Multilayer Perceptron-Based DFE with Lattice Structure

Azzedine Zerguine, *Member, IEEE*, Ahmar Shafi, and Maamar Bettayeb

***Abstract*—The severely distorting channels limit the use of linear equalizers and the use of the nonlinear equalizers then becomes justifiable. Neural-network-based equalizers, especially the multilayer perceptron (MLP)-based equalizers, are computationally efficient alternative to currently used nonlinear filter realizations, e.g., the Volterra type. The drawback of the MLP-based equalizers is, however, their slow rate of convergence, which limit their use in practical systems. In this work, the effect of whitening the input data in a multilayer perceptron-based decision feedback equalizer (DFE) is evaluated. It is shown from computer simulations that whitening the received data employing adaptive lattice channel equalization algorithms improves the convergence rate and bit error rate performances of multilayer perceptron-based DFEs. The adaptive lattice algorithm is a modification to the one developed by Ling and Proakis. The consistency in performance is observed in both time-invariant and time varying channels. Finally, it is found in this work that, for time-invariant channels, the MLP DFE out performs the least mean squares (LMS)-based DFE. However, for time-varying channels comparable performance is obtained for the two configurations.**

***Index Terms*—Decision feedback equalizer (DFE), lattice filters, multilayer perceptron (MLP).**

I. Introduction

A SERIOUS limitation in attempting to achieve a high transmission rate through a particular band-limited channel is the time dispersion suffered by the signal at the receiving end of this channel [1]. In data transmission, the time dispersion imparted on the transmitted signal results in a time overlap between successive symbols, known as intersymbol interference (ISI). Equalizers have been used to describe filters used to compensate for such distortions in the amplitude and delay characteristics of the channel.

Nonlinear equalizers [1], [2] are superior to linear ones in applications where the channel distortion is too severe for a linear equalizer to handle. In particular, a linear equalizer does not perform well on channels with deep spectral nulls in their amplitude characteristics or with nonlinear distortion.

A decision feedback equalizer (DFE) is a nonlinear equalizer that is widely used in situations where the ISI is very large. It has been proved theoretically and experimentally that the DFE performs significantly better than a linear equalizer of equivalent complexity [1]. The basic idea of DFE is that if the values of the symbols already detected are assumed to be correct, then the ISI contributed by these symbols can be canceled exactly by subtracting past symbol values with appropriate weighting from the equalizer output [2].

To further enhance the performance of the DFE, the multilayer perceptron (MLP) has been incorporated to the DFE. It is shown that the MLP-based DFE (MLP DFE) [3], using the back-propagation (BP) algorithm [4], gives a significantly improved performance over the simple DFE [2].

To extend the applicability of the MLP DFE to areas involving fast time-varying channels, e.g., mobile communication channels, the whitening procedure is a necessary preprocessing stage for the MLP DFE to be able to track variations in these environments. The whitening process can be achieved using a lattice filter as an input stage to the MLP DFE equalizer. In this case, the effect of the eigenvalue spread will be reduced substantially in both time-invariant and time-varying channels. The whitening procedure in this case can be considered as an approximation to second-order methods [5] since it is used for a nonlinear filter, that is the MLP DFE.

In this work, the performance of the MLP DFE using a whitening scheme in its input is evaluated, where the BP algorithm, rather than the recursive least squares (RLS) [6], is used here to update the MLP DFE as the former has a lower complexity than the latter one. It is shown that a great improvement in performance is obtained through the use of this technique over both the simple DFE and the MLP DFE in both time-invariant and time-varying channels. The whitening process is carried out through a simple modification of the adaptive lattice DFE originally proposed in [7]. Preliminary results of this work were reported in [8]. In this work we report on further results covering the aspect of the behavior of the proposed technique to time-varying channels and other parameters related to the length of the lattice DFE filter and the size of the MLP configuration in time-invariant channels.

The rest of this paper is organized as follows. In Section II, a brief review of artificial neural networks is given where the proposed new structure of the MLP-based DFE is presented along with the modification to the adaptive lattice DFE [7]. The computational complexity of the proposed algorithm is treated in Section III, while its performance is demonstrated by the simulation results of Section IV. Section V is concerned with the discussion of the results and conclusions.

II. Artificial Neural Networks

Because of the capabilities of artificial neural networks in efficiently modeling arbitrary nonlinearities, there has been

Manuscript received December 7, 1999; revised July 7, 2000, December 6, 2000, and February 7, 2001. This work was supported by KFUPM.

A. Zerguine and A. Shafi are with the Electrical Engineering Department, King Fahd University of Petroleum and Minerals, Dhahran 31261, Saudi Arabia (e-mail: azzedine@kfupm.edu.sa; ahmar@kfupm.edu.sa).

M. Bettayeb is with the Electrical and Electronics Engineering Department, University of Sharjah, Sharjah, United Arab Emirates (e-mail: maamarab@sharjah.ac.ae).

Publisher Item Identifier S 1045-9227(01)03570-6.

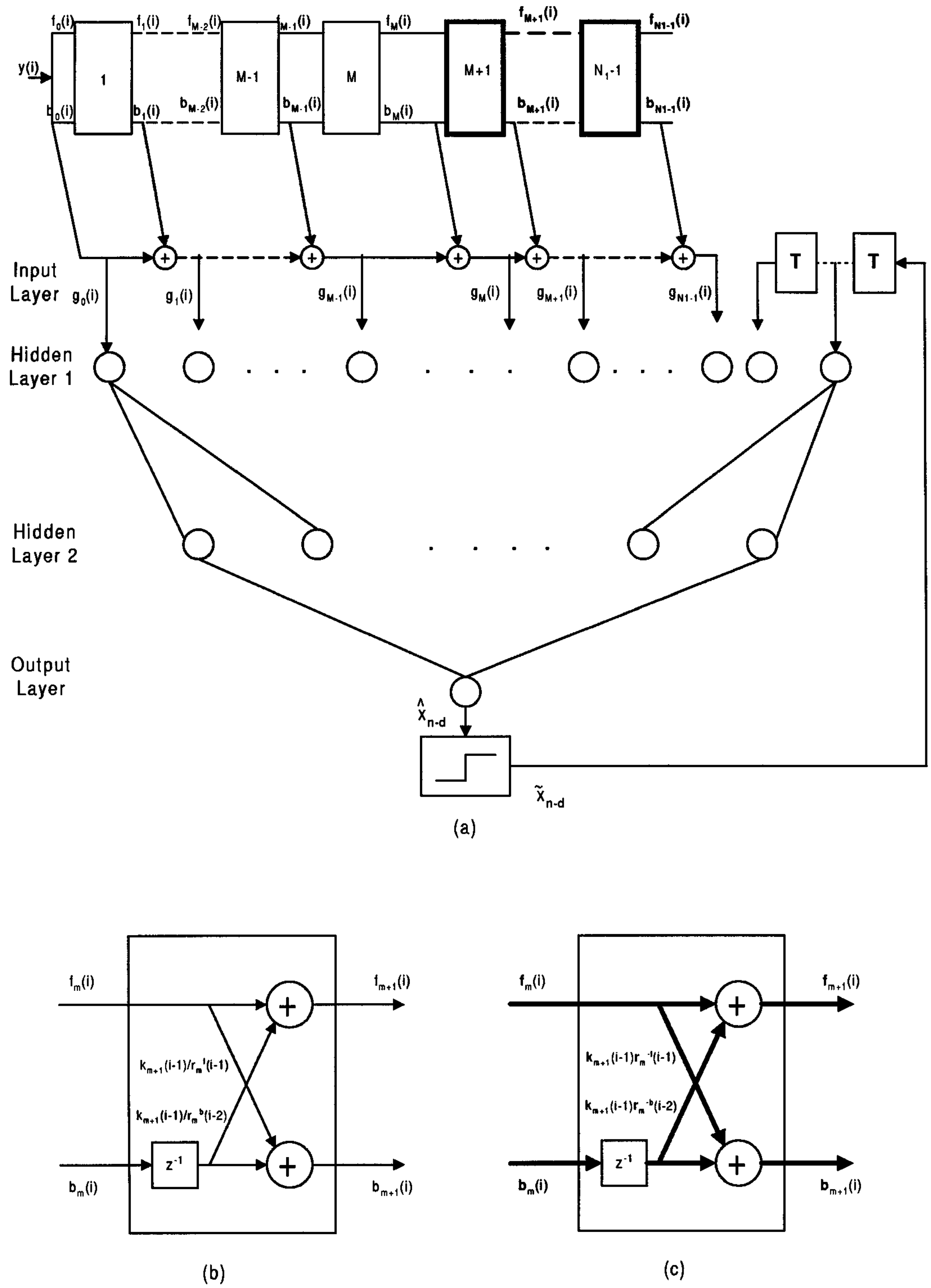


Fig. 1. (a) Block diagram of overall system lattice filter MLP DFE. (b) Single channel lattice (stage m). (c) Two channel lattice (stage m).

recent interest in employing them in adaptive equalization for data communication channels [3], [9], [10]. In this case, the linear adaptive filter is replaced by a neural network. Different artificial neural-network architectures such as multilayer perceptron, radial basis functions, and recurrent neural networks have all been proposed in the literature for channel equalization [11]. Among all these structures, the most commonly and widely used is the MLP structure. The popularity of MLP-based equalizers is due in part to their computational simplicity, finite parameterization, stability, and smaller structure size for a particular problem as compared to other structures.

A multilayer perceptron consists of several hidden layers of neurons that are capable of performing complex, nonlinear mappings between the input and output layer. The hidden layers provide the capability to use the nonlinear sigmoid function (to be defined later) to create intricately curved partitions of space with complex nonlinear decision boundaries [12]. Furthermore, it has

TABLE I
COMPUTATIONAL COMPLEXITY OF ADAPTIVE DFE ALGORITHMS

	DFE	MLP DFE	MLP DFE with lattice filter
Total number of additions and multiplications	$2(N_1+N_2)+1$	$N_1+N_2+11L_1$ $+17L_2+5(N_1+N_2)\times L_1$ $+6L_1\times L_2+9$	$19N_1+40N_2+11L_1$ $+17L_2+5(N_1+N_2)\times L_1$ $+6L_1\times L_2-30$
Divisions	0	$2L_1+2L_2+2$	$2N_1+2L_1+2L_2+2$

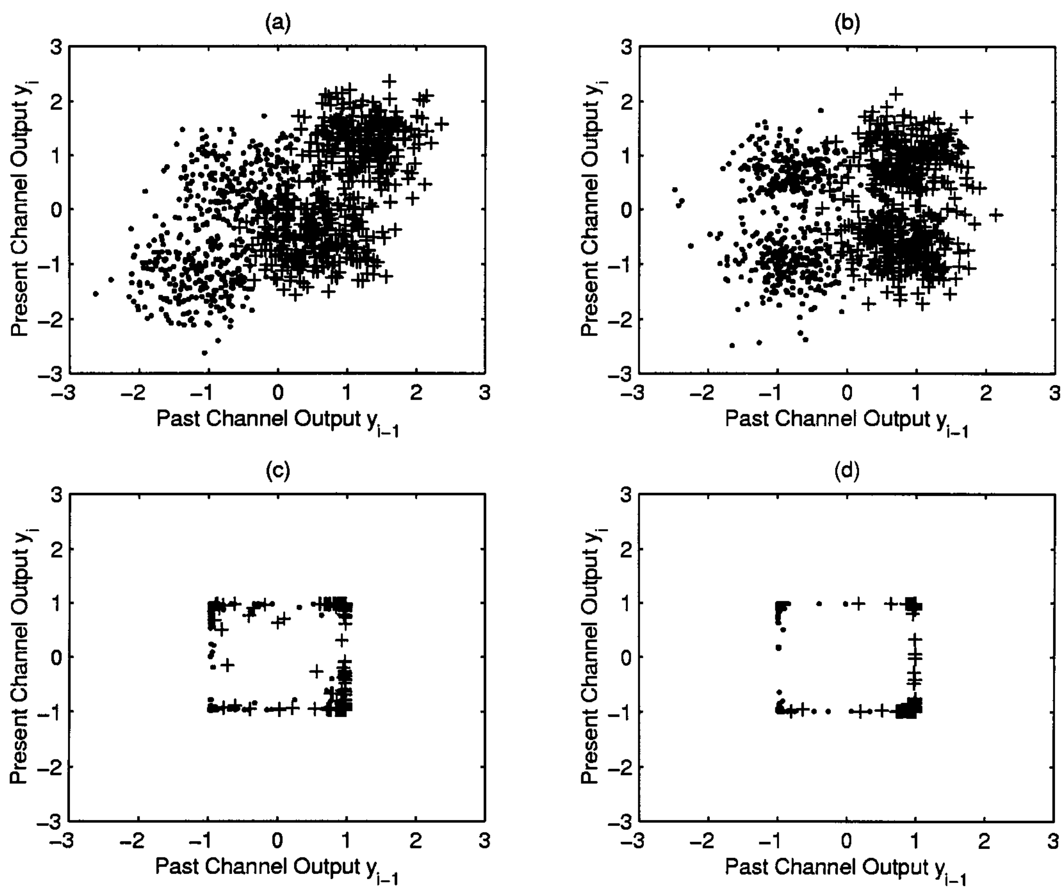


Fig. 2. Scattering diagrams of different types of equalizers with channel $H_1(z) = 0.3482 + 0.8704z^{-1} + 0.3482z^{-2}$ under SNR $= 10$ dB: (a) Unequalized data. (b) Equalization using the simple DFE. (c) Equalization using the MLP DFE. (d) Equalization using the lattice-based MLP DFE.

been shown that only three layers are needed by the MLP to generate these boundaries [13].

The basic element of the multilayer perceptron is the neuron. Each neuron in the layer has primary local connections and is characterized by a set of real weights $[w_{1j}, w_{2j}, \ldots, w_{Nj}]$ applied to the previous layer to which it is connected and a real threshold level I_j. The jth neuron in the pth layer accepts real inputs $v_h^{(p-1)}(h = 1, 2, \ldots, N)$ from the $(p-1)$th layer and produces an output $v_j^{(p)}$, which is also a real scalar, expressed in the following way:

$$v_j^{(p)} = f_j\left(\sum_{h=1}^{N} w_{hj} v_h^{(p-1)} + I_j^{(p)}\right). \tag{1}$$

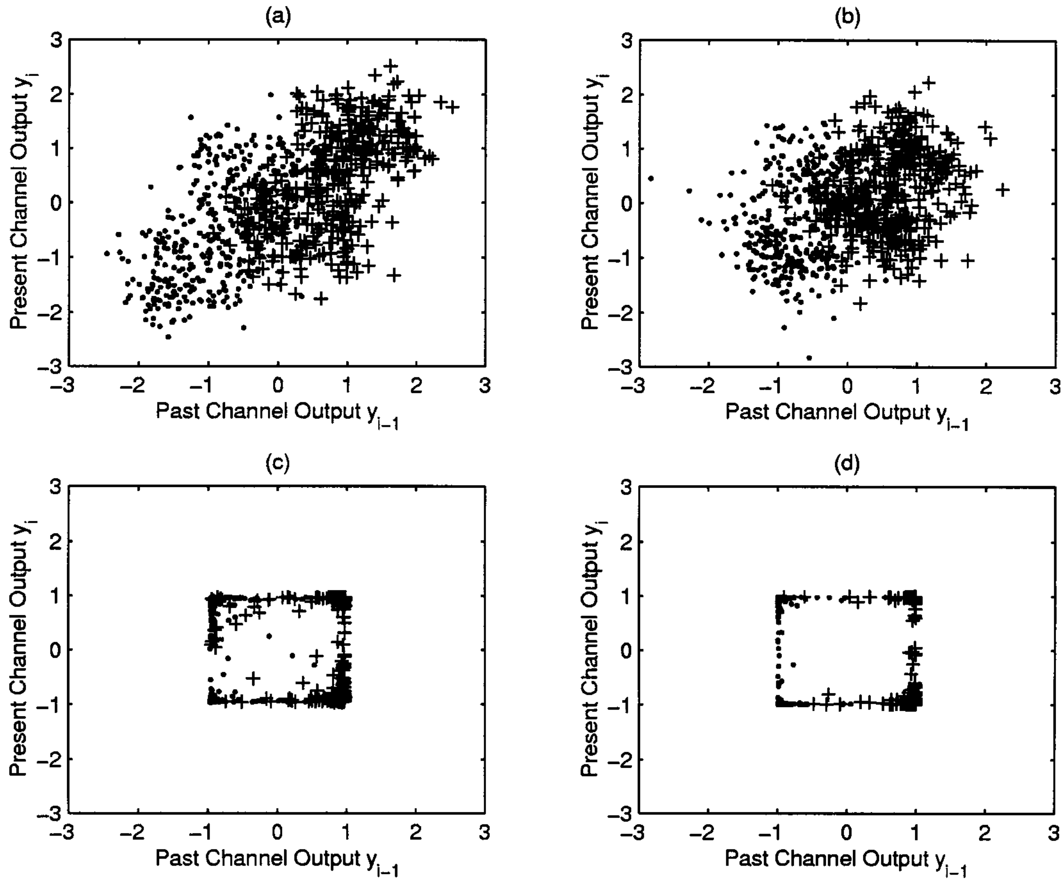


Fig. 3. Scattering diagrams of different types of equalizers with channel $H_2(z) = 0.408 + 0.816z^{-1} + 0.408z^{-2}$ under SNR $= 10$ dB: (a) Unequalized data. (b) Equalization using the simple DFE. (c) Equalization using the MLP DFE. (d) Equalization using the lattice-based MLP DFE.

This output value $v_j^{(p)}$ serves also as input to the $(p+1)$th layer (next layer) to which the neuron is connected. In the above expression, $f_j(.)$ represents the nonlinearity function. The most commonly used one in the perceptron is of the sigmoid type, defined as [13]

$$f_j(x) = \frac{1-e^{-x}}{1+e^{-x}} \tag{2}$$

where $f_j(x)$ is always in the range $[-1, 1]$, $\forall x \in \mathbf{R}$ (the set of real numbers). The weights $\{w_{hj}\}$ and thresholds levels $\{I_j\}$ are updated during training [3].

The MLP did not receive much attention in applications until the introduction of the BP algorithm [14]. The BP algorithm was used in both linear equalizers [15] and nonlinear equalizers (DFE) [3], and it was found that in both cases, the neural—network-based configuration out performed its nonneural network-based counterpart. In this work, however, only the MLP DFE [3] will be considered as it is more advantageous than its linear counterpart.

A. The Learning Phase

In the BP algorithm, the output value is compared with the desired output, resulting in an error signal. This error signal is fed back through the network whose weights are adjusted to minimize this error. The increments used in updating the weights, Δw_{hj}, and threshold levels, ΔI_j, of the pth ($p \in [1, 2, \ldots, P]$) layer are updated, respectively, according to the following relations:

$$\Delta w_{hj}^{(p)}(i+1) = \eta\delta_j^{(p)}v_j^{(p-1)} + \alpha\Delta w_{hj}^{(p)}(i) \tag{3}$$

and

$$\Delta I_j^{(p)}(i+1) = \beta\delta_j^{(p)} \tag{4}$$

where

- η learning gain;
- α momentum parameter;
- β threshold level adaptation gain.

The error signal $\delta_j^{(p)}$ for layer p is calculated starting from the output layer P, as

$$\delta_j^{(P)} = \frac{\left(z_j - v_j^{(P)}\right)\left(1 - v_j^{2(P)}\right)}{2} \tag{5}$$

and is then recursively backpropagated to lower layers ($p \in [1, 2, , P-1]$) according to

$$\delta_j^{(p)} = \left(1 - v_j^{2(p)}\right)\sum_l \frac{\delta_l^{(p+1)}w_{jl}^{(p+1)}}{2}, \tag{6}$$

where l is over all neurons above the neuron j in the $(p+1)-st$ layer and z_j is the desired output. To allow for a rapid learning,

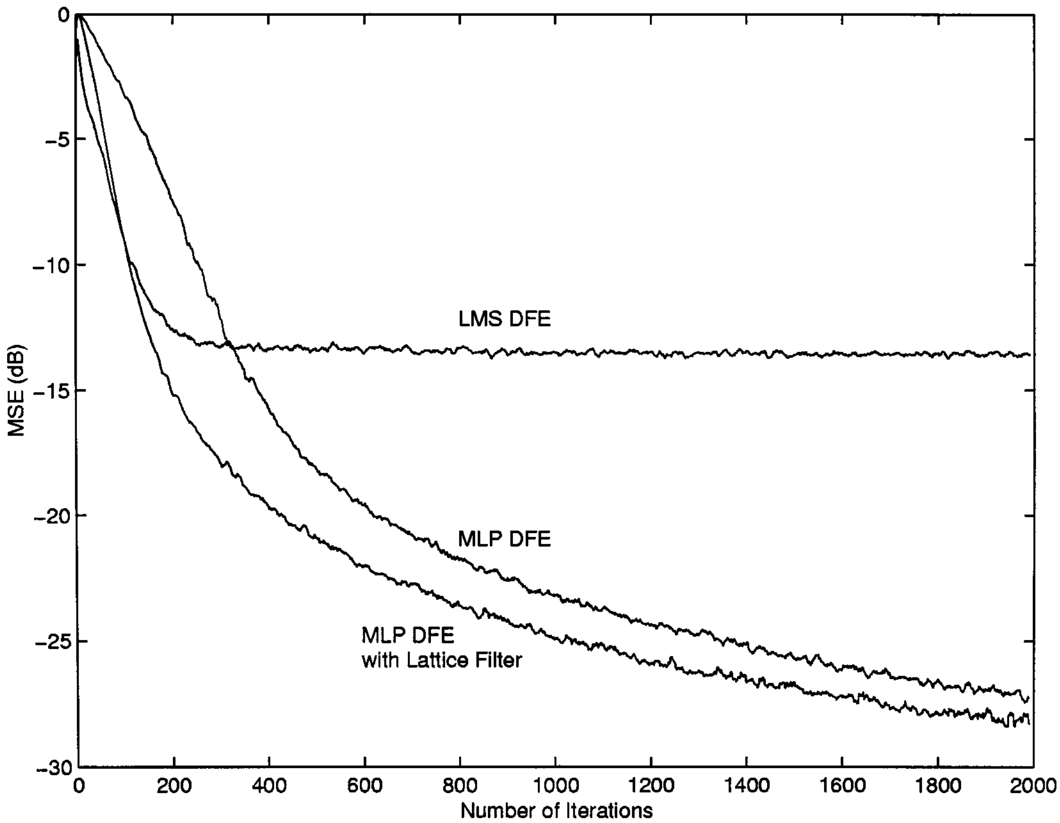


Fig. 4. Learning curves of different types of equalizers with channel $H_1(z) = 0.3482 + 0.8704z^{-1} + 0.3482z^{-2}$ under SNR $= 20$ dB.

a momentum term, $\Delta w_{hj}^{(p)}(i)$, scaled by α, is used to filter out high-frequency variations of the weight vector. Consequently, the convergence rate is much faster and the fast weight changes are smoothed out.

B. Proposed Scheme

Because the BP algorithm is no more than a generalized least mean squares (LMS) algorithm [16], it then suffers from the same problems as the LMS algorithm, particularly the slow rate of convergence when applied to channels with spectral nulls in their frequency responses. These channels are known to yield a large eigenvalue spread of the autocorrelation matrix of the signal at their outputs. This kind of channel characteristics is often encountered in time-variant channels. Moreover, as seen later, the BP-based MLP DFE has an equal performance to that of the simple DFE in time varying channels. The MLP DFE equalizer can then hardly be justified for the equalization of such channels. Eventually, a whitening procedure must be used to improve the performance of the MLP DFE with the BP algorithm as an updating scheme.

Since lattice equalizers are known to be insensitive to the eigenvalue spread of the channel autocorrelation matrix [17], [18], adaptive lattice (AL) algorithms can generate a set of orthogonal signal components that can be used as inputs to the equalizer lending themselves to fast convergence. Ling and Proakis [7] derived a generalized least squares multichannel lattice DFE (L-DFE) algorithm, implemented and investigated it on both time-invariant and time-varying channels. The results showed that the L-DFE has a much better performance than the simple DFE.

If this scheme is used as a prewhitening process to the MLP DFE, the performance of the MLP DFE in terms of convergence rate, BER, and steady-state MSE will therefore be expected to improve on both time-invariant and time-varying channels. Fig. 1 gives the details of this new configuration where the L-DFE [7] is used as a whitening scheme for the MLP DFE [3]. The L-DFE consists of a lattice predictor part and a joint estimator part. The predictor has $M = N_1 - N_2$ scalar stages followed by $N_2 - 1$ two-dimensional (2-D) stages. The L-DFE input is the received signal $y(i)$, while its outputs $\{g_m(i)\}$ and delayed versions of the decision device output are used as inputs to the MLP DFE which has one input layer, two hidden layers and one output layer. As shown in Fig. 1, the outputs of the L-DFE are used as inputs to the feedforward section of the lattice-based MLP DFE, and depending on the values of N_1 and N_2, the L-DFE will then consist of either a scalar stage or a 2-D one. However, it should be noted that the neural network carries out solely the estimation of the signal.

The outputs $\{g_m(i)\}$ that account for the modification to L-DFE [7] are detailed as follows. During the initialization process, it is defined as

$$g_o(i) = 0. \tag{7}$$

As for the scalar stages ($0 < m \leq N_1 - N_2$), they are given by

$$g_m(i) = g_{m-1}(i) + k_m^g(i-1)b_{m-1}(i)/r_{m-1}^b(i-1). \tag{8}$$

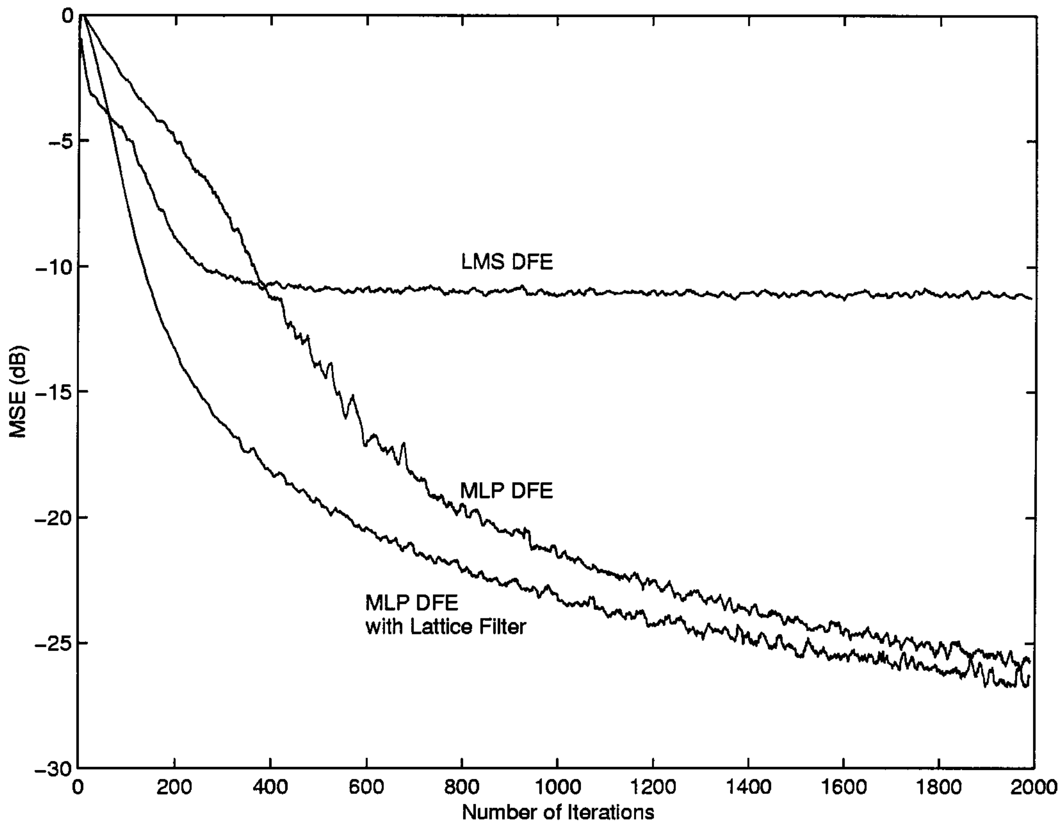


Fig. 5. Learning curves of different types of equalizers with channel $H_2(z) = 0.408 + 0.816z^{-1} + 0.408z^{-2}$ under SNR $= 20$ dB.

Finally, in the 2-D stages $(N_1 - N_2 < m < N_1)$, they are expressed as

$$g_m(i) = g_{m-1}(i) + \mathbf{k}_m^{gT}(i-1)\mathbf{r}_{m-1}^{-b}(i-1)\mathbf{b}_{m-1}(i) \quad (m \leq N_1) \tag{9}$$

where all the terms in the above expressions are defined in the Appendix which gives all the details of the algorithm.

III. Computational Complexity of the Algorithms

Most common operations used in adaptive filtering algorithms are additions, multiplications, and divisions. When digitally implemented, the last two operations are known to be more computationally costly than the first one and are very prone to causing instability specially in fixed-point computations. Even in systems equipped with floating-point arithmetic, these two operations can sometime lead to overflow, hence causing the algorithm to diverge. To avoid this problem of overflow, floating-point arithmetic is preferred. From a computational viewpoint and whenever possible, algorithms involving fewer multiplications and divisions should always sought after, especially in applications involving tracking in fast-changing environments, e.g., wireless communication.

Table I gives the computational complexity of the three algorithms used in this work, namely, the simple DFE, the MLP DFE, and the lattice-based MLP DFE. The notations used in Table I are defined as follows: N_1 is the number of feedforward taps, N_2 is the number of feedback taps, $(N_1 + N_2)$ is the number of inputs to the MLP, L_1 and L_2 are the numbers of neurons in the first and second hidden layers of the MLP, respectively, and finally $L_3 = 1$ is the number of neurons in the output layer of the MLP.

From Table I, we see that the LMS algorithm is the simplest of all and by using the lattice filter at the input of the MLP DFE, a load of $(20N_1 + 39N_2 - 39)$ computations is added. Thus, a tradeoff between performance and complexity has to be made when implementing the equalizer.

IV. Simulation Results

The performance of the new structure (MLP DFE with lattice structure) is compared to that of the simple DFE and MLP DFE. Both the simple DFE and the MLP DFE structures use four samples in the feedforward section and one sample in the feedback section. For the latter structure, this results in five input samples in its input layer. The number of neurons in the first and second hidden layer, and the output layer are 9, 3, and 1, respectively, and is denoted by the triplet (9, 3, 1) for ease of reference. Finally, the L-DFE was initially set up with $N_1 = 4$ and $N_2 = 1$ giving rise to four outputs $(g_m(i),\ 0 \leq m \leq 3)$. These four outputs and the delayed output from the decision device will provide the five inputs to the MLP DFE. In this case, a (6, 3, 1) configuration for the lattice-based MLP DFE is used such that the two MLP configurations have similar complexity. The back propagation algorithm is used to update the MLP DFE where the learning gain parameter η, momentum parameter α, and threshold level adaptation gain β, have been chosen as in [3], namely 0.07, 0.3, and 0.05, respectively. The weighting factor for the lattice algorithm is 0.99 and the step size for the LMS algorithm used in the simple DFE is 0.035. The digital message

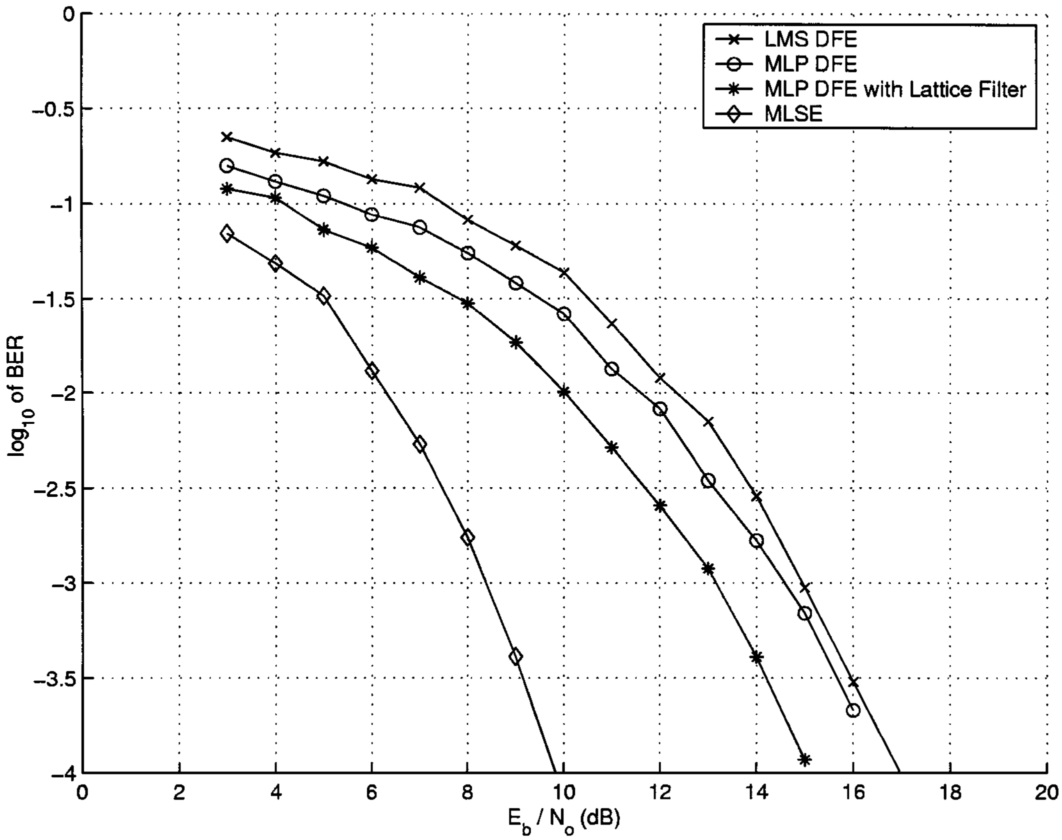


Fig. 6. BER performance of different types of equalizers with channel $H_1(z) = 0.3482 + 0.8704z^{-1} + 0.3482z^{-2}$.

applied to the channel is made of uniformly distributed bipolar random numbers $(-1, 1)$. The channel noise is taken to be an additive white Gaussian noise.

The simulation results reported in this paper were all obtained using MATLAB. Neither instability problem nor coefficient divergence were observed in these simulations even for long runs of 10^7 samples.

The system performance will be evaluated using both a time-invariant and a time-varying channel, as discussed below.

A. System Performance in Time-Invariant Channels

Two time-invariant channel models are used in the simulation and are described by their transfer functions $H_1(z) = 0.3482 + 0.8704z^{-1} + 0.3482z^{-2}$, and $H_2(z) = 0.408 + 0.816z^{-1} + 0.408z^{-2}$. The eigenvalue spreads for the first and second channel are 25 and 81, respectively. Also, it should be pointed out here that the first channel model matches the one used in [3], while the second channel matches one of the time-invariant channel models used in [7].

During this part of the simulations, the performance measure is obtained through the use of scatter diagrams, learning curves, and BER curves.

Scatter diagrams are plotted in terms of the present channel-output symbol as a function of the past channel-output symbol. Figs. 2 and 3 show these diagrams for the three equalizers' outputs after 10 000 iterations for first channel and second channel, respectively, and with a signal to noise ratio (SNR) of 10 dB. In both figures, part (a) shows the unequalized data, whereas parts (b), (c), and (d) represent the effect of equalization by the simple DFE, MLP DFE and lattice-based MLP DFE, respectively. As can be observed from these diagrams, the simple DFE tries to remove ISI from the received data, but with little success against the contaminating noise. However, the MLP DFE not only removes ISI, but also reduces the area of the noise cloud by restricting the equalized symbols within a small area. Finally, when whitening the received data with the lattice filter, the equalizer's performance is much improved and the symbols are seen to converge closer to their original positions. The positive effect of the lattice filter is clear.

In the case of the learning curves, these are obtained by averaging 600 independent runs. Each run has a different random sequence and random starting weights for the perceptron structure, and an SNR of 20 dB is used. Fig. 4 depicts the convergence behavior of the three algorithms for the first channel. This figure shows a clear improvement in both the convergence time and the steady-state MSE when whitening the received data with the lattice algorithm. This result illustrates also that even though the MLP DFE configuration converges more slowly than that of the simple DFE, it nevertheless results in a lower steady-state MSE value than that of the latter. It should also be clear from this figure that the steady-state MSE of both MLP DFE configurations is below the noise level. This results from the nonlinear nature of the equalizer transfer function [3]. Furthermore, the MLP DFE equalizer is capable of generating highly nonlinear decision regions, in contrast to the LMS DFE equalizer which only forms a hyperplane decision boundary [10]. Also, it is generally

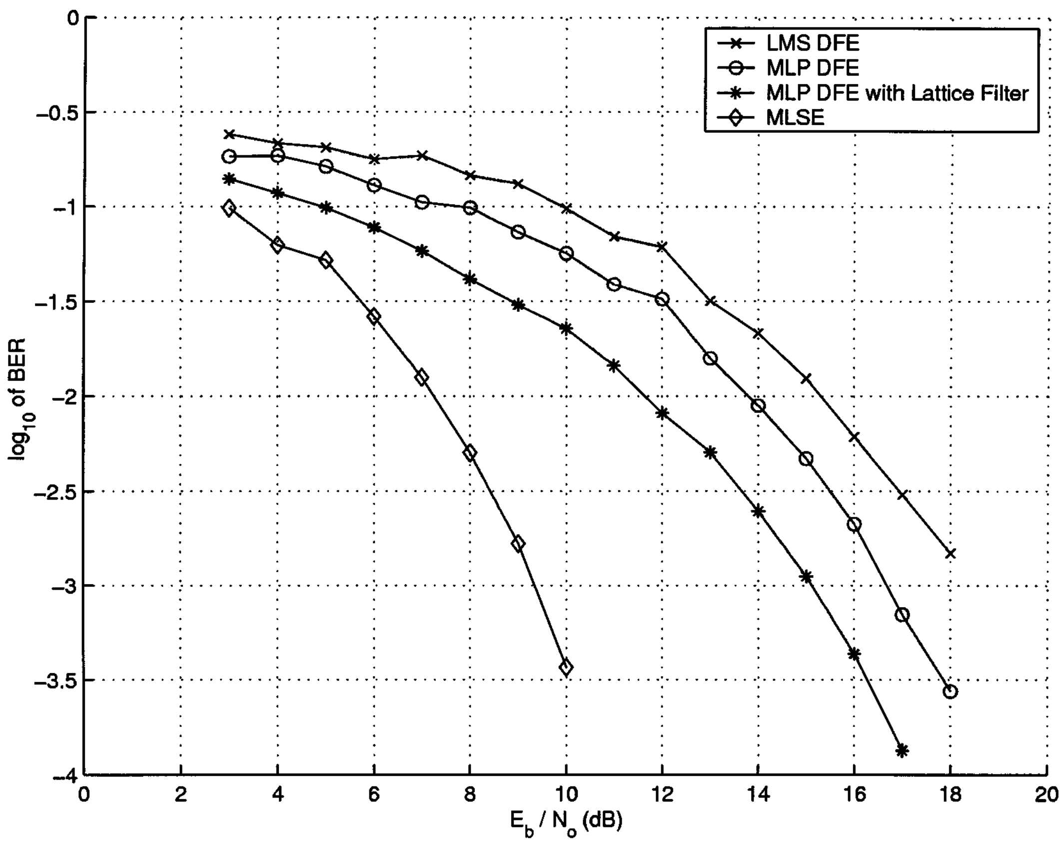


Fig. 7. BER performance of different types of equalizers with channel $H_2(z) = 0.408 + 0.816z^{-1} + 0.408z^{-2}$.

understood in linear signal processing that the l_2 norm error criterion produces a parabolic error surface with no local minima and has a continuously differential nature. However, the l_2 norm error criterion for perceptrons will not generally produce a parabolic error surface owing to its nonlinear nature [19].

A similar improvement is also obtained for the second channel despite its larger eigenvalue spread, as shown in Fig. 5. The difference in convergence time between the MLP DFE and the MLP DFE using lattice algorithm is now more pronounced. The insensitivity of the lattice algorithm to the eigenvalue spread is very clear.

Figs. 6 and 7 illustrate the BER curves for simple DFE, MLP DFE, and lattice-based MLP DFE for the first and second channel, respectively. The consistency in performance of the proposed structure in both channels is very distinctive. An improvement with the proposed lattice-based structure of 1–2 dB over the MLP DFE and of 2-dB over the simple DFE at a BER of 10^{-3} has been clearly achieved for the first channel. Again, the difference is more distinct for the case of the second channel. This makes the use of the lattice structure more justifiable when the ISI is large.

To analyze further the performance of the lattice-based MLP DFE previously discussed, its performance is compared to that of the maximum-likelihood sequence estimation (MLSE) [20] which comprises a channel impulse response estimator and a maximum likelihood sequence estimator. In the simulation, no channel estimation was performed and therefore the exact channel impulse response was used in the MLSE. The results are reported in Figs. 6 and 7 for the first and second channel, respectively. The effectiveness of MLSE over the lattice-based MLP DFE is very clear from these figures; however, the computational complexity of the MLSE technique (including the channel estimator) severely limits its applications [7].

To further enhance the performance of the lattice-based MLP DFE as far as the size of the lattice filter is concerned, it has been tested when $N_1 = 4$ and $N_2 = 1$ (4, 1) against the one with $N_1 = 5$ and $N_2 = 2$ (5, 2) for the second channel. The (5, 2) configuration has now five inputs entering the MLP DFE whereas the (4, 1) configuration has only four inputs. The results are depicted in Fig. 8. The (5, 2) configuration gives a better performance over the (4, 1) configuration; more than 1 dB improvement at a BER of 10^{-3} is obtained. This improvement in performance is of course reached at a slightly higher complexity, since the (5, 2) configuration process more information and therefore is expected to yield an enhancement in performance.

Finally, an experiment was conducted for both channels to observe the changes in performance of the lattice-based MLP DFE configuration by reducing its size from (9, 3, 1) to (3, 2, 1). The learning curves for the lattice-based MLP DFE configuration for the first and the second channel are shown in Fig. 9. The signal to noise ratio is set to 10 dB. From this figure, it is observed that the decrease in the network size increases the convergence time of the neural equalizer but the steady-state MSE remains unaffected. This is because it takes a longer time for the neurons of the small-size MLP DFE to converge to the optimum weights but once they converge, the mean-squared error is the same as that of the comparatively large-sized network. But this trend does not hold for any lower size of the network, i.e., re-

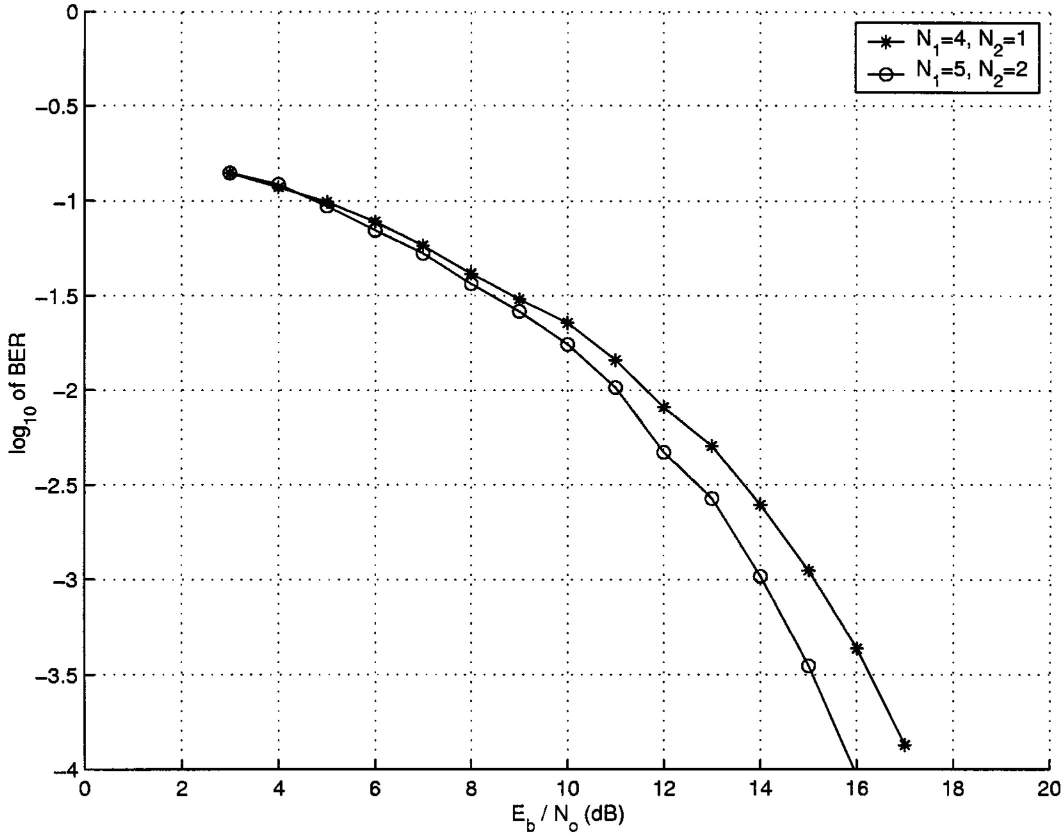


Fig. 8. BER performance of lattice-based MLP DFE with channel $H_2(z) = 0.408 + 0.816z^{-1} + 0.408z^{-2}$ comparing (4, 1) and (5, 2) configurations.

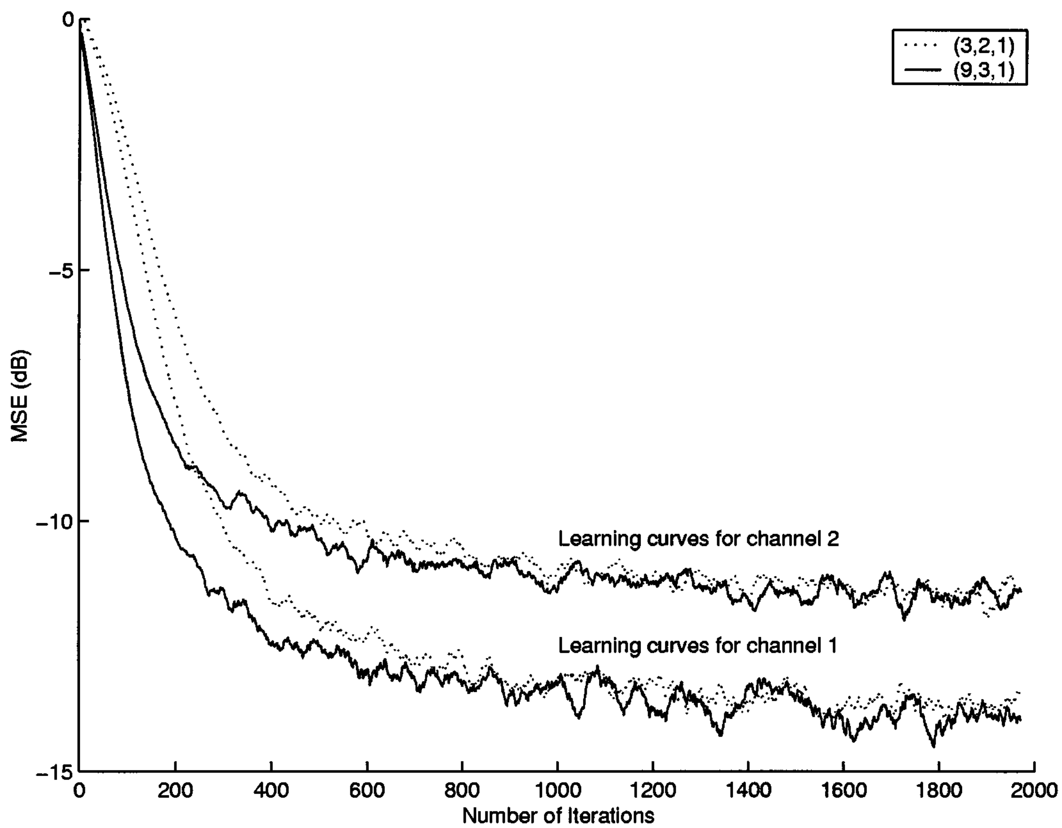


Fig. 9. Learning curves of lattice-based MLP DFE with channel 1 [$H_1(z) = 0.3482 + 0.8704z^{-1} + 0.3482z^{-2}$] and channel 2 [$H_2(z) = 0.408 + 0.816z^{-1} + 0.408z^{-2}$] for (9, 3, 1) and (3, 2, 1) configurations under SNR $= 10$ dB.

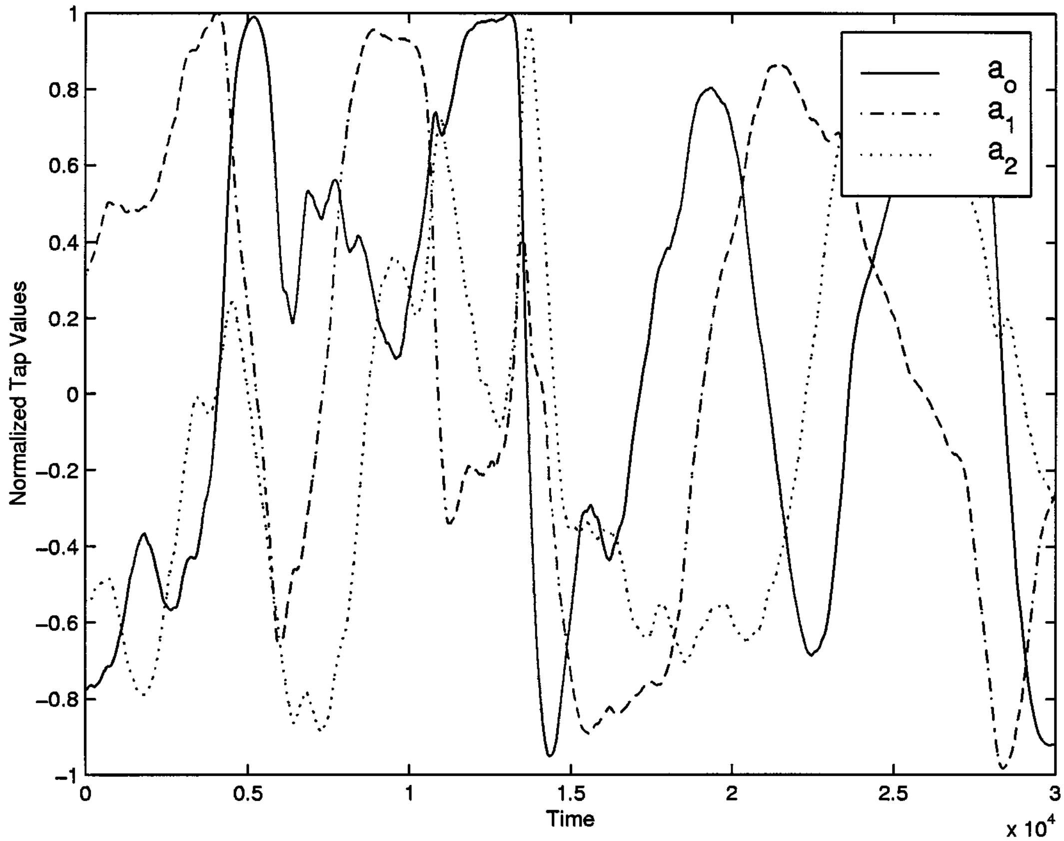


Fig. 10. Tap coefficients of time-variant channel with low-pass filter bandwidth of 0.5 Hz.

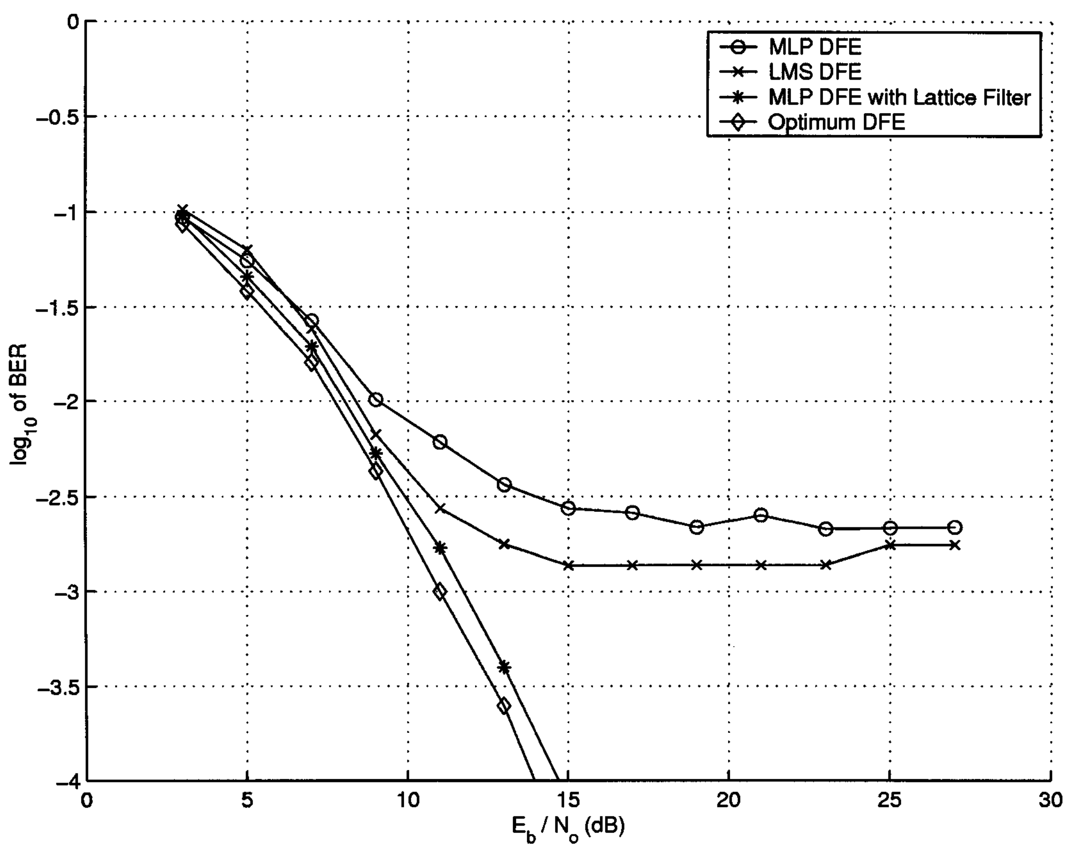


Fig. 11. BER performance of different types of equalizers for time varying channel with low-pass filter bandwidth of 0.1 Hz.

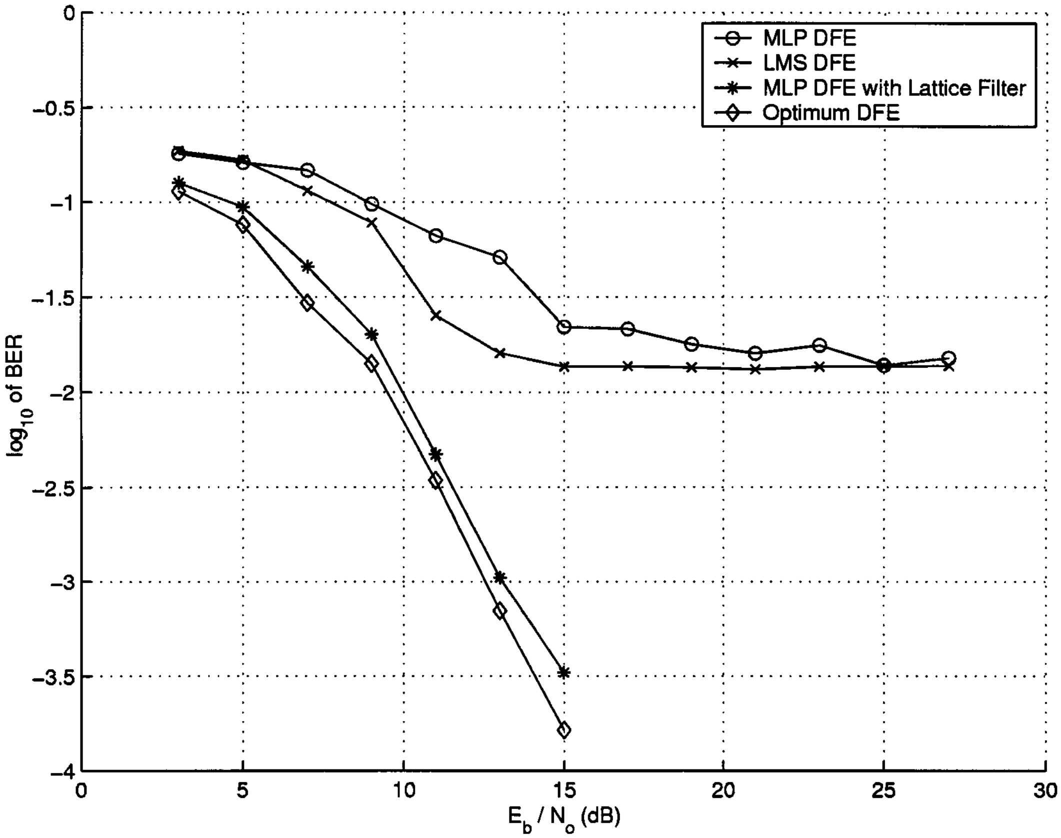


Fig. 12. BER performance of different types of equalizers for time varying channel with lowpass filter bandwidth of 0.5 Hz.

ducing the network size further will deteriorate the behavior of the equalizer, since the equalizer will lack the minimum necessary computational power (number of neurons) to be able to track variations in the weights caused by changes in the environment.

B. System Performance in Time-Varying Channels

The time-variant fading channel was used to evaluate the capability of the equalizer to track the changes in a time-varying dispersive channel. The discrete-time channel model for time-varying fading channel is described by the following transfer function $H(z) = a_0(t) + a_1(t)z^{-1} + a_2(t)z^{-2}$, where $a_0(t)$, $a_1(t)$, and $a_2(t)$ are the time-varying coefficients of the channel impulse response. These are generated by passing white Gaussian noise through a low-pass filter of a specified bandwidth [7]. If we assume that we have a nominal 3 kHz HF channel, the signaling rate is 2400 symbols/s, and the low-pass filter is a two-pole Butterworth filter, then the 3-dB bandwidth of the low-pass filter can be used as a parameter to control the rate of variation of the channel impulse response. The curves representing the time variation of the coefficients are depicted in Fig. 10 for bandwidth of 0.5 Hz.

Figs. 11 and 12 show the BER performance of the three equalizers for the time variations of the coefficients for bandwidths of 0.1 Hz and 0.5 Hz, respectively. The results illustrate the superiority of the lattice-based MLP DFE. In both figures, the other two equalizers' performances are not attractive. Moreover, as can be noticed from these figures that comparable performance is obtained for both MLP DFE and LMS DFE configurations; however, for the time-invariant channels, it was found that the MLP DFE gives better performance over the LMS DFE.

To further investigate the consistency in performance of the lattice-based MLP DFE, the rate of variation of the channel impulse response is increased. This is done by increasing the bandwidth of the low-pass filter, e.g., bandwidth of 1.0 Hz. There is a deterioration in the performance of both the LMS DFE and the MLP DFE and the difference between them and the lattice-based MLP DFE will increase as clearly shown in Fig. 13. The lattice-based MLP DFE attains lower error floor than both of the LMS DFE and the MLP DFE. The LMS DFE and the MLP DFE saturate after the SNR of 13 dB to approximately to same BER. Again in this case the MLP DFE and the LMS DFE have comparable performance.

For comparison purposes, the theoretically optimum DFE [7] was simulated. This is a transversal DFE with coefficients selected according to the exact characteristics of the channel at each time instant, to yield minimum MSE. The comparison between the lattice-based MLP DFE and the theoretically optimum DFE for channel variations characterized by bandwidths 0.1 Hz, 0.5 Hz and 1.0 Hz were used in simulation yielding results shown in Figs. 11–13, respectively. Comparable performances were obtained in Figs. 11 and 12. However, when the variation of the channel impulse response is increased, the performance of the lattice-based MLP DFE degrades a little bit.

The saturation effect (error-floor) for both LMS DFE and MLP DFE in Fig. 11 is due mainly to when E_b/N_0 is larger than 10 dB, noise has little influence on the BER, because at

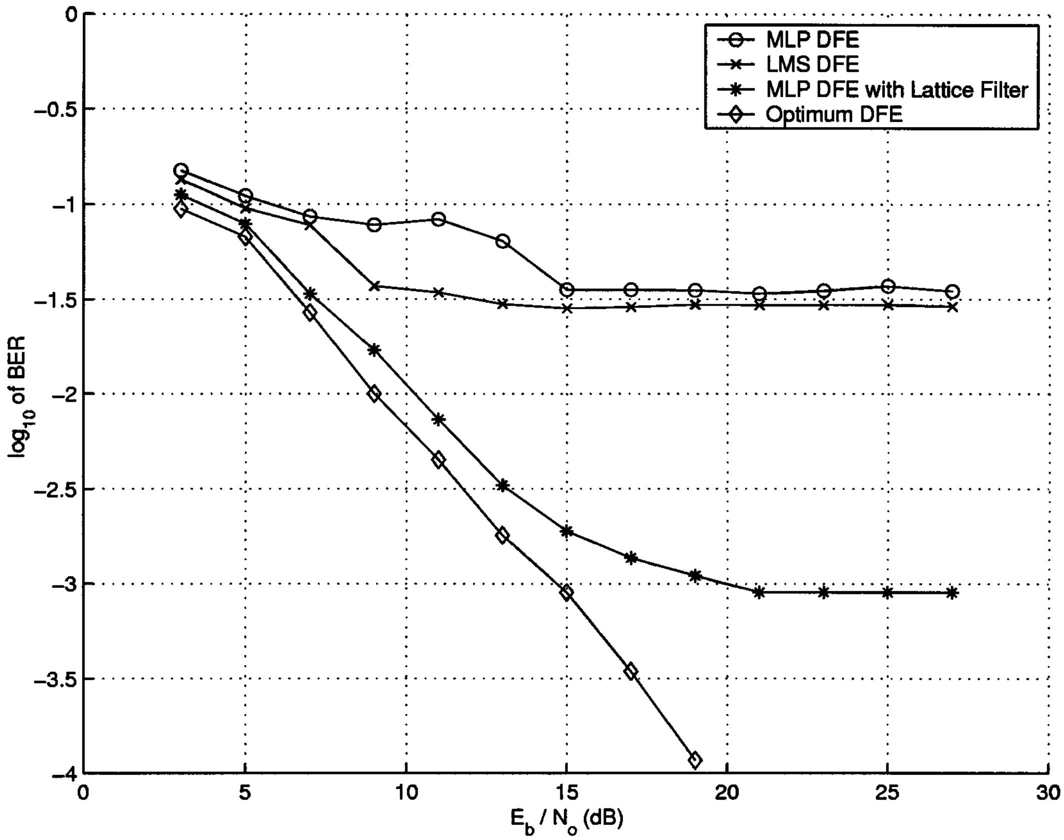


Fig. 13. BER performance of different types of equalizers for time varying channel with lowpass filter bandwidth of 1.0 Hz.

that time, errors are mainly caused by tap-gain lag due to the variation of channels. As the channel variation increases, the saturation effect gets worse, this is depicted in Figs. 12 and 13.

Finally, as explained earlier, it was made sure that the two MLP configurations have similar complexity.

V. Conclusion

The results of our study can be summarized briefly as follows:

1) The use of the lattice filter as a whitening scheme for the MLP DFE equalizer's input data results in substantial improvements in terms of convergence rate, steady-state MSE, and BER.
2) The convergence rate of the lattice-based MLP DFE is insensitive of the eigenvalue spread of the channel correlation matrix for time-invariant channels.
3) The whitening scheme makes the lattice-based MLP DFE less sensitive to the size of the MLP down to a tolerable size.
4) Increasing the size of the lattice DFE enhances further the performance of the lattice-based MLP DFE.
5) The proposed lattice-based MLP DFE has better tracking capabilities than both the LMS DFE and MLP DFE in time-varying channels.
6) It is found that, for time-invariant channels, the MLP DFE out performs the LMS DFE. However, for time-varying channels comparable performance is obtained for the two configurations. The LMS and the BP algorithms have comparable tracking capabilities in time-varying channels.
7) In time-varying channels, on which the results of Figs. 11 and 12 are based, the performance of the lattice-based MLP DFE is close to that of the theoretically optimum DFE.
8) The simulation results indicated that the lattice-based MLP DFE is stable and no sign of instability was shown for both time-invariant and time-varying channels.
9) A separate study to be reported later shows the robustness of the proposed scheme with respect to nonlinearities.
10) Further work aimed at comparisons with LS DFE in both time-invariant and time-varying channels is currently being pursued.

Finally, this work has presented the improvements brought about by whitening the received data samples before they are applied to the MLP DFE. The lattice-based MLP DFE out performed both the LMS DFE and MLP DFE in both time-invariant and time-varying channels.

Appendix
The Least Squares Lattice DFE Algorithm

The algorithm stated below is adapted from [7], and is summarized for convenience. Note that bold faced characters represent matrices or vectors. Specifically, $\mathbf{f}_m(i)$, $\mathbf{b}_m(i)$ and $\mathbf{k}^g_m(i)$

are 2×1 vectors, while $\mathbf{r}_m^f(i)$, $\mathbf{r}_m^b(i)$ and $\mathbf{k}_m(i)$ are 2×2 matrices. We also denote the inverse of $\mathbf{r}_m^f(i)$ and $\mathbf{r}_m^b(i)$ by $\mathbf{r}_m^{-f}(i)$ and $\mathbf{r}_m^{-b}(i)$, respectively. All other quantities are scalars. Finally, $\tilde{x}_{n-d}$ is the detected sample. In this algorithm γ is chosen to be a very small positive number, while w is a positive number close to one (typically 0.95–1.0).

A. Initialization

$$
\begin{aligned}
\mathrm{b}_o(i) &= f_o(i) = y(i)\\
\mathrm{r}_o^f(i) &= r_o^b(i) = w * r_o^f(i-1) + |y(i)|^2\\
\mathrm{e}_o(i) &= \tilde{x}_{n-d}(i), \qquad g_o(i) = 0, \qquad \alpha_o(i) = 1\\
\mathrm{r}_m^f(0) &= r_m^b(0) = \gamma, \qquad k_m(0) = 0, \qquad k_m^g(0) = 0\\
&\qquad (m = 1, \ldots, N_1 - N_2 - 1)\\
\mathrm{k}_M^b(0) &= 0 \qquad (M = N_1 - N_2)\\
\mathbf{k}_m(0) &= \mathbf{0}, \qquad \mathbf{k}_m^g(0) = \mathbf{0}, \qquad \mathbf{r}_m^f(0) = \mathbf{r}_m^b(0) = \gamma\mathbf{I}\\
&\qquad (m = N_1 - N_2, \ldots, N_1).
\end{aligned}
$$

B. Scalar Stages: ($0 < m \le N_1 - N_2$) Unless Otherwise Specified

$$
\begin{aligned}
\mathrm{f}_m(i) &= f_{m-1}(i) - k_m(i-1)b_{m-1}(i-1)/r_{m-1}^b(i-2)\\
\mathrm{b}_m(i) &= b_{m-1}(i-1) - k_m(i-1)f_{m-1}(i)/r_{m-1}^f(i-1)\\
\mathrm{k}_m(i) &= w * k_m(i-1) + \alpha_{m-1}(i-1)f_{m-1}(i)b_{m-1}(i-1)\\
\mathrm{r}_m^f(i) &= r_{m-1}^f(i) - |k_m(i)|^2/r_{m-1}^b(i-1) \quad (m < N_1 - N_2)\\
\mathrm{r}_m^b(i) &= r_{m-1}^b(i-1) - |k_m(i)|^2/r_{m-1}^f(i) \quad (m < N_1 - N_2)\\
g_m(i) &= g_{m-1}(i) + k_m^g(i-1)b_{m-1}(i)/r_{m-1}^b(i-1)\\
\alpha_m(i) &= \alpha_{m-1}(i) - |b_{m-1}(i)\alpha_{m-1}(i)|^2/r_{m-1}^b(i)\\
\mathrm{e}_m(i) &= \tilde{x}_{n-d}(i) - g_m(i) \qquad (0 \le m \le N_1 - N_2)\\
\mathrm{k}_m^g(i) &= w * k_m^g(i-1) + \alpha_{m-1}(i)e_{m-1}(i)b_{m-1}(i).
\end{aligned}
$$

C. Transitional Stage: ($M = N_1 - N_2$)

$$
\begin{aligned}
\mathrm{b}_M^*(i) &= e_{M-1}(i-1) - k_M^b(i-1)f_{M-1}(i)/r_{M-1}^f(i-1)\\
\mathrm{k}_M^b(i) &= w * k_M^b(i-1) + \alpha_{M-1}(i-1)f_{M-1}(i-1)e_{M-1}(i-1)\\
\mathbf{f}_M(i) &= [f_M(i)\ e_M(i-1)]^T\\
\mathbf{b}_M(i) &= [b_M(i)\ b_M^*(i)]^T\\
\mathbf{r}_M^f(i) &= w * \mathbf{r}_M^f(i-1) + \alpha_M(i-1)\mathbf{f}_M(i)\mathbf{f}_M^T(i)\\
\mathbf{r}_M^b(i) &= w * \mathbf{r}_M^b(i-1) + \alpha_M(i)\mathbf{b}_M(i)\mathbf{b}_M^T(i).
\end{aligned}
$$

D. Two-Dimensional Stages: ($N_1 - N_2 < m < N_1$) Unless Otherwise Specified

$$
\begin{aligned}
\mathbf{f}_m(i) &= \mathbf{f}_{m-1}(i) - \mathbf{k}_m(i-1)\mathbf{r}_{m-1}^{-b}(i-2)\mathbf{b}_{m-1}(i-1)\\
\mathbf{b}_m(i) &= \mathbf{b}_{m-1}(i-1) - \mathbf{k}_m(i-1)\mathbf{r}_{m-1}^{-f}(i-1)\mathbf{f}_{m-1}(i)\\
\mathbf{k}_m(i) &= w * \mathbf{k}_m(i-1) + \alpha_{m-1}(i-1)\mathbf{b}_{m-1}(i-1)\mathbf{f}_{m-1}^T(i)\\
\mathbf{r}_m^f(i) &= w * \mathbf{r}_m^f(i-1) + \alpha_m(i-1)\mathbf{f}_m(i)\mathbf{f}_m^T(i)\\
\mathbf{r}_m^b(i) &= w * \mathbf{r}_m^b(i-1) + \alpha_m(i)\mathbf{b}_m(i)\mathbf{b}_m^T(i)\\
g_m(i) &= g_{m-1}(i) + \mathbf{k}_m^{gT}(i-1)\mathbf{r}_{m-1}^{-b}(i-1)\mathbf{b}_{m-1}(i)\\
&\qquad (m < N_1)\\
\alpha_m(i) &= \alpha_{m-1}(i) - \alpha_{m-1}^2(i)\mathbf{b}_{m-1}^T(i)\mathbf{r}_{m-1}^{-b}(i)\mathbf{b}_{m-1}(i)\\
\mathrm{e}_m(i) &= \tilde{x}_{n-d}(i) - g_m(i)\\
\mathbf{k}_m^g(i) &= w * \mathbf{k}_m^g(i-1) + \alpha_{m-1}(i)\mathbf{b}_{m-1}(i)e_{m-1}(i)\\
&\qquad (m < N_1).
\end{aligned}
$$

Acknowledgment

The authors would like to thank the anonymous reviewers for their constructive suggestions which has helped improve the paper and Dr. L. Cheded for the discussions on this topic.

References

[1] S. Qureshi, "Adaptive equalization," *Proc. IEEE*, vol. 73, no. 9, pp. 1349–1387, 1985.

[2] J. G. Proakis, *Digital Commun.*. New York: McGraw-Hill, 1989.

[3] S. Siu, G. J. Gibson, and C. F. N. Cowan, "Decision feedback equalization using neural network structures and performance comparison with standard architecture," *Proc. Inst. Elect. Eng.*, vol. 137, pp. 221–225, Aug. 1990.

[4] R. P. Lippmann, "An introduction to computing with neural nets," *IEEE ASSP Mag.*, vol. 4, Apr. 1987.

[5] C. M. Bishop, *Neural Networks for Pattern Recognition*. Oxford, U.K.: Clarendon, 1996.

[6] S. Haykin, *Adaptive Filter Theory*. Englewood Cliffs, NJ: Prentice-Hall, 1991.

[7] F. Ling and J. G. Proakis, "Adaptive lattice decision-feedback equalizers: Their performance and application to time-variant multipath channels," *IEEE Trans. Comm.*, vol. COM-33, pp. 348–356, April 1985.

[8] A. Shafi, A. Zerguine, and M. Bettayeb, "Neural-network-based decision feedback equaliser with lattice structure," *Electron. Lett.*, vol. 35, pp. 1705–1707, September 1999.

[9] S. Chen, G. J. Gibson, C. F. N. Cowan, and P. M. Grant, "Adaptive equalization of finite nonlinear channels using multilayer perceptrons," *Signal Processing*, vol. 20, pp. 107–119, 1990.

[10] G. J. Gibson, S. Siu, and C. Cowan, "The application of nonlinear structures to the reconstruction of binary signals," *IEEE Trans. Signal Processing*, vol. 39, pp. 1877–1884, Aug. 1991.

[11] B. Mulgrew, "Applying radial basis functions," *IEEE ASSP Mag.*, vol. 13, pp. 50–65, Mar. 1996.

[12] A. Wieland and R. Leighton, "Geometric analysis of neural-network capabilities," in *1st Int. Conf. Neural Networks*, June 1987, pp. 385–393.

[13] S. Haykin, *Neural Networks: A Comprehensive Foundation*. New York: Macmillan, 1994.

[14] D. E. Rumelhart and J. L. McClelland, *Parallel Distribution Processing: Explorations in the Microstructure of Cognition*. Cambridge, MA: MIT Press, 1986, vol. 1.

[15] G. J. Gibson, S. Siu, and C. Cowan, "Multilayer perceptron structures applied to adaptive equalizers for data communications," in *Proc. ICASSP*, May 1989, pp. 1183–1186.

[16] B. Widrow, J. M. McCool, M. G. Larimore, and C. R. Johnson, "Stationary and nonstationary learning characteristics of the lms adaptive filter," *Proc. IEEE*, vol. 64, pp. 1151–1162, Aug. 1976.

[17] E. H. Satorius and S. T. Alexandre, "Channel equalization using adaptive lattice algorithms," *IEEE Trans. Commun.*, vol. COM-27, pp. 899–905, June 1979.

[18] J. Makhoul, "A class of all-zero lattice digital filters: Properties and applications," *IEEE Trans. Acoust., Speech, and Signal Processing*, vol. ASSP-26, pp. 304–314, Aug. 1978.

[19] S. Siu and C. F. N. Cowan, "Performance analysis of the l_p norm back-propagation algorithm for adaptive equalization," *Proc. Inst. Elect. Eng.*, pt. F, vol. 140, no. 1, pp. 43–47, Feb. 1993.

[20] G. D. Forney Jr., "Maximum-likelihood sequence estimation of digital sequences in the presence of intersymbol interferences," *IEEE Trans. Inform. Theory*, vol. IT-18, pp. 363–378, May 1972.

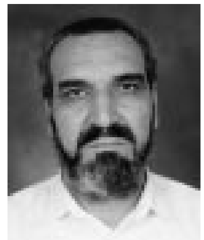
Azzedine Zerguine (S'93–M'96) received the B.Sc. degree from Case Western Reserve University, Cleveland, OH, in 1981, the M.Sc. degree from King Fahd University of Petroleum and Minerals (KFUPM), Dhahran, Saudi Arabia, in 1990, and the Ph.D. degree from Loughborough University, Loughborough, U.K., in 1996, all in electrical engineering.

From 1981 to 1987, he was working with different Algerian state-owned companies. From 1987 to 1990, he was a Research and Teaching Assistant, Electrical Engineering Department, KFUPM. In 1990, he joined the Physics Department at KFUPM as a Lecturer, and then from 1997 to 1998 as an Assistant Professor. In 1998, he joined the Department of Electrical Engineering at KFUPM where he is presently an Assistant Professor working in the area of Signal Processing and Communications. His research interests include signal processing, adaptive filtering, neural networks, and digital communications.

Ahmar Shafi received the B.Sc. degree in electrical engineering from N.E.D. University of Engineering and Technology, Karachi, Pakistan, and the M.Sc. degree in telecommunications from King Fahd University of Petroleum and Minerals (KFUPM), Dhahran, Saudi Arabia, in 1997 and 1999, respectively.

From September 1997 to August 1999, he worked in KFUPM as a Research and Teaching Assistant in Electrical Engineering Department. Currently, he is working as a Lecturer in the Electrical Engineering Department of KFUPM. He is also serving in the Telecommunications Division of KFUPM as a Telecom Engineer. There, his responsibilities include the commissioning, planning, management and trouble-shooting of the new state-of-the-art digital PABX system for 14 000 users of the university. He is also responsible for maintaining the web site for the Projects and Maintenance Department of the university. His areas of interest are neural networks, channel equalization, and wireless communications.

Mr. Shafi is a lifetime member of Pakistan Engineering Council.



Maamar Bettayeb received the B.S., M.S., and Ph.D. degrees in electrical engineering from University of Southern California, Los Angeles, in 1976, 1978 and 1981, respectively.

He worked as a Research Engineer at the Bellaire Research Center at Shell Oil Development Company, Houston, in the development of deconvolution seismic signal processing algorithms for the purpose of Gas and Oil exploration. From 1982 to 1988, he directed the Instrumentation and Control Laboratory of the High Commission for Research in Algeria, where he led various research and development projects in the field of modeling, simulation, and control design of large scale energy systems, specifically, model reduction, identification and decomposition, composite decentralized control and simulation of computer-controlled systems with applications to nuclear, solar and electric power systems. In 1988, he joined the Electrical Engineering Department at King Fahd University of Petroleum and Minerals (KFUPM), Dhahran, Saudi Arabia. He has been Professor at University of Sharjah, UAE since August 2000. He has been consulting for the Petrochemical Industries and has also been involved in various R&D funded projects in the areas of process control and signal processing applications to ultrasonic nondestructive testing for defect evaluation. His recent research interest is in H_∞ optimal control, rational approximation, signal processing, process control, artificial intelligence, and industrial applications.