SYMMETRIC RADIAL BASIS FUNCTION NETWORK EQUALISER

S. Chen[1], *A. Wolfgang*[1], *S. Benedetto*[2], *P. Duhamel*[3], *L. Hanzo*[1]

1 - School of ECS - University of Southampton, UK
2 - Dipartimento di Elettronica - Politecnico di Torino, Italy
3 - CNRS/LSS - Supelec, France

ABSTRACT

The paper investigates nonlinear equalisation using a novel symmetric radial basis function (RBF) network. By explicitly exploiting the inherently symmetric structure of the optimal Bayesian equaliser, the proposed symmetric RBF equaliser can be determined from the received noisy training data. Both a block-data based and a sample-by-sample adaptive algorithm are designed for this novel symmetric RBF equaliser. Simulation results are also provided to demonstrate the efficiency of the proposed symmetric RBF network equaliser.

1. INTRODUCTION

In this paper, we re-visit nonlinear equalisation using the radial basis function (RBF) network [1]-[5]. It is well-known that equalisation can be viewed as a classification problem and the optimal solution for this classification problem is known to be the Bayesian equaliser [1],[2]. We first show that the Bayesian nonlinear equalisation solution has an inherent symmetry, because the signal states corresponding to the different signal classes are distributed symmetrically [6]. This symmetry is hard to infer from noisy training data using the traditional RBF network. We propose a novel RBF network that is capable of exploiting the signal constellation's symmetric structure and demonstrate that the proposed symmetric RBF network is capable of approaching the optimal Bayesian equalisation performance.

A block-data based algorithm is developed for the construction of the symmetric RBF equaliser using the orthogonal forward selection (OFS) procedure combined with the Fisher ratio of class separability measure (FRCSM) [7]-[9]. It is shown that by explicitly exploiting the symmetry of the underlying signal constellation, the proposed symmetric RBF equaliser becomes capable of effectively realising the Bayesian equalisation solution. A novel nonlinear least bit error rate (NLBER) algorithm is also proposed, which enables a sample-by-sample adaptation of the symmetric RBF equaliser. The NLBER adaptive algorithm has its roots in the Parzen window density estimation technique of [10]-[13].

2. NONLINEAR CHANNEL EQUALISATION

Consider a binary phase shift keying (BPSK) modulation scheme communicating over a dispersive communication channel. The symbol-rate channel output samples can be expressed as

$$x(k) = \sum_{i=0}^{n_c-1} c_i b(k-i) + n(k), \tag{1}$$

where c_i represents the channel taps, n_c the channel impulse duration, $b(k) \in \{\pm 1\}$ and $n(k)$ the Gaussian white noise associated with $E[|n(k)|^2] = \sigma_n^2$. A finite-memory equaliser is employed, which has the input vector $\mathbf{x}(k) = [x(k)\ x(k-1)\cdots x(k-n_e+1)]^T$ in order to detect the transmitted symbols $b(k-\tau)$, where n_e is the equaliser's order and τ is the decision delay. The vector $\mathbf{x}(k)$ is given by

$$\mathbf{x}(k) = \mathbf{C}\mathbf{b}(k) + \mathbf{n}(k) = \bar{\mathbf{x}}(k) + \mathbf{n}(k), \tag{2}$$

where $\mathbf{n}(k) = [n(k)\ \ n(k-1)\cdots n(k-n_e+1)]^T$, $\mathbf{b}(k) = [b(k)\ \ b(k-1)\cdots b(k-L+1)]^T$ and $\mathbf{C}$ is the $(n_e \times L)$-dimensional channel matrix having $L = n_c + n_e - 1$.

Denote the $N_b = 2^L$ combinations of $\mathbf{b}(k)$ as $\mathbf{b}_q$, $1 \le q \le N_b$. Furthermore, denote the $(\tau+1)$-th element of $\mathbf{b}_q$ as $b_{q,\tau}$. The noiseless channel output $\bar{\mathbf{x}}(k)$ assumes legitimate values from the signal state set

$$\mathscr{X} \stackrel{\triangle}{=} \{\bar{\mathbf{x}}_q = \mathbf{C}\mathbf{b}_q,\ 1 \le q \le N_b\}. \tag{3}$$

The decision variable of the optimal Bayesian equaliser is then defined as [1],[2]

$$y_{Bay}(k) = \sum_{q=1}^{N_b} \mathrm{sgn}(b_{q,\tau})\beta_q e^{-\frac{\|\mathbf{x}(k)-\bar{\mathbf{x}}_q\|^2}{2\sigma_n^2}}, \tag{4}$$

with the optimal decision given by

$$\hat{b}(k-\tau) = \mathrm{sgn}(y_{Bay}(k)) = \begin{cases} +1, & y_{Bay}(k) \ge 0, \\ -1, & y_{Bay}(k) < 0, \end{cases} \tag{5}$$

where we have $\beta_q = \frac{1}{N_b(2\pi\sigma_n^2)^{L/2}}$.

The signal state set $\mathscr{X}$ can be divided into the following two subsets conditioned on the value of $b(k-\tau)$

$$\mathscr{X}^{(\pm)} \stackrel{\triangle}{=} \{\bar{\mathbf{x}}_i \in \mathscr{X}, 1 \le i \le N_{sb} :\ b(k-\tau) = \pm 1\}, \tag{6}$$

where the sizes of $\mathscr{X}^{(+)}$ and $\mathscr{X}^{(-)}$ are both $N_{sb} = N_b/2$. It may readily be visualised that the two subsets $\mathscr{X}^{(+)}$ and $\mathscr{X}^{(-)}$ are distributed symmetrically with respect to each other [6]. More explicitely, given an appropriate constellation point indexing, for any phasor $\bar{\mathbf{x}}_i^{(+)} \in \mathscr{X}^{(+)}$ there exists a symmetrical positioned phasor $\bar{\mathbf{x}}_i^{(-)} \in \mathscr{X}^{(-)}$ so that

The financial support of the EU under the auspices of the Newcom projects is gratefully acknowledged.

we have $\bar{\mathbf{x}}_i^{(-)} = -\bar{\mathbf{x}}_i^{(+)}$. Upon exploiting this symmetry, the Bayesian equaliser (4) can be rewritten as

$$y_{Bay}(k) = \sum_{q=1}^{N_{sb}} \beta_q \left(e^{-\frac{\|\mathbf{x}(k)-\bar{\mathbf{x}}_q^{(+)}\|^2}{2\sigma_n^2}} - e^{-\frac{\|\mathbf{x}(k)+\bar{\mathbf{x}}_q^{(+)}\|^2}{2\sigma_n^2}} \right), \quad (7)$$

where we have $\bar{\mathbf{x}}_q^{(+)} \in \mathscr{X}^{(+)}$. Note that the symmetry of the Bayesian equaliser, as seen in (7) is hard to both recognise and to exploit using a traditional RBF network.

3. SYMMETRIC RBF NETWORK EQUALISER

Consider the problem of training a RBF network $y_{RBF}(\mathbf{x}) : \mathscr{R}^{n_e} \to \{\pm 1\}$ based on a training data set $D_K = \{\mathbf{x}(k), d(k)\}_{k=1}^K$, where $d(k) \in \{\pm 1\}$ is the class type for data sample $\mathbf{x}(k)$. We adopt the RBF network of the form

$$\hat{d}(k) = \mathrm{sgn}(y_{RBF}(k)) \ \text{ with } \ y_{RBF}(k) = \sum_{i=1}^{M} \theta_i \phi_i(\mathbf{x}(k)), \quad (8)$$

where $\hat{d}(k)$ is the estimated class label for $\mathbf{x}(k)$, $\phi_i(\bullet)$ denotes the i-th RBF node, θ_i are the RBF weights and M is the number of RBF centres. We propose to adopt the following symmetric RBF construction

$$\phi_i(\mathbf{x}) \stackrel{\triangle}{=} \varphi(\mathbf{x}; \mu_i, \rho^2) - \varphi(\mathbf{x}; -\mu_i, \rho^2), \quad (9)$$

where $\mu_i \in \mathscr{R}^{n_e}$ represents the RBF centres, ρ^2 the RBF variance and $\varphi(\bullet)$ the usual RBF function. In this study we adopt the Gaussian RBF function of

$$\varphi(\mathbf{x}; \mu_i, \rho^2) = e^{-\frac{\|\mathbf{x}-\mu_i\|^2}{2\rho^2}}. \quad (10)$$

We now consider both a block-data based and a sample-by-sample adaptive algorithm for constructing this symmetric RBF equaliser.

3.1 Block-Data Based Algorithm

We apply the OFS procedure based on the FRCSM [8],[9] to construct a sparse symmetric RBF equaliser using the training data set D_k. We consider every training data point $\mathbf{x}(i)$ as a candidate RBF centre, hence we have $M = K$ in the RBF model (8) and $\mu_i = \mathbf{x}(i)$ for $1 \le i \le K$ as well as a given RBF variance ρ^2. Let us now define $\varepsilon(i) = d(i) - y_{RBF}(i)$ as the modelling residual sequence. Then the model (8) constructed from the training data set D_K can be written in matrix form as

$$\mathbf{d} = \mathbf{\Phi}\theta + \varepsilon, \quad (11)$$

where $\mathbf{d} = [d(1)\ d(2) \cdots d(K)]^T$, $\varepsilon = [\varepsilon(1)\ \varepsilon(2) \cdots \varepsilon(K)]^T$, $\theta = [\theta_1\ \theta_2 \cdots \theta_M]^T$, and

$$\mathbf{\Phi} = [\phi_1\ \phi_2 \cdots \phi_M] \in \mathscr{R}^{K \times M} \quad (12)$$

is the regression matrix having the column vectors $\phi_i = [\phi_i(\mathbf{x}(1))\ \phi_i(\mathbf{x}(2)) \cdots \phi_i(\mathbf{x}(K))]^T$, $1 \le i \le M$. Let an orthogonal decomposition of $\mathbf{\Phi}$ be $\mathbf{\Phi} = \mathbf{\Omega}\mathbf{A}$, where

$$\mathbf{A} = \begin{bmatrix} 1 & \alpha_{1,2} & \cdots & \alpha_{1,M} \\ 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \alpha_{M-1,M} \\ 0 & \cdots & 0 & 1 \end{bmatrix} \quad (13)$$

and

$$\mathbf{\Omega} = [\omega_1\ \omega_2 \cdots \omega_M] = \begin{bmatrix} \omega_{1,1} & \omega_{1,2} & \cdots & \omega_{1,M} \\ \omega_{2,1} & \omega_{2,2} & \cdots & \omega_{2,M} \\ \vdots & \vdots & \vdots & \vdots \\ \omega_{K,1} & \omega_{K,2} & \cdots & \omega_{K,M} \end{bmatrix}, \quad (14)$$

where $\mathbf{\Omega}$ has orthogonal columns that satisfy $\omega_i^T \omega_l = 0$, if $i \ne l$. The model (11) can alternatively be expressed as

$$\mathbf{d} = \mathbf{\Omega}\gamma + \varepsilon, \quad (15)$$

where $\gamma = [\gamma_1\ \gamma_2 \cdots \gamma_M]^T = \mathbf{A}\theta$ is the weight vector in the orthogonal space defined by $\mathbf{\Omega}$.

A sparse M_{spa}-term RBF network can be selected by incrementally maximising the FRCSM using the OFS procedure, outlined in [8],[9]. Let us define the two class sets $\mathbf{X}_\pm = \{\mathbf{x}(k) : d(k) = \pm 1\}$, and let the number of points in $\mathbf{X}_\pm$ be $K_\pm$, respectively, with $K_+ + K_- = K$. The means and variances of the training samples belonging to class $\mathbf{X}_+$ and class $\mathbf{X}_-$ in the direction of the basis ω_l are given by

$$m_{+,l} = \frac{1}{K_+} \sum_{k=1}^{K} \delta(d(k)-1)\omega_{k,l}, \quad (16)$$

$$\sigma_{+,l}^2 = \frac{1}{K_+} \sum_{k=1}^{K} \delta(d(k)-1)\left(\omega_{k,l} - m_{+,l}\right)^2, \quad (17)$$

and

$$m_{-,l} = \frac{1}{K_-} \sum_{k=1}^{K} \delta(d(k)+1)\omega_{k,l}, \quad (18)$$

$$\sigma_{-,l}^2 = \frac{1}{K_-} \sum_{k=1}^{K} \delta(d(k)+1)\left(\omega_{k,l} - m_{-,l}\right)^2, \quad (19)$$

respectively, where we have

$$\delta(x) = \begin{cases} 1, & x = 0, \\ 0, & x \ne 0. \end{cases} \quad (20)$$

To elaborate a little further, the Fisher ratio is defined as the ratio of the inter-class difference and the intra-class spread encountered in the direction of ω_l, which is given by [14]

$$F_l = \frac{\left(m_{+,l} - m_{-,l}\right)^2}{\sigma_{+,l}^2 + \sigma_{-,l}^2}. \quad (21)$$

Based on the FRCSM, the significant RBF units can be selected with the aid of an OFS procedure. At the l-th stage, a candidate unit is chosen as the l-th RBF unit in the selected model, if it produces the largest F_l ratio among the $M - l + 1$ candidate units ω_i. The procedure is terminated with a sparse M_{spa}-term model when we have

$$\frac{F_{M_{\text{spa}}}}{\sum_{l=1}^{M_{\text{spa}}} F_l} < \xi, \quad (22)$$

where the threshold ξ determines the grade of sparsity for the model selected. The appropriate value of ξ depends on the application concerned and it must be determined empirically. The least squares solution for the corresponding sparse model weight vector $\theta_{M_{\text{spa}}} = [\theta_1\ \theta_2 \cdots \theta_{M_{\text{spa}}}]^T$ is readily available, given the least squares solution of $\gamma_{M_{\text{spa}}} = [\gamma_1\ \gamma_2 \cdots \gamma_{M_{\text{spa}}}]^T$. The detailed construction algorithm based on the Gram-Schmidt orthogonalisation can be found in [8],[9] and hence it will not be repeated here.

3.2 Sample-by-Sample Adaptive Algorithm

Let us quantify the dependency of the RBF network's output on its parameters by using the general notation

$$y_{RBF}(k;\mathbf{w}) = \sum_{i=1}^{M} \theta_i \left(\varphi(\mathbf{x}(k);\mu_i,\rho_i^2) - \varphi(\mathbf{x}(k);-\mu_i,\rho_i^2)\right), \tag{23}$$

where the parameter vector $\mathbf{w}$ includes all the RBF centres μ_i, variances ρ_i^2 and weights θ_i. Let us define the signed decision variable as $y_s(k) = \mathrm{sgn}(d(k))y_{RBF}(k;\mathbf{w})$ and denote the probability density function (PDF) of $y_s(k)$ as $p_y(y_s)$. Then the error probability of the RBF equaliser (23) is

$$P_E(\mathbf{w}) = \mathrm{Prob}\{y_s(k) < 0\} = \int_{-\infty}^{0} p_y(y_s)\,dy_s. \tag{24}$$

The minimum bit error rate (MBER) solution is defined as the parameter vector $\mathbf{w}$ that directly minimises $P_E(\mathbf{w})$ [13]. Although the PDF of $y_s(k)$ is unknown, it may be estimated sufficiently accurately using the Parzen window method. Specifically, given a block of training data D_K, a Parzen window estimate [10] of $p_y(y_s)$ is given as

$$\tilde{p}_y(y_s) = \frac{1}{K\sqrt{2\pi}\sigma} \sum_{k=1}^{K} e^{-\frac{(y_s - \mathrm{sgn}(d(k))y_{RBF}(k;\mathbf{w}))^2}{2\sigma^2}}, \tag{25}$$

where σ^2 is the kernel variance chosen. With this estimated PDF, the estimated or approximate BER is given by

$$\tilde{P}_E(\mathbf{w}) = \int_{-\infty}^{0} \tilde{p}_y(y_s)\,dy_s = \frac{1}{K}\sum_{k=1}^{K} Q(\tilde{g}_k(\mathbf{w})), \tag{26}$$

where $Q(\bullet)$ is the usual Gaussian error function and

$$\tilde{g}_k(\mathbf{w}) = \frac{\mathrm{sgn}(d(k))y_{RBF}(k;\mathbf{w})}{\sigma}. \tag{27}$$

An approximate MBER solution for $\mathbf{w}$ can be obtained by minimising $\tilde{P}_E(\mathbf{w})$ using a gradient-based optimisation, commencing for example from the minimim mean squared error (MMSE) solution. To derive a sample-by-sample adaptive algorithm, consider a single-sample PDF "estimate" of $p_y(y_s)$ given by

$$\tilde{p}_y(y_s,k) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(y_s - \mathrm{sgn}(d(k))y_{RBF}(k;\mathbf{w}))^2}{2\sigma^2}}. \tag{28}$$

Conceptually, given this instantaneous PDF "estimate", we arrive at the single-sample BER "estimate" $\tilde{P}_E(\mathbf{w},k) = Q(\tilde{g}_k(\mathbf{w}))$. Using the instantaneous gradient $\nabla\tilde{P}_E(\mathbf{w},k)$ gives rise to the following stochastic gradient algorithm

$$\begin{aligned}\mathbf{w}(k) \;&=\; \mathbf{w}(k-1) + \frac{\xi}{\sqrt{2\pi}\sigma} e^{-\frac{y^2_{RBF}(k;\mathbf{w}(k-1))}{2\sigma^2}} \\ &\quad \times \mathrm{sgn}(d(k)) \frac{\partial y_{RBF}(k;\mathbf{w}(k-1))}{\partial \mathbf{w}},\end{aligned} \tag{29}$$

which we refer to as the NLBER algorithm, where the step size ξ and kernel width σ should be carefully chosen in order to ensure a fast convergence and small steady-state BER misadjustment.

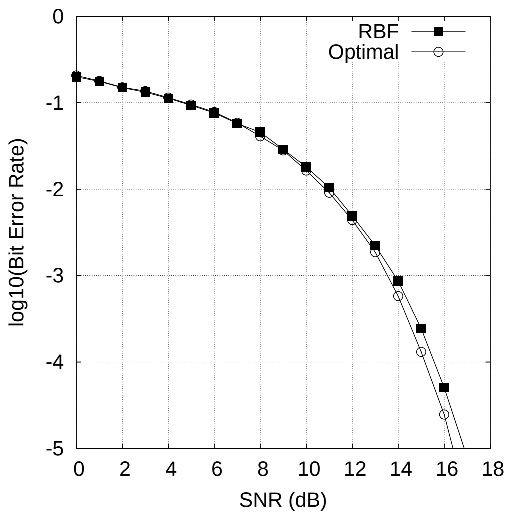

Figure 1: BER performance of the symmetric RBF and the Bayesian equaliser for transmission over a two-tap channel using an equaliser order $n_e = 2$ and a decision delay $\tau = 1$. The OFS aided FRCSM algorithm was used.

For the RBF equaliser (23) using the Gaussian basis function of (10), the derivatives of the RBF network's output with respect to the RBF equaliser's parameters are given by

$$\begin{aligned}\frac{\partial y_{RBF}}{\partial \theta_i} &= e^{-\frac{\|\mathbf{x}(k)-\mu_i\|^2}{\rho_i^2}} - e^{-\frac{\|\mathbf{x}(k)+\mu_i\|^2}{\rho_i^2}}, \\ \frac{\partial y_{RBF}}{\partial \rho_i^2} &= \theta_i \left(e^{-\frac{\|\mathbf{x}(k)-\mu_i\|^2}{\rho_i^2}} \frac{\|\mathbf{x}(k)-\mu_i\|^2}{(\rho_i^2)^2} - e^{-\frac{\|\mathbf{x}(k)+\mu_i\|^2}{\rho_i^2}} \frac{\|\mathbf{x}(k)+\mu_i\|^2}{(\rho_i^2)^2} \right), \\ \frac{\partial y_{RBF}}{\partial \mu_i} &= \theta_i \left(e^{-\frac{\|\mathbf{x}(k)-\mu_i\|^2}{\rho_i^2}} \frac{\mathbf{x}(k)-\mu_i}{\rho_i^2} + e^{-\frac{\|\mathbf{x}(k)+\mu_i\|^2}{\rho_i^2}} \frac{\mathbf{x}(k)+\mu_i}{\rho_i^2} \right),\end{aligned} \tag{30}$$

for $1 \le i \le M$.

4. SIMULATION STUDY

Example 1. The two-tap CIR used was given by $x(k) = 0.5b(k) + 1.0b(k-1) + n(k)$, the equaliser order was set to $n_e = 2$ and the decision delay to $\tau = 1$. Fig. 1 shows the BER of the optimal Bayesian equaliser as a function of the

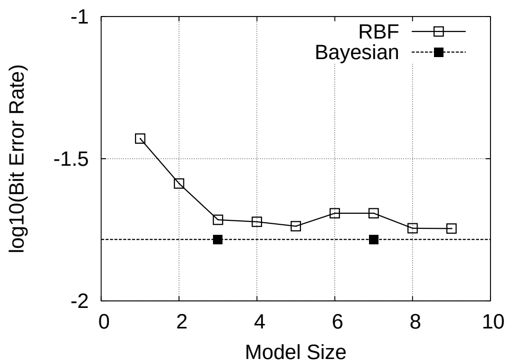

Figure 2: Influence of the model size on the BER performance for transmission over a two-tap channel in conjunction with $n_e = 2$ and $\tau = 1$, given SNR= 10 dB. The OFS aided FRCSM algorithm was used.

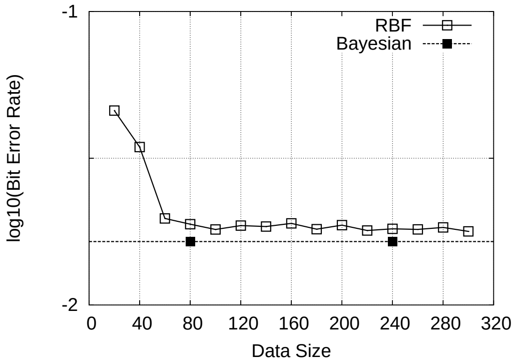

Figure 3: Influence of the data length on the BER performance for transmission over the two-tap channel in conjunction with $n_e = 2$ and $\tau = 1$, given SNR= 10 dB. The OFS aided FRCSM algorithm was used.

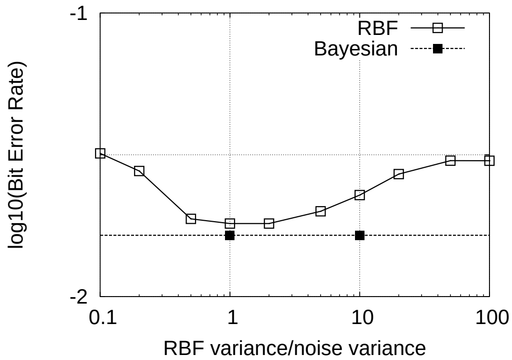

Figure 4: Influence of the RBF variance on the BER performance for transmission over the two-tap channel in conjunction with $n_e = 2$ and $\tau = 1$, given SNR= 10 dB. The OFS aided FRCSM algorithm was used.

signal to noise ratio (SNR). For this example, the size of the Bayesian equaliser was defined by $N_{sb} = 4$. For each SNR, a symmetric RBF equaliser having a $M_{\text{spa}} = 4$ was constructed from the training data set of length $K = 160$ using the OFS aided FRCSM algorithm, and the BER performance of the resultant RBF equaliser is also plotted in Fig. 1. Given SNR= 10 dB and a training data length of $K = 160$, Fig. 2 depicts the influence of the RBF model size M_{spa} on the attainable BER performance, where the RBF variance ρ^2 was tuned for each model size and was in the range of σ_n^2 to $4\sigma_n^2$. At SNR= 10 dB, a model size of $M_{\text{spa}} = 4$ and the RBF variance of $\rho^2 = 2\sigma_n^2$, the influence of the data length K is plotted in Fig. 3, while Fig. 4 illustrates the influence of the RBF variance ρ^2 given SNR= 10 dB, $M_{\text{spa}} = 4$ and $K = 160$. Fig. 5 compares the optimal Bayesian decision boundary with that of the RBF equaliser, given SNR= 10 dB.

Example 2. The three-tap CIR was given by $x(k) = 0.3b(k) + 0.8b(k-1) + 0.3b(k-2) + n(k)$, the equaliser order was set to $n_e = 4$ and the decision delay was $\tau = 2$. The BER of the optimal Bayesian equaliser is depicted in Fig. 6. For this example, the size of the Bayesian equaliser was defined by $N_{sb} = 32$. Given SNR= 13 dB and a model size of $M = 30$, Fig. 7 shows the learning curve $\tilde{P}_E(\mathbf{w}(k))$ of the NLBER algorithm averaged over 10 runs, where the BER was estimated using a block size of $K = 500$ symbols and a kernel variance of $\sigma^2 = \sigma_n^2$. The NLBER algorithm having a step size $\xi = 0.1$ was initialised with the first 30 data points as the initial RBF centres, and we had $\rho_i^2(0) = 4.0\sigma_n^2$ and $\theta_i(0) = \frac{1}{30}$ for $1 \leq i \leq 30$. The true BER $P_E(\mathbf{w}(k))$ was also calculated using simulations for $K = 0, 400, 800, 1200, 1600, 2000$. Fig. 8 depicts the influence of the model size on the RBF equaliser's performance, where the NLBER algorithm had the same settings as those used for obtaining the results of Fig. 7. The BER of the RBF equaliser using the model size of $M = 30$, trained by the NLBER algorithm over $K = 2000$ samples is also shown in Fig. 6.

The OFS aided FRCSM algorithm was also used to select the RBF equaliser having a model size of $M_{\text{spa}} = 30$ in conjunction with the training block length of $K = 600$ and the RBF variance of $\rho^2 = \sigma_n^2$. The BER performance of the resultant RBF equaliser is depicted in Fig. 9, in comparison to the performance of the Bayesian equaliser.

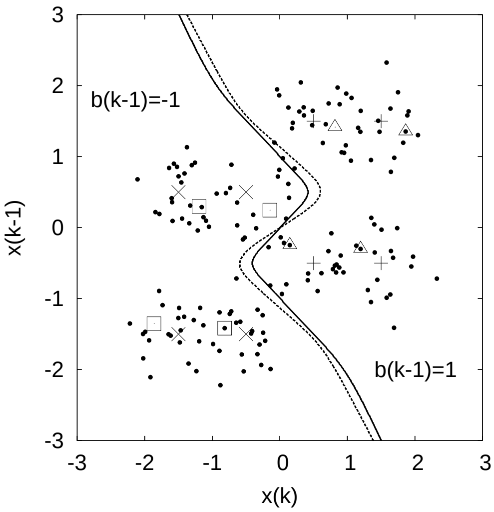

Figure 5: Comparison of the decision boundaries (solid: Bayesian and dashed: RBF) for the two-tap channel in conjunction with $n_e = 2$ and $\tau = 1$, given SNR= 10 dB, where pairs of cross and plus symbols indicate the positions of symmetric state pairs, and the pairs of square and triangular symbols indicate the positions of RBF centre pairs. The OFS aided FRCSM algorithm was used.

5. CONCLUSIONS

In this paper, we have investigated a novel symmetric RBF network structure designed for channel equalisation application. As a benefit of the underlying symmetry property of the optimal Bayesian equalisation solution, we were able to approach the optimal Bayesian performance using noisy training data. Both a block-data based and a sample-by-sample adaptive training algorithm have been proposed for this symmetric RBF equaliser, and the simulation results have indicated that they are capable of approaching the optimum Bayesian performance.

REFERENCES

[1] S. Chen and B. Mulgrew, "Overcoming co-channel interference using an adaptive radial basis function equaliser," *Signal Processing*, vol.28, no.1, pp.91–107, 1992.

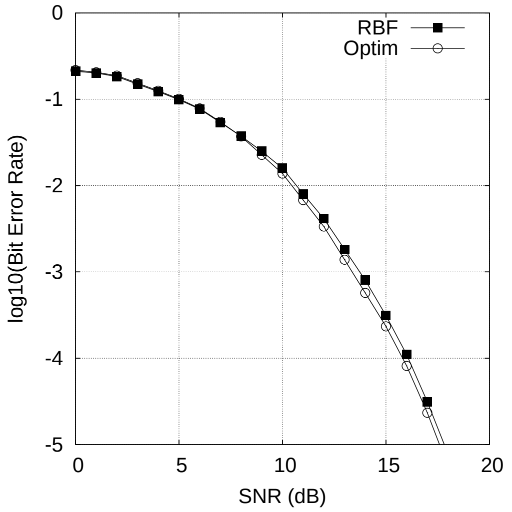

Figure 6: Comparison of BER performance for the three-tap channel with equaliser order $n_e = 4$ and decision delay $\tau = 2$. The NLBER algorithm was used.

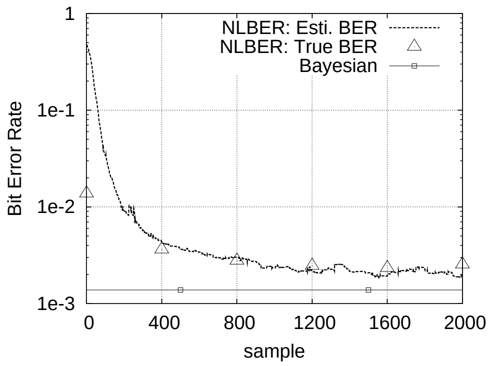

Figure 7: Learning curve of the NLBER algorithm averaged over 10 runs for the three-tap channel with $n_e = 4$ and $\tau = 2$, given SNR= 13 dB.

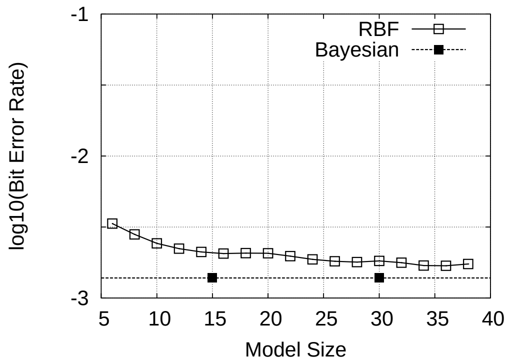

Figure 8: Influence of the model size to the BER performance for the three-tap channel with $n_e = 4$ and $\tau = 2$, given SNR= 13 dB. The NLBER algorithm was used.

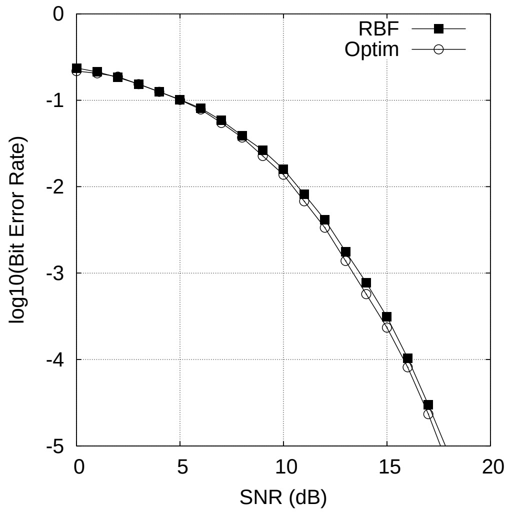

Figure 9: Comparison of BER performance for the three-tap channel with equaliser order $n_e = 4$ and decision delay $\tau = 2$. The OFS with FRCSM algorithm was used.

[2] 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.

[3] S. Chen, S. McLaughlin, B. Mulgrew and P.M. Grant, "Adaptive Bayesian decision feedback equaliser for dispersive mobile radio channels," *IEEE Trans. Communications*, vol.43, no.5, pp.1937–1946, 1995.

[4] 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.

[5] A. Wolfgang, S. Chen and L. Hanzo, "Radial basis function network assisted space-time equalisation for dispersive fading environments," *Electronics Letters*, vol.40, no.16, pp.1006–1007, 2004.

[6] S. Chen, B. Mulgrew and L. Hanzo, "Asymptotic Bayesian decision feedback equalizer using a set of hyperplanes," *IEEE Trans. Signal Processing*, vol.48, no.12, pp.3493–3500, 2000.

[7] K.Z. Mao, "RBF neural network center selection based on fisher ratio class separability measure," *IEEE Trans. Neural Networks*, vol.13, no.5, pp.1211–1217, 2002.

[8] S. Chen, L. Hanzo and A. Wolfgang, "Kernel-based nonlinear beamforming construction using orthogonal forward selection with Fisher ratio class separability measure," *IEEE Signal Processing Letters*, vol.11, no.5, pp.478–481, 2004.

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

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

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

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

[13] S. Chen, B. Mulgrew and L. Hanzo, "Least bit error rate adaptive nonlinear equalizers for binary signalling," *IEE Proc. Communications*, vol.150, no.1, pp.29–36, 2003.

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