ADAPTIVE BAYESIAN EQUALIZER WITH SUPERIMPOSED TRAINING FOR MIMO CHANNELS

R. Krummenauer[1], *F. de S. Chaves*[1], *R. Ferrari*[1], *M. Uliani Neto*[1,2], *J. M. T. Romano*[1], *A. Lopes*[1]

[1]School of Electrical and Computer Engineering, FEEC - UNICAMP, Campinas-SP, Brazil
[2]Foundation CPqD, Campinas-SP, Brazil
Emails: {rkrummen, fabiano, rferrari, romano, amauri}@decom.fee.unicamp.br, uliani@cpqd.com.br

ABSTRACT

The Bayesian equalizer is implementable by a proper employment of a radial basis function (RBF) neural network, with the inverse filtering problem posed as a classification problem. The proposed approach allows that the transmission of information and the RBF training be accomplished in a simultaneous and uninterrupted way. Moreover, the channel estimation procedure remains an unimodal optimization problem. Simulation results confirm the effectiveness of the proposed MIMO equalizer.

Index Terms— Bayesian equalizer, radial basis function (RBF) neural networks, superimposed training, MIMO channels

1. INTRODUCTION

Communication systems have evolved towards the achievement of strict service requirements as low error rates and high transmission rates. In this context, the capability of multiple-input multiple-output (MIMO) channels to increase the achievable system capacity has motivated intense research in the field of space-time signal processing [1, 2, 3].

Space-time equalization has been employed to suppress both intersymbol interference (ISI) and co-channel interference (CCI) in MIMO communication systems. From the viewpoint of signal detection, there are basically two categories of equalizers, namely sequence-estimation and symbol-by-symbol-decision equalizers. In the first category, the optimum equalizer is the maximum likelihood sequence estimator (MLSE) [4], which uses the Viterbi algorithm and presents high computational complexity. In contrast, the addressed symbol-by-symbol-decision equalizers are typically based on adaptive linear filter design, yielding simpler computational requirements.

In the context of single-input single-output (SISO) channels, linear equalization may not be able to compensate for channel effects in received signals and it becomes necessary to employ a nonlinear processing to recover the transmitted information [5]. Actually, the optimum equalizer structure without decision feedback is nonlinear and is given by the maximum *a posteriori* (MAP) equalizer, also known as Bayesian equalizer [6]. It is known that the Bayesian equalization solution can be implemented using radial basis function (RBF) neural networks, where the original inverse filtering or deconvolution problem is posed as a classification problem [7, 8].

The same Bayesian approach can be extended to MIMO communication systems, where RBF networks can implement MAP processing for multi-user detection (MUD) [9] or space-time equalization [10]. The training of an RBF network to perform MAP equalization can be carried out efficiently by exploiting the underlying data structure. The RBF parameters must be estimated and, for this purpose, it is common to use a training sequence. Two strategies are possible. The first one consists in executing a supervised clustering procedure to find the RBF centers according to data distribution. This strategy is suited for both linear and nonlinear channels, but the number of parameters to be estimated, and the computational complexity, grows exponentially with channel memory, number of transmitted signals and number of equalizer inputs. For linear channels, a second strategy is applicable, where the MIMO channel estimation is performed by a conventional adaptive algorithm and used to calculate the RBF centers [10]. In this approach, the number of parameters to be estimated grows only linearly with channel memory and a smaller training set is needed.

In this work, we explore a superimposed training approach [11, 12] for adapting parameters of the RBF-MAP equalizer. A low-power periodic training sequence is added (superimposed) to the information sequence at the transmitter. Unlike traditional training methods, no time slots are allocated for training. Then, there is no loss in information rate, although some useful power is wasted in the superimposed sequence. This approach permits MIMO channel estimation and, as consequence, the calculation of channel states, which is required to implement MAP equalization.

The rest of the paper is organized as follows. In Section 2, the conventional MIMO equalization model is pre-

R. Krummenauer, F. de S. Chaves and R. Ferrari are supported by PhD scholarships from CNPq, FAPESP and CAPES, respectively.

sented. Section 3 introduces the Bayesian MIMO equalizer and shows that it has the same structure of an RBF neural network. The proposed scheme of superimposed training for the RBF network is explained in Section 4. In Section 5, simulation results illustrate the proposed MIMO equalizer performance. Conclusions of this work are given in Section 6.

2. SYSTEM MODEL

A generic MIMO communication system is illustrated in Fig. 1. Subchannels $\mathbf{h}_{ij} = [h_{ij}^1, \dots, h_{ij}^L]^T$, $i = \{1, \dots, N\}$ and $j = \{1, \dots, M\}$, represent L paths of the link between the transmission sensor j and the reception sensor i. Each transmission sensor transmits an independent identically distributed (i.i.d.) sequence of BPSK symbols, $s_j(k)$, belonging to the set $\mathbb{A} = \{-1; +1\}$. Signals that arrive at reception sensors, $x_i(k)$, are added to a white gaussian noise, $\eta_i(k)$, and the resulting signals, $y_i(k)$, $i = \{1, ..., N\}$, represent the receiver input signals. Receiver output must provide estimates of transmitted signals, $\hat{s}_j(k - d_j)$, $j = \{1, ..., M\}$, given an equalization delay d_j.

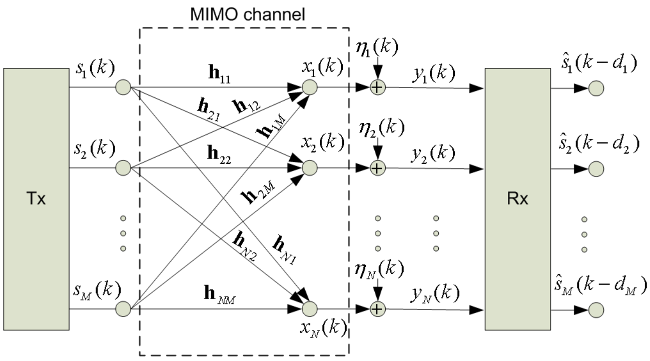

Fig. 1. MIMO communication system.

MIMO communication systems are classified according to the levels of coordination between their sensors [13]. Our interest in this work lies in systems where reception sensors are fully coordinated, that is, received signals can be jointly processed to recover the signals transmitted by each transmission sensor. MIMO single-user and MIMO multi-user multi-access systems are examples of such systems. In this work the focus is on MIMO single-user systems.

2.1. MIMO Equalization

MIMO equalizers are designed to exploit diversity in space and time domains to combat effects of MIMO channels. In order to model MIMO equalization, the channel output or equalizer input can be expressed by the concatenated vector:

$$\mathbf{y}_c(k) = [\mathbf{y}^T(k), \dots, \mathbf{y}^T(k - N_{in} + 1)]^T \in \mathbb{R}^{N \cdot N_{in} \times 1},$$

where $\mathbf{y}(k) = [y_1(k), \dots, y_N(k)]^T \in \mathbb{R}^{N \times 1}$ is a column vector containing the receiver input signals at instant k and N_{in} is the number of samples to be taken into account for equalization.

The channel convolution matrix is defined as:

$$\mathbf{H} = \begin{bmatrix} \mathbf{H}_0 & \cdots & \mathbf{H}_{L-1} & \mathbf{0} & \cdots & \mathbf{0} \\ & & \ddots & & \ddots & \\ \mathbf{0} & \cdots & \mathbf{0} & \mathbf{H}_0 & \cdots & \mathbf{H}_{L-1} \end{bmatrix}, \quad (1)$$

with $\mathbf{H} \in \mathbb{R}^{N \cdot N_{in} \times M \cdot (L + N_{in} - 1)}$. Matrices $\mathbf{H}_l \in \mathbb{R}^{N \times M}$ are given by:

$$\mathbf{H}_l = \begin{bmatrix} h_{11}^l & \cdots & h_{1M}^l \\ \vdots & & \vdots \\ h_{N1}^l & \cdots & h_{NM}^l \end{bmatrix}, \quad (2)$$

where h_{ij}^l is the gain of subchannel $\mathbf{h}_{ij}$ at the l-th path, $l = \{0, \dots, L-1\}$. From these definitions, the following model holds for the channel output:

$$\mathbf{y}_c(k) = \mathbf{H}\mathbf{s}_c(k) + \boldsymbol{\eta}_c(k), \quad (3)$$

with

$$\mathbf{s}_c(k) = [\mathbf{s}^T(k), \dots, \mathbf{s}^T(k - L - N_{in} + 1)]^T \quad (4)$$
$$\boldsymbol{\eta}_c(k) = [\boldsymbol{\eta}^T(k), \dots, \boldsymbol{\eta}^T(k - N_{in} + 1)]^T, \quad (5)$$

where $\mathbf{s}_c(k) \in \mathbb{A}^{M \cdot (L + N_{in} - 1) \times 1}$ is the concatenated vector containing the vectors of transmitted signals $\mathbf{s}(k) = [s_1(k), \dots, s_M(k)$ and $\boldsymbol{\eta}_c(k) \in \mathbb{R}^{N \cdot N_{in} \times 1}$ is the concatenated vector containing the vectors of sensor noise $\boldsymbol{\eta}(k) = [\eta_1(k), \dots, \eta_N(k)]^T \in \mathbb{R}^{N \times 1}$.

3. BAYESIAN MIMO EQUALIZER

In the proposed equalization context, the main objective is to recover each transmitted symbol $s_j(k - d_j)$, $j = \{1, \dots, M\}$, where d_j denotes the equalization delay for the j-th source. The equalizer can be seen as a classifier that divides the space spanned by the data vector $\mathbf{y}_c(k)$ into S partitions corresponding to each of the possible values of $s_j(k - d_j)$. The optimum decision boundaries depend on the equalization delay, data noise and channel states. The channel states are defined as the possible values of received signals in the absence of noise. The p-th state associated with the transmitted sequence of symbols $\mathbf{s}_p$ is defined as

$$\mathbf{c}_p = \mathrm{E}\{\mathbf{y}_c | \mathbf{s}_p\}, \quad (6)$$

where $\mathrm{E}\{\cdot\}$ denotes the expectation operator and $\mathbf{s}_p \in \mathbb{A}^{M \cdot (L + N_{in} - 1)}$. The number of channel states is equal to the number of possible different sequences $\mathbf{s}_p$ and, thus, we have $N_s = 2^{M \cdot (L + N_{in} - 1)}$ states, since the transmitted symbols belong to a binary alphabet and the length of the sequence $\mathbf{s}_p$ is $M \cdot (L + N_{in} - 1)$.

Note that it is not dependent on the number of reception sensors.

The channel states of order N_{in} are given by the columns of the matrix:

$$\mathbf{C}_L = \mathbf{H}\,\mathbf{S} = \left[\begin{array}{cccc} \mathbf{c}_1 & \mathbf{c}_2 & \cdots & \mathbf{c}_{N_s} \end{array} \right] , \quad (7)$$

where $\mathbf{C}_L \in \mathbb{R}^{N \cdot N_{in} \times N_s}$, $\mathbf{H}$ is the channel convolution matrix given by (1), and $\mathbf{S} \in \mathbb{A}^{M \cdot (L+N_{in}-1) \times N_s}$ is a matrix whose N_s columns are formed by all possible distinct symbol sequences $\mathbf{s}_p$.

As mentioned above, the Bayesian equalizer is derived from the maximum *a posteriori* (MAP) criterion [6]. We shall refer to this equalizer as maximum *a posteriori* space-time equalizer (MAP-STE). The MAP-STE decision function is determined by the conditional probability density function of vector $\mathbf{y}_c(k)$ with respect to channel states, $p(\mathbf{y}_c(k)|\mathbf{c}_p)$. Thus, to minimize the probability of classification error, it is needed to maximize the *a posteriori* probability of transmitted symbol, $P(s_j(k-d_j)|\mathbf{y}_c(k))$. Since the symbol alphabet is binary, this maximization process can be carried out by using the decision function below:

$$\begin{aligned} f_{B,j}(\mathbf{y}_c(k)) \quad = \quad & P(s_j(k-d_j) = +1|\mathbf{y}_c(k)) \\ & -P(s_j(k-d_j) = -1|\mathbf{y}_c(k)) \,, \quad (8) \end{aligned}$$

where the value of $f_{B,j}(\mathbf{y}_c(k))$ is positive when $P(s_j(k-d_j) = +1|\mathbf{y}_c(k)) \geq P(s_j(k-d_j) = -1|\mathbf{y}_c(k))$ and negative when $P(s_j(k-d_j) = +1|\mathbf{y}_c(k)) < P(s_j(k-d_j) = -1|\mathbf{y}_c(k))$.

Therefore, symbol detection in MAP-STE is accomplished as follows:

$$\begin{aligned} \hat{s}_j(k-d_j) \quad &= \quad \mathrm{sgn}(f_{B,j}(\mathbf{y}_c(k))) \\ &= \quad \left\{ \begin{array}{ll} +1 & \text{if } f_{B,j}(\mathbf{y}_c(k)) \geq 0 \\ -1 & \text{if } f_{B,j}(\mathbf{y}_c(k)) < 0 \end{array} \right. . \quad (9) \end{aligned}$$

It can be shown that the MAP-STE decision function for additive white gaussian noise (AWGN) channels can be written in the form [10]:

$$f_{B,j}(\mathbf{y}_c(k)) = \sum_{p=1}^{N_s} w_{p,j} \exp\left(-\frac{\|\mathbf{y}_c(k) - \mathbf{c}_p\|^2}{2\sigma^2} \right) , \quad (10)$$

where $w_{p,j} = +1$ if $\mathbf{c}_p$ belongs to the set of states for which $s_j(k-d_j) = +1$, and $w_{p,j} = -1$ if $\mathbf{c}_p$ belongs to the set of states for which $s_j(k-d_j) = -1$. Then, the MAP-STE decision function given in (10) is nonlinear and completely defined by channel states and noise statistics.

The idea of implementing the MAP-STE by using an RBF neural network comes from the mathematical similarity between the MAP-STE decision function and the input-output mapping of an RBF network [10]:

$$\begin{aligned} f_{RBF,j}(\mathbf{y}_c(k)) \quad &= \quad \sum_{p=1}^{N_s} w_{p,j} \varphi_p(\mathbf{y}_c(k)) \\ &= \quad \sum_{p=1}^{N_s} w_{p,j} \exp\left(-\frac{\|\mathbf{y}_c(k) - \boldsymbol{\mu}_p\|^2}{\rho} \right) , \end{aligned}$$

where the radial basis functions $\varphi_p(\cdot)$ are Gaussian, with dispersion $\rho = 2\sigma^2$ and centered at the points $\boldsymbol{\mu}_p$ associated with the respective channel states $\mathbf{c}_p$. The weights $w_{p,j}$ of the output layer are set to ± 1 according to the value of $s_j(k-d_j)$ associated to the respective channel state c_p.

Therefore, the RBF network structure corresponds to the Bayesian equalizer. Perfect knowledge of the channel and noise statistics allows the implementation of the optimum MAP-STE through an RBF network. However, such information is not available *a priori* and the RBF must be trained to estimate its parameters. In next section, the scheme of network training is addressed.

4. SUPERIMPOSED TRAINING-BASED MIMO CHANNEL ESTIMATION

In this work, the Bayesian equalizer for MIMO channels is implemented through an RBF neural network, where MIMO channel estimates are used to calculate the RBF centers. Channel estimation is commonly performed by conventional supervised techniques, where the transmission of information is periodically interrupted for the transmission of a training sequence. Regarding the process of optimization, supervised techniques are characterized by unimodal cost functions during the training period. In contrast, unsupervised techniques can perform a continuous process of training without interruption of information transmission. However, such techniques are faced with local optima. Our proposal for estimating the MIMO channel is based on the transmission of an uninterrupted pilot sequence superimposed to each information sequence. The advantage of this approach is that it combines the best features from both supervised and unsupervised techniques: it permits uninterrupted and simultaneous processes of information transmission and channel estimation, and the cost function to be optimized is unimodal. On the other hand, some power must be used in the superimposed training sequence.

The scheme of MIMO channel estimation based on superimposed training is illustrated in Fig. 2, where M transmission sensors and N reception sensors are considered. The MIMO channel to be estimated is composed of subchannels $\mathbf{h}_{ij} = [h_{ij}^1, \ldots, h_{ij}^L]^T$, with $i = \{1, \ldots, N\}$, $j = \{1, \ldots, M\}$, and L paths in each link. Using a spread spectrum technique, a binary pilot sequence $\mathbf{m}_j(k) = [m_j(k), \ldots, m_j(k-L+1)]^T$ is added to information sequence $\mathbf{s}_j(k) = [s_j(k), \ldots, s_j(k-$

$L+1)]^T$ at each transmission sensor. Therefore, the sequence to be transmitted becomes $\mathbf{t}_j(k) = \mathbf{s}_j(k) + \mathbf{m}_j(k)$, for $j = \{1, \dots, M\}$. The pilot sequence is a periodic nonrandom signal uncorrelated with the information signal.

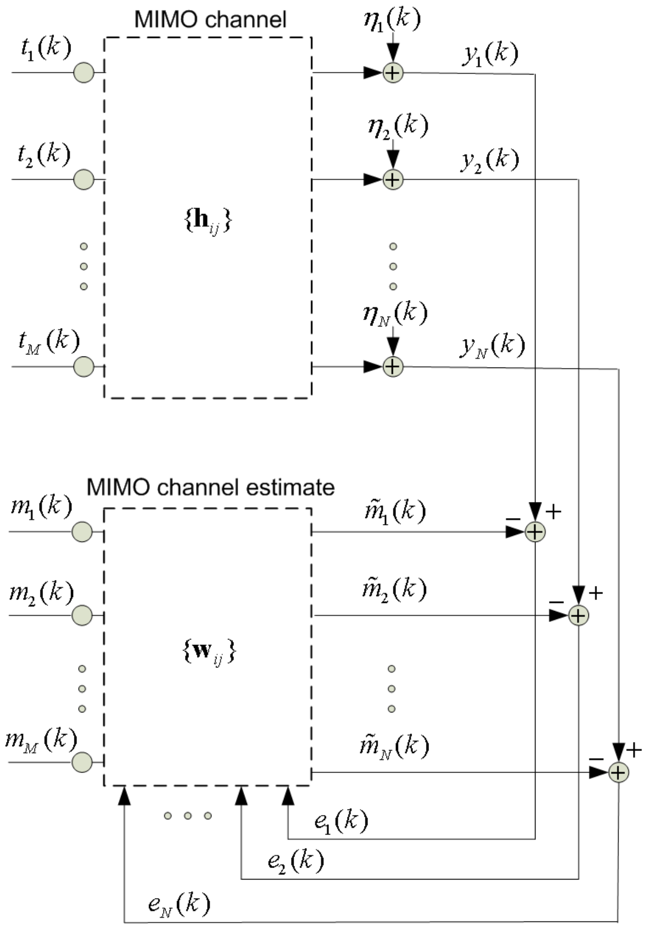

Fig. 2. MIMO channel estimation based on superimposed training.

At each symbol interval, each receiver input $y_i(k)$, $i = \{1, \dots, N\}$, can be expressed as:

$$y_i(k) = \left[\mathbf{h}_{i1}^T, \dots, \mathbf{h}_{iM}^T\right] \begin{bmatrix} \mathbf{s}_1(k) + \mathbf{m}_1(k) \\ \vdots \\ \mathbf{s}_M(k) + \mathbf{m}_M(k) \end{bmatrix} + \eta_i(k). \tag{11}$$

In order to obtain the MIMO channel estimates $\mathbf{w}_{ij} = [w_{ij}^1, \dots, w_{ij}^L]^T$, $i = \{1, \dots, N\}$, $j = \{1, \dots, M\}$, one can use the error signals $e_i(k)$,

$$e_i(k) = \tilde{m}_i(k) - y_i(k), \tag{12}$$

where $\tilde{m}_i(k)$, $i = \{1, \dots, N\}$, are the received pilot signals passed through the estimated channel:

$$\tilde{m}_i(k) = \left[\mathbf{w}_{i1}^T, \dots, \mathbf{w}_{iM}^T\right] \begin{bmatrix} \mathbf{m}_1(k) \\ \vdots \\ \mathbf{m}_M(k) \end{bmatrix}. \tag{13}$$

According to the Wiener criterion, a cost function $J = [e_i^2(k)]$ is defined as the expectation of the squared error for each subchannel vector to be estimated as follows:

$$\begin{aligned} J(\mathbf{w}_{ci}) = {} & \mathrm{E}\left[y_i^2(k)\right] - \mathbf{w}_{ci}^T \mathrm{E}\left[\mathbf{m}_c(k)\mathbf{m}_c^T(k)\mathbf{h}_{ci}\right] \\ & - \mathrm{E}\left[\mathbf{h}_{ci}^T \mathbf{m}_c(k)\mathbf{m}_c^T(k)\right]\mathbf{w}_{ci} \\ & + \mathbf{w}_{ci}^T \mathrm{E}\left[\mathbf{m}_c(k)\mathbf{m}_c^T(k)\right]\mathbf{w}_{ci}, \end{aligned} \tag{14}$$

where $\mathbf{h}_{ci} = [\mathbf{h}_{i1}^T, \dots, \mathbf{h}_{iM}^T]$, $\mathbf{w}_{ci} = [\mathbf{w}_{i1}^T, \dots, \mathbf{w}_{iM}^T]$ and $\mathbf{m}_c = [\mathbf{m}_1(k), \dots, \mathbf{m}_M(k)]^T$. Since the insertion of the pilot sequence into the signal to be transmitted is carried out by a spread spectrum procedure, signals $\mathbf{s}_j$ and $\mathbf{m}_j$ are independent. Then, if they are zero-mean sequences, the following relations can be stated:

$$\begin{aligned} \mathrm{E}\left[y_i^2(k)\right] &= \sigma_{y_i}^2 \\ \mathrm{E}\left[\mathbf{m}_c(k)\mathbf{m}_c^T(k)\mathbf{h}_{ci}\right] = \begin{bmatrix} \mathbf{p}_{i1} \\ \vdots \\ \mathbf{p}_{iM} \end{bmatrix} &= \mathbf{p}_{ci} \\ \mathrm{E}\left[\mathbf{m}_c(k)\mathbf{m}_c^T(k)\right] = \begin{bmatrix} \mathbf{R}_1 & 0 & \cdots & 0 \\ 0 & \mathbf{R}_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \mathbf{R}_M \end{bmatrix} &= \mathbf{R}_c, \end{aligned}$$

where $\sigma_{y_i}^2$ is the power of receiver input signal y_i; $\mathbf{p}_{ci}$ and $\mathbf{R}_c$ represent the cross-correlation and the autocorrelation of each pilot sequence $\mathbf{m}_c$, respectively.

From (14) and relations above, one can write the Wiener error surface in canonical form for the MIMO channel estimation using a pilot sequence:

$$J(\mathbf{w}_{ci}) = \sigma_{y_i}^2 - \mathbf{w}_{ci}^T \mathbf{p}_{ci} - \mathbf{p}_{ci}^T \mathbf{w}_{ci} + \mathbf{w}_{ci}^T \mathbf{R}_c \mathbf{w}_{ci}. \tag{15}$$

Then, the Wiener solution applies, with the optimum estimates $\mathbf{w}_{ci}^o$ given by:

$$\mathbf{w}_{ci}^o = \mathbf{R}_c^{-1}\mathbf{p}_{ci}, \tag{16}$$

and as consequence, the minimum value of cost function J is expressed as:

$$\min_{\mathbf{w}_{ci}} J(\mathbf{w}_{ci}) = \sigma_{y_i}^2 - \mathbf{p}_{ci}^T \mathbf{R}_c^{-1}\mathbf{p}_{ci}. \tag{17}$$

It can be observed that the optimum solution to this MIMO channel estimation problem depends only on the pilot sequence. The information signal has influence only on the minimum value of the cost function, but not on the optimum solution $\mathbf{w}_{ci}^o$.

In practice, the MIMO channel estimation task can be carried out by conventional adaptive algorithms. Because of its convergence properties, we employ the recursive least squares (RLS) algorithm. The RLS algorithm is executed for each reception sensor i, where the input signal is y_i. As can be observed in (11), the subchannels related to sensor i are arranged in a concatenated vector, $\left[\mathbf{h}_{i1}^T, \ldots, \mathbf{h}_{iM}^T\right]$. Then, the corresponding concatenated vector of subchannels estimates, $\left[\mathbf{w}_{i1}^T, \ldots, \mathbf{w}_{iM}^T\right]$, can be obtained by using the RLS algorithm.

5. SIMULATION RESULTS

In this section, we evaluate the proposed adaptive equalizer for MIMO channels as an implementation of the Bayesian (MAP-STE) equalizer through computer simulations. We consider a communication system with $M = 2$ transmission sensors and $N = 2$ reception sensors. The considered MIMO channel has $L = 2$ paths and is expressed below:

$$\mathbf{H}_0 = \begin{bmatrix} 1 & 0.5 \\ 0.8 & 0.6 \end{bmatrix}; \ \mathbf{H}_1 = \begin{bmatrix} 0.6 & 1.2 \\ 0.3 & 0.9 \end{bmatrix}. \qquad (18)$$

As exposed in Section 4, the RLS algorithm uses the superimposed training sequence to perform the MIMO channel estimation. Fig. 3 shows the mean squared error (MSE) of channel estimation along the RLS iterations for three values of information-to-training power ratio (ITR), namely 10 dB, 15 dB and 20 dB. Different ITR values are accomplished by adjusting the training sequence power, since the information sequence power level is fixed. The presented MSE is calculated as follows:

$$\text{MSE} = \frac{\sum_{l=0}^{L-1}\sum_{j=1}^{M}\sum_{i=1}^{N}\left|h_{i,j}^l - w_{i,j}^l\right|^2}{L \cdot M \cdot N}, \qquad (19)$$

and these curves represent the average of 5000 simulations. The signal-to-noise power ratio (SNR) is 20 dB.

It can be observed that lower values of ITR provide better channel estimates. The estimation task becomes easier for lower ITR values, since they mean training signals with higher power. On the other hand, it is noted that ITR does not influence the speed of convergence.

In order to evaluate the proposed equalizer as an implementation of the MAP-STE, we present in Fig. 4 the bit error rate (BER) curves of the MAP-STE and of three configurations of the proposed equalizer. Such curves represent the overall BER, i.e., they are related to the two transmitted sequences. The same equalization delay is considered for both

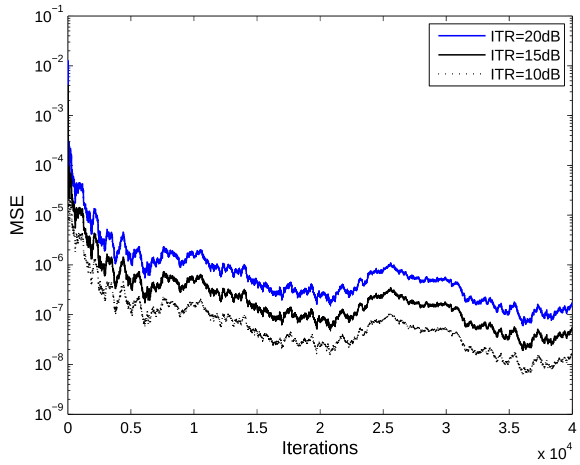

Fig. 3. MSE of superimposed training-based channel estimation for three values of ITR and SNR $= 20$ dB.

received signals: $d_1 = d_2 = 1$ symbol interval. Moreover, the number of samples of the received vectors used as equalizer inputs is $N_{in} = 2$. The BER curve of the Wiener minimum mean squared error (MMSE) linear equalizer is also illustrated in Fig. 4. In this case, $N_{in} = 15$ equalizer inputs were considered. It is noted that the performance of the Wiener equalizer is dramatically poor for the considered channel.

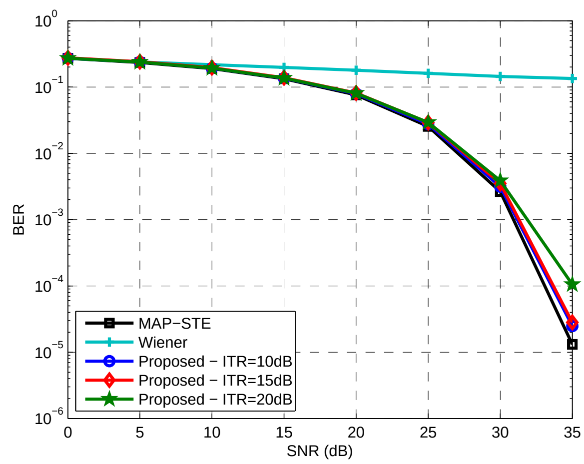

Fig. 4. BER performance of MAP-STE, proposed equalizer, and Wiener linear equalizer.

For MAP-STE, the available power is entirely used in the information signal transmission, since it is the optimum receiver, where perfect channel information is assumed. Different configurations of the proposed equalizer are defined by sharing the available transmission power between information and training signals, such that the desired ITR values be obtained. Towards a fair comparison, the same amount of power is available in all simulations.

It is expected that the proposed equalizer performs better for lower values of ITR, since in such scenarios the quality of channel estimates is better. This behavior is confirmed, but it is more evident for high SNR values, where the equalization process is more influenced by the quality of channel estimates. In scenarios with low or moderate values of SNR, the different configurations of the proposed equalizer in terms of ITR present BER curves very close to that of MAP-STE. It is worth noting that from an ITR about 15 dB, the proposed equalizer practically attains the BER performance of the MAP-STE, even for high SNR.

6. CONCLUSIONS

The implementation of the Bayesian equalizer for MIMO channels by means of radial basis function (RBF) neural networks is addressed. It is known that RBF networks must be trained to perform their tasks properly. In this context, it is common to use a training sequence in supervised procedures of clustering or channel estimation to define the RBF parameters.

In this work, we propose the use of a low-power superimposed pilot sequence to perform the MIMO channel estimation and the subsequent adaptation of the RBF parameters. This approach has the advantage to allow an uninterrupted information transmission and network training, both in parallel. Moreover, the channel estimation problem corresponds to the optimization of a unimodal cost function that leads to the optimal solution according to the Wiener criterion. Simulation results reveal that the performance of the proposed equalizer comes close to that of the Bayesian equalizer, even for low-power superimposed pilot sequences.

7. REFERENCES

[1] E. Telatar, "Capacity of multi-antenna gaussian channels," *European Trans. Telecommunications*, vol. 10, no. 6, pp. 585–596, Nov. 1999.

[2] J. H. Winters, "Optimum combining in digital mobile radio with cochannel interference," *IEEE Trans. Vehicular Technology*, vol. 3, pp. 144–155, Aug. 1984.

[3] A. J. Paulraj and C. B. Papadias, "Space-time processing for wireless communications," *IEEE Signal Processing Magazine*, vol. 14, no. 6, pp. 49–83, Nov. 1997.

[4] W. van Etteny, "Maximum-likelihood receiver for multiple channel transmission systems," *IEEE Trans. Communications*, vol. 24, no. 2, pp. 276283, Feb. 1976.

[5] J. Montalvo, B. Dorizzi, and J. C. M. Mota, "Some theoretical limits of efficiency of linear and nonlinear equalizers," *Journal of the Brazilian Telecommunications Society*, vol. 14, no. 2, pp. 85–92, Nov. 1999.

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

[7] 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, no. 4, pp. 570–579, Jul. 1993.

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

[9] K. Ko, S. Choi, C. Kang, and D. Hong, "RBF-based multi-user detector with channel estimation capability in a synchronous MC-CDMA system," *IEEE Trans. Neural Networks*, vol. 12, no. 6, pp. 1536–1539, Nov. 2001.

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

[11] J. K. Tugnait and W. Luo, "On channel estimation using superimposed training and first-order statistics," *IEEE Communications Letters*, vol. 7, no. 9, pp. 413–415, Sep. 2003.

[12] A. R. Varma, L. L. H. Andrew, C. R. N. Athaudage, and J. H. Manton, "Iterative algorithms for channel identification using superimposed pilots," *Proc. Australian Communications Theory Workshop*, Feb. 2004.

[13] J. R. Barry, E. A. Lee, and D. G. Messerschmitt, *Digital Communication*, Kluwer Academic Publishers, 3rd edition, 2004.