A Robust Nonlinear Beamforming Assisted Receiver for BPSK Signalling

S. Chen, A. Wolfgang and L. Hanzo

School of ECS, Univ. of Southampton, SO17 1BJ, U.K; lh@ecs.soton.ac.uk

ABSTRACT

Nonlinear beamforming designed for wireless communications is investigated. We derive the *optimal nonlinear beamforming assisted receiver* **designed for binary phase shift keying (BPSK) signalling. It is shown that this optimal Bayesian beamformer significantly outperforms the classic linear minimum mean square error (LMMSE) beamformer at the expense of an increased complexity. Hence the achievable user capacity of the wireless system invoking the proposed beamformer is substantially enhanced. In particular, when the angular separation between the desired and interfering signals is below a certain threshold, a linear beamformer will fail while a nonlinear beamformer can still perform adequately. Block-adaptive implementation of the optimal Bayesian beamformer can be realized using a Radial Basis Function network based on the Relevance Vector Machine (RVM) for classification, and a recursive sample-by-sample adaptation is proposed based on an enhanced κ-means clustering aided recursive least squares algorithm.**

I. INTRODUCTION

Multiple antenna assisted adaptive beamformers have the potential of substantially increasing the user capacity of wireless systems [1]-[3]. Adaptive beamforming based spatial signal processing is classically carried out by forming a linear combination of the signals received from the different elements of an antenna array. We will refer to this classic beamforming approach as *linear* beamforming. Recent work [4],[5] has investigated a linear beamforming technique directly minimizing the system's bit error rate (BER) and developed an adaptive algorithm for realizing the linear minimum BER (LMBER) beamformer. The results of [4],[5] have demonstrated that LMBER beamforming is capable of providing considerable performance gains in terms of a reduced BER in comparison to linear minimum mean square error (LMMSE) beamforming.

The angle of arrival (AOA) separation between the desired signal and the closest (in terms of its AOA) interfering signal dominates the system's BER performance. When this AOA separation is below a certain threshold, linear beamforming results in a high residual BER, since the interference-contaminated phasor constellation becomes linearly nonseparable [6],[7]. Even if it remains linearly separable, the phasor constellation points may be close to the decision boundary and hence nonlinear receivers [12] are typically capable of providing a better performance than their linear counterparts, although at the cost of an increased complexity. These considerations motivate this study on a nonlinear beamforming technique. We derive the optimal nonlinear beamformer, which is referred to as the Bayesian beamforming solution. It is shown that this Bayesian solution has an identical form to a radial basis function (RBF) network. A block-data based adaptive RBF beamformer is proposed, which employs the relevance vector machine (RVM) principle for classification [8],[9]. For adaptive sample-by-sample based weight adaptation an enhanced κ-means clustering technique and the recursive least squares (CRLS) algorithm [10],[11] are considered.

II. SYSTEM MODEL

It is assumed that the system supports M users, and each user transmits a binary phase shift keying (BPSK) signal on the same carrier frequency $\omega = 2\pi f$. The baseband signal of user i is given by

$$m_i(k) = A_i b_i(k), \ b_i(k) \in \{\pm 1\}, \ 1 \leq i \leq M, \quad (1)$$

where $|A_i|^2$ denotes the received signal power of user i and $b_i(k)$ represents the k-th bit of user i. Source 1 is assumed to be the desired user and the rest of the sources represent the interfering users. A linear antenna array is considered [1], which consists of L uniformly spaced antenna elements, and the signals arriving at the L-element antenna array are given by

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

for $1 \leq l \leq L$, where $t_l(\theta_i)$ is the relative time delay at element l for source i, θ_i is the AOA for source i, and $n_l(k)$ is the complex-valued additive white Gaussian noise (AWGN) having zero mean and variance of $E[|n_l(k)|^2] = 2\sigma_n^2$. The desired user's signal to noise ratio is defined as SNR$= A_1^2/2\sigma_n^2$, and his signal to interferer i ratio is given by SIR$_i = A_1^2/A_i^2$ for $i = 2, \cdots M$. The array input can be written in vectorial form as

$$\mathbf{x}(k) = [x_1(k) \cdots x_L(k)]^T = \bar{\mathbf{x}}(k) + \mathbf{n}(k) = \mathbf{Pb}(k) + \mathbf{n}(k) \quad (3)$$

where $E[\mathbf{n}(k)\mathbf{n}^H(k)] = 2\sigma_n^2\mathbf{I}_L$ with $\mathbf{I}_L$ denoting the $L \times L$-dimansional identity matrix, while the system matrix is $\mathbf{P} = [A_1\mathbf{s}_1 \; A_2\mathbf{s}_2 \cdots A_M\mathbf{s}_M]$, the steering vector of source i is given by

$$\mathbf{s}_i = [\exp(j\omega t_1(\theta_i)) \; \exp(j\omega t_2(\theta_i)) \cdots \exp(j\omega t_L(\theta_i))]^T \tag{4}$$

and the bit vector by $\mathbf{b}(k) = [b_1(k) \cdots b_M(k)]^T$.

III. LINEAR BEAMFORMER

The output of the linear beamformer is given by

$$y(k) = \mathbf{w}^H\mathbf{x}(k) = \mathbf{w}^H\bar{\mathbf{x}}(k) + \mathbf{w}^H\mathbf{n}(k) = \bar{y}(k) + e(k) \tag{5}$$

where $\mathbf{w}$ is the complex-valued beamformer weight vector, and $e(k)$ is Gaussian distributed with zero mean and variance of $E[|e(k)|^2] = 2\sigma_n^2\mathbf{w}^H\mathbf{w}$. The decision regarding the transmitted bit $b_1(k)$ is made according to: $\hat{b}_1(k) = +1$ if $y_R(k) \geq 0$ and $\hat{b}_1(k) = -1$ if $y_R(k) < 0$, where we have $y_R(k) = \Re[y(k)]$. The classic LMMSE beamforming solution is given by $\mathbf{w}_{\text{MMSE}} = (\mathbf{P}\mathbf{P}^H + 2\sigma_n^2\mathbf{I}_L)^{-1}\mathbf{p}_1$, with $\mathbf{p}_1$ being the first column of $\mathbf{P}$. The LMMSE beamforming solution can be realized adaptively using the recursive least squares (RLS) algorithm [12]. Recently we have developed the LMBER beamforming solution and its adaptive realization based on a stochastic-gradient technique referred to as the least bit error rate (LBER) algorithm [4],[5].

For the linear beamformer to operate adequately, the phasor constellation at the channel's output must be linearly separable. Denote the $N_b = 2^M$ possible M-user phasor combinations $\mathbf{b}(k)$ 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 $b_{q,1}$. The noiseless channel output at the array input $\bar{\mathbf{x}}(k)$ only takes values from the signal state set of

$$\mathcal{X} \triangleq \{\bar{\mathbf{x}}_q = \mathbf{P}\mathbf{b}_q, \; 1 \leq q \leq N_b\}. \tag{6}$$

Similarly, $\bar{y}(k)$ takes values from the scalar set $\mathcal{Y} \triangleq \{\bar{y}_q = \mathbf{w}^H\bar{\mathbf{x}}_q, 1 \leq q \leq N_b\}$. Thus, $\bar{y}_R(k)$ only takes values from the set $\mathcal{Y}_R \triangleq \{\bar{y}_{R,q} = \Re[\bar{y}_q], \; 1 \leq q \leq N_b\}$, which can be divided into the two subsets conditioned on the value of $b_1(k)$

$$\mathcal{Y}_R^{(\pm)} \triangleq \{\bar{y}_{R,q}^{(\pm)} \in \mathcal{Y}_R : b_1(k) = \pm 1\}. \tag{7}$$

The linear separability implies that there exists a weight vector $\mathbf{w}$ such that $\mathcal{Y}_R^{(-)}$ and $\mathcal{Y}_R^{(+)}$ are separable using a linear decision boundary. When the minimum AOA separation between the desired user and interfering users is below a certain threshold, the channel's output phasor points inevitably become linearly nonseparable - even in the absence of noise. In such a situation, the linear bermformer will have a high irreducible BER floor, and a nonlinear receiver has to be adopted for achieving an adequate performance. We will show that a nonlinear beamformer is capable achieving a better BER performance than a linear one, regardless, whether the system is linearly separable or not.

IV. BAYESIAN BEAMFORMING SOLUTION

Given the observation vector $\mathbf{x}(k)$, the optimal solution to the beamforming problem considered is the maximum *a posteriori* probability solution, which we derive below. The state set $\mathcal{X}$ can be divided into two subsets conditioned on $b_1(k)$:

$$\mathcal{X}^{(\pm)} \triangleq \{\bar{\mathbf{x}}_q^{(\pm)} \in \mathcal{X}, 1 \leq q \leq N_{sb} : b_1(k) = \pm 1\} \tag{8}$$

where $N_{sb} = N_b/2$. The *a posteriori* probabilities or decision variables for $b_1(k) = \pm 1$ given $\mathbf{x}(k)$ are formulated as

$$\eta^{(\pm)}(k) = \sum_{q=1}^{N_{sb}} \frac{\xi_q^{(\pm)}}{(2\pi\sigma_n^2)^L} \exp\left(-\frac{\|\mathbf{x}(k) - \bar{\mathbf{x}}_q^{(\pm)}\|^2}{2\sigma_n^2}\right) \tag{9}$$

where $\xi_q^{(\pm)}$ are the *a priori* probabilities of $\bar{\mathbf{x}}_q^{(\pm)}$. Assume that all values of $\bar{\mathbf{x}}_q^{(\pm)}$ are equiprobable, yielding $\xi_q^{(\pm)} = \frac{1}{N_b}$. The optimal decision is given by: $\hat{b}_1(k) = +1$ if $\eta^{(+)}(k) \geq \eta^{(-)}(k)$, and $\hat{b}_1(k) = -1$ otherwise. Let us now redefine a single decision variable as

$$y_B(k) = \sum_{q=1}^{N_b} \alpha_q \exp\left(-\frac{\|\mathbf{x}(k) - \bar{\mathbf{x}}_q\|^2}{2\sigma_n^2}\right) \tag{10}$$

where $\alpha_q = \text{sgn}(b_{q,1})/(N_b (2\pi\sigma_n^2)^L)$. Then the optimal decision is equivalent to: $\hat{b}_1(k) = +1$ for $y_B(k) \geq 0$ and $\hat{b}_1(k) = -1$ for $y_B(k) < 0$. Note that (10) characterises an RBF network associated with a Gaussian kernel function [12].

V. BLOCK ADAPTIVE RBF AND RVM AIDED BEAMFORMER

Given a block of N training data $\{\mathbf{x}(k), b_1(k)\}_{k=1}^N$, consider the nonlinear beamformer of the form

$$y(\mathbf{x}) = \sum_{l=1}^N \beta_l K_l(\mathbf{x}) \tag{11}$$

where β_l represents the beamformer weights and $K_l(\mathbf{x}) = K(\mathbf{x}, \mathbf{x}(l))$ the RBF kernel basis functions [12]. In our beamforming application, $K(\bullet, \bullet)$ can be chosen as the Gaussian kernel function of the form [12]

$$K(\mathbf{x}, \mathbf{x}(l)) = \exp\left(-\frac{\|\mathbf{x} - \mathbf{x}(l)\|^2}{2\rho^2}\right) \tag{12}$$

where ρ^2 is an estimate of the noise variance σ_n^2. Our aim is to minimize the number of kernel functions to be used by the beamformer for the sake of reducing its complexity. Hence the RVM method [8],[9] can be invoked for constructing a so-called sparse beamformer having a reduced number of N_{spa} terms. Denote $\mathbf{t} = [t_1 \cdots t_N]^T = [b_1(1) \cdots b_1(N)]^T$ and

$\boldsymbol{\beta} = [\beta_1 \cdots \beta_N]^T$. The *a posteriori* probability of $\boldsymbol{\beta}$ is given by [8]

$$p(\boldsymbol{\beta}|\mathbf{t}, \boldsymbol{\alpha}) = \frac{p(\mathbf{t}|\boldsymbol{\beta}, \boldsymbol{\alpha})p(\boldsymbol{\beta}|\boldsymbol{\alpha})}{p(\mathbf{t}|\boldsymbol{\alpha})} \tag{13}$$

where $p(\boldsymbol{\beta}|\boldsymbol{\alpha})$ is the *a priori* probability of $\boldsymbol{\beta}$, with $\boldsymbol{\alpha} = [\alpha_1 \cdots \alpha_N]^T$ denoting the vector of hyperparameters, while $p(\mathbf{t}|\boldsymbol{\beta}, \boldsymbol{\alpha})$ is the so-called *likelihood* and $p(\mathbf{t}|\boldsymbol{\alpha})$ is the *evidence*. According to the Bayesian classification principles [8],[9] the *likelihood* is expressed as

$$p(\mathbf{t}|\boldsymbol{\beta}, \boldsymbol{\alpha}) = \prod_{l=1}^{N} (f(y(\mathbf{x}(l))))^{(t_l+1)/2} (1 - f(y(\mathbf{x}(l))))^{(1-t_l)/2} \tag{14}$$

where $f(y) = 1/(1 + \exp(-y))$. Assuming that the *a priori* probability is Gaussian distributed, we have

$$p(\boldsymbol{\beta}|\boldsymbol{\alpha}) = \prod_{l=1}^{N} \frac{\sqrt{\alpha_l}}{\sqrt{2\pi}} \exp\left(-\frac{\alpha_l \beta_l^2}{2}\right). \tag{15}$$

With a fixed given $\boldsymbol{\alpha}$, the maximum *a posteriori* probability (MAP) solution $\hat{\boldsymbol{\beta}}$ can be obtained by maximizing $\log(p(\boldsymbol{\beta}|\mathbf{t}, \boldsymbol{\alpha}))$ or, equivalently, by minimizing the following cost function

$$J(\boldsymbol{\beta}|\mathbf{t}, \boldsymbol{\alpha}) = \sum_{l=1}^{N} \left(\frac{\alpha_l \beta_l^2}{2} - \frac{t_l+1}{2} \log(f(y(\mathbf{x}(l))))\right.$$
$$\left. - \frac{1-t_l}{2} \log(1 - f(y(\mathbf{x}(l))))\right). \tag{16}$$

It may be readily shown that the gradient of J with respect to $\boldsymbol{\beta}$ is

$$\nabla J = \mathbf{A}\boldsymbol{\beta} + \boldsymbol{\Phi}^T \left(\mathbf{f} - \frac{1}{2}(\mathbf{t} + \mathbf{1}_N)\right), \tag{17}$$

where $\mathbf{A} = \text{diag}\{\alpha_1, \cdots, \alpha_N\}, \mathbf{f} = [f(y(\mathbf{x}(1))) \cdots f(y(\mathbf{x}(N)))]^T$, $\mathbf{1}_N = [1 \cdots 1]^T$ and the matrix $\boldsymbol{\Phi}$ has elements $\phi_{i,j} = K(\mathbf{x}(i), \mathbf{x}(j))$. The Hessian of J is given by

$$\mathbf{H} = \nabla^2 J = \boldsymbol{\Phi}^T \mathbf{B}\boldsymbol{\Phi} + \mathbf{A} \tag{18}$$

where $\mathbf{B} = \text{diag}\{f(y(\mathbf{x}(1)))(1-f(y(\mathbf{x}(1)))), \cdots, f(y(\mathbf{x}(N)))(1 - f(y(\mathbf{x}(N))))\}$. The hyperparameters $\boldsymbol{\alpha}$ are updated using

$$\alpha_i^{\text{new}} = \frac{1 - \alpha_i^{\text{old}}\gamma_{i,i}}{\hat{\beta}_i^2}, \tag{19}$$

with $\gamma_{i,i}$ being the diagonal elements of $\boldsymbol{\Gamma}$, which is defined by

$$\boldsymbol{\Gamma} = \left(\mathbf{H}|_{\hat{\boldsymbol{\beta}}}\right)^{-1}. \tag{20}$$

During the optimization process, many of the hyperparameters α_i are driven to large values, so that the corresponding model weights β_i are effectively pruned out [8] and the related model terms $K_i(\bullet)$ are removed from the trained model. The following iterative procedure can be adopted for constructing a sparse RVM beamformer:

Initialization. The $(N \times N_{\text{spa}})$-dimensional kernel matrix $\boldsymbol{\Phi}$ is initialized with $N_{\text{spa}} = N$. Each weight β_i is initially associated with the same value of the hyperparameter α_i.

Step 1. Given the current value $\boldsymbol{\alpha}$, find $\hat{\boldsymbol{\beta}}$ by minimizing the cost function (16) using a simplified conjugate gradient algorithm [13].

Step 2. The hyperparameters are updated using (19). If a $\alpha_i > L_g$, where L_g is a preset large positive value, we set $N_{\text{spa}} := N_{\text{spa}} - 1$, the corresponding column of $\boldsymbol{\Phi}$ is removed, and thus the weight β_i and model term $K_i(\bullet)$ are pruned out from the model.

Test. If $\boldsymbol{\alpha}$ remains sufficiently unchanged in two successive iterations (no removal of hyperparameters) or a pre-set maximum number of iterations is reached, stop; otherwise go to *Step 1.*

VI. RECURSIVE ADAPTIVE RBF BEAMFORMER USING THE CLUSTERING AND RLS ALGORITHM

Consider again the RBF beamformer of the form:

$$y(\mathbf{x}(k)) = \sum_{i=1}^{N_c} \beta_i K(\mathbf{x}(k), \mathbf{c}_i), \tag{21}$$

where $\mathbf{c}_i$ represents the kernel centers and the number of kernel centers N_c is assumed to be given. For a recursive sample-by-sample adaptation of this RBF beamformer, the enhanced κ-means clustering and RLS algorithm [10],[11] can be used.

The enhanced κ-means clustering algorithm [14], which recursively updates the RBF centers, is described by:

$$\mathbf{c}_i(k) = \mathbf{c}_i(k-1) + \mathcal{M}_i(\mathbf{x}(k)) (g_c (\mathbf{x}(k) - \mathbf{c}_i(k-1))) \tag{22}$$

for $1 \leq i \leq N_c$, where $0 < g_c < 1.0$ define the learning rate, the membership function $\mathcal{M}_i(\mathbf{x}(k))$ is defined as

$$\mathcal{M}_i(\mathbf{x}) = \begin{cases} 1, & \text{if } v_i\|\mathbf{x} - \mathbf{c}_i\|^2 \leq v_l\|\mathbf{x} - \mathbf{c}_l\|^2 \text{ for all } l \neq i, \\ 0, & \text{otherwise,} \end{cases} \tag{23}$$

and v_i is the variation of the i-th cluster. In order to estimate the variation v_i, the following updating rule is used:

$$v_i(k) = g_v v_i(k-1) + (1-g_v) \left(\mathcal{M}_i(\mathbf{x}(k))\|\mathbf{x}(k) - \mathbf{c}_i(k-1)\|^2\right), \tag{24}$$

where g_v is a constant slightly less than 1.0. The initial variations, $v_i(0), 1 \leq i \leq N_c$, are set to the same small number. The learning rate g_c can either be set to a fixed small positive number or be self-adjusting based on an entropy formula [14]. The RBF weights β_i are updated using the usual RLS algorithm [12].

The above clustering algorithm is an unsupervised one. To take the full advantage of training, it can be reformulated as
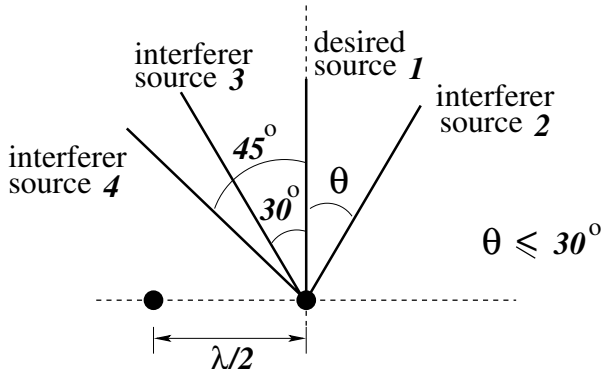
Fig. 1. Locations of the desired source and the interfering sources with respect to the two-element linear antenna array having $\lambda/2$ element spacing, where λ is the wavelength.

a supervised one. Let the RBF center set be divided into the two subsets: $\mathcal{C}^{(+)} = \{\mathbf{c}_i, \ 1 \leq i \leq N_c/2\}$ and $\mathcal{C}^{(-)} = \{\mathbf{c}_i, \ 1 + N_c/2 \leq i \leq N_c\}$, corresponding to the two classes $b_1(k) = \pm 1$. During the training instance k, if $b_1(k) = +1$, the enhanced κ-means clustering algorithm is applied only to the center subset $\mathcal{C}^{(+)}$; otherwise, if $b_1(k) = -1$, it is applied to $\mathcal{C}^{(-)}$. This supervised clustering is more efficient in dealing with linearly nonseparable cases.

VII. A SIMULATION EXAMPLE

In the example considered four users were supported with the aid of a two-element antenna array. Fig. 1 shows the locations of the desired source and the three interfering sources graphically. The desired user and all the three interfering users had equal signal power. The minimum AOA separation observed between the desired user and interferer 2, which was $\theta \leq 30°$. Fig. 2 compares the BERs of the LMMSE, LMBER and Bayesian beamformers for $\theta = 30°$. It can be seen from Fig. 2 that the LMMSE beamformer was unable to achieve linear separability and hence exhibited a high BER floor. By contrast, the LMBER beamformer achieved linear separability and had a better BER performance than the LMMSE beamformer. The Bayesian beamformer provided the optimal performance and had a 4 dB SNR gain at a BER of 10^{-3}, compared to the LMBER beamformer. When the AOA separation was reduced to $\theta = 10°$, the system became linearly nonseparable, as seen in Fig 3. While the linear beamformer failed in this linearly nonseparable situation, the Bayesian beamformer still performed adequately, and this is demonstrated clearly in Fig. 3. For this example, the number of states that defined the Bayesian beamformer was $N_b = 16$.

At an AOA of $\theta = 30°$, the RVM algorithm was used for constructing a RBF beamformer. The number of training data used at each SNR was $N = 160$, and the number of RBF centers identified by the algorithm ranged from $N_{\mathrm{spa}} = 14$ to 20, with an average value of $N_{\mathrm{spa}} = 18$. Recursive sample-by-sample adaptation of the RBF beamformer given $N_c = 16$ was also performed using the enhanced κ-means clustering and RLS algorithm, where the training data length was $N = 1000$ with $g_c = 0.2$ and $g_v = 0.995$. Fig. 4 de-

picts the BERs of both the RVM RBF beamformer and that of the CRLS RBF beamformer in comparison to the optimal Bayesian performance. For the CRLS RBF beamformer, the results obtained using both the unsupervised and supervised clustering algorithms were similar, and the unsupervised clustering results are shown in Fig. 4.

For the AOA of $\theta = 10°$, the BERs of the two adaptive beamformers, the RVM RBF beamformer and the CRLS RBF beamformer, are compared to the optimal Bayesian performance in Fig. 5. The training data lengths and adaptive algorithm parameters used were identical to the case of $\theta = 30°$. For the CRLS RBF beamformer, the result shown was obtained using the supervised clustering algorithm. In this linearly inseparable case, it was observed that the supervised clustering performed better than the unsupervised one.

VIII. CONCLUSIONS

A robust nonlinear beamforming technique has been designed for wireless communications. The optimal solution of the nonlinear beamforming assisted receiver has been derived for BPSK signalling. It has been shown that this optimal Bayesian beamformer significantly outperforms the family of classic linear beamformers, although at the expense of an increased complexity. This suggests that significant wireless system capacity enhancements may be achieved by employing nonlinear, rather than linear beamforming. Both block-based and sample-by-sample adaptive implementations of the optimal Bayesian beamformer have been considered using a RBF network based on the RVM algorithm as well as on the enhanced κ-means clustering and RLS algorithm, respectively.

REFERENCES

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

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

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

[4] S. Chen, L. Hanzo and N.N. Ahmad, "Adaptive minimum bit error rate beamforming assisted receiver for wireless communications," in *Proc. ICASSP-2003*, April 6-10, 2003, pp.640–643.

[5] S. Chen, N.N. Ahmad and L. Hanzo, "Adaptive minimum bit error rate beamforming," submitted to *IEEE Trans. Wireless Communications*, 2002.

[6] S. Chen, B. Mulgrew and P.M. Grant, "A clustering technique for digital communications channel equalisation using radial basis function networks," *IEEE Trans. Neural Networks*, Vol.4, No.4, pp.570–579, 1993.
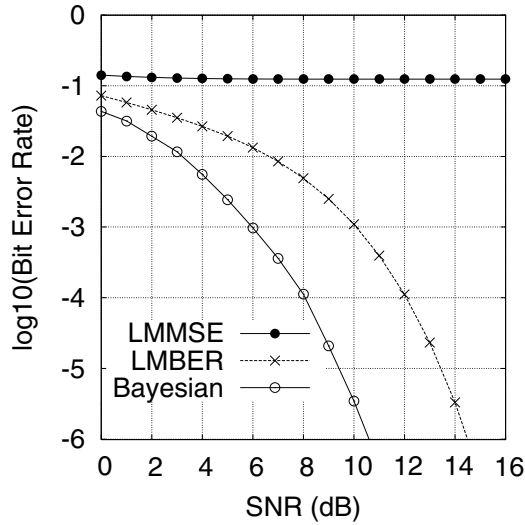
[7] S. Chen, B. Mulgrew and S. McLaughlin, "Adap-

Fig. 2. Comparison of the bit error rates of three beamformers for $\theta = 30°$.
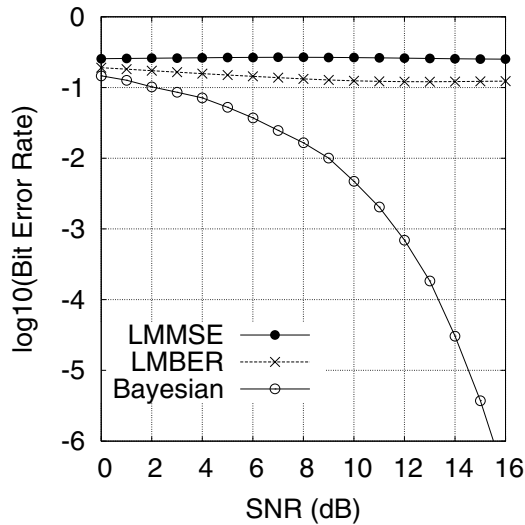

Fig. 3. Comparison of the bit error rates of three beamformers for $\theta = 10°$.
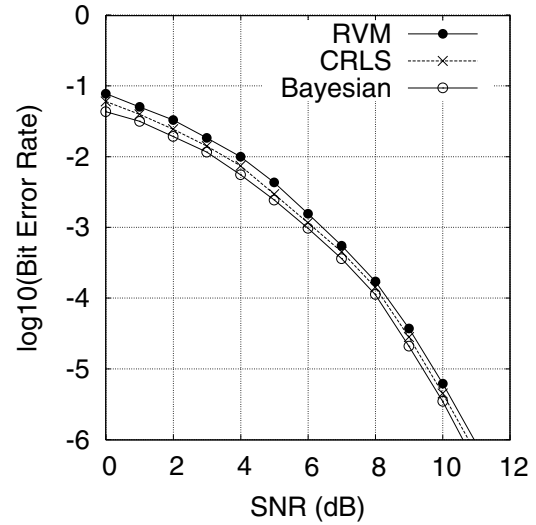

Fig. 4. Performance comparison of the Bayesian beamformer, of the RBF beamformer based on the block-adaptive RVM algorithm as well as of the recursive clustering and RLS algorithm for $\theta = 30°$. The block-length of the RVM algorithm was $N = 160$, while the training length of the recursive supervised clustering and RLS algorithm was $N = 1000$.
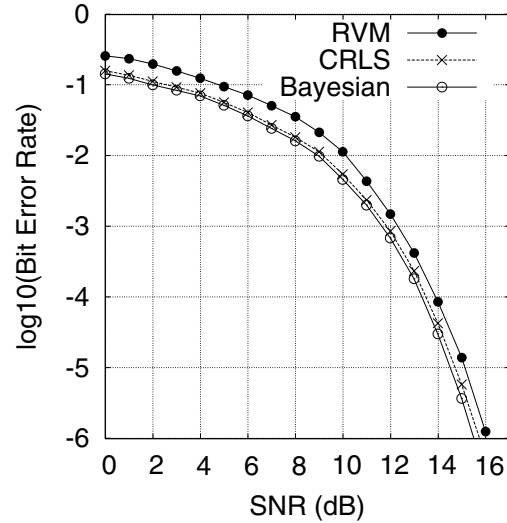

Fig. 5. Performance comparison of the Bayesian beamformer, of the RBF beamformer based on the block-adaptive RVM algorithm as well as of the recursive clustering and RLS algorithm for $\theta = 10°$. The block-length of the RVM algorithm was $N = 160$, while the training length of the recursive supervised clustering and RLS algorithm was $N = 1000$.

tive Bayesian equaliser with decision feedback," *IEEE Trans. Signal Processing*, Vol.41, No.9, pp.2918–2927, 1993.

[8] M.E. Tipping, "Sparse Bayesian learning and the relevance vector machine," *J. Machine Learning Research*, Vol.1, pp.211–244, 2001.

[9] S. Chen, S.R. Gunn and C.J. Harris, "The relevance vector machine technique for channel equalization application," *IEEE Trans. Neural Networks*, Vol.12, No.6, pp.1529–1532, 2001.

[10] S. Chen, "Nonlinear time series modelling and prediction using Gaussian RBF networks with enhanced clustering and RLS learning," *Electronics Letters*, Vol.31, No.2, pp.117–118, 1995.

[11] S. Chen, A.K. Samingan and L. Hanzo, "Support vector machine multiuser receiver for DS-CDMA signals in multipath channels," *IEEE Trans. Neural Networks*, Vol.12, No.3, pp.604-611, 2001.

[12] L. Hanzo, C.H. Wong, M.S. Yee: Adaptive Wireless Transceivers: Turbo-Coded, Turbo-Equalised and Space-Time Coded TDMA, CDMA and OFDM Systems, John Wiley, March 2002, ISBN 0-470-84689-5 **752 pages**.

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

[14] C. Chinrungrueng and C.H. Séquin, "Optimal adaptive κ-means algorithm with dynamic adjustment of learning rate," *IEEE Trans. Neural Networks*, Vol.6, No.1, pp.157–169, 1995.