MBER Space-Time Decision Feedback Equalization Assisted Multiuser Detection for Multiple Antenna Aided SDMA Systems

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

Abstract—This paper proposes a space-time decision feedback equalization (ST-DFE) assisted multiuser detection (MUD) scheme for multiple receiver antenna aided space division multiple access systems. A minimum bit error rate (MBER) design is invoked for the MUD, which is shown to be capable of improving the achievable bit error rate performance and enhancing the attainable system capacity over that of the standard minimum mean square error (MMSE) design. An adaptive implementation of the MBER ST-DFE assisted MUD is proposed using a stochastic gradient-based least bit error rate algorithm, which is demonstrated to consistently outperform the classical least mean square (LMS) algorithm, while achieving a lower computational complexity than the LMS algorithm for the binary signalling scheme. Our simulation results demonstrate that the MBER ST-DFE assisted MUD is more robust to channel estimation errors as well as to potential error propagation imposed by decision feedback errors, compared to the MMSE ST-DFE assisted MUD.

Index Terms—Decision feedback equalizer, minimum bit error rate, minimum mean square error, multiple antennas, multiple-input multiple-output, multiuser detection, space-division multiple access, space-time processing.

I. INTRODUCTION

IN an effort to further increase the achievable system capacity, antenna arrays can be employed for supporting multiple users in a space-division multiple access (SDMA) communications scenario [1]–[12]. We investigate a space-time (ST) decision feedback equalization (DFE) assisted multiuser detection (MUD) scheme designed for multiple receiver antenna aided SDMA systems. To interpret the multiuser-supporting capability of such a novel SDMA system [13], it is useful to relate it to 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 a simplistic but conceptually appealing interpretation, the unique user-specific CIR plays the role of a user-specific CDMA signature. In this analogy, the CIR-signatures are not orthogonal to each other,

but this is not a serious limitation, because even orthogonal spreading codes become non-orthogonal upon convolution by the CIR. However, owing to the non-orthogonal nature of the CIRs, an effective multiuser receiver is required for separating the users in an SDMA system.

The most popular SDMA-receiver design is constituted by the minimum mean square error (MMSE) MUD [5], [10]–[14]. However, as recognized by [15] in a CDMA context and by [16] in an adaptive beamforming-based MUD scenario, a better strategy is to choose the detector's coefficients by directly minimizing the system's bit error ratio (BER). For the single-user single-antenna system, the minimum BER (MBER) equalization design has become popular [17]–[24], and it has been shown that the MBER DFE is less sensitive to the error propagation due to decision feedback errors compared to the MMSE DFE [24]. For the base station employing multiple transmit antennas, an MBER multiuser transmission scheme has been proposed in [25], while for the multiple antenna assisted receiver, an MBER rake receiver has been discussed in [26]. Against this background, the novelty of this paper is that the MBER ST DFE assisted MUD (ST-DFE-MUD) is proposed for the first time in the literature in the context of SDMA. In addition to the theoretically MBER ST-DFE-MUD, which is unachievable in practice, the adaptive least bit error rate (LBER) aided ST-DFE-MUD is proposed for its practical implementation and characterized in terms of its steady-state BER and convergence performance.

In this paper, it is shown that the MBER ST-DFE-MUD design results in an enhanced BER performance in comparison to the standard MMSE design. Moreover, unlike the MMSE design, whose performance degrades significantly owing to decision feedback errors in the presence of multiuser feedback loops, the MBER ST-DFE-MUD is robust to the error propagation, as will be demonstrated by our simulation study. An adaptive implementation of the MBER ST-DFE-MUD is considered based on a stochastic gradient learning algorithm referred to as the LBER technique. It is demonstrated that this LBER ST-DFE-MUD consistently outperforms the least mean square (LMS)-based ST-DFE-MUD and yet has a lower computational complexity than the latter in the case of the binary phase-shift keying (BPSK) modulation scheme. Simulation results are also provided in support of the theoretical analysis. In this paper, we restrict attention to binary signalling. However, the technique can readily be extended to the quadrature phase-shift keying (QPSK) scheme [27] and other modulation schemes with multiple bits per symbol [24].
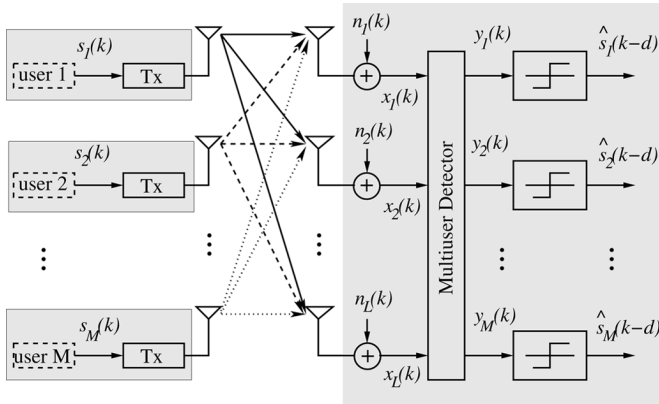
Fig. 1. Schematic of an antenna array aided SDMA uplink scenario, where each of the M users is equipped with a single transmit antenna and the base station's receiver is assisted by an L-element antenna array.

II. SYSTEM MODEL

Consider the multiple antenna aided SDMA system supporting M 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 received signal samples $x_l(k)$ at the symbol rate for $1 \le l \le L$ are given by

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

where $n_l(k)$ is an independently identically distributed 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 $s_m(k) \in \{\pm 1\}$. The physical interpretation of the CIRs associated with the model (1) in the generic content of multiple-input multiple-output systems can be found, for example, in [10].

A bank of the M ST-DFEs constitutes the MUD, and the soft outputs of the M ST-DFEs are given by

$$y_m(k) = \sum_{l=1}^{L} \sum_{i=0}^{n_F-1} w_{i,l,m}^* x_l(k-i)$$
$$+ \sum_{q=1}^{M} \sum_{i=1}^{n_B} b_{i,q,m}^* \hat{s}_q(k-d-i) \tag{2}$$

for $1 \le m \le M$, where $\hat{s}_m(k)$ denotes the decision for the transmitted symbol $s_m(k)$ and $\mathbf{w}_{l,m} = [w_{0,l,m} \; w_{1,l,m} \cdots w_{n_F-1,l,m}]^T$ denotes the feedforward filter weight vector of the mth user's detector associated with the lth receive antenna, while $\mathbf{b}_{q,m} = [b_{1,q,m} \; b_{2,q,m} \cdots b_{n_B,q,m}]^T$ denotes the mth user's detector feedback filter weight vector

associated with the qth user detector's feedback signal. Again, for notational simplicity, we have assumed that each of the M ST-DFEs has the same decision delay d, all the feedforward filters have the same order n_F, and all the feedback filters have the same order n_B. The M detectors' decisions are defined by

$$\hat{s}_m(k-d) = \text{sgn}\,(y_{R_m}(k)),\; 1 \le m \le M \tag{3}$$

where $y_{R_m}(k) = \Re[y_m(k)]$ denotes the real part of $y_m(k)$ and $\text{sgn}(\bullet)$ the sign function. Let us define $\mathbf{x}_l(k) = [x_l(k)\; x_l(k-1)\cdots x_l(k-n_F+1)]^T$, $\hat{\mathbf{s}}_{B_q}(k) = [\hat{s}_q(k-d-1)\cdots \hat{s}_q(k-d-n_B)]^T$, and

$$\mathbf{w}_m = \left[\mathbf{w}_{1,m}^T \; \mathbf{w}_{2,m}^T \cdots \mathbf{w}_{L,m}^T\right]^T \tag{4}$$

$$\mathbf{x}(k) = \left[\mathbf{x}_1^T(k) \; \mathbf{x}_2^T(k)\cdots \mathbf{x}_L^T(k)\right]^T \tag{5}$$

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

$$\hat{\mathbf{s}}_B(k) = \left[\hat{\mathbf{s}}_{B_1}^T(k) \; \hat{\mathbf{s}}_{B_2}^T(k)\cdots \hat{\mathbf{s}}_{B_M}^T(k)\right]^T. \tag{7}$$

Then the output of the mth ST-DFE can be written as

$$y_m(k) = \sum_{l=1}^{L} \mathbf{w}_{l,m}^H \mathbf{x}_l(k) + \sum_{q=1}^{M} \mathbf{b}_{q,m}^H \hat{\mathbf{s}}_{B_q}(k)$$
$$= \mathbf{w}_m^H \mathbf{x}(k) + \mathbf{b}_m^H \hat{\mathbf{s}}_B(k). \tag{8}$$

The ST-DFE structure's parameters are chosen as follows: $d = n_C - 1$, $n_F = n_C$, and $n_B = n_C - 1$. For the single-antenna single-user case, this particular choice of the DFE structure's parameters is sufficient for guaranteeing that the subsets of noise-free signal states are always linearly separable at the detector's output and therefore they guarantee an adequate performance [18], [19]. This linearly separable property is also valid in the present multiuser system. This choice of the ST-DFE structure's parameters may not always be optimal, but it does guarantee an adequate performance. Upon stipulating $n_F = n_C$ and $d = n_B = n_C - 1$, let us first introduce the two overall CIR matrices as

$$\mathbf{C}_F = \begin{bmatrix} \mathbf{C}_{F_1} \\ \mathbf{C}_{F_2} \\ \vdots \\ \mathbf{C}_{F_L} \end{bmatrix} \text{ and } \mathbf{C}_B = \begin{bmatrix} \mathbf{C}_{B_1} \\ \mathbf{C}_{B_2} \\ \vdots \\ \mathbf{C}_{B_L} \end{bmatrix} \tag{9}$$

where the $n_F \times M(d+1)$ and $n_F \times M n_B$ dimensional CIR matrices $\mathbf{C}_{F_l}$ and $\mathbf{C}_{B_l}$ are given by

$$\mathbf{C}_{F_l} = [\mathbf{C}_{F_l,1} \quad \mathbf{C}_{F_l,2} \quad \cdots \quad \mathbf{C}_{F_l,M}] \tag{10}$$

and

$$\mathbf{C}_{B_l} = [\mathbf{C}_{B_l,1} \quad \mathbf{C}_{B_l,2} \quad \cdots \quad \mathbf{C}_{B_l,M}] \tag{11}$$

respectively, with the $n_F \times (d+1)$ and $n_F \times n_B$ dimensional CIR matrices $\mathbf{C}_{F_{l,m}}$ and $\mathbf{C}_{B_{l,m}}$ defined by

$$\mathbf{C}_{Fl,m} = \begin{bmatrix} c_{0,l,m} & c_{1,l,m} & \cdots & c_{n_C-1,l,m} \\ 0 & c_{0,l,m} & \ddots & \vdots \\ \vdots & \ddots & \ddots & c_{1,l,m} \\ 0 & \cdots & 0 & c_{0,l,m} \end{bmatrix} \quad (12)$$

and

$$\mathbf{C}_{Bl,m} = \begin{bmatrix} 0 & \cdots & 0 \\ c_{n_C-1,l,m} & \ddots & \vdots \\ \vdots & \ddots & 0 \\ c_{1,l,m} & \cdots & c_{n_C-1,l,m} \end{bmatrix} \quad (13)$$

respectively. Let us define furthermore

$$\mathbf{s}_F(k) = [\mathbf{s}_{F_1}^T(k)\, \mathbf{s}_{F_2}^T(k) \cdots \mathbf{s}_{F_M}^T(k)]^T \quad (14)$$
$$\mathbf{s}_B(k) = [\mathbf{s}_{B_1}^T(k)\, \mathbf{s}_{B_2}^T(k) \cdots \mathbf{s}_{B_M}^T(k)]^T \quad (15)$$
$$\mathbf{n}(k) = [\mathbf{n}_1(k)\, \mathbf{n}_2(k) \cdots \mathbf{n}_L(k)]^T \quad (16)$$

where $\mathbf{s}_{F_m}(k) = [s_m(k)\, s_m(k-1) \cdots s_m(k-d)]^T$, $\mathbf{s}_{B_m}(k) = [s_m(k-d-1) \cdots s_m(k-d-n_B)]^T$, and $\mathbf{n}_l(k) = [n_l(k)\, n_l(k-1) \cdots n_l(k-n_F+1)]^T$. Then the received signal vector $\mathbf{x}(k)$ is modeled as

$$\mathbf{x}(k) = \mathbf{C}_F \mathbf{s}_F(k) + \mathbf{C}_B \mathbf{s}_B(k) + \mathbf{n}(k). \quad (17)$$

Under the assumption that the past decisions are correct, $\hat{\mathbf{s}}_B(k) = \mathbf{s}_B(k)$ and the received signal vector can be expressed as $\mathbf{x}(k) = \mathbf{C}_F \mathbf{s}_F(k) + \mathbf{C}_B \hat{\mathbf{s}}_B(k) + \mathbf{n}(k)$. Thus, as argued in [18] and [28], the decision feedback can be viewed as a translation of the original observation space $\mathbf{x}(k)$ into a new space $\mathbf{r}(k)$

$$\mathbf{r}(k) \triangleq \mathbf{x}(k) - \mathbf{C}_B \hat{\mathbf{s}}_B(k) = \mathbf{C}_F \mathbf{s}_F(k) + \mathbf{n}(k) = \bar{\mathbf{r}}(k) + \mathbf{n}(k). \quad (18)$$

In the translated space $\mathbf{r}(k)$, the original ST-DFE described by (8) is "translated" into a ST "linear equalizer" described as

$$y_m(k) = \mathbf{w}_m^H \mathbf{r}(k) = \mathbf{w}_m^H(\bar{\mathbf{r}}(k) + \mathbf{n}(k)) = \bar{y}_m(k) + e_m(k) \quad (19)$$

where $e_m(k)$ is Gaussian distributed, having a zero mean and $E[|e_m(k)|^2] = 2\mathbf{w}_m^H \mathbf{w}_m \sigma_n^2$. Note that we have $\mathbf{r}(k) = [\mathbf{r}_1^T(k)\, \mathbf{r}_2^T(k) \cdots \mathbf{r}_L^T(k)]^T$ with $\mathbf{r}_l(k) = [r_l(k)\, r_l(k-1) \cdots r_l(k-n_F+1)]^T$. By extending the results of [18] and [19], the elements of $\mathbf{r}_l(k)$ can be computed recursively according to

$$r_l(k-i) = z^{-1} r_l(k-i+1)$$
$$- \sum_{m=1}^{M} c_{n_C-i,l,m} \hat{s}_m(k-d-1),$$
$$\text{for } i = n_F-1, n_F-2, \ldots, 1$$
$$r_l(k) = x_l(k) \quad (20)$$

where z^{-1} defines the unit delay operator. Note that the recursion (20) realizes exactly the space translation (18) in a com-
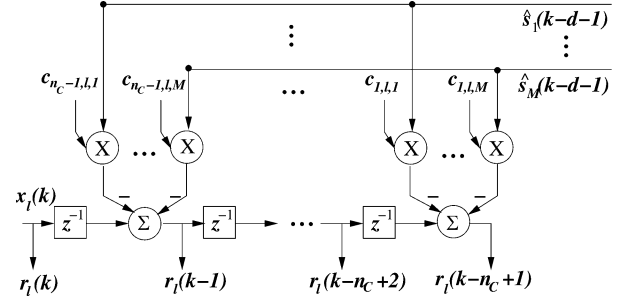

Fig. 2. Schematic of the observation space translation with $n_F = n_C$, $n_B = n_C - 1$, and $d = n_C - 1$.

putationally more efficient manner. A schematic of this space translation (20) is illustrated in Fig. 2.

It is important to emphasize that using the detector structure of (19) with the space translation (20) is exactly the same as using the original DFE structure (8). The feedback coefficient vector $\mathbf{b}_m$ does not simply "disappear." It has in fact been set to its "optimal value," which is $\mathbf{b}_m = -\mathbf{C}_B^H \mathbf{w}_m$. Thus, in an adaptive implementation, one has to estimate the coefficients of the CIRs, rather than estimating the coefficients of the feedback filters, when adopting the detector structure of (19) and (20). Moreover, when the MMSE solution, as given in (21), is used in the detector (19), the feedback filter coefficient vector in (8) is automatically set to its MMSE solution, which is $\mathbf{b}_{(\text{MMSE})m} = -\mathbf{C}_B^H \mathbf{w}_{(\text{MMSE})m}$. Similarly, if $\mathbf{w}_m$ in (19) is set to the MBER solution, as discussed in the next section, $\mathbf{b}_m$ in (8) is automatically set to the corresponding MBER solution.

III. MINIMUM BIT ERROR RATE MULTIUSER DETECTION

Classically, the mth detector's weight vector $\mathbf{w}_m$ is determined by minimizing the MSE criterion $E[|s_m(k-d) - y_m(k)|^2]$, which leads to the following MMSE solution:

$$\mathbf{w}_{(\text{MMSE})m} = \left(\mathbf{C}_F \mathbf{C}_F^H + 2\sigma_n^2 \mathbf{I}_{Ln_F} \right)^{-1} \mathbf{C}_{F|m(d+1)} \quad (21)$$

for $1 \leq m \leq M$, where $\mathbf{I}_{Ln_F}$ is the $Ln_F \times Ln_F$ dimensional identity matrix and $\mathbf{C}_{F|m(d+1)}$ denotes the $m(d+1)$th column of $\mathbf{C}_F$. An adaptive implementation of the MMSE solution can be realized, for example, using the LMS algorithm. The main contribution of this paper is to derive the MBER solution for the weight vectors of the ST-DFEs and develop an adaptive MBER multiuser detector for the SDMA systems. Let us denote the $N_s = 2^{M(d+1)}$ number of possible transmitted symbol sequences of $\mathbf{s}_F(k)$ as $\mathbf{s}^{(q)}$, $1 \leq q \leq N_s$. Denote furthermore the $m(d+1)$th element of $\mathbf{s}^{(q)}$, corresponding to the symbol $s_m(k-d)$, as $s_{m,d}^{(q)}$. The noise-free part of the mth detector input signal, namely, $\bar{\mathbf{r}}(k)$, assumes values from the finite signal set defined as

$$\mathcal{R}_m \triangleq \{ \bar{\mathbf{r}}^{(q)} = \mathbf{C}_F \mathbf{s}^{(q)}, 1 \leq q \leq N_s \}. \quad (22)$$

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

$$\mathcal{R}_m^{(\pm)} \triangleq \{ \bar{\mathbf{r}}^{(q,\pm)} \in \mathcal{R}_m : s_m(k-d) = \pm 1 \}. \quad (23)$$

Similarly, the noise-free part of the mth detector's output $\bar{y}_m(k)$ assumes values from the set

$$\mathcal{Y}_m \triangleq \{\bar{y}_m^{(q)} = \mathbf{w}_m^H \bar{\mathbf{r}}^{(q)}, 1 \le q \le N_s\}. \tag{24}$$

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

$$\mathcal{Y}_{R_m} \triangleq \{\bar{y}_{R_m}^{(q)} = \Re[\bar{y}_m^{(q)}], 1 \le q \le N_s\} \tag{25}$$

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

$$\mathcal{Y}_{R_m}^{(\pm)} \triangleq \{\bar{y}_{R_m}^{(q,\pm)} \in \mathcal{Y}_{R_m} : s_m(k-d) = \pm 1\}. \tag{26}$$

The conditional probability density function (pdf) of $y_{R_m}(k)$ given $s_m(k-d) = +1$ is a Gaussian mixture given by [15], [16]

$$p_m(y_R|+1)$$
$$= \frac{1}{N_{sb}} \sum_{q=1}^{N_{sb}} \frac{1}{\sqrt{2\pi\sigma_n^2 \mathbf{w}_m^H \mathbf{w}_m}} e^{-\left(y_R - \bar{y}_{R_m}^{(q,+)}\right)^2 / 2\sigma_n^2 \mathbf{w}_m^H \mathbf{w}_m} \tag{27}$$

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

$$P_E(\mathbf{w}_m) = \frac{1}{N_{sb}} \sum_{q=1}^{N_{sb}} Q\left(g_R^{(q,+)}(\mathbf{w}_m)\right) \tag{28}$$

where

$$Q(u) = \frac{1}{\sqrt{2\pi}} \int_u^\infty e^{-v^2/2} dv \tag{29}$$

and

$$g_R^{(q,+)}(\mathbf{w}_m) = \frac{\mathrm{sgn}(s_{m,d}^{(q)})\bar{y}_{R_m}^{(q,+)}}{\sigma_n \sqrt{\mathbf{w}_m^H \mathbf{w}_m}} = \frac{\mathrm{sgn}(s_{m,d}^{(q)})\Re[\mathbf{w}_m^H \bar{\mathbf{r}}^{(q,+)}]}{\sigma_n \sqrt{\mathbf{w}_m^H \mathbf{w}_m}}. \tag{30}$$

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

The MBER solution for the mth detector is then defined as the weight vector that minimizes the error probability (28), namely

$$\mathbf{w}_{(\text{MBER})m} = \arg \min_{\mathbf{w}_m} P_E(\mathbf{w}_m). \tag{31}$$

This MBER solution may be found by setting the derivative of $P_E(\mathbf{w}_m)$ to zero. The gradient of $P_E(\mathbf{w}_m)$ with respect to $\mathbf{w}_m$ is given by

$$\nabla P_E(\mathbf{w}_m) = \frac{1}{2N_{sb}\sqrt{2\pi}\sigma_n \sqrt{\mathbf{w}_m^H \mathbf{w}_m}}$$
$$\cdot \sum_{q=1}^{N_{sb}} e^{-\left(\bar{y}_{R_m}^{(q,+)}\right)^2 / 2\sigma_n^2 \mathbf{w}_m^H \mathbf{w}_m} \times \mathrm{sgn}\left(s_{m,d}^{(q)}\right)\left(\frac{\bar{y}_{R_m}^{(q,+)}\mathbf{w}_m}{\mathbf{w}_m^H \mathbf{w}_m} - \bar{\mathbf{r}}^{(q,+)}\right). \tag{32}$$

Given the gradient expression (32), the optimization problem (31) can be solved iteratively by commencing the iterations from an appropriate initialization point using a gradient algorithm. Since the BER is invariant to a positive scaling of $\mathbf{w}_m$, it is computationally advantageous to normalize $\mathbf{w}_m$ to a unit length after every iteration, so that the gradient expression (32) can be simplified to

$$\nabla P_E(\mathbf{w}_m) = \frac{1}{2N_{sb}\sqrt{2\pi}\sigma_n}$$
$$\cdot \sum_{q=1}^{N_{sb}} e^{-\left(\bar{y}_{R_m}^{(q,+)}\right)^2 / 2\sigma_n^2} \mathrm{sgn}\left(s_{m,d}^{(q)}\right) \times \left(\bar{y}_{R_m}^{(q,+)}\mathbf{w}_m - \bar{\mathbf{r}}^{(q,+)}\right). \tag{33}$$

The simplified conjugate gradient algorithm of [29] and [15] provides an efficient means of finding an MBER solution for the optimization problem formulated in (31).

IV. ADAPTIVE MINIMUM BIT ERROR RATE IMPLEMENTATION

The evaluation of the error probability requires the knowledge of the pdf of the ST-DFE's output signal $y_{R_m}(k)$. The pdf of $y_{R_m}(k)$ can be explicitly expressed as

$$p_m(y_R) = \frac{1}{N_s\sqrt{2\pi}\sigma_n \sqrt{\mathbf{w}_m^H \mathbf{w}_m}} \sum_{q=1}^{N_s} e^{-\left(y_R - \bar{y}_{R_m}^{(q)}\right)^2 / 2\sigma_n^2 \mathbf{w}_m^H \mathbf{w}_m} \tag{34}$$

and the associated BER can alternatively be calculated according to

$$P_E(\mathbf{w}_m) = \frac{1}{N_s} \sum_{q=1}^{N_s} Q\left(g_R^{(q)}(\mathbf{w}_m)\right) \tag{35}$$

with

$$g_R^{(q)}(\mathbf{w}_m) = \frac{\mathrm{sgn}(s_{m,d}^{(q)})\bar{y}_{R_m}^{(q)}}{\sigma_n \sqrt{\mathbf{w}_m^H \mathbf{w}_m}} \tag{36}$$

where the summation is carried out over the N_s number of elements $\bar{y}_{R_m}^{(q)} \in \mathcal{Y}_{R_m}$. In reality, the pdf of $y_{R_m}(k)$ is channel dependent and hence is unknown. Hence, some form of pdf estimation is required for supporting the adaptive implementation of the MBER ST-DFE assisted MUD.

A. Block-Data Based Gradient Adaptive MBER ST-DFE

The Parzen window method [30]–[32] constitutes an efficient means of estimating a pdf. Specifically, the Parzen window method estimates a pdf using a window or block of the ST-DFE output signal $y_{R_m}(k)$ by placing a symmetric unimodal kernel function centered on each $y_{R_m}(k)$ sample and averaging over all the data points. This density estimation technique is capable of producing reliable pdf estimates with the aid of short data records and is natural when dealing with Gaussian mixtures, such as the one given in (34). In our application, it is convenient to choose a Gaussian kernel function having a kernel width of $\rho_n \sqrt{\mathbf{w}_m^H \mathbf{w}_m}$ that is similar to the noise standard deviation of $\sigma_n \sqrt{\mathbf{w}_m^H \mathbf{w}_m}$. Given a block of K training samples

$\{\mathbf{r}(k), s_m(k-d)\}_{k=1}^{K}$, a Parzen window density estimate of the pdf in (34) takes the form

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

(37)

where the radius parameter ρ_n is related to the standard deviation σ_n of the system's noise. The accuracy analysis of the Parzen window density estimate is well documented in the literature [30]–[32]. The pdf estimate (37) is known to possess a mean integrated square error convergence rate having an order of K^{-1} [30].

Based on the estimated pdf of (37), the estimated BER is given by

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

(38)

with

$$\hat{g}_R^{(k)}(\mathbf{w}_m) = \frac{\text{sgn}(s_m(k-d))y_{R_m}(k)}{\rho_n\sqrt{\mathbf{w}_m^H\mathbf{w}_m}}.$$

(39)

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

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

(40)

Upon substituting $\nabla P_E(\mathbf{w}_m)$ by $\nabla\hat{P}_E(\mathbf{w}_m)$ in the simplified conjugate gradient updating mechanism, for example, a block-data-based adaptive algorithm is obtained [15], where the step size μ and the radius parameter ρ_n are two algorithmic parameters that control the rate of convergence. The radius parameter ρ_n also has an influence on the accuracy of the pdf and hence on that of the BER estimate.

B. Stochastic Gradient-Based Adaptive MBER ST-DFE

Our aim is to develop a sample-by-sample adaptive implementation of the MBER ST-DFE. In the Parzen window estimate (37), the kernel width $\rho_n\sqrt{\mathbf{w}_m^H\mathbf{w}_m}$ explicitly depends on the detector's weight vector $\mathbf{w}_m$. However, the BER is invariant to $\mathbf{w}_m^H\mathbf{w}_m$, and a constant kernel width ρ_n may also be adopted in the density estimate. An advantage of using a constant kernel width ρ_n, rather than $\rho_n\sqrt{\mathbf{w}_m^H\mathbf{w}_m}$, in the density estimate is that the gradient of the resultant estimated BER has a simpler form, which leads to a considerable computational complexity reduction. Adopting this approach, an alternative fixed kernel-width-based Parzen window estimate of the true pdf (34) is given by

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

(41)

and the resultant approximate BER formula becomes

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

(42)

where we have

$$\tilde{g}_R^{(k)}(\mathbf{w}_m) = \frac{\text{sgn}(s_m(k-d))y_{R_m}(k)}{\rho_n}.$$

(43)

This approximation is an adequate one, provided that the width ρ_n is chosen appropriately.

In order to derive a sample-by-sample adaptive algorithm for updating the detector's weight vector $\mathbf{w}_m$, consider a single-sample estimate of $p_m(y_R)$, namely

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

(44)

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

$$\nabla\tilde{P}_E(\mathbf{w}_m, k) = -\frac{\text{sgn}(s_m(k-d))}{2\sqrt{2\pi}\rho_n}e^{-y_{R_m}^2(k)/2\rho_n^2}\mathbf{r}(k)$$

(45)

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

$$\mathbf{w}_m(k+1) = \mathbf{w}_m(k) + \mu\frac{\text{sgn}(s_m(k-d))}{2\sqrt{2\pi}\rho_n}e^{-y_{R_m}^2(k)/2\rho_n^2}\mathbf{r}(k).$$

(46)

The adaptive gain μ as well as the kernel width ρ_n are the two algorithmic parameters that have to be set appropriately. Specifically, they are chosen to ensure adequate performance in terms of both the achievable convergence rate and the steady-state BER misadjustment. Note that there is no need to normalize the weight vector to a unit length after each update.

The CIR taps $c_{i,l,m}$ for $0 \leq i \leq n_C - 1, 1 \leq l \leq L$, and $1 \leq m \leq M$, which are needed to perform the space transformation (20), can be estimated using the standard LMS algorithm. It is interesting to compare the LBER ST-DFE (46) with the standard LMS ST-DFE, which is given by

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

(47)

with

$$\epsilon(k) = s_m(k-d) - y_m(k).$$

(48)

It can readily be shown that for the BPSK case, the LBER ST-DFE is computationally simpler than the LMS ST-DFE, imposing about half the computational complexity required by the LMS algorithm [16]. It can also be shown that for QPSK modulation, the LBER ST-DFE has a similar computational complexity as the LMS ST-DFE [27].

TABLE I
SYSTEM'S CIRS FOR A FOUR-ANTENNA FOUR-USER TIME-INVARIANT SDMA SYSTEM

$C_{l,m}(z)$	$m = 1$	$m = 2$	$m = 3$	$m = 4$
$l = 1$	$(0.6 + j0.7)$ $+(0.8 + j0.5)z^{-1}$ $+(0.3 + j0.4)z^{-2}$	$(-0.1 - j0.2)$ $+(0.4 + j0.5)z^{-1}$ $+(0.3 - j0.2)z^{-2}$	$(0.7 + j0.5)$ $+(0.6 + j0.4)z^{-1}$ $+(0.5 + j0.5)z^{-2}$	$(0.8 - j0.4)$ $+(-0.6 + j0.5)z^{-1}$ $+(0.3 + j0.3)z^{-2}$
$l = 2$	$(0.1 + j0.2)$ $+(0.4 + j0.3)z^{-1}$ $+(0.5 + j0.4)z^{-2}$	$(0.9 + j0.2)$ $+(0.3 + j0.7)z^{-1}$ $+(0.2 + j0.2)z^{-2}$	$(-0.3 - j0.3)$ $+(0.4 + j0.2)z^{-1}$ $+(-0.2 + j0.4)z^{-2}$	$(0.3 + j0.3)$ $+(0.4 + j0.4)z^{-1}$ $+(0.5 + j0.5)z^{-2}$
$l = 3$	$(-0.1 + j0.3)$ $+(0.6 - j0.5)z^{-1}$ $+(0.2 + j0.4)z^{-2}$	$(0.5 + j0.6)$ $+(-0.3 - j0.4)z^{-1}$ $+(0.2 + j0.4)z^{-2}$	$(0.2 - j0.3)$ $+(0.4 - j0.5)z^{-1}$ $+(0.6 + j0.3)z^{-2}$	$(0.1 + j0.8)$ $+(0.7 + j0.6)z^{-1}$ $+(0.8 + j0.5)z^{-2}$
$l = 4$	$(0.8 + j0.9)$ $+(0.6 + j0.5)z^{-1}$ $+(0.5 + j0.3)z^{-2}$	$(0.4 + j0.4)$ $+(0.4 + j0.4)z^{-1}$ $+(0.4 + j0.4)z^{-2}$	$(0.1 + j0.2)$ $+(-0.3 - j0.4)z^{-1}$ $+(0.3 + j0.2)z^{-2}$	$(0.4 + j0.6)$ $+(0.5 + j0.3)z^{-1}$ $+(0.2 + j0.3)z^{-2}$

V. SIMULATION STUDY

In our simulation investigations, unless otherwise stated, perfect channel estimates are assumed in performing the space translation (20). Hence our attention is focused on the performance of the adaptive MBER and MMSE designs, rather than on the adaptive channel estimator, which is well documented in the classic adaptive signal-processing literature [33]. The effect of imperfect channel estimates on the performance of a ST-DFE-MUD, however, was investigated in our simulation study. In order to avoid obfuscating the prevalent MBER/MMSE performance trends by asynchronous transmissions, we assumed synchronous communications and an identical CIR dispersion for all users.

A. Time-Invariant System

The system used in our simulations supported $M = 4$ users with the aid of $L = 4$ receiver antennas. All the four users had an equal transmit power. The resultant 16 CIRs are listed in Table I, each having $n_C = 3$ taps. In the simulations, all the 16 CIRs were normalized according to $C_{l,m}(z)/|C_{l,m}(z)|$ for the sake of providing a channel gain of unity. Since the length of the CIRs was $n_C = 3$, the ST-DFE structure was defined by the parameters of $n_F = 3$, $d = 2$, and $n_B = 2$. The theoretical BER curves of the MMSE and MBER ST-DFE-MUDs, computed using the BER expression of (28), are plotted in Fig. 3 over a range of signal-to-noise ratio (SNR) conditions, where the MMSE solutions were calculated using (21) while the MBER solutions were obtained numerically using the simplified conjugate gradient optimization algorithm. It can be seen that for all four users, the MBER ST-DFE-MUD provided a better BER performance than the MMSE ST-DFE-MUD. The BER calculated using the BER expression (28) represents the theoretical best case performance, since it was obtained assuming that the correct symbols were fed back in the ST-DFE-MUD's feedback loop. In reality, the ST-DFE-MUD may produce erroneous symbol feedback.

For the sake of investigating the effects of decision feedback induced error propagation, the BERs of the MMSE and MBER ST-DFE-MUDs were also calculated using simulations with the error-prone detected symbols being fed back, and the results are also depicted in Fig. 3, in comparison to the corresponding theoretical best case performance. It is interesting to see that the BER performance degradation owing to DFE-induced error propagation is less serious for the MBER ST-DFE-MUD than for the MMSE ST-DFE-MUD. Note that this phenomenon has
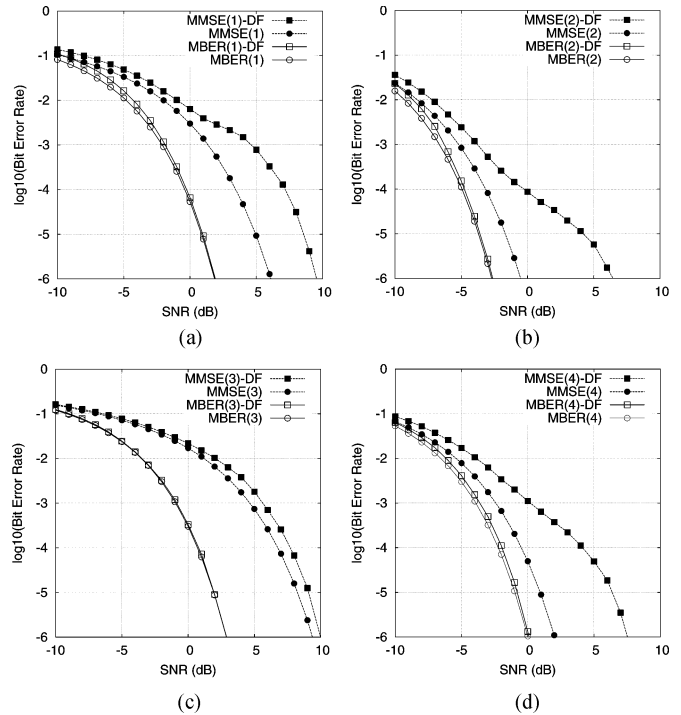

Fig. 3. Theoretical and simulated bit error rate comparison of the MMSE and MBER ST-DFE-MUDs for the four-user four-antenna time-invariant system, where DF indicates simulated BER with detected symbols being fed back: (a) user 1, (b) user 2, (c) user 3, and (d) user 4.

been observed before in the single-antenna single-user scenario of [24]. In the multiuser case, the effects of error propagation are aggravated for the MMSE ST-DFE-MUD to some degree owing to the multiple-user feedback effects. In particular, the users originally benefitting from the best performance owing to their specific CIRs suffer the most serious degradation, as seen by comparing user 2 and user 3 in Fig. 3. By contrast, the MBER ST-DFE-MUD appears to be significantly more robust to error propagation. This has important implications, especially in decision-directed (DD) adaptation.

The effect of imperfect channel estimates to the performance of a ST-DFE-MUD was next investigated. We added the Gaussian white noise with standard deviation 0.1 to each tap of the CIRs to represent channel estimation errors. The resultant "estimated" CIRs were then used to perform the space translation (20) as well as to calculate the MMSE and MBER
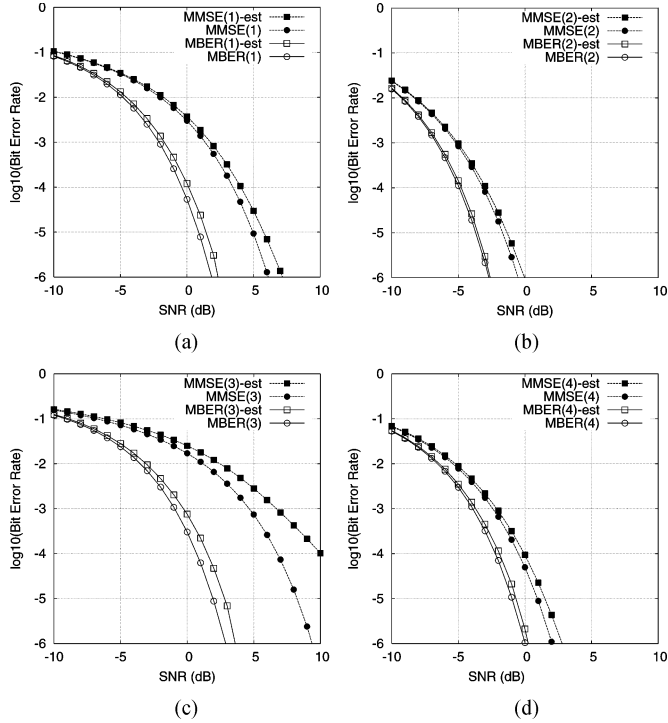
Fig. 4. Theoretical bit error rate comparison of the MMSE and MBER ST-DFE-MUDs for the four-user four-antenna time-invariant system, where "est" indicates imperfect channel estimates were used: (a) user 1, (b) user 2, (c) user 3, and (d) user 4.
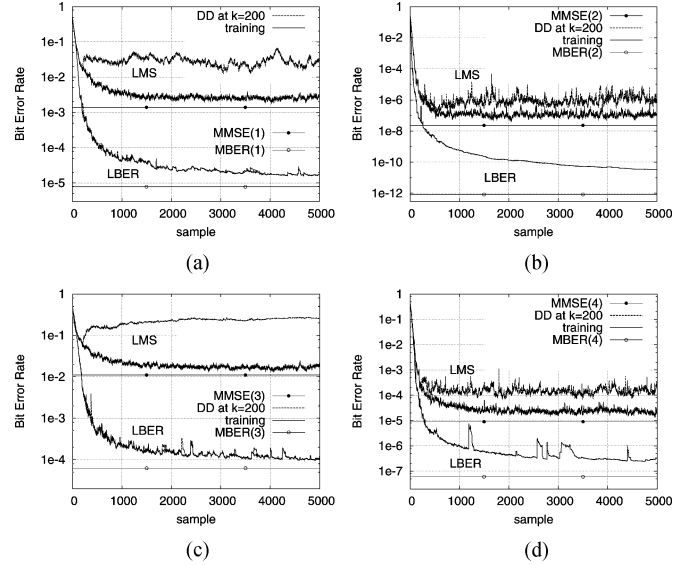

Fig. 5. Learning curves of the LMS and LBER ST-DFE-MUDs in terms of theoretical BERs averaged over 20 runs for the four-user four-antenna time-invariant system with SNR = 1 dB and given $\mathbf{w}_m(0)$ with the first element being $0.1 - j0.1$ and the rest of elements all being $0.0 + j0.0$ for $1 \leq m \leq 4$, where DD denotes the decision directed adaptation with $\hat{s}_m(k - d)$ substituting $s_m(k - d)$: (a) user 1, (b) user 2, (c) user 3, and (d) user 4. For the LMS algorithm, $\mu = 0.001$ was used for all four users while for the LBER algorithm, $\mu = 0.2$ and $\rho_n^2 = 16\sigma_n^2 \approx 6.34$ for user 1, $\mu = 2.0$ and $\rho_n^2 = 100\sigma_n^2 \approx 39.71$ for user 2, $\mu = 0.2$ and $\rho_n^2 = 16\sigma_n^2 \approx 6.34$ for user 3, and $\mu = 0.4$ and $\rho_n^2 = 20\sigma_n^2 \approx 7.94$ for user 4. In the LBER case, the training and DD learning curves are mostly indistinguishable.

solutions. The theoretical BERs of the MMSE and MBER ST-DFE-MUDs so obtained based on the "estimated" CIRs were averaged over ten "estimations" and the results are illustrated in Fig. 4, in comparison with the performance derived using the perfect channel knowledge. It can be seen from Fig. 4 that the performance degradation due to imperfect channel estimates is less serious for the MBER ST-DFE-MUD than for the MMSE one. We conclude that MBER design is more robust to channel estimation errors than the MMSE design. This agrees with the previous conclusion regarding the robustness to error propagation due to decision feedback errors.

The LMS and LBER ST-DFE-MUDs were also investigated under the condition of SNR= 1 dB. Fig. 5 depicts the learning curves of the two adaptive MUDs, averaged over 20 runs and started from an initial condition of $\mathbf{w}_m(0)$, $1 \leq m \leq 4$. The adaptive MUD operated in two modes, namely, the training mode in which the transmitted symbols $s_m(k - d)$ were known to the receiver and the DD adaptation mode, in which the detected symbols $\hat{s}_m(k - d)$ were used for substituting $s_m(k - d)$. The adaptive algorithmic parameters, namely, the adaptive gain μ of the LMS algorithm, and the adaptive gain μ as well as the kernel variance ρ_n^2 of the LBER technique, were empirically tuned for achieving the best performance in terms of the achievable convergence speed and steady-state BER misadjustment. For example, $\mu = 0.001$ was found to be appropriate for the LMS ST-DFE-MUD, and if μ was increased to 0.005, divergence occurred. It can be seen that the LBER ST-DFE-MUD consistently outperformed the LMS

ST-DFE-MUD. In particular, the LBER ST-DFE-MUD was capable of taking full advantage of DD adaptation, as is clearly demonstrated in Fig. 5, where it can also be seen that the DD LMS ST-DFE-MUD clearly failed to converge to the MMSE solution for user 3. A range of different initial conditions, including $\mathbf{w}_m(0) = \mathbf{w}_{(\mathrm{MMSE})_m}$ for $1 \leq m \leq 4$, were tested, and the same conclusions were observed.

B. Slow Fading System

The system again supported four users with four receive antennas. However, fading channels were simulated and each of the 16 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^{-6}, which for a carrier of 900 MHz and a symbol rate of 3 Msymbols/s corresponded to a user velocity of 1 m/s (3.6 km/h). Continuously fluctuating fading was used, which provided a different fading magnitude and phase for each transmitted symbol. The ST-DFE structure parameters were set to $d = 2$, $n_F = 3$, and $n_B = 2$. 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 $\rho_n^2 = 9\sigma_n^2$. The transmission frame structure consisted of 50 training symbols followed by 450 data symbols. The BER of an adaptive ST-DFE-MUD was calculated using Monte Carlo simulation. Fig. 6 compares the BERs of the LBER ST-DFE-MUD for the four users with those of the LMS-based ones. It can be seen from
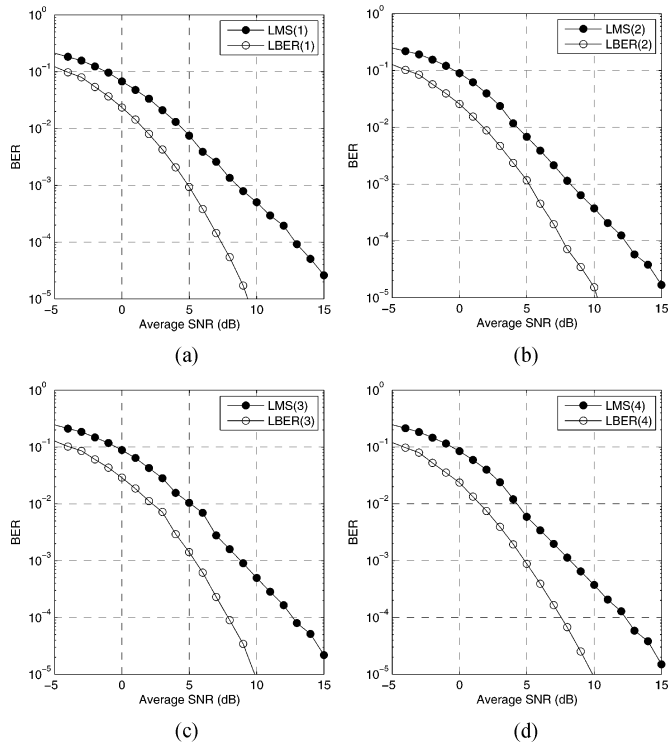
Fig. 6. Bit error rate comparison of the LMS and LBER ST-DFE-MUDs for the four-user four-antenna slow fading system: (a) user 1, (b) user 2, (c) user 3, and (d) user 4.

Fig. 6 that the LBER ST-DFE-MUD consistently outperformed the LMS ST-DFE-MUD for all four users.

VI. CONCLUSIONS

A novel minimum bit error rate design has been proposed for the space-time decision feedback equalization assisted multiuser detector employed in multiple antenna aided space-division multiple-access systems. It has been demonstrated that this MBER design is capable of achieving better performance and hence of improving the attainable system capacity, compared to the classic MMSE design. An adaptive implementation of the MBER ST-DFE-MUD has also been derived based on the LBER algorithm, which has been shown to consistently outperform the classic LMS algorithm and yet maintain a lower computational complexity than the latter for BPSK modulation. Another interesting result observed in this paper is that the MBER ST-DFE-MUD is significantly more robust against the error propagation caused by error-prone detected symbols used in the MUD's feedback loop, in comparison to the standard MMSE ST-DFE-MUD. This is an important advantage, especially in situations, where decision-directed adaption has to be employed.

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, pp. 1740–1751, 1994.

[2] 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.

[3] 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.

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

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

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

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

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

[9] S. Bellofiore, C. A. Balanis, J. Foutz, and A. S. Spanias, "Smart-antenna systems for mobile communication networks—Part 1: Overview and antenna design," *IEEE Antennas Propag. Mag.*, vol. 44, no. 3, pp. 145–154, 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, Synchronisation, 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, pp. 198–218, 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, 1999.

[15] 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, pp. 1240–1247, 2001.

[16] 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, 2005.

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

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

[19] S. Chen and B. Mulgrew, "The minimum-SER linear-combiner decision feedback equalizer," *Proc. Inst. Elect. Eng. Commun.*, vol. 146, no. 6, pp. 347–353, 1999.

[20] 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, 2000.

[21] B. Mulgrew and S. Chen, "Stochastic gradient minimum-BER decision feedback equalisers," in *Proc. IEEE Symp. Adaptive Systems Signal Process., Commun. Contr.*, 2000, pp. 93–98.

[22] K. A. Phillips, J. H. Reed, and W. H. Tranter, "Minimum BER adaptive filtering," in *Proc. ICC'2000*, 2000, vol. 3, pp. 1675–1680.

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

[24] 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, pp. 2092–2101, 2004.

[25] R. Irmer, R. Habendorf, W. Rave, and G. Fettweis, "Overloaded TDD-CDMA cells with multiuser transmission," in *Proc. 2004 ITG Workshop Smart Antennas*, 2004, pp. 235–242.

[26] J. Wehinger, V. R. Anreddy, C. F. Mecklenbräuker, S. Paul, and C. Antón-Haro, "Adaptive minimum bit error rate space-time rake receiver for the uplink of UMTS frequency division duplex mode," in *Proc. 3rd IEEE Int. Symp. Signal Process. Inf. Technol.*, 2003, pp. 294–297.

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

[28] L. Hanzo, C. H. Wong, and M. S. Yee, *Adaptive Wireless Transceivers: Turbo-Coded, Turbo-Equalised and Space-Time Coded TDMA, CDMA and OFDM Systems*. Chichester, U.K.: Wiley/IEEE Press, 2002.

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

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

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

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

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

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

Since 1999, he has been with the School of Electronics and Computer Science, the University of Southampton, U.K. He previously held research and academic appointments at the Universities of Sheffield, Edinburgh, and Portsmouth, U.K. His recent research works include adaptive signal processing, wireless communications, modelling and identification of nonlinear systems, neural network and machine learning, finite-precision digital controller design, evolutionary computation methods, and optimization. He has published more than 240 research papers.

Dr. Chen is listed among the most highly cited researchers in engineering by Institute for Scientific Information.
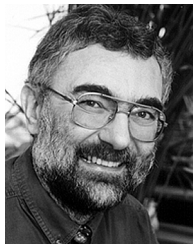
Lajos Hanzo (M'91–SM'92–F'04) received the master's degree in electronics in 1976 and the doctorate degree in 1983. He received the D.Sc. degree from the University of Southampton, U.K., in 2004.

During his 28-year career in telecommunications, he has held various research and academic posts in Hungary, Germany, and the United Kingdom. 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 on mobile radio communications, published in excess of 550 research papers, organized and chaired conference sessions, presented overview lectures, and received a number of distinctions. He is an enthusiastic supporter of industrial–academic liaison. He also offers a range of industrial research overview courses. He is a Non-Executive Director of the Virtual Centre of Excellence in mobile communications, U.K., and an Executive Board Member of the Pan-European Network of Excellence known as NEWCOM.

Prof. Hanzo is a Fellow of the Royal Academy of Engineering (FREng), U.K., and the Institution of Electrical Engineers, U.K. He is an IEEE Distinguished Lecturer of the Communications Society and the Vehicular Technology (VT) Society. He is a Governor of the IEEE VT society.

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

He has published two research papers in wireless communications.

Mr. Livingstone has received several prizes, including an IEEE Telecommunications Project Prize for his undergraduate project. He received the 2005 Lord Lloyd of Kilgerran Memorial Prize for undergraduate achievement, administered by the Institution of Electrical Engineers, U.K.