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) system. A novel minimum bit error ratio (MBER) design is invoked for the multiuser detector (MUD), which is shown to be capable of improving the attainable performance and enhancing system performance in comparison to that of the standard minimum mean square error (MMSE) design. The adaptive MUD coefficient adjustment procedure of the MBER space-time MUD is implemented using a stochastic gradient based least bit error rate (LBER) algorithm, which consistently outperforms the classic least mean square (LMS) algorithm, while maintaining a lower computational complexity than the latter.

II. SYSTEM MODEL

Consider a 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 BS’s receiver is assisted by an $L$-element antenna array. The symbol-rate received signal samples $x_l(k)$ for $1 \leq l \leq L$ are given by

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

where $n_l(k)$ is a complex-valued Gaussian white noise process associated with $E[|n_l(k)|^2] = 2\sigma_n^2$, $\tilde{x}_l(k)$ denotes the noise-free part of the $l$th receiver antenna’s output, $s_m(k)$ is the $k$th symbol of user $m$, while $c_{i,l,m}$ represents the CIR taps associated with the user $m$.

![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.](image-url)
Fig. 2 constitutes the MUD. The soft outputs of the antenna. The received signal vector is modeled by

\[ y_m(k) = \mathbf{w}_m^H(\mathbf{x}(k) + \mathbf{n}(k)) = \hat{y}_m(k) + e_m(k), \]

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 detector's weight vector \( \mathbf{w}_m \) is given by the following MMSE solution

\[ \mathbf{w}_{\text{MMSE}} = (\mathbf{C}\mathbf{C}^H + 2\sigma_n^2\mathbf{I})^{-1}\mathbf{C}(\mathbf{I} \cdots \mathbf{I})(\mathbf{I} \cdots \mathbf{I})(\mathbf{I} \cdots \mathbf{I}) \]

for \( 1 \leq m \leq M \), where \( \mathbf{I} \) denotes the \((L_F \times L_F)\)-dimensional identity matrix and \( \mathbf{C} \), the ith column of \( \mathbf{C} \). An adaptive MUD coefficient adjustment algorithms designed for the implementation of the MMSE solution can be realized using the LMS algorithm.

III. MINIMUM BIT ERROR RATE MULTIUSER DETECTION

Let us denote the \( N_s = 2^{M(n_F+n_C-1)} \) number of possible transmitted symbol sequences of \( s(k) \) as \( s^{(q)} \), \( 1 \leq q \leq N_s \). Denote furthermore the \((m-1)(n_F+n_C-1)\)th element of the symbol \( s_m(k-d) \), as \( s_m(k-d) \). The noise-free part of the mth detector input signal \( \mathbf{x}(k) \) assumes values from the set defined as: \( X_m \triangleq \{ \mathbf{x}(q) = \mathbf{C}s^{(q)} \mid 1 \leq q \leq N_s \} \). This set can be partitioned into two subsets, depending on the value of \( s_m(k-d) \), as follows: \( X_m^{(\pm)} \triangleq \{ \mathbf{x}(q) \in X_m : s_m(k-d) = \pm 1 \}. \) Similarly, the noise-free part of the mth detector's output \( \hat{y}_m(k) \) assumes values from the scalar set

\[ Y_m \triangleq \{ \hat{y}_m(q) = \mathbf{w}_m^H\mathbf{x}(q) \mid 1 \leq q \leq N_s \}. \]

Thus \( Y_m(k) = \mathbf{R}[\hat{y}_m(k)] \) can only take the values from the set

\[ Y_R \triangleq \{ \hat{y}_m(q) \in \mathbf{R}_m : 1 \leq q \leq N_s \}, \]

and \( Y_R \) can be divided into the two subsets conditioned on the value of \( s_m(k-d) \):

\[ Y_{R,m}^{(\pm)} \triangleq \{\hat{y}_m(q) \in Y_R : s_m(k-d) = \pm 1 \}. \]

It is readily seen that the conditional probability density function (PDF) of \( Y_{R,m}(k) \) given \( s_m(k-d) = +1 \) is

\[ p_m(y_m+1) = \frac{1}{N_s} \sum_{q=1}^{N_s} \frac{1}{\sqrt{2\pi\sigma_n^2}} e^{-\frac{(y_m-q^{(q)})^2}{2\sigma_n^2\mathbf{w}_m}}. \]

where \( \mathbf{w}_m^{(q)} \in Y_{R,m}^{(q)} \) and \( N_{sb} = N_s/2 \) is the number of the points in the set \( Y_{R,m}^{(q)} \). Thus the BER of the mth detector associated with the detector's weight vector \( \mathbf{w}_m \) is given by:

\[ P_E(\mathbf{w}_m) = \frac{1}{N_{sb}} \sum_{q=1}^{N_{sb}} Q\left( g^{(q)}(\mathbf{w}_m) \right), \]

where

\[ Q(u) = \frac{1}{\sqrt{2\pi}} \int_{u}^{\infty} e^{-\frac{v^2}{2}} dv \]

and

\[ g^{(q)}(\mathbf{w}_m) = \frac{\mathbf{sgn}(\mathbf{w}_m^T s_m(k-d) g^{(q)})}{\sigma_n \sqrt{\mathbf{w}_m^T \mathbf{w}_m}}. \]
Note that the BER is invariant to a positive scaling of $w_m$. Similarly, the BER may be calculated based on the other subset, namely on $Y_R(k)$:

The MBER solution for the $m$th detector is then defined as the weight vector that minimizes the error probability (13), namely

$$w_{(MBER) m} = \arg \min_{w_m} P_E(w_m).$$  \hspace{1cm} (16)

The gradient of $P_E(w_m)$ with respect to $w_m$ is given by:

$$\nabla P_E(w_m) = \frac{1}{2N_t \sqrt{2\pi}\sigma_n \sqrt{\sum_{q=1}^{N_q} w_m^H w_m}} \sum_{q=1}^{N_q} e^{\frac{(y_R - \tilde{y}_R^{(q)})^2}{2\sqrt{\sigma_n^2 w_m^H w_m}}} \
\times \text{sgn} \left( s_m^{(q)} \tilde{y}_R^{(q)} w_m - \tilde{x}^{(q)} \right).$$ \hspace{1cm} (17)

Given the gradient expression (17), the optimization problem (16) can be solved iteratively by commencing the iterations from an appropriate initialization point, such as the MMSE solution, using a gradient-based optimization algorithm. The simplified conjugate gradient algorithm [7] provides an efficient means of finding an MBER solution for the optimization problem formulated in (16).

### IV. ADAPTIVE MBER MULTIUSER DETECTION

The PDF of $y_{R_m}(k)$ can be shown to be explicitly given by:

$$p_m(y_R) = \frac{1}{N_t \sqrt{2\pi}\sigma_n \sqrt{\sum_{q=1}^{N_q} w_m^H w_m}} \sum_{q=1}^{N_q} e^{\frac{(y_R - \tilde{y}_R^{(q)})^2}{2\sqrt{\sigma_n^2 w_m^H w_m}}},$$ \hspace{1cm} (18)

and the BER of the $m$th detector can be calculated according to:

$$P_E(w_m) = \frac{1}{N_t} \sum_{q=1}^{N_q} Q \left( g^{(q)}(w_m) \right)$$ \hspace{1cm} (19)

with

$$g^{(q)}(w_m) = \frac{\text{sgn} \left( s_m^{(q)} \tilde{y}_R^{(q)} w_m \right)}{\sigma_n \sqrt{\sum_{q=1}^{N_q} w_m^H w_m}},$$ \hspace{1cm} (20)

where the summation is carried out over the $N_q$ number of elements $\tilde{y}_R^{(q)} \in Y_{R_m}$. In reality, the PDF of $y_{R_m}(k)$ is unknown. Therefore, some form of PDF estimation is required for supporting the adaptive implementation of the MBER MUD.

Given a block of $K$ training samples $\{x(k), s_m(k-d)\}_{k=1}^{K}$, a Parzen window estimate [8]-[10] of the PDF in (18) is given by:

$$\tilde{p}_m(y_R) = \frac{1}{K \sqrt{2\pi}\rho_n} \sum_{k=1}^{K} e^{\frac{(y_R - \tilde{s}_R^{(k)})^2}{2\sigma_n^2 \rho_n}},$$ \hspace{1cm} (21)

and the resultant approximate BER formula becomes

$$\tilde{P}_E(w_m) = \frac{1}{K} \sum_{k=1}^{K} Q \left( \tilde{g}_k(w_m) \right),$$ \hspace{1cm} (22)

where we have

$$\tilde{g}_k(w_m) = \frac{\text{sgn} \left( s_m(k-d) \tilde{s}_R^{(k)} \right)}{\rho_n}.$$ \hspace{1cm} (23)

In order to derive a sample-by-sample adaptive MUD coefficient adjustment algorithm for updating $w_m$, consider a single-sample estimate of $p_m(y_R)$, namely:

$$\hat{p}_m(y_R, k) = \frac{1}{\sqrt{2\pi}\rho_n} e^{\frac{(y_R - \tilde{s}_R^{(k)})^2}{2\rho_n^2 \sigma_n^2}}.$$ \hspace{1cm} (24)

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

$$\nabla \tilde{P}_E(w_m, k) = -\frac{\text{sgn} \left( s_m(k-d) \tilde{s}_R^{(k)} \right)}{2\sqrt{2\pi}\rho_n} \tilde{x}(k),$$ \hspace{1cm} (25)

gives rise to the LBER MUD coefficient adjustment algorithm of:

$$w_m(k+1) = w_m(k) + \mu \frac{\text{sgn} \left( s_m(k-d) \tilde{s}_R^{(k)} \right)}{2\sqrt{2\pi}\rho_n} \tilde{x}(k).$$ \hspace{1cm} (26)

The adaptive gain $\mu$ and kernel width $\rho_n$ have to be set appropriately to ensure adequate performance in terms of the achievable convergence rate and steady-state BER misadjustment. It can be shown that for the BPSK modulation scheme considered the LBER algorithm is slightly more complex than the LMS technique.

### V. SIMULATION STUDY

**Time-invariant system.** The system used in our simulations supported $M = 3$ users with the aid of $L = 4$ receiver antennas. All the three users had an equal power. The $3 \times 4 = 12$ CIRs are listed in Table I, each having $n_C = 2$ taps. In the actual simulation, all the 12 CIRs were normalized to provide unit channel energy. Each temporal filter had a length of $n_F = 3$ and the detector’s
decision delay was chosen to be \( d = 1 \). For the set of simulated CIRs Fig. 3 compares the BER performance of the MMSE and MBER MUDs. It can be seen that for all three users the MBER detectors had a better BER performance than the corresponding MMSE detectors. For the specific simulated channel conditions, the performance gap between the MBER and MMSE detectors was the smallest for user 3, exhibiting an approximately 1.5 dB SNR gain at the BER level of \( 10^{-4} \). At this BER level the MBER detector of user 2 had the SNR largest performance gain over the corresponding MMSE detector, which was in excess of 6.5 dB.

The MMSE and MBER solutions choose the detector’s weight vector very differently. Fig. 4 plots the full conditional PDFs \( p_m(y_j|+1) \), marginal conditional PDFs \( p_m(y_{jk}|+1) \), and the corresponding signal subsets \( Y_m^{(+)} \) and \( Y_{R_m}^{(+)} \) for user 2 at SNR=2 dB. Since the MMSE detector minimises the MSE term of \( E[|s_m(k-d) - y_m(k)|^2] \), the signal subset \( Y_m^{(+)} \) has to be distributed symmetrically with respect to both \( \Re[y] \) and \( \Im[y] \), and the associated conditional PDF \( p_{m}(y) \) has to be circular. However, the BER of the detector is mainly determined by the minimum distance of the subset \( Y_{R_m}^{(+)} \) from the decision threshold \( \Re[y] = 0 \). By contrast, the MBER detector is not constrained by the symmetric and circular considerations, it spreads \( Y_m^{(+)} \) significantly wider along \( \Im[y] \) and this leads to a significantly higher MSE value. However, this also doubles the minimum distance between \( Y_{R_m}^{(+)} \) and \( \Re[y] = 0 \), resulting in a substantially lower BER, compared to the MMSE solution. Clearly, the MBER design is more intelligent in utilising the detector’s resources.

The LMS and LBER MUDs were next investigated for user 2 at SNR=2 dB. Fig. 5 depicts the learning curves of the two adaptive detectors, averaged over 20 runs and started from two different initial conditions \( w_m(0) \). In these investigations the adaptive MUD operated in two different modes, namely using pilot-based training, when the transmitted symbols \( s_m(k-d) \) were known to the receiver and decision-directed (DD) adaptation, when the detected symbols \( s_m(k-d) \) were used to substitute \( s_m(k-d) \). It can be seen that the LBER MUD consistently outperformed the LMS MUD. Furthermore, the LBER MUD was capable of taking the full advantage of DD adaptation, as demonstrated in Fig. 5 (b), where it can also be seen that the DD LMS MUD failed to converge to the MMSE solution.

**Fading system.** The system setup and the structure of the MUDs were the same as in the time-invariant CIR-based example, but fading channels were simulated. The transmission frame structure consisted of 20 training symbols followed by 200 data symbols. The magnitudes of the CIR taps were i.i.d. Rayleigh processes, each having the root mean power of \( \sqrt{0.5} + j\sqrt{0.5} \). Idealistic fading conditions were simulated, where the CIR taps were subjected to frame-invariant fading. Hence the CIR taps were faded at the beginning of each transmission frame at a normalized Doppler frequency of \( 10^{-2} \), generating different channel magnitudes and phases for different frames, but they were kept constant within the transmission frame. Fig. 6 compares the bit error rate achieved by the LMS and LBER MUDs of user 1. The performance of the MUDs for the other two users were similar to those shown in Fig. 6.

**VI. CONCLUSIONS**

A novel minimum bit error rate design has been proposed for a space-time equalisation 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 a better BER performance and hence of enhancing the achievable system capacity, compared to the standard minimum mean square error design. An adaptive implementation of the MBER space-time MUD has also been derived based on the least bit error rate algorithm, which was shown to consistently outperform the classic least mean square algorithm, while maintaining a lower computational complexity in comparison to the latter.
TABLE I
CIRs for the 3-user 4-antenna time-invariant system. The actually simulated CIRs were $C_{l,m}(z)/|C_{l,m}(z)|$ to provide unit channel energy.

<table>
<thead>
<tr>
<th>$C_{l,m}(z)$</th>
<th>$m = 1$</th>
<th>$m = 2$</th>
<th>$m = 3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$l = 1$</td>
<td>$(0.5 + j0.6) + (0.8 - j0.7)z^{-1}$</td>
<td>$(1.0 + j0.8) + (0.7 + j0.4)z^{-1}$</td>
<td>$(0.6 - j0.8) + (0.5 + j0.3)z^{-1}$</td>
</tr>
<tr>
<td>$l = 2$</td>
<td>$(0.7 + j1.0) + (0.2 + j0.5)z^{-1}$</td>
<td>$(0.4 + j0.4) + (0.6 + j0.6)z^{-1}$</td>
<td>$(-0.5 + j0.3) + (0.9 + j0.2)z^{-1}$</td>
</tr>
<tr>
<td>$l = 3$</td>
<td>$(-0.5 + j0.7) + (-0.6 + j0.7)z^{-1}$</td>
<td>$(-0.3 - j0.6) + (0.5 - j0.7)z^{-1}$</td>
<td>$(0.7 + j0.8) + (-0.2 - j0.3)z^{-1}$</td>
</tr>
<tr>
<td>$l = 4$</td>
<td>$(0.3 - j0.6) + (0.5 - j0.2)z^{-1}$</td>
<td>$(0.8 - j0.5) + (0.6 + j0.2)z^{-1}$</td>
<td>$(0.2 + j0.1) + (0.9 + j0.2)z^{-1}$</td>
</tr>
</tbody>
</table>

Fig. 5. Learning curves of the adaptive detector of user-2 for the 3-user 4-antenna time-invariant system at SNR$^\text{(m)}$.

Fig. 6. Bit error rate comparison of the user-one LMS and LBER multiuser detectors for the 3-user 4-antenna fading system.

REFERENCES