ADAPTIVE MINIMUM-BER LINEAR MULTIUSER DETECTION

S. Chen†*, A.K. Samingan*†*, B. Mulgrew*‡ *and L. Hanzo*†

† Department of Electronics and Computer Science
University of Southampton, Southampton SO17 1BJ, U.K.

‡ Department of Electronics and Electrical Engineering
University of Edinburgh, Edinburgh EH9 3JL, U.K.

ABSTRACT

An adaptive minimum bit error rate (MBER) linear multiuser detector (MUD) is proposed for DS-CDMA systems. Based on the approach of kernel density estimation for approximating the bit error rate (BER) from training data, a least mean squares (LMS) style adaptive algorithm is developed for training linear MUDs. Computer simulation results show that this adaptive MBER linear MUD outperforms two existing LMS-style adaptive MBER algorithms.

1. INTRODUCTION

Within the class of linear MUDs, the minimum mean square error (MMSE) detector [1],[2] is popular, as it often performs adequately and has simple adaptive implementation. However, the BER of the MMSE MUD can be inferior to the MBER solution, and there exist two true stochastic gradient algorithms for realizing the MBER MUD [3],[4]. The algorithm of [3] uses a difference approximation to estimate the gradient of error probability. It does not assume the noise probability density function (p.d.f.) but has a complexity of $O(M^2)$, M being the detector dimension. This algorithm will be called DMBER. As it only adjusts the detector weights when an error occurs, the algorithm requires a very long training sequence to converge. The approximate MBER (AMBER) MUD of [4] is appealing due to its computational simplicity. It has a same form to the signed-error LMS algorithm [5], except in the vicinity of the decision boundary where it is modified to continue updating the weights when the signed-error LMS would not.

Adaptive MBER linear equalizers have been investigated for a longer time [6]-[8]. The LMS-style MBER equaliser of [7],[8], called least BER (LBER), has been shown to outperform the AMBER equalizer of [9], which is the counterpart of the AMBER linear MUD [4]. In this paper, we extend the LBER algorithm of [7],[8] to multiuser detection for DS-CDMA channels, and develop a new adaptive MBER MUD. Our study shows that this new LBER is superior in performance over the DMBER and AMBER.

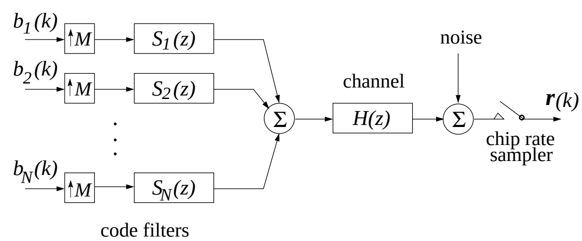

Figure 1: Discrete-time model of synchronous CDMA.

2. SYSTEM MODEL

The synchronous DS-CDMA system with N users and M chips per bit is depicted in Fig. 1, where $b_i(k) \in \{\pm 1\}$ denotes the k-th bit of user i, the signature sequence for user i $\bar{\mathbf{s}}_i = [\bar{s}_{i,1} \ \cdots \ \bar{s}_{i,M}]^T$ is normalized to have a unit length, and the channel impulse response (CIR) is

$$H(z) = \sum_{i=0}^{n_h-1} h_i z^{-i} \, . \tag{1}$$

The received signal sampled at chip rate is given by [10]:

$$\mathbf{r}(k) = \mathbf{H} \begin{bmatrix} \bar{\mathbf{S}}\mathbf{A} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \bar{\mathbf{S}}\mathbf{A} & \ddots & \vdots \\ \vdots & \ddots & \ddots & \mathbf{0} \\ \mathbf{0} & \cdots & \mathbf{0} & \bar{\mathbf{S}}\mathbf{A} \end{bmatrix} \begin{bmatrix} \mathbf{b}(k) \\ \mathbf{b}(k-1) \\ \vdots \\ \mathbf{b}(k-L+1) \end{bmatrix}$$

$$+\mathbf{n}(k) = \mathbf{P} \begin{bmatrix} \mathbf{b}(k) \\ \mathbf{b}(k-1) \\ \vdots \\ \mathbf{b}(k-L+1) \end{bmatrix} + \mathbf{n}(k) \, , \tag{2}$$

where the Gaussian noise vector $\mathbf{n}(k) = [n_1(k) \cdots n_M(k)]^T$ with $E[\mathbf{n}(k)\mathbf{n}^T(k)] = \sigma_n^2 \mathbf{I}$; the user bit vector $\mathbf{b}(k) = [b_1(k) \ \cdots \ b_N(k)]^T$; the user signature sequence matrix $\bar{\mathbf{S}} = [\bar{\mathbf{s}}_1 \ \cdots \ \bar{\mathbf{s}}_N]$; the diagonal user signal amplitude matrix $\mathbf{A} = \text{diag}\{A_1 \ \cdots \ A_N\}$; the $M \times LM$ CIR matrix $\mathbf{H}$

$$\mathbf{H} = \begin{bmatrix} h_0 & h_1 & \cdots & h_{n_h-1} & & & \\ & h_0 & h_1 & \cdots & h_{n_h-1} & & \\ & & \ddots & \ddots & \cdots & \ddots & \\ & & & h_0 & h_1 & \cdots & h_{n_h-1} \end{bmatrix} ; \tag{3}$$

and the system matrix $\mathbf{P}$ has a dimension $M \times LN$. The intersymbol interference span L depends on the length of the CIR related to the length of the chip sequence. For $n_h = 1$, $L = 1$; for $1 < n_h \leq M$, $L = 2$; for $M < n_h \leq 2M$, $L = 3$; and so on. Consider the linear MUD:

$$\hat{b}_i(k) = \mathrm{sgn}(y(k)) \text{ with } y(k) = \mathbf{w}^T \mathbf{r}(k) , \tag{4}$$

where $\mathbf{w} = [w_1 \cdots w_M]^T$ is the detector weight vector for user i. Let the $N_b = 2^{LN}$ possible sequences of $[\mathbf{b}^T(k)\, \mathbf{b}^T(k-1) \cdots \mathbf{b}^T(k-L+1)]^T$ be

$$\mathbf{b}^{(j)} = \begin{bmatrix} \mathbf{b}^{(j)}(k) \\ \mathbf{b}^{(j)}(k-1) \\ \vdots \\ \mathbf{b}^{(j)}(k-L+1) \end{bmatrix} , \; 1 \leq j \leq N_b , \tag{5}$$

and $b_i^{(j)}$ the ith element of $\mathbf{b}^{(j)}(k)$. Define the noise-free signal states $\mathbf{r}_j = \mathbf{P}\mathbf{b}^{(j)}$, $1 \leq j \leq N_b$, and the set:

$$y_j = \mathbf{w}^T \mathbf{r}_j , \; 1 \leq j \leq N_b . \tag{6}$$

3. THE MBER LINEAR MUD

The error probability of the linear detector (4) is:

$$P_E(\mathbf{w}) = \mathrm{Prob}\{\mathrm{sgn}(b_i(k)) y(k) < 0\} . \tag{7}$$

Following [7],[8], define the signed decision variable

$$y_s(k) = \mathrm{sgn}(b_i(k)) y(k) = \mathrm{sgn}(b_i(k)) y'(k) + n'(k) , \tag{8}$$

where

$$y'(k) = \mathbf{w}^T \mathbf{P} \begin{bmatrix} \mathbf{b}(k) \\ \mathbf{b}(k-1) \\ \vdots \\ \mathbf{b}(k-L+1) \end{bmatrix} \tag{9}$$

and $n'(k) = \mathrm{sgn}(b_i(k)) \mathbf{w}^T \mathbf{n}(k)$. Note that $y'(k)$ can only take the values from the set (6) and $n'(k)$ is Gaussian with zero mean and variance $\sigma_n^2 \mathbf{w}^T \mathbf{w}$. Assuming equiprobable $\mathbf{r}_j$, the p.d.f. of $y_s(k)$ is

$$p_y(y_s) = \frac{1}{N_b \sqrt{2\pi} \sigma_n \sqrt{\mathbf{w}^T \mathbf{w}}} \sum_{j=1}^{N_b} \exp\left(-\frac{(y_s - \mathrm{sgn}(b_i^{(j)}) y_j)^2}{2\sigma_n^2 \mathbf{w}^T \mathbf{w}} \right) . \tag{10}$$

Thus

$$P_E(\mathbf{w}) = \int_{-\infty}^{0} p_y(y_s)\, dy_s . \tag{11}$$

Given the gradient of $P_E(\mathbf{w})$ with respect to $\mathbf{w}$, $\nabla P_E(\mathbf{w})$, a MBER steepest-descent gradient algorithm is:

$$\mathbf{w}(l+1) = \mathbf{w}(l) - \mu \nabla P_E(\mathbf{w}(l)) . \tag{12}$$

As the orientation of $\mathbf{w}$ defines the decision boundary and thus the BER, not its size, it is computationally advantageous to normalize $\mathbf{w}$ to a unit-length after each iteration:

$$\mathbf{w} = \mathbf{w} / \sqrt{\mathbf{w}^T \mathbf{w}} . \tag{13}$$

A simplified conjugate gradient algorithm [11] can offer a better convergence rate:

Initialization. Choose $\mu > 0$ and termination scalar $\beta > 0$; given $\mathbf{w}(1)$ and $\mathbf{d}(1) = -\nabla P_E(\mathbf{w}(1))$; set $l = 1$.

Loop. If $\|\nabla P_E(\mathbf{w}(l))\| < \beta$: goto *Stop*.

$$\mathbf{w}(l+1) = \mathbf{w}(l) + \mu \mathbf{d}(l)$$

$$\mathbf{w}(l+1) = \mathbf{w}(l+1) / \|\mathbf{w}(l+1)\|$$

$$\phi_l = \|\nabla P_E(\mathbf{w}(l+1))\|^2 / \|\nabla P_E(\mathbf{w}(l))\|^2$$

$$\mathbf{d}(l+1) = \phi_l \mathbf{d}(l) - \nabla P_E(\mathbf{w}(l+1))$$

$l = l + 1$, goto *Loop*.

Stop. $\mathbf{w}(l)$ is the solution.

4. ADAPTIVE MBER LINEAR MUD

Kernel density estimation is known to produced reliable p.d.f. estimates with short data records [12]. Given a block of K training samples $\{\mathbf{r}(k), b_i(k)\}$, a kernel density estimate of the p.d.f. $p_y(y_s)$ is given by:

$$\hat{p}_y(y_s) = \frac{1}{K \sqrt{2\pi} \rho_n \sqrt{\mathbf{w}^T \mathbf{w}}} \times \sum_{k=1}^{K} \exp\left(-\frac{(y_s - \mathrm{sgn}(b_i(k)) y(k))^2}{2\rho_n^2 \mathbf{w}^T \mathbf{w}} \right) , \tag{14}$$

where the radius parameter ρ_n is related to the noise standard deviation σ_n [12]. From the estimated error probability $\hat{P}_E(\mathbf{w})$, $\nabla \hat{P}_E(\mathbf{w})$ can be calculated, and block adaptive gradient algorithms can similarly be developed.

Furthermore, a LMS-style adaptive algorithm with sample-by-sample adjustment, as in [7],[8], can be derived. At sample k, a point estimate of the p.d.f. is simply:

$$\hat{p}_y(y_s(k)) = \frac{1}{\sqrt{2\pi} \rho_n \sqrt{\mathbf{w}^T \mathbf{w}}} \times$$

$$\exp\left(-\frac{(y_s - \text{sgn}(b_i(k))y(k))^2}{2\rho_n^2 \mathbf{w}^T\mathbf{w}}\right) . \qquad (15)$$

Using the instantaneous or stochastic gradient

$$\nabla \hat{P}_E(k) = \frac{1}{\sqrt{2\pi}\rho_n}\left(\frac{\mathbf{w}\mathbf{w}^T - \mathbf{w}^T\mathbf{w}\mathbf{I}}{(\mathbf{w}^T\mathbf{w})^{\frac{3}{2}}}\right) \times \exp\left(-\frac{y^2(k)}{2\rho_n^2\mathbf{w}^T\mathbf{w}}\right)\text{sgn}(b_i(k))\mathbf{r}(k) \qquad (16)$$

and re-scaling after each update to ensure $\mathbf{w}^T(k)\mathbf{w}(k) = 1$ gives rise to a LMS style stochastic algorithm

$$\mathbf{w}(k+1) = \mathbf{w}(k) + \frac{\mu}{\sqrt{2\pi}\rho_n}\exp\left(-\frac{y^2(k)}{2\rho_n^2}\right) \times \text{sgn}(b_i(k))(\mathbf{r}(k) - \mathbf{w}(k)y(k)) . \qquad (17)$$

The motivation of this LBER MUD is different from that of the AMBER MUD [4]. The latter can be expressed as

$$\mathbf{w}(k+1) = \mathbf{w}(k) + \mu I(k)\text{sgn}(e(k))\mathbf{r}(k) \qquad (18)$$

with $e(k) = b_i(k) - y(k)$ and the indicator function

$$I(k) = \frac{1}{2}\left(1 - \text{sgn}(b_i(k)y(k) - \tau)\right) , \qquad (19)$$

where τ is a nonnegative threshold.

In the AMBER, non-zero τ defines a region around decision boundary where the algorithm will continue to update even when errors do not occur. In the LBER, the effect of the distance from the decision boundary is controlled by an exponential term. This can be viewed as a soft distance measure. The size of an update is a continuous and decreasing function of the distance from the boundary. Both algorithms have a complexity of $O(M)$ with two algorithm parameters that require tuning. Another existing adaptive algorithm, the DMBER [3], also has two tunable algorithm parameters, adaptive step size and differencing step size. The complexity of the DMBER, however, is at least $O(M^2)$.

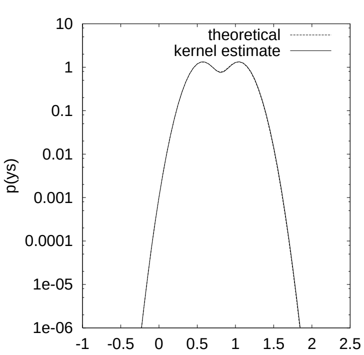

Figure 2: Distribution of the signed decision variable for user 1 of Example 1. $\text{SNR}_1 = \text{SNR}_2 = 16.5$ dB.

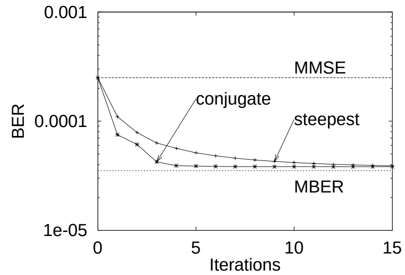

Figure 3: Convergence behaviours of the two block adaptive MBER algorithms for user 1 of Example 1. $\text{SNR}_1 = \text{SNR}_2 = 16.5$ dB.

5. SIMULATION EXAMPLES

Example 1. This was a two-user system with 4 chips per bit. The two code sequences were $(+1, +1, -1, -1)$ and $(+1, -1, -1, +1)$, respectively, and the CIR was

$$H(z) = 1.0 + 0.25z^{-1} + 0.5z^{-3} . \qquad (20)$$

The two users had equal signal power, that is, the user 1 signal to noise ratio SNR_1 was equal to SNR_2 of user 2. The BER difference between the MMSE and MBER solutions for user 1 is significant for the range of SNR_1 14 to 26 dB.

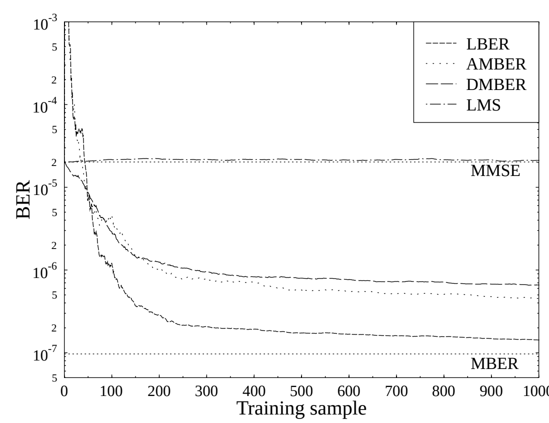

Figure 4: Learning curves of the three stochastic adaptive MBER algorithms and the LMS for user 1 of Example 1. $\text{SNR}_1 = \text{SNR}_2 = 19$ dB.

The kernel density estimate (14) constructed from 100 data samples at $\text{SNR}_1 = \text{SNR}_2 = 16.5$ dB is compared with the true p.d.f. (10) in Fig. 2. Using this constructed kernel density estimate, the block adaptive steepest-descent and conjugate gradient algorithms were applied to find a MBER solution, and the two iterative procedures are illustrated in Fig. 3. The three stochastic algorithms, LBER, AMBER and DMBER, were applied to user 1 with $\text{SNR}_1 = \text{SNR}_2 = 19$ dB. The convergence performance of these three algorithms together with that of the LMS are shown in Fig. 4, where the results were averaged on 100 runs. The two algorithm parameters for each stochastic adaptive MBER algorithm were tuned to give a best combined result of convergence rate and steady-state error. The initial weight vector was set to the MMSE solution.

Example 2. This was a 4-user system with 8 chips per symbol. The four code sequences were $(+1,+1,+1,+1,-1,-1,-1,-1)$, $(+1,-1,+1,-1,-1,+1,-1,+1)$, $(+1,+1,-1,-1,-1,-1,+1,+1)$ and $(+1,-1,-1,+1,-1,+1,+1,-1)$, respectively, and the CIR was

$$H(z) = 0.4 + 0.7z^{-1} + 0.4z^{-2}\,. \qquad (21)$$

The four users had equall power. The LBER, AMBER and DMBER were applied to user 1 with $\text{SNR}_i = 15$ dB, $1 \leq i \leq 4$. The convergence performance of these three algorithms are shown in Fig. 5, where the results were averaged over 50 runs. Again each stochastic adaptive MBER algorithm had its two algorithm parameters tuned to give a best combined result of convergence rate and steady-state error. The initial weight vector was set to the MMSE solution.

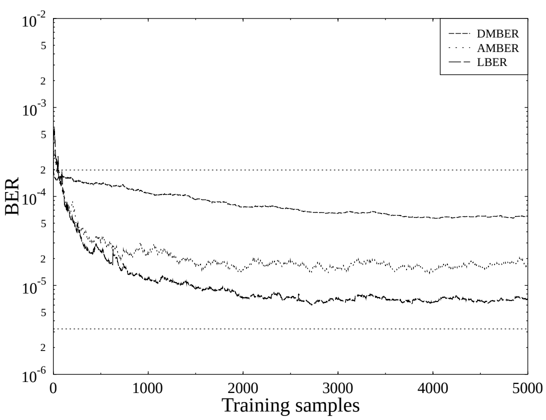

Figure 5: Learning curves of the three stochastic adaptive MBER algorithms for user 1 of Example 2. $\text{SNR}_i = 15$ dB, $1 \leq i \leq 4$.

6. CONCLUSIONS

Motivated from the kernel density estimation of the BER as a smooth function of the training data, block-based adaptive gradient algorithms have been developed to realize the MBER MUD. This has further led to the derivation of a LMS-style LBER MUD. A desired feature of this stochastic gradient algorithm is that the amount of the weight updating is a continuous and decreasing function of the distance from the decision boundary. Simulation results indicate that this adaptive LBER MUD outperforms two existing LMS-style adaptive MBER algorithms.

7. REFERENCES

[1] U. Madhow and M.L. Honig, "MMSE interference suppression for direct-sequence spread-spectrum CDMA," *IEEE Trans. Communications*, Vol.42, No.12, pp.3178–3188, 1994.

[2] G. Woodward and B.S. Vucetic, "Adaptive detection for DS-CDMA," *Proc. IEEE*, Vol.86, No.7, pp.1413–1434, 1998.

[3] I.N. Psaromiligkos, S.N. Batalama and D.A. Pados, "On adaptive minimum probability of error linear filter receivers for DS-CDMA channels," *IEEE Trans. Communications*, Vol.47, No.7, pp.1092–1102, 1999.

[4] C.C. Yeh, R.R. Lopes and J.R. Barry, "Approximate minimum bit-error rate multiuser detection," in *Proc. Globecom'98* (Sydney, Australia), Nov. 1998, pp.3590–3595.

[5] R. Sharma, W.A. Sethares and J.A. Bucklew, "Asymptotic analysis of stochastic gradient-based adaptive filtering algorithms with general cost functions," *IEEE Trans. Signal Processing*, Vol.44, No.9, pp.2186–2194, 1996.

[6] S. Chen, E.S. Chng, B. Mulgrew and G. Gibson, "Minimum-BER linear-combiner DFE," in *Proc. ICC'96* (Dallas, Texas), 1996, Vol.2, pp.1173–1177.

[7] B. Mulgrew and S. Chen, "Adaptive minimum-BER decision feedback equalisers for binary signalling," submitted to *Signal Processing*, 1999.

[8] B. Mulgrew and S. Chen, "Stochastic gradient minimum-BER decision feedback equalisers," in *Proc. IEEE Symp. AS-SPCC* (Lake Louise, Alberta, Canada), Oct.1-4, 2000, pp.93–98.

[9] C.C. Yeh and J.R. Barry, "Approximate minimum bit-error rate equalization for binary signaling," in *Proc. ICC'97* (Montreal, Canada), 1997, Vol.2, pp.1095-1099.

[10] B. Mulgrew, "Nonlinear signal processing for adaptive equalisation and multi-user detection," in *Proc. EUSIPCO 98* (Rhodes, Greece), September 1998, pp.537–544.

[11] M.S. Bazaraa, H.D. Sherali and C.M. Shetty, *Nonlinear Programming: Theory and Algorithms*. New York: John Wiley, 1993.

[12] B.W. Silverman, *Density Estimation*. London: Chapman Hall, 1996.