Adaptive MBER Decision Feedback Multiuser Receivers in Frequency Selective Fading Channels

Rodrigo C. de Lamare, *Student Member, IEEE*, and Raimundo Sampaio-Neto

***Abstract*—In this letter we investigate adaptive minimum bit error rate (BER) decision feedback multiuser receivers for DS-CDMA systems in fast frequency selective Rayleigh fading channels. We examine stochastic gradient adaptive algorithms and introduce fast algorithms for minimizing the BER cost function from training data.**

***Index Terms*—Adaptive algorithms, BER cost functions, decision feedback, multiuser detection.**

I. Introduction

Adaptive decision feedback (DFE) multiuser receivers (MUDs) employing the minimum mean squared error (MMSE) criterion usually show superior performance to linear MUDs and have simple adaptive implementation [1], [2]. However, it is well known that the mean squared error (MSE) cost function is not optimal in digital communications, and the most appropriate cost function is the bit error rate (BER) [3], [4]. The approximate minimum BER (AMBER) [3] and the least BER (LBER) [4] are two of the most successful and suitable cost functions for adaptive implementation, even though they usually require long training sequences to outperform MMSE techniques. In this work, we extend the AMBER and the LBER techniques to the DFE MUD case and investigate faster algorithms that can speed up the convergence of these receivers, requiring shorter training data. Finally, we assess the convergence, tracking and BER performance of these algorithms in fast frequency selective Rayleigh fading channels through computer simulations.

II. DS-CDMA System Model

We consider a synchronous DS-CDMA system with N users, PG chips per symbol and binary symbols $b_i(k)$, where the subscript denotes user i. These symbols are spread with signature sequences $\mathbf{c}_i = [\mathbf{c}_{i,1} \ldots \mathbf{c}_{i,PG}]^T$, modulated and transmitted through a communication channel characterized by: $H(z) = \sum_{i=0}^{n_h-1} h_i z^{-i}$, where the operator z^{-1} introduces a delay of one chip time in the transmitted signal. The received signal after filtering by a chip-pulse matched filter and sampled at the chip rate is given by $\mathbf{r}(k) = \mathbf{s}(k) + \mathbf{n}(k)$ or alternatively by

$$\mathbf{r}(k) = \mathbf{H}\begin{bmatrix} \mathbf{CA} & \mathbf{0} & \ldots\mathbf{0} & \\ \mathbf{0} & \mathbf{CA} & \ddots & \vdots \\ \vdots & \ddots & \ddots & \mathbf{0} \\ \mathbf{0} & \ldots & \mathbf{0} & \mathbf{CA} \end{bmatrix} \begin{bmatrix} \mathbf{b(k)} \\ \mathbf{b(k-1)} \\ \vdots \\ \mathbf{b(k-L+1)} \end{bmatrix} + \mathbf{n}(k) \quad (1)$$

where the Gaussian noise vector $\mathbf{n}(k) = [n_1(k) \ \ldots \ n_{PG}(k)]^T$ with $E[\mathbf{n}(k)\mathbf{n}^T(k)] = \sigma^2\mathbf{I}$, $\mathbf{s}(k)$ is the noise-free signal vector, the user bit vector is given by $\mathbf{b}(k) = [b_1(k) \ \ldots \ b_N(k)]^T$, the user signature sequence matrix is described by $\mathbf{C} = [\mathbf{c}_1 \ \ldots \ \mathbf{c}_N]$, the user signal amplitude matrix is represented by $\mathbf{A} = \mathrm{diag}\{A_1 \ \ldots \ A_N\}$, and the $PG \times (L \times PG)$ matrix $\mathbf{H}$ implements channel convolution as in [4]. The multiple access interference (MAI) is originated from the nonorthogonality between the user signature sequences, whereas the intersymbol interference (ISI) span L depends on the length of the channel response. For $n_h = 1$, $L = 1$ (no ISI), for $1 < n_h \leq PG$, $L = 2$, for $PG < n_h \leq 2PG$, $L = 3$.

III. DFE MUD

DFE receivers employ the decisions of associated users to cancel the MAI, outperforming linear MUDs and improving the performance of the system [1], [2]. The output of the one-shot DFE multiuser receiver is described by

$$x_i(k) = \mathbf{w}_i^T(k)\mathbf{r}(k) - \mathbf{f}_i^T(k)\hat{\mathbf{b}}(k) \quad (2)$$

where $\mathbf{r}(k)$ is the received vector and $\hat{\mathbf{b}}(k)$ is the $N \times 1$ vector of decisions. The feedforward matrix $\mathbf{w}(k)$ is $PG \times N$, the feedback matrix $\mathbf{f}(k)$ is $N \times N$ and is constrained to have zeros along the diagonal to avoid cancelling the desired symbols. In this work, we employ a full matrix $\mathbf{f}(k)$, except for the diagonal, which corresponds to parallel decision feedback [2]. The detected symbol for the DFE MUD is given by $\hat{b}_i(k) = \mathrm{sgn}(x_i(k))$, where $x_i(k)$ is the kth estimated symbol for user i.

IV. Minimum BER Cost Functions

Given a user i transmitted training sequence $\mathbf{d}_i$, the bit error probability $P(\epsilon|\mathbf{d}_i)$, for the DFE receiver, is expressed by

$$\begin{aligned} P\big(\epsilon|\mathbf{d}_i\big) &= P_{\epsilon_i} = P\big(d_i(k)\mathrm{sgn}(x_i(k)) = -1\big) \\ P_{\epsilon_i} &= P\big(\mathrm{sgn}(d_i(k)x_i(k)) = -1\big) \\ &= P\big(d_i(k)x_i(k) < 0\big) \end{aligned} \quad (3)$$

where $x_i(k)$ is given through (2) and $d_i(k)$ is the desired symbol taken from the training sequence for user i and symbol k.

Manuscript received September 5, 2002. The associate editor coordinating the review of this paper and approving it for publication was Dr. Y.-D. Yao. This work was supported by the Brazilian Council for Scientific and Technological Development (CNPq).

The authors are with CETUC/PUC-RIO, 22453-900, Rio de Janeiro, Brazil (e-mail: delamare@infolink.com.br; raimundo@cetuc.puc-rio.br).

Digital Object Identifier 10.1109/LCOMM.2002.808373

A. The AMBER Cost Function

The MUD solution for linear receivers ($\mathbf{f}_i(k) = 0$ in (2)) that minimizes the BER cost function via the AMBER algorithm [3] employs a vector function $g(\mathbf{w}_i(k))$ to approximate an expression for a coefficient vector $\mathbf{w}_i(k)$ that achieves a MBER performance in linear receivers as described by

$$g(\mathbf{w}_i(k)) = E\left[Q\left(\frac{d_i(k)\mathbf{w}_i^T(k)\mathbf{s}(k)}{\| \mathbf{w}_i(k) \| \sigma}\right)d_i(k)\mathbf{s}(k)\right] \quad (4)$$

where $d_i(k)$ is the desired transmitted symbol for user i, taken from the training sequence, $Q(.)$ is the Gaussian error function and $\mathbf{s}(k)$ are the received samples without noise taken from the outputs of chip-matched filters. For linear MUDs the quantity $Q\left(d_i(k)\mathbf{w}_i^T(k)\mathbf{s}(k)/\| \mathbf{w}_i(k) \| \sigma\right)$ inside the expected value operator in (6) corresponds to the conditional bit error probability given the product $d_i(k)\mathbf{s}(k)$. This quantity can be replaced in (6) by an error indicator function $i_{d_i}(k)$ given by [3] $i_{d_i}(k) = \frac{1}{2}(1 - \mathrm{sgn}(d_i(k)x_i(k)))$ where $x_i(k)$ is the estimated symbol. The cost function gradient reduces to

$$g(\mathbf{w}_i(k)) = E\Big[E\left[i_{d_i}(k) \mid d_i(k)\mathbf{s}(k)\right]d_i(k)\mathbf{s}(k)\Big]. \quad (5)$$

Since $\mathbf{s}(k) = \mathbf{r}(k) - \mathbf{n}(k)$, and $i_{d_i}(k)$ and $d_i(k)$ are statistically independent, we have $E[i_{d_i}(k)d_i(k)\mathbf{n}(k)] = E[d_i(k)]E[i_{d_i}(k)\mathbf{n}(k)] = 0$ and thus the instantaneous value of the function is given by $g(\mathbf{w}_i(k)) = i_{d_i}(k)d_i(k)\mathbf{r}(k)$.

B. The LBER Cost Function

The MUD BER depends on the distribution of the decision variable $x_i(k)$, which is a function of the weights of the receiver. The sign-adjusted decision variable for the DFE MUD $x_{s_i}(k) = d_i(k)x_i(k)$ is drawn from a Gaussian mixture, described by

$$x_{s_i}(k) = \mathrm{sgn}(d_i(k))\left(\mathbf{w}_i^T\mathbf{s}(k) - \mathbf{f}_i^T\hat{\mathbf{b}}(k) + \mathbf{w}_i^T\mathbf{n}(k)\right) \quad (6)$$

where the first term of (6) is the noise free sign-adjusted MUD output. A single point kernel density estimate [4] is given by

$$\hat{p}_{x_{s_i}}(x_{s_i}) = \frac{1}{\sqrt{2\pi}\rho\sqrt{\mathbf{w_i^T w_i}}} \times \exp\left(\frac{-(x_{s_i} - \mathrm{sgn}(d_i(k))x_i(k))^2}{2\rho^2\mathbf{w_i^T w_i}}\right) \quad (7)$$

where ρ is the radius parameter of the kernel density estimate. The probability of error for user i is estimated by

$$P_{\epsilon_i} = P(x_{s_i} < 0) = \int_{-\infty}^{0}\hat{p}_{x_{s_i}}(x_{s_i})dx_{s_i} = Q\left(\frac{\mathrm{sgn}(d_i(k)x_i(k)}{\rho(\mathbf{w_i^T w_i})^{1/2}}\right). \quad (8)$$

The gradient terms of P_ϵ are

$$\frac{\partial P_{\epsilon_i}}{\partial \mathbf{w}_i} = \frac{\exp\left(\frac{-x_i(k)^2}{2\rho^2\mathbf{w_i^T w_i}}\right)\mathrm{sgn}(d_i(k))}{\sqrt{2\pi}\rho}\left(\frac{-\mathbf{r}(k)}{(\mathbf{w_i^T w_i})^{1/2}} + \frac{\mathbf{w_i}x_i(k)}{(\mathbf{w_i^T w_i})^{3/2}}\right) \quad (9)$$

$$\frac{\partial P_{\epsilon_i}}{\partial \mathbf{f}_i} = \frac{1}{\sqrt{2\pi}\rho\sqrt{\mathbf{w_i^T w}_i}}\exp\left(\frac{-x_i(k)^2}{2\rho^2\mathbf{w_i^T w}_i}\right)\mathrm{sgn}(d_i(k))\hat{\mathbf{b}}(k). \quad (10)$$

V. Adaptive Algorithms

In this section, we describe stochastic gradient (SG) and Gradient-Newton (GN) (fast) algorithms that adjust the parameters of the MUD's based on the minimization of the MSE and the BER cost functions. We have chosen the GN technique because the error surfaces of the MBER cost functions exhibit local minima and with the GN approach one can control the rate of convergence by carefully tuning the step size.

A. SG Algorithms

The AMBER algorithm for the DFE MUD, devised as an extension of the linear case, is obtained via a SG optimization and is expressed by

$$\mathbf{w}_i(k+1) = \mathbf{w}_i(k) + \mu i_{d_i}(k)d_i(k)\mathbf{r}(k) \quad (11)$$

$$\mathbf{f}_i(k+1) = \mathbf{f}_i(k) - \mu i_{d_i}(k)d_i(k)\hat{\mathbf{b}}(k) \quad (12)$$

where μ and μ_f are the algorithm step sizes, $i_{d_i}(k) = 1/2(1 - \mathrm{sgn}(d_i(k)x_i(k) - \tau))$ is a modified error indicator function whose threshold τ increases the rate of convergence [3].

The LBER algorithm for the DFE MUD is obtained substituting the gradient terms in $\mathbf{w}_i(k+1) = \mathbf{w}_i(k) - \mu\left[\partial P_{\epsilon_i}/\partial \mathbf{w}_i\right]_k$ and $\mathbf{f}_i(k+1) = \mathbf{f}_i(k) - \mu_f\left[\partial P_{\epsilon_i}/\partial \mathbf{f}_i\right]_k$ and adjusting the receiver weights such that $\mathbf{w}_i^T(k)\mathbf{w}_i(k) = 1$

$$\mathbf{w}_i(k+1) = \mathbf{w}_i(k) + \mu\frac{1}{\sqrt{2\pi}\rho}\exp\left(\frac{-(x_i(k))^2}{2\rho^2}\right)\mathrm{sgn}(d_i(k)) \times (\mathbf{r}(k) - \mathbf{w}_i(k)x_i(k)) \quad (13)$$

$$\mathbf{f}_i(k+1) = \mathbf{f}_i(k) - \mu_f\frac{1}{\sqrt{2\pi}\rho}\exp\left(\frac{-(x_i(k))^2}{2\rho^2}\right) \times \mathrm{sgn}(d_i(k))\hat{\mathbf{b}}(k) \quad (14)$$

where ρ is related to the noise standard deviation σ.

B. GN Algorithms

GN algorithms [5] incorporate second-order statistics of input signals. They usually have a faster convergence rate than SG techniques, although they require a higher computational complexity. The update equation of Newton's method is given by $\mathbf{w}_i(k+1) = \mathbf{w}_i(k) - \delta\hat{\mathbf{R}}_u^{-1}(k)z_{\mathbf{w_i}}(k)$ where $\mathbf{R}_u(k)$ is the autocorrelation matrix of the observation vector $\mathbf{u}(k)$ and $z_{\mathbf{w_i}}(k)$ is the gradient vector to be minimized. To avoid the required inversion of $\hat{\mathbf{R}}(k)$, we use the matrix inversion lemma, described in [5]. The GN-AMBER solution for the DFE MUD is derived in an analogous form to its SG version and is given by

$$\mathbf{w}_i(k+1) = \mathbf{w}_i(k) + \mu\hat{\mathbf{R}}_r^{-1}(k)i_{d_i}(k)d_i(k)\mathbf{r}(k) \quad (15)$$

$$\mathbf{f}_i(k+1) = \mathbf{f}_i(k) - \mu\hat{\mathbf{R}}_{\hat{b}}^{-1}(k)i_{d_i}(k)d_i(k)\hat{\mathbf{b}}(k). \quad (16)$$

The GN-LBER algorithm for the DFE MUD is also devised in a similar way to its SG counterpart

$$\mathbf{w}_i(k+1) = \mathbf{w}_i(k) + \delta\hat{\mathbf{R}}_r^{-1}(k)\frac{1}{\sqrt{2\pi}\rho}\exp\left(\frac{-(x_i(k))^2}{2\rho^2}\right) \times \mathrm{sgn}(d_i(k))\left(\mathbf{r}(k) - \mathbf{w}_i(k)x_i(k)\right) \quad (17)$$

$$\mathbf{f}_i(k+1) = \mathbf{f}_i(k) - \delta_f\hat{\mathbf{R}}_{\hat{b}}^{-1}(k)\frac{1}{\sqrt{2\pi}\rho}\exp\left(\frac{-(x_i(k))^2}{2\rho^2}\right) \times \mathrm{sgn}(d_i(k))\hat{\mathbf{b}}(k) \quad (18)$$

where δ and δ_f are the step sizes.

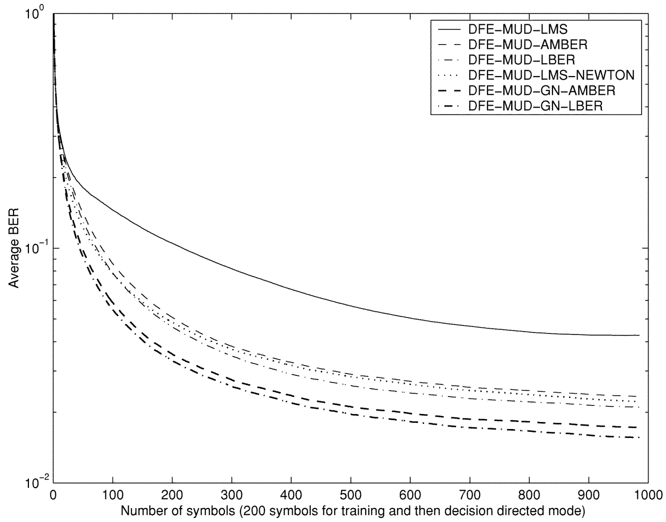

Fig. 1. Convergence performance of the algorithms at $E_b/N_0 = 10$ dB.

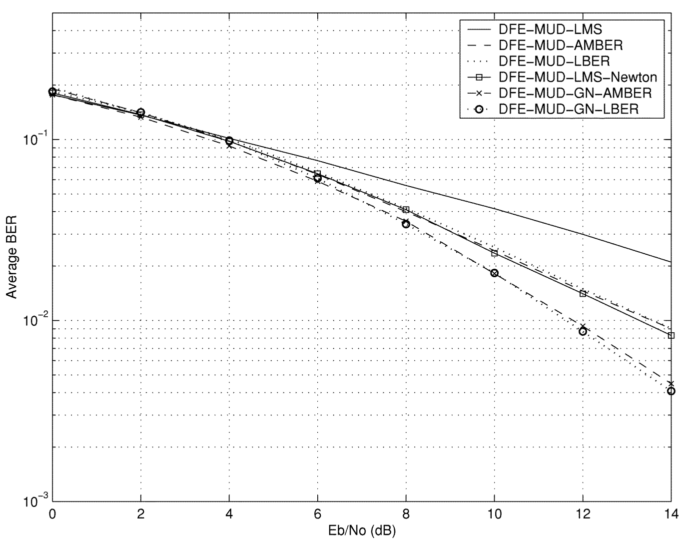

Fig. 2. BER performance of the DFE-MUD.

VI. Simulations

The performance of the DFE MUD's with the adaptive algorithms was evaluated in a three-path fast frequency selective Rayleigh fading channel with coefficients $h_i(k) = a_i\alpha_i(k)$ $(i = 0, 1, 2)$, where $\{\alpha_i(k)\}$, $i = 0, 1, 2$, are independent sequences of independent unit power Rayleigh random variables $(E[\alpha_i^2(k)] = 1)$, that change at each transmitted symbol. In all situations, the MUDs operate with Gold sequences of length $PG = 15$, process 200 symbols in training mode (TR) and then switch to the decision-directed (DD) mode. The optimized parameters of the algorithms are: $\mu = 0.0075$, $\mu_f = 0.0025$; $\delta = 0.000\,25$; $\delta_f = 0.0001$; $\alpha = 0.0025$; $\rho = 4\sigma^2$; and $\tau = 0.1$. The channel parameters are: $a_0 = 0.5$; $a_1 = 0.3$; and $a_2 = 0.2$.

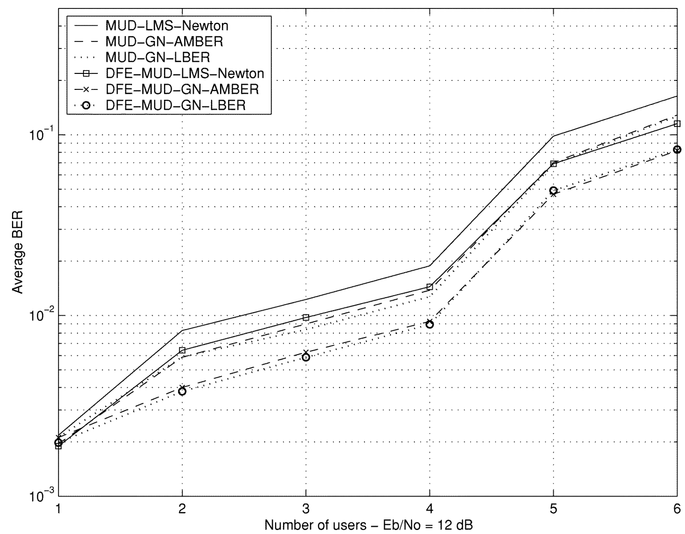

Fig. 3. BER performance with a varying number of users.

The convergence performance of the algorithms for a system with $N = 4$ users are shown in Fig. 1, where the MUDs process 200 symbols in TR and 800 symbols in DD, averaged over 100 independent experiments. The average BER performance versus E_b/N_0 is shown in Fig. 2 for a system with $N = 4$ users where each MUD processes 10^3 symbols averaged over 100 independent experiments. In Fig. 3 the average BER performance of the system with a varying number of users is shown. The results show that the MBER GN algorithms are superior to the other techniques, saving transmitting power and increasing system's capacity. It is also worth noting that the DFE-MUD-AMBER shows a very good trade-off between performance and complexity. Since complexity is comparable to the LMS with savings of almost 4 dB in E_b/N_0 for BER on the order of 10^{-2}.

VII. Conclusions

MBER algorithms for DFE MUDs have been proposed and evaluated in fast frequency selective Rayleigh fading channels. The fast algorithms have shown a performance superior to stochastic gradient MBER and previously reported MMSE algorithms.

References

[1] S. Verdu, *Multiuser Detection* Cambridge, 1998.
[2] M. Honig and M. K. Tsatsanis, "Adaptive techniques for multiuser CDMA receivers," *IEEE Signal Processing Mag.*, pp. 49–61, May 2000.
[3] C. Yeh, R. R. Lopes, and J. R. Barry, "Approximate minimum bit-error rate multiuser detection," in *GLOBECOM'98*, Sydney, Australia, Nov. 1998, pp. 3590–3595.
[4] S. Chen, A. K. Samingan, B. Mulgrew, and L. Hanzo, "Adaptive minimum BER linear multiuser detection," in *Proc. IEEE ICASSP*, Salt Lake City, UT, May 2001, pp. 2253–2256.
[5] P. S. R. Diniz, *Adaptive Filtering: Algorithms and Practical Implementations*. Boston, MA: Kluwer Academic, 1997.