methodology proposed in this brief can be applied to a wide range of other neural network architectures.

REFERENCES

[1] D. Nguyen and B. Widrow, "Neural networks for self-learning control system," *IEEE Control Syst. Mag.*, vol. 10, no. 3, pp. 18–23, Apr. 1990.

[2] L. Jain, *Recurrent Neural Networks*. Boca Raton, FL: CRC Press, 2000.

[3] D. V. Prokhorov, "Training recurrent neurocontrollers for real-time applications," *IEEE Trans. Neural Netw.*, vol. 18, no. 4, pp. 1003–1015, Jul. 2007.

[4] R. J. Williams and D. Zipser, "A learning algorithm for continually running fully recurrent neural networks," *Neural Comput.*, no. 1, pp. 270–280, 1989.

[5] D. Zipser, "A subgrouping strategy that reduces complexity and speeds up learning in recurrent networks," *Neural Comput.*, no. 1, pp. 552–558, 1989.

[6] N. Euliano and J. Principe, "Dynamic subgrouping in RTRL provides a faster $O(n^2)$ algorithm," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process.*, Istanbul, 2000, vol. 6, pp. 3418–3421.

[7] O. D. Jesus and M. T. Hagan, "Backpropagation algorithms for a broad class of dynamic networks," *IEEE Trans. Neural Netw.*, vol. 18, no. 1, pp. 14–27, Jan. 2007.

[8] J. Schmidhuber, "A fixed size storage $O(n^3)$ time complexity learning algorithm for fully recurrent continually running networks," *Neural Comput.*, vol. 4, pp. 243–248, 1992.

[9] G.-Z. Sun, H.-H. Chen, and Y.-C. Lee, "Green's function method for fast on-line learning algorithm of recurrent neural networks," in *Proc. Neural Inf. Process. Syst.*, 1991, pp. 333–340.

[10] V. S. Asirvadam, S. F. McLoone, and G. W. Irwin, "Memory efficient BFGS neural-network learning algorithms using MLP-network: A survey," in *Proc. IEEE Int. Conf. Control Appl.*, Sep. 2004, vol. 1, pp. 586–591.

[11] D. Budik and I. Elhanany, "TRTRL: a localized resource-efficient learning algorithm for recurrent neural networks," in *Proc. IEEE Midwest Symp. Circuits Syst.*, Puerto Rico, Aug. 2006, pp. 371–374.

[12] Z. Liu and I. Elhanany, "A scalable model-free recurrent neural network framework for solving POMDPs," in *Proc. IEEE Int. Symp. Approx. Dyn. Programm. Reinforcement Learn.*, Apr. 2007, pp. 119–126.

[13] S. E. Fahlman, "An empirical study of learning speed in back-propagation networks," Computer Sci. Tech. Rep., 1988.

[14] N. N. Schraudolph, "Local gain adaptation in stochastic gradient descent," in *Proc. 9th Int. Conf. Neural Netw.*, Sep. 1999, pp. 569–574.

[15] N. N. Schraudolph, J. Yu, and D. Aberdeen, "Fast online policy gradient learning with SMD gain vector adaptation," in *Proc. 19th Annu. Conf. Neural Inf. Process. Syst.*, Vancouver, BC, Canada, Dec. 2005, pp. 1185–1192.

[16] J. Kivinen and M. Warmuth, "Exponentiated gradient versus gradient descent for linear predictors," *Inf. Comput.*, no. 1, pp. 1–64, 1997.

[17] B. A. Pearlmutter, "Fast exact multiplication by the Hessian," *Neural Comput.*, vol. 6, no. 1, pp. 147–160, 1994.

[18] G. D. Magoulas, M. N. Vrahatis, and G. S. Androulakis, "Improving the convergence of the backpropagation algorithm using learning rate adaptation methods," *Neural Comput.*, vol. 11, no. 7, pp. 1769–1796, 1999.

[19] T. Tollenære, "Supersab: fast adaptive backpropagation with good scaling properties," *Neural Netw.*, vol. 3, no. 5, pp. 561–573, 1990.

[20] M. Mackey and L. Glass, "Oscillation and chaos in physiological control systems," *Science*, vol. 197, pp. 287–289, Jul. 1977.

[21] R. Croder, III, "Predicting the Mackey-Glass time-series with cascade-correlation learning," in *Connectionist Models Summer School Proceedings*, D. Touretzky, G. Hinton, and T. Sejnowski, Eds. Pittsburgh, PA: Carnegie Mellon Univ., 1990, pp. 117–123.

Symmetric Complex-Valued RBF Receiver for Multiple-Antenna-Aided Wireless Systems

Shenq Chen, Lajos Hanzo, and S. Tan

Abstract—A nonlinear beamforming assisted detector is proposed for multiple-antenna-aided wireless systems employing complex-valued quadrature phase shift-keying modulation. By exploiting the inherent symmetry of the optimal Bayesian detection solution, a novel complex-valued symmetric radial basis function (SRBF)-network-based detector is developed, which is capable of approaching the optimal Bayesian performance using channel-impaired training data. In the uplink case, adaptive nonlinear beamforming can be efficiently implemented by estimating the system's channel matrix based on the least squares channel estimate. Adaptive implementation of nonlinear beamforming in the downlink case by contrast is much more challenging, and we adopt a cluster-variation enhanced clustering algorithm to directly identify the SRBF center vectors required for realizing the optimal Bayesian detector. A simulation example is included to demonstrate the achievable performance improvement by the proposed adaptive nonlinear beamforming solution over the theoretical linear minimum bit error rate beamforming benchmark.

Index Terms—Beamforming, complex-valued radial basis function (RBF) network, multiple-antenna wireless system, symmetry, unsupervised clustering.

I. INTRODUCTION

The ever-increasing demand for an improved throughput in wireless communication has motivated the development of adaptive antenna-array-assisted spatial processing techniques [1]–[12] in order to further improve the achievable spectral efficiency. A specific technique that has shown real promise in achieving substantial capacity enhancements is constituted by adaptive beamforming. The concept of beamforming is traditionally defined as a linear spatial filtering. Upon appropriately combining the signals received by the antenna array linearly, adaptive beamforming is capable of separating user signals transmitted on the same carrier frequency, provided that the signal sources are sufficiently separated in the angular domain. The name beamforming comes from the classical interpretation of the beampattern of the linear spatial filter. The beampattern is basically the discrete Fourier transform (DFT) of the linear spatial filter's weights. The classical beamforming design, which has its root in other applications, such as radar and sonar, aims to ensure this DFT has a maximum response at the desired signal's direction and try to minimize its response at the interfering signals' directions. In order to create a maximum beam for the desired user and place nulls in the directions of the interfering users, it is necessary that the number of users supported is no more than the number of receive antenna elements. If this condition is not met, the system is referred to as rank-deficient or overloaded, and the classical beamforming will fail. Moreover, the classical beamforming is "zero forcing," which is well known to suffer from a serious noise enhancement problem. A better design for the linear beamformer is the minimum mean square error (L-MMSE) solution [1], [5], [7], [13], which trades off rejecting interference from amplifying noise. The optimal solution for the linear beamforming has been shown to be the minimum bit error rate (L-MBER) design [14], [15]. The L-MBER beamforming

outperforms the L-MMSE one, particularly in hostile rank-deficient scenarios.

It is well known that digital communication signal detection, in general, can be viewed as a classification problem [16]–[25], where the receiver detector simply classifies the received multidimensional channel-impaired signal into the most likely transmitted symbol constellation point or class. For the multiple-antenna-aided beamforming receiver, if one is willing to extend the beamforming process to nonlinear, substantial performance enhancement can be achieved over the linear beamforming at the cost of an increased complexity. The idea of nonlinear beamforming has recently been developed for wireless systems with the real-valued binary phase shift-keying modulation [26], [27], where a symmetric radial basis function (SRBF) network is proposed to adaptively implement the optimal nonlinear beamforming solution. This study extends nonlinear beamforming to wireless systems that employ the complex-valued quadrature phase-shift keying (QPSK) modulation scheme. For QPSK systems, the optimal Bayesian detection solution can be expressed as a complex-valued radial basis function (RBF) network [28], [29]. We further exploit the inherent symmetry of the optimal nonlinear beamforming solution and propose a novel complex-valued SRBF network for adaptively implementing the Bayesian beamforming solution. It is worth pointing out that our proposed nonlinear spatial filtering approach can be interpreted as a generalized beamforming. Instead of using the classical beampattern, which has a rather limited application, it is natural to interpret or to visualize the *a posterior* probability as the generalized "beampattern" of this nonlinear spatial filter. Thus, the optimal design is to maximize the *a posterior* probability for the desired user, which also implies to minimize the *a posterior* probabilities for the interfering users.

This contribution is organized as follows. In Section II, we present the QPSK beamforming signal model. Based on the system model of Section II, the optimal nonlinear beamforming solution is derived in Section III, where the inherent symmetric structure of the optimal Bayesian detection solution is discussed, while in Section IV, the novel complex-valued SRBF beamformer is presented and adaptive solutions are discussed for both uplink and downlink. For the uplink scenario, adaptive nonlinear beamforming can be realized effectively by estimating the system channel matrix using the least squares channel estimate (LSCE). For the downlink scenario, adaptive nonlinear beamforming is proposed by adopting an enhanced κ-means clustering algorithm [26], [30]. The achievable performance of this nonlinear beamforming approach is demonstrated in Section V, using a simulation example, where the two adaptive solutions are also investigated. In Section VI, we offer our conclusions.

II. MULTIPLE-ANTENNA-ASSISTED BEAMFORMING RECEIVER

Consider a coherent wireless communication system that supports M users, where each user transmits on the same angular carrier frequency of ω with a single transmit antenna. The receiver is equipped with a linear antenna array consisting of L uniformly spaced elements, in order to achieve user separation in the angular domain [7], [9]. Assume furthermore that the channel is nondispersive and hence it does not induce intersymbol interference. Then, the symbol-rate complex-valued received signal samples can be expressed as [1], [5]

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

for $1 \leq l \leq L$, where $t_l(\theta_i)$ is the relative time delay at array element l for source i, with θ_i being the direction of arrival for source i, $n_l(k)$ is the complex-valued Gaussian white noise with $E[|n_l(k)|^2] = 2\sigma_n^2$, $\bar{x}_l(k)$ denotes the noise-free part of the received signal, A_i is the complex-valued nondispersive channel coefficient of user i, and $b_i(k)$ is the

kth symbol of user i, which takes values from the QPSK symbol set, i.e.,

$$b_i(k) \in \Big\{ b^{[1]} = +1+j, \ b^{[2]} = -1+j,$$
$$b^{[3]} = -1-j, \ b^{[4]} = +1-j \Big\}. \quad (2)$$

Let source i be the desired user and the rest of the sources be the interfering users. The average signal-to-noise ratio (SNR) of the system is given by

$$\text{SNR} = \left(\frac{1}{M} \sum_{i=1}^{M} |A_i|^2 \right) \frac{\sigma_b^2}{2\sigma_n^2} \quad (3)$$

where σ_b^2 is the QPSK symbol energy, and the desired signal-to-interferer q ratio (SIR) is defined by $\text{SIR}_{i,q} = |A_i|^2/|A_q|^2$, for $q \neq i$. The received signal vector $\mathbf{x}(k) = [x_1(k) \ x_2(k) \cdots x_L(k)]^T$ can be expressed as

$$\mathbf{x}(k) = \mathbf{P}\mathbf{b}(k) + \mathbf{n}(k) = \bar{\mathbf{x}}(k) + \mathbf{n}(k) \quad (4)$$

where $\mathbf{n}(k) = [n_1(k) \ n_2(k) \cdots n_L(k)]^T$, the system's channel matrix $\mathbf{P}$ is given by

$$\mathbf{P} = [A_1\mathbf{s}_1 \ A_2\mathbf{s}_2 \cdots A_M\mathbf{s}_M] \quad (5)$$

with the steering vector of source i given by

$$\mathbf{s}_i = \left[e^{j\omega t_1(\theta_i)} \ e^{j\omega t_2(\theta_i)} \cdots e^{j\omega t_L(\theta_i)} \right]^T \quad (6)$$

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

Although a structure of uniformly spaced linear antenna array is assumed for beamforming, the results can be extended to other antenna array structures. In fact, the discussion is applicable to the generic multiple-input–multiple-output (MIMO) communication system [7], [9], where the (l,q)th element of the system matrix $\mathbf{P}$ represents the channel coefficient connecting the qth transmit antenna to the lth receive antenna. An implicit assumption for the signal model (4) is that all the users are symbol-synchronized. For the downlink scenario, synchronous transmission of the users is guaranteed. By contrast, in an uplink scenario, the different 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 (GSM) [31]. The GSM system has a timing-advance control accuracy of 0.25-bit duration. Since synchronous systems perform better than their asynchronous counterparts, the new generation systems, such as those proposed by the third generation partnership research consortium (3GPP) also consider the employment of timing-advance control.

Traditionally, a linear beamformer is adopted to detect the desired user's signal [1], [5]. The linear beamformer for user i is defined by $y_{\text{Lin}}(k) = \boldsymbol{\alpha}_i^H \mathbf{x}(k)$, where $\boldsymbol{\alpha}_i = [\alpha_{1,i} \ \alpha_{2,i} \cdots \alpha_{L,i}]^T$ is the complex-valued ith linear beamformer's weight vector. The decision regarding the transmitted symbol $b_i(k)$ is given by $\hat{b}_i(k) = \text{csgn}(y_{\text{Lin}}(k))$ with

$$\text{csgn}(y) = \begin{cases} b^{[1]} = +1+j, & y_R \geq 0 \text{ and } y_I \geq 0 \\ b^{[2]} = -1+j, & y_R < 0 \text{ and } y_I \geq 0 \\ b^{[3]} = -1-j, & y_R < 0 \text{ and } y_I < 0 \\ b^{[4]} = +1-j, & y_R \geq 0 \text{ and } y_I < 0 \end{cases} \quad (7)$$

where $y_R = \Re[y]$ and $y_I = \Im[y]$ denote the real and imaginary parts of y, respectively. The optimal weight vector designed for the linear beamformer is known to be the L-MBER solution [14], [15], which directly minimizes the BER of the linear beamformer. However, we will

show that the true optimal solution for the beamforming-aided detector is nonlinear.

III. OPTIMAL BAYESIAN BEAMFORMING SOLUTION

Denote the $N_b = 4^M$ legitimate combinations of $\mathbf{b}(k)$ as $\mathbf{b}_q$, $1 \leq q \leq N_b$. The noiseless channel output $\bar{\mathbf{x}}(k)$ takes values from the vector state set

$$\bar{\mathbf{x}}(k) \in \mathcal{X} \triangleq \{\bar{\mathbf{x}}_q = \mathbf{P}\mathbf{b}_q, \ 1 \leq q \leq N_b\}. \tag{8}$$

The signal state set $\mathcal{X}$ can be divided into the four subsets conditioned on the value of $b_i(k)$ as follows:

$$\mathcal{X}^{[m]} \triangleq \{\bar{\mathbf{x}}_q^{[m]} \in \mathcal{X}, 1 \leq q \leq N_{sb} : b_i(k) = b^{[m]}\} \tag{9}$$

for $1 \leq m \leq 4$, where the size of $\mathcal{X}^{[m]}$ is $N_{sb} = 4^{M-1}$. Denote the conditional probabilities of receiving $\mathbf{x}(k)$ given $b_i(k) = b^{[m]}$ as $p^{[m]}(\mathbf{x}(k)) = p(\mathbf{x}(k)|b_i(k) = b^{[m]})$. According to Bayes' decision theory [32], the optimal detection strategy is

$$\hat{b}_i(k) = b^{[m^*]} \tag{10}$$

where

$$m^* = \arg \max_{1 \leq m \leq 4} p^{[m]}(\mathbf{x}(k)). \tag{11}$$

If we introduce the following complex-valued Bayesian decision variable [29]:

$$y_{\mathrm{Bay}}(k) \triangleq \sum_{m=1}^{4} b^{[m]} \cdot p^{[m]}(\mathbf{x}(k)) \tag{12}$$

the optimal Bayesian detection rule (10) and (11) is equivalent to $\hat{b}_i(k) = \mathrm{csgn}(y_{\mathrm{Bay}}(k))$.

The conditional probability $p^{[m]}(\mathbf{x}(k))$ can be expressed as

$$p^{[m]}(\mathbf{x}(k)) = \sum_{q=1}^{N_{sb}} \beta_q e^{-\|\mathbf{x}(k) - \bar{\mathbf{x}}_q^{[m]}\|^2 / 2\sigma_n^2} \tag{13}$$

where $\bar{\mathbf{x}}_q^{[m]} \in \mathcal{X}^{[m]}$, and β_q is proportional to the *a priori* probability of $\bar{\mathbf{x}}_q^{[m]}$. Since all the $\bar{\mathbf{x}}_q^{[m]}$ are equiprobable, all the coefficients $\beta_q = 1/N_{sb}(2\pi\sigma_n^2)^L$. It can be seen from (13) that the optimal Bayesian decision variable (12) takes the structure of a complex-valued RBF network [28] with a Gaussian RBF function. The state subsets $\mathcal{X}^{[m]}$, $1 \leq m \leq 4$, are distributed symmetrically with respect to each other as summarized in the following lemma.

Lemma: The four subsets $\mathcal{X}^{[m]}$, $1 \leq m \leq 4$, satisfy

$$\mathcal{X}^{[2]} = +j \cdot \mathcal{X}^{[1]} \tag{14}$$
$$\mathcal{X}^{[3]} = -1 \cdot \mathcal{X}^{[1]} \tag{15}$$
$$\mathcal{X}^{[4]} = -j \cdot \mathcal{X}^{[1]}. \tag{16}$$

Proof: Consider any $\bar{\mathbf{x}}_q^{[1]} = \mathbf{P}\mathbf{b}_q^{[1]} \in \mathcal{X}^{[1]}$, where the ith element of $\mathbf{b}_q^{[1]}$ is $b^{[1]} = +1 + j$. Noting $j \cdot b^{[1]} = b^{[2]}$

$$j \cdot \bar{\mathbf{x}}_q^{[1]} = \mathbf{P}\left(j \cdot \mathbf{b}_q^{[1]}\right) \in \mathcal{X}^{[2]}. \tag{17}$$
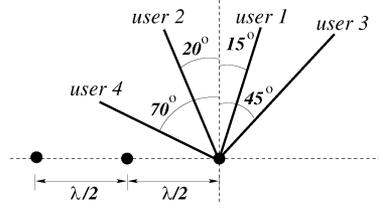
Fig. 1. Angular locations of the four QPSK users with respect to the three-element linear antenna array having $\boldsymbol{\lambda/2}$ spacing, where $\boldsymbol{\lambda}$ is the wavelength.

This proves the relationship (14). The proofs of the other two relationships are similar. Given this symmetry, the optimal Bayesian solution (12) can alternatively be expressed as

$$\begin{aligned}
y_{\mathrm{Bay}}(k) = \sum_{q=1}^{N_{sb}} \Big\{ & b^{[1]}\beta \cdot e^{-\|\mathbf{x}(k) - \bar{\mathbf{x}}_q^{[1]}\|^2 / 2\sigma_n^2} \\
& + b^{[2]}\beta \cdot e^{-\|\mathbf{x}(k) - j \cdot \bar{\mathbf{x}}_q^{[1]}\|^2 / 2\sigma_n^2} \\
& + b^{[3]}\beta \cdot e^{-\|\mathbf{x}(k) + \bar{\mathbf{x}}_q^{[1]}\|^2 / 2\sigma_n^2} \\
& + b^{[4]}\beta \cdot e^{-\|\mathbf{x}(k) + j \cdot \bar{\mathbf{x}}_q^{[1]}\|^2 / 2\sigma_n^2} \Big\}
\end{aligned} \tag{18}$$

where $\bar{\mathbf{x}}_q^{[1]} \in \mathcal{X}^{[1]}$ and $\beta > 0$ is any positive constant.

IV. SYMMETRIC RADIAL BASIS FUNCTION NETWORK

Consider the problem of realizing the optimal beamforming solution using a generic RBF network. The symmetry of the Bayesian solution (18) should be explicitly exploited, and we propose to use the following SRBF network for the detection of user i data:

$$\begin{aligned}
& y_{\mathrm{SRBF}}(k) \\
& = \sum_{q=1}^{N_c} \Big\{ \alpha_q^{[1]}\varphi(\mathbf{x}(k); \mathbf{c}_q, \sigma_q^2) + \alpha_q^{[2]}\varphi(\mathbf{x}(k); j\mathbf{c}_q, \sigma_q^2) \\
& \quad + \alpha_q^{[3]}\varphi(\mathbf{x}(k); -\mathbf{c}_q, \sigma_q^2) + \alpha_q^{[4]}\varphi(\mathbf{x}(k); -j\mathbf{c}_q, \sigma_q^2) \Big\}
\end{aligned} \tag{19}$$

with the decision $\hat{b}_i(k) = \mathrm{csgn}(y_{\mathrm{SRBF}}(k))$, where N_c is the number of RBF centers, $\mathbf{c}_q$ are the complex-valued RBF center vectors, $\alpha_q^{[m]}$ for $1 \leq m \leq 4$ are complex-valued RBF weights, σ_q^2 are the RBF variances, and $\varphi(\bullet)$ is the real-valued RBF. In this study, we adopt the Gaussian function of the form

$$\varphi(\mathbf{x}; \mathbf{c}_q, \sigma^2) = e^{-\|\mathbf{x} - \mathbf{c}_q\|^2 / 2\sigma^2}. \tag{20}$$

Since the number of users is usually known, the number of RBF centers can be set to $N_c = N_{sb}$. To further exploit the structure of the optimal Bayesian solution (18), the complex-valued RBF weights are set to $\alpha_q^{[m]} = \beta b^{[m]}$, $1 \leq m \leq 4$, where $\beta > 0$ is a constant. Obviously, specific value of β does not affect detection performance. This is simply because multiplying $y_{\mathrm{Bay}}(k)$ or $y_{\mathrm{SRBF}}(k)$ by a positive number does not alter its BER performance. Furthermore, all the RBF variances can be set to $\sigma_q^2 = \hat{\sigma}_n^2$, where $\hat{\sigma}_n^2$ is an estimate of the noise variance. Thus, adaptation of the SRBF network (19) becomes the task of finding appropriately the RBF center vectors $\mathbf{c}_q$.

Note that the standard complex-valued RBF network [28], [29] in the form of

$$y_{\mathrm{RBF}}(k) = \sum_{q=1}^{N_b} \alpha_q \varphi(\mathbf{x}(k); \mathbf{c}_q, \sigma_q^2) \tag{21}$$
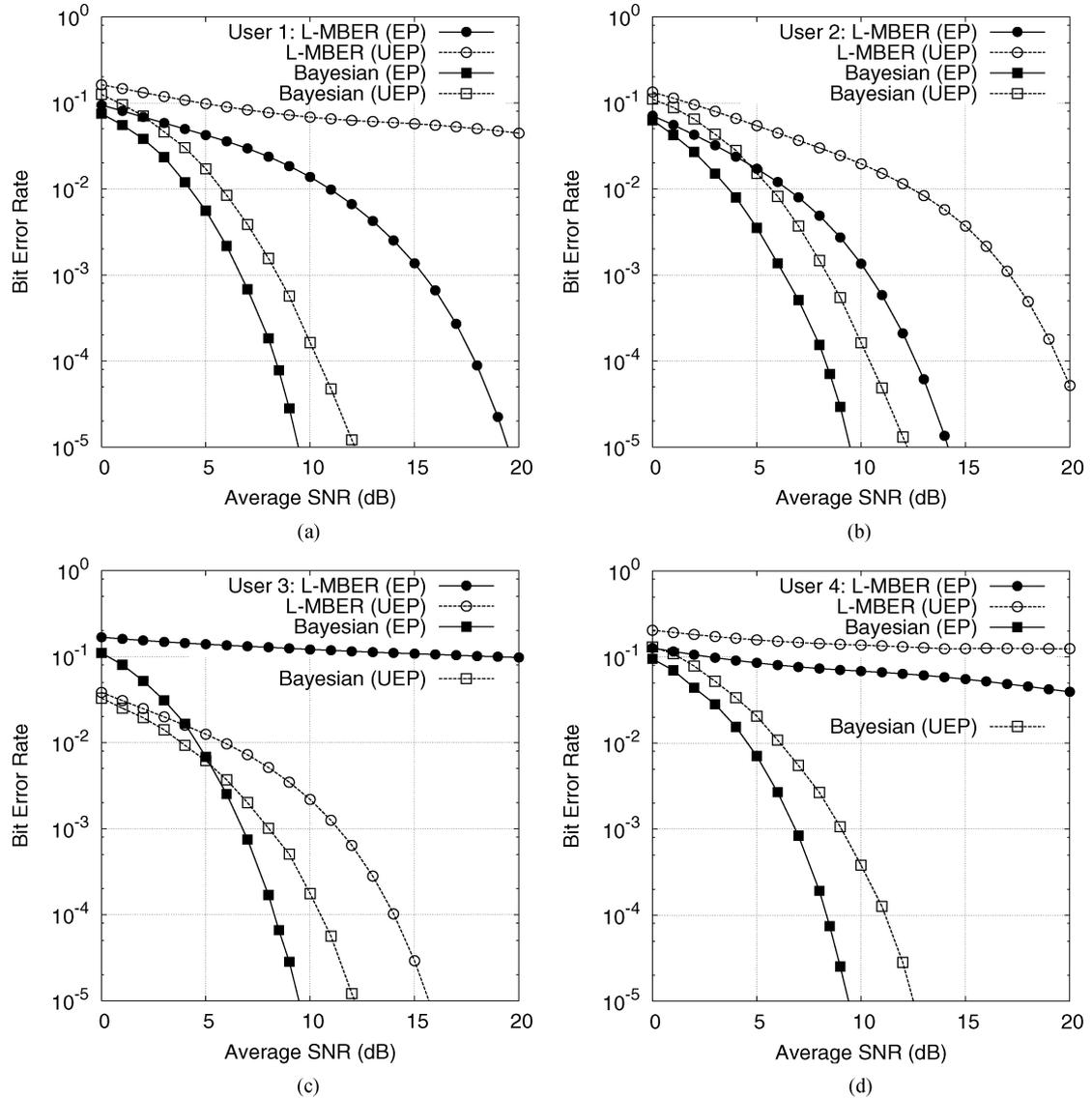
Fig. 2. BER performance comparison of the optimal nonlinear beamforming with the optimal linear beamforming, for the three-element array system supporting four QPSK users. In the equal power case, all four users have the same power, while in the near–far case, users 1, 2, and 4 have the same power but user 3 has 6 dB more power than users 1, 2, and 4. The total signal power, however, is kept the same for both the EP and UEP cases. (a) User 1. (b) User 2. (c) User 3. (d) User 4.

does not guarantee to posses the same symmetry property of the optimal Bayesian solution (18), particularly when the RBF centers $\mathbf{c}_q$ are obtained directly from the channel-impaired observation data. By contrast, the SRBF network (19) is guaranteed to have the same symmetry property of the optimal Bayesian solution (18).

A. Uplink Detection

In the uplink scenario, the receiver has to detect all the users' data, and it has access to the training symbols of all the users. The most effective way of adaptive implementation of the SRBF network detector (19) is to estimate the system's channel matrix $\mathbf{P}$ first and then use it to calculate the state subset $\mathcal{X}^{[1]}$, which specifies the optimal Bayesian solution. Given the training data set $\{\mathbf{x}(k), \mathbf{b}(k)\}_{k=1}^{K}$, where K is the number of training symbols, define the training symbol matrix and the corresponding observation matrix as $\mathbf{B}_K = [\mathbf{b}(1)\,\mathbf{b}(2)\cdots\mathbf{b}(K)]$ and $\mathbf{X}_K = [\mathbf{x}(1)\,\mathbf{x}(2)\cdots\mathbf{x}(K)]$, respectively. Then, the LSCE for $\mathbf{P}$ is given by

$$\hat{\mathbf{P}} = \mathbf{X}_K \mathbf{B}_K^H \left(\mathbf{B}_K \mathbf{B}_K^H \right)^{-1}. \tag{22}$$

An estimated channel noise variance

$$\hat{\sigma}_n^2 = \frac{1}{2K} \|\mathbf{X}_K - \hat{\mathbf{P}}\,\mathbf{B}_K\|^2 \tag{23}$$

is also produced as a byproduct of the LSCE.

Note that the supervised clustering algorithm of [17], [18], [21]–[25] can in theory be applied to directly identifying the channel states $\{\bar{\mathbf{x}}_q\}$ in the uplink case. In practice, however, this is too inefficient in comparison with our proposed approach of identifying the channel matrix $\mathbf{P}$. A rule of thumb for a supervised clustering is that each cluster requires at least ten data samples to converge. For the case of four QPSK users, this requires well over 2000 training symbol vectors $\mathbf{b}(k)$ to be transmitted. It is far simpler to estimate the channel matrix as we suggested, which requires less than 100 training symbol vectors.

B. Downlink Detection

In the downlink scenario, the task of receiver is to detect the data of the single desired user i. During training, the receiver has the training data of the reference user i, $\{\mathbf{x}(k), b_i(k)\}_{k=1}^{K}$, but the receiver does

not have access to the interfering users' data $\{b_q(k)\}$, $q \neq i$. Thus, estimating the system's channel matrix $\mathbf{P}$ is a challenging task. It is more feasible using the channel-impaired training data to directly adjust the SRBF network (19) and hence to approximate the optimal Bayesian solution. We propose to use the cluster-variation-assisted clustering algorithm [26], [30] to adapt the RBF center vectors. Specifically, during training, the RBF centers are adjusted according to

$$\mathbf{c}_l(k) = \mathbf{c}_l(k-1) + \mu_c \mathcal{M}_l(\check{\mathbf{x}}(k))(\check{\mathbf{x}}(k) - \mathbf{c}_l(k-1)) \qquad (24)$$

where

$$\check{\mathbf{x}}(k) = \begin{cases} +1 \cdot \mathbf{x}(k), & b_i(k) = b^{[1]} \\ -j \cdot \mathbf{x}(k), & b_i(k) = b^{[2]} \\ -1 \cdot \mathbf{x}(k), & b_i(k) = b^{[3]} \\ +j \cdot \mathbf{x}(k), & b_i(k) = b^{[4]} \end{cases} \qquad (25)$$

μ_c is the step size and the membership function $\mathcal{M}_l(\mathbf{x})$ is defined as

$$\mathcal{M}_l(\mathbf{x}) = \begin{cases} 1, & \text{if } \bar{v}_l \|\mathbf{x} - \mathbf{c}_l\|^2 \leq \bar{v}_q \|\mathbf{x} - \mathbf{c}_q\|^2, \forall q \neq l, \\ 0, & \text{otherwise} \end{cases} \qquad (26)$$

with $\bar{v}_l$ being the variation of the lth cluster. In order to estimate the associated variation $\bar{v}_l$, the following updating rule is used:

$$\bar{v}_l(k) = \mu_v \bar{v}_l(k-1) + (1 - \mu_v)\mathcal{M}_l(\check{\mathbf{x}}(k))\|\check{\mathbf{x}}(k) - \mathbf{c}_l(k-1)\|^2 \quad (27)$$

where μ_v is a constant slightly less than 1.0. The initial variations $\bar{v}_l(0)$, $\forall l$, are set to the same small number.

Note that this cluster-variation-assisted clustering algorithm is an unsupervised learning algorithm. Because the underlying symmetric property is exploited, we only need to find a smaller set of the RBF centers which are related to the signal subset $\mathcal{X}^{[1]}$. In order to do so, the desired user's symbols $b_i(k)$ are used to modify the noise observation data $\mathbf{x}(k)$ according to (25). It is known that this cluster-variation-enhanced clustering algorithm is capable of obtaining the optimal cluster partitioning structure and all the cluster variations converge to the same value [30]. Specifically, in our particular application, the RBF center vectors $\{\mathbf{c}_q\}$ converge to the subset of the noise-free signal states $\{\bar{\mathbf{x}}_q^{[1]}\}$ and all the cluster variations $\{\bar{v}_q\}$ converge to the noise variance $2\sigma_n^2$.

V. SIMULATION STUDY

A three-element antenna array was designed to support four QPSK users. Fig. 1 shows the angular positions of the four users . The simulated narrowband channels were $A_i = \eta_i(1 + j0), 1 \leq i \leq 4$, where η_i^2 specified the power of user i. First, we demonstrated the performance improvement achievable by the optimal nonlinear beamforming over the optimal linear one. Two cases of user power distribution were considered. In the equal power (EP) case, all four users had the same signal power, and therefore, all the SIRs were 0 dB. In the unequal power (UEP) case, users 1, 2, and 4 had the same power but user 3 had 6 dB more power than users 1, 2, and 4. The total signal power, however, was kept the same for both the EP and UEP cases. Fig. 2 compares the BER performance of the Bayesian beamforming and the L-MBER beamforming. As expected, the Bayesian beamforming achieved much better BER performance over the optimal linear beamforming. This performance gain was of course obtained at the cost of an increased complexity. Note that the horizontal axis in Fig. 2 is the average SNR as defined in (3), not the desired user SNR.
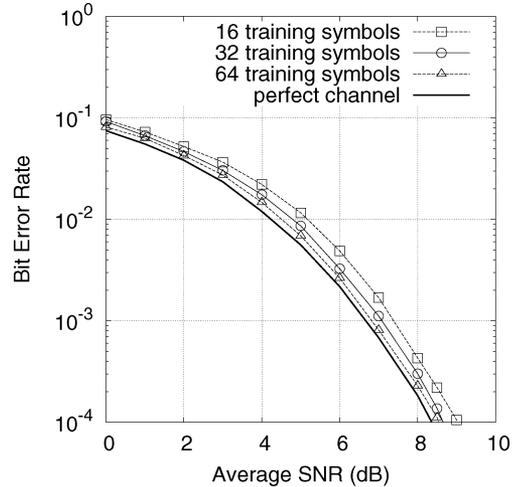
Fig. 3.   User 1 BER performance of the LSCE-based adaptive SRBF beamformer for the EP case, given $\hat{\sigma}_n^2 = \sigma_n^2$.

From Fig. 2, it can be seen that in the EP case the performance of the individual linear beamformer depended on the particular user's angular position as well as the other users' locations. By contrast, all the four optimal Bayesian beamformers had the similar performance. Below a plausible explanation for this enhanced robustness of Bayesian beamforming is offered, where the signal space is the same one, namely, $\mathcal{X}$, but different users partition $\mathcal{X}$ into different four subsets. In the EP user case, the amount of interference is similar for each user. However, the different users require a different partitioning of $\mathcal{X}$. For users 3 and 4, this underlying partitioning is linearly nonseparable and, hence, inherent error floors are observed for the L-MBER beamformers of users 3 and 4. By contrast, the Bayesian beamformer, by virtue of being the optimal nonlinear detection solution, is capable of successfully operating in both the linearly separable and linearly nonseparable cases. Because the amount of interference is similar for each user, the Bayesian beamformers of the different users exhibit a similar performance. Moreover, the results of the UEP case shown in Fig. 2 also confirm that the nonlinear beamforming was much more robust to the near–far effect than the linear beamforming. Because of this remarkable robustness property, we only concentrated on the user 1 in the EP case when investigating adaptive implementation of the nonlinear beamforming.

First, the LSCE-based adaptive implementation was investigated, and Fig. 3 depicts the user 1 BER performance of the adaptive SRBF beamformer with the different numbers of training symbols K, given $\hat{\sigma}_n^2 = \sigma_n^2$, in comparison with the case of the perfect channel knowledge. It can be seen from Fig. 3 that the LSCE-based adaptive implementation required $K = 64$ training symbols to closely approach the optimal Bayesian performance. The performance shown in Fig. 3 was obtained by setting the RBF variance $\hat{\sigma}_n^2$ to the true noise variance σ_n^2. The influence of the RBF variance $\hat{\sigma}_n^2$ used to the BER performance of the LSCE-based adaptive SRBF beamformer is illustrated in Fig. 4, given the SNR value of 7 dB and $K = 64$ training symbols. It can be seen from Fig. 4 that the performance of the LSCE-based adaptive SRBF beamformer is not sensitive at all to the value of the RBF variance and there exists a large range of $\hat{\sigma}_n^2$ values that enable the LSCE-based adaptive SRBF beamformer to match the Bayesian performance.

The clustering-based adaptive SRBF beamforming was then studied. For this example, the number of the subset channel states was $N_{sb} = 64$, and we used the first 64 data points $\check{\mathbf{x}}(k), 1 \leq k \leq 64$, as the initial RBF centers. The initial cluster variations were set to $\bar{v}_l(0) = 0.1$ for
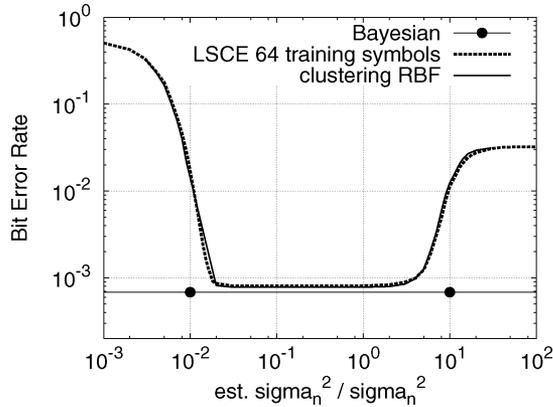
Fig. 4. Influence of the RBF variance on the BER performance of the LSCE- and clustering-based adaptive SRBF beamformers for user one of the EP case, given SNR = 7 dB.
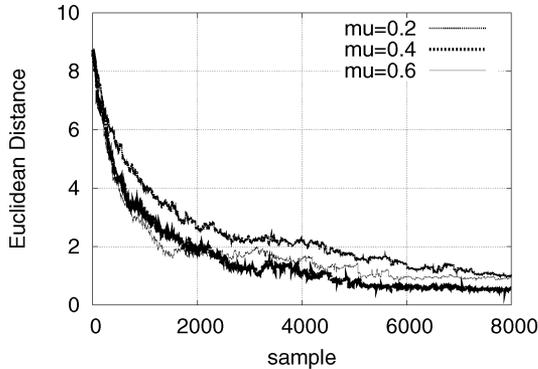
Fig. 5. Learning curves of the cluster algorithm for user 1 of the EP case, in terms of Euclidean distance between the RBF centers and true channel states averaged over ten runs, given SNR = 7 dB.

$1 \leq l \leq N_{sb}$, and the adaptive constant for updating the cluster variations was chosen to be $\mu_v = 0.995$. Note that the general rule is that all the initial cluster variations $\bar{v}_l(0)$ should be set to the same small positive number and μ_v should be set to a constant slightly less than 1.0. Convergence performance of the cluster-variation-enhanced clustering algorithm was assessed in the simulation based on the Euclidean distance between the set of the RBF centers $\{\mathbf{c}_l\}_{l=1}^{N_{sb}}$ and the set of the true subset channel states $\{\bar{\mathbf{x}}_l^{[1]}\}_{l=1}^{N_{sb}}$ defined as

$$ \mathrm{ED}(k) = \frac{1}{N_{sb}} \sum_{l=1}^{N_{sb}} \|\mathbf{c}_l(k) - \bar{\mathbf{x}}_l^{[1]}\|^2. \qquad (28) $$

Given SNR = 7 dB, Fig. 5 plots the learning curves of the clustering algorithm in terms of the Euclidean distance (28) averaged over ten different random runs for the three values of the adaptive gain μ_c. It is seen from Fig. 5 that for this example the best convergence performance was achieved with $\mu_c = 0.4$. The robustness of the clustering-based adaptive SRBF beamforming with respect to the value of the RBF variance $\hat{\sigma}_n^2$ used is also demonstrated in Fig. 4, while Fig. 6 compares the BER performance of the clustering-based adaptive SRBF beamformer for user one after convergence with that of the optimal Bayesian beamformer, given the RBF variance $\hat{\sigma}_n^2 = \sigma_n^2$.

VI. CONCLUSION

A nonlinear-beamforming-based detector has been extended to multiple-antenna-assisted wireless systems that employ the complex-valued QPSK modulation. It has been demonstrated that nonlinear
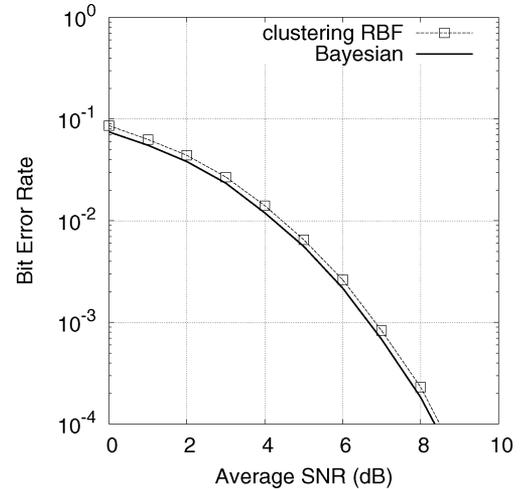
Fig. 6. User 1 BER performance of the clustering-based adaptive SRBF beamformer for the EP case, given $\hat{\sigma}_n^2 = \sigma_n^2$, in comparison with the optimal Bayesian beamforming performance.

beamforming is capable of substantially improving the achievable system performance and significantly increasing user capacity over the traditional linear beamforming, at the cost of an increased computational complexity. By explicitly exploiting the inherent symmetry of the optimal Bayesian solution, a novel complex-valued SRBF network has been proposed for adaptive nonlinear beamforming. It has been shown that in the uplink scenario adaptive SRBF beamforming can be implemented efficiently by estimating the system's channel matrix based on the LSCE. For the much more challenging downlink scenario, the cluster-variation enhanced clustering algorithm has been adopted to implement the adaptive SRBF beamforming. The robustness of the adaptive SRBF beamformer with respect to the RBF variance used has been verified in the simulation. Our future research will focus on the extension of this complex-valued SRBF network to the generic MIMO system in the framework of nonlinear space-time equalization, where frequency selective channels are encountered. Further work will also be carried out to investigate decision-directed adaptation for the clustering algorithm in order to shorten the training length.

REFERENCES

[1] J. Litva and T. K. Y. Lo, *Digital Beamforming in Wireless Communications*. London, U.K.: Artech House, 1996.
[2] J. H. Winters, "Smart antennas for wireless systems," *IEEE Personal Commun.*, vol. 5, no. 1, pp. 23–27, Feb. 1998.
[3] X. Wang and H. V. Poor, "Robust adaptive array for wireless communications," *IEEE J. Sel. Areas Commun.*, vol. 16, no. 8, pp. 1352–1366, Aug. 1998.
[4] P. Vandenameele, L. van Der Perre, and M. Engels, *Space Division Multiple Access for Wireless Local Area Networks*. Boston, MA: Kluwer, 2001.
[5] J. S. Blogh and L. Hanzo, *Third Generation Systems and Intelligent Wireless Networking – Smart Antennas and Adaptive Modulation*. Chichester, U.K.: Wiley, 2002.
[6] W. M. Younis, A. H. Sayed, and N. Al-Dhahir, "Efficient adaptive receivers for joint equalization and interference cancellation in multiuser space-time block-coded systems," *IEEE Trans. Signal Process.*, vol. 51, no. 11, pp. 2849–2862, Nov. 2003.
[7] A. Paulraj, R. Nabar, and D. Gore, *Introduction to Space-Time Wireless Communications*. Cambridge, U.K.: Cambridge Univ. Press, 2003.
[8] A. J. Paulraj, D. A. Gore, R. U. Nabar, and H. Bölcskei, "An overview of MIMO communications – A key to gigabit wireless," *Proc. IEEE*, vol. 92, no. 2, pp. 198–218, Feb. 2004.
[9] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge, U.K.: Cambridge Univ. Press, 2005.

[10] H. Li and H. V. Poor, "Spectral efficiency of equal-rate DS-CDMA systems with multiple transmit antennas," *IEEE Trans. Wireless Commun.*, vol. 5, no. 12, pp. 3680–3688, Dec. 2006.

[11] A. M. Tulino, A. Lozano, and S. Verdú, "Capacity-achieving input covariance for single-user multi-antenna channels," *IEEE Trans. Wireless Commun.*, vol. 5, no. 3, pp. 662–671, Mar. 2006.

[12] M. Sadek, A. Tarighat, and A. H. Sayed, "Active Antenna selection in multiuser MIMO communications," *IEEE Trans. Signal Process.*, vol. 55, no. 4, pp. 1498–1510, Apr. 2007.

[13] D. N. C. Tse and S. V. Hanly, "Linear multiuser receivers: Effective interference, effective bandwidth and user capacity," *IEEE Trans. Inf. Theory*, vol. 45, no. 2, pp. 641–657, Mar. 1999.

[14] S. Chen, N. N. Ahmad, and L. Hanzo, "Adaptive minimum bit error rate beamforming," *IEEE Trans. Wireless Commun.*, vol. 4, no. 2, pp. 341–348, Mar. 2005.

[15] S. Chen, L. Hanzo, N. N. Ahmad, and A. Wolfgang, "Adaptive minimum bit error rate beamforming assisted receiver for QPSK wireless communication," *Digital Signal Process.*, vol. 15, no. 6, pp. 545–567, 2005.

[16] K. Abend and B. D. Fritchman, "Statistical detection for communication channels with intersymbol interference," *Proc. IEEE*, vol. 58, no. 5, pp. 779–785, May 1970.

[17] S. Chen, B. Mulgrew, and P. M. Grant, "A clustering technique for digital communications channel equalization using radial basis function networks," *IEEE Trans. Neural Netw.*, vol. 4, no. 4, pp. 570–579, Jul. 1993.

[18] S. Chen, S. McLaughlin, B. Mulgrew, and P. M. Grant, "Bayesian decision feedback equalizer for overcoming co-channel interference," *Proc. Inst. Electr. Eng. Commun.*, vol. 143, no. 4, pp. 219–225, 1996.

[19] S. Verdú, *Multiuser Detection*. New York: Cambridge Univ. Press, 1998.

[20] S. Chen, A. K. Samingan, and L. Hanzo, "Support vector machine multiuser receiver for DS-CDMA signals in multipath channels," *IEEE Trans. Neural Netw.*, vol. 12, no. 3, pp. 604–611, May 2001.

[21] Y. Kopsinis and S. Theodoridis, "A novel cluster based MLSE equalizer for M-PAM signaling schemes," *Signal Process.*, vol. 83, no. 9, pp. 1905–1918, 2003.

[22] Y. Kopsinis and S. Theodoridis, "An efficient low-complexity technique for MLSE equalizers for linear and nonlinear channels," *IEEE Trans. Signal Process.*, vol. 51, no. 12, pp. 3236–3248, Dec. 2003.

[23] S. Chen, L. Hanzo, and A. Wolfgang, "Nonlinear multi-antenna detection methods," *EURASIP J. Appl. Signal Process.*, vol. 2004, no. 9, pp. 1225–1237, 2004.

[24] E. Kofidis, Y. Kopsinis, and S. Theodoridis, "On the least-squares performance of a novel efficient center estimation method for clustering-based MLSE equalization," *IEEE Trans. Signal Process.*, vol. 54, no. 4, pp. 1459–1471, Apr. 2006.

[25] E. Kofidis, V. Dalakas, Y. Kopsinis, and S. Theodoridis, "A novel efficient cluster-based MLSE equalizer for satellite communication channels with M-QAM signaling," *EURASIP J. Appl. Signal Process.*, vol. 2006, p. 16, 2006.

[26] S. Chen, K. Labib, and L. Hanzo, "Clustering-based symmetric radial basis function beamforming," *IEEE Signal Process. Lett.*, vol. 14, pp. 589–592, 2007.

[27] S. Chen, A. Wolfgang, C. J. Harris, and L. Hanzo, "Symmetric RBF classifier for nonlinear detection in multiple-antenna-aided systems," *IEEE Trans. Neural Netw.*, vol. 19, no. 5, pp. 737–745, May 2008.

[28] S. Chen, S. McLaughlin, and B. Mulgrew, "Complex-valued radial basis function network—Part I: Network architecture and learning algorithms," *Signal Process.*, vol. 35, pp. 19–31, 1994.

[29] S. Chen, S. McLaughlin, and B. Mulgrew, "Complex-valued radial basis function network—Part II: Application to digital communications channel equalization," *Signal Process.*, vol. 36, pp. 175–188, 1994.

[30] C. Chinrungrueng and C. H. Séquin, "Optimal adaptive κ-means algorithm with dynamic adjustment of learning rate," *IEEE Trans. Neural Netw.*, vol. 6, no. 1, pp. 1873–1896, Jan. 1995.

[31] R. Steele and L. Hanzo, *Mobile Radio Communications*. Piscataway, NJ: IEEE Press, 1999.

[32] R. O. Duda and P. E. Hart, *Pattern Classification and Scene Analysis*. New York: Wiley, 1973.