MULTI-LEVEL RADIAL BASIS FUNCTION NETWORK BASED EQUALISERS FOR RAYLEIGH CHANNEL

M. S. Yee, L. Hanzo

Dept. of Electronics and Computer Science,
University of Southampton, SO17 1BJ, UK.
Tel: +44-703-593 125, Fax: +44-703-594 508
e-mail: lh@ecs.soton.ac.uk

Abstract – Radial Basis Function (RBF) network based channel equalisers have a close relationship with Bayesian schemes. Decision feedback is introduced in the design of the RBF equaliser in order to reduce its computational complexity. The RBF Decision Feedback Equaliser (DFE) was found to give similar performance to the conventional mean square error (MSE) DFE over Gaussian channels using various Quadrature Amplitude Modulation (QAM) schemes, while requiring a lower feedforward and feedback order. Over Rayleigh-fading channels similar findings were valid for binary modulation, while for higher order modems the RBF based DFE requires increased feedforward and feedback orders in order to outperform the conventional MSE DFE scheme.

I. RADIAL BASIS FUNCTION NETWORK EQUALISER WITH DECISION FEEDBACK

Background and Formalism

The background to neural networks has been documented for example by Haykin [1] and Bishop [2]. Various so-called Radial Basis Function (RBF) based equalisers were studied by Chen, McLaughlin, Mulgrew and Grant [3, 4] and the reader is referred to the above references for background reading. Below we introduce our basic formalism in the context of the RBF Decision Feedback Equaliser (DFE) following the approach of Chen, Mulgrew and McLaughlin [3, 5], while extending it to $\mathcal{M}$-ary modems. The characteristics of the transmitted sequence can be exploited by capitalising on the finite state nature of the channel and by considering equalisation as a geometric classification problem related to the complex signal constellation. We assume that the transmitted symbols are selected from the set of $\mathcal{M}$ complex values, $\mathcal{I}_i, i = 1, 2, \ldots, \mathcal{M}$, with equal probability. The symbol-spaced channel output can be defined by

$$\begin{aligned} v_k &= \sum_{n=0}^{L} f_n I_{k-n} + \eta_k \\ &= \tilde{v}_k + \eta_k \qquad -\infty \leq k \leq \infty, \end{aligned} \tag{1}$$

where $\{\eta_k\}$ is the additive Gaussian noise sequence with variance σ_η^2, $\{f_n\}$, $n = 0, 1, \ldots, L$ is the channel impulse response, $\{I_k\}$ is the channel input sequence and $\{\tilde{v}_k\}$ is the noise-free channel output.

The channel output observed by the linear m-th order equaliser can be written in vector form as $\mathbf{v}_k = \begin{bmatrix} v_k & v_{k-1} & \ldots & v_{k-m+1} \end{bmatrix}^T$, and hence we can say that the equaliser has an m-dimensional channel output observation space. For a channel impulse response of length $L+1$, there are hence $n_s = \mathcal{M}^{L+m}$ possible combinations of the channel input sequence

$$\mathbf{I}_k = \begin{bmatrix} I_k & I_{k-1} & \ldots & I_{k-m-L+1} \end{bmatrix}^T \tag{2}$$

that produce n_s different possible values of the noise-free channel output vector

$$\tilde{\mathbf{v}}_k = \begin{bmatrix} \tilde{v}_k & \tilde{v}_{k-1} & \ldots & \tilde{v}_{k-m+1} \end{bmatrix}^T. \tag{3}$$

The possible values or particular points in the observation space will be referred to as the desired channel states. Expounding further, we denote each of the n_s combinations of channel input sequence $\mathbf{I}_k$ of length $L+m$ symbols as $\mathbf{s}_i, 1 \leq i \leq n_s$. Corresponding to the channel input state $\mathbf{s}_i$ is the desired channel output state $\mathbf{r}_i$, where $i = 1, 2, \ldots, n_s$, i.e.,

$$\tilde{\mathbf{v}}_\mathbf{k} = \mathbf{r}_i \qquad \text{if } \mathbf{I}_k = \mathbf{s}_i, \qquad i = 1, 2, \ldots, n_s. \tag{4}$$

The desired channel states $\mathbf{r}_i$ can be partitioned into $\mathcal{M}$ classes according to the $\mathcal{M}$-ary value of the transmitted symbol $I_{k-\tau}$, as seen below:

$$\begin{aligned} &\mathbf{r}_j^i \in V_{m,\tau}^i = \{\tilde{\mathbf{v}}_k | I_{k-\tau} = \mathcal{I}_i\}, \\ &1 \leq i \leq \mathcal{M}, \qquad 1 \leq j \leq n_s^i, \end{aligned} \tag{5}$$

where the quantities n_s^i represent the number of channel states $\mathbf{r}_j^i$ in the set $V_{m,\tau}^i$ and τ is the equaliser decision delay.

For an equaliser with a feedback order of n, the n-symbol long feedback sequence denoted by the vector $\tilde{\mathbf{I}}_{feedback,k-\tau} = \begin{bmatrix} \tilde{I}_{k-\tau-1} & \ldots & \tilde{I}_{k-\tau-n} \end{bmatrix}^T$, where $\tilde{I}_{k-\tau}$ is the delayed detected symbol, can be employed to partition further the channel output states according to the

0-7803-5565-2/99/$10.00 © 1999 IEEE

$n_f = \mathcal{M}^n$ possible feedback state $\mathbf{s}_{f,j}, j = 1, \ldots, n_f$ as follows:

$$\mathbf{r}^i_{j,l} \in V^i_{m,\tau,j} = \{\tilde{\mathbf{v}}_k | I_{k-\tau} = \mathcal{I}_i \cap \tilde{\mathbf{I}}_{feedback,k-\tau} = \mathbf{s}_{f,j}\},$$
$$1 \le i \le \mathcal{M}, \qquad 1 \le j \le n_f, \qquad 1 \le l \le n^i_{s,j}. \quad (6)$$

The equaliser that minimizes the average error probability in symbol detection must operate according to Bayes' decision rule and thus the optimal Bayesian equalisation solution with the knowledge of the feedback state is given by [5]

$$\tilde{I}_{k-\tau} = \mathcal{I}^*_i, \qquad \text{if } \zeta^*_i(k) = \max\{\zeta_i(k), 1 \le i \le \mathcal{M}\}, \quad (7)$$

where $\zeta_i(k)$ is the decision variable based on the conditional density function given by:

$$\begin{aligned} \zeta_i(k) &= P(\mathbf{v}_k | I_{k-\tau} = \mathcal{I}_i \cap \tilde{\mathbf{I}}_{feedback,k-\tau} = \mathbf{s}_{f,j}) \\ &= \sum_{l=1}^{n^i_{s,j}} p^i_{j,l} p(\mathbf{v}_k - \mathbf{r}^i_{j,l}), \\ & \quad 1 \le i \le M, \qquad 1 \le j \le n_f. \end{aligned} \quad (8)$$

The term $p^i_{j,l}$ represents the *a priori* probability of appearance of each desired state $\mathbf{r}^i_{j,l} \in V^i_{m,\tau,j}$ and $p(\cdot)$ is the probability density function. If the noise distribution is Gaussian, the probability density function is given by:

$$p(\mathbf{x}) = (2\pi\sigma^2_\eta)^{-m/2} exp(-\|\mathbf{x}\|^2 / 2\sigma^2_\eta). \quad (9)$$

The relationship between the RBF network, detailed in References [1, 2], and the Bayesian equalisation solution expressed in Equation 7, 8 and 9 can be given explicitly. There are $\mathcal{M}$ RBF 'subnets' for $\mathcal{M}$ decision variables $\zeta_i(k), i = 1, \ldots, \mathcal{M}$. The term $p^i_{j,l}$ is the weight of the network, the function $p(\cdot)$ is the basis function of the network, $\mathbf{r}^i_{j,l}$ is the RBF's center, $2\sigma^2_\eta$ is the RBF's width, $n^i_{s,j}$ is the number of RBF hidden nodes and $\mathbf{v}_k$ is the input to the RBF network. Having introduced our basic formalism, let us now evaluate the performance effects of the equaliser paramters in the next Section.

Effect of Equaliser Parameters

The oldest corresponding symbol that influences the decision at the kth signalling instant, which produces the detected symbol $\tilde{I}_{k-\tau}$ is $I_{k-m-L+1}$, as seen in Equation 2. The oldest feedback symbol is $\tilde{I}_{k-\tau-n}$. Therefore, it is sufficient to employ a feedback order of $n = \log_2 n_f = L + m - 1 - \tau$, because this will enable us to influence decisions over the memory duration of the concatenated channel and the feedforward RBF section. Assuming $m > \tau + 1$ and $n = L + m - 1 - \tau$, Chen *et. al.* [5] mathematically proved that the Bayesian DFE of a feedforward order of $m = \tau + 1$ has the same conditional decision variables as those with feedforward order $m > \tau + 1$. Thus, given the equaliser decision delay τ, the feedforward order $m = \tau + 1$ is sufficient [5]. Overall, the equaliser delay τ specifies the number of channel states n_s, required for computing the decision variables and thus determines the computational complexity. A pragmatic rule is to set the delay to $\tau = L$ [5]. However, note that increasing the decision delay τ and feedforward order m will improve the performance of the RBF equaliser, as demonstrated in Figure 1 for Binary Phase Shift Keying (BPSK) at the expense of increasing the computational complexity exponentially, since the number of desired channel states n_s increases exponentially with m.

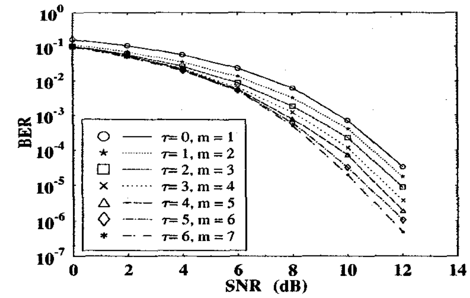
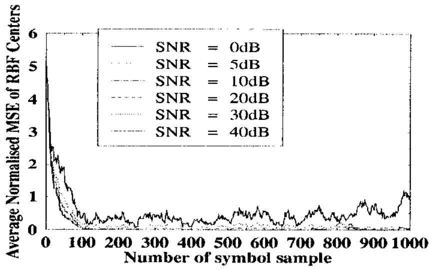


Figure 1: Bit error ratio (BER) versus SNR performance of the BPSK RBF equaliser with correct decision feedback upon varying the decision delay τ over the dispersive two-path AWGN channel with transfer function $F(z) = 0.707 + 0.707z^{-1}$. The equaliser feedforward order m was fixed to $m = \tau + 1$ and the feedback order n was varied according to $n = L + m - 1 - \tau$.

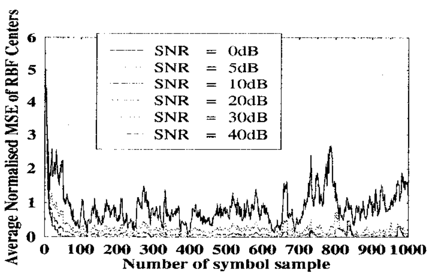
II. CHANNEL STATE ESTIMATION BY CLUSTERING

The knowledge of the noise-free channel outputs is essential for the determination of the decision function. Thus the channel state has to be 'learned' or inferred during the equaliser training period, when the transmitted symbols are known to the receiver. This can be achieved typically in two ways [5] – by invoking channel estimation methods and using the resulting channel estimate to calculate the channel states or by employing clustering algorithms that identify the channel states directly. Chen *et. al.* [3] uses the second approach, but they estimate the 'scalar channel states' with the clustering algorithm instead of the m-element vector channel states adopted in [5]. In this paper we use Chen *et. al.*'s approach [3], which we will refer to as the scalar center clustering algorithm. Let us first introduce the scalar channel states.

Referring back to Equation 1 and Equation 3, the lth element $\tilde{v}_{k-l}$ of the noise-free channel output vector $\tilde{\mathbf{v}}_k$, where $l = 0, 1, \ldots, m-1$, corresponds to the convolution of a sequence of $L+1$ transmitted symbols and the $L+1$ channel impulse response taps. The number of transmitted symbols contributing to the value of $\tilde{v}_{k-l}$ is $L+1$ and we represent these transmitted symbol as an $L+1$ element vector $\mathbf{I}_{scalar,k} = \begin{bmatrix} I_k & \ldots & I_{k-L} \end{bmatrix}^T$.



(a) $\mu_c = 0.1$



(b) $\mu_c = 0.2$

Figure 2: The MSE of the BPSK RBF equaliser centers versus transmitted symbol index for various learning rates μ_c of the centers using the scalar center clustering algorithm over the two-path channel environment with transfer function $F(z) = 0.707 + 0.707z^{-1}$. The equaliser had 5 feedforward taps and a decision delay of 2 symbols, and the number of training symbols was 700.

The number of different possible combinations of the symbol sequence in $\mathbf{I}_{scalar,k}$ is $n_{s,f} = \mathcal{M}^{L+1}$ for an $\mathcal{M}$-ary modulation scheme. We represent these transmitted symbol combinations as the channel input states $\mathbf{s}_{scalar,i}$, where $i = 1, 2, \ldots, n_{s,f}$. Each of these channel input states $\mathbf{s}_{scalar,i}$ corresponds to a *scalar channel output state* $r_i, i = 1, 2, \ldots, n_{s,f}$. Thus, the noise-free channel output $\tilde{v}_k$ can assume any of the $n_{s,f}$ scalar channel output states r_i, depending on $\mathbf{I}_{scalar,k}$, i.e., $\tilde{v}_k = r_i$ if $\mathbf{I}_{scalar,k} = \mathbf{s}_{scalar,i}$. The scalar states $r_i, i = 1, 2, \ldots, n_{s,f}$, can be suitably combined to form the vector states $\mathbf{r}_j, j = 1, 2, \ldots, n_s = \mathcal{M}^{m+L}$, mentioned in Equation 4.

The scalar channel state r_i can be estimated using a clustering procedure as follows [3]:

$$\text{if } \mathbf{I}_{scalar,k} = \mathbf{s}_{scalar,i} \text{ then} \quad r_{i,k} = r_{i,k-1} + \mu_c \cdot (v_k - r_{i,k-1}), \tag{10}$$

where μ_c is the learning rate for the channel states. For time-varying channels, it is necessary to continuously update r_i during data transmission. This can be achieved using the following decision-directed version of Equation 10:

$$\text{if } \tilde{\mathbf{I}}_{scalar,k-\tau} = \mathbf{s}_{scalar,i}, \text{ then} \quad r_{i,k} = r_{i,k-1} + \mu_c \cdot (v_{k-\tau} - r_{i,k-1}), \tag{11}$$

where $\tilde{\mathbf{I}}_{scalar,k-\tau} = \begin{bmatrix} \tilde{I}_{k-\tau} & \ldots & \tilde{I}_{k-\tau-L} \end{bmatrix}^T$. In order to quantify the centers' training performance, we define the average normalised MSE of the vector centers at signalling interval k as:

$$\mathrm{MSE}(\mathbf{c}, k) = \frac{1}{n_s \sigma_{\tilde{v}}^2} \sum_{i=1}^{n_s} \|\mathbf{c}_{i,k} - \mathbf{c}_{i,opt}\|^2, \tag{12}$$

where n_s is the number of centers, $\sigma_{\tilde{v}}^2$ is the variance of the noise-free received signal, while $\mathbf{c}_{i,k}, i = 1, \ldots, n_s, k = 0, \ldots, \infty$ represents the ith assumed RBF center at signalling interval k and $\mathbf{c}_{i,opt}$ is the vector associated with the ith desired or assumed 'true' RBF center.

Figure 2 shows the average normalised MSE of the RBF centers using the scalar clustering algorithm. As the learning rate μ_c is increased from 0.1 to 0.2, the centers converge faster to their desired positions, but as expected, the MSE curves of the centers becomes more spurious, especially at low SNRs, as we can see from Figure 2. We recommend using a variable learning rate μ_c, where μ_c is set to a higher rate during the training mode so that it converges faster and it is set to lower rate during the decision-directed learning mode to reduce the spuriousity of the center MSE.

We note that the number of channel states $n_{s,f} = \mathcal{M}^{L+1}$ grows exponentially with the number of symbol levels $\mathcal{M}$ used in the modulation scheme. Thus, the convergence rate is dependent on the type of modulation scheme used. This presents a problem, when we use the scalar center clustering algorithm to train the RBF DFE, since the number of $\mathcal{M}$-ary training symbols required for higher order QAM increases exponetially, as more scalar centers $n_{s,j}$ have to be trained. A solution is to use channel estimation methods to train the equaliser, where the number of training symbols required will only be dependent on the channel impulse response length. Having considered the issues of RBF center-training, let us now focus our attention on performance issues in the next Section.

III. SIMULATION RESULTS

Performance of the RBF Equaliser with Square-QAM over Gaussian Channels

In this Section the performance of the RBF DFE is analysed in conjunction with multilevel modulation schemes in a Gaussian environment. We used the so-called maximum-minimum distance square-shaped QAM constellations [6]. In all our simulations, the transmitted symbol I_k was an equiprobable $\mathcal{M}$-QAM symbol. Therefore the weights of the RBF network were fixed to +1

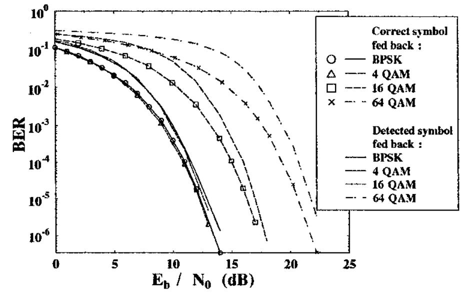


Figure 3: BER versus E_b/N_0 performance of the RBF DFE over the dispersive two-path AWGN channel with transfer function $F(z) = 0.707 + 0.707z^{-1}$ for different $\mathcal{M}$-QAM schemes.

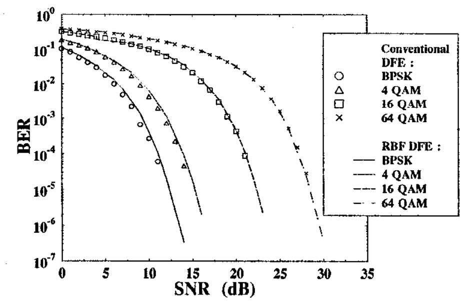


Figure 4: BER versus SNR performance of the RBF DFE and conventional DFE over the dispersive **two-path AWGN channel** with transfer function $F(z) = 0.707+0.707z^{-1}$ for different $\mathcal{M}$-QAM schemes. Correct symbols were fed back.

[4]. The noise variance σ_η^2 was fixed to unity, while the power of the transmitted symbol was varied according to the SNR. Initially the centers of the RBF network were positioned at the desired channel states. The width of the RBF network was set to $2\sigma_\eta^2$. The RBF DFE used in our simulations had a feedforward order of $m = 2$, feedback order of $n = 1$ and decision delay of $\tau = 1$.

Figure 3 shows the BER performance of the RBF DFE for BPSK, 4-, 16- and 64-QAM schemes with correct and actual detected symbol feedback. The SNR degradation due to decision errors is about 0.5dB for BPSK and 4-QAM, 1dB for 16-QAM and 1.5dB for 64-QAM at a BER of 10^{-4} and thus has little effect at low BERs. Note however the E_b/N_0 degradation increase as the BER increases, which becomes more significant for higher order QAM. Figure 4 shows the performance comparison between the conventional DFE and the RBF DFE over the dispersive two-path Gaussian channel with transfer function $F(z) = 0.707+0.707z^{-1}$. The coefficients of the conventional DFE were optimised using the MSE criterion [7] and assumming perfect channel estimation. The parameters of the conventional DFE, i.e. m, n and τ, were chosen such that it exhibited the best performance for our simulation scenario and hence any increase of the feedforward order would not give significant performance improvement. Explicitly, the conventional DFE used in our simulations had a feedforward order of $m = 7$, feedback order of $n = 1$ and decision delay of $\tau = 7$ symbols. The RBF DFE is found to give similar performance with a feedforward order as low as 2, feedback order of 1 and decision delay of 1 symbol. The performance of the RBF DFE can still be further improved quite significantly by increasing both the decision delay τ and the feedforward order m, as it was demonstrated in Figure 1 for BPSK.

Performance of the RBF Equaliser over wideband Rayleigh Fading Channels

In this Section, we analyse the performance of the $\mathcal{M}$-QAM RBF DFE over dispersive Rayleigh fading channels. The baseband multipath fading channel was time-invariant, represented as follows:

$$c(t) = \sum_{i=0}^{n_c} f_i(t)\delta_{t-\tau_i(t)}, \tag{13}$$

where n_c is the number of fading paths, $f_i(t)$ is the complex-valued ith channel tap at time t, $\tau_i(t)$ is the excess delay at time t and δ_t is a delta function located at signalling instant t. The multipath components $f_i(t)$ have independent Rayleigh fading statistics, they are uncorrelated and are scaled by their designed weights. For a more in depth charaterization of Rayleigh fading channels, the reader is referred to [6]. The fading channel parameters used in our simulations are given in Table 1 and we employed two symbol-spaced fading paths with the weights given by $0.707 + 0.707z^{-1}$. The training symbol sequence was implemented as a preamble. In our simulations, the number of training symbols L_T was set to 27 and the number of data symbols L_D was set to 144.

Figures 4 and 5 provide the BER performance comparisons between the conventional DFE and the RBF DFE for different $\mathcal{M}$-QAM schemes over AWGN and Rayleigh channels, respectively. The conventional DFE assumed perfect channel estimation and its equaliser coefficients were optimised using the MSE criterion [7]. The centers of the RBF DFE were positioned at the desired channel states for the *perfect* RBF DFE. In the Rayleigh-fading simulations, the channel impulse response was kept constant for the duration of the transmitted burst. From Figure 5, we note that for BPSK, the RBF DFE with a feedforward order as low as $m = 2$, feedback order of $n = 1$ and decision delay of $\tau = 1$ symbol was found to give similar performance to the conventional DFE with a feedforward order of $m = 7$, feedback order of $n = 1$ and decision delay of $\tau = 7$ symbols. For 4 QAM, 16 QAM and 64 QAM, the RBF DFE with the same parameters gave inferior performance compared to the conventional DFE in the two-path Rayleigh fading channel scenario.

Transmission Frequency	1800MHz
Transmission Rate	133kBd
Vehicular Speed	30 mph
Normalised Doppler Frequency	6×10^{-4}

Table 1: Simulation parameters of the Rayleigh fading channel

This is unlike the performance of the two-path AWGN channel scenario shown in Figure 4. The performance degradations for the higher order modulation schemes are higher under fading channel conditions, than over AWGN channels, despite perfect channel estimation, because QAM schemes are more sensitive to fades due to their reduced Euclidean distance between their neighbouring constellation points. Nevertheless, the performance of the RBF DFE can be improved by increasing both the decision delay τ and the feedforward order m, as we discussed in previous Section, at the expense of increased computational complexity.

The adaptive performance of the RBF DFE was investigated over the 2-path Rayleigh fading channel for BPSK. In our simulations for the adaptive RBF DFE, we used a variable center learning rate μ_c, where we had $\mu_c = 0.3$ during the training mode and $\mu_c = 0.1$ during the decision-directed learning mode. The channel fade was kept constant during a transmission burst, or in other words, the fading was frame-invariant. We assigned a sequence of $L_T = 27$ pseudo-random binary symbols to the training symbol sequence in the transmission burst. The performance curves of the BPSK RBF DFE in Figure 5 provide the comparison of the RBF DFE with perfect channel estimation and that when the adaptive RBF DFE was trained with the aid of the scalar center clustering algorithm. The clustering algorithm was able to train the RBF equaliser effectively in the studied simulation scenario.

IV. CONCLUSION

This paper presented the design of a RBF DFE based on the Bayesian equaliser solution. The RBF DFE was found to give equivalent BER performance to the conventional MSE DFE with a lower equaliser order and delay for stationary Gaussian channels. The BER performance of the RBF DFE can be further improved by increasing both the equaliser order and decision delay, at the expense of exponentially increasing the computational complexity. We also note that the performance degradations for the higher order modulation schemes are higher than those for BPSK under fading channel conditions, even with perfect channel estimation, since these constellations are more sensitive to fades due to their reduced Euclidean distance between their neighbouring constellation points. Our future work is targeted at improving the system's performance using novel residual number system based error correction codes, turbo codes and at invoking the RBF-DFE in joint-detection code division multiple access schemes.

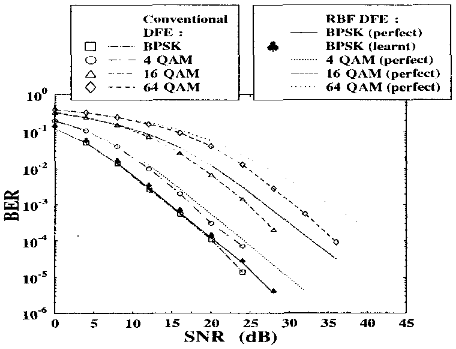


Figure 5: BER versus SNR performance of the conventional DFE and the RBF DFE over the **two equal weight, symbol-spaced path Rayleigh fading channel** of Table 1 for different $\mathcal{M}$-QAM modulation schemes. The conventional DFE assumed perfect channel estimation. The centers of the RBF DFE were set to the desired channel states for the *perfect* RBF DFE curve and were 'learnt' using scalar clustering algorithm for the *learnt* RBF DFE curve.

V. ACKNOWLEDGEMENT

The financial support of the following organisations is gratefully acknowledged: EPSRC, UK, the Comission of the European Communities, Brussels, Belgium and Motorola ECID, Swindon, UK and the Virtual Centre of Excellence (VCE) in Mobile Communications.

REFERENCES

(1) S. Haykin, *Neural Networks : A Comprehensive Foundation.* Macmillan Publishing Company, 1994.

(2) C. M. Bishop, *Neural Networks for Pattern Recognition.* Oxford University Press, 1995.

(3) S. Chen, S. McLaughlin, and B. Mulgrew, "Complex-valued radial basis function network, part ii: Application to digital communications channel equalisation," *EURASIP Signal Processing*, vol. 36, pp. 175–188, March 1994.

(4) S. Chen, B. Mulgrew, and P. M. Grant, "A clustering technique for digital communications channel equalization using radial basis function networks," *IEEE Trans. Neural Networks*, vol. 4, pp. 570–579, July 1993.

(5) S. Chen, B. Mulgrew, and S. McLaughlin, "Adaptive bayesian equalizer with decision feedback," *IEEE Trans. Signal Processing*, vol. 41, pp. 2918–2927, September 1993.

(6) W. T. Webb and L. Hanzo, *Modern Quadrature Amplitude Modulation: Principles and Applications for Fixed and Wireless Channels.* London: IEEE – John-Wiley, 1994.

(7) J. G. Proakis, *Digital Communications.* McGraw-Hill, Inc., 1995.