Minimum Bit-Error Rate Design for Space–Time Equalization-Based Multiuser Detection

Sheng Chen, *Senior Member, IEEE*, Andrew Livingstone, and Lajos Hanzo, *Fellow, IEEE*

***Abstract*—A novel minimum bit-error rate (MBER) space–time-equalization (STE)-based multiuser detector (MUD) is proposed for multiple-receive-antenna-assisted space-division multiple-access systems. It is shown that the MBER-STE-aided MUD significantly outperforms the standard minimum mean-square error design in terms of the achievable bit-error rate (BER). Adaptive implementations of the MBER STE are considered, and both the block-data-based and sample-by-sample adaptive MBER algorithms are proposed. The latter, referred to as the least BER (LBER) algorithm, is compared with the most popular adaptive algorithm, known as the least mean square (LMS) algorithm. It is shown that in case of binary phase-shift keying, the computational complexity of the LBER-STE is about half of that required by the classic LMS-STE. Simulation results demonstrate that the LBER algorithm performs consistently better than the classic LMS algorithm, both in terms of its convergence speed and steady-state BER performance.**

***Index Terms*—Adaptive algorithm, minimum bit-error rate (MBER), multiuser detection (MUD), space–time processing.**

I. Introduction

SMART-antenna-aided space–time processing is capable of substantially improving the achievable wireless system capacity, coverage, and quality by suppressing the effects of both intersymbol interference (ISI) and cochannel interference (CCI) [1]–[12]. In this paper, we consider a space-division multiple-access (SDMA) uplink scheme, where each transmitter employs a single antenna, while the basestation (BS) receiver has multiple antennas. To interpret the multiuser supporting capability of such an SDMA system [13], it is informative to compare it with classic code-division multiple-access (CDMA) multiuser systems [11]. In a CDMA system, each user is separated by a unique user-specific spreading code. By contrast, an SDMA system differentiates each user by the associated unique user-specific channel impulse response (CIR) encountered at the receiver antennas. In this analogy, the unique user-specific CIR plays the role of a user-specific CDMA signature. However, owing to the nonorthogonal nature of the CIRs, an effective multiuser detection (MUD) is required for separating the users in an SDMA system. We investigate a space–time-equalization (STE)-based MUD designed for SDMA systems. The most popular SDMA receiver design is the minimum mean-square error (MMSE) MUD [6], [10]–[14], which leads to simple and effective adaptive implementation using the least mean square (LMS) algorithm [15]. We consider an alternative design for the STE-aided MUD based on the minimum bit-error rate (MBER) criterion. Time-only processing, i.e., channel equalization, based on the MBER design has been considered before [16]–[22]. Recently, we have also proposed the MBER design for space-only processing, i.e., the narrowband-beamforming-assisted receiver [23], [24]. In this paper, we extend the MBER design to the STE-aided MUD operated in a generic multiple-antenna-assisted SDMA system.

This work is very different from that reported in [24], where the antenna array spacing was assumed to be half a wavelength, the uplink channel was frequency nonselective, and the receiver adopted a beamforming structure, which simply combined the output signals of the antenna elements. By contrast, in this paper, we do not impose any restrictions on the antenna array structure and we consider frequency-selective CIRs. Furthermore, the receiver employs a generic STE structure. The contribution of this paper is two-fold. First, it is shown that the MBER STE-based MUD is superior in comparison with the MMSE design in terms of its achievable bit-error rate (BER). This is significant, since the MMSE design is often considered to be the state-of-the-art technique in multiple-antenna-assisted systems [6], [10]–[14]. Our study thus demonstrates that the system capacity can further be enhanced beyond that of the MMSE solution. Second, we propose effective adaptive implementations of the MBER design. Both block-data-based and sample-by-sample adaptations of the MBER STE-based MUD weights are considered. The sample-by-sample adaptive algorithm is referred to here as the least BER (LBER) algorithm. It is interesting to see that this LBER STE-based MUD has, in fact, a lower complexity than the LMS-based one in the case of binary phase-shift keying (BPSK) modulation. Our simulation results also show that the LBER STE-aided MUD consistently outperforms the LMS-based one, both in terms of its convergence speed and its achievable BER.

II. System Model

Consider the multiple-antenna-aided SDMA system supporting M active users as depicted in Fig. 1, where each of the M users is equipped with a single transmit antenna and the receiver is assisted by an L-element antenna array. The

Paper approved by R. A. Kennedy, the Editor for Data Communications Modulation and Signal Design of the IEEE Communications Society. Manuscript received October 1, 2003; revised October 24, 2003 and April 15, 2005. This work was supported by the European Union under the auspices of the Phoenix and Newcom projects.

The authors are with School of Electronics and Computer Science, University of Southampton, Southampton SO17 1BJ, U.K. (e-mail: sqc@ecs.soton.ac.uk).

Digital Object Identifier 10.1109/TCOMM.2006.873999

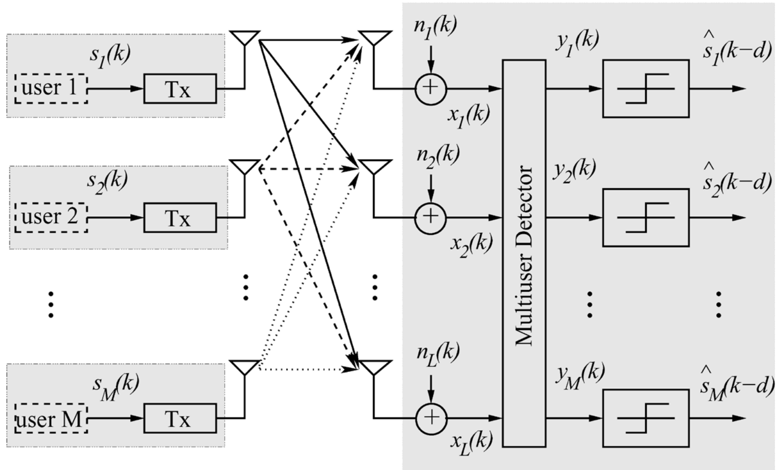


Fig. 1. Schematic of an antenna-array-aided SDMA system, where each of the M users is equipped with a single transmit antenna, and the receiver is assisted by an L-element antenna array.

symbol-rate received signal samples $x_l(k)$ for $1 \leq l \leq L$ are given by [10]

$$x_l(k) = \sum_{m=1}^{M} \sum_{i=0}^{n_C-1} c_{i,l,m} s_m(k-i) + n_l(k) = \bar{x}_l(k) + n_l(k) \tag{1}$$

where $n_l(k)$ is a complex-valued Gaussian white noise process with $E[|n_l(k)|^2] = 2\sigma_n^2$, $\bar{x}_l(k)$ denotes the noise-free part of the lth receive antenna's output, $s_m(k)$ is the kth transmitted symbol of user m, and $\mathbf{c}_{l,m} = [c_{0,l,m} c_{1,l,m} \cdots c_{n_C-1,l,m}]^T$ denotes the tap vector of the CIR connecting the user m and the lth receive antenna. For notational simplicity, we have assumed that each of the $M \times L$ CIRs has the same length of n_C. We assume furthermore that BPSK modulation is employed, and hence, $s_m(k) \in \{\pm 1\}$. We point out that this work can be extended to the quadrature phase-shift keying (QPSK) and other modulation schemes with multiple bits per symbol [25], [26]. For this multiuser system, the user m received signal-to-noise ratio (SNR) is defined as

$$\mathrm{SNR}(m) = \frac{\sum_{l=1}^{L} \mathbf{c}_{l,m}^H \mathbf{c}_{l,m} \sigma_s^2}{2L\sigma_n^2} \tag{2}$$

and the user m received signal-to-interference ratio (SIR) with respect to interfering user i, $i \neq m$, is given by

$$\mathrm{SIR}_i(m) = \frac{\sum_{l=1}^{L} \mathbf{c}_{l,m}^H \mathbf{c}_{l,m} \sigma_s^2}{\sum_{l=1}^{L} \mathbf{c}_{l,i}^H \mathbf{c}_{l,i} \sigma_s^2} \tag{3}$$

where $\sigma_s^2 = 1$ is the transmitted symbol energy.

A bank of the M STEs, as shown in Fig. 2, constitutes the MUD. The soft outputs of the M detectors are given by

$$y_m(k) = \sum_{l=1}^{L} \sum_{i=0}^{n_F-1} w_{i,l,m}^* x_l(k-i) \tag{4}$$

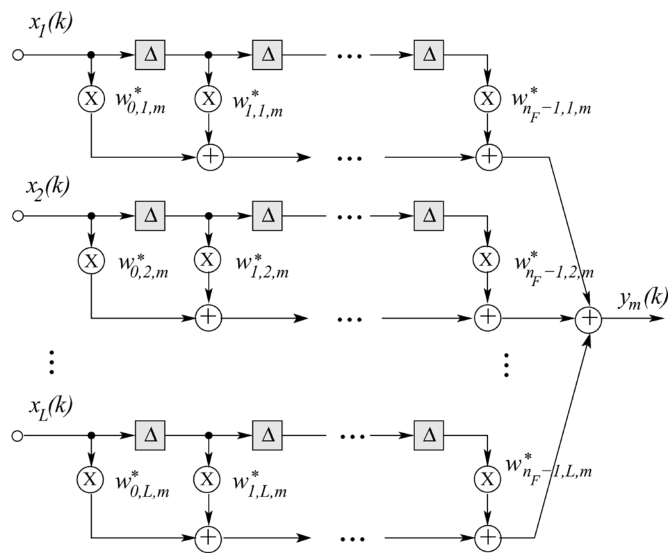


Fig. 2. Space–time-equalizer-assisted MUD for user m, where Δ denotes the symbol-spaced delay, L is the number of receive antennas, $1 \leq m \leq M$, and M is the number of users.

for $1 \leq m \leq M$, where $\mathbf{w}_{l,m} = [w_{0,l,m} w_{1,l,m} \cdots w_{n_F-1,l,m}]^T$ denotes the mth user detector's equalizer weight vector associated with the lth receive antenna. The M user detectors' decisions are defined by

$$\hat{s}_m(k-d) = \mathrm{sgn}\left(y_{R_m}(k)\right), \quad 1 \leq m \leq M \tag{5}$$

where $\hat{s}_m(k-d)$ is the estimate of $s_m(k-d)$, $y_{R_m}(k) = \Re[y_m(k)]$ denotes the real part of $y_m(k)$, and $\mathrm{sgn}(\bullet)$ the sign function. Again for notational simplicity, we assume that each of the M detectors has the same decision delay d, and all the temporal equalizer filters have the same order n_F. Obviously, $0 \leq d \leq n_F + n_C - 2$. Let us define

$$\mathbf{w}_m = \left[\mathbf{w}_{1,m}^T \, \mathbf{w}_{2,m}^T \, \cdots \, \mathbf{w}_{L,m}^T\right]^T \tag{6}$$

$$\mathbf{x}_l(k) = [x_l(k) \, x_l(k-1) \, \cdots \, x_l(k-n_F+1)]^T \tag{7}$$

$$\mathbf{x}(k) = \left[\mathbf{x}_1^T(k) \, \mathbf{x}_2^T(k) \, \cdots \, \mathbf{x}_L^T(k)\right]^T. \tag{8}$$

Then the output of the mth detector can be written as

$$y_m(k) = \sum_{l=1}^{L} \mathbf{w}_{l,m}^H \mathbf{x}_l(k) = \mathbf{w}_m^H \mathbf{x}(k). \quad (9)$$

Let us define the $n_F \times (n_F + n_C - 1)$ CIR convolution matrix associated with the user m and lth receive antenna as shown in (10) at the bottom of the page, and further introduce the overall system CIR convolution matrix as

$$\mathbf{C} = \begin{bmatrix} \mathbf{C}_{1,1} & \mathbf{C}_{1,2} & \cdots & \mathbf{C}_{1,M} \\ \mathbf{C}_{2,1} & \mathbf{C}_{2,2} & \cdots & \mathbf{C}_{2,M} \\ \vdots & \vdots & \cdots & \vdots \\ \mathbf{C}_{L,1} & \mathbf{C}_{L,2} & \cdots & \mathbf{C}_{L,M} \end{bmatrix}. \quad (11)$$

Then the received signal vector $\mathbf{x}(k)$ can be expressed by

$$\mathbf{x}(k) = \mathbf{C}\mathbf{s}(k) + \mathbf{n}(k) = \bar{\mathbf{x}}(k) + \mathbf{n}(k) \quad (12)$$

where

$$\mathbf{n}(k) = [\mathbf{n}_1(k)\ \mathbf{n}_2(k)\ \cdots\ \mathbf{n}_L(k)]^T \quad (13)$$

with $\mathbf{n}_l(k) = [n_l(k)\ n_l(k-1) \cdots n_l(k - n_F + 1)]^T$, and

$$\mathbf{s}(k) = \left[\mathbf{s}_1^T(k)\ \mathbf{s}_2^T(k)\ \cdots\ \mathbf{s}_M^T(k)\right]^T \quad (14)$$

with $\mathbf{s}_m(k) = [s_m(k) s_m(k-1) \cdots s_m(k - n_F - n_C + 2)]^T$. Note that the output of the mth detector can be expressed as

$$y_m(k) = \mathbf{w}_m^H (\bar{\mathbf{x}}(k) + \mathbf{n}(k)) = \bar{y}_m(k) + e_m(k) \quad (15)$$

where $e_m(k)$ is Gaussian distributed, having a zero mean and $E[|e_m(k)|^2] = 2\mathbf{w}_m^H \mathbf{w}_m \sigma_n^2$.

Classically, the mth STE detector's weight vector $\mathbf{w}_m$ is given by the following MMSE solution [6], [10]–[14]:

$$\mathbf{w}_{(\mathrm{MMSE})m} = \left(\mathbf{C}\mathbf{C}^H + 2\sigma_n^2 \mathbf{I}\right)^{-1} \mathbf{C}_{|(m-1)(n_F+n_C-1)+(d+1)} \quad (16)$$

for $1 \le m \le M$, where $\mathbf{I}$ denotes the $Ln_F \times Ln_F$ identity matrix, and $\mathbf{C}_{|i}$ the ith column of $\mathbf{C}$. An adaptive implementation of the MMSE solution can readily be realized using the LMS algorithm. However, as recognized by [27] in a CDMA context, and by [24] in a beamforming-based MUD scenario, a better strategy is to choose the detector's coefficients by directly minimizing the system's BER. A main contribution of this paper is to derive the MBER solution for the STE-based MUD (9).

III. MBER Space–Time Equalizer

Following the derivations presented in [24] and [27], let us denote the $N_s = 2^{M(n_F+n_C-1)}$ number of possible transmitted symbol sequences of $\mathbf{s}(k)$ as $\mathbf{s}^{(q)}$, $1 \le q \le N_s$. Denote the $((m-1)(n_F + n_C - 1) + (d+1))$th element of $\mathbf{s}^{(q)}$, corresponding to the desired symbol $s_m(k-d)$, as $s_{m,d}^{(q)}$. The noise-free part of the mth detector input signal $\bar{\mathbf{x}}(k)$ assumes values from the signal set defined as

$$\mathcal{X}_m \triangleq \left\{\bar{\mathbf{x}}^{(q)} = \mathbf{C}\mathbf{s}^{(q)}, 1 \le q \le N_s\right\}. \quad (17)$$

This set can be partitioned into two subsets, depending on the value of $s_m(k-d)$, as follows:

$$\mathcal{X}_m^{(\pm)} \triangleq \left\{\bar{\mathbf{x}}^{(q,\pm)} \in \mathcal{X}_m : s_m(k-d) = \pm 1\right\}. \quad (18)$$

For a (linear) STE to perform adequately, $\mathcal{X}_m^{(+)}$ and $\mathcal{X}_m^{(-)}$ must be linearly separable. Otherwise, a nonlinear STE is required to achieve adequate performance, a situation that is similar to the case of single-user single-antenna channel equalization [28]–[30]. Similarly, by noting the STE (15), the noise-free part of the mth detector's output $\bar{y}_m(k)$ assumes values from the scalar set

$$\mathcal{Y}_m \triangleq \left\{\bar{y}_m^{(q)} = \mathbf{w}_m^H \bar{\mathbf{x}}^{(q)}, \quad 1 \le q \le N_s\right\}. \quad (19)$$

Thus $\bar{y}_{R_m}(k) = \Re[\bar{y}_m(k)]$ can only take the values from the set

$$\mathcal{Y}_{R_m} \triangleq \left\{\bar{y}_{R_m}^{(q)} = \Re\left[\bar{y}_m^{(q)}\right], \quad 1 \le q \le N_s\right\} \quad (20)$$

and $\mathcal{Y}_{R_m}$ can be divided into the two subsets conditioned on the value of $s_m(k-d)$

$$\mathcal{Y}_{R_m}^{(\pm)} \triangleq \left\{\bar{y}_{R_m}^{(q,\pm)} \in \mathcal{Y}_{R_m} : s_m(k-d) = \pm 1\right\}. \quad (21)$$

$$\mathbf{C}_{l,m} = \begin{bmatrix} c_{0,l,m} & c_{1,l,m} & \cdots & c_{n_C-1,l,m} & 0 & \cdots & 0 \\ 0 & c_{0,l,m} & c_{1,l,m} & \cdots & c_{n_C-1,l,m} & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & c_{0,l,m} & c_{1,l,m} & \cdots & c_{n_C-1,l,m} \end{bmatrix} \quad (10)$$

The conditional probability density function (PDF) of $y_{R_m}(k)$ given $s_m(k-d)=+1$ is a Gaussian mixture defined by

$$p_m(y_R|+1)=\frac{1}{N_{sb}}\sum_{q=1}^{N_{sb}}\frac{1}{\sqrt{2\pi\sigma_n^2\mathbf{w}_m^H\mathbf{w}_m}}e^{-\frac{\left(y_R-\bar{y}_{R_m}^{(q,+)}\right)^2}{2\sigma_n^2\mathbf{w}_m^H\mathbf{w}_m}} \quad (22)$$

where $\bar{y}_{R_m}^{(q,+)}\in\mathcal{Y}_{R_m}^{(+)}$ and $N_{sb}=N_s/2$ is the number of the points in $\mathcal{Y}_{R_m}^{(+)}$. Thus the BER of the mth detector associated with the detector's weight vector $\mathbf{w}_m$ is given by [24], [27]

$$P_E(\mathbf{w}_m)=\frac{1}{N_{sb}}\sum_{q=1}^{N_{sb}}Q\left(g^{(q,+)}(\mathbf{w}_m)\right) \quad (23)$$

where

$$Q(u)=\frac{1}{\sqrt{2\pi}}\int_u^\infty e^{-\frac{v^2}{2}}dv \quad (24)$$

$$g^{(q,+)}(\mathbf{w}_m)=\frac{\operatorname{sgn}\left(s_{m,d}^{(q)}\right)\bar{y}_{R_m}^{(q,+)}}{\sigma_n\sqrt{\mathbf{w}_m^H\mathbf{w}_m}}. \quad (25)$$

Note that the BER is invariant to a positive scaling of $\mathbf{w}_m$. Alternatively, the BER may be calculated based on the other subset $\mathcal{Y}_{R_m}^{(-)}$.

The MBER solution for the mth STE detector is then defined as the weight vector that minimizes the error probability (23), namely

$$\mathbf{w}_{(\text{MBER})m}=\arg\min_{\mathbf{w}_m}P_E(\mathbf{w}_m). \quad (26)$$

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

$$\nabla P_E(\mathbf{w}_m)=\frac{1}{2N_{sb}\sqrt{2\pi}\sigma_n\sqrt{\mathbf{w}_m^H\mathbf{w}_m}}\sum_{q=1}^{N_{sb}}e^{-\frac{\left(\bar{y}_{R_m}^{(q,+)}\right)^2}{2\sigma_n^2\mathbf{w}_m^H\mathbf{w}_m}} \times\operatorname{sgn}\left(s_{m,d}^{(q)}\right)\left(\frac{\bar{y}_{R_m}^{(q,+)}\mathbf{w}_m}{\mathbf{w}_m^H\mathbf{w}_m}-\bar{\mathbf{x}}^{(q,+)}\right). \quad (27)$$

Given the gradient expression (27), the optimization problem (26) can be solved iteratively by commencing the iterations from an appropriate initial point using a gradient-based optimization algorithm, such as the simplified conjugate gradient algorithm [24], [27], [31]. Because the BER (23) is invariant to a positive scaling of $\mathbf{w}_m$, it is computationally advantageous to normalize $\mathbf{w}_m$ to a unit-norm after every iteration, so that the gradient can be simplified as

$$\nabla P_E(\mathbf{w}_m)=\frac{1}{2N_{sb}\sqrt{2\pi}\sigma_n}\sum_{q=1}^{N_{sb}}e^{-\frac{\left(\bar{y}_{R_m}^{(q,+)}\right)^2}{2\sigma_n^2}} \times\operatorname{sgn}\left(s_{m,d}^{(q)}\right)\left(\bar{y}_{R_m}^{(q,+)}\mathbf{w}_m-\bar{\mathbf{x}}^{(q,+)}\right). \quad (28)$$

In general, unlike for the MMSE solution (16), there exists no closed-form MBER solution, and therefore, a numerical solution has to be sought. Previous results involving time-only and space-only processing [23]–[27] have suggested that the simplified conjugated gradient algorithm performs well, and it is capable of finding a global minimum of $P_E(\mathbf{w}_m)$. In our extensive investigations, we found no cases of converging to a local minimum of the BER surface.

IV. Adaptive MBER Space–Time Equalizer

To derive adaptive implementation of the MBER STE, it is more convenient to consider the PDF of $y_{R_m}(k)$ explicitly, which is given by

$$p_m(y_R)=\frac{1}{N_s\sqrt{2\pi\sigma_n^2\mathbf{w}_m^H\mathbf{w}_m}}\sum_{q=1}^{N_s}e^{\frac{-\left(y_R-\bar{y}_{R_m}^{(q)}\right)^2}{2\sigma_n^2\mathbf{w}_m^H\mathbf{w}_m}} \quad (29)$$

where $\bar{y}_{R_m}^{(q)}\in\mathcal{Y}_{R_m}$. Thus the BER of the mth STE with weight vector $\mathbf{w}_m$ can alternatively be calculated by

$$P_E(\mathbf{w}_m)=\frac{1}{N_s}\sum_{q=1}^{N_s}Q\left(g^{(q)}(\mathbf{w}_m)\right) \quad (30)$$

with

$$g^{(q)}(\mathbf{w}_m)=\frac{\operatorname{sgn}\left(s_{m,d}^{(q)}\right)\bar{y}_{R_m}^{(q)}}{\sigma_n\sqrt{\mathbf{w}_m^H\mathbf{w}_m}}. \quad (31)$$

In general, however, the system CIR matrix $\mathbf{C}$ is unavailable, and therefore, the PDF of $y_{R_m}(k)$ is unknown. The key to adaptive implementation of the MBER STE-based MUD is an effective estimate of the PDF (29). A widely used approach to approximate a PDF is known as the Parzen window estimate [32]–[34]. The Parzen window method estimates a PDF using a block of $y_{R_m}(k)$ by placing a symmetric unimodal kernel function on each $y_{R_m}(k)$. This Parzen window density estimation is capable of producing reliable PDF estimates with short data records, and, in particular, is natural when dealing with Gaussian mixtures, such as (29).

A. Block-Data Gradient Adaptive MBER Space–Time Equalizer

Given a block of K training samples $\{\mathbf{x}(k),s_m(k-d)\}_{k=1}^K$, a Parzen window density estimate of the PDF (29) is readily given by

$$\hat{p}_m(y_R)=\frac{1}{K\sqrt{2\pi\rho_n^2\mathbf{w}_m^H\mathbf{w}_m}}\sum_{k=1}^{K}e^{-\frac{\left(y_R-y_{R_m}(k)\right)^2}{2\rho_n^2\mathbf{w}_m^H\mathbf{w}_m}} \quad (32)$$

where the kernel function is chosen as Gaussian, and the kernel width $\rho_n\sqrt{\mathbf{w}_m^H\mathbf{w}_m}$ is related to the standard deviation

$\sigma_n \sqrt{\mathbf{w}_m^H \mathbf{w}_m}$ for the noise $\Re[e_m(k)]$. From this estimated PDF, the estimated BER is then given by

$$\hat{P}_E(\mathbf{w}_m) = \frac{1}{K} \sum_{k=1}^{K} Q\left(\hat{g}^{(k)}(\mathbf{w}_m)\right) \tag{33}$$

with

$$\hat{g}^{(k)}(\mathbf{w}_m) = \frac{\operatorname{sgn}\left(s_m(k-d)\right) y_{R_m}(k)}{\rho_n \sqrt{\mathbf{w}_m^H \mathbf{w}_m}}. \tag{34}$$

The gradient of $\hat{P}_E(\mathbf{w}_m)$ is

$$\nabla \hat{P}_E(\mathbf{w}_m) = \frac{1}{2K\sqrt{2\pi}\rho_n \sqrt{\mathbf{w}_m^H \mathbf{w}_m}} \sum_{k=1}^{K} e^{-\frac{y_{R_m}^2(k)}{2\rho_n^2 \mathbf{w}_m^H \mathbf{w}_m}} \times \operatorname{sgn}\left(s_m(k-d)\right) \left(\frac{y_{R_m}(k)\mathbf{w}_m}{\mathbf{w}_m^H \mathbf{w}_m} - \mathbf{x}(k) \right). \tag{35}$$

By substituting $\nabla P_E(\mathbf{w}_m)$ with $\nabla \hat{P}_E(\mathbf{w}_m)$ in the simplified conjugate gradient updating mechanism, a block-data gradient adaptive algorithm is readily obtained [24], [27].

In this block-data-based adaptive STE, the step size μ and the radius parameter ρ_n are the two algorithmic parameters that need to be chosen appropriately. The step size μ and, to an extent, the radius parameter ρ_n, control the rate of convergence. The accuracy of the solution is mainly determined by the radius parameter ρ_n, which is related to the noise standard deviation σ_n for $n_l(k)$ [33]. In practice, ρ_n can often be chosen from a large range of values.

B. Stochastic Gradient Adaptive MBER Space–Time Equalizer

In the Parzen window estimate (32), the kernel width $\rho_n \sqrt{\mathbf{w}_m^H \mathbf{w}_m}$ depends on the STE weight vector $\mathbf{w}_m$. Such a choice is based on the observation of the "width" $\sigma_n \sqrt{\mathbf{w}_m^H \mathbf{w}_m}$ in the true density (29). In general, however, there is no reason why the kernel width has to be chosen in such a form. For the purpose of deriving a stochastic adaptive algorithm, it is advantageous to choose a constant width ρ_n in density estimate, as this leads to a much simpler form for the gradient of the estimated BER. Adopting this approach, an alternative Parzen window density estimate to the true PDF (29) is given by

$$\tilde{p}_m(y_R) = \frac{1}{K\sqrt{2\pi}\rho_n} \sum_{k=1}^{K} e^{-\frac{\left(y_R - y_{R_m}(k)\right)^2}{2\rho_n^2}}. \tag{36}$$

This approximation is valid, provided that the constant kernel width ρ_n is chosen appropriately. With this Parzen window density estimate, an approximate BER is then given by

$$\tilde{P}_E(\mathbf{w}_m) = \frac{1}{K} \sum_{k=1}^{K} Q\left(\tilde{g}^{(k)}(\mathbf{w}_m)\right) \tag{37}$$

with

$$\tilde{g}^{(k)}(\mathbf{w}_m) = \frac{\operatorname{sgn}\left(s_m(k-d)\right) y_{R_m}(k)}{\rho_n}. \tag{38}$$

The gradient of $\tilde{P}_E(\mathbf{w}_m)$ has a much simpler form

$$\nabla \tilde{P}_E(\mathbf{w}_m) = -\frac{1}{2K\sqrt{2\pi}\rho_n} \sum_{k=1}^{K} e^{-\frac{y_{R_m}^2(k)}{2\rho_n^2}} \operatorname{sgn}(s_m(k-d))\, \mathbf{x}(k). \tag{39}$$

In order to derive a sample-by-sample adaptive algorithm, adopt a single-sample estimate of $p_m(y_R)$, namely

$$\tilde{p}_m(y_R, k) = \frac{1}{\sqrt{2\pi}\rho_n} e^{-\frac{\left(y_R - y_{R_m}(k)\right)^2}{2\rho_n^2}}. \tag{40}$$

Conceptually, from this one-sample PDF "estimate," we have a one-sample or instantaneous BER "estimate" $\tilde{P}_E(\mathbf{w}_m, k)$. Using the instantaneous stochastic gradient formula of

$$\nabla \tilde{P}_E(\mathbf{w}_m, k) = -\frac{\operatorname{sgn}\left(s_m(k-d)\right)}{2\sqrt{2\pi}\rho_n} e^{-\frac{y_{R_m}^2(k)}{2\rho_n^2}} \mathbf{x}(k) \tag{41}$$

gives rise to a stochastic gradient adaptive algorithm, which we refer to as the LBER algorithm

$$\mathbf{w}_m(k+1) = \mathbf{w}_m(k) + \mu \frac{\operatorname{sgn}\left(s_m(k-d)\right)}{2\sqrt{2\pi}\rho_n} e^{-\frac{y_{R_m}^2(k)}{2\rho_n^2}} \mathbf{x}(k). \tag{42}$$

The adaptive gain μ and the kernel width ρ_n are the two algorithmic parameters that have to be set appropriately to ensure a fast convergence rate and small steady-state BER misadjustment. Note that there is no need to normalize the weight vector to a unit-norm after each adaptation.

Our previous empirical results using this LBER algorithm in time-only and space-only processing [23]–[27] have suggested that the algorithm behaves well and has a reasonably fast convergence rate. Note that this LBER algorithm belongs to the general stochastic gradient-based adaptive algorithm investigated in [35]. Therefore, the results of convergence analysis presented in [35] is applicable here. It is also interesting to compare this LBER algorithm with the LMS algorithm, which is given by

$$\mathbf{w}_m(k+1) = \mathbf{w}_m(k) + \mu \mathbf{x}(k)\epsilon^*(k) \tag{43}$$

where

$$\epsilon(k) = s_m(k-d) - y_m(k). \tag{44}$$

It is well known that the computational requirements of the LMS algorithm are $8 \times N_w + 2$ multiplications and $8 \times N_w - 1$ additions per weight update, where $N_w = Ln_F$ is the dimension of the weight vector $\mathbf{w}_m$. It can be shown that the LBER algorithm has a complexity of $4 \times N_w + 4$ multiplications, $4 \times N_w - 1$ additions, and $1e^{(\bullet)}$ evaluation [24]. For the BPSK modulation, it is seen that the LBER algorithm needs half of the computations

TABLE I
CIRs FOR THE 3-USER 4-ANTENNA STATIONARY SYSTEM. ACTUALLY SIMULATED CIRs WERE $C_{l,m}(z)/|C_{l,m}(z)|$ TO PROVIDE UNIT CHANNEL ENERGY

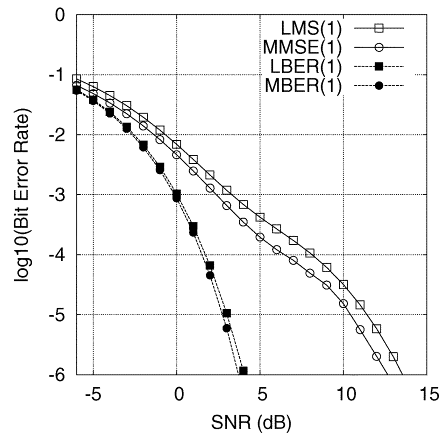
$C_{l,m}(z)$	$m=1$	$m=2$	$m=3$
$l=1$	$(-0.5+j0.4)+(0.7+j0.6)z^{-1}$	$(-0.1-j0.2)+(0.7+j0.6)z^{-1}$	$(-0.7+j0.9)+(0.6+j0.4)z^{-1}$
$l=2$	$(0.5-j0.4)+(-0.8-j0.3)z^{-1}$	$(-0.3+j0.5)+(-0.7-j0.9)z^{-1}$	$(-0.6+j0.8)+(-0.6-j0.7)z^{-1}$
$l=3$	$(0.4-j0.4)+(-0.7-j0.8)z^{-1}$	$(-0.1-j0.2)+(0.7+j0.6)z^{-1}$	$(0.3-j0.5)+(0.9+j0.1)z^{-1}$
$l=4$	$(0.5+j0.5)+(0.6-j0.9)z^{-1}$	$(-0.6-j0.4)+(0.9-j0.4)z^{-1}$	$(-0.6-j0.6)+(0.8+j0.0)z^{-1}$

required by the LMS algorithm. The $e^{(\bullet)}$ function evaluation can be implemented as a lookup table, in practice.

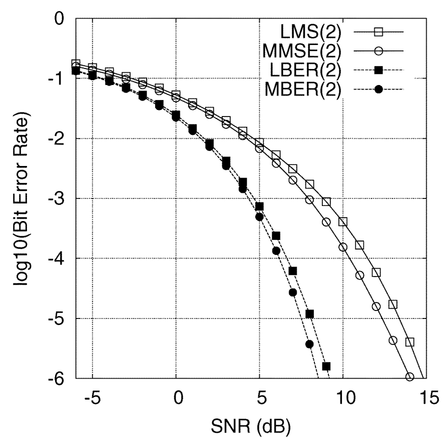
V. Simulation Study

Stationary System: The system used in our simulation supported $M = 3$ users with $L = 4$ receiver antennas. All three users had an equal transmit power. The $ML = 12$ CIRs are listed in Table I, each CIR having $n_C = 2$ taps. The CIRs used in both the stationary and fading channels are extensions of the often-used single-input single-output (SISO) CIRs proposed by Proakis in his book, which were extended to the MIMO scenario considered. In the actual simulation, all 12 CIRs were normalized to provide unit channel energy, i.e., $\|\mathbf{c}_{l,m}\|^2 = 1$ for all l and m. Thus, $\mathrm{SIR}_i(m) = 0$ dB for all m and i. Each equalizer temporal filter had a length of $n_F = 3$, and the detector decision delay was chosen to be $d = 1$. For this stationary system, Fig. 3 compares the BER performance of the MMSE and MBER STE-based MUDs. The BER of an STE-based MUD was computed using the theoretic BER formula (23), the MMSE STE weight vector was calculated using the formula (16), and the MBER STE solution was computed numerically using the simplified conjugate gradient algorithm. It can be seen that for all three users, the MBER STE detectors had better BER performance than the corresponding MMSE detectors. For the specific simulated channel conditions, the performance gap between the MBER and MMSE STE detectors was the smallest for user 3, with the MBER solution achieving above 1.0 dB gain in SNR at the BER level of 10^{-4}. At this BER level, the MBER STE detector for user 1 had the largest performance gain over the corresponding MMSE STE detector, above 5.0 dB gain in SNR. The performance of the block-data gradient adaptive MBER algorithm employing the simplified conjugate gradient updating mechanism, as described in Section IV-A, was investigated. Our simulation results show that with a block size $K = 600$ the block-data-based adaptive MBER STE can closely match the theoretical MBER STEs performance, and the algorithm typically converged within 20 iterations. Space limitation precludes the inclusion of these simulation results.

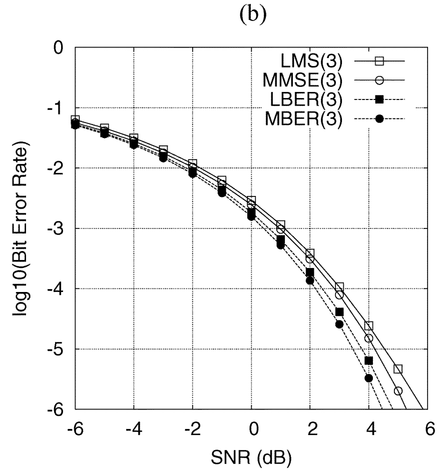
Rather, we concentrate on presenting the comparison of the LMS and LBER adaptive STE-based MUDs. The initial weight vector $\mathbf{w}_m(0)$ for the two adaptive algorithms was chosen by setting the $(d+1)$th element of $\mathbf{w}_{l,m}$ to $1/L + j0.0$, and its rest of elements to $0.0 + j0.0$ for $1 \le l \le L$. The step size μ of the LMS algorithm should be chosen to ensure fast convergence and small steady-state error, and it was found empirically that $\mu = 0.001$ was appropriate for this simulated stationary system.



(a)



(b)



(c)

Fig. 3. BER comparison of the theoretical MMSE and MBER as well as adaptive LMS and LBER STE-based MUDs for the 3-user 4-antenna stationary system. (a) User 1. (b) User 2. (c) User 3.

Similarly, the two algorithmic parameters of the LBER algorithm were chosen empirically to be $\mu = 0.2$ and $\rho_n^2 = 10\sigma_n^2$.

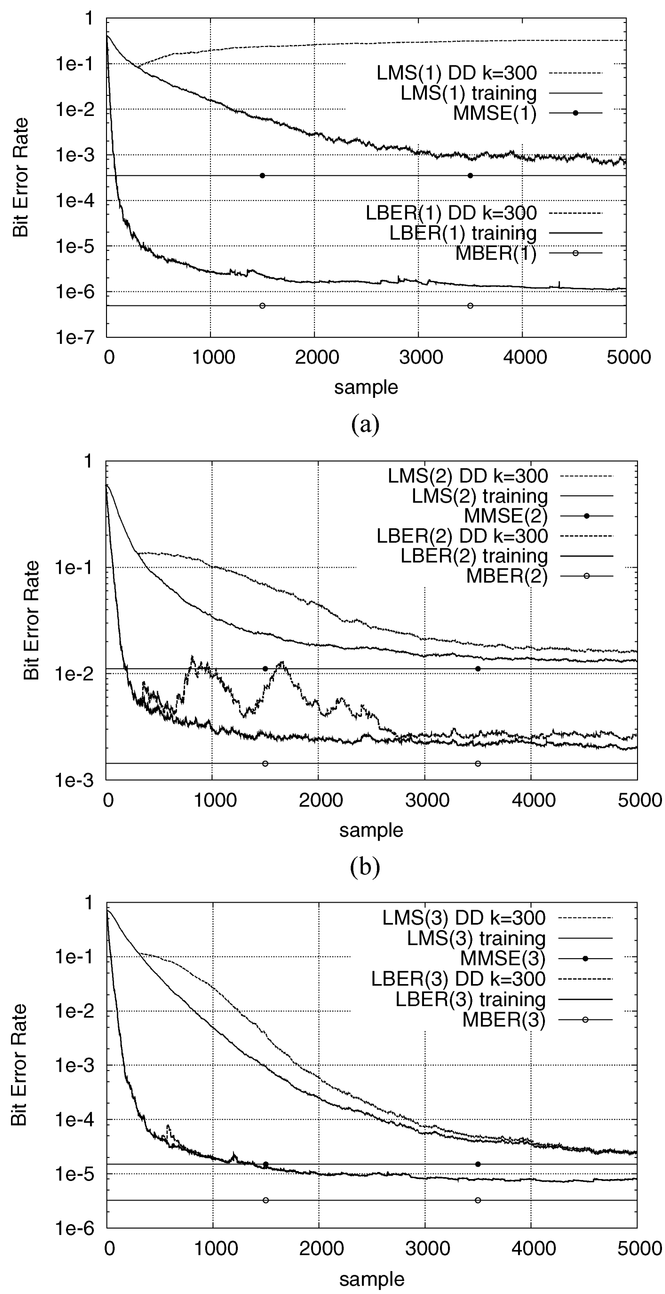


Fig. 4. Learning curves of the LMS and LBER STE-based MUDs averaged over 20 runs and given $\mathrm{SNR}(m) = 4$ dB for all m. (a) User 1. (b) User 2. (c) User 3, where DD denotes the decision-directed adaptation starting from $k = 300$ with $\hat{s}_m(k-d)$ substituting $s_m(k-d)$. For the LMS algorithm, the step size $\mu = 0.001$; and for the LBER algorithm, the step size $\mu = 0.2$ and the kernel variance $\rho_n^2 = 10\sigma_n^2 \approx 2.0$. The learning curve of the DD LBER algorithm for user 1 is indistinguishable from the training performance.

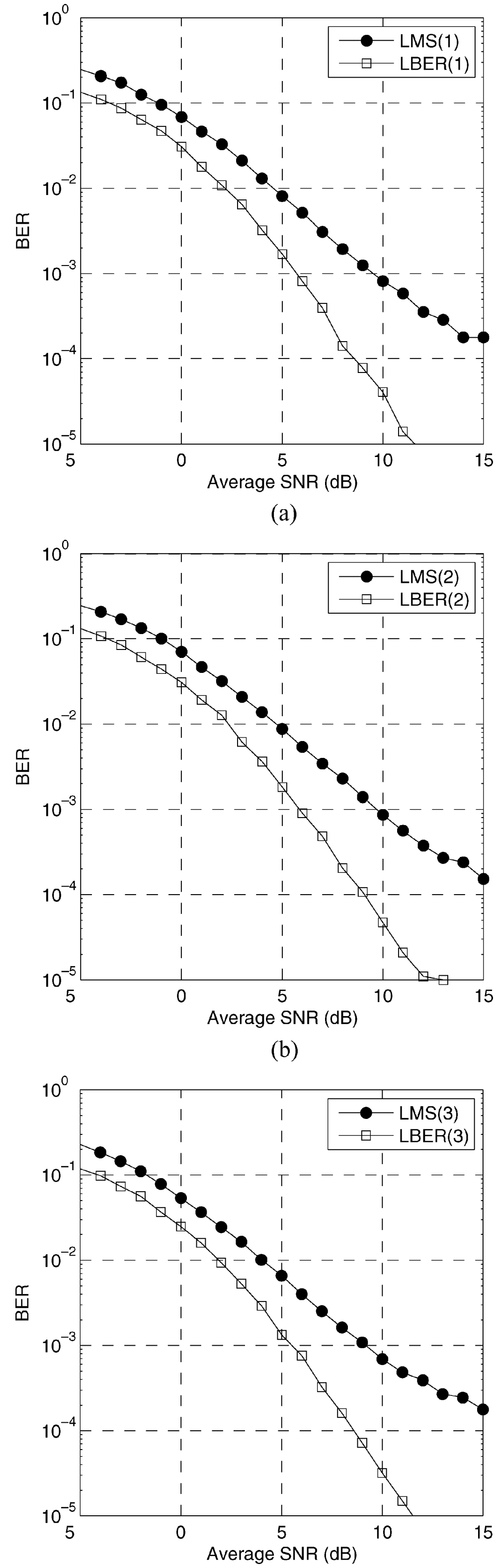


Fig. 5. BER comparison of the adaptive LMS and LBER STE-based MUDs for the 3-user 4-antenna slow fading system. (a) User 1. (b) User 2. (c) User 3.

With a training length of 5000 symbols and averaging over 20 runs, the BERs of the adaptive LMS and LBER STE-based MUDs are also given in Fig. 3, in comparison with the corresponding MMSE and MBER performance. Fig. 4 shows the learning curves of the two stochastic gradient-based adaptive algorithms averaged over 20 runs and given $\mathrm{SNR}(m) = 4$ dB for all m. From Fig. 4, it can be seen that the LBER algorithm had a faster convergence speed and achieved a smaller steady-state BER than the LMS algorithm for all three users. Having a training sequence of a few thousand symbols is, of course, impractical, and this difficulty may be avoided by considering a decision-directed (DD) adaptation in which $s_m(k-d)$

is substituted by the STE's decision $\hat{s}_m(k-d)$. We also trained the LMS and LBER STE-based MUDs, first with 300 training symbols, and then switched them to the DD mode. The resulting learning curves are also plotted in Fig. 4. It can be seen that the DD LBER algorithm operated successfully and, in the cases of users 1 and 3, its performance was indistinguishable from the related training performance. It can also be seen from Fig. 4 that the DD adaptation caused the LMS STE detector for user 1 to diverge. The ability for the STE MUD to operate successfully in a DD adaptation under adverse channel conditions is a significant advantage of the LBER design over the LMS design.

Slow Fading System: The system again supported three users with four receive antennas. However, fading channels were simulated, and moreover, each of the 12 CIRs had $n_C = 3$ taps. Magnitudes of the CIR taps were uncorrelated Rayleigh processes, each having the root mean power of $\sqrt{0.5} + j\sqrt{0.5}$. The normalized Doppler frequency for the simulated system was 10^{-5}, which for a carrier of 900 MHz and a symbol rate of 3 Msymbols/s corresponded to a user velocity of 10 m/s (36 km/h). Continuously fluctuating fading was used, which provided a different fading magnitude and phase for each transmitted symbol. Each equalizer temporal filter had a length of $n_F = 5$, and the detector decision delay was set to be $d = 2$. The step size for the LMS algorithm was chosen as $\mu = 0.005$, while for the LBER algorithm, the step size $\mu = 0.1$ and kernel variance $\rho_n^2 = 16\sigma_n^2$. The transmission frame structure consisted of 50 training symbols followed by 450 data symbols. The BER of an adaptive STE-based MUD was calculated using Monte Carlo simulation. Fig. 5 compares the BERs of the LBER STE-based MUDs for three users with those of the LMS-based ones. It can be seen from Fig. 5 that the LBER STE-based MUD consistently outperformed the LMS STE-based MUD for all three users.

VI. Conclusions

MUD based on the STE has been investigated for multiple-antenna-aided SDMA systems. A novel MBER design has been derived for the STE-based MUD. It has been shown that the MBER STE-assisted MUD can obtain significant performance gains over the standard MMSE design, in terms of the achievable system BER. Adaptive implementation of the MBER STE-assisted MUD has been considered based on a classical Parzen window density-estimation approach. Both the block-data-based and sample-by-sample adaptive MBER STE-assisted MUDs have been presented. The stochastic gradient adaptive MBER algorithm, referred to as the LBER, has some interesting properties. It requires half of the computational complexity needed by the LMS algorithm for the BPSK signaling. Our simulation results have demonstrated that the adaptive LBER STE-assisted MUD converges faster and consistently achieves better BER performance, compared with the LMS STE-assisted MUD.

Acknowledgment

The authors gratefully acknowledge the constructive critique of the anonymous reviewers.

References

[1] J. H. Winters, J. Salz, and R. D. Gitlin, "The impact of antenna diversity on the capacity of wireless communication systems," *IEEE Trans. Commun.*, vol. 42, no. 2–4, pp. 1740–1751, Feb.–Apr. 1994.

[2] G. J. Foschini, "Layered space-time architecture for wireless communication in a fading environment when using multiple antennas," *Bell Labs Tech. J.*, vol. 1, no. 2, pp. 41–59, 1996.

[3] A. J. Paulraj and C. B. Papadias, "Space-time processing for wireless communications," *IEEE Signal Process. Mag.*, vol. 14, no. 6, pp. 49–83, 1997.

[4] G. Tsoulos, M. Beach, and J. McGeehan, "Wireless personal communications for the 21st century: European technological advances in adaptive antennas," *IEEE Commun. Mag.*, vol. 35, no. 9, pp. 102–109, 1997.

[5] J. H. Winters, "Smart antennas for wireless systems," *IEEE Pers. Commun. Mag.*, vol. 5, no. 1, pp. 23–27, 1998.

[6] A. J. Paulraj and B. C. Ng, "Space-time modems for wireless personal communications," *IEEE Pers. Commun. Mag.*, vol. 5, no. 1, pp. 36–48, 1998.

[7] A. J. Paulraj and E. Lindskog, "Taxonomy of space-time processing for wireless networks," *IEE Proc. Radar, Sonar, Navigat.*, vol. 145, no. 1, pp. 25–31, 1998.

[8] P. Vandenameele, L. van Der Perre, and M. Engels, *Space Division Multiple Access for Wireless Local Area Networks*. Boston, MA: Kluwer, 2001.

[9] J. S. Blogh and L. Hanzo, *Third Generation Systems and Intelligent Wireless Networking—Smart Antenna and Adaptive Modulation*. Chichester, U.K.: Wiley, 2002.

[10] A. Paulraj, R. Nabar, and D. Gore, *Introduction to Space-Time Wireless Communications*. Cambridge, U.K.: Cambridge Univ. Press, 2003.

[11] L. Hanzo, L.-L. Yang, E.-L. Kuan, and K. Yen, *Single- and Multi-Carrier DS-CDMA: Multi-User Detection, Space-Time Spreading, Synchronization, Standards and Networking*. New York: IEEE Press/Wiley, 2003.

[12] A. J. Paulraj, D. A. Gore, R. U. Nabar, and H. Bölcskei, "An overview of MIMO communications—A key to gigabit wireless," *Proc. IEEE*, vol. 92, no. 2, pp. 198–218, Feb. 2004.

[13] L. Hanzo, M. Münster, B. J. Choi, and T. Keller, *OFDM and MC-CDMA*. West Sussex, U.K.: Wiley/IEEE Press, 2003.

[14] D. N. C. Tse and S. V. Hanly, "Linear multiuser receivers: Effective interference, effective bandwidth, and user capacity," *IEEE Trans. Inf. Theory*, vol. 45, no. 2, pp. 641–657, Feb. 1999.

[15] S. Haykin, *Adaptive Filter Theory*, 3rd ed. Upper Saddle River, NJ: Prentice-Hall, 1996.

[16] 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. IEEE Int. Conf. Commun.*, 1974, pp. 25F1–25F5.

[17] S. Chen, E. S. Chng, B. Mulgrew, and G. Gibson, "Minimum-BER linear-combiner DFE," in *Proc. IEEE Int. Conf. Commun.*, Dallas, TX, 1996, vol. 2, pp. 1173–1177.

[18] C. C. Yeh and J. R. Barry, "Approximate minimum bit-error rate equalization for binary signaling," in *Proc. IEEE Int. Conf. Commun.*, Montreal, QC, Canada, 1997, vol. 2, pp. 1095–1099.

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

[20] B. Mulgrew and S. Chen, "Stochastic gradient minimum-BER decision feedback equalisers," in *Proc. IEEE Symp. Adapt. Syst. Signal Process., Commun., Control*, Lake Louise, AB, Canada, Oct. 2000, pp. 93–98.

[21] C. C. Yeh and J. R. Barry, "Adaptive minimum bit-error rate equalization for binary signaling," *IEEE Trans. Commun.*, vol. 48, no. 7, pp. 1226–1235, Jul. 2000.

[22] B. Mulgrew and S. Chen, "Adaptive minimum-BER decision feedback equalisers for binary signalling," *Signal Process.*, vol. 81, no. 7, pp. 1479–1489, 2001.

[23] S. Chen, L. Hanzo, and N. N. Ahmad, "Adaptive minimum bit error rate beamforming assisted receiver for wireless communications," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, Hong Kong, China, Apr. 2003, vol. 4, pp. 640–643.

[24] S. Chen, N. N. Ahmad, and L. Hanzo, "Adaptive minimum bit-error rate beamforming," *IEEE Trans. Wireless Commun.*, vol. 4, no. 2, pp. 341–348, Feb. 2005.

[25] S. Chen, L. Hanzo, N. N. Ahmad, and A. Wolfgang, "Adaptive minimum bit error rate beamforming assisted QPSK receiver," in *Proc. IEEE Int. Conf. Commun.*, 2004, vol. 6, pp. 3389–3393.

[26] S. Chen, L. Hanzo, and B. Mulgrew, "Adaptive minimum symbol-error-rate decision feedback equalization for multi-level pulse-amplitude modulation," *IEEE Trans. Signal Process.*, vol. 52, no. 7, pp. 2092–2101, Jul. 2004.

[27] S. Chen, A. K. Samingan, B. Mulgrew, and L. Hanzo, "Adaptive minimum-BER linear multiuser detection for DS-CDMA signals in multipath channels," *IEEE Trans. Signal Process.*, vol. 49, no. 6, pp. 1240–1247, Jun. 2001.

[28] S. Chen, B. Mulgrew, and P. M. Grant, "A clustering technique for digital communications channel equalization using radial basis function networks," *IEEE Trans. Neural Netw.*, vol. 4, no. 4, pp. 570–579, Jul. 1993.

[29] S. Chen, B. Mulgrew, and S. McLaughlin, "Adaptive Bayesian equaliser with decision feedback," *IEEE Trans. Signal Process.*, vol. 41, no. 9, pp. 2918–2927, Sep. 1993.

[30] S. Chen, S. McLaughlin, B. Mulgrew, and P. M. Grant, "Adaptive Bayesian decision feedback equaliser for dispersive mobile radio channels," *IEEE Trans. Commun.*, vol. 43, no. 5, pp. 1937–1946, May 1995.

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

[32] E. Parzen, "On estimation of a probability density function and mode," *Ann. Math. Statist.*, vol. 33, pp. 1066–1076, 1962.

[33] B. W. Silverman, *Density Estimation*. London, U.K.: Chapman-Hall, 1996.

[34] A. W. Bowman and A. Azzalini, *Applied Smoothing Techniques for Data Analysis*. Oxford, U.K.: Oxford Univ. Press, 1997.

[35] 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 Process.*, vol. 44, no. 9, pp. 2186–2194, Sep. 1996.

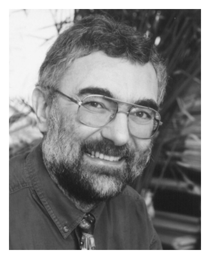
Sheng Chen (M'90–SM'97) received the B.Eng. degree from the East China Petroleum Institute, Dongying, China, in 1982, and the Ph.D. degree from the City University, London, U.K., in 1986, both in control engineering. He also received the D.Sc. degree in 2005 from the University of Southampton, Southampton, U.K.

Since 1999, he has been with the School of Electronics and Computer Science of the University of Southampton. He previously held research and academic appointments at the Universities of Sheffield, Edinburgh. and Portsmouth, all in the U.K. His recent research works include adaptive signal processing, wireless communications, modeling and identification of nonlinear systems, neural network and machine learning, finite-precision digital controller design, evolutionary computation methods, and optimization. He has published over 260 research papers.

Dr. Chen is on the list of the highly cited researchers in the engineering category in the database of the world's most highly cited researchers in various disciplines, compiled by the Institute for Scientific Information (ISI) of the USA (see http://www.ISIHighlyCited.com).

Andrew Livingstone is currently working toward the M.Eng. degree in electronics with the School of Electronics and Computer Science, University of Southampton, Southampton, U.K.

Mr. Livingstone has published four research papers in wireless communications. He has received several prizes for his excellent academic performance, including an IEEE Telecommunications Project Prize for his undergraduate project. He also received the 2005 Lord Lloyd of Kilgerran Memorial Prize for undergraduate achievement, administered by the IEE, U.K.



Lajos Hanzo (M'91–SM'92–F'04) received the master degree in electronics in 1976 and the doctorate degree in 1983, both from the Technical University of Budapest, Hungary. In 2004, he was awarded the D.Sc. degree by the University Southampton, Southampton, U.K.

During his 28-year career in telecommunications, he has held various research and academic posts in Hungary, Germany, and the U.K. Since 1986, he has been with the School of Electronics and Computer Science, University of Southampton, where he holds the Chair in Telecommunications. He has coauthored 11 Wiley/IEEE Press books totalling about 9000 pages on mobile radio communications, published in excess of 550 research papers, organized and chaired conference sessions, presented overview lectures, and has been awarded a number of distinctions.

Dr. Hanzo is a Fellow of the Royal Academy of Engineering (FREng), U.K. He is an IEEE Distinguished Lecturer of both the Communications Society and the Vehicular Technology Society, as well as a Fellow of the IEE. He is a nonexecutive director of the Virtual Centre of Excellence (VCE) in mobile communications, U.K., a governor of the IEEE Vehicular Technology Society, and an executive board member of the Pan-European Network of Excellence, known as NEWCOM.