On the Convergence of Blind DD-LMS Decision-Feedback Equalizers Using a Joint Equalization & Decoding Procedure

C. M. Panazio, A. O. Neves, and J. M. T. Romano

DSPCom, Faculty of Electrical and Computer Engineering, State University of Campinas, SP, Brazil,
panazio@decom.fee.unicamp.br, aline@decom.fee.unicamp.br, romano@decom.fee.unicamp.br
DECOM – FEEC – UNICAMP, CP 6101, CEP 13083-970
Phone: +55-19-37883703, Fax: +55-19-32891395

ABSTRACT

This paper shows that it is possible to achieve convergence to the optimum solution in a blind equalization framework with the use of least mean square algorithm in decision-directed mode (DD-LMS). In linear equalizers structures, the attainment of the optimum solution strongly depends on the filter weights initialization. We also show that decision-feedback equalizers in DD mode (DD-DFE) converges to undesired local minima, when all its weights are initialized with zeros, for a certain class of channels. However, it is possible to improve convergence and remove such local minimum by making use of joint equalization and convolutional codes. The major contribution of the present work is the convergence study of the mentioned configuration by means of a comparative analysis of the performance surface and the dynamical adaptation of both conventional DFE and joint DFE & decoding structures.

1. INTRODUCTION

It is well known that intersymbol interference (ISI) is one of the major impairments to achieve a higher capacity or a data rate improvement in communication systems. Adaptive equalizer is a classical and efficient technique for mitigating ISI in unknown or time varying channels. The most conventional approach employs a training sequence to adapt the equalizer weights into an opened-eye condition, normally using the LMS adaptive algorithm. Then the equalizer is changed to the so-called decision-directed (DD) mode, in which the effective information is transmitted . However, in some specific systems, the use of a training sequence may not be practical. The adaptation process is then said to be unsupervised and some more robust (blind) algorithms are used. In this case, only some statistics characteristics of the transmitted data symbols are known a priori.

The first objective of the present paper is to make clear the possibilities of carrying out a blind equalizer by means of a continuous employment of the decision-directed stochastic gradient algorithm (DD-LMS). We recall the limits of such technique for the conventional linear transversal (LTE) equalizer and then analyze the case of the non-linear decision-feedback (DFE) structure.

The second and main motivation is the joint application of both equalization and decoding process in the receiver, by taking into account that a convolutional error correcting code is typically used in transmission. The DFE equalizer is used in such alternative configuration since its recursive nature is particularly suitable for dealing with the corrected symbols.

The major original contribution of the work is the convergence study of the mentioned configuration, which is carried out in two steps. First we pose the performance surface for DD-DFE equalizers showing that the joint use of the decoder avoids ill-convergence in the most critical cases of “bad channels”, as defined in [1].

As far as previous works are concerned, the present one extends the results presented by Casas [1,6] regarding the class of bad channels for which the DD-DFE equalizer converges to undesired local minima, when all its weights are initialized with zeros. We also complement the work in [5], where the joint equalization-decoding structure was proposed in a supervised (non blind) context and evaluated by standard results on bit error rates, without any discussion about convergence issues.

Our work shows that the convergence to the desired global minimum in a blind DD context is possible if we take benefice of the convolutional code. This approach is shown to be even more effective than other alternative solutions proposed in order to improve convergence in DFE equalizers such as the use soft decision devices[1,7].

In order to assess such results, this paper is organized as follows. Some backgrounds on adaptive equalization are recalled in section 2. The DD-DFE structure is described in section 3, where the joint DFE & decoding procedure is also introduced. In section 4 we present and discuss the results. The conclusions and perspectives are briefly posed in section 5.

2. BACKGROUND ON DD EQUALIZERS

The DD-LMS algorithm is obtained by the minimization of the following cost function:

$$J_{DD} = E\left\{\frac{1}{2}\left|z(n) - y(n)\right|^2\right\}, \qquad (1)$$

where $z(n)$ is the decided symbol when $y(n)$ is the equalizer output. The optimization of the equalizer parameters may be carried out by the application of a steepest descent algorithm over the functional (1). Then a stochastic approximation provides the DD-LMS algorithm, given by:

$$y(n) = \mathbf{w}^H(n)\mathbf{u}(n)$$
$$e(n) = z(n) - y(n) \qquad (2)$$
$$\mathbf{w}(n+1) = \mathbf{w}(n) + \mu e^*(n)\mathbf{u}(n)$$

where $\mathbf{u}(n)$ is the equalizer tap-input vector consisting of $[u(n), u(n-1),..., u(n-N+1)]^T$, $\mathbf{w}(n) = [w_0(n)\ w_1(n)\ ...\ w_{N-1}(n)]^T$ is the equalizer tap vector, *e(n)* is the decision (DD) error signal and the constant μ is the adaptation step size.

Generally DD algorithms are used after a training sequence when the eye-diagram is sufficiently opened, in order to refine the equalization and track channel variations. In this case, an important result reported in [2] shows that DD algorithm converges to the optimal solution if a linear invariant noiseless channel is considered. Nonetheless, if the training sequence is not sufficient to open the eye, convergence may take place towards other stationary points resulting in high error rates.

The reference [3] deals with the case where a previous training sequence is not used, i.e., DD blind equalization, for binary and multilevel data. It concludes that the convergence to the global minima depends on the initialization. An illustrative example is shown in figure 1 where the contours of the performance surface of a DD linear transversal equalizer (LTE), with two coefficients, are represented. In this case the transfer function of the channel is given by $h(z) = 1+0.6z^{-1}$.

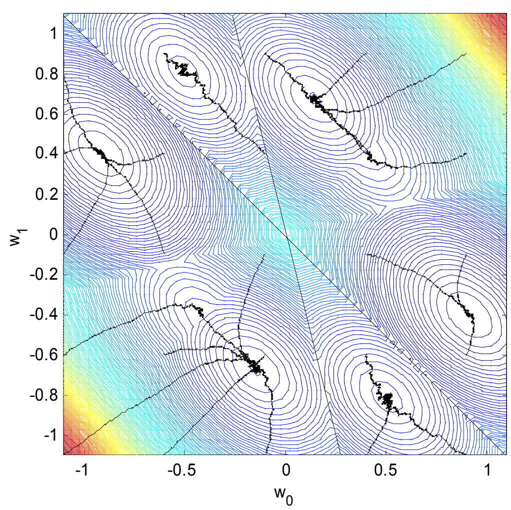

Figure 1: Trajectories of the DD-LMS algorithm over the contours of the performance surface.

We can observe the existence of four local minima located on [0.14 0.66], [-0.49 0.82] and their corresponding symmetric points. A global minimum is found in [0.91 –0.40] as well as in its symmetric position.

A number of trajectories of the LMS algorithm is also presented in figure 1 from several initial conditions into their corresponding critical points. A sufficiently small adaptation step size, ($\mu = 0.001$) was used so that the stochastic gradient could be considered a good approximation of the true one.

The two straight lines indicate a region for which the convolution channel-equalizer leads to a closed-eye condition. Such region contains the local minimum [-0.49 0.82] and its symmetrical point. It is worth pointing out that convergence to these two local minima implies in wrong decisions and such errors are fed back in the update algorithm as it can be observed in (2). Hence, if it is possible to reduce the quantity of errors generated by this closed-eye condition, we can plane or even eliminate the minima associated to these errors. This is the basic idea of using error correcting code together with the equalization process.

The same principle can be applied in DD-DFE equalizers. In this case, the approach is still more justified, since there is a kind of double feedback: via the algorithm as well as the recursive structure itself. Such case is studied in the next section.

3. THE DD-DFE EQUALIZERS

In spite of its simplicity and suitability in several applications, it is well known that linear equalizers suffer from important limitations, among which a critical one is the noise enhancement problem, in cases when the zeros of the channel are close to the unit circle. Due to its non-linear nature, DFE equalizers are an interesting alternative in such cases. Its recursive structure is also appropriate in other contexts, for instance when the channel presents a long impulse response. On the other hand, the performance of DFE structure can be affected by the phenomenon of error propagation. Figure 2 illustrates the DFE structure. The DD-LMS algorithm to be used in this case is given by:

$$y(n) = x(n) + \mathbf{b}^H(n)\mathbf{a}(n)$$
$$e(n) = \hat{a}(n) - y(n) \qquad (3)$$
$$\mathbf{b}(n+1) = \mathbf{b}(n) + \mu e^*(n)\mathbf{a}(n)$$

where *x(n)* and *y(n)* are the equalizer input and output respectively; $\hat{a}(n)$ is the decided symbol, so that its past values compose the vector $\mathbf{a}(n)=[\hat{a}(n-1)\ \hat{a}(n-2)\ ...\ \hat{a}(n-N)]^T$, which feeds the recursive and adaptive filter defined, at instant *n*, by the weight vector $\mathbf{b}(n)=[b_1(n)\ b_2(n)...\ b_{TD}(n)]^T$. The weights are updated by means of the decision error *e(n)* ,as given in (3).

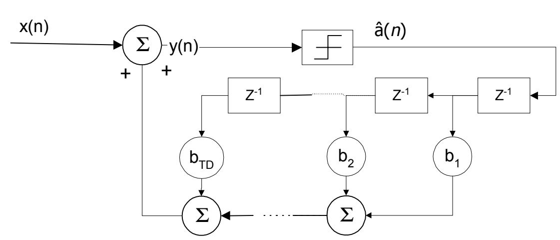

Figure 2: DFE structure.

Analysis of convergence properties of the DD-DFE can be found in [1], [4], [6] and [7]. References [1] and [6] provide a class of channels that results in ill convergence when the feedback filter coefficients are initialized with zeros. Note that this is the most natural value to be used as initialization, since it guarantees convergence when the channel has an opened-eye condition. In this work, we deal with this class of bad channels to show that error correcting codes provide better decisions to the DD-DFE so that it can converge to global minimum.

The following assumptions were made:

Assumption 1:

The source alphabet is QPSK {±1 ±j} obtained by the output of a convolutional encoder with a rate R=½. This encoder is fed by an independent and identically distributed bit sequence with p(0)= ½.

Assumption 2:

The channel has a finite impulse response and the feedback filter matches the length of the channel postcursor response. The channel has no precursor and the leading tap dominates. Hence, we can define the vector $\mathbf{h}=[h_0\ h_1\ \dots\ h_{M-1}]$, composed by M elements of the channel impulse response, where $h_0=1$ and $|h_i|<1$ for $i=1,2,\dots M-1$. Note that such condition does not imply in a minimum phase characteristic. Finally the channel is also considered to be noiseless.

Assumption 2 is very limiting for practical channels, where precursors are normally present. As a consequence of this assumption, the feedforward (FF) filter is not useful, and thereby it can be discarded. Therefore we can restrict our analysis to the feedback (FB) filter and to local minima associated with error propagation. Nevertheless the assumption is justified since a full theoretical analysis was developed in [4], where both FF and FB filters were considered,. However, this work did not take into account the impact of error correcting codes in the joint adaptation, which is the interest point of our work. Such analysis is not trivial and Assumption 2 makes it more feasible. Thereby, further studies should be done in order to include precursor ISI and corresponding FF equalizer.

The joint DFE and convolutional decoder structure to be used in this work is shown in figure 3. It can be seen that the feedback filter is divided in two parts. This is done because the output of the decoder is expected to be more reliable then the output of the decision device. However, the decoder has an intrinsic delay. For that reason, its output can only be used as the input of the feedback filter if such delay is considered.

This scheme was firstly used in [5] with the objective of reducing theerror propagation effect after a training period. In this work, we use the same idea but without a training period, i.e. in blind operation. The algorithm used for adaptation is the DD-LMS given by (3).

We tested two different convolutional codes with R= ½ , the first one with a polynomial generator [5 7] and the other with [64 74] (octal representation) [8]. A convolutional Viterbi decoder, the metric of which is a quadratic Euclidean distance, has been used in the receiver. The obtained results are presented in the sequel where different aspects of the method are discussed.

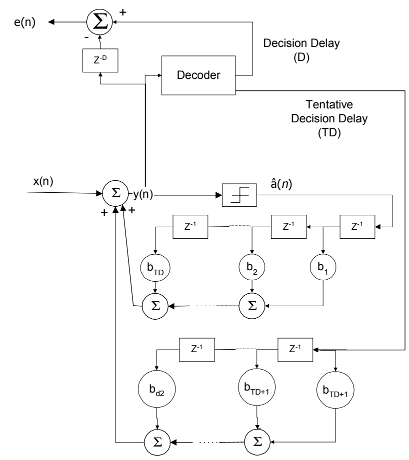

Figure 3: Joint DFE and decoding structure.

4. SIMULATION RESULTS

4.1 Avoiding Local Minimum in the Performance Surface

The channel to be considered in our first simulation is given by $h(z)=1+0.9z^{-1}-0.8z^{-2}$. It belongs to the class of bad channels defined in [6]. The error surfaces were obtained by fixing the two equalizer weights and transmitting a long sequence of about 3500 symbols. If ergodicity is assumed, a time average of the quadratic error can be calculated and used as an approximation of the cost function for each fixed pair of weights. The procedure is then repeated for a sufficient number of distinct weights so that the surface contours be refined enough to render possible the analysis of critical points.

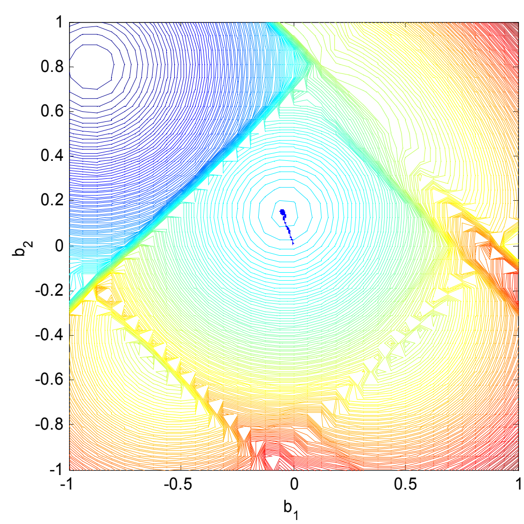

Figure 4: Contours of the performance surface for the conventional DFE.

Figure 4 shows the result obtained using the DFE-structure shown in figure 2. The transmitted symbols were obtained from the output of a convolutional encoder with a polynomial generator [5 7]. It can be observed that there is an undesired local minimum close to the point (0.0 0.0). Then the choice of such usual initialization results in ill convergence. This is illustrated in the trajectory of the DD-LMS algorithm from (0.0 0.0) to the local minimum in ≈ (-0.033,0.145) with a step-size of μ=1x10^{-4}.

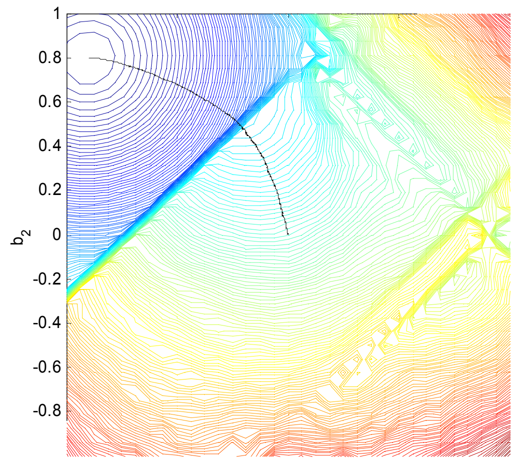

Figure 5: Contours of the performance surface for joint DFE & decoding. Convolutional code generated by polynomial [5 7], D=7, TD=1.

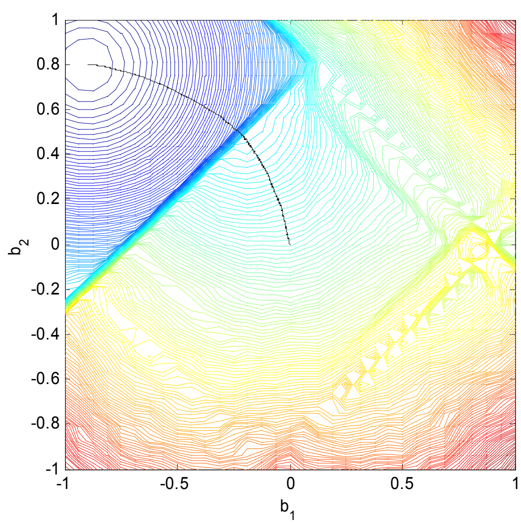

Figure 6: Contours of the performance surface for joint DFE & decoding. Convolutional code generated by polynomial [64 74], D=9, TD=1.

The results in figure 5 and 6 are obtained by using the joint DFE and decoding structure given in figure 3. The codes are generated by polynomials [5 7] and [64 74] respectively. As it can be observed, there is not any undesirable local minimum in the path from the zero initialization to the global minimum in (-0.9,0.8), so that a convergence to the optimum and desired solution is attained. The trajectories correspond to the DD-LMS with step-size μ=1x10^{-4}.

4.2 The effects of the convolutional decoder

The local minimum shown in figure 4 is a result of error propagation in the DFE. The use of codes makes it possible to reduce such error propagation and eliminate the corresponding local minimum. A complete theoretical description of such behavior is not trivial due to the difficulty in the code analysis. However, as a practical justification, it is interesting to point out the effects of the decoding process in a closed-eye condition. Hence some results are obtained and presented in table 1 in order to make clear the error correction capability even in such conditions.

Figure 7 depicts the structure used to obtain the results in table 1. The feedback filter weights were kept constant and equal to zero, i.e. there was no feedback filter. Table 1 shows the error rate at the output of the decoder for various tentative decision delays using the same channel as in subsection 4.1 . Clearly the error rate is 0.5 at the decision device output $\hat{a}(n)$.

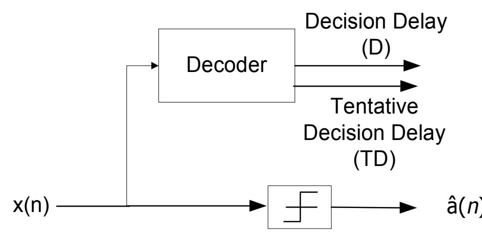

Figure 7: Structure used to evaluate the effect of decoding in closed-eye conditions

	Error Rate	
(Tentative) Decision Delay	**[5 7]**	**[64 74]**
0	**0.3448**	**0.3850**
1	**0.4302**	**0.3780**
2	**0.3508**	**0.4396**
3	**0.3475**	**0.4201**
4	**0.3599**	**0.4184**
5	**0.3373**	**0.4147**
6	**0.3340**	**0.4187**
7	**0.3458**	**0.4084**
8	**0.3427**	**0.4045**
9	**0.3418**	**0.4010**

Table 1: Error rate as a function of tentative decision delay

Firstly, we can see that the increase of the tentative decision delay does not necessarily bring lower error rates. This is because the Viterbi decoder was not designed for closed-eye condition and the survivor paths do not converge as the decision delay increases. Another possibility that accounts for such behavior is that the code was not designed for burst errors. Nevertheless, as the decision delay increases the oscillation of the error rate is smaller. Hence, by our simulations, a decision delay of three times the code constraint length should be enough to achieve a good level of error correction in this situation.

Secondly, it can be observed that the increase in the code constraint length did not yield lower error rates, as is expected in AWGN case. Again, this anomalous behavior is a result of the extreme situation generated by the closed-eye condition. Nonetheless, further studies must be done with other constraint lengths, aiming a more general conclusion about this behavior.

4.3 Convergence rate assessment

Another important aspect that must be taken into account is the value of the adaptation step size μ. It is known that the algorithm may escape from the local minimum by increasing the value of the step size. In order to compare the performance of conventional DFE and the joint DFE and decoding in terms of the adaptation step size, an example is shown in figure 8. The convolutional code was generated by polynomial [5 7] and the channel was $h(z)=1+0.4z^{-1}-0.2z^{-2}+0.8z^{-3}-0.7z^{-4}+0.1z^{-5}$. The same step-size was used for both structures in each case.

As it can be observed, the step size in figure 8a is small enough for conventional DFE to stay seized in the local minimum, while the joint structure converges toward the optimal solution. On the other hand, figure 8b shows a case were conventional DFE escapes from the local minimum. Nevertheless we note that the joint DFE and decoding structure converges significantly faster. Additionally, figure 8c shows the worst case of gradient noise due to the higher value of step size, which causes the code to lose performance. Thus, joint DFE and decoding converges slower than conventional DFE. Note, however, that the value used for the adaptation step size is very high and close to the limit for which convergence is guaranteed.

Others solutions have been proposed in the literature in order to search for the convergence to the desired global minimum in DFE. An interesting one is the use of soft decision like a saturation device [1,7]. We tested such approach with two types of bad channels and observed that it converges to the desired minimum for $h(z)=1+0.9z^{-1}-0.8z^{-2}$. However it does not converge for a more critical channel like $h(z)=1+0.95z^{-1}-0.9z^{-2}$. In this same case the joint DFE and decoding still converges to the desired solution as shown in figure 9, where 55000 iterations was executed using μ=1x10^{-4},as the step-size.

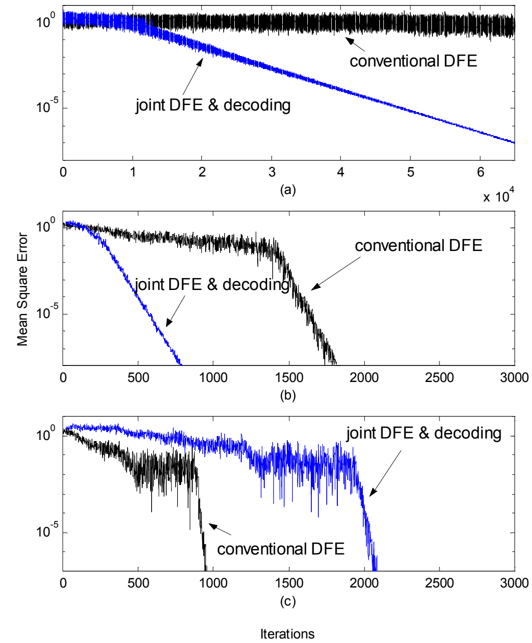

Figure 8: Dynamical comparison between conventional DD-DFE and the joint DFE & decoding structures: (a) μ=1x10^{-4}, (b) μ=0.01, (c) μ=0.05

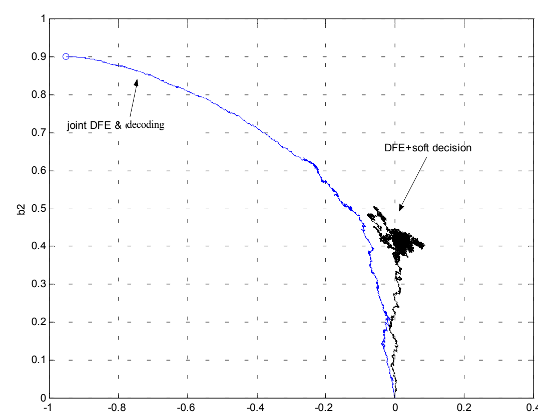

Figure 9: Weights trajectories for joint DFE & decoding vs. DFE + soft decision.

5. CONCLUSION & PERSPECTIVES

In this paper we have shown that it is possible to achieve convergence to the desirable global minimum with the use of the DD-LMS. In a linear equalizer, such convergence is strongly dependent on the filter weights initialization. Using the DFE structure together with DD-LMS, there is a class of channels that presents ill convergence when the feedback filter weights are initialized with zeros. We have shown that the use of error correcting codes renders possible the convergence to the desirable global minimum, even for channels belonging to this class.

There are others advantages in the use of joint equalization and decoding, besides the improvement in the convergence of blind equalizers. Reference [5] shows that the use of this technique gives an improvement of about 3dB and can reach more than 6dB if spatial and temporal diversity are applied, in comparison with conventional techniques, in a peer-to-peer situation.

It is also important to notice that this is a preliminary work in the subject. Further studies are being done to achieve a more general and reliable result. Analytical results are extremely difficult to obtain due to the technical difficulties imposed by the convolutional code. The study of other blind algorithms as CMA is also in course.

Acknowledgements

We would like to express our appreciation to the Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) for supporting our graduate education.

References

[1] R. A. Casas, C. R. Johnson, Jr., R. A. Kennedy, Z. Ding and R. Malamut, 'Blind Adaptive Decision Feedback Equalization: A Class of Channels Resulting in Ill-Convergence from a Zero Initialization', Int. J. Adapt. Control Signal Process., vol. 12, 173-193, 1998.

[2] O. Macchi and E. Eweda, 'Convergence Analysis of Self-Adaptive Equalizers', IEEE Trans. Inf. Theory, vol. 30, no. 2, 161-176, March 1984.

[3] J. E. Mazo, 'Analysis of Decision-Directed Equalizer Convergence', Bell System Technical Journal, vol. 59, no. 10, 1857-1876, December 1980.

[4] R. Kennedy, B. Anderson, R. Bitmead, 'Blind Adaptation of Decision Feedback Equalizers: Gross Convergence Properties', Int. J. Adapt. Control Signal Process., vol. 7, 497-523, 1993

[5] S. L. Ariyavisitakul e G. M. Durant; 'A Broadband Wireless Packet Technique Based on Coding, Diversity, and Equalization', IEEE Communications Magazine, 110-115, July 1998.

[6] R. A. Casas, 'Blind Adaptive Decision Feedback Equalization: A Class of Bad Channels', MS Thesis, Cornell University, 1996

[7] S. Marcos, S. Cherif, M. Jaidane, 'Blind Cancellation of Intersymbol Interference in Decision Feedback Equalizers', Proc. Int. Conf. on Acoustics, Speech and Signal Processing, 1073-1076, May 1995

[8] S. Lin e D. J. Costello Jr.; 'Error Control Coding: Fundamentals and Applications', Prentice-Hall, Englewood Cliffs, NJ, 1983.