with a complexity like LMS. The cost function is related to that of the Shortening SNR solution of Melsa, Younce, and Rohrs, and simulations are provided to demonstrate the performance of the algorithm.

1 Introduction

Multicarrier systems, such as OFDM or DMT, have less stringent equalization requirements than single carrier systems. If the channel is shorter than the cyclic prefix, then the effect of the channel is just a complex scaling for each carrier. However, if the channel exceeds this length, then inter-channel interference (ICI) and inter-symbol interference (ISI) will be present. The standard solution is to use a channel-shortening (time-domain) equalizer, or TEQ.

There are currently many methods which, when given a channel, can compute the optimal equalizer (for some metric) [1], [2], [3]. There are also several suboptimal and/or adaptive approaches, such as [4], [5], [6], [7]. Most approaches to TEQ design are non-adaptive, have high complexity, and require training or a channel estimate. While there are methods for blind channel identification for multicarrier systems, there is only one blind, adaptive method that directly equalizes the channel [7]. However, [7] performs complete equalization rather than channel shortening. This is not the desired criterion, so the overall performance is expected to be worse. Furthermore, [7] requires two matrix-vector multiplies per update, which is more computationally intensive than the proposed method.

*Jai Balakrishnan was at Cornell University when this work was performed.

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

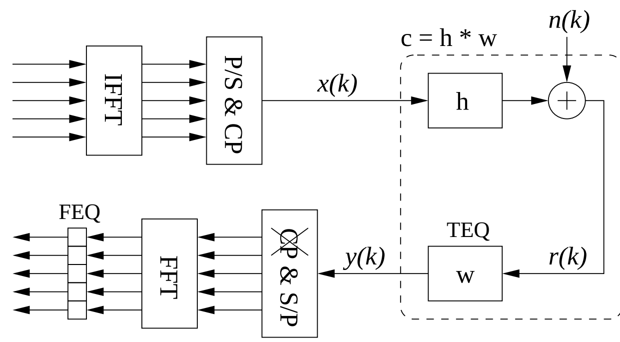

Figure 1: System model. (I)FFT: (inverse) fast Fourier transform, P/S: parallel to serial, CP: add cyclic prefix, xCP: remove cyclic prefix.

We propose a blind, adaptive channel shortening algorithm, which has significantly lower complexity than [7] and is intended for channel shortening (rather than equalization to a single spike).

2 System model

The (baseband) system model is shown in Figure 1. Each of the N frequency bins is modulated with a QAM signal, although often some bins are left as null carriers [7]. Modulation is performed via an inverse fast Fourier transform (IFFT), and demodulation is accomplished via an FFT. Channel shortening is performed by a TEQ, and the resulting shortened effective channel is equalized by a frequency-domain equalizer (FEQ), a bank of complex scalars.

After the CP is added, the last ν samples are identical to the first ν samples in the symbol, i.e.

$$x\left(Mk+i\right) = x\left(Mk+i+N\right), \quad i \in \left\{1, \ldots, \nu\right\}, \quad (1)$$

where $M = N + \nu$ is the total symbol duration and k is the symbol index. Fig. 2 shows an example of this, with $N = 8$, $\nu = 2$, and $M = N + \nu = 10$, and the symbol pictured is for $k = 0$. The received data $\mathbf{r}$ is

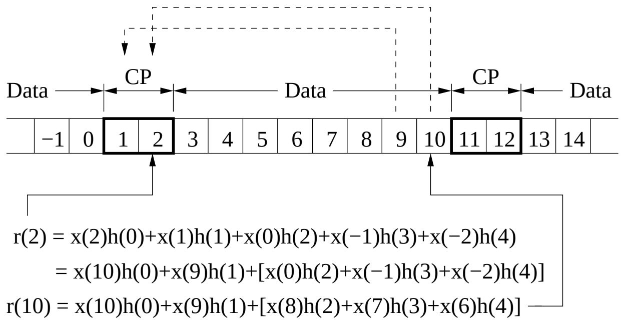

Figure 2: Illustration of the difference in the ISI at the received CP and at the end of the received symbol. $x(i)$, $h(i)$, and $r(i)$ are the transmitted data, channel, and received data, respectively.

obtained from $\mathbf{x}$ by

$$r(Mk+i) = \sum_{l=0}^{L_h} h(l) \cdot x(Mk+i-l) + n(Mk+i), \quad (2)$$

and the equalized data $\mathbf{y}$ is obtained from $\mathbf{r}$ by

$$y(Mk+i) = \sum_{j=0}^{L_w} w(j) \cdot r(Mk+i-j). \quad (3)$$

The channel has $L_h + 1$ taps, the TEQ has $L_w + 1$ taps, and the effective channel $\mathbf{c} = \mathbf{h} \star \mathbf{w}$ has $L_c + 1$ taps, where $L_c = L_h + L_w$.

3 A Merry algorithm

The channel destroys the relationship analogous to (1) in the received data, because the ICI & ISI that affect the CP are different from the ICI & ISI that affect the last ν samples in the symbol. Consider the example in Figure 2. The transmitted samples 2 and 10 are identical. However, at the receiver, the interfering samples before sample 2 are not all equal to their counterparts before sample 10. If $h(2)$, $h(3)$, and $h(4)$ were zero, then $r(2) = r(10)$. If we try to force $r(2) = r(10)$, we may force $h(2) = h(3) = h(4) = 0$, forcing the channel to be as short as the CP. The location of the window of ν non-zero taps can be varied by comparing $r(3)$ to $r(11)$, or $r(4)$ to $r(12)$, etc.

In general, if the channel length $L_h + 1 \leq \nu$, then the last sample in the CP should match the last sample in the symbol. One cost function that reflects this is

$$J_\Delta = \mathrm{E}\left[|y(Mk+\nu+\Delta) - y(Mk+\nu+N+\Delta)|^2 \right], \quad \Delta \in \{0, \ \ldots, \ M-1\}, \quad (4)$$

where Δ is the symbol synchronization parameter, which represents the desired delay of the channel-TEQ combination. The choice of Δ affects the cost function.

A stochastic gradient descent of (4) leads to a blind, adaptive TEQ, since the transmitted data need not be known. The proposed algorithm, "Multicarrier Equalization by Restoration of RedundancY" (MERRY), performs a stochastic gradient descent of (4), with a constraint to avoid the trivial solution $\mathbf{w} = \mathbf{0}$. The MERRY algorithm is:

For symbol $k = 0, 1, 2, \ldots,$

$$\begin{aligned} \tilde{\mathbf{r}}(k) &= \mathbf{r}(Mk+\nu+\Delta) - \mathbf{r}(Mk+\nu+N+\Delta) \\ e(k) &= \mathbf{w}^T(k)\ \tilde{\mathbf{r}}(k) \\ \hat{\mathbf{w}}(k+1) &= \mathbf{w}(k) - \mu\ e(k)\ \tilde{\mathbf{r}}^*(k) \\ \mathbf{w}(k+1) &= \frac{\hat{\mathbf{w}}(k+1)}{\|\hat{\mathbf{w}}(k+1)\|_2} \end{aligned} \quad (5)$$

where $\mathbf{r}(i) = [r(i), r(i-1), \ldots, r(i-L_w)]^T$, and * denotes complex conjugation.

Observe that MERRY is a simple vector update rule, with the added complexity of a renormalization. Due to the fact that MERRY compares the CP to the end of the symbol, only one update is possible per symbol. Alternate implementations of the constraint include fixing one tap to unity, maintaining a channel estimate and renormalizing to enforce $\|\mathbf{c}\| = 1$ instead of $\|\mathbf{w}\| = 1$, or including a penalty term in the cost function to enforce the norm constraint.

MERRY can also be implemented in transmitter-zero OFDM (TZ-OFDM) systems [8], as opposed to cyclic prefix OFDM (CP-OFDM) systems. TZ-OFDM systems transmit zeros during the guard period that is used for the cyclic prefix in CP-OFDM. This is equivalent to assuming that the samples in the CP ($x(1)$ and $x(2)$ in Fig. 2) are zero, rather than copies of the data. The MERRY cost function and update in this case are formed in the same fashion.

4 Properties of the solution

We now analyze the cost function, and relate the minima of (4) to the Shortening SNR solution of [3]. Throughout, we assume that:

1. The IFFT input is zero-mean, white, and wide sense stationary (implying that the output bins of the IFFT are uncorrelated),
2. $N \geq L_c + 1$ (the FFT size is at least as large as the length of the effective channel),
3. the noise autocorrelation function $\mathbf{R}_n(\delta) = 0$ for $\delta \geq N - L_w$,
4. the noise is uncorrelated with the data.

4.1 The cost function and constraints

The following theorem relates our work to the non-adaptive maximum shortening SNR (MSSNR) TEQ design [3]. It says that MERRY attempts to produce a "don't care" region with a width of ν taps.

Theorem 1 *For CP-OFDM systems, the cost function (4) simplifies to*

$$J_\Delta = 2\ \sigma_x^2 \left(\sum_{j=0}^{\Delta-1} |c_j|^2 + \sum_{j=\nu+\Delta}^{L_c} |c_j|^2 \right) + 2\ \mathbf{w}^T \mathbf{R}_n \mathbf{w}^*, \tag{6}$$

and for TZ-OFDM systems [8], (4) simplifies to

$$J_\Delta = \sigma_x^2 \left(\sum_{j=0}^{\Delta-1} |c_j|^2 + \sum_{j=\nu+\Delta}^{L_c} |c_j|^2 \right) + \|\mathbf{c}\|_2^2 + 2\ \mathbf{w}^T \mathbf{R}_n \mathbf{w}^* \tag{7}$$

where $\mathbf{R}_n = E\left[\mathbf{n}_i\ \mathbf{n}_i^H\right]$*, and* $\mathbf{c} = \mathbf{h} \star \mathbf{w}$.

Proof: Consider the following definitions:

$$\begin{aligned} \mathbf{x}_j &= [x(j), x(j-1), \ldots, x(j-L_c)]^T \\ \tilde{\mathbf{x}}_j &= \mathbf{x}_j - \mathbf{x}_{j+N} \\ \mathbf{n}_j &= [n(j), n(j-1), \ldots, n(j-L_w)]^T \\ \tilde{\mathbf{n}}_j &= \mathbf{n}_j - \mathbf{n}_{j+N}. \end{aligned} \tag{8}$$

For simplicity, drop the symbol index k (or assume $k=0$). Then (4) simplifies to

$$\begin{aligned} J_\Delta &= \mathrm{E}\left[\left| \mathbf{c}^T \mathbf{x}_{\nu+\Delta} - \mathbf{c}^T \mathbf{x}_{\nu+N+\Delta} \right. \right. \\ &\qquad \left. \left. + \mathbf{w}^T \mathbf{n}_{\nu+\Delta} - \mathbf{w}^T \mathbf{n}_{\nu+N+\Delta} \right|^2 \right] \\ &= \mathrm{E}\left[\left| \mathbf{c}^T \tilde{\mathbf{x}}_{\nu+\Delta} \right|^2 \right] + \mathrm{E}\left[\left| \mathbf{w}^T \tilde{\mathbf{n}}_{\nu+\Delta} \right|^2 \right] \\ &= \mathbf{c}^T \underbrace{\mathrm{E}\left[\tilde{\mathbf{x}}_{\nu+\Delta} \tilde{\mathbf{x}}_{\nu+\Delta}^H \right]}_{\mathbf{A}_x^{\nu+\Delta}} \mathbf{c}^* + \mathbf{w}^T \underbrace{\mathrm{E}\left[\tilde{\mathbf{n}}_{\nu+\Delta} \tilde{\mathbf{n}}_{\nu+\Delta}^H \right]}_{\mathbf{A}_n^{\nu+\Delta}} \mathbf{w}^*. \end{aligned} \tag{9}$$

In going to the second line, we have assumed that the noise and the data are uncorrelated (assumption 4). The matrix $\mathbf{A}_n^{\nu+\Delta}$ simplifies to

$$\begin{aligned} \mathbf{A}_n^{\nu+\Delta} &= \mathrm{E}\left[\mathbf{n}_{\nu+\Delta} \mathbf{n}_{\nu+\Delta}^H \right] - \mathrm{E}\left[\mathbf{n}_{\nu+\Delta} \mathbf{n}_{\nu+N+\Delta}^H \right] \\ &\quad - \mathrm{E}\left[\mathbf{n}_{\nu+N+\Delta} \mathbf{n}_{\nu+\Delta}^H \right] + \mathrm{E}\left[\mathbf{n}_{\nu+N+\Delta} \mathbf{n}_{\nu+N+\Delta}^H \right]. \end{aligned}$$

If assumption 3 holds, then the middle two terms are zero. The remaining two terms each equal the noise autocorrelation matrix, so $\mathbf{A}_n^{\nu+\Delta} = 2\mathbf{R}_n$.

To simplify $\mathbf{A}_x^{\nu+\Delta}$, observe that

$$\begin{aligned} \tilde{\mathbf{x}}_{\nu+\Delta} = [&x(\nu+\Delta) - x(\nu+\Delta+N), \\ &x(\nu+\Delta-1) - x(\nu+\Delta+N-1),\ \cdots, \\ &x(\nu+\Delta-L_c) - x(\nu+\Delta-L_c+N)]^T. \end{aligned} \tag{10}$$

The values of x that enter additively have a highest index of $(\nu+\Delta)$, whereas the values of x that enter with a minus sign have a smallest index of $(\nu+\Delta-L_c+N)$. If $N \geq L_c+1$ (assumption 2), then the highest index in the first group will be lower than the lowest index in the second group. If this holds, and if the IFFT output is uncorrelated (assumption 1), then the elements of $\tilde{\mathbf{x}}_{\nu+\Delta}$ will be uncorrelated. This means that $\mathbf{A}_x^{\nu+\Delta}$ will be diagonal.

In CP-OFDM systems, the middle ν elements in (10) are all zero, due to (1). Thus, for CP-OFDM,

$$\mathbf{A}_x^{\nu+\Delta} = 2\sigma_x^2 \left[\mathrm{diag}\left(\mathbf{1}_{\Delta\times 1},\ \mathbf{0}_{\nu\times 1},\ \mathbf{1}_{(L_c+1-\nu-\Delta)\times 1} \right) \right], \tag{11}$$

where diag($\mathbf{v}$) is a diagonal matrix with diagonal elements equal to the elements of vector $\mathbf{v}$.

In TZ-OFDM systems, the middle ν elements in (10) are not zero, but one summand of each of the middle $\nu+1$ pairs will be zero. Thus, for TZ-OFDM,

$$\begin{aligned} \mathbf{A}_x^{\nu+\Delta} =& \sigma_x^2 \left[\mathrm{diag}\left(\mathbf{1}_{\Delta\times 1},\ \mathbf{1}_{\nu\times 1},\ \mathbf{1}_{(L_c+1-\nu-\Delta)\times 1} \right) \right] + \\ & \sigma_x^2 \left[\mathrm{diag}\left(\mathbf{1}_{\Delta\times 1},\ \mathbf{0}_{\nu\times 1},\ \mathbf{1}_{(L_c+1-\nu-\Delta)\times 1} \right) \right]. \end{aligned} \tag{12}$$

(12) was written in this fashion to emphasize the contrast with (11). Substituting into (9), for CP-OFDM we get (6), and for TZ-OFDM we get (7). Thus J_Δ is proportional to the energy of the combined impulse response outside of a length ν window, plus a noise gain term, plus (for TZ-OFDM) the norm of the effective channel. This completes the proof. ∎

Theorem 1 suggests that MERRY finds a solution similar to the one found by the MSSNR design [3]. MERRY minimizes the energy outside of a length ν window plus the energy of the filtered noise, subject to a constraint (e.g. $\|\mathbf{w}\| = 1$ or $\|\mathbf{c}\| = 1$). In contrast, [3] minimizes the energy of the combined impulse response outside of a window of length $\nu+1$ (rather than ν), subject to the constraint

$$\|\mathbf{c}_{win}\|_2^2 = \sum_{j=\Delta}^{\Delta+\nu} |c_j|^2 = 1. \tag{13}$$

In the absence of noise, if MERRY uses the constraint $\|\mathbf{c}_{win}\|_2^2 = 1$, its optimum solution is the same as the MSSNR solution, albeit with a window size of ν instead of the usual $\nu+1$. The next theorem says that even if the constraint MERRY uses is $\|\mathbf{c}\| = 1$, the MERRY solution matches the MSSNR solution.

Theorem 2 *For CP-OFDM, the minimization of (4) under the constraint* $E\left[y^2(k)\right] = 1$ *(which is* $\|\mathbf{c}\|_2^2 = 1$ *in the noiseless case) yields the same solution as the minimization of (4) under the constraint* $\|\mathbf{c}_{win}\|_2^2 = 1$.

Proof: Define

$$\mathbf{A} = \mathbf{H}_{wall}^T \mathbf{H}_{wall} + \mathbf{R}_n \quad (14)$$

$$\mathbf{B} = \mathbf{H}_{win}^T \mathbf{H}_{win} \quad (15)$$

$$\mathbf{C} = \mathbf{H}^T \mathbf{H} + \mathbf{R}_n = \mathbf{A} + \mathbf{B}, \quad (16)$$

where $\mathbf{Hw}$, $\mathbf{H}_{win}\mathbf{w}$, and $\mathbf{H}_{wall}\mathbf{w}$ form the effective channel, the windowed effective channel, and the effective channel outside the window, respectively, as in [3]. The minimization of (4) with $\|\mathbf{c}_{win}\|_2^2 = 1$ is

$$\mathbf{w}_1 = \arg \min_{\mathbf{w}^T\mathbf{Bw}=1} \mathbf{w}^T \mathbf{Aw}. \quad (17)$$

The solution must satisfy

$$\mathbf{Aw}_1 = \lambda \mathbf{Bw}_1, \quad (18)$$

where λ is the smallest generalized eigenvalue of the matrix pair $(\mathbf{A}, \mathbf{B})$. The minimization of (4) under the constraint $\mathrm{E}\left[y^2(n)\right] = 1$ is

$$\mathbf{w}_2 = \arg \min_{\mathbf{w}^T\mathbf{Cw}=1} \mathbf{w}^T \mathbf{Aw}. \quad (19)$$

The solution must satisfy

$$\mathbf{Aw}_2 = \hat{\lambda} \mathbf{Cw}_2, \quad (20)$$

where $\hat{\lambda}$ is the smallest generalized eigenvalue of the matrix pair $(\mathbf{A}, \mathbf{C})$. If (18) is satisfied, then

$$\mathbf{Aw}_1 = \lambda \mathbf{Cw}_1 - \lambda \mathbf{Aw}_1, \quad (21)$$

so

$$\mathbf{Aw}_1 = \left(\frac{\lambda}{1+\lambda} \right) \mathbf{Cw}_1. \quad (22)$$

Similarly, if (20) is satisfied, then

$$\mathbf{Aw}_2 = \left(\frac{\hat{\lambda}}{1-\hat{\lambda}} \right) \mathbf{Bw}_2. \quad (23)$$

Comparing (18) to (22) and (20) to (23), there is a one-to-one mapping between the eigenvalues of $(\mathbf{A}, \mathbf{B})$ and the eigenvalues of $(\mathbf{A}, \mathbf{C})$. Now note from (20) that

$$\hat{\lambda} = \frac{\mathbf{w}_2^T \mathbf{Aw}_2}{\mathbf{w}_2^T \mathbf{Cw}_2} = \frac{\mathbf{w}_2^T \mathbf{Aw}_2}{\mathbf{w}_2^T \left(\mathbf{A}+\mathbf{B}\right) \mathbf{w}_2},$$

so $0 < \hat{\lambda} < 1$ (since $\mathbf{A}$ and $\mathbf{B}$ are positive semi-definite). A similar argument shows that $\lambda > 0$. Note that $x/(1-x)$ is monotonically increasing for $0 < x < 1$ and $x/(1+x)$ is monotonically increasing for $x > 0$. Thus, if λ is the smallest eigenvalue of (18), it produces the smallest value of $\lambda/(1+\lambda)$, which is the smallest eigenvalue in (22). Making explicit use of the one-to-one mapping between the eigenvalues of $(\mathbf{A}, \mathbf{B})$ and the eigenvalues of $(\mathbf{A}, \mathbf{C})$, we can say that if $\mathbf{w}_1$ is the eigenvector corresponding to the smallest eigenvalue of $(\mathbf{A}, \mathbf{B})$, then it is the eigenvector corresponding to the smallest eigenvalue of $(\mathbf{A}, \mathbf{C})$. The converse can be shown in a similar fashion, so (17) and (19) have the same solution. ■

The result of Theorem 2 is that if we enforce a constraint by monitoring the TEQ output energy $\mathrm{E}\left[y^2(k)\right]$ and forcing it to unity, then MERRY will converge to the MSSNR solution in the noiseless case. Future work will show that MERRY converges to the MMSE solution in the noisy case.

4.2 The minima

This section outlines a proof of global convergence of a gradient descent of (4). Define

$$\begin{aligned} \mathbf{r}_j &= \left[r(j), r(j-1), \ldots, r(j-L_w)\right]^T \\ \tilde{\mathbf{r}}_i &= \mathbf{r}_j - \mathbf{r}_{j+N} \end{aligned} \quad (24)$$

Adding a Lagrangian constraint to the cost function,

$$\begin{aligned} J_\Delta &= \mathrm{E}\left[\left| y\left(\nu+\Delta\right) - y\left(\nu+\Delta+N\right) \right|^2 \right] + \lambda\left(1 - \mathbf{w}^H\mathbf{w}\right) \\ &= \mathrm{E}\left[\left| \mathbf{w}^T \mathbf{r}_{\nu+\Delta} - \mathbf{w}^T \mathbf{r}_{\nu+\Delta+N} \right|^2 \right] + \lambda\left(1 - \mathbf{w}^H\mathbf{w}\right) \\ &= \mathbf{w}^T \underbrace{\mathrm{E}\left[\tilde{\mathbf{r}}_{\nu+\Delta} \tilde{\mathbf{r}}_{\nu+\Delta}^H \right]}_{2\mathbf{A}} \mathbf{w}^* + \lambda\left(1 - \mathbf{w}^H\mathbf{w}\right). \end{aligned}$$

The gradient $\bigtriangledown_{\mathbf{w}} J_\Delta = 0$ if and only if $(\lambda, \mathbf{w})$ are an eigenpair of $\mathbf{A}$. Then the Hessian $\mathbf{A} - \lambda\mathbf{I}$ is positive definite only if we choose λ to be the smallest eigenvalue. If the smallest eigenvalue is repeated, then there will be multiple minima but all will have the same cost (equal to the eigenvalue). This proves global convergence of the gradient descent algorithm. The proof is similar for the constraints $\mathbf{w}^T\mathbf{Bw} = 1$, $\|\mathbf{c}\| = 1$, and $\mathrm{E}\left[y^2(k)\right] = 1$.

5 Simulations

Figure 3 shows simulation results using CSA (carrier serving area) loop 1, a standard DSL test channel [5]. The Matlab code is available at [9]. The FFT size was 512, the CP length was 32, the TEQ had 16 taps, and $\sigma_x^2 \|\mathbf{h}\|^2 / \sigma_n^2 = 40$ dB. (Robustness to crosstalk will be considered in future work.) Initialization was a single spike. The DSL performance metric is the achievable bit rate for a fixed probability of error,

$$B = \sum_i \ln\left(1 + \frac{SNR_i}{\Gamma}\right),$$

where SNR_i is the signal to interference and noise ratio in frequency bin i. (We assume a 6 dB margin

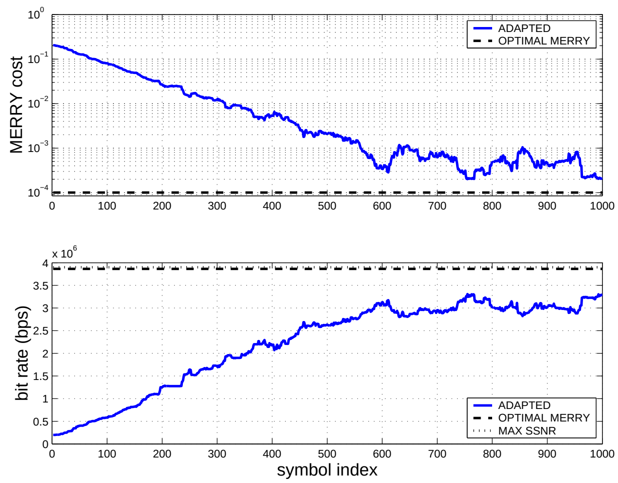

Figure 3: Achievable bit rate vs. time.

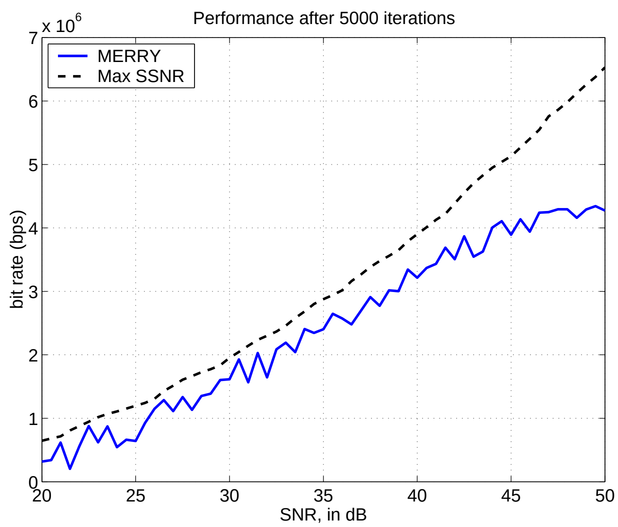

Figure 4: Achievable bit rate vs. SNR.

and 4.2 dB coding gain; for more details, refer to [1].) Figure 3 shows that MERRY can rapidly provide a solution approaching the maximum SSNR solution and the optimal MERRY solution.

Figure 4 shows a plot of the bit rate vs. SNR. The bit rate for this plot was computed by running for 5000 symbols and gradually decreasing the step size over time. For all these SNR values, MERRY approaches the max SSNR solution.

6 Conclusions

The MERRY algorithm performs blind, adaptive channel shortening for a multicarrier signal with a cyclic prefix, which has not hitherto been attempted (although de Courville, *et al.* [7] attempt full equalization for a multicarrier system without a cyclic prefix). The MERRY algorithm is low complexity and globally convergent. Future work will involve improving the convergence speed of the algorithm and expanding upon the analysis of the cost function.

Acknowledgments

The authors thank the authors of the TEQ design toolbox [10], which was used in the simulations.

References

[1] 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.

[2] 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.

[3] 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.

[4] 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.

[5] B. Farhang-Boroujeny and M. Ding, "Design Methods for Time-Domain Equalizers in DMT Transceivers," *IEEE Trans. on Comm.*, vol. 49, no. 3, pp. 554–562, Mar. 2001.

[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] A. Scaglione, G. B. Giannakis, and S. Barbarossa, "Redundant Filterbank Precoders and Equalizers, Part I: Unification and Optimal Designs," *IEEE Trans. on Signal Processing*, vol. 47, pp. 1988–2006, July 1999.

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

[10] 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.