Adaptive Minimum Symbol Error Rate Beamforming Assisted Receiver for Quadrature Amplitude Modulated Systems

S. Chen, H.-Q. Du and L. Hanzo

School of ECS, University of Southampton, SO17 1BJ, U.K.
E-mails: {sqc,lh,hd404}@ecs.soton.ac.uk

Abstract **An adaptive beamforming assisted receiver is proposed for multiple antenna aided multiuser systems that employ bandwidth efficient quadrature amplitude modulation (QAM). A novel minimum symbol error rate (MSER) design is proposed for the beamforming assisted receiver, where the system's symbol error rate is directly optimized. Hence the MSER approach provides a significant symbol error ratio performance enhancement over the classic minimum mean square error design. A sample-by-sample adaptive algorithm, referred to as the least symbol error rate (LBER) technique, is derived for allowing the adaptive implementation of the system to arrive from its initial beamforming weight solution to MSER beamforming solution.**

I. Introduction

The ever-increasing demand for mobile communication capacity has motivated the development of adaptive antenna array assisted spatial processing techniques [1]–[10] in order to further improve the achievable spectral efficiency of wireless systems. A particular technique that has shown real promise in achieving substantial capacity enhancements is the employment of adaptive beamforming with the aid of antenna arrays, which create angularly selective beam maxima towards the desired user and a null towards a limited number of dominant interferers. By appropriately combining the signals received by the different elements of an antenna array to form a single output, adaptive beamforming becomes capable of separating signals transmitted on the same carrier frequency, provided that they arrive from sufficiently different angular directions. Thus bemaforming becomes capable of supporting multiple users in an 'angular division multiple access' scenario. Classically, the beamforming process is carried out by adjusting the beamforming array weights upon minimizing the mean square error (MSE) between the desired output of the array, which is typically the most likely legitimate transmitted symbol and the actual array output. In other words, we set the partial derivative of the array output to zero with respect to the array weights.

However, in most communications systems it is the bit error ratio (BER) or symbol error ratio (SER) that really matters. Hence adaptive beamforming based on directly minimizing the system's BER has been proposed for both binary phase shift keying and quadrature phase shift keying modulation schemes in [11],[12].

In recent years the family of high-throughput quadrature amplitude modulation (QAM) schemes [13] has become predominant in numerous wireless local area network (WLAN) standards, such as the IEEE 802.11 standards. Adaptive minimum SER (MSER) equalization has been investigated in the context of a single-antenna single-user system, when using either a pulse-amplitude modulation scheme [14] or a QAM scheme [15]. *Against this backcloth, in this paper, we derive the MSER beamforming design for a multiple antenna assisted multiuser system employing QAM signalling. We show that the MSER design is capable of providing significant SER performance gains over the traditional minimum MSE (MMSE) design. An attractive adaptive implementation of the MSER beamforming solution is also proposed, which step-by-step adjusts the array weights, commencing from an adequate initial solution using the classic stochastic gradient algorithm (SGA), which we refer to as the least symbol error rate (LSER) technique.* Our proposed solution is substantially different from the method proposed in [15], since the adaptive LSER algorithm invoked has its roots in the so-called Parzen window based density estimation [16]-[18]. In this sense, the proposed adaptive MSER technique is an extension of the method proposed in [14] for an interference-limited multiuser system to a more bandwidth-efficient QAM scheme.

II. System Model

The system supports S users, and each user transmits an M-QAM signal on the same carrier frequency $\omega = 2\pi f$. The receiver is equipped with a linear antenna array consisting of L uniformly spaced elements. We assume that the channel does not induce intersymbol interference (ISI). Then the symbol-rate received signal samples can be expressed as

$$x_l(k) = \sum_{i=1}^{S} A_i b_i(k) e^{j\omega t_l(\theta_i)} + n_l(k) = \bar{x}_l(k) + n_l(k), \quad (1)$$

for $1 \leq l \leq L$, where $t_l(\theta_i)$ is the relative time delay at element l for source i with θ_i being the direction of arrival

The financial support of the EPSRC, Swindon UK and that of the EU under the auspices of the PHOENIX as well as NEWCOM projects is gratefully acknowledged.

0-7803-9392-9/06/$20.00 (c) 2006 IEEE

for source i, $n_l(k)$ is the complex-valued Additive White Gaussian Noise (AWGN) with $E[|n_l(k)|^2] = 2\sigma_n^2$, A_i is the complex-valued non-dispersive channel coefficient of user i, while $b_i(k)$ is the kth M-QAM symbol of user i

$$\mathcal{B} \triangleq \{b_{l,q} = u_l + ju_q,\ 1 \leq l, q \leq \sqrt{M}\} \quad (2)$$

with $u_l = 2l - \sqrt{M} - 1$ and $u_q = 2q - \sqrt{M} - 1$. Source 1 is the desired user and the rest of the sources are interfering users.The desired-user's signal to noise ratio is SNR= $|A_1|^2\sigma_b^2/2\sigma_n^2$ and the desired signal to interferer i ratio is SIR$_i$ $= A_1^2/A_i^2$, for $2 \leq i \leq S$, where σ_b^2 denotes the M-QAM symbol's energy. The received signal vector $\mathbf{x}(k) = [x_1(k)\ x_2(k) \cdots x_L(k)]^T$ is given by

$$\mathbf{x}(k) = \mathbf{P}\mathbf{b}(k) + \mathbf{n}(k) = \bar{\mathbf{x}}(k) + \mathbf{n}(k), \quad (3)$$

where $\mathbf{n}(k) = [n_1(k)\ n_2(k) \cdots n_L(k)]^T$, the system matrix $\mathbf{P} = [A_1\mathbf{s}_1\ A_2\mathbf{s}_2 \cdots A_S\mathbf{s}_S]$ with the steering vector of source i given by $\mathbf{s}_i = [e^{j\omega t_1(\theta_i)}\ e^{j\omega t_2(\theta_i)} \cdots e^{j\omega t_L(\theta_i)}]^T$ and that of the transmitted QAM symbol vector by $\mathbf{b}(k)$ $= [b_1(k)\ b_2(k) \cdots b_S(k)]^T$.

A linear beamformer's soft output is given by

$$y(k) = \mathbf{w}^H\mathbf{x}(k) = \mathbf{w}^H(\bar{\mathbf{x}}(k) + \mathbf{n}(k)) = \bar{y}(k) + e(k) \quad (4)$$

where $\mathbf{w} = [w_1\ w_2 \cdots w_L]^T$ is the beamformer's weight vector and $e(k)$ is Gaussian distributed with zero mean and $E[|e(k)|^2] = 2\sigma_n^2\mathbf{w}^H\mathbf{w}$. We define the combined impulse response of the beamformer and the channel as $\mathbf{w}^H\mathbf{P} = \mathbf{w}^H[\mathbf{p}_1\ \mathbf{p}_2 \cdots \mathbf{p}_S] = [c_1\ c_2 \cdots c_S]$. The beamformer's output can alternatively be expressed as

$$y(k) = c_1b_1(k) + \sum_{k=2}^{S} c_ib_i(k) + e(k). \quad (5)$$

Provided that $c_1 = c_{R_1} + jc_{I_1}$ satisfies $c_{R_1} > 0$ and $c_{I_1} = 0$, the symbol decision $\hat{b}_1(k) = \hat{b}_{R_1}(k) + j\hat{b}_{I_1}(k)$ can be decoupled into

$$\hat{b}_{R_1}(k) = \begin{cases} u_1, & \text{if } y_R(k) \leq c_{R_1}(u_1 + 1) \\ u_l, & \text{if } c_{R_1}(u_l - 1) < y_R(k) \leq c_{R_1}(u_l + 1) \\ & \text{for } 2 \leq l \leq \sqrt{M} - 1 \\ u_{\sqrt{M}}, & \text{if } y_R(k) > c_{R_1}(u_{\sqrt{M}} - 1) \end{cases} \quad (6)$$

$$\hat{b}_{I_1}(k) = \begin{cases} u_1, & \text{if } y_I(k) \leq c_{R_1}(u_1 + 1) \\ u_q, & \text{if } c_{R_1}(u_q - 1) < y_I(k) \leq c_{R_1}(u_q + 1) \\ & \text{for } 2 \leq q \leq \sqrt{M} - 1 \\ u_{\sqrt{M}}, & \text{if } y_I(k) > c_{R_1}(u_{\sqrt{M}} - 1) \end{cases} \quad (7)$$

where $y(k) = y_R(k) + jy_I(k)$ and $\hat{b}_1(k)$ is the estimate of $b_1(k) = b_{R_1}(k) + jb_{I_1}(k)$. Fig. 1 depicts the decision thresholds associated with the decision $\hat{b}_1(k) = b_{l,q}$. In general, $c_1 = \mathbf{w}^H\mathbf{p}_1$ is complex-valued and the rotating operation

$$\mathbf{w}^{\text{new}} = \frac{c_d^{\text{old}}}{|c_d^{\text{old}}|}\mathbf{w}^{\text{old}} \quad (8)$$

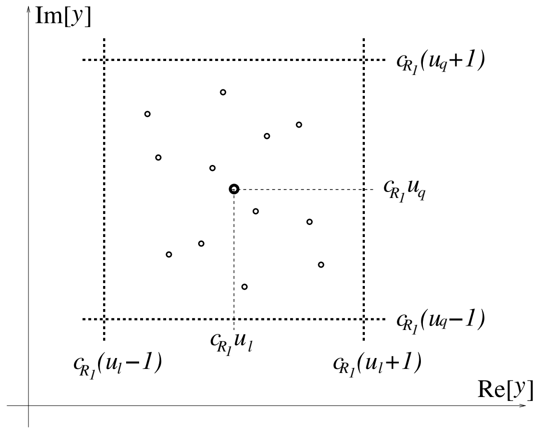


Fig. 1. Decision thresholds associated with point $c_1b_{l,q}$ assuming $c_{R_1} > 0$ and $c_{I_1} = 0$, illustrating the symmetry of the distribution of $\mathcal{Y}_{l,q}$ around $c_1b_{l,q}$.

can be used to render c_1 real-valued and positive. This rotation is a linear transformation and does not alter the system's SER. Thus the desired user's channel A_1 and steering vector $\mathbf{s}_1$ are required at the receiver in order to apply the decision rules of (6) and (7).

III. Minimum Symbol Error Rate Beamforming

The classic MMSE solution for the beamformer of (4) is given by

$$\mathbf{w}_{\text{MMSE}} = \left(\mathbf{P}\mathbf{P}^H + \frac{2\sigma_n^2}{\sigma_b^2}\mathbf{I}_L\right)^{-1}\mathbf{p}_1, \quad (9)$$

Since the SER is the ultimate performance indicator, it is desirable to find the optimal MSER beamforming weight solution. We denote the $N_b = M^S$ number of possible sequences of $\mathbf{b}(k)$ as $\mathbf{b}_i$, $1 \leq i \leq N_b$. Then $\bar{\mathbf{x}}(k)$ can only assume values from the finite signal set defined by $\mathcal{X} \triangleq \{\bar{\mathbf{x}}_i = \mathbf{P}\mathbf{b}_i,\ 1 \leq i \leq N_b\}$. The set $\mathcal{X}$ can be partitioned into M subsets, depending on the value of $b_1(k)$

$$\mathcal{X}_{l,q} \triangleq \{\bar{\mathbf{x}}_i \in \mathcal{X}:\ b_1(k) = b_{l,q}\},\ 1 \leq l, q \leq \sqrt{M}. \quad (10)$$

The noise-free component of the beamformer's output $\bar{y}(k)$ only assumes values from the scalar set $\mathcal{Y} \triangleq \{\bar{y}_i = \mathbf{w}^H\bar{\mathbf{x}}_i,\ 1 \leq i \leq N_b\}$, and $\mathcal{Y}$ can be divided into M subsets conditioned on the value of $b_1(k)$

$$\mathcal{Y}_{l,q} \triangleq \{\bar{y}_i \in \mathcal{Y}:\ b_1(k) = b_{l,q}\},\ 1 \leq l, q \leq \sqrt{M}. \quad (11)$$

Lemma 1: The subsets $\mathcal{Y}_{l,q}$, $1 \leq l, q \leq \sqrt{M}$, satisfy the shifting properties

$$\mathcal{Y}_{l+1,q} = \mathcal{Y}_{l,q} + 2c_1,\ 1 \leq l \leq \sqrt{M} - 1, \quad (12)$$

$$\mathcal{Y}_{l,q+1} = \mathcal{Y}_{l,q} + j2c_1,\ 1 \leq q \leq \sqrt{M} - 1, \quad (13)$$

$$\mathcal{Y}_{l+1,q+1} = \mathcal{Y}_{l,q} + (2+j2)c_1, \ 1 \le l, q \le \sqrt{M} - 1. \quad (14)$$

The proof of Lemma 1 is straightforward.

Lemma 2: The points of $\mathcal{Y}_{l,q}$ are distributed symmetrically around the symbol point $c_1 b_{l,q}$.

Lemma 2 is a direct consequence of the symmetry of the symbol constellation (2). This symmetric property is also illustrated in Fig. 1. Note that the distribution of $\mathcal{Y}_{l,q}$ is symmetric with respect to the two vertical decision thresholds $c_{R_1}(u_l \pm 1)$ and with respect to the two horizontal decision threshold $c_{R_1}(u_q \pm 1)$.

For the beamformer having a weight vector $\mathbf{w}$ we introduce the notation

$$P_E(\mathbf{w}) = \text{Prob}\{\hat{b}_1(k) \neq b_1(k)\}, \quad (15)$$

$$P_{E_R}(\mathbf{w}) = \text{Prob}\{\hat{b}_{R_1}(k) \neq b_{R_1}(k)\}, \quad (16)$$

$$P_{E_I}(\mathbf{w}) = \text{Prob}\{\hat{b}_{I_1}(k) \neq b_{I_1}(k)\}. \quad (17)$$

It is then readily seen that the SER is given by

$$P_E(\mathbf{w}) = P_{E_R}(\mathbf{w}) + P_{E_I}(\mathbf{w}) - P_{E_R}(\mathbf{w})P_{E_I}(\mathbf{w}). \quad (18)$$

The conditional probability density function (PDF) of $y(k)$ given $b_1(k) = b_{l,q}$ is a Gaussian mixture defined by

$$p(y|b_{l,q}) = \frac{1}{N_{sb} 2\pi \sigma_n^2 \mathbf{w}^H \mathbf{w}} \sum_{i=1}^{N_{sb}} e^{-\frac{|y - \bar{y}_i^{(l,q)}|^2}{2\sigma_n^2 \mathbf{w}^H \mathbf{w}}}, \quad (19)$$

where $N_{sb} = N_b/M$ is the size of $\mathcal{Y}_{l,q}$, $\bar{y}_i^{(l,q)} = \bar{y}_{R_i}^{(l,q)} + j\bar{y}_{I_i}^{(l,q)} \in \mathcal{Y}_{l,q}$, and $y = y_R + jy_I$. Noting that c_1 is real-valued and positive, as well as taking into account the symmety of the distribution of $\mathcal{Y}_{l,q}$ (lemma 2), for $2 \le l \le \sqrt{M} - 1$, the conditional error probability of $\hat{b}_{R_1}(k) \neq u_l$ given $b_{R_1}(k) = u_l$ can be shown to be

$$P_{E_R,l}(\mathbf{w}) = \frac{2}{N_{sb}} \sum_{i=1}^{N_{sb}} Q(g_{R_i}^{(l,q)}(\mathbf{w})), \quad (20)$$

where

$$Q(u) = \frac{1}{\sqrt{2\pi}} \int_u^\infty e^{-\frac{z^2}{2}} dz, \quad (21)$$

$$g_{R_i}^{(l,q)}(\mathbf{w}) = \frac{\bar{y}_{R_i}^{(l,q)} - c_{R_1}(u_l - 1)}{\sigma_n \sqrt{\mathbf{w}^H \mathbf{w}}}. \quad (22)$$

Furthermore, taking into account the shifting property (lemma 1), it is straightforward to show that we have

$$P_{E_R}(\mathbf{w}) = \gamma \frac{1}{N_{sb}} \sum_{i=1}^{N_{sb}} Q(g_{R_i}^{(l,q)}(\mathbf{w})), \quad (23)$$

where $\gamma = \frac{2\sqrt{M}-2}{\sqrt{M}}$. It is seen that P_{E_R} can be evaluated using (the real part of) any single subset $\mathcal{Y}_{l,q}$. Similarly, P_{E_I} can be evaluated using (the imaginary part of) any single subset $\mathcal{Y}_{l,q}$ as

$$P_{E_I}(\mathbf{w}) = \gamma \frac{1}{N_{sb}} \sum_{i=1}^{N_{sb}} Q(g_{I_i}^{(l,q)}(\mathbf{w})) \quad (24)$$

with

$$g_{I_i}^{(l,q)}(\mathbf{w}) = \frac{\bar{y}_{I_i}^{(l,q)} - c_{R_1}(u_q - 1)}{\sigma_n \sqrt{\mathbf{w}^H \mathbf{w}}}. \quad (25)$$

Note that the SER is invariant to a positive scaling of $\mathbf{w}$.

The MSER solution $\mathbf{w}_{\text{MSER}}$ is defined as the one that minimizes the upper bound of the SER given by

$$P_{E_B}(\mathbf{w}) = P_{E_R}(\mathbf{w}) + P_{E_I}(\mathbf{w}), \quad (26)$$

that is,

$$\mathbf{w}_{\text{MSER}} = \arg\min_{\mathbf{w}} P_{E_B}(\mathbf{w}). \quad (27)$$

The upper bound $P_{E_B}(\mathbf{w})$ is very tight, i.e. very close to the true SER $P_E(\mathbf{w})$. The gradients of $P_{E_R}(\mathbf{w})$ and $P_{E_I}(\mathbf{w})$ with respect to $\mathbf{w}$ can be shown to be

$$\nabla P_{E_R}(\mathbf{w}) = \frac{\gamma}{2N_{sb}\sqrt{2\pi}\sigma_n\sqrt{\mathbf{w}^H\mathbf{w}}} \sum_{i=1}^{N_{sb}} e^{-\frac{\left(\bar{y}_{R_i}^{(l,q)} - c_{R_1}(u_l-1)\right)^2}{2\sigma_n^2 \mathbf{w}^H \mathbf{w}}} \times \left(\frac{\bar{y}_{R_i}^{(l,q)} - c_{R_1}(u_l - 1)}{\mathbf{w}^H \mathbf{w}} \mathbf{w} - \bar{\mathbf{x}}_i^{(l,q)} + (u_l - 1)\mathbf{p}_1 \right), \quad (28)$$

$$\nabla P_{E_I}(\mathbf{w}) = \frac{\gamma}{2N_{sb}\sqrt{2\pi}\sigma_n\sqrt{\mathbf{w}^H\mathbf{w}}} \sum_{i=1}^{N_{sb}} e^{-\frac{\left(\bar{y}_{I_i}^{(l,q)} - c_{R_1}(u_q-1)\right)^2}{2\sigma_n^2 \mathbf{w}^H \mathbf{w}}} \times \left(\frac{\bar{y}_{I_i}^{(l,q)} - c_{R_1}(u_q - 1)}{\mathbf{w}^H \mathbf{w}} \mathbf{w} + j\bar{\mathbf{x}}_i^{(l,q)} + (u_q - 1)\mathbf{p}_1 \right), \quad (29)$$

where $\bar{\mathbf{x}}_i^{(l,q)} \in \mathcal{X}_{l,q}$. With the gradient $\nabla P_{E_B}(\mathbf{w}) = \nabla P_{E_R}(\mathbf{w}) + \nabla P_{E_I}(\mathbf{w})$, the optimization problem (27) can be solved iteratively using a gradient optimization algorithm, such as the simplified conjugate gradient algorithm [11]. The rotating operation (8) should be applied after each iteration, to ensure that we have a real and positive c_1 value.

The PDF $p(y)$ of $y(k)$ can be estimated using the Parzen window estimate based on a block of training data. This leads to an estimated SER for the beamformer. Minimizing this estimated SER based on a gradient optimization yields an approximated MSER solution. To derive a sample-by-sample adaptive algorithm, consider a single-sample "estimate" of $p(y)$

$$\tilde{p}(y,k) = \frac{1}{2\pi\rho_n^2} e^{-\frac{|y - y(k)|^2}{2\rho_n^2}} \quad (30)$$

and the corresponding one-sample SER "estimate" $\tilde{P}_{E_B}(\mathbf{w}, k)$. Using the instantaneous stochastic gradient of $\nabla \tilde{P}_{E_B}(\mathbf{w}, k) = \nabla \tilde{P}_{E_R}(\mathbf{w}, k) + \nabla \tilde{P}_{E_I}(\mathbf{w}, k)$ with

$$\nabla \tilde{P}_{E_R}(\mathbf{w}, k) = \frac{\gamma}{2\sqrt{2\pi}\rho_n} e^{-\frac{\left(y_R(k) - \hat{c}_{R_1}(k)(b_{R_1}(k)-1)\right)^2}{2\rho_n^2}} \times \left(-\mathbf{x}(k) + (b_{R_1}(k) - 1)\hat{\mathbf{p}}_1\right) \quad (31)$$

and

$$\nabla \tilde{P}_{E_I}(\mathbf{w}, k) = \frac{\gamma}{2\sqrt{2\pi}\rho_n} e^{-\frac{\left(y_I(k) - \hat{c}_{R_1}(k)(b_{I_1}(k)-1)\right)^2}{2\rho_n^2}} \times \left(j\mathbf{x}(k) + (b_{I_1}(k) - 1)\hat{\mathbf{p}}_1\right) \quad (32)$$

gives rise to the following stochastic gradient adaptive algorithm, which we refer to as the LSER algorithm

$$\mathbf{w}(k+1) = \mathbf{w}(k) + \mu\left(-\nabla \tilde{P}_{E_B}(\mathbf{w}(k), k)\right), \quad (33)$$

$$\hat{c}_1(k+1) = \mathbf{w}^H(k+1)\hat{\mathbf{p}}_1, \quad (34)$$

$$\mathbf{w}(k+1) = \frac{\hat{c}_1(k+1)}{|\hat{c}_1(k+1)|}\mathbf{w}(k+1), \quad (35)$$

where $\hat{\mathbf{p}}_1$ is an estimate of $\mathbf{p}_1$. The step size μ and the kernel width ρ_n are the two algorithmic parameters that should be set appropriately in order to attain an adequate performance in terms of both the convergence rate and steady-state SER misadjustment.

IV. Simulation Study

Stationary system. Our prototype system supported four users with the aid of a three-element antenna array. Fig. 2 shows the locations of both the desired source and the interfering sources. The channel coefficients were $A_i = 1 + j0$, $1 \le i \le 4$. Thus we had $\text{SIR}_i = 0$ dB for $2 \le i \le 4$. The modulation scheme was 16-QAM. Fig. 3 compares the SER performance of the MSER solution to that of the MMSE solution under three different conditions: **(a)** the minimum spatial separation between the desired user 1 and the interfering user 4 was $\theta = 32°$ **(b)** $\theta = 30°$, and **(c)** $\theta = 28°$. For

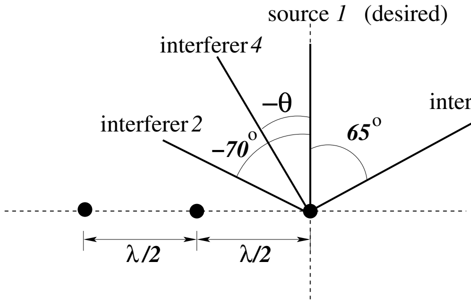
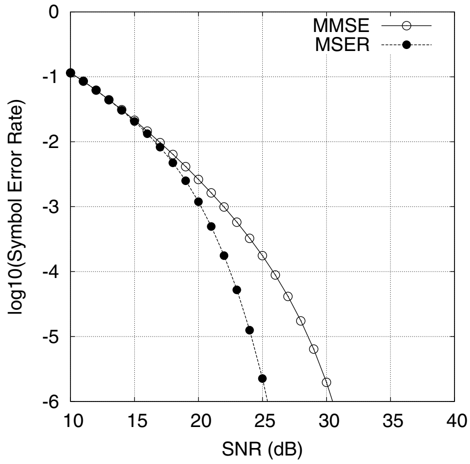
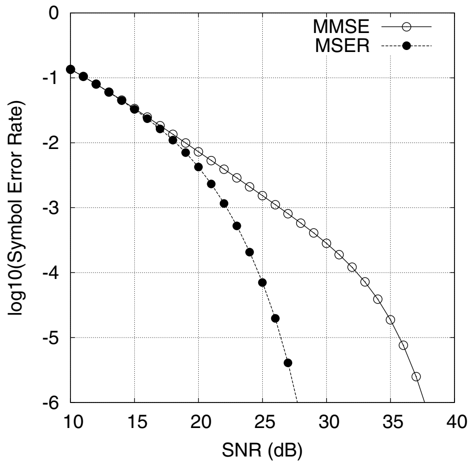


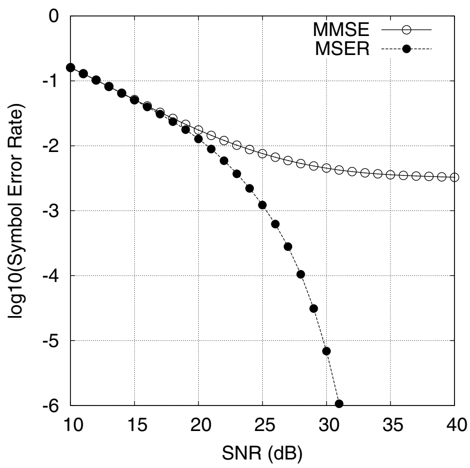
Fig. 2. Locations of the desired source and the interfering sources with respect to the three-element linear array with $\lambda/2$ element spacing, λ being the wavelength.



(a) $\theta = 32°$

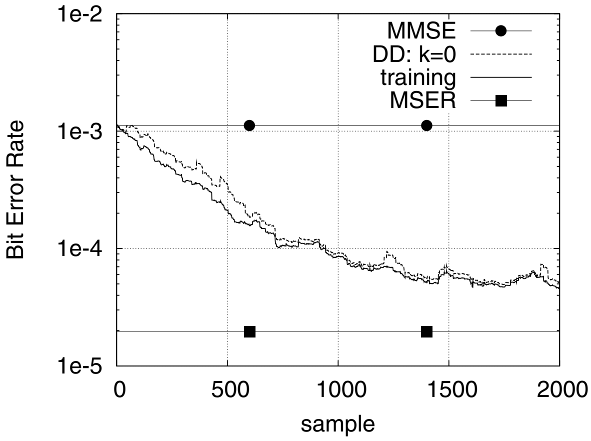


(b) $\theta = 30°$

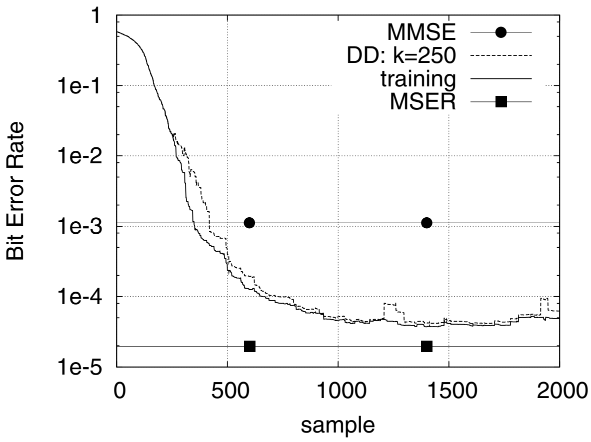


(c) $\theta = 28°$

Fig. 3. SER performance over non-fading channels.



(a) $\mathbf{w}(0) = \mathbf{w}_{\text{MMSE}}$



(b) $\mathbf{w}(0) = [0.1 + j0.1 \; 0.1 - j0.01 \; 0.1 - j0.1]^T$

Fig. 4. Learning curves of the stochastic gradient adaptive LSER algorithm for the stationary system averaged over 20 runs, given $\theta = 30^\circ$ and SNR$=$ 26 dB, where DD denotes decision-directed adaptation with $\hat{b}_1(k)$ substituting for $b_1(k)$. The step size $\mu = 0.001$ and kernel width $\rho_n = \sigma_n$.

this example, the MSER beamformer achieved a significantly better performance than the MMSE beamformer.

The erformance of the adaptive LSER algorithm was investigated using the system associated with $\theta = 30^\circ$ and SNR$=$ 26 dB. Given $\mathbf{w}(0) = \mathbf{w}_{\text{MMSE}}$, the step size of $\mu = 0.001$ and the kernel width $\rho_n = \sigma_n$, Fig. 4 (a) depicts the learning curves of the LSER algorithm, where DD denotes the decision-directed adaptation with $\hat{b}_1(k)$ substituting for $b_1(k)$. Fig. 4 (b) portrays the associated learning curves of the LSER algorithm under the same conditions, except for $\mathbf{w}(0) = [0.1{+}j0.1 \; 0.1{-}j0.01 \; 0.1{-}j0.1]^T$. It can be seen from Fig. 4 that the LSER beamformer had a reasonable convergence speed. It can also be seen that the initial condition $\mathbf{w}(0)$ had some influence on the achievable convergence rate.

V. Conclusions

An adaptive MSER beamforming technique has been proposed for multiple antenna aided multiuser wireless communication systems using 16QAM signalling. It has been demonstrated that the MSER beamforming design is capable of providing significant SER performance improvements over the standard MMSE beamforming design. An adaptive implementation of the MSER beamforming solution has been proposed, namely the stocastic gradient adaptive algorithm, which was referred to here as the LSER technique. Our future research will consider similar schemes employing error correction codecs and iterative receivers.

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. Communications*, Vol.42, No.2, pp.1740–1751, February/March/April 1994.

[2] J. Litva and T. K.Y. Lo, *Digital Beamforming in Wireless Communications*. London: Artech House, 1996.

[3] L. C. Godara, "Applications of antenna arrays to mobile communications, Part I: Performance improvement, feasibility, and system considerations," *Proc. IEEE*, Vol.85, No.7, pp.1031–1060, 1997.

[4] A.J. Paulraj and C.B. Papadias, "Space-time processing for wireless communications," *IEEE Signal Processing Magazine*, Vol.14, No.6, pp.49–83, 1997.

[5] R. Kohno, "Spatial and temporal communication theory using adaptive antenna array," *IEEE Personal Communications*, Vol.5, No.1, pp.28–35, 1998.

[6] P. Petrus, R.B. Ertel and J.H. Reed, "Capacity enhancement using adaptive arrays in an AMPS system," *IEEE Trans. Vehicular Technology*, Vol.47, No.3, pp.717–727, 1998.

[7] J.H. Winters, "Smart antennas for wireless systems," *IEEE Personal Communications*, Vol.5, No.1, pp.23–27, 1998.

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

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

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

[11] S. Chen, N.N. Ahmad and L. Hanzo, "Adaptive minimum bit error rate beamforming," *IEEE Trans. Wireless Communications*, Vol.4, No.2, pp.341–348, 2005.

[12] 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.

[13] L. Hanzo, S.X. Ng, T. Keller and W. Webb, *Quadrature Amplitude Modulation: From Basics to Adaptive Trellis-Coded, Turbo-Equalised and Space-Time Coded OFDM, CDMA and MC-CDMA Systems*. John Wiley and IEEE Press: Chichester, UK, 2004.

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

[15] C.-C. Yeh and J.R. Barry, "Adaptive minimum symbol-error rate equalization for quadrature-amplitude modulation," *IEEE Trans. Signal Processing*, Vol.51, No.12, pp.3263–3269, 2003.

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

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

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