

Approximate Minimum Bit-Error Rate Equalization for Binary Signaling

Chen-Chu Yeh and John R. Barry

School of Electrical and Computer Engineering
Georgia Institute of Technology, Atlanta, Georgia 30332-0250

Abstract — Although most linear and decision-feedback equalizers are designed to minimize a mean-squared error (MSE) performance metric, the equalizer that directly minimizes bit-error rate (BER) may significantly outperform the minimum-MSE equalizer, especially for binary antipodal signaling and its biorthogonal extensions, such as four quadrature-amplitude modulation. We show that the performance gain of the minimum-BER equalizer over the minimum-MSE equalizer is most pronounced when the number of equalizer coefficients is small relative to the severity of the intersymbol interference. We propose a simple stochastic gradient algorithm that approximately minimizes BER in the presence of linear intersymbol interference and white Gaussian noise. Computer simulations reveal that the proposed algorithm compares favorably to the popular least-mean-square algorithm in terms of both steady-state performance and complexity.

I. INTRODUCTION

We consider the design and adaptation of a finite-tap linear equalizer for combating intersymbol interference in the presence of additive white Gaussian noise, under the constraint that decisions are made on a symbol-by-symbol basis by quantizing the equalizer output. The most popular design strategy in this setting is the minimum mean-squared error (MMSE) equalizer, which can be realized using the least-mean square (LMS) algorithm. However, as recognized in [1–3], a better strategy is to choose the equalizer coefficients so as to minimize the error probability or bit-error rate (BER), not MSE. Unfortunately, stochastic gradient algorithms minimizing BER are significantly more complex than their MSE counterparts. Also, minimum-BER equalizers require that the noise power spectral density be estimated [3]. Finally, the BER surface of even a simple binary channel is highly irregular, and convergence to the global minimum cannot generally be guaranteed.

In this paper, we derive a simple adaptive algorithm achieving approximate minimum-BER (AMBER) performance for both linear and decision-feedback equalizers, specializing to binary and quadrature-amplitude modulation. The proposed AMBER algorithm has the following attributes: it closely approximates the minimum-BER equalizer; it does not require knowledge of the noise variance; it has low complexity (even lower than the least mean-square algorithm); and is guaranteed to converge to the global minimum of its cost function under the mild condition that the channel be equalizable.

This paper is organized as follows. In Sect. II, we present models for the channel and equalizer. In Sect. III, we discuss exact minimum-BER (EMBER) equalization. In Sect. IV, we

modify the minimum-BER cost function and propose the AMBER adaptive algorithm. In Sect. V, we present numerical results comparing the AMBER and MMSE equalizers.

II. MODELS FOR CHANNEL AND EQUALIZER

Consider the linear discrete-time binary signaling channel depicted in Fig. 1, where x_k is the binary input drawn from $\{\pm 1\}$, h_k is the channel impulse response with memory M , and n_k is white Gaussian noise with power spectral density σ^2 . The channel output r_k is:

$$r_k = s_k + n_k = \sum_{i=0}^M h_i x_{k-i} + n_k, \quad (1)$$

where s_k is the noiseless channel output.

Also shown in Fig. 1 is a linear equalizer with $N + 1$ coefficients described by the vector $\mathbf{c} = [c_0 \dots c_N]^T$. The equalizer output can be expressed as the inner product $y_k = \mathbf{c}^T \mathbf{r}_k$ between \mathbf{c} and a channel output vector $\mathbf{r}_k = [r_k \dots r_{k-N}]^T$, where:

$$\mathbf{r}_k = \mathbf{s}_k + \mathbf{n}_k = \mathbf{H}\mathbf{x}_k + \mathbf{n}_k, \quad (2)$$

where $\mathbf{x}_k = [x_k \dots x_{k-M-N}]^T$ is a vector of channel inputs, $\mathbf{n}_k = [n_k \dots n_{k-N}]^T$ is a vector of noise samples, $\mathbf{s}_k = [s_k \dots s_{k-N}]^T = \mathbf{H}\mathbf{x}_k$ is a vector of noiseless channel outputs, and \mathbf{H} is the $(N + 1) \times (M + N + 1)$ Toeplitz convolution matrix:

$$\mathbf{H} = \begin{bmatrix} h_0 & \dots & h_M & 0 & \dots & 0 \\ & & & \dots & & \\ 0 & \dots & 0 & h_0 & \dots & h_M \end{bmatrix}. \quad (3)$$

We make the restrictive assumption that the decision \hat{x}_{k-D} about the information symbol x_{k-D} is determined by the sign of the equalizer output, $\hat{x}_{k-D} = \text{sgn}(y_k)$, where D accounts for the delay of both the channel and the equalizer. Of course, this memoryless decision device is suboptimal; better BER performance can be achieved by performing maximum-likelihood sequence detection on the equalizer output.

Because of the binary alphabet for x_k , the bit-error probability after the linear equalizer \mathbf{c} has a particularly simple form:

$$\Pr[\hat{x}_{k-D} \neq x_{k-D}] = \mathbb{E} \left[Q \left(\frac{x_{k-D} \mathbf{c}^T \mathbf{s}_k}{\|\mathbf{c}\| \sigma} \right) \right], \quad (4)$$

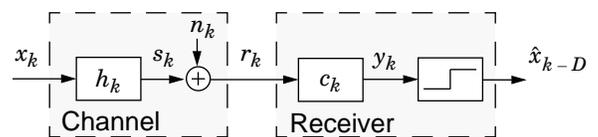


Fig. 1. Binary signaling channel and receiver.

where Q is the Gaussian error function, and the expectation is over all 2^{M+N+1} equally likely $\mathbf{s}_k = \mathbf{H}\mathbf{x}_k$ vectors. Observe that the error probability depends on the direction $\mathbf{c}/\|\mathbf{c}\|$ of \mathbf{c} only, and that the length of \mathbf{c} is irrelevant; this is because the receiver decisions are determined by the sign of the equalizer output only.

III. EXACT MINIMUM-BER EQUALIZATION

Unlike the coefficient vector \mathbf{c}_{MMSE} that minimizes MSE, there is no closed-form expression for a coefficient vector \mathbf{c}_{EMBER} that achieves exact minimum-BER (EMBER) performance. However, by setting to zero the gradient of (4) with respect to \mathbf{c} , it is straightforward to show that \mathbf{c}_{EMBER} must satisfy the following fixed-point relationship:

$$\mathbf{c}_{EMBER} = A f(\mathbf{c}_{EMBER}), \quad \text{for some } A > 0, \quad (5)$$

where $f: \mathbf{R}^{N+1} \rightarrow \mathbf{R}^{N+1}$ is defined by:

$$f(\mathbf{c}) = \mathbb{E} \left[\exp \left(\frac{-(x_{k-D} \mathbf{c}^T \mathbf{s}_k)^2}{2\|\mathbf{c}\|^2 \sigma^2} \right) x_{k-D} \mathbf{s}_k \right], \quad (6)$$

and again the expectation is over all 2^{M+N+1} equally likely $\mathbf{s}_k = \mathbf{H}\mathbf{x}_k$ vectors.

The function $f(\mathbf{c})$ plays an important role in our analysis, and it has a useful geometric interpretation. If we let $\mathbf{s}^{(1)}, \mathbf{s}^{(2)}, \dots, \mathbf{s}^{(L)}$ denote the $L = 2^{M+N+1}$ equally likely $x_{k-D} \mathbf{s}_k = x_{k-D} \mathbf{H}\mathbf{x}_k$ vectors, then $f(\mathbf{c})$ can be expressed as a weighted sum of $\mathbf{s}^{(i)}$ vectors:

$$f(\mathbf{c}) = \frac{1}{L} \left(e^{-\alpha_1^2/2} \mathbf{s}^{(1)} + e^{-\alpha_2^2/2} \mathbf{s}^{(2)} + \dots + e^{-\alpha_L^2/2} \mathbf{s}^{(L)} \right), \quad (7)$$

where $\alpha_i = \mathbf{c}^T \mathbf{s}^{(i)} / (\|\mathbf{c}\| \sigma)$ is a normalized inner product of $\mathbf{s}^{(i)}$ with \mathbf{c} . Because $\exp(\cdot)$ is an exponentially decreasing function, $f(\mathbf{c})$ is dictated by only those $\mathbf{s}^{(i)}$ vectors whose inner products with \mathbf{c} are relatively small. But the inner product of $\mathbf{s}^{(i)}$ with \mathbf{c} is the noiseless equalizer output when a one is transmitted; thus, a small inner product is equivalent to a nearly closed eye diagram. Thus, $f(\mathbf{c})$ will be very nearly a linear combination of the few $\mathbf{s}^{(i)}$ vectors for which the eye diagram is most closed.

The existence of at least one unit-length $\tilde{\mathbf{c}}$ satisfying (5) can be intuitively explained: The hyper-sphere of all unit-length vectors $\tilde{\mathbf{c}}$ is closed, continuous, and bounded. Each point on the sphere is mapped to a real value via the differentiable and bounded BER function of (4) and forms another closed, continuous, and bounded surface. The resultant surface is differentiable everywhere and has at least one local minimum, due to its closed and bounded properties. In general, there exist more than one local minima, *i.e.* more than one unit-length $\tilde{\mathbf{c}}$ satisfying (5).

Although the BER cost function in general is not convex, a gradient algorithm may still be used to search for its global minimum. In particular, a gradient algorithm based on (4) yields:

$$\begin{aligned} \mathbf{c}_{k+1} &= \mathbf{c}_k - \mu \nabla_{\mathbf{c}_k} \Pr[\hat{x}_{k-D} \neq x_{k-D}] \\ &= \left(1 - \frac{\mu \mathbf{c}_k^T f(\mathbf{c}_k)}{\|\mathbf{c}_k\|^3} \right) \left(\mathbf{c}_k + \frac{\mu}{1 - \mu \mathbf{c}_k^T f(\mathbf{c}_k) / \|\mathbf{c}_k\|^3} f(\mathbf{c}_k) \right). \end{aligned} \quad (8)$$

Recall that the length of \mathbf{c} has no impact on BER, and observe that the first bracketed factor in (8) represents an adjustment of the length of \mathbf{c}_{k+1} . Eliminating this factor leads to the so-called EMBER deterministic gradient algorithm (EMBER-DGA):

$$\mathbf{c}_{k+1} = \mathbf{c}_k + \mu f(\mathbf{c}_k). \quad (9)$$

Although the transformation from (8) to (9) affects the convergence rate and the steady-state norm $\|\mathbf{c}_\infty\|$, it has no effect on the steady-state direction $\mathbf{c}_\infty/\|\mathbf{c}_\infty\|$, and thus no effect on the steady-state BER performance. In fact, it can be proven that the EMBER-DGA of (9) is globally convergent for a certain class of channels, as described in the following theorem:

Theorem 1. If the maximum angle between any pair of $x_{k-D} \mathbf{s}_k$ vectors is less than $\pi/2$, then the EMBER-DGA of (9) is guaranteed to converge to the minimum-BER solution, regardless of its initial condition.

IV. APPROXIMATE MINIMUM-BER EQUALIZATION

We now propose the *approximate* minimum-BER deterministic gradient algorithm (AMBER-DGA) by making a simple modification to the EMBER-DGA update equation (9). We then propose a low-complexity stochastic update equation; propose a modification for faster convergence; and extend it to quadrature amplitude modulation.

A. Deterministic AMBER

The error function $Q(\alpha)$ is upper bounded and approximated by $\exp(-\alpha^2/2)/(\sqrt{2\pi}\alpha)$ [4], so that $f(\mathbf{c})$ of (6) and (7) can be approximated by:

$$\begin{aligned} f(\mathbf{c}) &\approx \frac{\sqrt{2\pi}}{L} \left(\alpha_1 Q(\alpha_1) \mathbf{s}^{(1)} + \alpha_2 Q(\alpha_2) \mathbf{s}^{(2)} + \dots + \alpha_L Q(\alpha_L) \mathbf{s}^{(L)} \right) \quad (10) \\ &\approx \sqrt{2\pi} \alpha_{\min} g(\mathbf{c}), \end{aligned} \quad (11)$$

where $\alpha_{\min} = \min\{\alpha_i\}$, and where we have introduced the vector function $g: \mathbf{R}^{N+1} \rightarrow \mathbf{R}^{N+1}$:

$$\begin{aligned} g(\mathbf{c}) &= \frac{1}{L} \left(Q(\alpha_1) \mathbf{s}^{(1)} + Q(\alpha_2) \mathbf{s}^{(2)} + \dots + Q(\alpha_L) \mathbf{s}^{(L)} \right) \\ &= \mathbb{E} \left[Q \left(\frac{x_{k-D} \mathbf{c}^T \mathbf{s}_k}{\|\mathbf{c}\| \sigma} \right) x_{k-D} \mathbf{s}_k \right]. \end{aligned} \quad (12)$$

Comparing (6) and (12), we see that $g(\mathbf{c})$ has the same form as $f(\mathbf{c})$, but with $Q(\alpha)$ replacing $\exp(-\alpha^2/2)$. The approximation in (11) is valid because only the terms in (10) for which $\alpha_i \approx \alpha_{\min}$ are relevant, the other terms have negligible impact. Using $g(\mathbf{c})$ to approximate $f(\mathbf{c})$ in (9) leads to the following *approximate* minimum-BER DGA (AMBER-DGA):

$$\mathbf{c}_{k+1} = \mathbf{c}_k + \mu g(\mathbf{c}_k). \quad (13)$$

Because of the approximations in (10) and (11), the AMBER DGA no longer minimizes BER exactly. However, the approximations are well-justified, and the simulation results of Sect. V verify that the equalizer minimizing the AMBER cost function very nearly minimizes BER.

Fortunately, the AMBER-DGA of (13) is globally convergent for a much broader class of channels than the EMBER-DGA:

Theorem 2. If the eye of the channel can be opened by the equalizer, the AMBER algorithm of (13) is guaranteed to converge to the global minimum of its cost function.

B. Stochastic AMBER

At first glance, the AMBER-DGA of (13) is no less complicated than the EMBER-DGA of (9). However, we can introduce an *error indicator* function I_{xy} , where

$$I_{xy} = \frac{1}{2} (1 - \text{sgn}[x_{k-D}y_k]). \quad (14)$$

In other words, $I_{xy} = 1$ when an error is made ($\hat{x}_{k-D} \neq x_{k-D}$), and $I_{xy} = 0$ when no error is made ($\hat{x}_{k-D} = x_{k-D}$). We can use this indicator function to simplify the AMBER-DGA of (13), through the following series of straightforward equalities:

$$\begin{aligned} \mathbf{c}_{k+1} &= \mathbf{c}_k + \mu \mathbb{E} \left[Q \left(\frac{x_{k-D} \mathbf{c}_k^T \mathbf{s}_k}{\|\mathbf{c}_k\| \sigma} \right) x_{k-D} \mathbf{s}_k \right] \\ &= \mathbf{c}_k + \mu \mathbb{E} \left[\mathbb{E}[I_{xy} | x_{k-D} \mathbf{s}_k] \mathbb{E}[x_{k-D} \mathbf{s}_k] \right] \\ &= \mathbf{c}_k + \mu \mathbb{E}[I_{xy} x_{k-D} \mathbf{s}_k] \\ &= \mathbf{c}_k + \mu \mathbb{E}[I_{xy} x_{k-D} \mathbf{r}_k]. \end{aligned} \quad (15)$$

The last equality follows because $\mathbf{r}_k = \mathbf{s}_k + \mathbf{n}_k$, and I_{xy} and x_{k-D} are statistically independent, so that:

$$\mathbb{E}[I_{xy} x_{k-D} \mathbf{n}_k] = \mathbb{E}[x_{k-D}] \mathbb{E}[I_{xy} \mathbf{n}_k] = 0. \quad (16)$$

A simple and asymptotically unbiased stochastic gradient update algorithm consequently can be formed by removing the expectation in (15):

$$\mathbf{c}_{k+1} = \mathbf{c}_k + \mu I_{xy} x_{k-D} \mathbf{r}_k. \quad (17)$$

We refer to this stochastic algorithm as AMBER. The equalizer is updated only when an error is made. AMBER has a form similar to the familiar LMS algorithm: both are described by (17), the only difference being that $I_{xy} = 1 - x_{k-D}y_k$ for LMS, and $I_{xy} = (1 - \text{sgn}[x_{k-D}y_k])/2$ for AMBER.

Because AMBER-SGA is an unbiased stochastic version of AMBER-DGA, it can be shown that, for a sufficiently small step size, the AMBER-SGA converges to a solution satisfying the following fixed-point relation:

$$\mathbf{c}_{AMBER} = A \mathbf{g}(\mathbf{c}_{AMBER}), \quad \text{for some } A > 0. \quad (18)$$

Observe the similarity between (18) and (5).

The simple AMBER-SGA of (17) has an insightful geometric interpretation. Recall that the noiseless equalizer output when a one is transmitted is the inner product of \mathbf{c} with $x_{k-D} \mathbf{s}_k = \mathbf{s}^{(i)}$. Most errors occur when this inner product is small, *i.e.* when the eye is nearly closed. The AMBER update of (17) dictates that each time an error is made, the equalizer coefficient vector \mathbf{c} takes a small step in space towards the $\mathbf{s}^{(i)}$ vector that resulted in the error. Therefore, the next time the input bits conspire to produce the same $\mathbf{s}^{(i)}$ vector, its inner product with \mathbf{c} will be larger.

In other words, the eye opening will be larger. Thus, we can view (17) as a heuristic algorithm for maximizing the eye opening.

At steady state, as shown in (18), \mathbf{c}_{AMBER} is a weighted combination of $\mathbf{s}^{(i)}$ vectors, with weights proportional to the conditional error probability. And since these conditional error probabilities are dominated by the few $\mathbf{s}^{(i)}$ vectors that close the eye the most, \mathbf{c}_{AMBER} will be very nearly a linear combination of the few $\mathbf{s}^{(i)}$ vectors for which the eye diagram is most closed.

C. Threshold Modification

When a training sequence is not available, it may be tempting to operate the AMBER algorithm in a decision-directed manner by replacing x_{k-D} by \hat{x}_{k-D} in (17). Unfortunately, this would cause the indicator function (14) to be zero always, stalling adaptation from the start, because $\hat{x}_{k-D}y_k = \text{sgn}(y_k)y_k = |y_k| > 0$. To overcome this problem, we introduce a positive threshold τ into the AMBER adaptation process. The modified update equation for both training and data modes becomes:

$$\mathbf{c}_{k+1} = \mathbf{c}_k + \mu \hat{I}_{xy} \hat{x}_{k-D} \mathbf{r}_k, \quad (19)$$

where the modified indicator function is $\hat{I}_{xy} = 1$ if $y_k \hat{x}_{k-D} \leq \tau$ and $\hat{I}_{xy} = 0$ otherwise. Because the norm of \mathbf{c} may grow with time, we define τ relative to this norm using $\tau = \alpha \|\mathbf{c}\|$ for some fixed constant $\alpha \in (0, 1)$. As discussed in Sect. V, the threshold modification changes the steady state performance of the equalizer.

Besides its utility as a decision-directed algorithm, a second advantage of the threshold modification is that it increases significantly the rate of convergence of the AMBER algorithm with training. The original AMBER algorithm updates only when an error is made, while the modified AMBER algorithm updates when an error is made and also when an error is almost made. Thus, the threshold modification causes \mathbf{c} to be updated much more frequently.

Similar to the decision-directed LMS update process, we can expect the modified AMBER update process to be valid when the BER is reasonably small, perhaps 10^{-1} or 10^{-2} or less.

D. Approximate Minimum-BER Equalization for 4-QAM

Although the original AMBER algorithm is defined for binary signaling only, it can be generalized to four quadrature-amplitude modulation (4-QAM) as follows. Using superscripts R and I to denote real and imaginary parts, respectively, of a complex 4-QAM system, the BER (not symbol error rate) for 4-QAM signaling with Gray coding is:

$$\text{BER}_{QAM} = \frac{1}{2} \mathbb{E} \left[Q \left(\frac{x_{k-D}^R (\mathbf{c}_k^T \mathbf{s}_k)^R}{\sigma \|\mathbf{c}_k\|} \right) + Q \left(\frac{x_{k-D}^I (\mathbf{c}_k^T \mathbf{s}_k)^I}{\sigma \|\mathbf{c}_k\|} \right) \right]. \quad (20)$$

We assume n_k^R and n_k^I are white and independent with PSD σ^2 . Following the AMBER-SGA derivation for binary signaling, the following AMBER update equations for 4-QAM can be derived:

$$\begin{aligned} \mathbf{c}_{k+1}^R &= \mathbf{c}_k^R + \mu (\hat{I}_{x^R y^R} \hat{x}_{k-D}^R \mathbf{r}_k^R + \hat{I}_{x^I y^I} \hat{x}_{k-D}^I \mathbf{r}_k^I) \\ \mathbf{c}_{k+1}^I &= \mathbf{c}_k^I + \mu (-\hat{I}_{x^R y^R} \hat{x}_{k-D}^R \mathbf{r}_k^I + \hat{I}_{x^I y^I} \hat{x}_{k-D}^I \mathbf{r}_k^R). \end{aligned} \quad (21)$$

V. NUMERICAL RESULTS

To demonstrate the effectiveness of the proposed algorithms, we now present simulation results for both binary and 4-QAM channels. All BER simulation results are obtained with $5 \cdot 10^5$ training data bits and 10^7 data bits. All equalizers use a step size of $\mu = 0.001$ during training and no adaptation occurs in data mode. For binary signaling channels, EMBER curves are plotted via numerical evaluation of (4) with ideal EMBER filter coefficients. We use a small fixed threshold of $\tau = 0.1$ (not normalized relative to $\|\mathbf{c}\|$) for the AMBER algorithm to increase its convergence rate for all BER simulations. In every example we choose the equalizer delay D to be optimal in the MSE sense.

A. Channel A: Linear Equalization for Binary Signaling

Here we consider binary signaling and linear equalization for a channel with transfer function $H(z) = 1.2 + 1.1z^{-1} - 0.2z^{-2}$. In Fig. 2 we plot BER versus $\text{SNR} = \sum_k h_k^2 / \sigma^2$, considering both MMSE and AMBER equalizers of length three and five. The figure shows that with 3 filter taps and a delay of $D = 2$, the AMBER equalizer has a more than 6.5 dB gain over the MMSE equalizer. With 5 filter taps and a delay of $D = 4$, AMBER still has a nearly 2 dB gain over MMSE. Observe that the AMBER (solid) and EMBER (dashed) curves are nearly indistinguishable.

The results of Fig. 2 show that the improvement of AMBER over MMSE drops from 6.5 dB to 2 dB as the equalizer length increases from three to five, suggesting that MMSE approaches AMBER as the equalizer length increases. In Fig. 3 we plot SNR

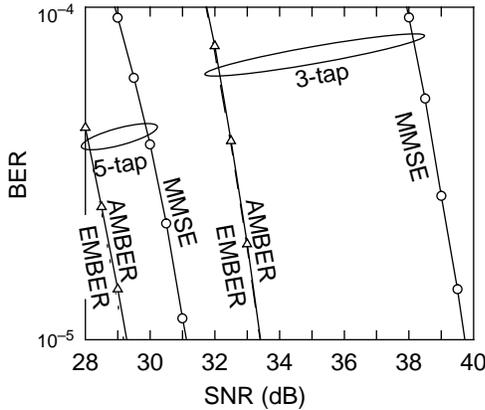


Fig. 2. Performance of linear equalization for channel A.

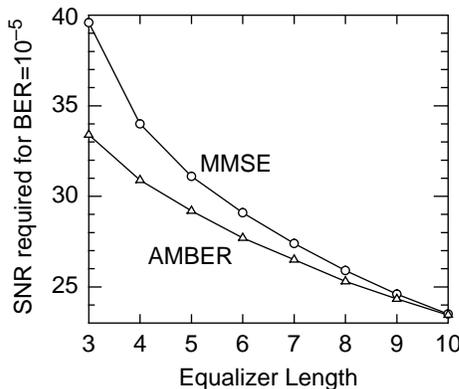


Fig. 3. SNR requirement vs. equalizer length for channel A.

required to achieve $\text{BER} = 10^{-5}$ versus equalizer length for both AMBER and MMSE. We see that MMSE approaches AMBER as the length of the equalizer increases.

Fig. 4 depicts the “artificial” noiseless eye patterns (obtained by interpolating all possible noiseless equalizer outputs with a triangular pulse shape) for EMBER, AMBER, and MMSE with five equalizer taps and $\text{SNR} = 30$ dB. All three equalizer are normalized to have identical norm (and thus identical noise enhancement). The EMBER and AMBER eye patterns are virtually identical, while the MMSE eye pattern is drastically different. In particular, the EMBER and AMBER eye patterns exhibit a larger eye opening than the MMSE eye pattern.

The convergence rates of the AMBER and MMSE 3-tap equalizers are compared at $\text{SNR} = 30$ dB in Fig. 5, assuming a fixed threshold of $\tau = 0.1$. Observe that AMBER converges much slower than LMS. To partially mitigate this problem, a gear-shift algorithm combining LMS and AMBER is used here. The LMS algorithm is used to first quickly reach a good initial condition, and then the adaptation process is switched to the AMBER algorithm to approach the minimum BER solution.

In Fig. 6 we plot both steady-state BER and convergence time for the AMBER-DGA as a function of $\alpha = \tau / \|\mathbf{c}\|$, assuming $\text{SNR} = 32$ dB. Observe that for α less than 0.1, the AMBER steady state BER is very close to the true minimum BER, whereas the BER degrades significantly for larger thresholds. Meanwhile, the convergence rate (right-hand scale) improves dramatically as α increases from 0 to 0.1, but not appreciably for α greater than 0.1. (Here, we define convergence as when the BER is within 20% of its steady-state value.) Comparing the two curves, we conclude that choosing $\alpha = 0.1$ provides a good balance between low steady-state BER and fast convergence.

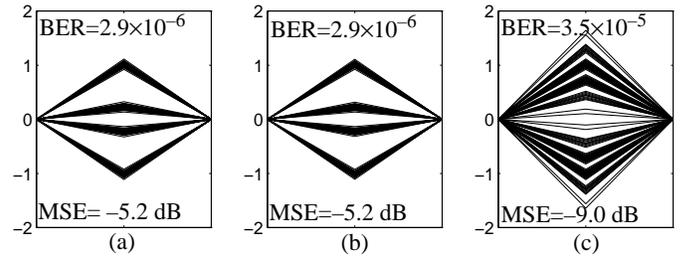


Fig. 4. Equalized eye patterns for (a) EMBER; (b) AMBER; and (c) MMSE.

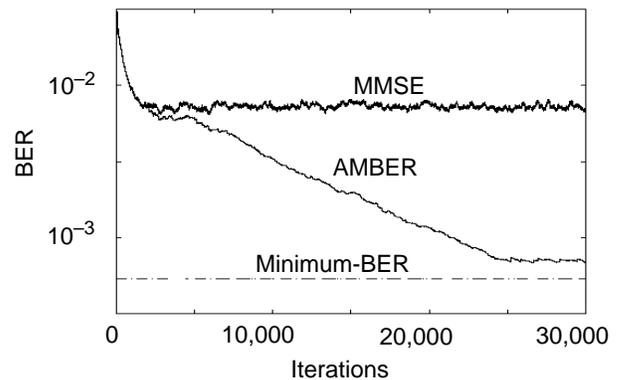


Fig. 5. BER versus time for MMSE and AMBER SGAs.

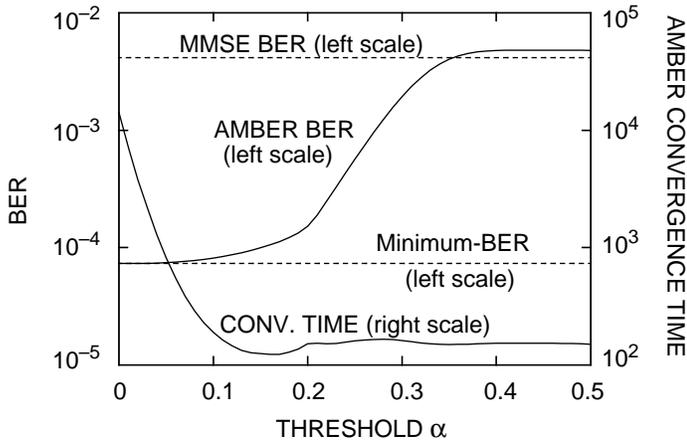


Fig. 6. The effect of AMBER threshold on steady-state BER (left scale) and convergence time (right scale).

B. Channel B: Decision-Feedback Equalization

Consider another simple channel with transfer function $H(z) = 0.35 + 0.8z^{-1} + z^{-2} + 0.8z^{-3}$, again with binary signaling, but this time with decision-feedback equalization. In Fig. 7 we compare the BER performance of AMBER to MMSE. For a five-tap DFE (3 forward and 2 feedback taps), AMBER has a more than 5 dB gain over MMSE at $\text{BER} = 10^{-5}$. For a seven-tap DFE (4 forward and 3 feedback taps), AMBER outperforms MMSE by about 1.8 dB. Observe that the 5-tap AMBER DFE outperforms the 7-tap MMSE DFE.

C. Channel C: Linear Equalization for 4-QAM Signaling

Here we consider 4-QAM $\{\pm 1 \pm j\}$ with linear equalization and $H(z) = (0.7 - 0.2j) + (0.4 - 0.5j)z^{-1} + (-0.2 + 0.3j)z^{-2}$, and $\text{SNR} = \sum_k |h_k|^2 / \sigma^2$. As shown in Fig. 8, the 4-tap ($D=3$) AMBER linear equalizer outperforms MMSE equalizer by about 18 dB. With five taps, the gain drops to slightly more than 2 dB.

In Fig. 9 we present the noiseless constellation diagrams for the 5-tap AMBER and MMSE linear equalizers. Observe the interesting structure of the AMBER constellation clouds; they result in a higher MSE than the MMSE clouds (which appear roughly Gaussian), but the edges of the AMBER clouds are further apart.

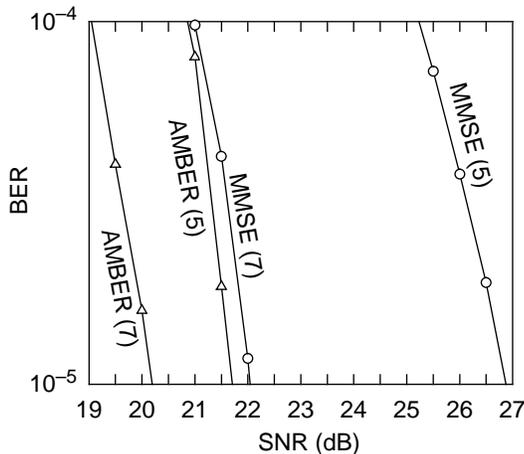


Fig. 7. BER performance for DFE on Channel B.

VI. CONCLUSION

We have demonstrated the potentially dramatic gains of a linear equalizer that exactly minimizes BER as compared to a minimum-MSE equalizer for both binary and 4-QAM signaling. We have proposed the AMBER algorithm that is less complex than the LMS algorithm but achieves far superior BER performance on severely distorted channels, especially when the equalizer length is short. Under the mild condition that the equalizer can open the channel eye diagram, the AMBER update algorithm is globally convergent, regardless of its initial condition. Simulation results confirm that the proposed algorithm very nearly minimizes BER. On the other hand, the proposed algorithm generally needs more training data to extract the ensemble average information from a time average. Improving the convergence rate of AMBER is a topic for future research.

VII. REFERENCES

- [1] E. Shamash and K. Yao, "On the Structure and Performance of a Linear Decision Feedback Equalizer Based on the Minimum Error Probability Criterion," *ICC '74*, pp. 25F1-25F5, 1974.
- [2] P. Galko and S. Pasupathy, "Optimal Linear Receiver Filters for Binary Digital Signals," *ICC '82*, pp. 1H.6.1-1H.6.5, 1982.
- [3] S. Chen, E. Chng, B. Mulgrew, and G. Gibson, "Minimum-BER Linear-Combiner DFE," *ICC '96*, pp. 1173-1177, 1996.
- [4] E. A. Lee and D. G. Messerschmitt, *Digital Communication*, Second Edition, Kluwer, Boston, 1994.

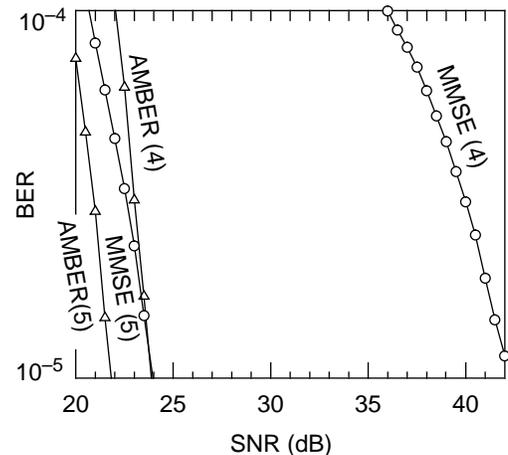


Fig. 8. BER comparison for linear equalizer on channel C.

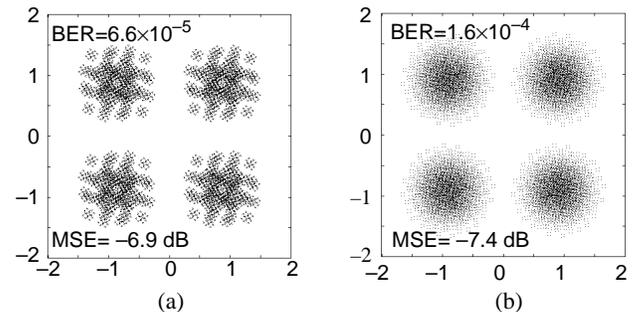


Fig. 9. Noiseless equalized constellations of 5-tap (a) AMBER and (b) MMSE at $\text{SNR} = 20$ dB on channel C.