Adaptive Minimum-BER Linear Multiuser Detection for DS-CDMA Signals in Multipath Channels

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

Abstract

The problem of constructing adaptive minimum bit error rate (MBER) linear multiuser detectors is considered for direct-sequence code division multiple access (DS-CDMA) signals transmitted through multipath channels. Based on the approach of kernel density estimation for approximating the bit error rate (BER) from training data, a least mean squares (LMS) style stochastic gradient adaptive algorithm is developed for training linear multiuser detectors. Computer simulation is used to study the convergence speed and steady-state BER misadjustment of this adaptive MBER linear multiuser detector, and the results show that it outperforms an existing LMS-style adaptive MBER algorithm first presented at Globecom'98 by Yeh, Lopes and Barry.

Keywords

Linear multiuser detectors, minimum mean square error, minimum bit error rate, adaptive algorithms, stochastic gradient algorithms.

I. Introduction

Within the class of linear multiuser detectors for DS-CDMA signals, the minimum mean square error (MMSE) detector [1]-[5] is the most popular one, as it provides good performance and can readily be implemented using standard adaptive filter techniques such as the LMS and recursive least squares algorithms [3],[5]. However, it is well known that the MMSE solution is not always optimal in this application, and the BER of the MMSE linear multiuser detector can in certain situations be distinctly inferior to the optimal MBER solution [6]-[9]. Gradient optimization for obtaining the theoretical MBER linear multiuser detector is considered in [6][1] for narrowband Gaussian CDMA channels (i.e. channels which do not introduce intersymbol interference (ISI)). Gradient optimization to achieve the theoretical MBER solution is also addressed in [7], where constraints are added to ensure a global solution by gradient-based algorithms at the expense that the solution obtained may not be the true MBER solution. There are two stochastic gradient adaptive algorithms for realizing the MBER linear multiuser detector in the literature [8],[9].

SP EDICS: 3-ACCS multiuser and multiaccess communication

S. Chen, A.K. Samingan and L. Hanzo are with the Department of Electronics and Computer Science, the University of Southampton, Highfield, Southampton SO17 1BJ, U.K. (Contact author: S. Chen, Tel./Fax: +44 (0)23 8095 6660/4508; E-mail: sqc@ecs.soton.ac.uk.)

B. Mulgrew is with the Department of Electronics and Electrical Engineering, the University of Edinburgh, King's Buildings, Edinburgh EH9 3JL, U.K.

[1]The "adaptive" algorithm in [6] is in fact non-adaptive, as it requires the received signal minus the noise component. This is the same to an off-line optimization with the perfect channel knowledge and all user bits.

The adaptive algorithm given in [8] uses a difference approximation to estimate the gradient of one-sample error probability and moves the detector weights in the negative direction of the estimated stochastic gradient. The algorithm only adjusts the detector weight vector when the detector makes an error[2]. The main drawback of this algorithm is therefore a very slow convergence, particularly when the error rate is very low. For the sake of distinguishing this stochastic adaptive MBER algorithm from others, it will be called the Difference approximation adaptive MBER (DMBER) algorithm. The adaptive algorithm reported in [9], called the Approximate MBER (AMBER) detector, is appealing due to its computational simplicity. It is a stochastic gradient algorithm that is identical to the signed-error LMS algorithm [10], except in the vicinity of the decision boundary where it is modified to continue updating the weights when the signed-error LMS algorithm would not. The AMBER algorithm therefore can continuously update when the detector weight vector has reached the regions of very low error rate.

Adaptive MBER linear equalisers have been investigated for a longer time [11]-[16]. In particular, the adaptive MBER equaliser presented in [15],[16] is a LMS-style stochastic gradient algorithm and has been shown to outperform the approximate MBER linear equaliser first reported in [17], which is the counterpart of the AMBER linear multiuser detector of [9]. In this study, we extend the adaptive MBER algorithm of [15],[16] to multiuser detection for DS-CDMA channels and develop a new adaptive MBER linear multiuser detector. For the purpose of distinguishing it from the two above-mentioned adaptive algorithms, this new LMS-style stochastic gradient adaptive algorithm will be referred to as the Least BER (LBER) algorithm. Our investigation involving simulation shows that this new LBER linear multiuser detector is superior in performance over the AMBER linear multiuser detector of [9].

The paper is organized as follows. Section II presents the DS-CDMA system model used and provides the necessary notations and definitions. Section III is devoted to formulating the MBER solution for the linear multiuser detector and developing a gradient search algorithm. In Section IV, the proposed adaptive MBER multiuser detector is derived. Kernel density estimation is employed to approximate the BER as a smooth function of training data, and this leads to the formulation of a LMS-style stochastic gradient adaptive algorithm, called the LBER. Comparison with the two existing stochastic adaptive algorithms, the DMBER and the AMBER, is also given in this section. Section V gives some computer simulation results, and the paper concludes in Section VI.

[2]More precisely, when the difference of the detector decisions corresponding to the weights perturbed in opposite directions is nonzero.

II. System model

The discrete-time baseband model of the synchronous DS-CDMA downlink system with N users and M chips per symbol is depicted in Fig. 1, where $b_i(k) \in \{\pm 1\}$ denotes the k-th symbol of user i, the unit-length signature sequence for user i is

$$\bar{\mathbf{s}}_i = [\bar{s}_{i,1} \ \cdots \ \bar{s}_{i,M}]^T \,, \tag{1}$$

and the transfer function of the channel impulse response (CIR) is

$$H(z) = \sum_{i=0}^{n_h-1} h_i z^{-i} \,. \tag{2}$$

The baseband model for received signal sampled at chip rate is given by [4],[18]:

$$\begin{aligned} \mathbf{r}(k) &= \mathbf{H} \begin{bmatrix} \bar{\mathbf{S}}\mathbf{A} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \bar{\mathbf{S}}\mathbf{A} & \ddots & \vdots \\ \vdots & \ddots & \ddots & \mathbf{0} \\ \mathbf{0} & \cdots & \mathbf{0} & \bar{\mathbf{S}}\mathbf{A} \end{bmatrix} \begin{bmatrix} \mathbf{b}(k) \\ \mathbf{b}(k-1) \\ \vdots \\ \mathbf{b}(k-L+1) \end{bmatrix} + \mathbf{n}(k) \\ &= \mathbf{P} \begin{bmatrix} \mathbf{b}(k) \\ \mathbf{b}(k-1) \\ \vdots \\ \mathbf{b}(k-L+1) \end{bmatrix} + \mathbf{n}(k) = \bar{\mathbf{r}}(k) + \mathbf{n}(k) \,, \end{aligned} \tag{3}$$

where $\bar{\mathbf{r}}(k)$ denotes the noise-free received signal; the white Gaussian noise vector

$$\mathbf{n}(k) = [n_1(k) \ \cdots \ n_M(k)]^T \quad \text{with} \quad E[\mathbf{n}(k)\mathbf{n}^T(k)] = \sigma_n^2 \mathbf{I} \,; \tag{4}$$

the user symbol vector

$$\mathbf{b}(k) = [b_1(k) \ \cdots \ b_N(k)]^T \,; \tag{5}$$

the user signature sequence matrix

$$\bar{\mathbf{S}} = [\bar{\mathbf{s}}_1 \ \cdots \ \bar{\mathbf{s}}_N] \,; \tag{6}$$

the diagonal user signal amplitude matrix

$$\mathbf{A} = \mathrm{diag}\{A_1 \ \cdots \ A_N\} \,, \tag{7}$$

that is, A_i^2 is the user i signal power; the $M \times LM$ CIR matrix $\mathbf{H}$ has the form

$$\mathbf{H} = \begin{bmatrix} h_0 & h_1 & \cdots & h_{n_h-1} & & & \\ & h_0 & h_1 & \cdots & h_{n_h-1} & & \\ & & \ddots & \ddots & \cdots & \ddots & \\ & & & h_0 & h_1 & \cdots & h_{n_h-1} \end{bmatrix} ; \tag{8}$$

and the $M \times LN$ system matrix

$$\mathbf{P} = \mathbf{H} \begin{bmatrix} \bar{\mathbf{S}}\mathbf{A} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \bar{\mathbf{S}}\mathbf{A} & \ddots & \vdots \\ \vdots & \ddots & \ddots & \mathbf{0} \\ \mathbf{0} & \cdots & \mathbf{0} & \bar{\mathbf{S}}\mathbf{A} \end{bmatrix} . \tag{9}$$

The channel ISI span L depends on the length of the CIR, n_h, related to the length of the chip sequence, M. 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. The model (3) adopted in this study is a wideband synchronous DS-CDMA channel and it can be extended to the case of asynchronous DS-CDMA downlink systems.

We will study the linear multiuser detector of the form:

$$\hat{b}_i(k) = \text{sgn}(y(k;\mathbf{w})) \quad \text{with} \quad y(k;\mathbf{w}) = \mathbf{w}^T \mathbf{r}(k) \,, \tag{10}$$

where

$$\mathbf{w} = [w_1 \ \cdots \ w_M]^T \tag{11}$$

denotes the detector weight vector for user i. Let the $N_b = 2^{LN}$ possible combinations or sequences of $[\mathbf{b}^T(k) \ \mathbf{b}^T(k-1) \ \cdots \ \mathbf{b}^T(k-L+1)]^T$ be

$$\mathbf{b}^{(j)} = \begin{bmatrix} \mathbf{b}^{(j)}(k) \\ \mathbf{b}^{(j)}(k-1) \\ \vdots \\ \mathbf{b}^{(j)}(k-L+1) \end{bmatrix} , \ 1 \leq j \leq N_b \,, \tag{12}$$

and $b_i^{(j)}$ the ith element of $\mathbf{b}^{(j)}(k)$. Define the N_b noise-free received signal states

$$\mathbf{r}_j = \mathbf{P}\mathbf{b}^{(j)}, \ 1 \leq j \leq N_b \,, \tag{13}$$

and the set of N_b scalars

$$y_j = \mathbf{w}^T \mathbf{r}_j, \ 1 \leq j \leq N_b \,. \tag{14}$$

III. The MBER linear multiuser detector

In this section, we derive the theoretical MBER solution for the linear multiuser detector and present a gradient search algorithm for finding this MBER solution. This will provide some insight into the development of our LMS-style stochastic gradient adaptive algorithm. The error probability of the linear detector (10) is:

$$P_E(\mathbf{w}) = \text{Prob}\{\text{sgn}(b_i(k))y(k;\mathbf{w}) < 0\} \,. \tag{15}$$

Following [15],[16], define the signed decision variable

$$y_s(k;\mathbf{w}) = \mathrm{sgn}(b_i(k))y(k;\mathbf{w}) = \mathrm{sgn}(b_i(k))y'(k;\mathbf{w}) + n'(k)\,, \tag{16}$$

where

$$y'(k;\mathbf{w}) = \mathbf{w}^T\mathbf{P}\begin{bmatrix}\mathbf{b}(k)\\ \mathbf{b}(k-1)\\ \vdots \\ \mathbf{b}(k-L+1)\end{bmatrix} \tag{17}$$

and

$$n'(k) = \mathrm{sgn}(b_i(k))\mathbf{w}^T\mathbf{n}(k)\,. \tag{18}$$

Notice that $y'(k;\mathbf{w})$ can only take the values from the set defined in (14) and $n'(k)$ is Gaussian with zero mean and variance $\sigma_n^2\mathbf{w}^T\mathbf{w}$. Under the assumption that $\mathbf{r}_j$, $1 \le j \le N_b$, are equiprobable, the probability density function (p.d.f.) of $y_s(k;\mathbf{w})$ is

$$p_y(y_s;\mathbf{w}) = \frac{1}{N_b\sqrt{2\pi}\sigma_n\sqrt{\mathbf{w}^T\mathbf{w}}}\sum_{j=1}^{N_b}\exp\left(-\frac{(y_s - \mathrm{sgn}(b_i^{(j)})y_j)^2}{2\sigma_n^2\mathbf{w}^T\mathbf{w}}\right)\,. \tag{19}$$

Thus

$$P_E(\mathbf{w}) = \int_{-\infty}^{0} p_y(y_s;\mathbf{w})\,dy_s = \frac{1}{N_b}\sum_{j=1}^{N_b} Q\left(c_j(\mathbf{w})\right)\,, \tag{20}$$

where

$$Q(y) = \frac{1}{\sqrt{2\pi}}\int_y^{\infty}\exp\left(-\frac{x^2}{2}\right)\,dx \tag{21}$$

and

$$c_j(\mathbf{w}) = \frac{\mathrm{sgn}(b_i^{(j)})y_j}{\sigma_n\sqrt{\mathbf{w}^T\mathbf{w}}} = \frac{\mathrm{sgn}(b_i^{(j)})\mathbf{w}^T\mathbf{r}_j}{\sigma_n\sqrt{\mathbf{w}^T\mathbf{w}}}\,. \tag{22}$$

The gradient of $P_E(\mathbf{w})$ with respect to $\mathbf{w}$ is

$$\nabla P_E(\mathbf{w}) = \frac{1}{N_b\sqrt{2\pi}\sigma_n}\left(\frac{\mathbf{w}\mathbf{w}^T - \mathbf{w}^T\mathbf{w}\mathbf{I}}{(\mathbf{w}^T\mathbf{w})^{\frac{3}{2}}}\right)\sum_{j=1}^{N_b}\exp\left(-\frac{y_j^2}{2\sigma_n^2\mathbf{w}^T\mathbf{w}}\right)\mathrm{sgn}(b_i^{(j)})\mathbf{r}_j\,. \tag{23}$$

The following steepest-descent gradient algorithm can be used to find the MBER solution

$$\mathbf{w}(l+1) = \mathbf{w}(l) - \mu\nabla P_E(\mathbf{w}(l))\,, \tag{24}$$

where l indicates the iteration and μ is an adaptive step size. Notice that the orientation of the weight vector $\mathbf{w}$ defines the decision boundary and thus the BER, not the size of $\mathbf{w}$. It is computationally advantageous to normalize $\mathbf{w}$ to a unit-length after every iteration:

$$\mathbf{w} = \frac{\mathbf{w}}{\sqrt{\mathbf{w}^T\mathbf{w}}}\,. \tag{25}$$

With this re-scaling, the gradient can be simplified as

$$\nabla P_E(\mathbf{w}) = \frac{1}{N_b\sqrt{2\pi}\sigma_n}\sum_{j=1}^{N_b}\exp\left(-\frac{y_j^2}{2\sigma_n^2}\right)\mathrm{sgn}(b_i^{(j)})(\mathbf{w}y_j - \mathbf{r}_j)\,. \tag{26}$$

The steepest-descent gradient algorithm (24) may converge slowly and a Gauss-Newton algorithm is computationally expensive. The conjugate gradient method [19] offers a better alternative. A simplified conjugate gradient algorithm is summarized:

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

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

$$\begin{aligned}
\mathbf{w}(l+1) &= \mathbf{w}(l) + \mu\mathbf{d}(l)\\
\mathbf{w}(l+1) &= \frac{\mathbf{w}(l+1)}{\|\mathbf{w}(l+1)\|}\\
\phi_l &= \frac{\|\nabla P_E(\mathbf{w}(l+1))\|^2}{\|\nabla P_E(\mathbf{w}(l))\|^2}\\
\mathbf{d}(l+1) &= \phi_l\mathbf{d}(l) - \nabla P_E(\mathbf{w}(l+1))
\end{aligned}$$

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

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

At a local minimum, $\|\nabla P_E(\mathbf{w})\| = 0$. So the small positive scalar β determines the accuracy of the solution obtained. A gradient algorithm can in general find a local minimum of $P_E(\mathbf{w})$. To ensure a unique global solution, certain constraints can be added to the optimization problem, as shown in [7], at a possible cost that the solution found may deviate slightly from the true MBER solution.

In general, the global MBER solutions are infinitely many, and they form a half hyperplane in the M-dimensional weight space. The normalization (25) has an effect of fixing the solution to a unique unit-length one. Consider the following simplest example with two equal-power users and two chips per bit. The two chip codes are $(+1, +1)$ and $(+1, -1)$, respectively, and the transfer function of the CIR is $H(z) = 1 + 0.8z^{-1} + 0.6z^{-2}$. The signal to noise ratio for user 1, SNR_1, is 25 dB. The $\log_{10}(BER)$ surface for user 1 is plotted in Fig. 2. For this example, there are only global minimum solutions, and all the MBER solutions form a half line (half hyperplane in the two-dimensional space). The point marked in the MBER solution half line is the unit-length one. The MMSE solution for this example is also depicted in Fig. 2, where it can be seen that the MMSE solution is very different from the MBER ones. In fact, for the MMSE solution, $\log_{10}(BER) = -3.88$, while for a MBER one, $\log_{10}(BER) = -5.56$.

IV. Adaptive MBER linear multiuser detector

The key to developing an effective adaptive algorithm is the p.d.f. $p_y(y_s; \mathbf{w})$ of the signed decision variable. Kernel density estimation is known to produced reliable p.d.f. estimates with short data records and in particular extremely natural in dealing with Gaussian mixtures [20],[21]. Given a block of K training samples $\{\mathbf{r}(k), b_i(k)\}$, a kernel density estimate of the p.d.f. is given by:

$$\hat{p}_y(y_s; \mathbf{w}) = \frac{1}{K\sqrt{2\pi}\rho_n\sqrt{\mathbf{w}^T\mathbf{w}}} \sum_{k=1}^{K} \exp\left(-\frac{(y_s - \mathrm{sgn}(b_i(k))y(k; \mathbf{w}))^2}{2\rho_n^2 \mathbf{w}^T\mathbf{w}}\right) , \tag{27}$$

where the radius parameter ρ_n is related to the noise standard deviation σ_n. A lower bound of ρ_n is [20]:

$$\rho_n = \left(\frac{4}{3K}\right)^{\frac{1}{5}} \sigma_n \, . \tag{28}$$

From the estimated error probability

$$\hat{P}_E(\mathbf{w}) = \int_{-\infty}^{0} \hat{p}_y(y_s; \mathbf{w}) \, dy_s \, , \tag{29}$$

$\nabla \hat{P}_E(\mathbf{w})$ can be calculated:

$$\nabla \hat{P}_E(\mathbf{w}) = \frac{1}{\sqrt{2\pi}\rho_n} \left(\frac{\mathbf{w}\mathbf{w}^T - \mathbf{w}^T\mathbf{w}\mathbf{I}}{(\mathbf{w}^T\mathbf{w})^{\frac{3}{2}}}\right) \sum_{k=1}^{K} \exp\left(-\frac{y^2(k; \mathbf{w})}{2\rho_n^2 \mathbf{w}^T\mathbf{w}}\right) \mathrm{sgn}(b_i(k))\mathbf{r}(k) \, . \tag{30}$$

Thus block adaptive gradient algorithms can readily be developed by substituting $\nabla P_E(\mathbf{w})$ with $\nabla \hat{P}_E(\mathbf{w})$ in the steepest-descent or conjugate gradient updating mechanisms.

A. LBER stochastic adaptive algorithm

Our aim is however to develop a LMS-style adaptive algorithm with sample-by-sample adjustment, as in [15],[16]. The LMS algorithm can be viewed as replacing the ensemble average of the gradient in its related steepest-descent gradient algorithm by a single data-point estimate of the gradient. In a similar manner, at sample k, a point estimate of the p.d.f. is simply:

$$\hat{p}_y(y_s, k; \mathbf{w}) = \frac{1}{\sqrt{2\pi}\rho_n\sqrt{\mathbf{w}^T\mathbf{w}}} \exp\left(-\frac{(y_s - \mathrm{sgn}(b_i(k))y(k; \mathbf{w}))^2}{2\rho_n^2 \mathbf{w}^T\mathbf{w}}\right) . \tag{31}$$

Using the instantaneous or stochastic gradient

$$\nabla \hat{P}_E(k; \mathbf{w}) = \frac{1}{\sqrt{2\pi}\rho_n} \left(\frac{\mathbf{w}\mathbf{w}^T - \mathbf{w}^T\mathbf{w}\mathbf{I}}{(\mathbf{w}^T\mathbf{w})^{\frac{3}{2}}}\right) \exp\left(-\frac{y^2(k; \mathbf{w})}{2\rho_n^2 \mathbf{w}^T\mathbf{w}}\right) \mathrm{sgn}(b_i(k))\mathbf{r}(k) \tag{32}$$

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

$$\mathbf{w}(k+1) = \mathbf{w}(k) + \frac{\mu}{\sqrt{2\pi}\rho_n} \exp\left(-\frac{y^2(k; \mathbf{w}(k))}{2\rho_n^2}\right) \mathrm{sgn}(b_i(k))(\mathbf{r}(k) - \mathbf{w}(k)y(k; \mathbf{w}(k))) \, , \tag{33}$$

which we refer to as the LBER algorithm.

Two important issues for any stochastic gradient adaptive MBER algorithm are the convergence speed and steady-state BER misadjustment (with respect to the optimal MBER). A theoretical analysis of these two properties for the LBER (33) is extremely complex and is still under investigation. We will use computer simulation to study these two properties.

B. Comparison with the DMBER and AMBER adaptive algorithms

The motivation of the above LBER multiuser detector is different from those of the existing DMBER and AMBER detectors [8],[9]. For the purpose of a comparison, we present a modified version of the DMBER adaptive algorithm reported in [8]. Define the one-sample decision distortion measure[3]

$$\zeta(k;\mathbf{w}) = \frac{1}{2}|b_i(k) - \hat{b}_i(k)| \,. \tag{34}$$

Notice that this distortion measure is a function of $\mathbf{w}$, as $\hat{b}_i(k) = \text{sgn}(\mathbf{w}^T\mathbf{r}(k))$. Obviously, $\zeta(k;\mathbf{w}) = \hat{P}_E(k;\mathbf{w})$ is the one-sample error probability. The following difference approximation for the gradient of $\zeta(k;\mathbf{w})$ is adopted:

$$\mathbf{x}(k;\mathbf{w}) = [x_1(k;\mathbf{w}) \ \cdots \ x_M(k;\mathbf{w})]^T \tag{35}$$

with

$$x_j(k;\mathbf{w}) = \frac{1}{2c_k}\left(\zeta(k;\mathbf{w}+c_k\mathbf{e}_j) - \zeta(k;\mathbf{w}-c_k\mathbf{e}_j)\right) \,, \quad 1 \le j \le M \,, \tag{36}$$

where $c_k = c_0 k^{-1/4}$ for some $c_0 > 0$ and $\mathbf{e}_j$ denotes the jth coordinate unit vector. The DMBER adaptive algorithm takes the form:

$$\mathbf{w}(k+1) = \mathbf{w}(k) - \alpha\mathbf{x}(k;\mathbf{w}(k)) \,, \tag{37}$$

where α is an adaptive step size.

In theory, the DMBER should work, as it attempts to minimize the BER directly. As a difference approximation of the stochastic gradient is used, the algorithm does not rely on the assumption of a Gaussian channel noise at a cost of increased complexity (The DMBER has a complexity of $O(M^2)$ while the AMBER and LBER has a complexity of $O(M)$.). However, in practice, its rate of convergence is very slow after the weight vector has reached the region of small error rate. This is obvious, since $\mathbf{x}(k;\mathbf{w})$ will be zero most of the time.

[3]In [8], the distortion measure is defined in terms of two samples, one for $b_i(k) = -1$ and the other for $b_i(k+1) = 1$, which lead to some complications in sample-by-sample adaptation. Our modification has a nature and simple adaptive implementation, and is equivalent to the original formulation.

A more efficient stochastic gradient adaptive algorithm is the AMBER [9], which can be expressed as

$$\mathbf{w}(k+1) = \mathbf{w}(k) + \mu I(k;\mathbf{w}(k))\mathrm{sgn}(e(k;\mathbf{w}(k)))\mathbf{r}(k) \tag{38}$$

with the error signal

$$e(k;\mathbf{w}) = b_i(k) - y(k;\mathbf{w}) \tag{39}$$

and the indicator function

$$I(k;\mathbf{w}) = \frac{1}{2}\left(1 - \mathrm{sgn}(b_i(k)y(k;\mathbf{w}) - \tau)\right) , \tag{40}$$

where τ is a nonnegative threshold parameter. In terms of the algorithm tuning requirements, the two adaptive algorithms, the LBER (33) and AMBER (38), are similar. The former requires the tuning of the adaptive gain μ and kernel width ρ_n, while the latter needs the tuning of the adaptive gain μ and threshold parameter τ.

The following comparison of the two adaptive mechanisms, the LBER and AMBER, can be made [15],[16]. The algorithm (38) in its simplest form has $\tau = 0$ and the algorithm only updates when a decision error is observed. When the algorithm is initialized, it is unlikely to separate all the noise-free states $\mathbf{r}_j$ correctly. Thus the indicator function $I(k;\mathbf{w})$ will be on most of the time, in which case it is equivalent to the signed-error LMS algorithm [10]. When the algorithm has converged to a point where it can separate the noise-free states correctly, the probability of the algorithm updating may be low, as in this region errors will be predominated by noise, and hence further convergence may be slow. Introducing the threshold τ essentially defines a region around decision boundary where the algorithm will continue to update even when errors do not occur. This region is defined by:

$$|\mathbf{w}^T\mathbf{r}(k)| < \tau\,. \tag{41}$$

In the algorithm (33), the effect of the distance from the decision boundary is controlled by the exponential term $\exp(-y^2(k;\mathbf{w})/(2\rho_n^2))$. 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. The distance is scaled by the kernel width ρ_n which in turn is a function of the noise root mean square σ_n.

V. Simulation results

Computer simulation was conducted to investigate the convergence speed and steady-state BER misadjustment for the three stochastic adaptive algorithms. For the DMBER algorithm, a fixed step size α was used with a time-varying difference step $c_k = c_0 k^{-1/4}$ (as was used in [8]). For the AMBER algorithm, the threshold τ was fixed with a time-varying adaptive step size μ given by

$$\mu_k = \mu_0 k^{-1} . \tag{42}$$

The LBER algorithm had a constant width ρ_n and employed the time-varying adaptive step size as given by (42). The two algorithm parameters, c_0 and α for the DMBER, μ_0 and τ for the AMBER, and μ_0 and ρ_n for the LBER, were chosen to give an adequate combined result of convergence rate and steady-state error for the respective algorithm.

Example 1. A two-user system with 4 chips per symbol was used in the simulation. The code sequences of the two users were $(+1, +1, -1, -1)$ and $(+1, -1, -1, +1)$, respectively, and the transfer function of the CIR was

$$H(z) = 1.0 + 0.25z^{-1} + 0.5z^{-3} . \tag{43}$$

The two users had equal signal power, that is, SNR_1 was equal to SNR_2. Figs. 3 and 4 depict the linear detector BERs for the two users, respectively. The BER formula (20) was used with the detector weight vector set to the MMSE and MBER solutions, respectively, to produce the corresponding error rate curves. For this example, the difference between the MMSE and MBER solutions for user 1 is significant for the range of SNR_1 14 to 26 dB.

The kernel density estimate (27) constructed from 100 data samples at $\mathrm{SNR}_1 = \mathrm{SNR}_2 = 16.5$ dB is compared with the true p.d.f. (19) in Fig. 5 for user 1, where the detector weight vector was set to the MBER solution. For this example, the kernel estimate and the true density are indistinguishable. Using the constructed kernel density estimate with the detector weight vector $\mathbf{w}$, the block adaptive steepest-descent and conjugate gradient algorithms were applied to find the MBER solutions, and the two iterative procedures are illustrated in Fig. 6. It can be seen that, starting from the MMSE solution, the block adaptive conjugate gradient algorithm took 4 iterations to converge to the MBER solution.

The three stochastic gradient adaptive algorithms, the DMBER, AMBER and LBER, were applied to user 1 with $\mathrm{SNR}_1 = \mathrm{SNR}_2 = 19$ dB and the initial weight vector set to the MMSE solution. For the DMBER, $c_0 = 2.0$ and $\alpha = 0.01$ was used. The AMBER used $\mu_0 = 0.5$ and $\tau = 0.8$, while the LBER had $\mu_0 = 1.0$ and $\rho_n^2 = 0.11$ ($\approx 9\times$noise variance). These algorithm parameters were found in simulation to be adequate for the respective algorithm. The convergence performance of these three algorithms are shown in Fig. 7, where the results were averaged on 100 runs. The standard LMS with an adaptive gain 0.01 is also depicted in Fig. 7. The value of the learning curve at k is the true BER $P_E(\mathbf{w}(k))$ of (20) for the given weight vector $\mathbf{w}(k)$, not any approximation. For this example, it can be seen that the proposed LBER algorithm is superior over the other two adaptive MBER algorithms in terms of convergence speed and steady-state error.

Example 2. This was a 4-user system with 8 chips per symbol. The code sequences for the four users 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 transfer function of the CIR was

$$H(z) = 0.4 + 0.7z^{-1} + 0.4z^{-2} . \tag{44}$$

The four users had equal signal power. Figs. 8 to 10 depict the linear detector BERs for the four users, respectively. For users 2 and 4, the MMSE and MBER solutions are indistinguishable.

Fig. 11 shows the kernel estimate and true p.d.f. of the signed decision variable for user 1 at $\text{SNR}_i = 16$ dB, $1 \le i \le 4$. The p.d.f. estimate was constructed from 1500 data samples, and the detector weight vector was set to the MBER solution. It can be seen that, for this example, the kernel estimate approximates the true density reasonably well. Based on the kernel density estimate, the iterative procedures of the block adaptive steepest-descent and conjugate gradient algorithms are depicted in Fig. 12. It can be seen that, starting from the MMSE solution, the block adaptive conjugate gradient algorithm took 5 iterations to converge to a near MBER solution.

The three stochastic gradient adaptive algorithms were tested for user 1 at $\text{SNR}_i = 15$ dB, $1 \le i \le 4$ with the initial weight vector set to the MMSE solution. The two algorithm parameters were found to be $c_0 = 2.0$ and $\alpha = 0.01$ for the DMBER, $\mu_0 = 0.05$ and $\tau = 0.8$ for the AMBER. For this example, the LBER used $c_k = 0.2k^{-1/4}$ and $\rho_n^2 = 0.06$ ($\approx 2\times$noise variance). The learning curves of these three algorithms are given in Fig. 13, where the results were averaged on 50 runs. For this example, the very slow convergence of the DMBER is apparent when the BER is below 10^{-4}. It can also be seen that the LBER algorithm (33) has a superior performance over the AMBER algorithm (38).

VI. Conclusions

The MBER linear multiuser detection has been considered for DS-CDMA systems. 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 realized the MBER linear multiuser detector. This has further led to the derivation of a LMS-style adaptive MBER algorithm for linear multiuser detectors. 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 MBER detector outperforms an existing LMS-style MBER algorithm, called the AMBER, in terms of convergence speed and steady-state BER misadjustment. Theoretical analysis of convergence and steady-state misadjustment for the proposed new adaptive MBER algorithm is still under investigation.

References

[1] Z. Xie, R.T. Short and C.K. Rushforth, "A family of suboptimum detectors for coherent multiuser communications," *IEEE J. Selected Areas in Communications*, Vol.8, No.4, pp.683–690, 1990.

[2] 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.

[3] S.L. Miller, "An adaptive direct-sequence code-division multiple-access receiver for multiuser interference rejection," *IEEE Trans. Communications*, Vol.43, No.2/3/4, pp. 1746–1755, 1995.

[4] H.V. Poor and S. Verdú, "Probability of error in MMSE multiuser detection," *IEEE Trans. Information Theory*, Vol.43, No.3, pp.858–871, 1997.

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

[6] N.B. Mandayam and B. Aazhang, "Gradient estimation for sensitivity analysis and adaptive multiuser interference rejection in code-division multi-access systems," *IEEE Trans. Communications*, Vol.45, No.7, pp.848–858, 1997.

[7] X.F. Wang, W.S. Lu and A. Antoniou, "Constrained minimum-BER multiuser detection," in *Proc. ICASSP* (Phoenix, USA), May 14-18, 1999, Vol.5, pp.2603–2606.

[8] 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.

[9] 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.

[10] 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.

[11] E. Shamash and K. Yao, "On the structure and performance of a linear decision feedback equalizer based on the minimum error probability criterion," in *Proc. ICC'74*, 1974, pp.25F1–25F5.

[12] 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.

[13] S. Chen, B. Mulgrew, E.S. Chng and G. Gibson, "Space translation properties and the minimum-BER linear-combiner DFE," *IEE Proc. Communications*, Vol.145, No.5, pp.316–322, 1998.

[14] S. Chen and B. Mulgrew, "The minimum-SER linear-combiner decision feedback equalizer," *IEE Proc. Communications*, Vol.146, No.6, pp.347–353, 1999.

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

[16] B. Mulgrew and S. Chen, "Stochastic gradient minimum-BER decision feedback equalisers," in *Proc. IEEE Symposium on Adaptive Systems for Signal Processing, Communication and Control* (Lake Louise, Alberta, Canada), Oct.1-4, 2000, pp.93–98.

[17] 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.

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

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

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

[21] A.W. Bowman and A. Azzalini, *Applied Smoothing Techniques for Data Analysis.* Oxford University Press, 1997.

Sheng Chen (SM'97) received the PhD degree in control engineering from the City University, London in 1986. From 1986 to 1999, he held various research and academic appointments at the Universities of Sheffield, Edinburgh and Portsmouth, UK. Since 1999, he has been with the Department of Electronics and Computer Science, the University of Southampton, UK, where he currently holds a post of Readership in Communications Signal Processing. Dr. Chen is a Senior Member of the Institute of Electrical and Electronics Engineers. His research interests include modelling and identification of nonlinear systems, adaptive nonlinear signal processing, artificial neural network research, finite-precision digital controller design, evolutionary computation methods and optimization. He has published over 140 research papers.

Ahmad Kamsani Samingan received the B.Eng. (Hons.) degree in electronic engineering from the University of Southampton, UK in 1999 with an exceptional First Class honours. He was named a Zepler Prize winner in 1998 due to his excellent academic achievement. He is currently with the Communication group, the Department of Electronic and Computer Science, the University of Southampton, and is studying for his PhD degree which is due to complete in 2002. His research interest covers the field of CDMA multiuser detection, particularly in the area of adaptive multiuser detection and Support Vector Machine (SVM) based multiuser detection.

Bernard Mulgrew received the B.Sc. degree in Electrical & Electronic Engineering in 1979 from Queen's University Belfast. After graduation, he worked for 4 years as a Development Engineer in the Radar Systems Department at GEC-Marconi Avionics, Edinburgh. From 1983-1986 he was a research associate in the Department of Electrical Engineering at the University of Edinburgh. He became a lecturer in 1986 and received his Ph.D. in 1987. Promotion to Senior Lectureship and Readership followed in 1994 and 1996 respectively. He was elected to the Personal Chair in Signals and Systems in 1999. His research interests are in adaptive signal processing and estimation theory and in their application to radar, audio and communications systems. Professor Mulgrew is a co-author of three books on signal processing, and has published over 150 research papers. He is a member of the IEEE, a member of EURASIP, a member of the IEE and a member of the Audio Engineering Society.

Lajos Hanzo has held various research and academic posts in Hungary, Germany and the UK. Since 1986 he has been with the Department of Electronics and Computer Science, University of Southampton, UK, where he currently holds the chair in Telecommunications. He co-authored five books on mobile radio communications, published over 200 research papers, organised and chaired conference sessions, presented overview lectures and was awarded a number of distinctions. Currently he is managing a research team, working on a range of research projects in the field of wireless multimedia communications under the auspices of the Engineering and Physical Sciences Research Council (EPSRC) UK, the European IST Programme and the Mobile Virtual Centre of Excellence (VCE). For further information on research in progress and associated publications please refer to http://www-mobile.ecs.soton.ac.uk.

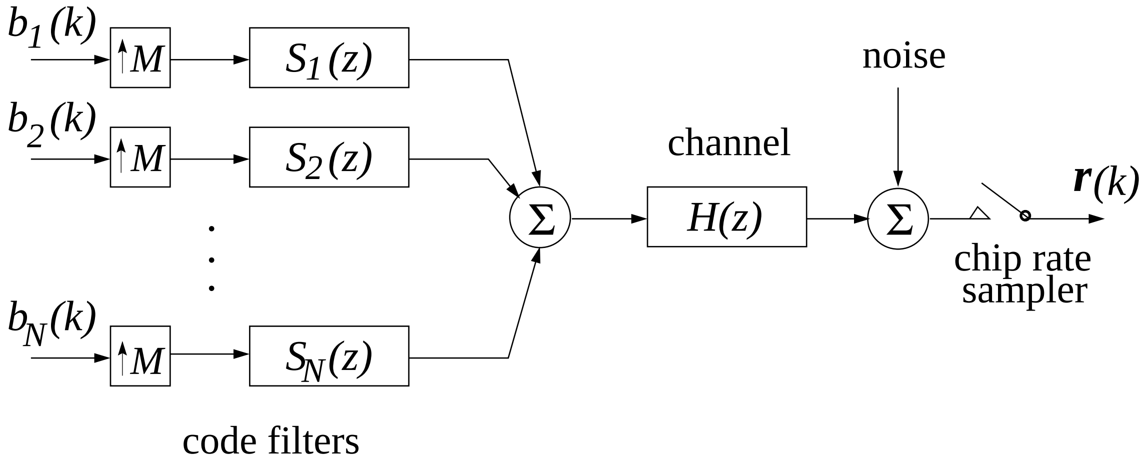

Fig. 1. Discrete-time model of synchronous CDMA downlink.

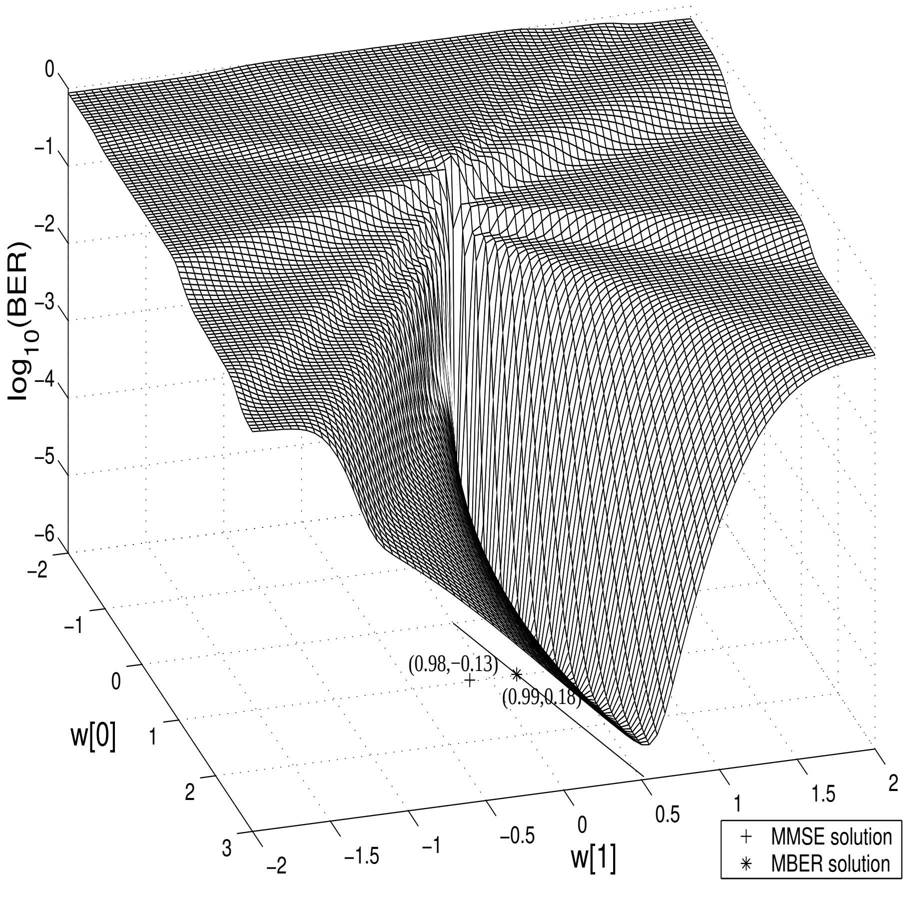

Fig. 2. Bit error rate surface of user-one detector for a simple two-user system with two chips per bit and SNR_1=25 dB.

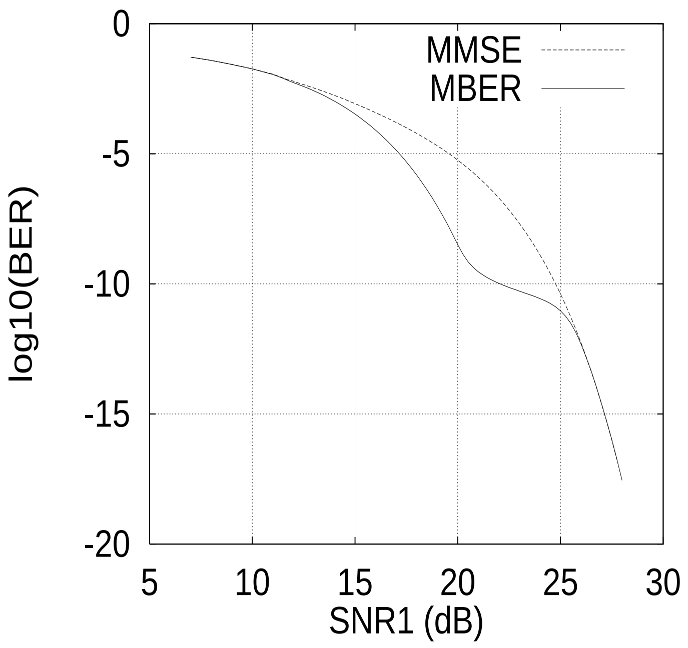

Fig. 3. Linear detector BERs for user 1 of Example 1. $\mathrm{SNR}_1 = \mathrm{SNR}_2$.

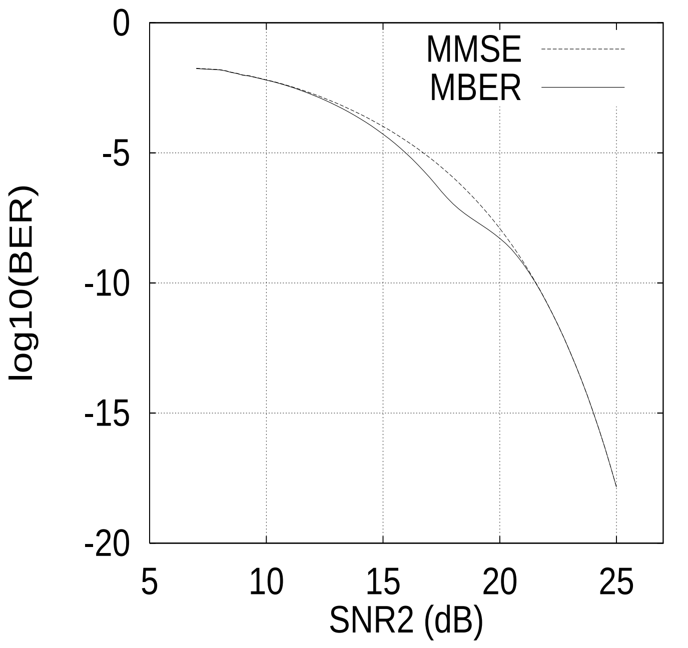

Fig. 4. Linear detector BERs for user 2 of Example 1. $\mathrm{SNR}_1 = \mathrm{SNR}_2$.

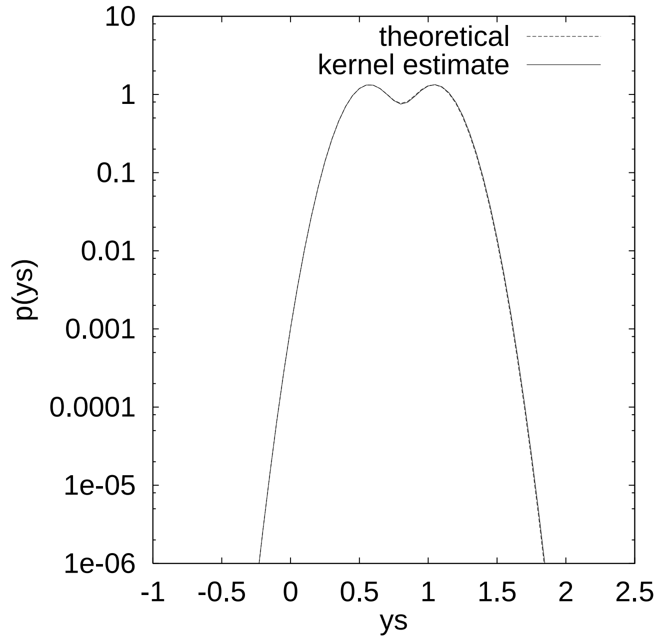

Fig. 5. Distribution of the signed decision variable for user 1 of Example 1. $\mathrm{SNR}_1 = \mathrm{SNR}_2 = 16.5$ dB, the detector weight vector was set to the MBER solution, and the kernel estimate was constructed from 100 data samples.

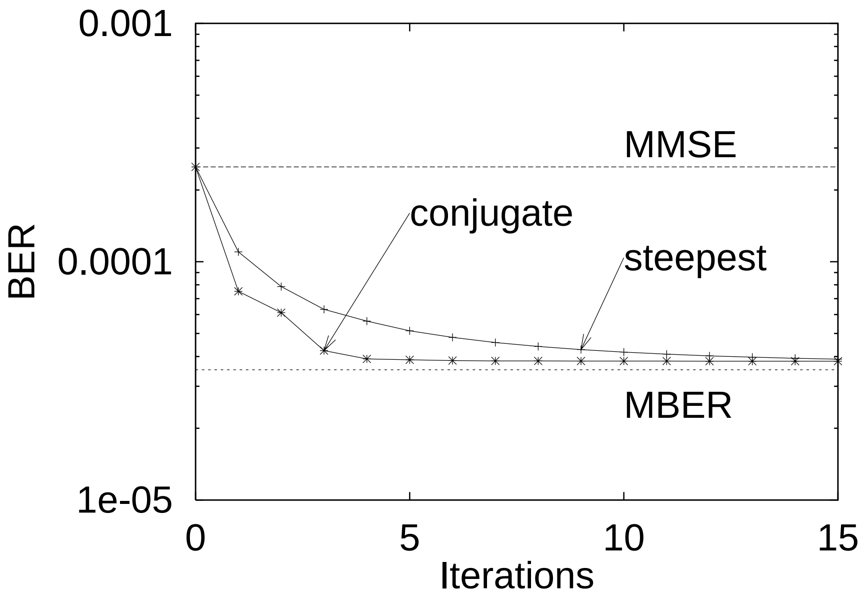

Fig. 6. Convergence behaviours of the two block adaptive MBER algorithms for user 1 of Example 1. Length of the data block is 100, and $\mathrm{SNR}_1 = \mathrm{SNR}_2 = 16.5$ dB.

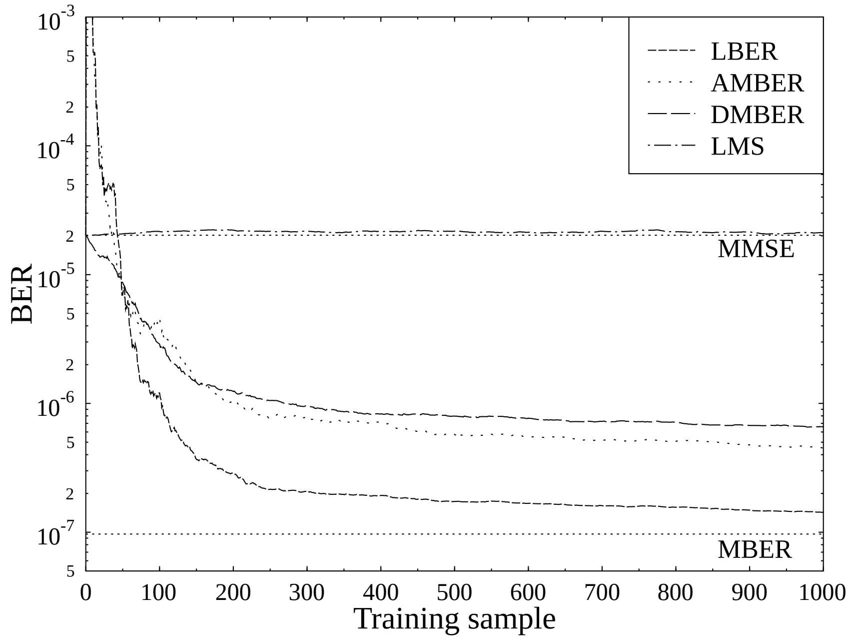

Fig. 7. Learning curves of the three stochastic adaptive MBER algorithms and the LMS algorithm for user 1 of Example 1. $\mathrm{SNR}_1 = \mathrm{SNR}_2 = 19$ dB.

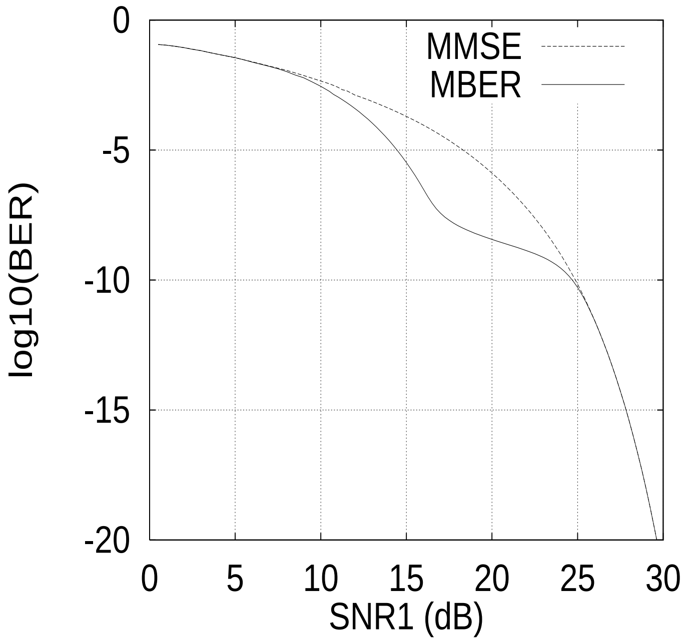

Fig. 8. Linear detector BERs for user 1 of Example 2. SNR_i, $1 \le i \le 4$, are identical.

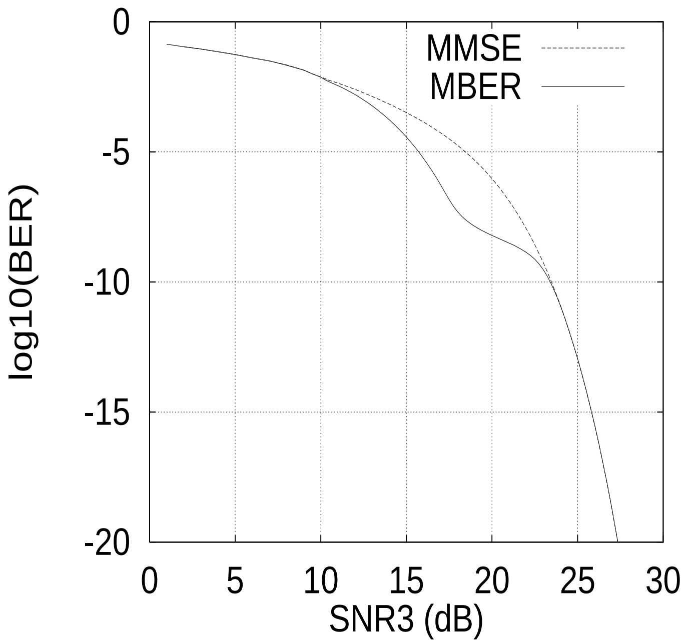

Fig. 9. Linear detector BERs for user 3 of Example 2. SNR_i, $1 \le i \le 4$, are identical.

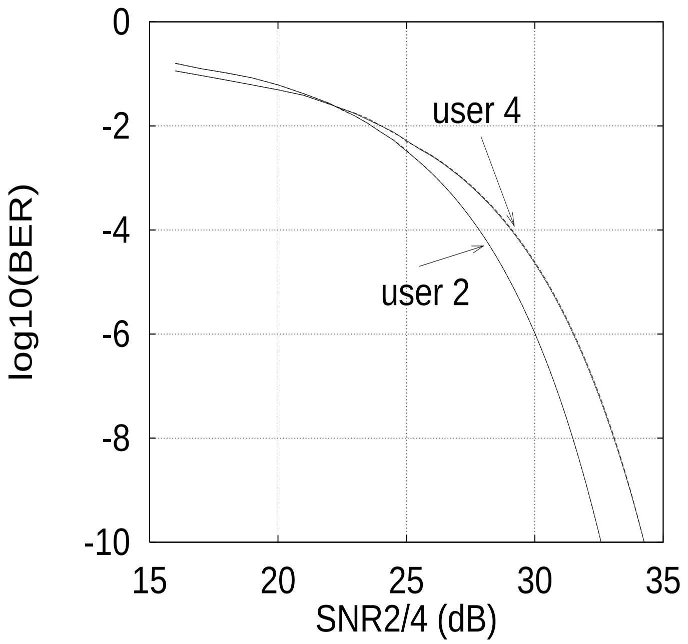

Fig. 10. Linear detector BERs for users 2 and 4 of Example 2. SNR_i, $1 \le i \le 4$, are identical. The MMSE and MBER solutions for these two users are indistinguishable.

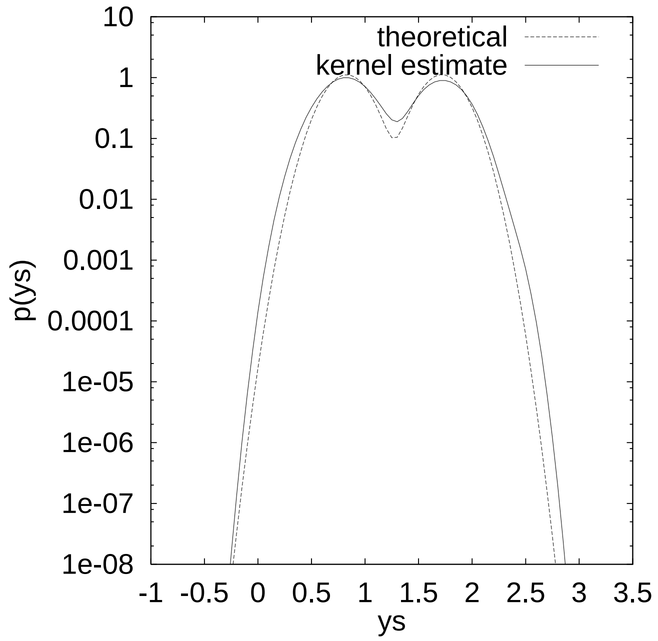

Fig. 11. Distribution of the signed decision variable for user 1 of Example 2. $\mathrm{SNR}_i = 16$ dB, for $1 \leq i \leq 4$, the detector weight vector was set to the MBER solution, and the kernel estimate was constructed from 1500 data samples.

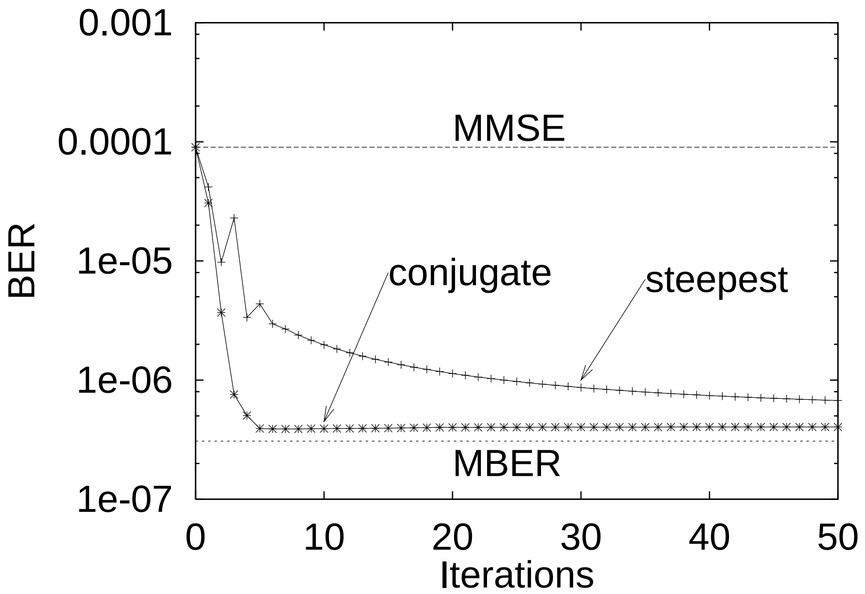

Fig. 12. Convergence behaviours of the two block adaptive MBER algorithms for user 1 of Example 2. Length of the data block is 1500, and $\mathrm{SNR}_i = 16$ dB, for $1 \leq i \leq 4$.

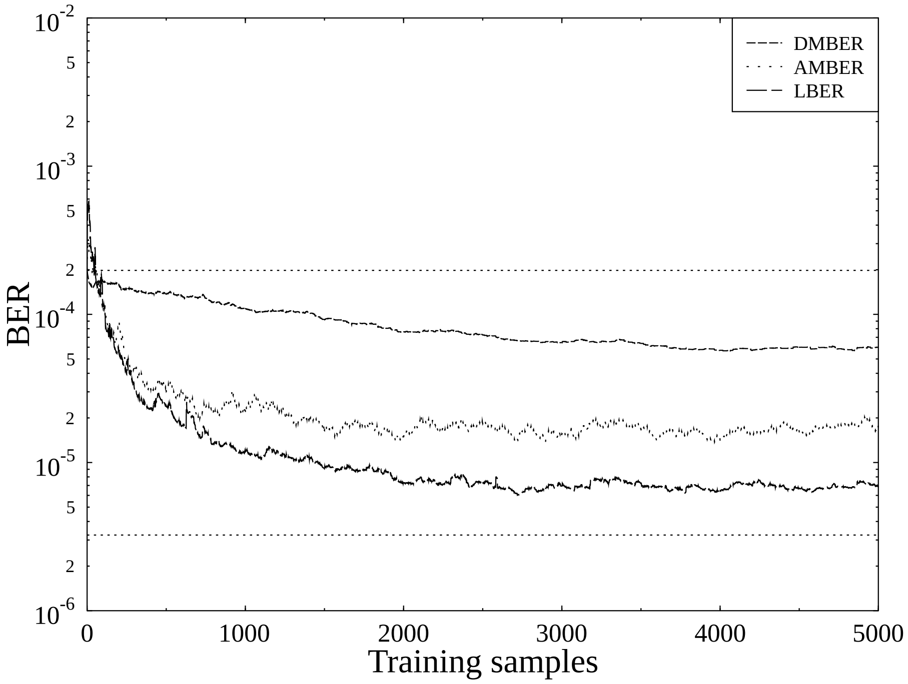

Fig. 13. Learning curves of the three stochastic adaptive MBER algorithms for user 1 of Example 2. $\mathrm{SNR}_i = 15$ dB, for $1 \leq i \leq 4$.