Adaptive Decision Feedback Multiuser Detectors with Recurrent Neural Networks for DS-CDMA in Fading Channels

Rodrigo C. de Lamare and Raimundo Sampaio-Neto

CETUC/PUC-RIO, 22453-900, Rio de Janeiro - Brazil
E-mails: delamare@infolink.com.br, raimundo@cetuc.puc-rio.br

Abstract. In this work we propose adaptive decision feedback (DF) multiuser detectors (MUDs) for DS-CDMA systems using recurrent neural networks (RNN). A DF CDMA receiver structure is presented with dynamically driven RNNs in the feedforward section and finite impulse response (FIR) linear filters in the feedback section for performing interference cancellation. A stochastic gradient (SG) algorithm is developed for estimation of the parameters of the proposed receiver structure. A comparative analysis of adaptive minimum mean squared error (MMSE) receivers operating with SG algorithms is carried out for linear and DF receivers with FIR filters and neural receiver structures with and without DF. Simulation experiments including fading channels show that the DF neural MUD outperforms DF MUDs with linear FIR filters, linear receivers and the neural receiver without interference cancellation.

1 Introduction

In third generation wideband direct-sequence code-division multiple access (DS-CDMA) systems high data rate users can be accomodated by reducing the processing gain N and using a low spreading factor [1]. In these situations, the multiacess interference (MAI) is relatively low (small number of users), but the intersymbol interference (ISI) can cause significant performance degradation. The deployment of non-linear structures, such as neural networks and decision feedback (DF), can mitigate more effectively ISI, caused by the multipath effect of radio signals, and MAI, which arises due to the non-orthogonality between user signals. Despite the increased complexity over conventional multiuser receivers with FIR linear filters, the deployment of neural structures is feasible for situations where the spreading factor is low and the number of high data rate users is small. In this case, the trade-off between computational complexity and superior performance is quite attractive. Neural networks have recently been used in the design of DS-CDMA multiuser receivers [2]-[5]. Neural MUDs employing the minimum MMSE [2]-[5] criterion usually show good performance and have simple adaptive implementation, at the expense of a higher computational complexity. In the last few years, different artificial neural networks structures have been used in the design of MUDs: multilayer perceptrons (MLP) [2], radial-basis functions (RBF) [3], and RNNs [4, 5]. These neural MUDs employ non-linear functions to create decision boundaries for the detection of transmitted symbols, whilst conventional MUDs use linear functions to form such decision regions. In this work, we present an adaptive DF MUD, using dynamically driven RNNs in the feedforward section and FIR linear filters in the feedback section, and develop stochastic gradient (SG) algorithms for the proposed DF receiver structure. Adaptive DF and linear MMSE MUDs are examined with the LMS algorithm and compared to the DF and non-DF neural MUDs operating with SG algorithms. Computer simulation experiments with fading channels show that the DF neural MUDs outperforms linear and DF receivers with the LMS and the conventional single user detector (SUD), which corresponds to the matched filter.

2 DS-CDMA system model

Let us consider the uplink of a symbol synchronous DS-CDMA system with K users, N chips per symbol and L_p propagation paths. The baseband signal transmitted by the k-th active user to the base station is given by:

$$x_k(t) = A_k \sum_{i=-\infty}^{\infty} b_k(i) s_k(t - iT) \tag{1}$$

where $b_k(i) \in \{\pm 1\}$ denotes the i-th symbol for user k, the real valued spreading waveform and the amplitude associated with user k are $s_k(t)$ and A_k, respectively. The spreading waveforms are expressed by $s_k(t) = \sum_{n=1}^{N} a_k(i)\phi(t - nT_c)$, where $a_k(i) \in \{\pm 1/\sqrt{N}\}$, $\phi(t)$ is the chip waveform, T_c is the chip duration and $N = T/T_c$ is the processing gain. Assuming that the receiver is synchronised with the main path, the coherently demodulated composite received signal is

$$r(t) = \sum_{k=1}^{K} \sum_{l=0}^{L_p-1} h_{k,l}(t) x_k(t - \tau_{k,l}) \tag{2}$$

where $h_{k,l}(t)$ and $\tau_{k,l}$ are, respectively, the channel coefficient and the delay associated with the l-th path and the k-th user. Assuming that $\tau_{k,l} = lT_c$ and that the channel is constant during each symbol interval, the received signal $r(t)$ after filtering by a chip-pulse matched filter and sampled at chip rate yields the N dimensional received vector

$$\mathbf{r}(i) = \sum_{k=1}^{K} A_k \mathbf{H}_k(i) \mathbf{C}_k \mathbf{b}_k(i) + \mathbf{n}(i) \tag{3}$$

where the Gaussian noise vector is $\mathbf{n}(i) = [n_1(i) \ \ldots \ n_N(i)]^T$ with $E[\mathbf{n}(k)\mathbf{n}^T(i)] = \sigma^2\mathbf{I}$, where $(.)^T$ denotes matrix transpose and $E[.]$ stands for expected value, the k-th user symbol vector is given by $\mathbf{b}_k(i) = [b_k(i) \ \ldots \ b_k(i - L_s + 1)]^T$, where L_s is the ISI span. The $(L_s \times N) \times L_s$ user k code matrix $\mathbf{C}_k$ is described by

$$\mathbf{C}_k = \begin{bmatrix} \mathbf{s}_k & 0 & \ldots & 0 \\ 0 & \mathbf{s}_k & \ddots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \ldots & \ldots & \mathbf{s}_k \end{bmatrix} \tag{4}$$

where $\mathbf{s}_k = [a_k(1) \dots a_k(N)]^T$ is the signature sequence for the k-th user, and the $N \times (L_s \times N)$ channel matrix $\mathbf{H}_k(i)$ for the user k is expressed by

$$\mathbf{H}_k(i) = \begin{bmatrix} h_{k,0}(i) \dots h_{k,L_p-1}(i) & \dots & 0 & 0 \\ \vdots \quad \ddots \quad \ddots & & \ddots \quad \ddots & \vdots \\ 0 \quad\quad 0 \quad\quad \dots & & h_{k,0}(i) \dots h_{k,L_p-1}(i) \end{bmatrix} \tag{5}$$

The MAI comes from the non-orthogonality between the received signature sequences, whereas the ISI span L_s depends on the length of the channel response, which is related to the length of the chip sequence. For $L_p = 1,\ L_s = 1$ (no ISI), for $1 < L_p \leq N, L_s = 2$, for $N < L_p \leq 2N, L_s = 3$.

3 Conventional Decision Feedback MUD

Consider a one shot DF MUD (the receiver observes and detects only one symbol at each time instant), whose observation vector $\mathbf{u}(i)$ is formed from the outputs of a bank of matched filters , where $\mathbf{u}(i) = \mathbf{S}^T\mathbf{r}(i)$ with $\mathbf{S} = [\mathbf{s}_1 \ \dots \ \mathbf{s}_K]$. The observation vector is represented by:

$$\mathbf{u}(i) = [u_1 \dots u_K]^T \tag{6}$$

The detected symbols for an one shot DF multiuser receiver using FIR linear filters are given by :

$$\hat{b}_k(i) = sgn(\mathbf{w}_k^T(i)\mathbf{u}(i) - \mathbf{f}_k^T\hat{\mathbf{b}}(i)) \tag{7}$$

where $\mathbf{w}_k(i) = [w_1 \dots w_K]^T$ and $\mathbf{f}_k(i) = [f_1 \dots f_K]^T$ are, respectively the feedforward and feedback weight vectors for user k for the i-th symbol in a system with K users. The feedforward matrix $\mathbf{w}(k)$ is $K \times K$, the feedback matrix $\mathbf{f}(k)$ is $K \times K$ and is constrained to have zeros along the diagonal to avoid cancelling the desired symbols. In this work, we employ a full matrix $\mathbf{f}(k)$, except for the diagonal, which corresponds to parallel DF [6]. The MMSE solution for this MUD can be obtained via an adaptive algorithm, such as the LMS algorithm [7], which uses the error signal $e_k(i) = b_k(i) - \mathbf{w}_k^T(i)\mathbf{u}(i) + \mathbf{f}_k^T(i)\hat{\mathbf{b}}(i)$, and is described by:

$$\mathbf{w}_k(i+1) = \mathbf{w}_k(i) + \mu_w e_k(i)\mathbf{u}(i) \tag{8}$$

$$\mathbf{f}_k(i+1) = \mathbf{f}_k(i) - \mu_f e_k(i)\hat{\mathbf{b}}(i) \tag{9}$$

where $b_k(i)$ is the desired signal for the k-th user taken from a training sequence, $\mathbf{u}(i)$ is the observation vector, $\hat{\mathbf{b}}(i) = [\hat{b}_1(i) \ \dots \ \hat{b}_K(i)]$ is the vector with the decisions, μ_w and μ_f are the algorithm step sizes.

4 Proposed DF Neural Receiver

In this section we present a decision feedback CDMA receiver that employs recurrent neural networks as its feedforward section and linear FIR filters, similar to the DF MUD described in the preceeding section, as its feedback section. The decision feedback CDMA receiver structure, depicted in Fig. 1, employs dynamically driven RNNs in the feedforward section for suppressing MAI and ISI and FIR linear filters in the feedback section for cancelling the associated users in the system. With respect to the structure, RNNs have one or more feedback connections, where each artificial neuron is connected to the others. These neural networks are suitable to channel equalisation and multiuser detection applications, since they are able to cope with channel transfer functions that exhibit deep spectral nulls, forming optimal decision boundaries [8].

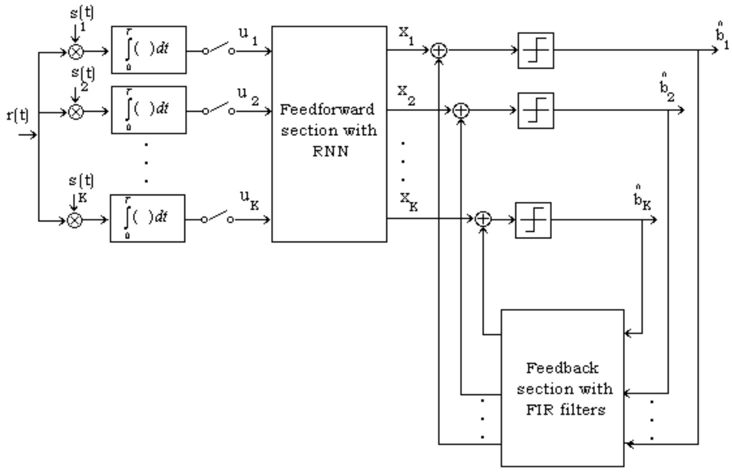

Fig. 1. Proposed DF neural receiver: RNNs are employed in the feedforward section for MAI and ISI rejection and FIR filters are used in the feedback section for cancelling associated interferers.

To describe the proposed neural DF system we use a state-space approach, where the $K \times 1$ vector $\mathbf{x}_k(i)$ corresponds to the K states of the artificial neural network for user k, the $K \times 1$ vector $\mathbf{u}(i)$ to the channel K user symbols output observation vector and the output of the DF neural MUD $\hat{b}_k(i)$ is given by:

$$\xi_k(i) = \left[\mathbf{x}_k^T(i-1)\ \mathbf{u}^T(i)\right]^T \tag{10}$$

$$\mathbf{x}_k(i) = \varphi(\mathbf{W}_k^T(i)\xi_k(i)) \tag{11}$$

$$\hat{b}_k(i) = sgn(\mathbf{D}\mathbf{x}_k(i) - \mathbf{f}_k^T(i)\hat{\mathbf{b}}(i)) \tag{12}$$

where the matrix $\mathbf{W}_k(i) = [\mathbf{w}_{k,1}(i)\ \ldots\ \mathbf{w}_{k,j}\ \ldots\ \mathbf{w}_{k,K}]$ has dimension $2K \times K$ and whose K columns $\mathbf{w}_{k,j}(i)$, with $j = 1,\ 2,\ \ldots, K$ have dimension $K \times 1$ and contain the coefficients of the RNN receiver for user k, $\mathbf{D} = \left[1\ 0\ ...\ 0\right]$ is the $1 \times K$ matrix that defines the number of outputs of the network, $\varphi(.)$ is the activation function of the neural network, the feedback matrix $\mathbf{f}(k) = [\mathbf{f}_1(i)\ \ldots\ \mathbf{f}_K(i)]$ is $K \times K$ and as in the conventional MUD case is constrained to have zeros along the diagonal to avoid cancelling the desired symbols. Note that this type of interference cancellation provides uniform performance over the user population. In particular, we have only one output $\hat{b}_k(i)$ per observation vector $\mathbf{u}(i)$, which corresponds to the one shot approach.

5 Adaptive Algorithms for the neural DF receiver

In order to derive an SG adaptive algorithm that minimizes the mean squared error (MSE) for the proposed DF receiver structure, we consider the following cost function:

$$J(\mathbf{W}_k(i), \mathbf{f}_k(i)) = E[e_k^2(i)] = E[(b_k(i) - (\mathbf{D}\mathbf{x}_k(i) - \mathbf{f}_k^T(i)\hat{\mathbf{b}}(i)))^2] \quad (13)$$

where $e_k(i) = b_k(i) - (\mathbf{D}\mathbf{x}_k(i) - \mathbf{f}_k^T(i)\hat{\mathbf{b}}(i))$. A stochastic gradient algorithm can be developed by computing the gradient terms with respect to $\mathbf{w}_{k,j}$, $j = 1, 2, \ldots, K$, and $\mathbf{f}_k$ and using their instantaneous values. Firstly, we consider the first partial derivative of $J(\mathbf{W}_k(i), \mathbf{f}_k(i))$ with respect to the parameter vector $\mathbf{w}_{k,j}(i)$ with dimension $2K \times 1$, which forms the matrix $\mathbf{W}_k$:

$$\frac{\partial J(\mathbf{W}_k(i), \mathbf{f}_k(i))}{\partial \mathbf{w}_{k,j}(i)} = \left(\frac{\partial e_k(i)}{\partial \mathbf{w}_{k,j}(i)}\right) e_k(i) = -\mathbf{D}\left(\frac{\partial \mathbf{x}_k(i)}{\partial \mathbf{w}_{k,j}(i)}\right) e_k(i) = -\mathbf{D}\boldsymbol{\Lambda}_{k,j}(i) e_k(i) \quad (14)$$

where the $K \times 2K$ matrix $\boldsymbol{\Lambda}_{k,j}(i)$ contains the partial derivatives of the state vector $\mathbf{x}_k(i)$ with respect to $\mathbf{w}_{k,j}(i)$. To obtain the expressions for the updating of the matrix $\boldsymbol{\Lambda}_{k,j}(i)$, we consider the update equations for the state vector $\mathbf{x}_k(i)$ given through (10) and (11). Using the chain rule of calculus in (11), we obtain the following recursion that describes the dynamics of the learning process of the neural receiver:

$$\boldsymbol{\Lambda}_{k,j}(i+1) = \boldsymbol{\Phi}_k(i)\Big(\mathbf{W}_k^{1:K}(i)\boldsymbol{\Lambda}_{k,j}(i) + \mathbf{U}_{k,j}(i)\Big),\ j = 1, 2, \ldots, K \quad (15)$$

where the $K \times K$ matrix $\mathbf{W}_k^{1:K}$ denotes the submatrix of $\mathbf{W}_k$ formed by the first K rows of $\mathbf{W}_k$, the $K \times K$ matrix $\boldsymbol{\Phi}_k(i)$ for user k has a diagonal structure where the elements correspond to the partial derivative of the activation function $\varphi(.)$ with respect to the argument in $\mathbf{w}_{k,j}^T(i)\boldsymbol{\xi}_k(i)$ as expressed by:

$$\boldsymbol{\Phi}_k(i) = diag\left(\varphi'(\mathbf{w}_{k,1}^T(i)\boldsymbol{\xi}_k(i)), \ldots, \varphi'(\mathbf{w}_{k,j}^T(i)\boldsymbol{\xi}_k(i)), \ldots, \varphi'(\mathbf{w}_{k,K}^T(i)\boldsymbol{\xi}_k(i))\right) \quad (16)$$

and the $K \times 2K$ matrix $\mathbf{U}_{k,j}(i)$ has all the rows with zero elements, except for the j-th row that is equal to the vector $\boldsymbol{\xi}_k(i)$):

$$\mathbf{U}_{k,j}(i) = \begin{bmatrix} \mathbf{0}^T \\ \boldsymbol{\xi}_k(i) \\ \mathbf{0}^T \end{bmatrix}, \quad j = 1, 2, \ldots, K \quad (17)$$

The update equation for the feedforward parameter vector $\mathbf{w}_{k,j}$ of the decision feedback receiver is obtained via a stochastic gradient optimisation that uses the expression obtained in (14) to update the parameters using the gradient rule $\mathbf{w}_{k,j}(i+1) =$

$\mathbf{w}_{k,j}(i) - \mu_n \frac{\partial J(\mathbf{W}_k(i), \mathbf{f}_k(i))}{\partial \mathbf{w}_{k,j}(i)}$ which yields the recursion for the neural section of the receiver:

$$\mathbf{w}_{k,j}(i+1) = \mathbf{w}_{k,j}(i) + \mu_n \mathbf{D} \boldsymbol{\Lambda}_{k,j}(i) e_k(i) \tag{18}$$

where μ_n is the step size for the algorithm that adjusts the feedforward section of the proposed MUD. To compute the update rule for the feedback parameter vector $\mathbf{f}_k$ of the decision feedback receiver, we compute the gradient of $J(\mathbf{W}_k(i), \mathbf{f}_k(i))$ with respect to $\mathbf{f}_k$ and obtain the following gradient-type recursion:

$$\mathbf{f}_k(i+1) = \mathbf{f}_k(i) - \mu_f e_k(i) \hat{\mathbf{b}}(i) \tag{19}$$

where μ_f is the step size of the algorithm that updates the feedback section.

6 Simulation Experiments

In this section we assess the BER and the convergence performance of the adaptive receivers. The DS-CDMA system employs Gold sequences of length $N = 15$. The carrier frequency of the system was chosen to be 1900 MHz. It is assumed here that the channels experienced by different users are statistically independent and identically distributed. The channel coefficients for each user k ($k = 1, \ldots, K$) are $h_{k,l}(i) = p_l |\alpha_{k,l}(i)|$, where $\alpha_{k,l}(i)$ ($l = 0, 1, \ldots, L_p - 1$) is a complex Gaussian random sequence obtained by passing complex white Gaussian noise through a filter with approximate transfer function $\beta / \sqrt{1 - (f/f_d)^2}$ where β is a normalization constant, $f_d = v/\lambda$ is the maximum Doppler shift, λ is the wavelength of the carrier frequency, and v is the speed of the mobile [9]. For each user, say user k, this procedure corresponds to the generation of L_p independent sequences of correlated, unit power ($E[|\alpha_{k,l}(i)|^2 = 1$), Rayleigh random variables and has a bandwidth of 4.84 MHz, which corresponds to the data rate of 312.2 kbps. The simulations assess and compare the BER performance of the DF and linear receivers operating with the LMS, DF and non-DF neural MUDs operating with the algorithms of Section V, the SUD and the single user bound (SU-Bound), which corresponds to the SUD in a system with a single user (no MAI). Note that for the neural receiver without DF we make $\mathbf{f} = \mathbf{0}$ in the structure and algorithms of Sections IV and V. The parameters of the algorithms are optimised for each situation and we assume perfect power control in the DS-CDMA system. The activation function $\varphi(.)$ for the neural receiver is the hyperbolic tangent (i.e. $tanh(.)$) in all simulations. The receivers process 10^3 data symbols, averaged over 100 independent experiments in a scenario where the mobile terminals move at 80km/h. The algorithms are adjusted with 200 training data symbols during the training period and then switch to decision directed mode in all experiments. We remark that the BER performance shown in the results refers to the average BER amongst the K users.

6.1 Flat Rayleigh fading channel performance

The BER and the BER convergence performance of the receivers were evaluated in a flat Rayleigh fading channel ($L_p = 1$, $p_0 = 1$) with additive white gaussian noise (AWGN). The BER convergence performance of the MUDs is shown in Fig. 2, where the proposed DF neural MUD achieves the best performance, followed by the neural receiver without DF, the DF MUD with linear FIR filters, the linear receiver and the SUD. In Figs. 3 and 4 the BER performance versus the number of users (K) and versus E_b/N_0, respectively, is illustrated. The results show that the novel DF neural MUD achieves the best BER performance, outperforming the neural MUD, the DF MUD, the linear MUD and the SUD.

6.2 Multipath Rayleigh fading channel performance

The BER and the BER convergence performance of the receivers are now assessed in a frequency selective Rayleigh fading channel with AWGN. The channel is modeled as a two-path ($L_p = 2$) and the parameters are $p_0 = 0.895$ and $p_1 = 0.447$ for each user. The BER convergence performance is shown in Fig. 5, whereas the BER performance versus the number of users (K) and versus E_b/N_0 is depicted in Figs. 6 and 7, respectively. The results show that the new DF neural MUD has the best BER performance, outperforming the neural MUD, the DF MUD, the linear MUD and the SUD.

7 Concluding Remarks

In this paper we proposed adaptive DF multiuser receivers for DS-CDMA systems using recurrent neural networks. We developed SG adaptive algorithms for estimating the parameters of the feedforward and feedback sections of the new receiver. A comparative analysis through computer simulation experiments was carried out for evaluating the BER performance of the proposed receiver and algorithms. The results have shown that the novel DF neural MUD receiver outperforms the other analysed structures.

References

1. "Physical Layer Standard for CDMA2000 Spread Spectrum Systems (Release C)," Telecommunications Industry Association, TIA/EIA/IS- 2002.2-C, May 2002.
2. B. Aazhang, B. P. Paris and G. C. Orsak,"Neural Networks for Multiuser Detection in Code-Division-Multiple-Access Communications," *IEEE Transactions on Communications*, vol. 40, No. 7, July 1992.
3. U. Mitra and H. V. Poor, "Neural Network Techniques for Adaptive Multi-user Demodulation," *IEEE Journal on Selected Areas of Communications*, Vol. 12, No. 9, December 1994.
4. W. G. Teich, M. Seidl and M. Nold," Multiuser Detection for DS-CDMA Communication Systems based on Recurrent Neural Networks Structures", *Proc. IEEE ISSSTA98*, 02-04 September, Sun City, South Africa, pp. 863-867, 1998.
5. R. C. de Lamare and R. Sampaio-Neto, "Analysis of Adaptive Multiuser Receivers for DS-CDMA using Recurrent Neural Networks", *Proc. SBrT/IEEE International Telecommunications Symposium*, September 2002.
6. R. C. de Lamare, R. Sampaio-Neto, "Adaptive MBER decision feedback multiuser receivers in frequency selective fading channels", *IEEE Communications Letters*, vol. 7, no. 2, Feb. 2003, pp. 73 - 75.

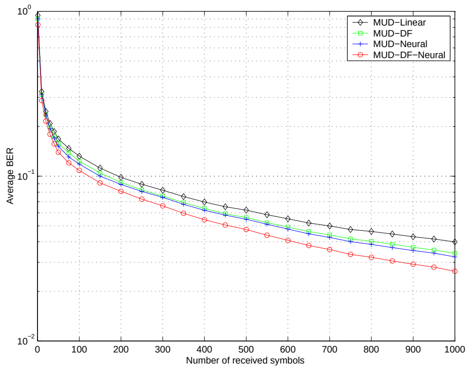

Fig. 2. BER convergence performance of the receivers in a flat Rayleigh fading channel with AWGN, $E_b/N_0 = 8$ dB and with $K = 4$ users. The parameters of the algorithms are $\mu_w = 0.005$, $\mu_n = 0.005$ and $\mu_f = 0.0015$.

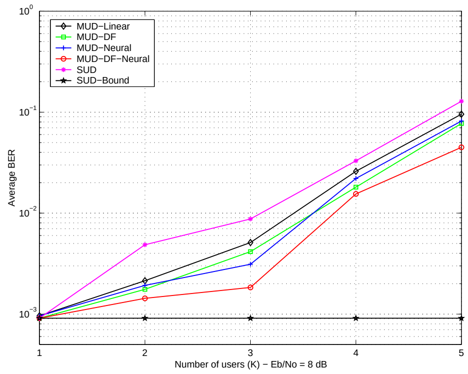

Fig. 3. BER performance of the receivers in a flat Rayleigh fading channel with AWGN, $E_b/N_0 = 8$ dB and with a varying number of users. The parameters of the algorithms are $\mu_w = 0.005$, $\mu_n = 0.005$ and $\mu_f = 0.0015$.

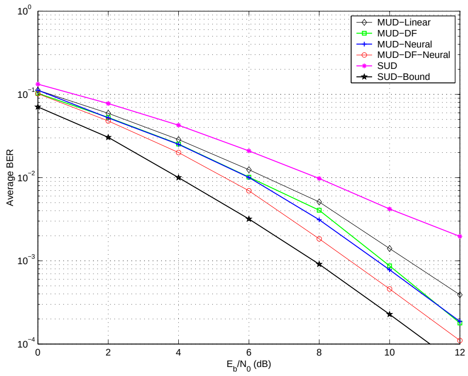

Fig. 4. BER performance versus E_b/N_0 for the receivers in a flat Rayleigh fading channel with AWGN and $K = 3$ users. The parameters of the algorithms are $\mu_w = 0.005$, $\mu_n = 0.005$ and $\mu_f = 0.0015$.

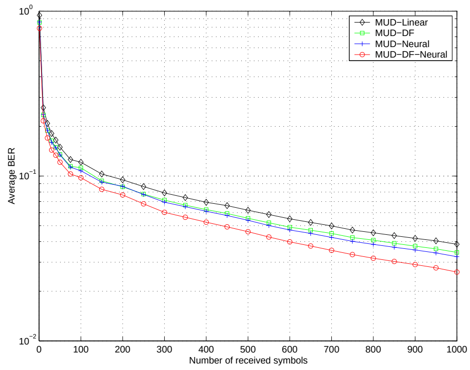

Fig. 5. BER convergence performance of the receivers in a two-path Rayleigh fading channel with AWGN, $E_b/N_0 = 10$ dB and with $K = 4$ users. The parameters are $\mu_w = 0.0025$, $\mu_n = 0.0025$ and $\mu_f = 0.0015$.

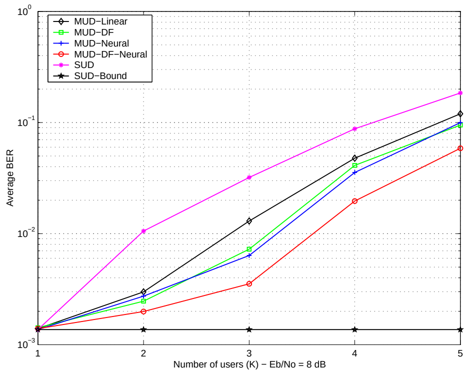

Fig. 6. BER performance versus number of users for the receivers in a two-path Rayleigh fading channel with AWGN, $E_b/N_0 = 8$ dB. The parameters of the algorithms are $\mu_w = 0.0025$, $\mu_n = 0.0025$ and $\mu_f = 0.0015$.

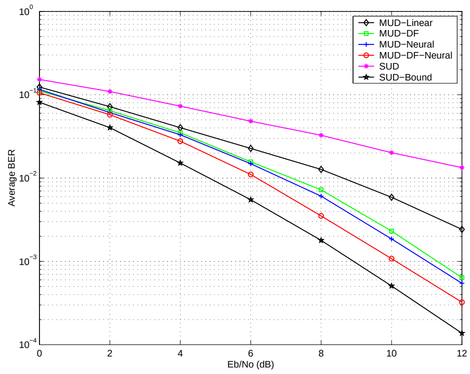

Fig. 7. BER performance versus E_b/N_0 for the receivers in a two-path Rayleigh fading channel with AWGN and $K = 3$ users. The parameters of the algorithms are $\mu_w = 0.0025$, $\mu_n = 0.0025$ and $\mu_f = 0.0015$.

7. G. Woodward, R. Ratasuk, M. Honig. and P. Rapajic, "Multistage decision-feedback detection for DS-CDMA," *Proc. IEEE Int. Conf. Communications (ICC)*, June 1999.
8. R. C. de Lamare and R. Sampaio-Neto, "Minimum Mean Squared Error Iterative Successive Parallel Arbitrated Decision Feedback Detectors for DS-CDMA Systems," *IEEE Transactions on Communications*, submitted.
9. S. Haykin, *Adaptive Filter Theory*, 3rd edition, Prentice-Hall, Englewood Cliffs, NJ, 1996.
10. S. Haykin, *Neural Networks: A Comprehensive Foundation*, 2nd Edition, Prentice-Hall, 1999.
11. T. S. Rappaport, *Wireless Communications*, Prentice-Hall, Englewood Cliffs, NJ, 1996.