Space-Time Decision Feedback Equalisation Using a Minimum Bit Error Rate Design for Single-Input Multi-Output Channels

S. Chen, A. Wolfgang, Y. Shi and L. Hanzo

School of Electronics and Computer Science
University of Southampton, Highfield
Southampton SO17 1BJ, U.K.

Abstract

This contribution proposes a minimum bit error rate (MBER) decision feedback equaliser (DFE) designed for single-input multiple-output (SIMO) systems employing a quadrature phase shift keying (QPSK) modulation scheme. It is shown that this MBER design is superior over the standard minimum mean square error DFE in the SIMO scenario considered, in terms of the achievable system bit error rate. A sample-by-sample adaptive implementation of this MBER DFE is derived, which is referred to as the least bit error rate (LBER) algorithm. It is shown that for SIMO systems using a QPSK scheme, the LBER algorithm has a similar computational complexity as the simple least mean square (LMS) algorithm. Simulation results demonstrate that the proposed adaptive LBER-based DFE outperforms the adaptive LMS-based DFE, in both stationary and fading cases.

Index Terms

Single-input multiple-output, multiple antennas, space-time processing, decision feedback equaliser, minimum mean square error, minimum bit error rate

I. Introduction

Smart antenna aided space-time (ST) processing plays an increasingly important role in wireless communications [1]-[8]. With the aid of smart antenna arrays and by exploiting both the space and time dimensions, ST processing is capable of effectively improving the achievable system capacity, coverage and quality of service by suppressing both intersymbol interference and co-channel interference. The family of single-input multiple-output (SIMO) systems has enjoyed popularity owing to its simplicity. A SIMO system consists of a single-antenna transmitter and a receiver equipped with multiple antennas. A ST equaliser [9]-[12] based on this SIMO structure is capable of mitigating the channel impairments arising from hostile multipath propagation. The standard ST equalisation design is based on the well-understood minimum mean square error (MMSE) criterion, which has been successfully employed in a range of detection problems, such as classic channel equalisers, multiuser detectors, beamformers, space-time equalisers, etc.

The financial support of the EU under the auspices of the Phoenix and NEWCOM projects is gratefully acknowledged.

However, for a communication system it is the system's achievable bit error rate (BER), not the mean square error (MSE) value, that really matters and minimising the MSE does not guarantee minimising the BER. Hence the introduction of the novel minimum BER (MBER) criterion opened a new chapter in the optimisation of communications receivers and its design trade-offs have to be documented in contrast to those of the classic but actually still unexhausted MMSE and other often-used optimisation criteria. We will demonstrate that in many respects the MBER optimisation criterion is significantly more powerful, than the MMSE criterion, but naturally, it requires more design attention. For single-input single-output (SISO) equalisation (time-only processing) [13]-[20] and adaptive beamforming (space-only processing) [21]-[24], it has been demonstrated that the receiver design based on the MBER criterion outperforms that based on the MMSE criterion in terms of the achievable BER. In this paper, we develop these ideas further to a combined ST equaliser and invoke the MBER design for ST decision feedback equalisers (ST-DFE) employed in SIMO systems.

In order to keep our notations and the associated concepts relatively simple, we have used a quadrature phase shift keying (QPSK) modulation scheme. However, the proposed approach may be extended to higher order multi-level modulation schemes, as it was demonstrated in the context of SISO equalisation [19],[20]. In simple conceptual terms the underlying detection approach exploits the following ideas. When communicating over a channel inflicting additive Gaussian noise, if the probability density function (PDF) of the received signal is known, closed-form formulae may be derived for the BER as a function of the ST-DFE weights. Given an explicit BER formula, we may set its derivatives to zero with respect to the ST-DFE weights, in order to find the MBER solution. However, when the PDF of the channel's output is unknown and hence no explicit BER formula is available, then it is possible to initialise the ST-DFE's weights for example to the MMSE solution and invoke a gradient-type adaptive algorithm for arriving at a near-MBER solution.

Therefore a novel contribution of this paper is that an efficient adaptive implementation of this MBER ST-DFE is investigated. More specifically, by adopting the classic Parzen window estimation technique for modelling the PDF [25]-[27] and using a stochastic approximation strategy [20],[23], a sample-by-sample adaptive algorithm, referred to as the least bit error rate (LBER) technique, is developed for training the ST-DFE. It is then shown that this LBER ST-DFE has a similarly low computational complexity to the least mean square (LMS) ST-DFE. Moreover, it is demonstrated in our simulation study that the LBER ST-DFE outperforms the classical LMS ST-DFE in fading environments, as it does not aim for minimising the system's MSE and therefore does not suffer from numerical ill-conditioning problems. By contrast, it is well-known that the performance of the LMS ST-DFE degrades considerably in hostile propagation environments.

The outline of the paper is as follows. Section II introduces the SIMO system model considered and defines the ST-DFE structure, while Section III derives the MBER design for the ST-DFE employed in SIMO systems. Adaptive implementation of the MBER ST-DFE is considered in Section IV, where both a block-based and a stochastic adaptive algorithms are portrayed. Finally, Section V describes our simulation studies and Section VI offers our conclusions.

II. System Model

Consider the SIMO system employing a single transmit antenna and L (> 1) receive antennas, as depicted in Fig. 1, where $s(t)$ is the transmitted signal, $x_l(t)$ denotes the lth receive antenna's output signal and $n_l(t)$ the lth channel's noise. The received signals are sampled at the symbol rate in order to obtain the L antennas' output samples $x_l(k)$, $1 \leq l \leq L$, which are passed to a ST-DFE, as shown in Fig. 2. The received signal sample $x_l(k)$ for the lth antenna can be expressed as

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

where $n_l(k)$ is a complex-valued Gaussian white noise with $E[|n_l(k)|^2] = 2\sigma_n^2$, $\bar{x}_l(k)$ denotes the noise-free part of the lth channel's output, the transmitted symbol sequence $s(k) = s_R(k) + js_I(k)$ takes values from the QPSK symbol set $\{\pm 1 \pm j\}$, and $c_{i,l}$ are the complex-valued taps of the lth channel impulse response (CIR) having a length of n_c. For notational simplicity, we have assumed that each of the L channels has the same length of n_c. The soft output of the ST-DFE is given by

$$y(k) = \sum_{l=1}^{L} \left(\sum_{i=0}^{m-1} w_{i,l}^* x_l(k-i) + \sum_{i=1}^{n_b} b_{i,l}^* \hat{s}(k-d-i) \right), \tag{2}$$

where $\hat{s}(k-d)$ is the estimate of $s(k-d)$, d is the decision delay, m and n_b are the feedforward and feedback filter orders, respectively, while $w_{i,l}$ and $b_{i,l}$ are the coefficients of the lth feedforward and feedback filters, respectively. Let

$$\begin{aligned} \mathbf{w}_l &= [w_{0,l}\ w_{1,l} \cdots w_{m-1,l}]^T, \\ \mathbf{x}_l(k) &= [x_l(k)\ x_l(k-1) \cdots x_l(k-m+1)]^T, \\ \mathbf{b}_l &= [b_{1,l}\ b_{2,l} \cdots b_{n_b,l}]^T, \\ \hat{\mathbf{s}}_b(k) &= [\hat{s}(k-d-1)\ \hat{s}(k-d-2) \cdots \hat{s}(k-d-n_b)]^T, \end{aligned} \tag{3}$$

and let us define furthermore

$$\begin{aligned} \mathbf{w} &= [\mathbf{w}_1^T\ \mathbf{w}_2^T \cdots \mathbf{w}_L^T]^T, \\ \mathbf{x}(k) &= [\mathbf{x}_1^T(k)\ \mathbf{x}_2^T(k) \cdots \mathbf{x}_L^T(k)]^T, \\ \mathbf{b} &= \sum_{l=1}^{L} \mathbf{b}_l. \end{aligned} \tag{4}$$

Then the ST-DFE output can be expressed as

$$y(k) = \sum_{l=1}^{L} \left(\mathbf{w}_l^H \mathbf{x}_l(k) + \mathbf{b}_l^H \hat{\mathbf{s}}_b(k) \right) = \mathbf{w}^H \mathbf{x}(k) + \mathbf{b}^H \hat{\mathbf{s}}_b(k). \tag{5}$$

We will choose the ST-DFE structure's parameters as follows: $d = n_c - 1$, $m = n_c$ and $n_b = n_c - 1$. For the SISO 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 and therefore they guarantee an adequate performance [14],[15]. Using $m = n_c$ and $d = n_b = n_c - 1$, the received signal vector of the lth channel can be expressed as

$$\mathbf{x}_l(k) = \bar{\mathbf{x}}_l(k) + \mathbf{n}_l(k) = \mathbf{C}_{F_l} \mathbf{s}_f(k) + \mathbf{C}_{B_l} \mathbf{s}_b(k) + \mathbf{n}_l(k), \tag{6}$$

where

$$\begin{aligned} \mathbf{s}_f(k) &= [s(k)\ s(k-1) \cdots s(k-d)]^T, \\ \mathbf{s}_b(k) &= [s(k-d-1)s(k-d-2) \cdots s(k-d-n_b)]^T, \\ \mathbf{n}_l(k) &= [n_l(k)\ n_l(k-1) \cdots n_l(k-m+1)]^T, \end{aligned} \tag{7}$$

and the $m \times (d+1)$ and $m \times n_b$ dimensional CIR matrices $\mathbf{C}_{F_l}$ and $\mathbf{C}_{B_l}$ are given by

$$\mathbf{C}_{F_l} = \begin{bmatrix} c_{0,l} & c_{1,l} & \cdots & c_{n_c-1,l} \\ 0 & c_{0,l} & \ddots & \vdots \\ \vdots & \ddots & \ddots & c_{1,l} \\ 0 & \cdots & 0 & c_{0,l} \end{bmatrix} \tag{8}$$

and

$$\mathbf{C}_{B_l} = \begin{bmatrix} 0 & \cdots & 0 \\ c_{n_c-1,l} & \ddots & \vdots \\ \vdots & \ddots & 0 \\ c_{1,l} & \cdots & c_{n_c-1,l} \end{bmatrix}, \tag{9}$$

respectively. Under the assumption that the past decisions are correct, we have $\hat{\mathbf{s}}_b(k) = \mathbf{s}_b(k)$ and the lth received signal vector may be expressed as $\mathbf{x}_l(k) = \mathbf{C}_{F_l} \mathbf{s}_f(k) + \mathbf{C}_{B_l} \hat{\mathbf{s}}_b(k) + \mathbf{n}_l(k)$. Thus, the decision feedback may be viewed as a translation of the original observation space $\mathbf{x}_l(k)$ into a new space $\mathbf{r}_l(k)$:

$$\mathbf{r}_l(k) \stackrel{\triangle}{=} \mathbf{x}_l(k) - \mathbf{C}_{B_l} \hat{\mathbf{s}}_b(k) = \mathbf{C}_{F_l} \mathbf{s}_f(k) + \mathbf{n}_l(k) = \bar{\mathbf{r}}_l(k) + \mathbf{n}_l(k). \tag{10}$$

Let us now define

$$\begin{aligned} \mathbf{r}(k) &= [\mathbf{r}_1^T(k)\ \mathbf{r}_2^T(k) \cdots \mathbf{r}_L^T(k)]^T, \\ \mathbf{n}(k) &= [\mathbf{n}_1^T(k)\ \mathbf{n}_2^T(k) \cdots \mathbf{n}_L^T(k)]^T. \end{aligned} \tag{11}$$

In the translated observation space $\mathbf{r}(k)$, the original ST-DFE (5) is "translated" into a ST "linear equaliser" described as:

$$y(k) = \mathbf{w}^H \mathbf{r}(k) = \mathbf{w}^H (\bar{\mathbf{r}}(k) + \mathbf{n}(k)) = \bar{y}(k) + e(k), \tag{12}$$

where $e(k)$ is Gaussian distributed, having a zero mean and $E[|e(k)|^2] = 2\mathbf{w}^H\mathbf{w}\sigma_n^2$. The elements of $\mathbf{r}(k)$ can be computed recursively according to [14],[15]:

$$\left.\begin{aligned} r_l(k-i) &= z^{-1}r_l(k-i+1) - c_{n_c-i,l}\hat{s}(k-d-1), \\ &\quad \text{for } i = m-1, m-2, \cdots, 1 \\ r_l(k) &= x_l(k) \end{aligned}\right\}, \tag{13}$$

where z^{-1} is interpreted as the unit delay operator. 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 equaliser structure of (12) and (13). This equivalent ST-DFE is illustrated in Fig. 3.

We define the overall signal to noise ratio (SNR) of the SIMO system under consideration as

$$\text{SNR} = \frac{1}{L\sigma_n^2}\sum_{l=1}^{L}\sum_{i=0}^{n_c-1}|c_{i,l}|^2. \tag{14}$$

The following decision rule is used for providing an estimate of $s(k-d)$:

$$\hat{s}(k-d) = \text{sgn}(y_R(k)) + j\text{sgn}(y_I(k)), \tag{15}$$

where $y_R(k) = \Re[y(k)]$ and $y_I(k) = \Im[y(k)]$ are the real and imaginary parts of $y(k)$, respectively, and sgn($\bullet$) is the sign function. Let us now define the following $(Lm) \times (d+1)$ dimensional overall CIR matrix

$$\mathbf{C}_F = \begin{bmatrix} \mathbf{C}_{F_1} \\ \vdots \\ \mathbf{C}_{F_L} \end{bmatrix} = [\mathbf{c}_{F,0}\ \mathbf{c}_{F,1} \cdots \mathbf{c}_{F,d}]\,. \tag{16}$$

Note that the last column of $\mathbf{C}_F$ is simply given by:

$$\mathbf{c}_{F,d} = [c_{n_c-1,1} \cdots c_{1,1}\ c_{0,1} \cdots c_{n_c-1,L} \cdots c_{1,L}\ c_{0,L}]^T\,. \tag{17}$$

Let us also define the combined impulse response of the channels and the equaliser as $\mathbf{f}$, which is given by

$$\mathbf{f}^T = [f_0\ f_1 \cdots f_d] = \mathbf{w}^H\mathbf{C}_F = \left[\mathbf{w}^H\mathbf{c}_{F,0}\ \mathbf{w}^H\mathbf{c}_{F,1} \cdots \mathbf{w}^H\mathbf{c}_{F,d}\right]. \tag{18}$$

The ST-DFE output can then be expressed as

$$y(k) = f_d s(k-d) + \sum_{i=0}^{d-1} f_i s(k-i) + e(k). \tag{19}$$

The first term in (19) is the desired signal, while the second term represents the residual ISI. Provided that f_d is real and positive, the decision rule (15) is optimal. We point out that in general f_d is complex-valued, and the rotation operation of

$$\mathbf{w}^{\text{new}} = \frac{f_d^{\text{old}}}{|f_d^{\text{old}}|}\mathbf{w}^{\text{old}} \tag{20}$$

may be applied to the weight vector $\mathbf{w}$ for the sake of rendering f_d real and positive. This rotation is a linear transformation and does not alter the BER, but it allows the simple optimal decision rule (15) to be adopted.

III. Minimum Bit Error Rate Design

Classically, the equaliser weight vector $\mathbf{w}$ is determined by minimising the MSE $E[|s(k-d)-y(k)|^2]$, which leads to the following MMSE solution

$$\mathbf{w}_{\text{MMSE}} = \left(\mathbf{C}_F\mathbf{C}_F^H + \sigma_n^2\mathbf{I}_{Lm}\right)^{-1}\mathbf{c}_{F,d}, \tag{21}$$

with $\mathbf{I}_{Lm}$ being an $Lm \times Lm$ dimensional identity matrix. An adaptive implementation of the MMSE solution may be realised for example using the LMS algorithm. The main contribution of this paper is to derive the MBER solution for the weight vector of the ST-DFE and develop an adaptive MBER ST-DFE for the SIMO systems. Let us denote the $N_s = 4^{d+1}$ number of possible transmitted symbol sequences of $\mathbf{s}_f(k)$ as $\mathbf{s}^{(q)}$, $1 \le q \le N_s$. Denote furthermore the last element of $\mathbf{s}^{(q)}$, corresponding to the symbol $s(k-d)$, as $s_d^{(q)}$. The noise-free part of the equaliser input signal, namely $\bar{\mathbf{r}}(k)$, assumes values from the finite signal set defined as:

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

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

$$\mathcal{R}_{\pm,\pm} \triangleq \{\bar{\mathbf{r}}^{(q)} \in \mathcal{R} : s(k-d) = \pm 1 \pm j\}. \tag{23}$$

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

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

and $\mathcal{Y}$ can be divided into the four subsets conditioned on the value of $s(k-d)$:

$$\mathcal{Y}_{\pm,\pm} \triangleq \{\bar{y}^{(q)} \in \mathcal{Y} : s(k-d) = \pm 1 \pm j\}. \tag{25}$$

It is readily seen that the conditional PDF of $y(k)$ given $s(k-d) = (+1+j)$ is:

$$p(y|+1+j) = \frac{1}{N_{sb}} \sum_{\bar{y}^{(q)} \in \mathcal{Y}_{+,+}} \frac{1}{2\pi\sigma_n^2\mathbf{w}^H\mathbf{w}} \exp\left(-\frac{\left|y - \bar{y}^{(q)}\right|^2}{2\sigma_n^2\mathbf{w}^H\mathbf{w}}\right), \tag{26}$$

where $N_{sb} = N_s/4$ is the number of the constellation points in $\mathcal{Y}_{+,+}$. With the notations $y = y_R + jy_I$ and $\bar{y}^{(q)} = \bar{y}_R^{(q)} + j\bar{y}_I^{(q)}$, the two marginal conditional PDFs are given by

$$p(y_R|+1+j) = \frac{1}{N_{sb}} \sum_{\bar{y}^{(q)} \in \mathcal{Y}_{+,+}} \frac{1}{\sqrt{2\pi\sigma_n^2\mathbf{w}^H\mathbf{w}}} \exp\left(-\frac{\left(y_R - \bar{y}_R^{(q)}\right)^2}{2\sigma_n^2\mathbf{w}^H\mathbf{w}}\right) \tag{27}$$

and

$$p(y_I|+1+j) = \frac{1}{N_{sb}} \sum_{\bar{y}^{(q)} \in \mathcal{Y}_{+,+}} \frac{1}{\sqrt{2\pi\sigma_n^2\mathbf{w}^H\mathbf{w}}} \exp\left(-\frac{\left(y_I - \bar{y}_I^{(q)}\right)^2}{2\sigma_n^2\mathbf{w}^H\mathbf{w}}\right), \tag{28}$$

respectively. Let us define

$$P_{E_R}(\mathbf{w}) \triangleq \text{Prob}(\Re[\hat{s}(k-d)] \neq \Re[s(k-d)]) = \text{Prob}(\hat{s}_R(k-d) \neq s_R(k-d)) \tag{29}$$

and

$$P_{E_I}(\mathbf{w}) \triangleq \text{Prob}(\Im[\hat{s}(k-d)] \neq \Im[s(k-d)]) = \text{Prob}(\hat{s}_I(k-d) \neq s_I(k-d)). \tag{30}$$

Then the BER of the ST-DFE associated with the equaliser weight vector $\mathbf{w}$ is given by:

$$P_E(\mathbf{w}) = \frac{1}{2}\left(P_{E_R}(\mathbf{w}) + P_{E_I}(\mathbf{w})\right). \tag{31}$$

Noting the decision rule (15) and the two marginal conditional PDFs given in (27) and (28), it can readily be shown that

$$P_{E_R}(\mathbf{w}) = \frac{1}{N_{sb}} \sum_{\bar{y}^{(q)} \in \mathcal{Y}_{+,+}} Q\left(g_R^{(q)}(\mathbf{w})\right) \tag{32}$$

and

$$P_{E_I}(\mathbf{w}) = \frac{1}{N_{sb}} \sum_{\bar{y}^{(q)} \in \mathcal{Y}_{+,+}} Q\left(g_I^{(q)}(\mathbf{w})\right), \tag{33}$$

where

$$Q(u) = \frac{1}{\sqrt{2\pi}} \int_u^\infty \exp\left(-\frac{v^2}{2}\right) dv, \tag{34}$$

$$g_R^{(q)}(\mathbf{w}) = \frac{\text{sgn}(\Re[s_d^{(q)}])\bar{y}_R^{(q)}}{\sigma_n\sqrt{\mathbf{w}^H\mathbf{w}}} = \frac{\text{sgn}(s_{R,d}^{(q)})\Re[\mathbf{w}^H\bar{\mathbf{r}}^{(q)}]}{\sigma_n\sqrt{\mathbf{w}^H\mathbf{w}}} \tag{35}$$

and

$$g_I^{(q)}(\mathbf{w}) = \frac{\text{sgn}(\Im[s_d^{(q)}])\bar{y}_I^{(q)}}{\sigma_n\sqrt{\mathbf{w}^H\mathbf{w}}} = \frac{\text{sgn}(s_{I,d}^{(q)})\Im[\mathbf{w}^H\bar{\mathbf{r}}^{(q)}]}{\sigma_n\sqrt{\mathbf{w}^H\mathbf{w}}}. \tag{36}$$

Note that the BER is invariant to a positive scaling of $\mathbf{w}$. Similarly, the BER may be calculated based on anyone of the other three subsets, namely on $\mathcal{Y}_{+,-}$, $\mathcal{Y}_{-,+}$ or $\mathcal{Y}_{-,-}$.

The ST-DFE MBER solution is then defined as the weight vector, minimising the error probability, which is formulated as:

$$\mathbf{w}_{\text{MBER}} = \arg\min_{\mathbf{w}} P_E(\mathbf{w}). \tag{37}$$

As in any optimisation problem, the MBER solution may be found by setting the derivative of $P_E(\mathbf{w})$ to zero. The gradient of $P_E(\mathbf{w})$ with respect to $\mathbf{w}$ is given by:

$$\nabla P_E(\mathbf{w}) = \frac{1}{2}\left(\nabla P_{E_R}(\mathbf{w}) + \nabla P_{E_I}(\mathbf{w})\right), \tag{38}$$

and it can be shown that

$$\nabla P_{E_R}(\mathbf{w}) = \frac{1}{2N_{sb}\sqrt{2\pi}\sigma_n\sqrt{\mathbf{w}^H\mathbf{w}}} \sum_{\bar{y}^{(q)} \in \mathcal{Y}_{+,+}} \exp\left(-\frac{\left(\bar{y}_R^{(q)}\right)^2}{2\sigma_n^2\mathbf{w}^H\mathbf{w}}\right) \text{sgn}\left(s_{R,d}^{(q)}\right)\left(\frac{\bar{y}_R^{(q)}\mathbf{w}}{\mathbf{w}^H\mathbf{w}} - \bar{\mathbf{r}}^{(q)}\right) \tag{39}$$

and

$$\nabla P_{E_I}(\mathbf{w}) = \frac{1}{2N_{sb}\sqrt{2\pi}\sigma_n\sqrt{\mathbf{w}^H\mathbf{w}}} \sum_{\bar{y}^{(q)}\in\mathcal{Y}_{+,+}} \exp\left(-\frac{\left(\bar{y}_I^{(q)}\right)^2}{2\sigma_n^2\mathbf{w}^H\mathbf{w}}\right) \mathrm{sgn}\left(s_{I,d}^{(q)}\right)\left(\frac{\bar{y}_I^{(q)}\mathbf{w}}{\mathbf{w}^H\mathbf{w}} + j\bar{\mathbf{r}}^{(q)}\right). \quad (40)$$

Given the gradient expressions (38)–(40), the optimisation problem (37) can be solved iteratively by commencing the iterations from an appropriate initialisation point, such as the MMSE solution, using a gradient-based optimisation algorithm. Since the BER is invariant to a positive scaling of $\mathbf{w}$, it is computationally advantageous to normalise $\mathbf{w}$ to a unit-length after every iteration, so that the gradient expressions (39) and (40) can be simplified to:

$$\nabla P_{E_R}(\mathbf{w}) = \frac{1}{2N_{sb}\sqrt{2\pi}\sigma_n} \sum_{\bar{y}^{(q)}\in\mathcal{Y}_{+,+}} \exp\left(-\frac{\left(\bar{y}_R^{(q)}\right)^2}{2\sigma_n^2}\right) \mathrm{sgn}\left(s_{R,d}^{(q)}\right)\left(\bar{y}_R^{(q)}\mathbf{w} - \bar{\mathbf{r}}^{(q)}\right) \quad (41)$$

and

$$\nabla P_{E_I}(\mathbf{w}) = \frac{1}{2N_{sb}\sqrt{2\pi}\sigma_n} \sum_{\bar{y}^{(q)}\in\mathcal{Y}_{+,+}} \exp\left(-\frac{\left(\bar{y}_I^{(q)}\right)^2}{2\sigma_n^2}\right) \mathrm{sgn}\left(s_{I,d}^{(q)}\right)\left(\bar{y}_I^{(q)}\mathbf{w} + j\bar{\mathbf{r}}^{(q)}\right). \quad (42)$$

The simplified conjugate gradient algorithm of [28],[29] provides an efficient means of finding a MBER solution for the optimisation problem formulated in (37).

IV. Adaptive Minimum Bit Error Rate Implementation

As usual, the evaluation of the error probability requires the knowledge of the PDF of the ST-DFE's output signal $y(k)$. The PDF of $y(k)$ can be expressed explicitly by:

$$p(y) = \frac{1}{N_s 2\pi\sigma_n^2\mathbf{w}^H\mathbf{w}} \sum_{q=1}^{N_s} \exp\left(-\frac{\left|y - \bar{y}^{(q)}\right|^2}{2\sigma_n^2\mathbf{w}^H\mathbf{w}}\right), \quad (43)$$

and the associated BER can alternatively be calculated with the aid of two "marginal" BERs given by

$$P_{E_R}(\mathbf{w}) = \frac{1}{N_s}\sum_{q=1}^{N_s} Q\left(g_R^{(q)}(\mathbf{w})\right) \quad (44)$$

and

$$P_{E_I}(\mathbf{w}) = \frac{1}{N_s}\sum_{q=1}^{N_s} Q\left(g_I^{(q)}(\mathbf{w})\right), \quad (45)$$

where the summations are carried out over all the N_s number of elements of the set $\bar{y}^{(q)} \in \mathcal{Y}$. In reality, the PDF of $y(k)$ is channel dependent and hence it is unknown. Some form of PDF estimation is required for supporting the adaptive implementation of the MBER ST-DFE.

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

The key to the efficient adaptive implementation of the MBER solution is generating an up-to-date and accurate estimate of the PDF (43). Parzen window or kernel density estimation method [25]–[27] constitutes an efficient method of estimating a PDF. Specifically, the Parzen window method estimates a PDF using a window or block of the ST-DFE output signal $y(k)$ by placing a symmetric unimodal kernel function on each $y(k)$ sample. This kernel density estimation technique is capable of producing reliable PDF estimates with the aid of short data records and it is particularly natural when dealing with Gaussian mixtures, such as the one given in (43). In our particular application, it is natural to choose a Gaussian kernel function having a kernel width of $\rho_n\sqrt{\mathbf{w}^H\mathbf{w}}$ that is similar to the noise standard deviation of $\sigma_n\sqrt{\mathbf{w}^H\mathbf{w}}$. Given a block of K training samples $\{\mathbf{r}(k), s(k-d)\}_{k=1}^{K}$, a kernel density estimate of the PDF in (43) is readily given by:

$$\hat{p}(y) = \frac{1}{K2\pi\rho_n^2\mathbf{w}^H\mathbf{w}}\sum_{k=1}^{K}\exp\left(-\frac{|y-y(k)|^2}{2\rho_n^2\mathbf{w}^H\mathbf{w}}\right), \tag{46}$$

where the radius or scaling parameter ρ_n is related to the standard deviation σ_n of the system's noise. The accuracy analysis of Parzen window density estimate is well documented in the literature [25]–[27]. The PDF estimate (46) is known to possess a mean integrated square error convergence rate at an order of K^{-1} [25].

Based on the estimated PDF of (46), the estimated BER is given by:

$$\hat{P}_E(\mathbf{w}) = \frac{1}{2}\left(\hat{P}_{E_R}(\mathbf{w}) + \hat{P}_{E_I}(\mathbf{w})\right) = \frac{1}{2K}\sum_{k=1}^{K}\left(Q\left(\hat{g}_R^{(k)}(\mathbf{w})\right) + Q\left(\hat{g}_I^{(k)}(\mathbf{w})\right)\right), \tag{47}$$

with

$$\hat{g}_R^{(k)}(\mathbf{w}) = \frac{\mathrm{sgn}(s_R(k-d))y_R(k)}{\rho_n\sqrt{\mathbf{w}^H\mathbf{w}}} \tag{48}$$

and

$$\hat{g}_I^{(k)}(\mathbf{w}) = \frac{\mathrm{sgn}(s_I(k-d))y_I(k)}{\rho_n\sqrt{\mathbf{w}^H\mathbf{w}}}. \tag{49}$$

The gradient of $\hat{P}_E(\mathbf{w})$ can readily be calculated with the aid of (39) and (40) as follows:

$$\nabla\hat{P}_{E_R}(\mathbf{w}) = \frac{1}{2K\sqrt{2\pi}\rho_n\sqrt{\mathbf{w}^H\mathbf{w}}}\sum_{k=1}^{K}\exp\left(-\frac{y_R^2(k)}{2\rho_n^2\mathbf{w}^H\mathbf{w}}\right)\mathrm{sgn}(s_R(k-d))\left(\frac{y_R(k)\mathbf{w}}{\mathbf{w}^H\mathbf{w}} - \mathbf{r}(k)\right) \tag{50}$$

and

$$\nabla\hat{P}_{E_I}(\mathbf{w}) = \frac{1}{2K\sqrt{2\pi}\rho_n\sqrt{\mathbf{w}^H\mathbf{w}}}\sum_{k=1}^{K}\exp\left(-\frac{y_I^2(k)}{2\rho_n^2\mathbf{w}^H\mathbf{w}}\right)\mathrm{sgn}(s_I(k-d))\left(\frac{y_I(k)\mathbf{w}}{\mathbf{w}^H\mathbf{w}} + j\mathbf{r}(k)\right). \tag{51}$$

Upon substituting $\nabla P_E(\mathbf{w})$ by $\nabla\hat{P}_E(\mathbf{w})$ in the simplified conjugate gradient updating mechanism, for example, a block-data based adaptive algorithm is obtained [29],[24], where the step size μ and the radius parameter ρ_n are two algorithmic parameters which 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

In this section our aim is to develop a sample-by-sample adaptive implementation of the MBER ST-DFE. In the Parzen window estimate (46), the kernel width $\rho_n\sqrt{\mathbf{w}^H\mathbf{w}}$ explicitly depends on the ST-DFE's weight vector $\mathbf{w}$, which was so arranged because the true density of (43) also depends on the weight vector $\mathbf{w}$. However, the BER is invariant to $\mathbf{w}^H\mathbf{w}$, and a constant kernel width ρ_n may also be adopted in the density estimate. A particular advantage of using a constant kernel width ρ_n, rather than $\rho_n\sqrt{\mathbf{w}^H\mathbf{w}}$ in the density estimate is that the gradient of the resultant estimated BER has a significantly simpler form, which leads to a considerable reduction in computational complexity. This is particular relevant in the derivation of stochastic gradient ST-DFE weight updating mechanisms. Adopting this approach, an alternative fixed kernel-width based Parzen window estimate of the true PDF (43) is given by

$$\tilde{p}(y) = \frac{1}{K2\pi\rho_n^2}\sum_{k=1}^{K}\exp\left(-\frac{|y-y(k)|^2}{2\rho_n^2}\right), \tag{52}$$

and the resultant approximate BER formula becomes

$$\tilde{P}_E(\mathbf{w}) = \frac{1}{2}\left(\tilde{P}_{E_R}(\mathbf{w}) + \tilde{P}_{E_I}(\mathbf{w})\right) = \frac{1}{2K}\sum_{k=1}^{K}\left(Q\left(\tilde{g}_R^{(k)}(\mathbf{w})\right) + Q\left(\tilde{g}_I^{(k)}(\mathbf{w})\right)\right), \tag{53}$$

where we have:

$$\tilde{g}_R^{(k)}(\mathbf{w}) = \frac{\mathrm{sgn}(s_R(k-d))y_R(k)}{\rho_n} \tag{54}$$

and

$$\tilde{g}_I^{(k)}(\mathbf{w}) = \frac{\mathrm{sgn}(s_I(k-d))y_I(k)}{\rho_n}. \tag{55}$$

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 ST-DFE's weight vector $\mathbf{w}$, consider a single-sample estimate of $p(y)$, namely:

$$\tilde{p}(y,k) = \frac{1}{2\pi\rho_n^2}\exp\left(-\frac{|y-y(k)|^2}{2\rho_n^2}\right). \tag{56}$$

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

$$\nabla\tilde{P}_E(\mathbf{w},k) = \frac{\left(-\mathrm{sgn}(s_R(k-d))\exp\left(-\frac{y_R^2(k)}{2\rho_n^2}\right) + j\mathrm{sgn}(s_I(k-d))\exp\left(-\frac{y_I^2(k)}{2\rho_n^2}\right)\right)}{4\sqrt{2\pi}\rho_n}\mathbf{r}(k) \tag{57}$$

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

$$\mathbf{w}(k+1) = \mathbf{w}(k) - \mu\nabla\tilde{P}_E(\mathbf{w}(k),k), \tag{58}$$

$$\hat{f}_d = \mathbf{w}^H(k+1)\hat{\mathbf{c}}_{F,d}, \tag{59}$$

$$\mathbf{w}(k+1) = \frac{\hat{f}_d}{|\hat{f}_d|}\mathbf{w}(k+1), \tag{60}$$

where $\hat{\mathbf{c}}_{F,d}$ is the estimated CIR $\mathbf{c}_{F,d}$ given in (17), and 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 steady-state BER misadjustment. Note that there is no need to normalise the weight vector to a unit-length after each update. The CIR $\mathbf{c}_{F,d}$, which is also needed for the sake of performing the space transformation of (13), may be estimated using the standard LMS algorithm.

This LBER ST-DFE has a similar computational complexity to the LMS ST-DFE. For the sake of a comparison, the weight updating equations of the LMS ST-DFE are reproduced here:

$$\mathbf{w}(k+1) = \mathbf{w}(k) + \mu \left(s(k-d) - y(k)\right)^* \mathbf{r}(k), \tag{61}$$

$$\hat{f}_d = \mathbf{w}^H(k+1)\hat{\mathbf{c}}_{F,d}, \tag{62}$$

$$\mathbf{w}(k+1) = \frac{\hat{f}_d}{|\hat{f}_d|}\mathbf{w}(k+1). \tag{63}$$

Let us assume that L number of LMS channel estimators are used for identifying the L channels, which has a computational requirement of $8 \times (Ln_c) + 2$ multiplications and $8 \times (Ln_c)$ additions per channel estimator update. Table I compares the total computational complexity of the LBER ST-DFE to that of the LMS ST-DFE. It is worth emphasising that the performance of the LMS ST-DFE is closely related to the conditioning number of the matrix $\mathbf{C}_F\mathbf{C}_F^H + \sigma_n^2\mathbf{I}_{Lm}$. In fading associated environments with relatively low levels of noise, this matrix may not always be invertible and hence the computationally simple LMS ST-DFE may suffer from serious performance degradation. In such situations, in order to realize the MMSE solution the recursive least squares algorithm may have to be used for the weight updating procedure, which has a significantly higher complexity. By contrast, the LBER ST-DFE appears to be robust, as it does not rely on any matrix inversion. This will be demonstrated in our simulation study.

V. Simulation Study

In all of our simulation based investigations, a perfect channel estimate was assumed in performing the space translation (13) and in calculating f_d. Hence our attention was focused on the performance of the adaptive MBER and MMSE designs, rather than on the adaptive channel estimator.

Time-invariant system: In our simulations, the number of receiver antennas was varied from $L = 1$ to $L = 4$, and each simulated channel had the same CIR of:

$$C_l(z) = (0.1 + j0.1) + (-0.2 - j0.2)z^{-1} + (0.4 + j0.4)z^{-2} + (-0.8 - j0.8)z^{-3}, \tag{64}$$

for $1 \leq l \leq L$. The actual simulated channel was normalised according to $C_l(z)/|C_l(z)|$ for the sake of maintaining unity channel gain. Since the length of the CIRs was $n_c = 4$, the ST-DFE was

characterised by the parameters of $m = 4$, $d = 3$ and $n_b = 3$. The theoretical BERs of the MMSE and MBER ST-DFEs, calculated using the expression (31), are given in Fig. 4 as a function of a varying number of the receiver antennas, ranging from $L = 1$ to $L = 4$. In Fig. 4, the MMSE weight solution was computed using the closed-form expression (21), while the MBER ST-DFE weights were calculated numerically by solving the optimisation problem formulated in (37) using the simplified conjugate gradient algorithm. It can be seen from Fig. 4 that as the number of antennas increased, the achievable performance of the ST-DFE improved. Moreover, the MBER ST-DFE had a substantially better performance than the MMSE ST-DFE, yielding a SNR gain in excess of 5 dB at the BER level of 10^{-4}.

The theoretical BERs shown in Fig. 4 represented the best-case performance, as they were obtained assuming that correct symbols were fed back by the ST-DFE's feedback loop. For the sake of investigating the effects of error propagation, the BERs of the MMSE and MBER ST-DFEs were also calculated using simulations with the error-prone detected symbol being fed back, and the corresponding results are depicted in Fig. 5 for the case of $L = 4$, in comparison to the theoretical BERs. It is interesting to see that for this example the BER performance degradation owing to error propagation was less serious for the MBER ST-DFE, than for the MMSE ST-DFE.

The performance of the LBER ST-DFE algorithm was studied next. For the case of $L = 4$ receiver antennas and for an SNR of 17 dB, Fig. 6 depicts the learning curves of the LBER algorithm in terms of the achievable theoretical BER averaged over 20 random runs, where the initial weight vector was chosen as $\mathbf{w}(0) = \mathbf{w}_{\rm MMSE}$ and the algorithmic parameters are set to $\mu = 0.05$ and $\rho_n^2 = 20\sigma_n^2 \approx 0.4$. The LBER ST-DFE algorithm operated in two modes, namely the training mode in which $s(k-d)$ was known and correct symbols were fed back, and the decision directed (DD) mode in which $\hat{s}(k-d)$ was used for replacing $s(k-d)$ and error-prone detected symbols were fed back. For the scenario investigated, the training-based and DD learning curves of the LBER algorithm were indistinguishable. For the sake of comparison, the training learning curve of the LMS algorithm using $\mu = 0.01$ is also shown in Fig. 6. The DD learning curve of the LMS algorithm, not shown here, was observed to diverge owing to catastrophic error propagation.

Fading system: $L = 4$ receiver antennas were used and each of the four CIRs had the same length of $n_c = 4$. The magnitudes of the complex-valued CIR tap weights $c_{i,l}$ for $0 \leq i \leq 3$ and $1 \leq l \leq 4$ were Rayleigh processes and the associated root mean power of each $c_{i,l}$ were $\sqrt{0.5} + j\sqrt{0.5}$. The ST-DFE structure was therefore defined by decision delay $d = 3$, feedforward order $m = 4$ and feedback order $n_b = 3$. The transmission frame structure consisted of 20 training symbols followed by 200 data

symbols. Frame fading was implemented, in which the CIR taps were faded at the beginning of each frame at a normalised Doppler frequency of 10^{-2} and were then kept constant within the frame. This provided a different fading magnitude and phase for each transmitted frame.

The performance of the LMS and LBER ST-DFEs are compared in Fig. 7, where the BERs were calculated using the actual detected symbols for feedback. From Fig. 7, it can be seen that the BER curve of the LMS ST-DFE first became flat as the noise level decreased and eventually it increased slightly, when the noise level became extremely low. A possible explanation of this performance degradation was as follows. The LMS ST-DFE was sensitive to the eigenvalue spread of the matrix $\mathbf{C}_F\mathbf{C}_F^H + \sigma_n^2\mathbf{I}_{Lm}$. In a fading scenario associated with a low noise level σ_n^2, the associated matrix became ill-conditioned or even non-invertible. This inflicted a substantial performance degradation of the LMS ST-DFE. Auspiciously, the LBER ST-DFE did not suffer from this numerical ill-conditioning problem and hence exhibited a superior performance in comparison to the LMS ST-DFE.

VI. Conclusions

A novel MBER ST-DFE design has been proposed for employment in SIMO systems. It has been demonstrated that this MBER ST-DFE generally outperforms the standard MMSE design in terms of the achievable BER, and therefore this design may be expected to increase the expected system capacity. An adaptive implementation of this MBER design has also been derived using the LBER algorithm. It has been shown that for a QPSK modulation scheme the resultant LBER ST-DFE has a similarly low computational complexity to the LMS ST-DFE. Our simulation results have demonstrated that the LBER ST-DFE is robust and does not suffer from numerical ill-conditioning problems in low-noise fading environments. Finally the advocated design outperforms the LMS ST-DFE. Our future work includes incorporating the proposed MBER ST-DFE design with channel coding to develop an iterative joint detection and decoding scheme for SIMO systems.

References

[1] Foschini, G.J.: 'Layered space-time architecture for wireless communication in a fading environment when using multiple antennas', *Bell Labs Tech. J.*, 1996, 1, (2), pp.41–59

[2] Paulraj, A.J., and Papadias, C.B.: 'Space-time processing for wireless communications', *IEEE Signal Processing Magazine*, 1997, 14, (6), pp.49–83

[3] Winters, J.H.: 'Smart antennas for wireless systems', *IEEE Personal Communications*, 1998, 5, (1), pp.23–27

[4] Gesbert, D.: 'Robust linear MIMO receivers: A minimum error-rate approach', *IEEE Trans. Signal Processing*, 2003, 51, (11), pp.2863–2871

[5] Mujtaba, S.A.: 'MIMO signal processing – the next frontier for capacity enhancement'. *Proc. IEEE 2003 Custom Integrated Circuits Conf.*, Sept. 21-24, 2003, pp.263–270

[6] Paulraj, A., Nabar, R., and Gore, D.: 'Introduction to Space-Time Wireless Communications' (Cambridge University Press, 2003)

[7] Hanzo, L., Yang, L-L., Kuan, E-L., and Yen, K.: 'Single- and Multi-Carrier DS-CDMA: Multi-User Detection, Space-Time Spreading, Synchronisation, Standards and Networking' (IEEE Press - John Wiley, 2003)

[8] Paulraj, A.J., Gore, D.A., Nabar, R.U., and Bolcskei, H.: 'An overview of MIMO communications – A key to gigabit wireless', *Proc. IEEE*, 2004, 92, (2), pp.198–218

[9] Trautwein, U., Sommerkorn, G., and Thomä, R.S.: 'A simulation study on space-time equalization for mobile broadband communication in an industrial indoor environment'. *Proc. VTC 1999-Spring*, Houston, USA, May 16-20, 1999, pp.511–515

[10] Trautwein, U., Hampicke, D., Sommerkorn, G., and Thomä, R.S.: 'Performance of space-time processing for ISI- and CCI-suppression in industrial scenarios'. *Proc. VTC 2000-Spring*, Tokyo, Japan, May 15-18, 2000, pp.1894–1898

[11] Stoica, P., Vikalo, H., and Hassibi, B.: 'Joint maximum-likelihood channel estimation and signal detection for SIMO channels'. *Proc. ICASSP'03*, Hong Kong, China, April 6-10, 2003, pp.13–16

[12] Zhu, X., and Murch, R.D.: 'Layered space-time equalization for wireless MIMO systems', *IEEE Trans. Wireless Communications*, 2003, 2, (6), pp.1189–1203

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

[14] Chen, S., Mulgrew, B., Chng, E.S., and Gibson, G.: 'Space translation properties and the minimum-BER linear-combiner DFE'. *IEE Proc. Communications*, 1998, 145, (5), pp.316–322

[15] Chen, S., and Mulgrew, B.: 'The minimum-SER linear-combiner decision feedback equalizer', *IEE Proc. Communications*, 1999, 146, (6), pp.347–353

[16] Yeh, C-C., and Barry, J.R.: 'Adaptive minimum bit-error rate equalization for binary signaling', *IEEE Trans. Communications*, 2000, 48, (7), pp.1226–1235

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

[18] Mulgrew, B., and Chen, S.: 'Adaptive minimum-BER decision feedback equalisers for binary signalling', *Signal Processing*, 2001, 81, (7), pp.1479–1489

[19] Yeh, C-C., and Barry, J.R.: 'Adaptive minimum symbol-error rate equalization for quadrature-amplitude modulation', *IEEE Trans. Signal Processing*, 2003, 51, (12), pp.3263–3269

[20] Chen, S., Hanzo, L., and Mulgrew, B.: 'Adaptive minimum symbol-error-rate decision feedback equalization for multi-level pulse-amplitude modulation', *IEEE Trans. Signal Processing*, 2004, 52, (7), pp.2092–2101

[21] Chen, S., Hanzo, L., and Ahmad, N.N.: 'Adaptive minimum bit error rate beamforming assisted receiver for wireless communications'. *Proc. ICASSP-2003*, Hong Kong, China, April 6-10, 2003, pp.640–643

[22] Wolfgang, A., Ahmad, N.N., Chen, S., and Hanzo, L.: 'Genetic algorithm assisted minimum bit error rate beamforming'. *Proc. VTC2004-Spring*, Milan, Italy, May 17-19, 2004, pp.142-146

[23] Chen, S., Hanzo, L., Ahmad, N.N., and Wolfgang, A.: 'Adaptive minimum bit error rate beamforming assisted QPSK receiver'. *Proc. ICC2004*, Paris, France, June 20-24, 2004, pp.3389–3393

[24] Chen, S., Ahmad, N.N., and Hanzo, L.: 'Adaptive minimum bit error rate beamforming', *IEEE Trans. Wireless Communications*, 2005, 4, (2), pp.341–348

[25] Parzen, E.: 'On estimation of a probability density function and mode', *The Annals of Mathematical Statistics*, 1962, 33, pp.1066–1076

[26] Silverman, B.W.: '*Density Estimation*' (Chapman Hall, 1996)

[27] Bowman, A.W., and Azzalini, A.: '*Applied Smoothing Techniques for Data Analysis*' (Oxford University Press, 1997)

[28] Bazaraa, M.S., Sherali, H.D., and Shetty, C.M.: '*Nonlinear Programming: Theory and Algorithms*' (John Wiley, 1993)

[29] Chen, S., Samingan, A.K., Mulgrew, B., and Hanzo, L.: 'Adaptive minimum-BER linear multiuser detection for DS-CDMA signals in multipath channels', *IEEE Trans. Signal Processing*, 2001, 49, (6), pp.1240–1247

TABLE I

COMPARISON OF COMPUTATIONAL COMPLEXITY PER UPDATE, WHERE L IS THE NUMBER OF THE CHANNELS, n_c IS THE LENGTH OF EACH CHANNEL, THE ST-DFE FEEDFORWARD AND FEEDBACK ORDERS ARE CHOSEN AS $m = n_c$ AND $n_b = n_c - 1$, RESPECTIVELY.

ST-DFE	multiplications	additions	$\exp(\bullet)$ evaluation	square root evaluation
LBER	$24 \times (Ln_c) + 12$	$22 \times (Ln_c) - 4$	2	1
LMS	$24 \times (Ln_c) + 8$	$22 \times (Ln_c) - 2$	—	1

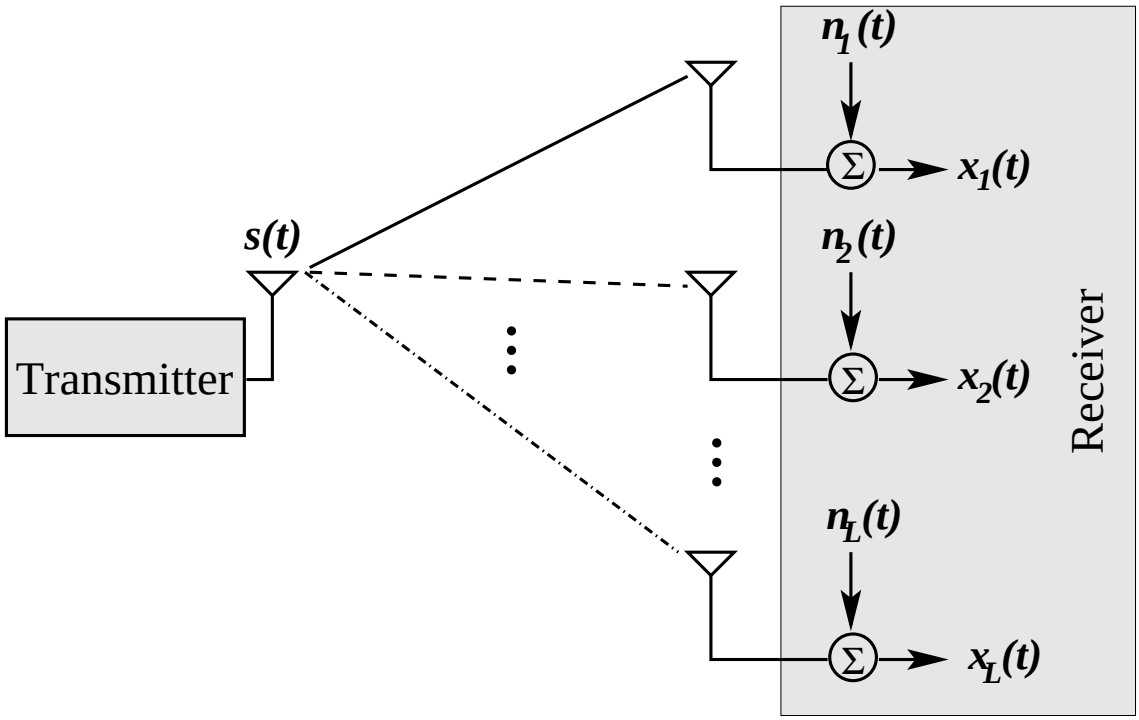

Fig. 1. Single-input multiple-output system employing multiple receive antennas.

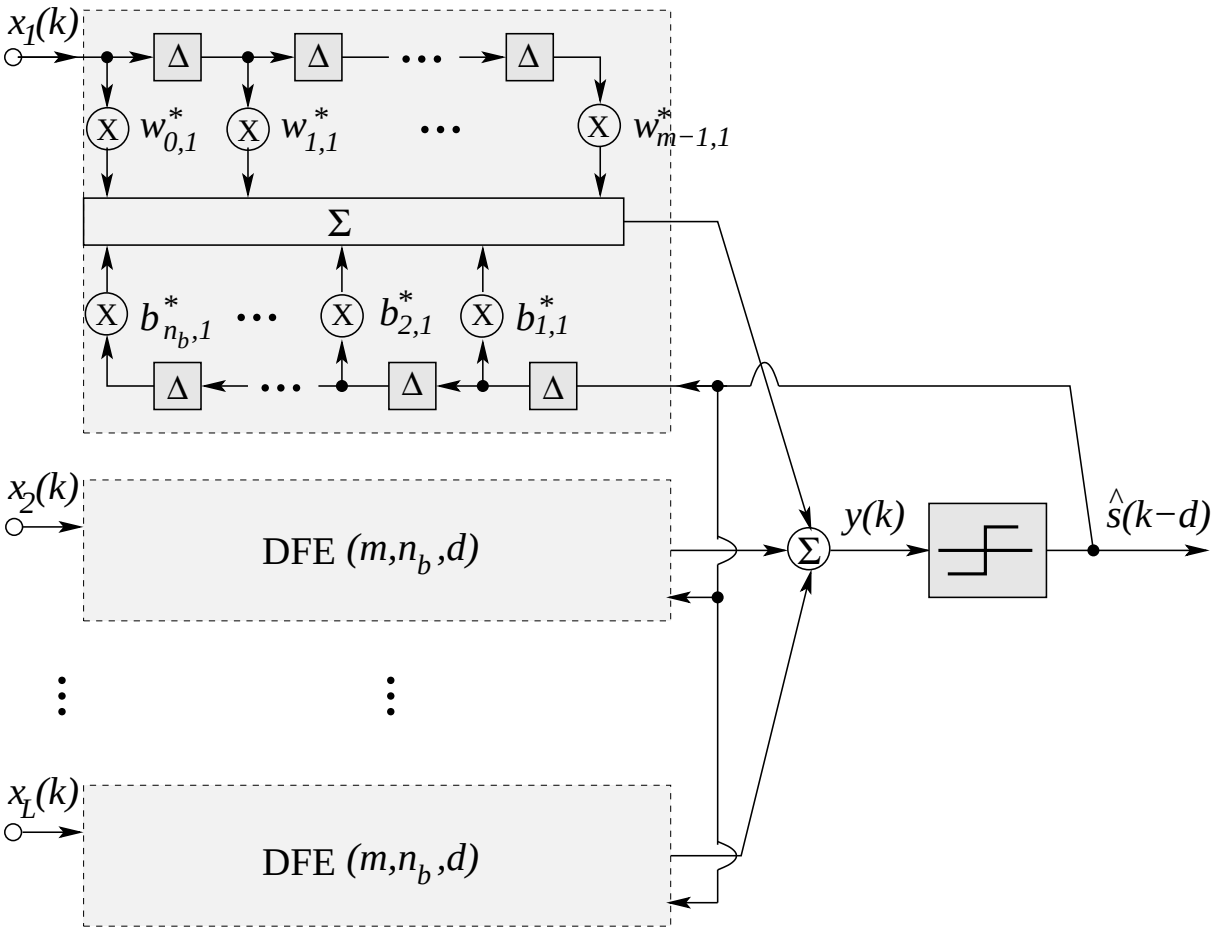

Fig. 2. Space-time decision feedback equalisation structure using Δ-spaced temporal filters, where $\Delta = T_s$ and T_s denotes the symbol period, m is the feedforward filter order, n_b the feedback filter order, and d the decision delay.

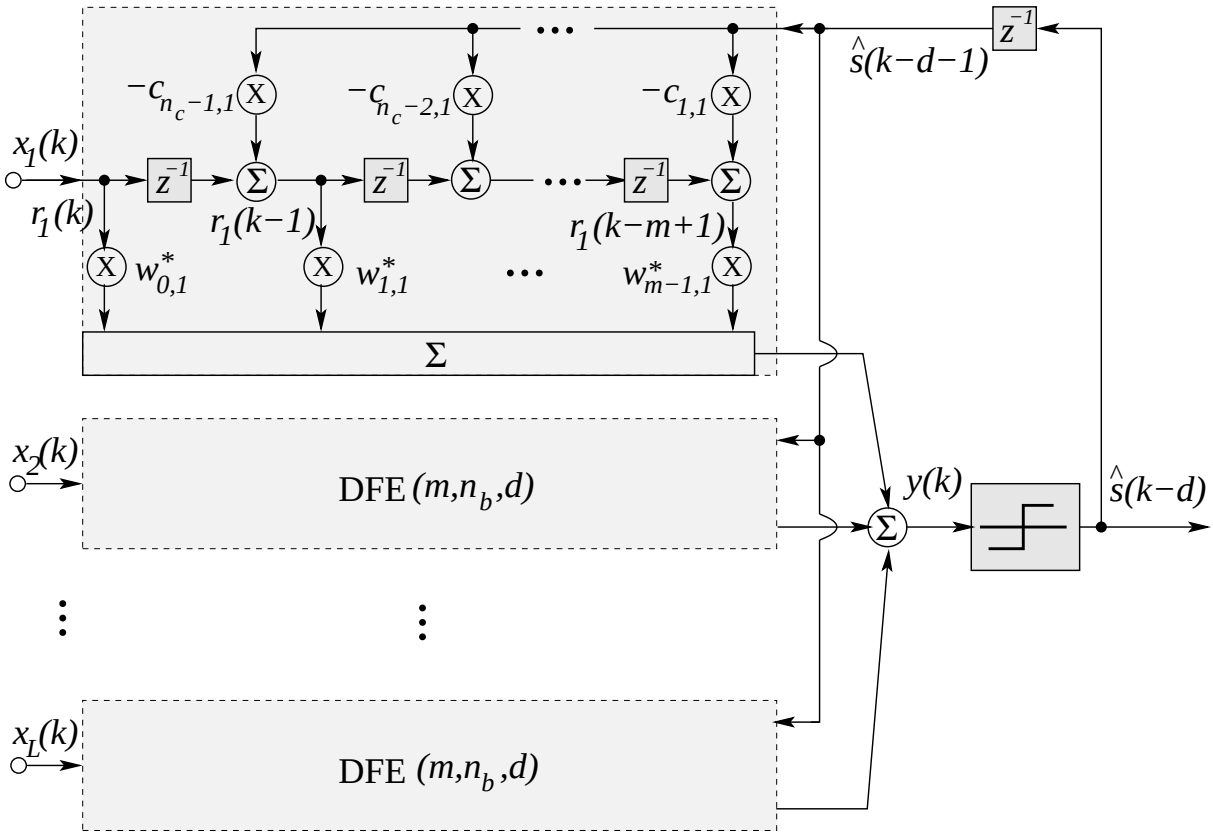

Fig. 3. Equivalent translated space-time decision feedback equaliser structure, where z^{-1} denotes the unit delay operator, n_c is the length of CIRs, and the DFE structure parameters are chosen as decision delay $d = n_c - 1$, feedforward order $m = n_c$ and feedback order $n_b = n_c - 1$.

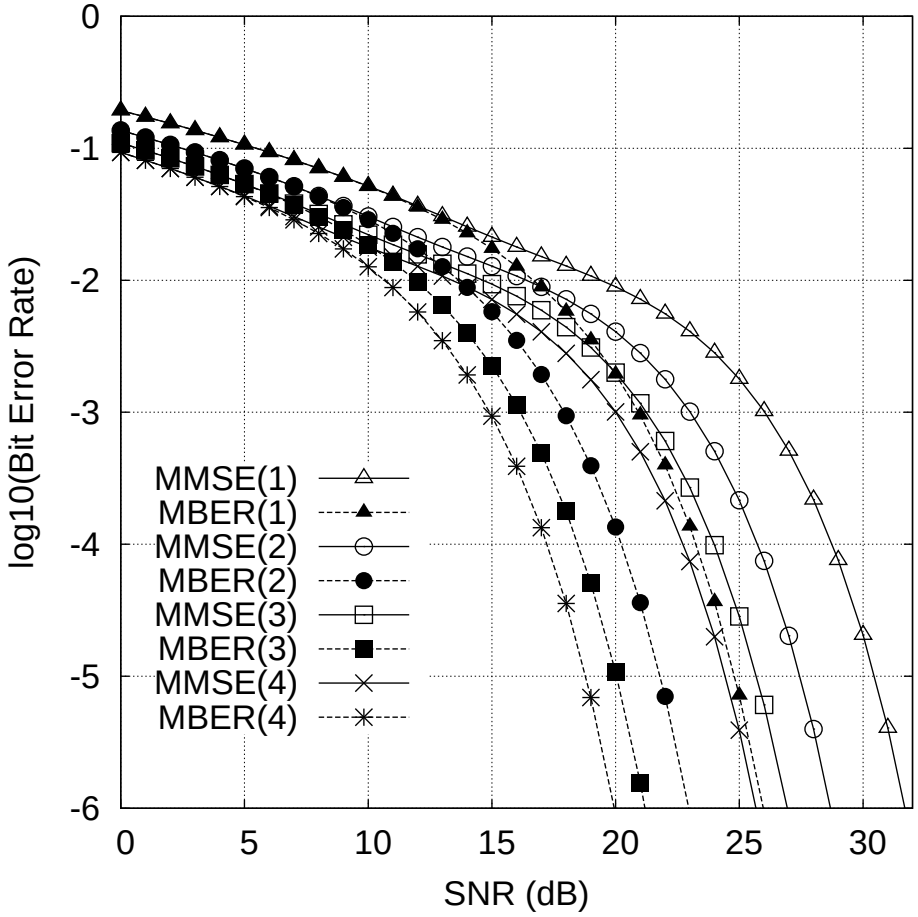

Fig. 4. Theoretical BER comparison of the MMSE and MBER ST-DFEs for the time-invariant system with varying number of antennas $L = 1$ to $L = 4$.

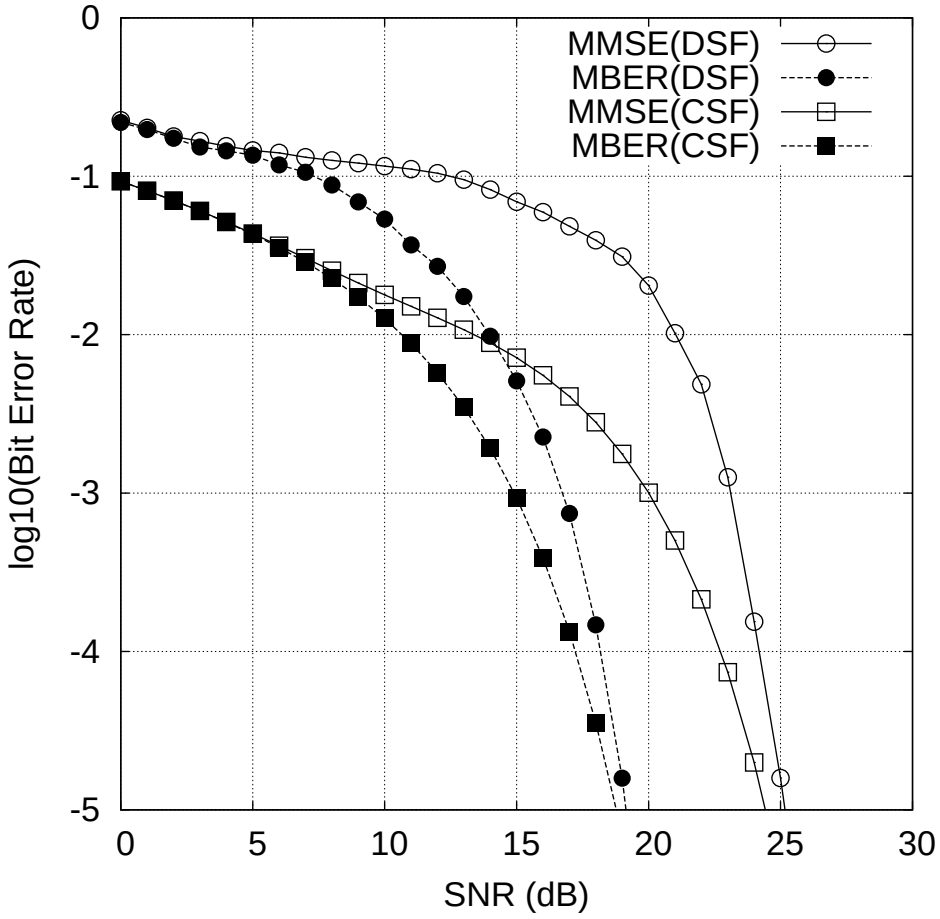

Fig. 5. Effects of error propagation on BER for the MMSE and MBER ST-DFE of the time-invariant system with number of antennas $L = 4$, where "DSF" denotes simulated BER with detected symbols being fed back, while "CSF" denotes theoretical BER with correct symbols being fed back.

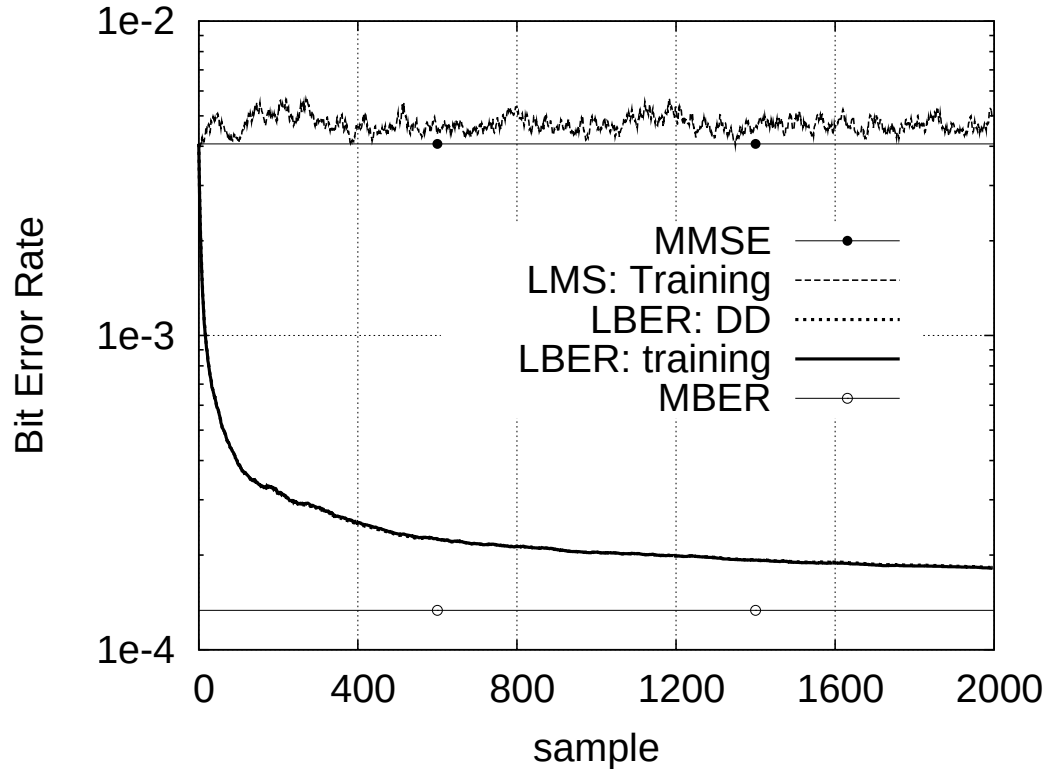

Fig. 6. Learning curves of the LBER algorithm in terms of the theoretical BER averaged over 20 runs for the time-invariant system with 4 receiver antennas and a SNR of 17 dB, given $\mathbf{w}(0) = \mathbf{w}_{\mathrm{MMSE}}$, $\mu = 0.05$ and $\rho_n^2 = 20\sigma_n^2 \approx 0.4$. DD denotes decision directed adaptation with $\hat{s}(k-d)$ substituting $s(k-d)$ and detected symbols being fed back. Note that for the LBER algorithm in this case the DD learning curve is indistinguishable from the training learning curve. For a comparison, the training learning curve for the LMS algorithm with $\mu = 0.01$ is also shown.

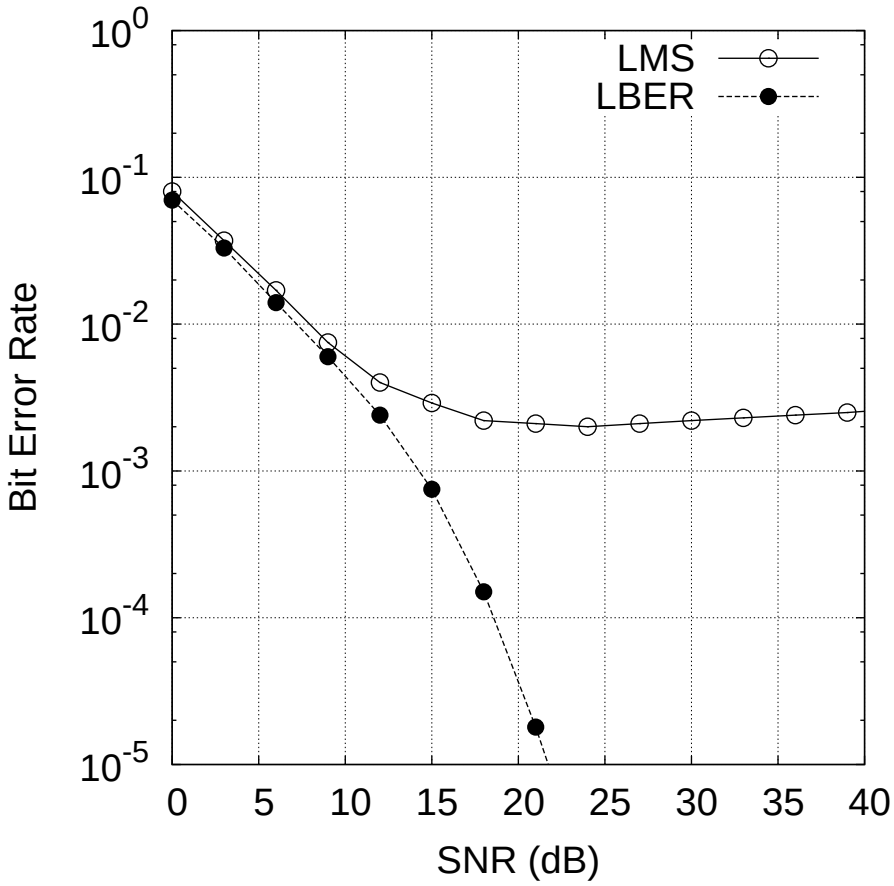

Fig. 7. BER comparison of the LBER and LMS ST-DFEs for the fading system with $L = 4$ receiver antennas. The BERs are obtained by simulation with actual detected symbols being fed back.