



Available online at www.sciencedirect.com

SCIENCE @ DIRECT®

Digital Signal Processing 15 (2005) 545–567

**Digital
Signal
Processing**

www.elsevier.com/locate/dsp

Adaptive minimum bit error rate beamforming assisted receiver for QPSK wireless communication

S. Chen^{*}, L. Hanzo, N.N. Ahmad, A. Wolfgang

School of Electronics and Computer Science, University of Southampton, Southampton, SO17 1BJ, UK

Available online 25 March 2005

Abstract

This paper considers interference limited communication systems where the desired user and interfering users are symbol-synchronized. A novel adaptive beamforming technique is proposed for quadrature phase shift keying (QPSK) receiver based directly on minimizing the bit error rate. It is demonstrated that the proposed minimum bit error rate (MBER) approach utilizes the system resource (antenna array elements) more intelligently, than the standard minimum mean square error (MMSE) approach. Consequently, an MBER beamforming assisted receiver is capable of providing significant performance gains in terms of a reduced bit error rate over an MMSE beamforming one. A block-data based adaptive implementation of the theoretical MBER beamforming solution is developed based on the classical Parzen window estimate of probability density function. Furthermore, a sample-by-sample adaptive implementation is also considered, and a stochastic gradient algorithm, called the least bit error rate, is derived for the beamforming assisted QPSK receiver.

© 2005 Elsevier Inc. All rights reserved.

Keywords: Adaptive beamforming; Quadrature phase shift keying; Mean square error; Least mean square; Bit error rate; Probability density function; Parzen window estimate; Least bit error rate

^{*} Corresponding author.

E-mail addresses: sqc@ecs.soton.ac.uk (S. Chen), lh@ecs.soton.ac.uk (L. Hanzo), nna00r@ecs.soton.ac.uk (N.N. Ahmad), aw03r@ecs.soton.ac.uk (A. Wolfgang).

1. Introduction

The ever-increasing demand for mobile communication capacity has motivated the development of adaptive antenna array assisted spatial processing techniques [1–12] in order to further improve the achievable spectral efficiency. A particular technique that has shown real promise in achieving substantial capacity enhancements is the use of adaptive beamforming with antenna arrays. Through appropriately combining the signals received by the different elements of an antenna array to form a single output, adaptive beamforming is capable of separating signals transmitted on the same carrier frequency, provided that they are separated sufficiently in the spatial domain. Classically, this has been achieved by minimizing the mean square error (MSE) between the desired output and the actual array output, and this principle is rooted in the traditional beamforming employed in sonar and radar systems. Adaptive implementation of the theoretical minimum MSE (MMSE) beamforming solution can readily be realized using temporal reference techniques [2–4,13–17]. Specifically, block-data based beamformer weight adaptation can be achieved using the so-called sample matrix inversion (SMI) algorithm [13,14], while sample-by-sample adaptation can be carried out using the least mean square (LMS) algorithm [15–17].

For a communication system, it is the achievable bit error rate (BER), not the MSE performance, that really matters. Ideally, the system design should be based directly on minimizing the BER, rather than the MSE. For applications in single-user channel equalization and code division multiple access (CDMA) multiuser detection, it has been shown that the MMSE solution can in certain situations be distinctly inferior in comparison to the minimum BER (MBER) solution, and several adaptive implementations of the theoretical MBER solution have been studied in Refs. [18–22]. The recent conference paper [23] of the authors proposed an MBER beamforming assisted receiver for binary phase shift keying (BPSK) communication systems, where the desired user and interfering users are symbol-synchronized. This paper derives an adaptive beamforming technique based on directly minimizing the system's BER for such interference limited systems employing quadrature phase shift keying (QPSK) modulation. It is demonstrated that the MBER solution utilizes the array weights more intelligently than the MMSE approach. The MBER beamforming appears to be "smarter" than the MMSE solution, since it directly optimizes the system's BER performance, rather than minimizing the MSE, where the latter strategy often turns out to be deficient. In particular, when facing strong interfering sources, the MMSE beamforming receiver may exhibit a high BER floor as the underlying signal classes become linearly inseparable, while the MBER beamforming receiver can often maintain the linear separability and hence avoids such a BER floor.

An adaptive implementation of the theoretical MBER beamforming technique is studied in this paper. The classic Parzen window or kernel density estimation technique [24–26] is adopted for approximating the probability density function (pdf) of the beamformer's output, and this naturally leads to a block-data adaptive MBER algorithm, which iteratively minimizes the estimated BER of the beamforming assisted receiver by adjusting the beamformer weights using a simplified conjugate gradient optimization method [22,27]. It is demonstrated in a simulation study that this block-data adaptive MBER algorithm converges rapidly and the length of the data block required for achieving an accurate approximation of the MBER solution is reasonably small. Sample-by-sample adaptation has

also been considered and an adaptive stochastic gradient MBER algorithm, referred to as the least bit error rate (LBER), is derived. This LBER algorithm has a low computational complexity which is comparable to that of the very simple LMS algorithm. Simulation results suggest that the LBER algorithm converges reasonably fast.

Before presenting our novel MBER QPSK beamforming technique, the assumption that the desired user and interfering signals are symbol-synchronized is elaborated on. For such a symbol-synchronized interference-limited QPSK system the non-Gaussian nature of the interfering signals is effectively exploited by the MBER beamforming receiver, resulting in an improved BER performance. For the downlink scenario synchronous transmission of the users is guaranteed. By contrast, in an uplink scenario the differently delayed asynchronous signals of the users are no longer automatically synchronized. However, the quasi-synchronous operation of the system may be achieved with the aid of adaptive timing advance control, as in the global system of mobile communications, known as GSM [28]. The GSM system has a timing-advance control accuracy of 0.25 bit duration. Since synchronous systems perform better than their asynchronous counterparts [29], the third-generation partnership research consortium known as 3GPP is also considering the employment of timing-advance control in next-generation systems. In general, when the number of users is large, the users are asynchronous and the idealistic assumption of perfect power control is stipulated, the performance gain of the (symbol-rate) MBER solution over the MMSE beamformer may be expected to diminish, since the interference becomes nearly Gaussian at the symbol-rate samples. One way of maintaining the benefits of the MBER solution for asynchronous systems is to perform a joint MBER detection and synchronization by sampling faster than the symbol rate. During each symbol period, several signal samples are taken and the receiver maintains several tentative MBER detectors. The detector having the smallest BER is chosen to perform symbol detection.

2. System model

It is assumed that the system supports M symbol-synchronized users, that is, there exist M synchronized signal sources, and each user transmits a QPSK modulated signal on the same carrier frequency of $\omega = 2\pi f$. The baseband complex-valued signal of user i , sampled at the symbol rate, is formulated as

$$m_i(k) = A_i b_i(k), \quad 1 \leq i \leq M, \quad (1)$$

where $b_i(k) \in \{\pm 1 \pm j\}$ are QPSK symbols, the complex-valued A_i is the channel coefficient for user i multiplying by the transmitted signal amplitude of user i , and therefore $2|A_i|^2$ denotes the received signal power of user i . Without any loss of generality, source 1 is assumed to be the desired user and the rest of the sources are the interfering users. The linear antenna array considered consists of L uniformly spaced elements, and the signals received by the L -element antenna array are given by

$$x_l(k) = \sum_{i=1}^M m_i(k) \exp(j\omega t_l(\theta_i)) + n_l(k) = \bar{x}_l(k) + n_l(k), \quad 1 \leq l \leq L, \quad (2)$$

where $t_l(\theta_i)$ is the relative time delay at array element l for source i , θ_i is the direction of arrival for source i , and $n_l(k)$ is the complex-valued white Gaussian noise having a zero mean and a variance of $E[|n_l(k)|^2] = 2\sigma_n^2$. The desired user's signal to noise ratio is defined as $\text{SNR} = |A_1|^2/\sigma_n^2$, the interference to noise ratio of interfering user i is given by $\text{INR}_i = |A_i|^2/\sigma_n^2$, the desired signal to interference ratio with respect to interfering user i is defined as $\text{SIR}_i = |A_1|^2/|A_i|^2$, for $i = 2, \dots, M$, and the desired signal to interference plus noise ratio is given by $\text{SINR} = |A_1|^2/(\sum_{i=2}^M |A_i|^2 + \sigma_n^2)$. In a vector form, the array input $\mathbf{x}(k) = [x_1(k) x_2(k) \dots x_L(k)]^T$ can be expressed as

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

where $\mathbf{n}(k) = [n_1(k) n_2(k) \dots n_L(k)]^T$ has a covariance matrix of $E[\mathbf{n}(k)\mathbf{n}^H(k)] = 2\sigma_n^2\mathbf{I}_L$ with \mathbf{I}_L representing the $L \times L$ identity matrix, the system matrix \mathbf{P} is given by

$$\mathbf{P} = [A_1\mathbf{s}_1 \ A_2\mathbf{s}_2 \ \dots \ A_M\mathbf{s}_M], \quad (4)$$

the steering vector for source i is formulated as

$$\mathbf{s}_i = [\exp(j\omega t_1(\theta_i)) \exp(j\omega t_2(\theta_i)) \dots \exp(j\omega t_L(\theta_i))]^T \quad (5)$$

and the transmitted QPSK symbol vector is $\mathbf{b}(k) = [b_1(k) b_2(k) \dots b_M(k)]^T$.

A standard linear beamformer is employed at the receiver, and the beamformer's output is given by

$$y(k) = \mathbf{w}^H \mathbf{x}(k) = \mathbf{w}^H \bar{\mathbf{x}}(k) + \mathbf{w}^H \mathbf{n}(k) = \bar{y}(k) + e(k), \quad (6)$$

where $\mathbf{w} = [w_1 \ w_2 \ \dots \ w_L]^T$ is the complex-valued beamformer weight vector, and $e(k)$ is Gaussian distributed having a zero mean and a variance of $E[|e(k)|^2] = 2\sigma_n^2 \mathbf{w}^H \mathbf{w}$. Define the combined impulse response of the beamformer and the system as

$$\mathbf{w}^H \mathbf{P} = [c_1 \ c_2 \ \dots \ c_M]. \quad (7)$$

The beamformer's output can alternatively be expressed as

$$y(k) = c_1 b_1(k) + \sum_{i=2}^M c_i b_i(k) + e(k), \quad (8)$$

where the first term is the desired signal and the second term the residual interference. Define the decision variable as $d(k) = y(k)/c_1$. Then the decision regarding the transmitted symbol $b_1(k)$ is made according to

$$\hat{b}_1(k) = \text{sgn}(d_R(k)) + j \text{sgn}(d_I(k)), \quad (9)$$

where $d_R(k) = \Re[d(k)]$ and $d_I(k) = \Im[d(k)]$ are the real and imaginary parts of $d(k)$, respectively, and $\text{sgn}(\cdot)$ the sign function. Noting $c_1 = \mathbf{w}^H \mathbf{p}_1$ and $\mathbf{p}_1 = A_1 \mathbf{s}_1$, we can see that the steering vector \mathbf{s}_1 and the channel A_1 of the desired user are required at the receiver in order to make the unbiased decision (9). This fact is often overlooked. Provided that c_1 is real and positive, the optimal unbiased decision (9) is equivalent to

$$\hat{b}_1(k) = \text{sgn}(y_R(k)) + j \text{sgn}(y_I(k)). \quad (10)$$

The following rotating operation

$$\mathbf{w}^{\text{new}} = \frac{c_1^{\text{old}}}{|c_1^{\text{old}}|} \mathbf{w}^{\text{old}} \quad (11)$$

can be used to make sure that c_1 is real and positive. This rotation is a linear transformation and does not alter the BER of the underlying system.

Classically, the beamformer's weight vector is determined by minimizing the MSE term of $E[|b_1(k) - y(k)|^2]$, which leads to the following MMSE solution:

$$\mathbf{w}_{\text{MMSE}} = (\mathbf{P}\mathbf{P}^H + \sigma_n^2 \mathbf{I}_L)^{-1} \mathbf{p}_1. \quad (12)$$

Although the system matrix \mathbf{P} is generally unknown, the MMSE solution can be readily realized using the block-data SMI algorithm or the least squares (LS) algorithm [13,14] if training is available. The MMSE solution can also be implemented via training by using the stochastic gradient algorithm known as the LMS algorithm [15–17].

3. Minimum bit error rate beamforming

Denote the $N_b = 4^M$ number of possible transmitted symbol sequences of $\mathbf{b}(k)$ as $\mathbf{b}^{(q)}$, $1 \leq q \leq N_b$. Denote furthermore the first element of $\mathbf{b}^{(q)}$, corresponding to the desired user, as $b_1^{(q)}$. The noise-free part of the array input signal, namely $\bar{\mathbf{x}}(k)$, only takes values from the finite signal set defined as

$$\mathcal{X} \triangleq \{\bar{\mathbf{x}}^{(q)} = \mathbf{P}\mathbf{b}^{(q)}, 1 \leq q \leq N_b\}. \quad (13)$$

This set can be partitioned into four subsets, depending on the specific value of $b_1(k)$, as follows:

$$\mathcal{X}_{\pm, \pm} \triangleq \{\bar{\mathbf{x}}^{(q)} \in \mathcal{X}: b_1(k) = \pm 1 \pm j\}. \quad (14)$$

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

$$\mathcal{Y} \triangleq \{\bar{y}^{(q)} = \mathbf{w}^H \bar{\mathbf{x}}^{(q)}, 1 \leq q \leq N_b\} \quad (15)$$

and \mathcal{Y} can be divided into the four subsets conditioned on the value of $b_1(k)$:

$$\mathcal{Y}_{\pm, \pm} \triangleq \{\bar{y}^{(q)} \in \mathcal{Y}: b_1(k) = \pm 1 \pm j\}. \quad (16)$$

For the linear beamformer (6) to perform adequately, an implicit assumption is that $\mathcal{X}_{\pm, \pm}$ are linearly separable, that is, there exists a weight vector \mathbf{w} such that the four scalar sets $\mathcal{Y}_{\pm, \pm}$ are completely separable by linear decision boundaries. Otherwise nonlinear beamforming is required, a situation that is similar to nonlinear single-user equalization and nonlinear CDMA multiuser detection [30–35]. In this study, we restrict to the linear beamformer (6), because it has a low computational complexity and can readily be implemented in downlink receivers. It is worth pointing out that, under severe interference situations which cause the MMSE beamforming to lose linear separability, the MBER beamforming to be derived here can often maintain linear separability. This is because the objective of

the MMSE solution is to minimize the MSE which does not necessarily require the separation of $\mathcal{Y}_{\pm,\pm}$, while the MBER solution will try to separate $\mathcal{Y}_{\pm,\pm}$ as far apart as possible. In this sense, the MBER beamforming is more intelligent than the MMSE one. This will also be demonstrated later in the simulation study.

Noting $y(k) = \bar{y}(k) + e(k)$, it is easily seen that the conditional pdf of $y(k)$ given $b_1(k) = +1 + j$ is

$$p(y|+1+j) = \frac{1}{N_{sb}} \sum_{\bar{y}^{(q)} \in \mathcal{Y}_{+,+}} \frac{1}{2\pi\sigma_n^2 \mathbf{w}^H \mathbf{w}} \exp\left(-\frac{|y - \bar{y}^{(q)}|^2}{2\sigma_n^2 \mathbf{w}^H \mathbf{w}}\right), \quad (17)$$

where $N_{sb} = N_b/4$ is the number of the 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

$$p(y_R|+1+j) = \frac{1}{N_{sb}} \sum_{\bar{y}^{(q)} \in \mathcal{Y}_{+,+}} \frac{1}{\sqrt{2\pi\sigma_n^2 \mathbf{w}^H \mathbf{w}}} \exp\left(-\frac{(y_R - \bar{y}_R^{(q)})^2}{2\sigma_n^2 \mathbf{w}^H \mathbf{w}}\right) \quad (18)$$

and

$$p(y_I|+1+j) = \frac{1}{N_{sb}} \sum_{\bar{y}^{(q)} \in \mathcal{Y}_{+,+}} \frac{1}{\sqrt{2\pi\sigma_n^2 \mathbf{w}^H \mathbf{w}}} \exp\left(-\frac{(y_I - \bar{y}_I^{(q)})^2}{2\sigma_n^2 \mathbf{w}^H \mathbf{w}}\right), \quad (19)$$

respectively. Define

$$P_{E_R}(\mathbf{w}) \triangleq \text{Prob}(\Re[\hat{b}_1(k)] \neq \Re[b_1(k)]) = \text{Prob}(\hat{b}_{R,1}(k) \neq b_{R,1}(k)) \quad (20)$$

and

$$P_{E_I}(\mathbf{w}) \triangleq \text{Prob}(\Im[\hat{b}_1(k)] \neq \Im[b_1(k)]) = \text{Prob}(\hat{b}_{I,1}(k) \neq b_{I,1}(k)). \quad (21)$$

Obviously, the BER of the beamformer associated with the weight vector \mathbf{w} is given by

$$P_E(\mathbf{w}) = \frac{1}{2}(P_{E_R}(\mathbf{w}) + P_{E_I}(\mathbf{w})). \quad (22)$$

Noting the decision rule (10) (assuming that c_1 is positive) and the two marginal conditional pdfs (18) and (19), it can easily be shown that

$$P_{E_R}(\mathbf{w}) = \frac{1}{N_{sb}} \sum_{\bar{y}^{(q)} \in \mathcal{Y}_{+,+}} Q(g_R^{(q)}(\mathbf{w})) \quad (23)$$

and

$$P_{E_I}(\mathbf{w}) = \frac{1}{N_{sb}} \sum_{\bar{y}^{(q)} \in \mathcal{Y}_{+,+}} Q(g_I^{(q)}(\mathbf{w})), \quad (24)$$

where

$$Q(u) = \frac{1}{\sqrt{2\pi}} \int_u^\infty \exp\left(-\frac{v^2}{2}\right) dv, \quad (25)$$

$$g_R^{(q)}(\mathbf{w}) = \frac{\text{sgn}(\Re[b_1^{(q)}])\bar{y}_R^{(q)}}{\sigma_n \sqrt{\mathbf{w}^H \mathbf{w}}} = \frac{\text{sgn}(b_{R,1}^{(q)})\Re[\mathbf{w}^H \bar{\mathbf{x}}^{(q)}]}{\sigma_n \sqrt{\mathbf{w}^H \mathbf{w}}} \quad (26)$$

and

$$g_I^{(q)}(\mathbf{w}) = \frac{\text{sgn}(\Im[b_1^{(q)}])\bar{y}_I^{(q)}}{\sigma_n\sqrt{\mathbf{w}^H\mathbf{w}}} = \frac{\text{sgn}(b_{I,1}^{(q)})\Im[\mathbf{w}^H\bar{\mathbf{x}}^{(q)}]}{\sigma_n\sqrt{\mathbf{w}^H\mathbf{w}}}. \quad (27)$$

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

The MBER beamforming solution is then defined as

$$\mathbf{w}_{\text{MBER}} = \arg \min_{\mathbf{w}} P_E(\mathbf{w}). \quad (28)$$

Unlike the MMSE solution (12), there exists no closed-form MBER solution, and a numerical optimization must be used in order to obtain an MBER solution. The gradient of $P_E(\mathbf{w})$ with respect to \mathbf{w} is

$$\nabla P_E(\mathbf{w}) = \frac{1}{2}(\nabla P_{E_R}(\mathbf{w}) + \nabla P_{E_I}(\mathbf{w})), \quad (29)$$

and it can be shown that

$$\begin{aligned} \nabla P_{E_R}(\mathbf{w}) &= \frac{1}{2N_{sb}\sqrt{2\pi}\sigma_n\sqrt{\mathbf{w}^H\mathbf{w}}} \\ &\times \sum_{\bar{y}^{(q)} \in \mathcal{Y}_{+,+}} \exp\left(-\frac{(\bar{y}_R^{(q)})^2}{2\sigma_n^2\mathbf{w}^H\mathbf{w}}\right) \text{sgn}(b_{R,1}^{(q)}) \left(\frac{\bar{y}_R^{(q)}\mathbf{w}}{\mathbf{w}^H\mathbf{w}} - \bar{\mathbf{x}}^{(q)}\right) \end{aligned} \quad (30)$$

and

$$\begin{aligned} \nabla P_{E_I}(\mathbf{w}) &= \frac{1}{2N_{sb}\sqrt{2\pi}\sigma_n\sqrt{\mathbf{w}^H\mathbf{w}}} \\ &\times \sum_{\bar{y}^{(q)} \in \mathcal{Y}_{+,+}} \exp\left(-\frac{(\bar{y}_I^{(q)})^2}{2\sigma_n^2\mathbf{w}^H\mathbf{w}}\right) \text{sgn}(b_{I,1}^{(q)}) \left(\frac{\bar{y}_I^{(q)}\mathbf{w}}{\mathbf{w}^H\mathbf{w}} + j\bar{\mathbf{x}}^{(q)}\right). \end{aligned} \quad (31)$$

Given the gradient (29)–(31), the optimization problem (28) can be solved for iteratively using a gradient-based optimization algorithm. Since the BER is invariant to a positive scaling of \mathbf{w} , it is computationally advantageous to normalize \mathbf{w} to a unit-length after every iteration, so that the gradient (30) and (31) can be simplified to

$$\nabla P_{E_R}(\mathbf{w}) = \frac{1}{2N_{sb}\sqrt{2\pi}\sigma_n} \sum_{\bar{y}^{(q)} \in \mathcal{Y}_{+,+}} \exp\left(-\frac{(\bar{y}_R^{(q)})^2}{2\sigma_n^2}\right) \text{sgn}(b_{R,1}^{(q)}) (\bar{y}_R^{(q)}\mathbf{w} - \bar{\mathbf{x}}^{(q)}) \quad (32)$$

and

$$\nabla P_{E_I}(\mathbf{w}) = \frac{1}{2N_{sb}\sqrt{2\pi}\sigma_n} \sum_{\bar{y}^{(q)} \in \mathcal{Y}_{+,+}} \exp\left(-\frac{(\bar{y}_I^{(q)})^2}{2\sigma_n^2}\right) \text{sgn}(b_{I,1}^{(q)}) (\bar{y}_I^{(q)}\mathbf{w} + j\bar{\mathbf{x}}^{(q)}). \quad (33)$$

The rotating operation (11) should also be applied after each iteration to ensure a real and positive c_1 . The following simplified conjugate gradient algorithm [22,27] provides an efficient means of finding a MBER solution.

Initialization. Choose a step size of $\mu > 0$ and a termination scalar of $\beta > 0$; given $\mathbf{w}(1)$ and $\mathbf{d}(1) = -\nabla P_E(\mathbf{w}(1))$; set the iteration index to $\iota = 1$.

Loop. If $\|\nabla P_E(\mathbf{w}(\iota))\| = \sqrt{(\nabla P_E(\mathbf{w}(\iota)))^H \nabla P_E(\mathbf{w}(\iota))} < \beta$, goto *Stop*. Else,

$$\mathbf{w}(\iota + 1) = \mathbf{w}(\iota) + \mu \mathbf{d}(\iota),$$

$$c_1 = \mathbf{w}^H(\iota + 1) \mathbf{p}_1,$$

$$\mathbf{w}(\iota + 1) = \frac{c_1}{|c_1|} \mathbf{w}(\iota + 1),$$

$$\mathbf{w}(\iota + 1) = \frac{\mathbf{w}(\iota + 1)}{\|\mathbf{w}(\iota + 1)\|},$$

$$\phi_\iota = \frac{\|\nabla P_E(\mathbf{w}(\iota + 1))\|^2}{\|\nabla P_E(\mathbf{w}(\iota))\|^2},$$

$$\mathbf{d}(\iota + 1) = \phi_\iota \mathbf{d}(\iota) - \nabla P_E(\mathbf{w}(\iota + 1))$$

$\iota = \iota + 1$, goto *Loop*.

Stop. $\mathbf{w}(\iota)$ is the solution.

At a minimum we have $\|\nabla P_E(\mathbf{w})\| = 0$. Hence the termination scalar β determines the accuracy of the solution obtained. The step size μ controls the rate of convergence. Typically, a much larger value of μ can be used compared to the steepest-descent gradient algorithm. As the BER surface $P_E(\mathbf{w})$ is highly nonlinear, occasionally the search direction \mathbf{d} may no longer be a good approximation to the conjugate gradient direction or may even point to the “uphill” direction, when the iteration index becomes large. It is thus advisable to periodically reset \mathbf{d} to the negative gradient in the above conjugate gradient algorithm. With this resetting mechanism, this simplified conjugate gradient algorithm has been shown to converge fast to the theoretical MBER solution, typically in tens of iterations, in many simulation studies. Although in theory there is no guarantee that the above conjugate gradient algorithm can always find a global minimum point of the BER surface $P_E(\mathbf{w})$, in practice we have found that the algorithm works well and we have never observed any occurrence of the algorithm being trapped at some local minimum solution. This is likely to be a consequence of the specific shape of the BER surface. Note that the BER is invariant to a positive scaling of \mathbf{w} , i.e. the size of \mathbf{w} does not matter (except zero size). Thus, the BER surface has an infinitely long valley, and any point at the bottom of this valley is a true global MBER solution. For an illustration, see the simple example given in Ref. [22]. Once a weight vector \mathbf{w} is near the edge of this infinitely long valley, convergence to the bottom is extremely fast, since the slope or gradient is large. Note that once we restrict to the unit-length \mathbf{w} , the MBER solution becomes unique. As alternatives to the simplified conjugate gradient algorithm, global optimization algorithms, such as the genetic algorithm [36,37] and adaptive simulated annealing [38,39], can be used to obtain a global minimum solution of $P_E(\mathbf{w})$, at an expense of considerably increased computational requirements.

4. Adaptive minimum bit error rate beamforming

Noting $y(k) = \bar{y}(k) + e(k)$ with $\bar{y}(k)$ taking values from \mathcal{Y} , the pdf of $y(k)$ can be shown to be explicitly given by

$$p(y) = \frac{1}{N_b 2\pi \sigma_n^2 \mathbf{w}^H \mathbf{w}} \sum_{q=1}^{N_b} \exp\left(-\frac{|y - \bar{y}^{(q)}|^2}{2\sigma_n^2 \mathbf{w}^H \mathbf{w}}\right) \quad (34)$$

and the BER can alternatively be calculated with the two “marginal” BERs given by

$$P_{E_R}(\mathbf{w}) = \frac{1}{N_b} \sum_{q=1}^{N_b} Q(g_R^{(q)}(\mathbf{w})) \quad (35)$$

and

$$P_{E_I}(\mathbf{w}) = \frac{1}{N_b} \sum_{q=1}^{N_b} Q(g_I^{(q)}(\mathbf{w})), \quad (36)$$

where the computation is over $\bar{y}^{(q)} \in \mathcal{Y}$. In reality, the pdf of $y(k)$ is unknown. Hence, we will adopt the temporal reference technique for supporting the adaptive implementation of the MBER beamforming.

4.1. A block-data based gradient adaptive MBER algorithm

The key to adaptive implementation of the MBER solution is an effective estimate of the pdf (34). Parzen window or kernel density estimate [24–26] is a well known method for estimating a probability distribution. Parzen window method estimates a pdf using a window or block of $y(k)$ by placing a symmetric unimodal kernel function on each $y(k)$. Kernel density estimation is capable of producing reliable pdf estimates with short data records and in particular is extremely natural when dealing with Gaussian mixtures, such as the one given in (34). In our particular application, it is obvious and natural to choose a Gaussian kernel function with a kernel width $\rho_n \sqrt{\mathbf{w}^H \mathbf{w}}$ that is similarly in form to the noise standard deviation $\sigma_n \sqrt{\mathbf{w}^H \mathbf{w}}$. Given a block of K training samples $\{\mathbf{x}(k), b_1(k)\}$, a kernel density estimate of the pdf (34) is readily given by

$$\hat{p}(y) = \frac{1}{K 2\pi \rho_n^2 \mathbf{w}^H \mathbf{w}} \sum_{k=1}^K \exp\left(-\frac{|y - y(k)|^2}{2\rho_n^2 \mathbf{w}^H \mathbf{w}}\right), \quad (37)$$

where the radius or scaling parameter ρ_n is related to the standard deviation σ_n of the system noise. Accuracy analysis of Parzen window density estimate is well documented in the literature. The pdf estimate (37) is known to possess a mean integrated square error convergence rate at order of K^{-1} [24]. Some examples of accurate pdf estimates using (37) with short data records can be seen in Refs. [21,22]. In Ref. [25], a lower bound of $\rho_n = (4/3K)^{1/5} \sigma_n$ is suggested. In practice, ρ_n can often be chosen from a large range of values.

From this estimated pdf (37), the estimated BER is given by

$$\hat{P}_E(\mathbf{w}) = \frac{1}{2}(\hat{P}_{E_R}(\mathbf{w}) + \hat{P}_{E_I}(\mathbf{w})) = \frac{1}{2K} \sum_{k=1}^K (Q(\hat{g}_R^{(k)}(\mathbf{w})) + Q(\hat{g}_I^{(k)}(\mathbf{w}))) \quad (38)$$

with

$$\hat{g}_R^{(k)}(\mathbf{w}) = \frac{\text{sgn}(b_{R,1}(k))y_R(k)}{\rho_n \sqrt{\mathbf{w}^H \mathbf{w}}} \quad (39)$$

and

$$\hat{g}_I^{(k)}(\mathbf{w}) = \frac{\text{sgn}(b_{I,1}(k))y_I(k)}{\rho_n \sqrt{\mathbf{w}^H \mathbf{w}}}. \quad (40)$$

The gradient of $\hat{P}_E(\mathbf{w})$ can readily be calculated with

$$\begin{aligned} \nabla \hat{P}_{E_R}(\mathbf{w}) &= \frac{1}{2K \sqrt{2\pi} \rho_n \sqrt{\mathbf{w}^H \mathbf{w}}} \\ &\times \sum_{k=1}^K \exp\left(-\frac{y_R^2(k)}{2\rho_n^2 \mathbf{w}^H \mathbf{w}}\right) \text{sgn}(b_{R,1}(k)) \left(\frac{y_R(k)\mathbf{w}}{\mathbf{w}^H \mathbf{w}} - \mathbf{x}(k)\right) \end{aligned} \quad (41)$$

and

$$\begin{aligned} \nabla \hat{P}_{E_I}(\mathbf{w}) &= \frac{1}{2K \sqrt{2\pi} \rho_n \sqrt{\mathbf{w}^H \mathbf{w}}} \\ &\times \sum_{k=1}^K \exp\left(-\frac{y_I^2(k)}{2\rho_n^2 \mathbf{w}^H \mathbf{w}}\right) \text{sgn}(b_{I,1}(k)) \left(\frac{y_I(k)\mathbf{w}}{\mathbf{w}^H \mathbf{w}} + j\mathbf{x}(k)\right). \end{aligned} \quad (42)$$

Upon substituting $\nabla P_E(\mathbf{w})$ by $\nabla \hat{P}_E(\mathbf{w})$ in the conjugate gradient updating mechanism, a block-data based adaptive algorithm is obtained. The step size μ and the radius parameter ρ_n are two algorithmic parameters. Again μ and ρ_n control the rate of convergence, and the radius parameter ρ_n also helps to determine the accuracy of the pdf and hence BER estimate.

4.2. A stochastic gradient based adaptive MBER algorithm

In the Parzen window estimate (37), the kernel width or scaling parameter $\rho_n \sqrt{\mathbf{w}^H \mathbf{w}}$ depends on the beamformer weight vector \mathbf{w} . In a general density estimate, there is no reason why the scaling parameter should be chosen in such a way except that we notice the dependency of the scaling parameter to \mathbf{w} in the true density (34). However, the BER is invariant to $\mathbf{w}^H \mathbf{w}$. To fully take advantage of this fact, we propose to use a constant width ρ_n in density estimate. One advantage of using a constant width ρ_n , rather than a variable one $\rho_n \sqrt{\mathbf{w}^H \mathbf{w}}$, in the density estimate is that the gradient of the resulting estimated BER has a much simpler form, which leads to considerable reduction in computational complexity. This is particularly relevant in the derivation of stochastic gradient updating

mechanisms. Adopting this approach, an alternative Parzen window estimate of the true pdf (34) is given by

$$\tilde{p}(y) = \frac{1}{K2\pi\rho_n^2} \sum_{k=1}^K \exp\left(-\frac{|y - y(k)|^2}{2\rho_n^2}\right) \quad (43)$$

and an approximation of the BER is

$$\tilde{P}_E(\mathbf{w}) = \frac{1}{2}(\tilde{P}_{E_R}(\mathbf{w}) + \tilde{P}_{E_I}(\mathbf{w})) = \frac{1}{2K} \sum_{k=1}^K (\mathcal{Q}(\tilde{g}_R^{(k)}(\mathbf{w})) + \mathcal{Q}(\tilde{g}_I^{(k)}(\mathbf{w}))) \quad (44)$$

with

$$\tilde{g}_R^{(k)}(\mathbf{w}) = \frac{\text{sgn}(b_{R,1}(k))y_R(k)}{\rho_n} \quad (45)$$

and

$$\tilde{g}_I^{(k)}(\mathbf{w}) = \frac{\text{sgn}(b_{I,1}(k))y_I(k)}{\rho_n}. \quad (46)$$

This approximation is valid, provided that the width ρ_n is chosen appropriately.

In order to derive a sample-by-sample adaptive algorithm, consider a single-sample estimate of $p(y)$, namely:

$$\tilde{p}(y, k) = \frac{1}{2\pi\rho_n^2} \exp\left(-\frac{|y - y(k)|^2}{2\rho_n^2}\right). \quad (47)$$

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 of

$$\nabla \tilde{P}_E(\mathbf{w}, k) = \frac{(-\text{sgn}(b_{R,1}(k)) \exp(-\frac{y_R^2(k)}{2\rho_n^2}) + j \text{sgn}(b_{I,1}(k)) \exp(-\frac{y_I^2(k)}{2\rho_n^2}))}{4\sqrt{2\pi}\rho_n} \mathbf{x}(k) \quad (48)$$

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

$$\mathbf{w}(k+1) = \mathbf{w}(k) + \mu \frac{(\text{sgn}(b_{R,1}(k)) \exp(-\frac{y_R^2(k)}{2\rho_n^2}) - j \text{sgn}(b_{I,1}(k)) \exp(-\frac{y_I^2(k)}{2\rho_n^2}))}{4\sqrt{2\pi}\rho_n} \mathbf{x}(k), \quad (49)$$

$$c_1 = \mathbf{w}^H(k+1)\mathbf{p}_1, \quad (50)$$

$$\mathbf{w}(k+1) = \frac{c_1}{|c_1|} \mathbf{w}(k+1), \quad (51)$$

where the adaptive gain μ and 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 convergence rate and steady-state BER misadjustment. Note that there is no need to normalize the weight vector to a unit-length after each update. This LBER

algorithm has a similar computational complexity to the low-complexity LMS algorithm, which has a weight updating equation given by

$$\mathbf{w}(k+1) = \mathbf{w}(k) + \mu(b_1(k) - y(k))^* \mathbf{x}(k). \quad (52)$$

Note that the rotation operation (50) and (51) are also required for the LMS beamforming in order to apply the decision rule (10).

5. Simulation study

5.1. Time-invariant system

The example consisted of four signal sources and a three-element linear antenna array. The array element spacing was $\lambda/2$ with λ being the wavelength. Figure 1 shows the locations of the desired source and three interfering sources graphically. The simulated channel conditions were $A_i = \alpha_i + j0$ for $1 \leq i \leq 4$, with $\alpha_i > 0$ so chosen to provide the required received signal powers. Figure 2 compares the BER performance of the MBER beamforming assisted receiver with that of the MMSE beamforming assisted receiver under four different conditions: (a) the desired user and all the three interfering sources had equal power, (b) the desired user and the interfering users 2 and 3 had equal power but the interfering source 4 had 6 dB more power than the desired user, (c) all the three interfering sources had 2 dB more power than the desired user, and (d) the interfering sources 2 and 4 had 2 dB more power, while the interferer 3 had 6 dB more power, than the desired user. The MMSE solution was calculated using (12) while the MBER solution was determined numerically using the simplified conjugate gradient algorithm presented in Section 3. For this example, the superior performance of the MBER beamforming technique over the MMSE scheme is evident. It can be seen from Figs. 2a–2c that, as the interference signals get stronger, the MMSE beamformer's performance deteriorates quickly and exhibits an irreducible BER floor. In contrast, the MBER solution shows some degree of robustness to the near-far effect. The first attempt to explain this phenomenon was made by examining the beam pattern used in traditional beamforming.

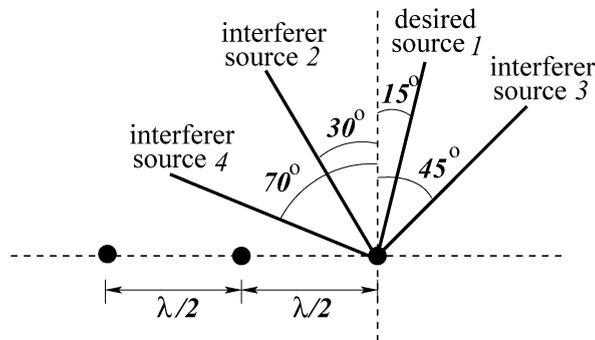


Fig. 1. Locations of the desired source and the three interfering sources with respect to the three-element linear array having $\lambda/2$ element spacing, where λ is the wavelength, for the time-invariant system.

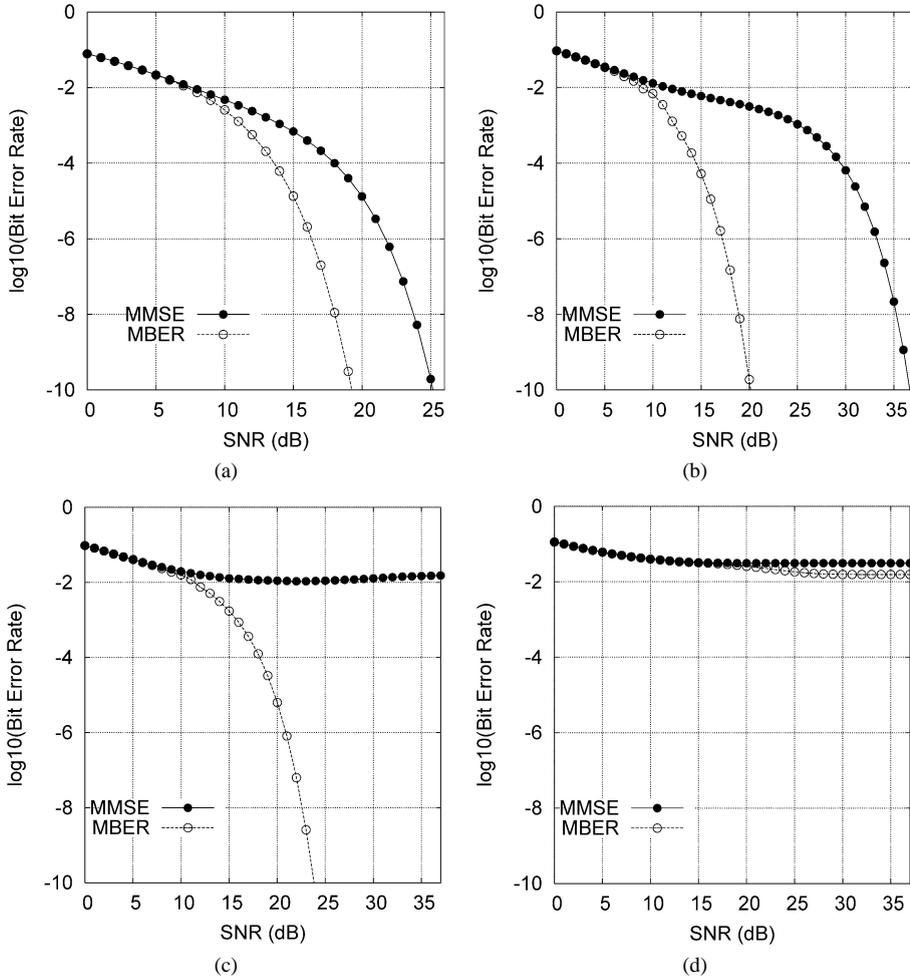


Fig. 2. Comparison of the bit error rates of the MMSE and MBER beamformers for the time-invariant system: (a) $SIR_i = 0$ dB for $i = 2, 3, 4$; (b) $SIR_2 = SIR_3 = 0$ dB and $SIR_4 = -6$ dB; (c) $SIR_i = -2$ dB for $i = 2, 3, 4$; and (d) $SIR_2 = SIR_4 = -2$ dB and $SIR_3 = -6$ dB.

The discrete Fourier transform of the beamformer weights, also referred to as the beam pattern, is given by

$$F(\theta) = \sum_{l=1}^L w_l \exp(-j\omega_l(\theta)), \tag{53}$$

which describes the response of the beamformer to the source arriving at angle θ . In traditional beamforming, the magnitude of $F(\theta)$ is used for characterizing the performance of a beamformer. Using the amplitude response alone, however, can be misleading, since both the magnitude and phase of $F(\theta)$ should be used together for characterizing the beamformer. Figure 3 shows the beam patterns for the MMSE and MBER beamform-

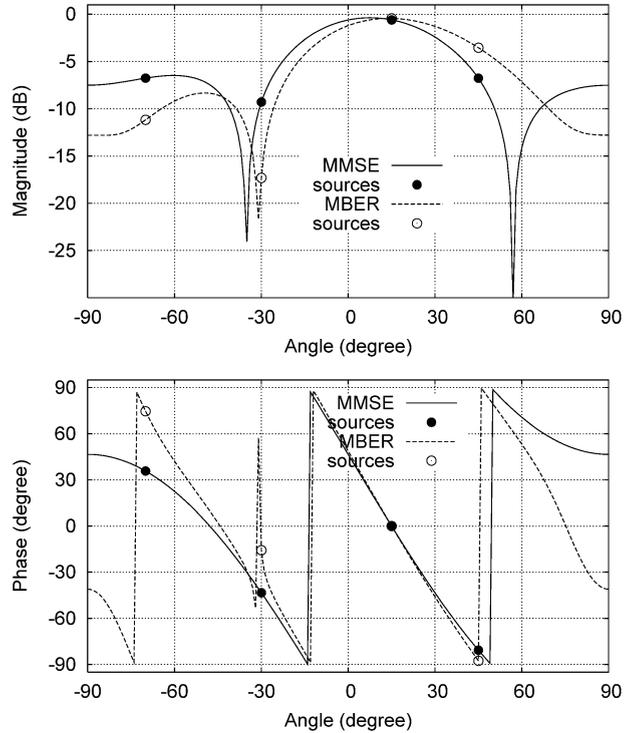


Fig. 3. Comparison of the MMSE and MBER beam patterns given $\text{SNR} = 15$ dB and $\text{SIR}_i = 0$ dB for $i = 2, 3, 4$ for the time-invariant system. The weight vector of the MBER solution is scaled to have the same length as the MMSE solution.

ers, respectively, given $\text{SNR} = 15$ dB, $\text{SIR}_i = 0$ dB for $i = 2, 3, 4$, which illustrates a condition represented in Fig. 2a. Figure 4 depicts the corresponding beam patterns given $\text{SNR} = 20$ dB, and $\text{SIR}_i = -2$ dB for $i = 2, 3, 4$, which represents a case of the conditions shown in Fig. 2c. The two beam patterns shown in Fig. 3 do not have big differences and thus it is difficult to explain from the beam pattern why the MBER solution has a much better BER performance than the MMSE scheme, as can be seen from Fig. 2a. Moreover, the beam patterns of Figs. 3 and 4 are similar, which could not explain why the MMSE scheme should have a high BER floor, as shown in Fig. 2c.

The pdf of the beamformer's output fully characterizes the true performance of the beamformer. Figures 5 and 6 depict the full conditional pdf $p(y|+1+j)$, the two marginal conditional pdfs $p(y_R|+1+j)$ and $p(y_I|+1+j)$ together with the subset $\mathcal{Y}_{+,+}$ for the MMSE and MBER beamformers under the same conditions as given in Figs. 3 and 4, respectively. In these two figures, the beamformer's weight vector has been normalized to a unit length. It can be seen from Fig. 5 that the minimum distance from $\mathcal{Y}_{+,+}$ to the decision boundaries for the MMSE case (approximately 0.3) is smaller than that for the MBER case (approximately 0.6), reflecting the fact that the MBER beamformer has a better BER performance than the MMSE one as shown in Fig. 2a. When the interference power increases to give rise to $\text{SIR}_i = -2$ dB for $i = 2, 3, 4$, the MMSE beamformer has

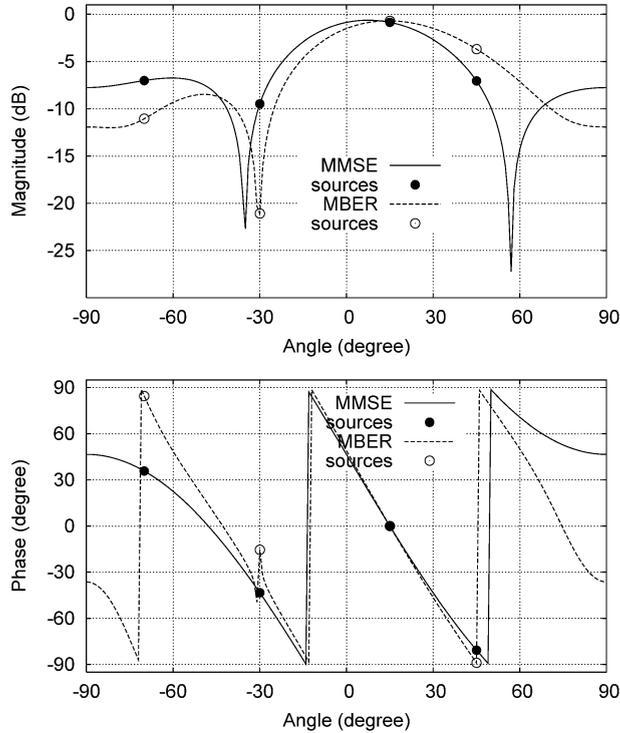
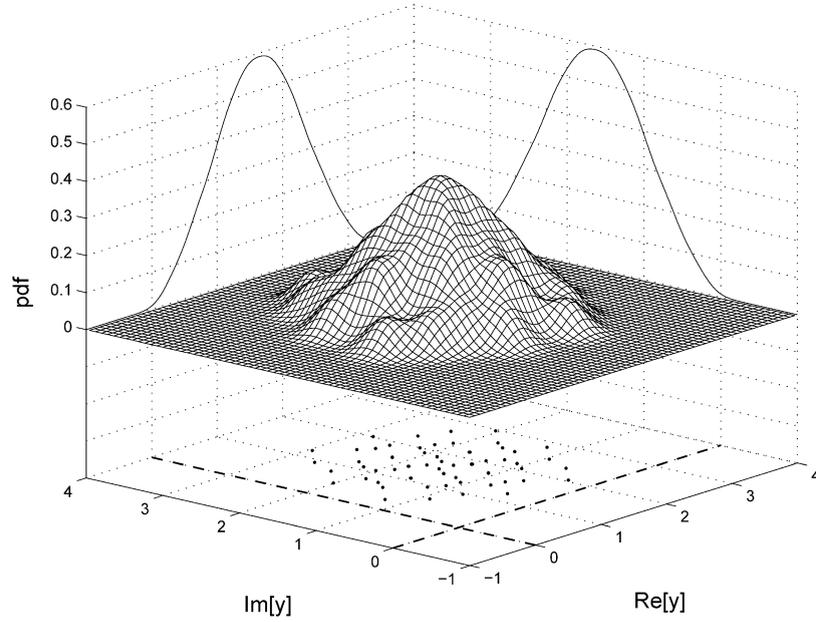


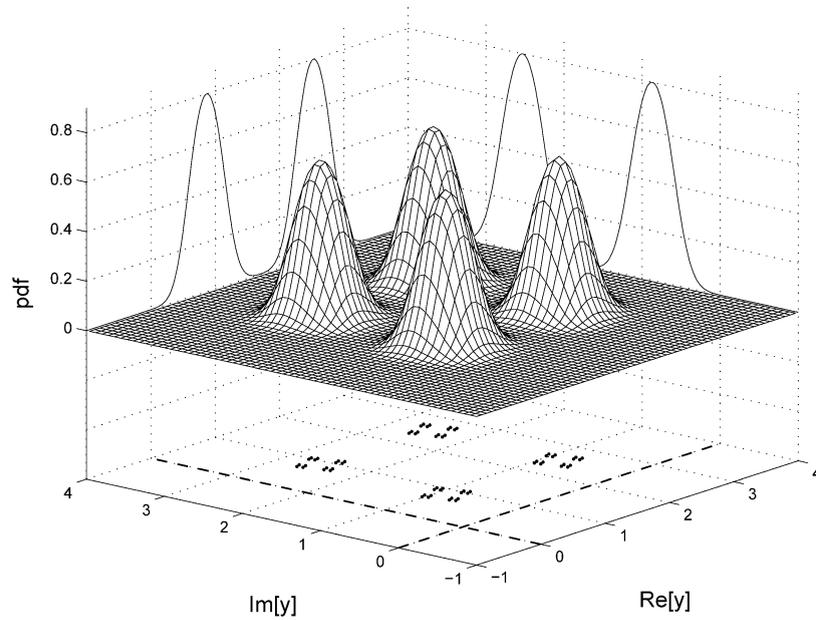
Fig. 4. Comparison of the MMSE and MBER beam patterns given $\text{SNR} = 20$ dB and $\text{SIR}_i = -2$ dB for $i = 2, 3, 4$ for the time-invariant system. The weight vector of the MBER solution is scaled to have the same length as the MMSE solution.

lost linear separability, as can be clearly seen in Fig. 6a, where the two circles mark the points of $\mathcal{Y}_{+,+}$ that have just crossed over to the wrong sides of the decision boundaries. This explains the irreducible high BER exhibited in Fig. 2c for the MMSE beamformer. In contrast, a desired linear separability is maintained for the MBER beamformer even under such an adverse condition. At the extremely adverse condition given in Fig. 2d, the underlying system becomes linearly inseparable, and any linear beamformer will exhibit a high BER floor. In such a situation, nonlinear beamforming may be employed to achieve an adequate performance at a cost of increased complexity [40].

Let us now study the performance of the block-data based gradient adaptive MBER algorithm employing the conjugate gradient updating mechanism presented in Section 4.1. The effect of the block size K on the performance of this block-data based adaptive MBER algorithm is investigated in Fig. 7, given the condition that the desired user and the interfering sources 2 and 3 had an equal power, while the interfering source 4 had a 6 dB higher power than the desired user. Note that for this example, the signal set \mathcal{X} contains 256 states, calculated using the formula $N_b = 4^M$ with $M = 4$. It is seen that with a short block length of $K = 100$, the BER performance of the block-data based adaptive MBER solution can closely approximate the performance of the theoretical MBER solution. Given $\text{SNR} = 17$ dB, $\text{SIR}_2 = \text{SIR}_3 = 0$ dB and $\text{SIR}_4 = -6$ dB and with two different initial

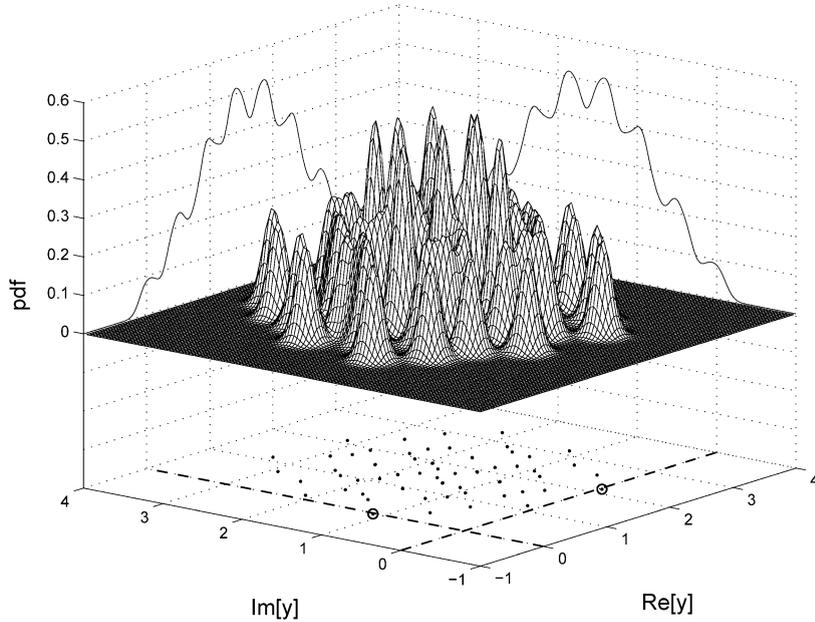


(a) MMSE beamforming

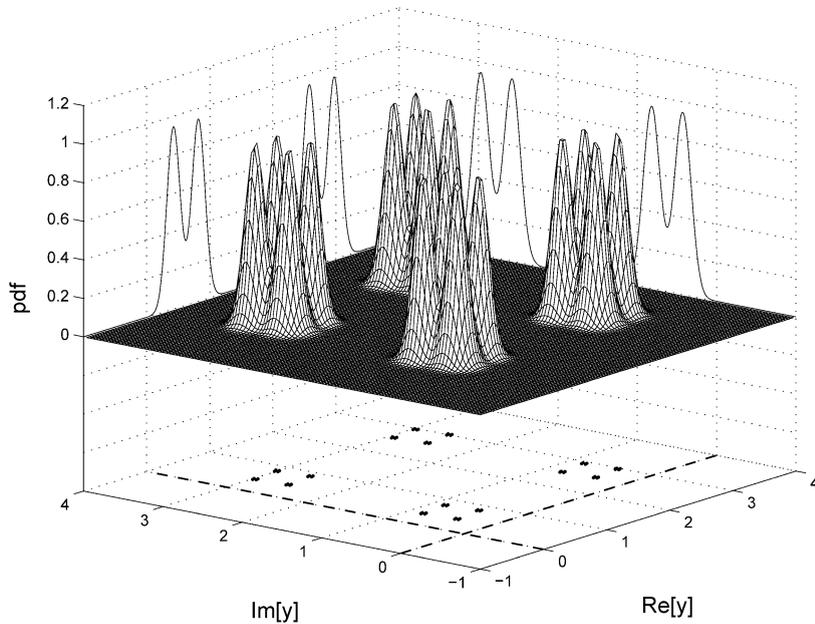


(b) MBER beamforming

Fig. 5. Conditional probability density functions $p(y|1+j)$ (surfaces), marginal conditional probability density functions $p(y_R|1+j)$ and $p(y_I|1+j)$ (curves), and signal subsets $\mathcal{Y}_{+,+}$ (dots) for the time-invariant system. SNR = 15 dB and SIR_{*i*} = 0 dB for $i = 2, 3, 4$. The beamformer's weight vector is normalized to unit length.



(a) MMSE beamforming



(b) MBER beamforming

Fig. 6. Conditional probability density functions $p(y|+1+j)$ (surfaces), marginal conditional probability density functions $p(y_R|+1+j)$ and $p(y_I|+1+j)$ (curves), and signal subsets $\mathcal{Y}_{+,+}$ (dots) for the time-invariant system. SNR = 20 dB and SIR_{*i*} = -2 dB for *i* = 2, 3, 4. The beamformer's weight vector is normalized to unit length. Circles in (a) indicate the points of $\mathcal{Y}_{+,+}$ that lie in the wrong sides of the decision boundaries.

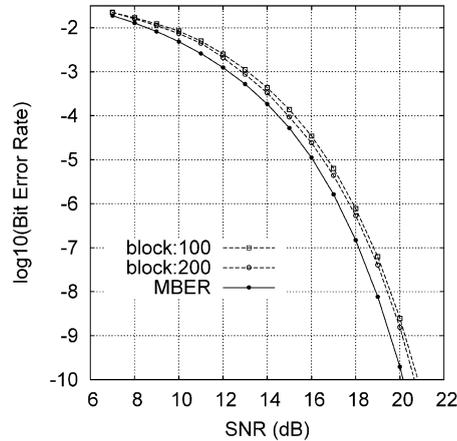


Fig. 7. Effect of block size on the performance of the block-data based gradient adaptive MBER algorithm of Section 4.1 for the time-invariant system. $SIR_2 = SIR_3 = 0$ dB and $SIR_4 = -6$ dB.

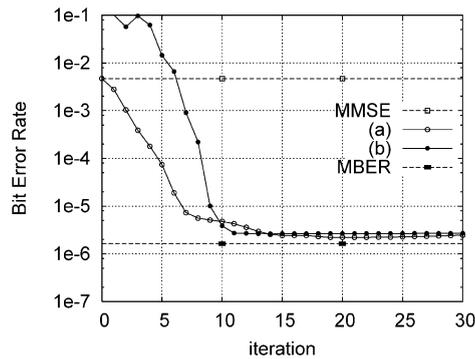


Fig. 8. Convergence rate of the block-data based gradient adaptive MBER algorithm of Section 4.1 for the time-invariant system with a block size of $K = 400$, $SNR = 17$ dB, $SIR_2 = SIR_3 = 0$ dB, and $SIR_4 = -6$ dB, and given (a): initial $\mathbf{w} = \mathbf{w}_{MMSE}$, $\mu = 0.3$ and $\rho_n^2 = 3\sigma_n^2 \approx 0.06$; and (b): initial $\mathbf{w} = [0.0 + j0.1 \ 0.1 + j0.0 \ 0.1 + j0.0]^T$, $\mu = 0.7$, and $\rho_n^2 = 3\sigma_n^2 \approx 0.06$.

weight conditions, Fig. 8 illustrates the convergence rates of the block-data based gradient adaptive MBER algorithm. From Fig. 8, it can be seen that this block-data based adaptive algorithm converges rapidly. The step size μ and radius parameter ρ_n used were chosen empirically to ensure a fast convergence speed. The influence of the scaling parameter ρ_n^2 on the performance of the block-data based adaptive MBER algorithm was investigated in Fig. 9 under the same condition given in Fig. 8b. It can be seen that the performance of the algorithm is not overly sensitive to a large range of ρ_n^2 values.

The performance of the stochastic gradient based adaptive MBER algorithm portrayed in Section 4.2 is investigated next. Figures 10a and 10b depict the learning curves of the LBER algorithm given two different initial weight conditions, respectively. Two kinds of learning curves are shown, obtained respectively by the LBER algorithm with training and with decision-directed (DD) adaptation in which $b_1(k)$ is substituted by its estimate $\hat{b}_1(k)$.

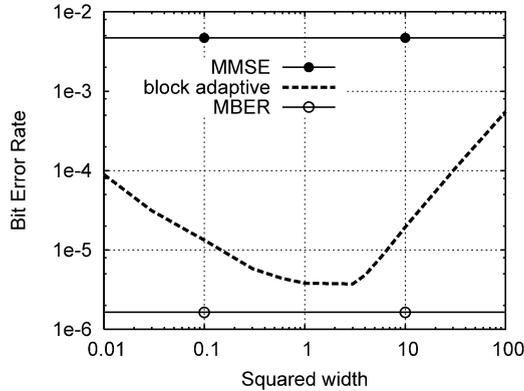


Fig. 9. Influence of the scaling parameter ρ_n^2 on the performance of the block-data based gradient adaptive MBER algorithm of Section 4.1 for the time-invariant system with a block size of $K = 400$, $\text{SNR} = 17$ dB, $\text{SIR}_2 = \text{SIR}_3 = 0$ dB, and $\text{SIR}_4 = -6$ dB, and given $\mathbf{w} = [0.0 + j0.1 \ 0.1 + j0.0 \ 0.1 + j0.0]^T$.

It can be seen that this stochastic gradient adaptive MBER algorithm converges reasonably fast. In Fig. 9a, the initial BER was lower than 10^{-2} , which was sufficient low for the DD adaptation. For the condition specified in Fig. 9b, however, 140 samples of training were used first to lower the BER before it switched to the DD adaptation. The adaptive gain μ and kernel width ρ_n were determined empirically to ensure a good performance in terms of convergence rate and steady-state BER misadjustment. As expected, initial weight condition affects convergence speed since the BER is a complicated nonlinear function of the weight vector. As a comparison, the learning curves of the LMS algorithm were also shown in Fig. 10. As expected, the BER of the LMS beamforming assisted receiver cannot be lower than that of the MMSE solution.

5.2. Slow-fading system

The locations of the four users were identical to those shown in Fig. 1 but the antenna array consisted of four $\lambda/2$ -spacing elements. The magnitudes of the 4 channels A_i , $1 \leq i \leq 4$, were independent Rayleigh processes and the associated root mean powers of A_i were $\sqrt{0.5} + j\sqrt{0.5}$, for $1 \leq i \leq 4$. Continuously fluctuating fading was used at a normalized Doppler frequency of 10^{-6} , providing a different fading magnitude and phase for each transmitted symbol. The transmission frame structure consisted of 40 training symbols followed by 400 data symbols. The performance of the LBER and LMS beamforming assisted receivers are compared in Fig. 11, where the superior performance of the LBER algorithm over the LMS one is evident. Note that this was not an over loaded system, since the number of the users was four and the number of the receiver antennas was also four.

6. Conclusions

An adaptive MBER beamforming technique has been developed for QPSK wireless systems, where the users transmissions are synchronized. It has been shown that the MBER

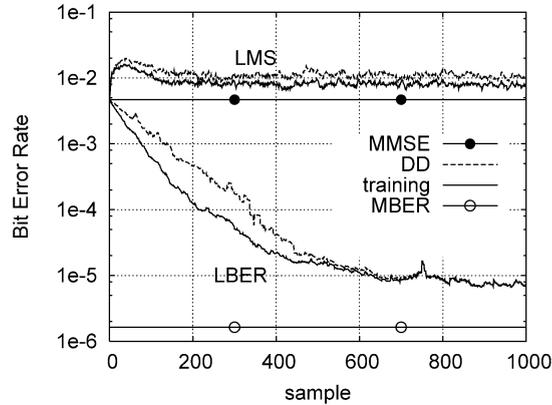
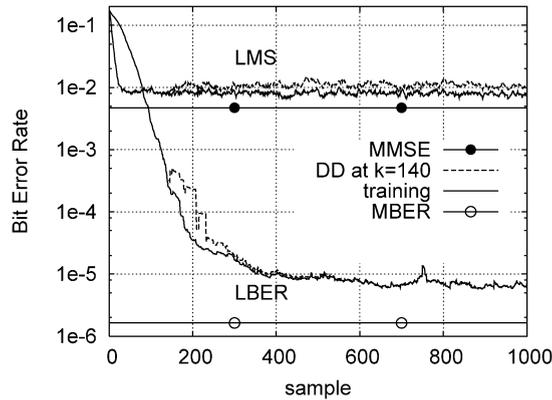
(a) $\mathbf{w}(0) = \mathbf{w}_{\text{MMSE}}$, $\mu = 0.03$, and $\rho_n^2 = 3\sigma_n^2 \approx 0.06$ (b) $\mathbf{w}(0) = [0.0 + j0.1 \ 0.1 + j0.0 \ 0.1 + j0.0]^T$, $\mu = 0.03$, and $\rho_n^2 = 3\sigma_n^2 \approx 0.06$

Fig. 10. Learning curves of the stochastic gradient adaptive MBER algorithm of Section 4.2 averaged over 100 runs for the time-invariant system, given $\text{SNR} = 17$ dB, $\text{SIR}_2 = \text{SIR}_3 = 0$ dB, and $\text{SIR}_4 = -6$ dB, where DD: decision-directed adaptation with $\hat{b}_1(k)$ substituting $b_1(k)$.

beamforming assisted receiver exploits the system's resources more intelligently than the standard MMSE beamforming assisted one and, consequently, can achieve a better performance in terms of a lower BER. Therefore, by employing the MBER beamforming assisted receiver, the achievable system capacity is enhanced. Simulation results also suggest that the MBER solution is more robust to the near-far effect, compared with the MMSE scheme. Adaptive implementation of the MBER beamformer has also been addressed. A block-data based conjugate gradient adaptive MBER algorithm has been shown to converge rapidly, while requiring a reasonably small block size for accurately approximating the theoretical MBER solution. A stochastic gradient adaptive MBER algorithm, called the LBER, have also been derived, which has been shown to have a reasonably fast convergence speed in simulation. Furthermore, this LBER algorithm has a low computational complexity that is comparable to the simple LMS algorithm. In this study, we assume narrow-band channels and deal with narrow-band beamforming (space processing only) assisted receiver. Further

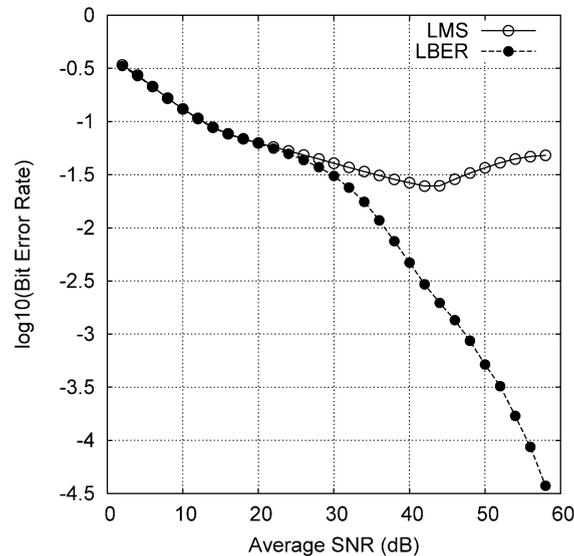


Fig. 11. Comparison of the bit error rates of the LMS and LBER beamforming assisted receivers for the slow-fading system.

work is required to investigate the general case of wideband channels and to study broadband beamforming (space-time processing) assisted receiver. Finally, we would also like to point out that the MBER beamforming solution derived for synchronous systems can be extended to asynchronous systems by considering a detection window of three symbols, where the two symbols of the asynchronous interferer overlap with the desired symbol of the reference user. Naturally, this increases the detection complexity. Note that this technique of using a window of three symbols, namely the previous, current and next symbol is a method widely adopted in asynchronous CDMA multiuser detection [41].

References

- [1] J.H. Winters, J. Salz, R.D. Gitlin, The impact of antenna diversity on the capacity of wireless communication systems, *IEEE Trans. Commun.* 42 (2) (1994) 1740–1751.
- [2] M.C. Wells, Increasing the capacity of GSM cellular radio using adaptive antennas, *IEE Proc. Commun.* 143 (5) (1996) 304–310.
- [3] J. Litva, T.K.Y. Lo, *Digital Beamforming in Wireless Communications*, Artech House, London, 1996.
- [4] L.C. Godara, Applications of antenna arrays to mobile communications, Part I: Performance improvement, feasibility, and system considerations, *Proc. IEEE* 85 (7) (1997) 1031–1060.
- [5] L.C. Godara, Applications of antenna arrays to mobile communications, Part II: Beam-forming and direction-of-arrival considerations, *Proc. IEEE* 85 (8) (1997) 1193–1245.
- [6] A.J. Paulraj, C.B. Papadias, Space-time processing for wireless communications, *IEEE Signal Process. Mag.* 14 (6) (1997) 49–83.
- [7] J.H. Winters, Smart antennas for wireless systems, *IEEE Personal Commun.* 5 (1) (1998) 23–27.
- [8] R. Kohno, Spatial and temporal communication theory using adaptive antenna array, *IEEE Personal Commun.* 5 (1) (1998) 28–35.

- [9] P. Petrus, R.B. Ertel, J.H. Reed, Capacity enhancement using adaptive arrays in an AMPS system, *IEEE Trans. Vehicular Technol.* 47 (3) (1998) 717–727.
- [10] G.V. Tsoulos, Smart antennas for mobile communication systems: benefits and challenges, *IEE Electron. Commun. J.* 11 (2) (1999) 84–94.
- [11] J.S. Blogh, L. Hanzo, *Third Generation Systems and Intelligent Wireless Networking—Smart Antennas and Adaptive Modulation*, Wiley, Chichester, 2002.
- [12] R.A. Soni, R.M. Buehrer, R.D. Benning, Intelligent antenna system for cdma2000, *IEEE Signal Process. Mag.* 19 (4) (2002) 54–67.
- [13] I.S. Reed, J.D. Mallett, L.E. Brennan, Rapid convergence rate in adaptive arrays, *IEEE Trans. Aerospace Electron. Syst.* AES-10 (1974) 853–863.
- [14] M.W. Ganz, R.L. Moses, S.L. Wilson, Convergence of the SMI and the diagonally loaded SMI algorithms with weak interference (adaptive array), *IEEE Trans. Antennas Propagation* 38 (3) (1990) 394–399.
- [15] B. Widrow, P.E. Mantey, L.J. Griffiths, B.B. Goode, Adaptive antenna systems, *Proc. IEEE* 55 (1967) 2143–2159.
- [16] L.J. Griffiths, A simple adaptive algorithm for real-time processing in antenna arrays, *Proc. IEEE* 57 (1969) 1696–1704.
- [17] S. Haykin, *Adaptive Filter Theory*, third ed., Prentice Hall, Upper Saddle River, NJ, 1996.
- [18] S. Chen, B. Mulgrew, E.S. Cheng, G. Gibson, Space translation properties and the minimum-BER linear-combiner DFE, *IEE Proc. Commun.* 145 (5) (1998) 316–322.
- [19] I.N. Psaromiligkos, S.N. Batalama, D.A. Pados, On adaptive minimum probability of error linear filter receivers for DS-CDMA channels, *IEEE Trans. Commun.* 47 (7) (1999) 1092–1102.
- [20] C.C. Yeh, J.R. Barry, Adaptive minimum bit-error rate equalization for binary signaling, *IEEE Trans. Commun.* 48 (7) (2000) 1226–1235.
- [21] B. Mulgrew, S. Chen, Adaptive minimum-BER decision feedback equalisers for binary signalling, *Signal Process.* 81 (7) (2001) 1479–1489.
- [22] S. Chen, A.K. Samingan, B. Mulgrew, L. Hanzo, Adaptive minimum-BER linear multiuser detection for DS-CDMA signals in multipath channels, *IEEE Trans. Signal Process.* 49 (6) (2001) 1240–1247.
- [23] S. Chen, L. Hanzo, N.N. Ahmad, Adaptive minimum bit error rate beamforming assisted receiver for wireless communications, in: *Proc. ICASSP 2003*, vol. IV, Hong Kong, China, April 2003, pp. 640–643.
- [24] E. Parzen, On estimation of a probability density function and mode, *Ann. Math. Statist.* 33 (1962) 1066–1076.
- [25] B.W. Silverman, *Density Estimation*, Chapman & Hall, London, 1996.
- [26] A.W. Bowman, A. Azzalini, *Applied Smoothing Techniques for Data Analysis*, Oxford Univ. Press, Oxford, 1997.
- [27] M.S. Bazaraa, H.D. Sherali, C.M. Shetty, *Nonlinear Programming: Theory and Algorithms*, Wiley, New York, 1993.
- [28] R. Steele, L. Hanzo, *Mobile Radio Communications*, IEEE Press, Piscataway, NJ, 1999.
- [29] S.-H. Hwang, L. Hanzo, Reverse-link performance of synchronous DS-CDMA systems in dispersive Rician multipath fading channels, *IEE Electron. Lett.* 39 (23) (2003) 1682–1684.
- [30] K. Abend, B.D. Fritchman, Statistical detection for communication channels with intersymbol interference, *Proc. IEEE* 58 (5) (1970) 779–785.
- [31] S. Chen, B. Mulgrew, P.M. Grant, A clustering technique for digital communications channel equalisation using radial basis function networks, *IEEE Trans. Neural Networks* 4 (4) (1993) 570–579.
- [32] S. Chen, B. Mulgrew, S. McLaughlin, Adaptive Bayesian equaliser with decision feedback, *IEEE Trans. Signal Process.* 41 (9) (1993) 2918–2927.
- [33] S. Chen, S. McLaughlin, B. Mulgrew, P.M. Grant, Adaptive Bayesian decision feedback equaliser for dispersive mobile radio channels, *IEEE Trans. Commun.* 43 (1995) 1937–1946.
- [34] S. Chen, S. McLaughlin, B. Mulgrew, P.M. Grant, Bayesian decision feedback equaliser for overcoming co-channel interference, *IEE Proc. Commun.* 143 (4) (1996) 219–225.
- [35] S. Chen, A.K. Samingan, L. Hanzo, Support vector machine multiuser receiver for DS-CDMA signals in multipath channels, *IEEE Trans. Neural Networks* 12 (3) (2001) 604–611.
- [36] D.E. Goldberg, *Genetic Algorithms in Search, Optimization and Machine Learning*, Addison–Wesley, Reading, MA, 1989.
- [37] K.F. Man, K.S. Tang, S. Kwong, *Genetic Algorithms: Concepts and Design*, Springer-Verlag, London, 1998.

- [38] L. Ingber, Simulated annealing: practice versus theory, *Math. Comput. Modeling* 18 (11) (1993) 29–57.
- [39] S. Chen, B.L. Luk, Adaptive simulated annealing for optimization in signal processing applications, *Signal Process.* 79 (1) (1999) 117–128.
- [40] S. Chen, L. Hanzo, A. Wolfgang, Kernel-based nonlinear beamforming construction using orthogonal forward selection with Fisher ratio class separability measure, *IEEE Signal Process. Lett.* 11 (5) (2004) 478–481.
- [41] L. Hanzo, L.-L. Yang, E.-L. Kuan, K. Yen, *Single and Multi-Carrier DS-CDMA: Multi-User Detection, Space-Time Spreading, Synchronisation, Networking and Standards*, Wiley, Chichester, UK, 2003.