On Complex Domain Decision Feedback Equalizer Based on Bell-Sejnowski Neuron

Vladimir R. Krstić
Communication department
Institute Mihailo Pupin
Belgrade, Serbia
vladak@kondor.imp.bg.ac.yu

Miroslav Dukić
Communication department
Faculty of Electrical Engineering, University of Belgrade
Belgrade, Serbia
dukic@etf.bg.ac.yu

***Abstract*—This paper considers the complex domain Bell-Sejnowski class neuron in the self-optimized decision feedback equalizer structure where an appropriate nonlinearity responds with the maximum entropy signal. The presented simulation results have proved the stability of the joint entropy maximization algorithm during the blind acquisition mode and also its ability to mitigate the error propagation effects that emerge when the cascade of linear components switches into the decision feedback mode.**

Keywords-Blind decision feedback equalization; Bell-Sejnowski class neuron model; joint entropy maximization algorithm.

I. INTRODUCTION

The ability of the Bell-Sejnowski neurons in the structural model to align to the statistical distributions of the input signals is one of the most interesting aspects of the independent component analysis theory [1]. This property transformed in different learning rules was successfully exploited in a number of signal processing applications such as blind separations, blind deconvolution and probability density function estimation [2]. In the field of blind equalization, which is of interest here, the Bell-Sejnowski maximum-entropy learning formula was applied to the decision feedback equalizer (DFE) and the class of joint entropy maximization (JEM) stochastic gradient algorithms was derived for the real-valued signals [3]. Also, more recently in [4] and [5], the maximum entropy learning approach was extended to the complex domain and combined with the self-optimized DFE scheme, which was originally proposed in [6]. This combination was resulting in a performance-efficient and low-complexity blind equalizer which can be applied to different systems in a severe intersymbol interference (ISI) environment.

In the self-optimized DFE scheme, the most critical component is its recursive part that changes position during the activation scenario: 1) in the blind acquisition mode the recursive part acts as an all-pole whitener (decorrelator) placed in front of a linear equalizer; and 2) after a well finished acquisition, it continues adaptation as the feedback filter of DFE (see Fig. 1). In this paper we consider the recursive part of DFE as a soft feedback (SFBF) filter of the Bell-Sejnowski class in the complex domain context. Referring to the complex stochastic gradient algorithm of the JEM type in [5], our main goal is to gain a deeper insight into the stability of the decorrelator in the blind acquisition mode and also to assess the performance of SFBF with respect to the error propagation phenomenon in symbol intervals after the structure switching.

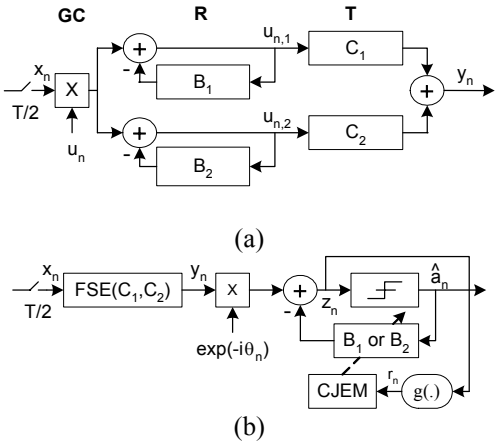

Figure 1. Soft-DFE structure in (a) blind acquisition mode and (b) soft transition modes: **GC**-gain control, **R**-decorrelator, **T**-transversal equalizer

II. SELF-OPTIMIZED DFE STRUCTURE

The activation of the proposed self-optimized DFE (Soft-DFE), the structure of which is depicted in Fig. 1, is characterized by three operation modes: blind acquisition, soft transition and tracking [4]. In the blind acquisition mode, the cascade **GC**+**R**+**T** effectively acts as a linear equalizer (LE). The equalizer **T** is a $T/2$ fractionally-spaced equalizer ($T/2$-FSE) controlled by constant modulus algorithm (CMA) [7], [8], where T is a symbol interval. Consequently, the decorrelator **R** consists of two all-pole whitening filters. Note one of them is coupled with the **GC**. In the soft transition mode, at the beginning of which the structure is switched, the joint equalization and carrier phase tracking take place as follows: 1) the linear part of the Soft-DFE is the decision-directed (DD) equalizer minimizing mean-square error (MMSE); and 2) the nonlinear part is the soft feedback filter maximizing joint entropy. At the moment of structure switching the recursive part of one of whiteners (based on some criterion) becomes the feedback filter of DFE and, hence, the **R** is suppressed. In the tracking mode, the Soft-DFE operates as a classical DD-MMSE DFE. It is worth noting that the Soft-DFE can continue to be in a soft transition mode (no

978-1-4244-1688-2/08/$25.00 ©2008 IEEE

tracking mode) in solutions where it is combined with some channel coding schemes [9], [10]. Finally, the switching control of different operation modes is provided by MSE performance monitoring device. This MSE estimator monitors the two thresholds: M_{TL-1} from blind acquisition to soft transition and M_{TL-2} from soft transition to tracking mode. The latter threshold is defined as a measure of successful activation of equalizer.

III. Complex Domain Soft Feedback Algorithms

A. DFE as the neuron of Bell-Sejnowski class

The SFBF filter in [3] was considered as a neural unit with the corresponding set of adjustable parameters, Fig 2. In this context, the input-output description of an activation function $g(\cdot)$, which is both a memoryless and a monotonically increasing nonlinearity, is given as follows: the neuron's net input z_n is the combination of the noiseless channel output x_n mixing present and past data symbols and the weighted sum of previous outputs r_{n-j}, $j=1,...,N$, where N is filter length. The objective of the activation function in the given structure model is to transform the input sequence z_n with an unknown probability density function to a maximum entropy sequence $r_n = g(z_n)$ using the Bell-Sejnowski class adapting formula. In other words, the coefficients $\{b_j\}$ are the subject of learning through an adaptation algorithm that maximizes the joint entropy defined by $H\left[r_{n,1},...,r_{n,N+1}\right] \triangleq -E\{\ln f_r(\mathbf{r}_n)\}$ where $f_r(\mathbf{r}_n)$ is the joint density function of the input vector $\mathbf{r}_n = \left[r_{n,1},...,r_{n,N+1}\right]^T$.

Based on assumptions commonly used for DFE analysis, that is, the input data a_n are independently identically distributed and zero-mean and the previous decisions are correct, $r_{n-j} = a_{n-j}$, the JEM criterion was suitably simplified as $\tilde{J}_{EM} = E\{\ln|J|\} = E\left\{\ln\left|\frac{\partial r_n}{\partial z_n}\right|\right\}$ where E is expectation and $|J| = \left|\frac{\partial r_n}{\partial z_n}\right|$ is the absolute value of the Jacobian of the transformation. Consequently, in the case of stochastic gradient ascent learning the corresponding instantaneous approximation of $\tilde{J}_{EM}$ can be used as the adaptation criterion, that is,

$$J_{EM} = \ln\left|\frac{\partial r_n}{\partial z_n}\right|. \quad (1)$$

The above cost function expresses the main idea underlying the use of information-theoretic concept in channel deconvolution applications: the statistical dependence between the current output of the SFBF and its previous outputs can be reduced by maximizing the joint entropy in an iterative manner and, hence, that leads to reduction of ISI.

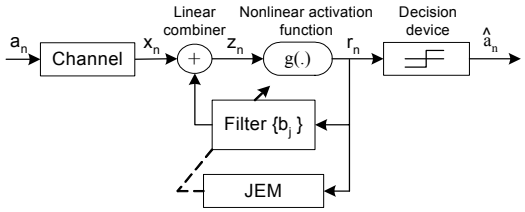

Figure 2. Model of transmission system with the basic scheme of soft feedback filter

B. Complex domain gradient of cost function J_{EM}

Let the neuron's complex-valued activation function be $g(z) = r_R(z_R, z_I) + ir_I(z_R, z_I)$ where the subscripts R and I indicate the real and imaginary parts of g, respectively, and $z = z_R + iz_I$, $i = \sqrt{-1}$; for convenience of notation, the symbol interval index n has been dropped. The function $g(z)$ is assumed to be an appropriate nonlinearity whose properties will be discussed later. In the complex-valued case of interest here, the input-output relationship of the SFBF can be defined by the net

$$z = z_R + iz_I = (x_R + ix_I) - \sum_{j=1}^{N}\left(b_{j,R} + ib_{j,I}\right)\left(r_{j,R} + ir_{j,I}\right). \quad (2)$$

It should be noted that for the given complex net the entropy J_{EM} is a real scalar function of unknown complex coefficients. In fact, this is an essential property of the function J_{EM} that simplifies deriving of its gradient with respect to complex coefficients. This simplification is done through use of a multidimensional complex gradient vector [11, Ch. 2], whose *j*th element is defined as follows

$$\nabla_{b_j} J_{EM} = \frac{\partial J_{EM}}{\partial b_{R,j}} + i\frac{\partial J_{EM}}{\partial b_{I,j}},\ j = 1,...N. \quad (3)$$

In our particular case of the SFBF's net, the gradient of J_{EM} with respect to coefficients $\{b_j\}$, which is derived in [5], is given by

$$\nabla_{b_j} J_{EM} = -\left[\left(r_R^{rr} + ir_R^{ri}\right)\delta_R + \left(r_I^{ir} + ir_I^{ii}\right)\delta_I\right]r_{n-j}^*, \quad (4)$$

where * denotes complex conjugation. The respective quantities in (4) are the corresponding first and second order partial derivatives of the activation function which are defined as follows: $\left(\frac{\partial r_R}{\partial z_R}\right)^{-1} = \delta_R$, $\left(\frac{\partial r_I}{\partial z_I}\right)^{-1} = \delta_I$, $\frac{\partial^2 r_R}{\partial z_R^2} = r_R^{rr}$,

$\frac{\partial^2 r_R}{\partial z_R \partial z_I} = r_R^{ri}$, $\frac{\partial^2 r_I}{\partial z_I \partial z_R} = r_I^{ir}$ and $\frac{\partial^2 r_I}{\partial z_I^2} = r_I^{ii}$. It is worth noting in (4) that the entropy gradient does not depend explicitly on an activation function, but it is governed by the quantities that are given as proportions of the corresponding second and first order partial derivatives of a complex activation function.

Let us discuss in more detail the properties of the complex function $g(z)$. This function need not necessarily be analytic [11, Ch. 17], but it is the suitable activation function in the sense of properties that provide a stable gradient learning [5] and can be summarized as follows:

P1. $g(z)$ is nonlinear in z_R and z_I.

P2. For all z in a bounded domain D, a suitable complex activation function $g(z)$ must have no singularities (especially no poles) and it must be bounded. This property is, in fact, the bounded-input bounded-output (BIBO) condition for a complex activation function. In other words, the system using the activation function $g(z)$ must be stable in the BIBO sense.

P3. The second-order partial derivatives r_R^{rr}, r_R^{ri}, r_I^{ii}, r_I^{ir} exist for all $z \in \mathbb{C}$, and the corresponding "normalized" derivatives $\delta_R r_R^{rr}$, $\delta_R r_R^{ri}$, $\delta_I r_I^{ii}$ and $\delta_I r_I^{ir}$ must be bounded since coefficients updating is in quantities proportional to the normalized second-order partial derivatives (see (4)). Obviously, the derivatives δ_R and δ_I must be different from zero to avoid an instability.

P4. The partial derivatives of $g(z)$ obey $r_R^{rr} r_I^{ii} \neq r_I^{ir} r_R^{ri}$. If this condition is not satisfied no learning state is possible for both nonzero input $z_n = (z_R, z_I)$ and $\delta_n = (\delta_R, \delta_I)$.

C. Stochastic gradient algorithms of JEM type

Let us consider a simple complex activation function which is described by the input-output equation

$$g(z) = r(z) = z(1 + \beta |z|^2) \tag{5}$$

where β is a real positive constant. This function has a property of mapping a point $z = z_R + iz_I = (z_R, z_I)$ on the complex plane to a unique point $g(z) = \left(z_R\left(1+\beta|z|^2\right)\right), \left(z_I\left(1+\beta|z|^2\right)\right)$ keeping the same phase angle. In addition, the magnitude of the complex function $g(z)$ is a paraboloid that has the following properties: the horizontal cross sections are actually circles, the bottom surface is located at the point $(0,0)$ and the parameter β modifies the shape of the surface. In fact, the parameter β affects the input-output mapping modifying in the same manner both the real and the imaginary parts of activation function. Finally, it can be easily proved that this function is not analytic, but it satisfies the properties **P1-P4**.

For the complex function defined by (5) the stochastic gradient algorithm of the JEM type (CJEM) is given by

$$b_{j,n+1} = b_{j,n} + \alpha \nabla_{b_{n,j}} J_{EM},$$

$$b_{j,n+1} = b_{j,n} - \mu z_n \left(1 - \beta |z_n|^2\right) r_{n-j}^*, \; j = 1, ..., N, \tag{6}$$

where α is a positive learning constant and $\mu = 8\alpha\beta$ is a step size. In Section V., the parameter β will be used as a tool to optimize the convergence characteristics of the CJEM algorithm. We have denoted β as a smoothing parameter taking into account its overall effects on the algorithm behavior in the Soft-DFE.

D. Modified CJEM algorithms

The CJEM algorithm is a key component of the Soft-DFE system. As it is mentioned in Section II, in this self-optimized scheme the filters B_1 and B_2 start their adaptation as all-pole whiteners in the blind acquisition mode and then, one of them, continues adaptation as a part of the SFBF. In this scenario, the two variants of the CJEM algorithm are of interest. In the first case, by the analogy with the Extended LMS algorithm (ELMS) [12], the algorithm (6) can be transformed into the corresponding decorrelation algorithm (EJEM) as follows

$$b_{j,n+1} = b_{j,n} - \mu_E u_n \left(1 - \beta_E |u_n|^2\right) u_{n-j}^*, \; j = 1, ..., N \tag{7}$$

where μ_E is a step size and β_E is a parameter that can be selected to provide the fast and stable convergence of the whiteners.

The second variant of the CJEM algorithm is the result of the structure modification in the basic scheme of the SFBF where the hard decision estimates of the transmitted symbols, $\hat{a}_{n-j}$, feed the feedback filter instead of the soft decisions r_{n-j} (see Fig. 1b). Consequently, the decision-directed variant of the CJEM algorithm (DD-CJEM) reads

$$b_{j,n+1} = b_{j,n} - \mu z_n \left(1 - \beta |z_n|^2\right) \hat{a}_{n-j}^*, \; j = 1, ..., N. \tag{8}$$

The applied modification is a heuristic approach that is motivated by the following reasons: 1) the decorrelator **R** is possibly correctly set up during the blind acquisition mode because it is an amplitude equalizer performing a speed convergence even if a channel is in a deep fade; 2) the equalizer **T** of sufficient length is globally convergent [8]; and 3) the DD-CJEM algorithm combines the soft error $z_n\left(1-\beta|z_n|^2\right)$ of the CJEM algorithm and the conventional decision-directed method.

IV. STABILITY OF DECORRELATOR

The study of stability of the decorrelation algorithm (7) is based on experimental results obtained by the software simulator. We have considered the BIBO stability of whiteners as a kind of less strict stability condition [12]. As it is known, the BIBO stability is a global property over all the time instants n which does not preclude a local instability when, at certain time instants n, the predictor exhibits poles outside the unit circle. When this happens, the innovation u_n will generally present a burst of samples with large amplitudes. A kind of this sort of instability is the self-stabilization (SS) phenomenon [12] encountered in the adaptive predictor with the ELMS algorithm. Hence, that is the motivation to examine the SS and BIBO stability of the EJEM algorithm taking into account the decorrelator **R** has a short-time task in blind acquisition mode after which it is suppressed. The results of two stability tests are presented as follows: 1) the decorrelator stability (without **GC**) in the case of a sinusoidal disturbance; and 2) the stability of the LE in the case of 4QAM signal. For simplicity reason, only one whitener is observed in the following tests.

A. Test 1

The SS phenomenon is checked for both ELMS and EJEM algorithms employing narrowband disturbance. The auto regressive model of a pure sinusoidal signal with a random phase exhibits two poles on the unit circle. Adding a weak noise these poles can be little pushed inside the unit circle but they are still a maximal risk of instability for the adaptive predictor. Thus, a sinusoidal signal is a simplest example of the input disturbing signal

$$s_n = s(nT) = A\cos(2\pi fnT + \xi) + \eta(nT) \tag{9}$$

where A is amplitude, f is frequency, ξ is a random phase with the flat distribution over $[\pi, -\pi]$ and $\eta(nT)$ is zero-mean noise. The parameters of the signal (9) are selected for test 1 as follows: amplitude $A = 1$, frequency $f = 1800Hz$, and zero-mean noise of variance $\delta_\eta^2 = 2.5 \cdot 10^{-4}$. The initial values of the sixth-order ($N = 6$) whitener are zero. The magnitudes of both the most critical coefficient $b_{1,n}$ and the innovation u_n have been observed through the test and presented in Fig. 3. It should be noted that in the ELMS case the bursting drift pushes $b_{1,n}$ towards the unstable region $|b_{1,n}| > 1$, and then the SS effect appears, which pushes $b_{1,n}$ back to the stable region $|b_{1,n}| < 1$. This phenomenon is followed by a kind of BIBO stability that is an intermediate situation between desirable (permanent) stability and evident (hopeless) instability. To the contrary, in the case of EJEM algorithm, the whitener shows the instability in this limiting case test; the unbounded output (overflow) has happened at $t = 10835T$ for $\beta_E = 5$. This kind of instability is also proved for the smaller values of β_E.

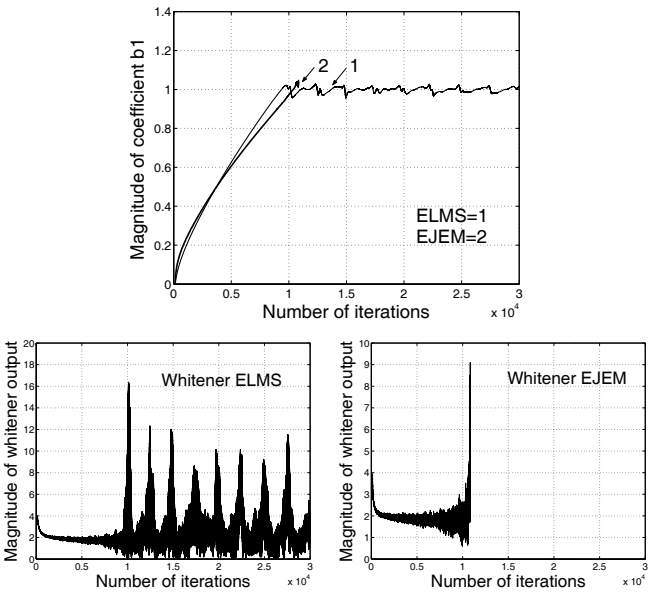

Figure 3. Sinusoidal disturbance self-stabilization test of whitener with algorithms ELMS and EJEM for $\beta_E = 5$

B. Test 2

In this test the 4QAM signal and severe multipath channel MP-E are applied (see Section V). The LE is in the blind mode of operation over the whole test of duration $60000T$. The results of this test, the evolution of the magnitude of the first coefficient $b_{1,n}$ and the MSE at the output of LE are presented in Fig. 4. Note the magnitude $|b_{1,n}|$ shows larger fluctuations with JEM-W algorithm than in the case of ELMS. However, this does not produce any instability: the whitener is BIBO stable and, as a consequence, the LE shows the stable MSE convergence. To provide a better insight into the stability of the whitener the locations of poles of one whitener are presented in Fig. 5; these poles are inside the unit circle and their positions are very similar for both algorithms at the end of the test of duration $60000T$ intervals. The corresponding coefficients of the sixth-order whitener are given in Table I.

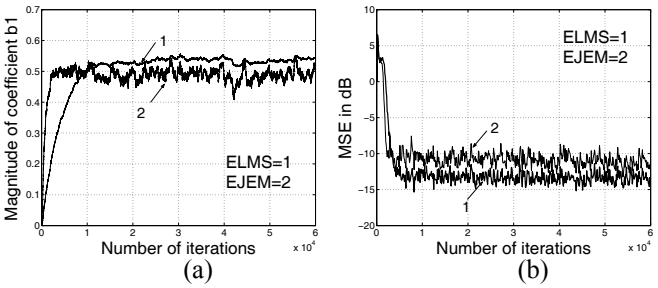

Figure 4. The comparison of ELMS and EJEM decorrelators in the blind mode in the case of 4QAM noiseless system: (a) Evolution of the magnitude of the first coefficient of **R** and (b) MSE convergence at the output of LE

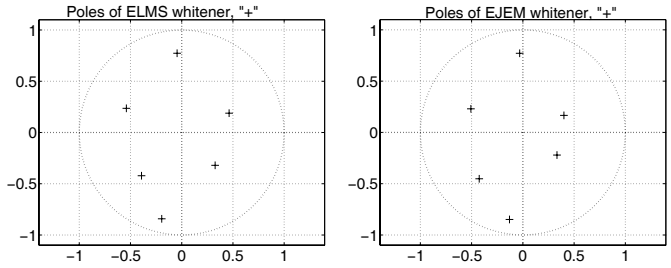

Figure 5. Poles of whitener with algorithms ELMS and EJEM for $\beta_E = 5$

TABLE I. THE COEFFICIENTS OF THE WHITENER AT THE END OF THE TEST

	Algorithm EJEM, $\beta_E = 5$		Algorithm ELMS	
b_k	$\mathrm{Re}(b_k)$	$\mathrm{Im}(b_k)$	$\mathrm{Re}(b_k)$	$\mathrm{Im}(b_k)$
1	0.359969	0.353165	0.383821	0.385971
2	0.463459	-0.028449	0.435103	-0.002116
3	0.043394	0.077096	0.021841	0.075368
4	-0.084899	0.000490	-0.080875	0.015623
5	-0.046674	-0.035240	-0.026929	-0.059371
6	0.039535	0.003342	0.051308	-0.007748

V. SIMULATION RESULTS

The Soft-DFE achievements presented in this paper are characterized in terms of symbol error rate (SER) and block error length (BEL) statistics which are observed in the soft transition mode. Also, the convergence characteristics of the Soft-DFE are compared with one of the Hard-DFE where the same structure is used but with the Extended LMS and conventional DD LMS algorithms in the blind and tracking modes, respectively, as in [6]. Both equalizers have the same lengths of $22T$ and $6T$ in their linear and recursive parts, respectively. Also, the same two-spike initialization of the FSE-CMA and step size are employed in both cases. The presented results are obtained using the 16QAM signal and multipath channels (MP). These channels are realized by the three-ray channel model [8] included in the transmitter filter. The amplitude characteristics of the channels for different propagation parameters are presented in Fig. 6.

Let us first describe the transition mode of the Soft-DFE by using its convergence characteristics obtained with the Mp channels and the signal-to-noise ratio $SNR = 25.0dB$, Fig. 7. The threshold levels $M_{TL-1} = 1.5dB$ and $M_{TL-2} = -8.0dB$ define the beginning and the end of the transition mode, respectively. The threshold level M_{TL-1} is selected to be sufficiently high so that, at the moment of configuration switching, the initial burst of errors is certainly present. In fact, this is the most critical phase of the Soft-DFE operation when several algorithms are directed and slightly coupled by the sequence of unreliable symbol estimates: 1) the DD-LMS algorithm that adjusts $T/2$-FSE; 2) phase-tracking loop that rotates a signal constellation; and 3) the DD-CJEM algorithm.

Fig. 8 presents the results of the SER versus the smoothing parameter β which are averaged over 200 independent runs. The SER is measured during 1500 symbol intervals period that, in the case of severe-ISI channels, Mp=(C,D,E,F,G,H), roughly corresponds to equalizer's convergence time between the

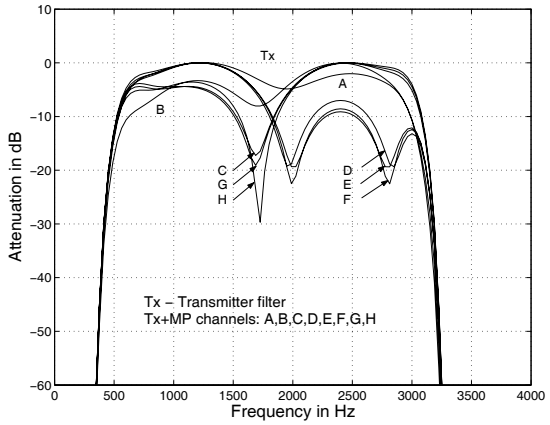

Figure 6. Amplitude response of MP channels (A,B,C,D,E,F,G,H)

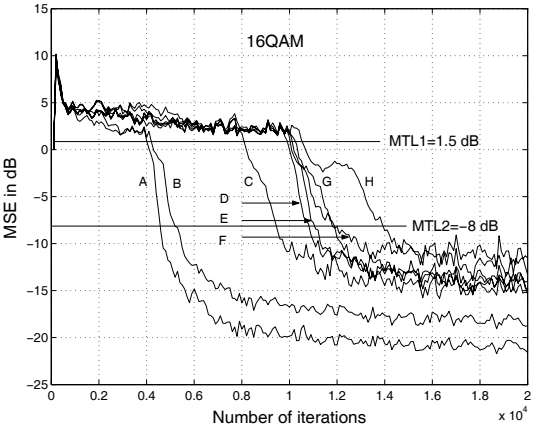

Figure 7. MSE convergence characteristics of the Soft-DFE with MP channels for $\beta_E = 1$ and $\beta = 10$

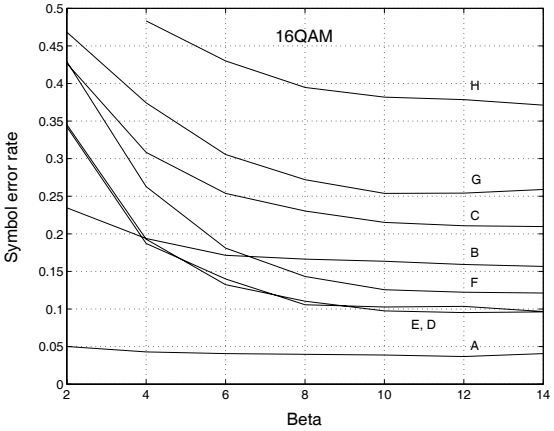

Figure 8. Symbol error rate versus β in transition mode: Mp=(A,B,C,D,E,F,G,H), 16QAM, SNR=25 dB

M_{TL-1} and M_{TL-2} thresholds. Obviously, for these channels, the SER decreases with increasing β in a similar manner. This behavior clearly indicates the capability of DD-CJEM to mitigate error propagation to some extent. These results are followed by the block error length statistics given in Table II

for the lengths from 2 to 8 in T intervals; the BEL=8 also includes longer burst errors. The numbers of failed activations (FAs) of equalizers during the Monte Carlo test are given in the last column of the table. It should be noted that the long block errors have disappeared and the number of short blocks is reduced for β in range $\{10,14\}$. Also, the SER and BEL statistics, which are introduced as a measure of the error propagation phenomenon, are not critically sensitive for a wide range of values of the smoothing parameter. On the other hand, the Hard-DFE yields similar SER results, but it suffers from the increased number of long burst errors the consequence of which is the activation fail of the equalizer.

The presented SER and BEL statistics have proved the ability of the new algorithms to mitigate the error propagation effects. Also, these results indicate that improved burst error statistics (no long burst errors) can be transformed into a larger channel coding gain in systems with separate equalization and decoding scheme.

TABLE II. BLOCK ERROR LENGTH STATISTICS FOR CHANNEL MP-E

BEL/ β	2	3	4	5	6	7	8	FA
2	70	29	13	7	3	2	2	40
4	39	15	6	3	1	0	1	4
6	26	10	4	2	1	0	0	0
8	22	8	3	1	1	0	0	1
10	20	6	2	1	0	0	0	0
12	19	6	2	1	0	0	0	1
14	20	5	2	1	0	0	0	1
Hard-DFE	22	2	0	0	0	0	0	4

In Fig. 9 the overall convergence characteristics of Soft-DFE are presented for different values of β. For the purpose of comparison, the convergence characteristics of the Hard-DFE and "trained" Soft-DFE solutions are also presented. The "trained" denotes the Soft-DFE which directly switches from the blind mode to the actually trained adaptation based on the LMS algorithm and desired symbols a_n. As we can see the

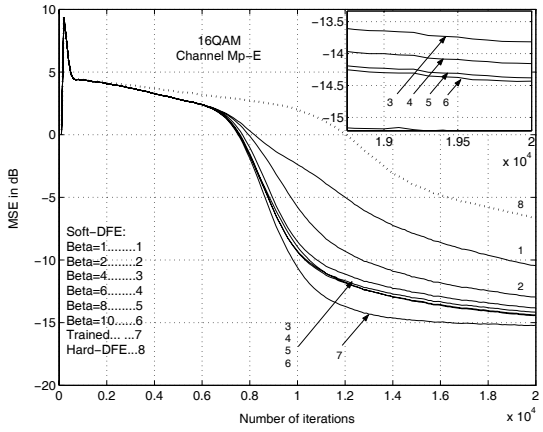

Figure 9. MSE convergence: Soft-DFE with DD-CJEM algorithm in transition mode for different values of β is compared with the "trained" Soft-DFE and the Hard-DFE

Soft-DE strikes the best MSE convergence characteristics with increasing β where they show a kind of saturation for $\beta > 8$. On the other hand, for smaller values of β, the performance of DD-CJEM algorithm deteriorates so that both the speed of convergence and MMSE become unacceptable. In the special case of $\beta = 1$, the performances of the DD-CJEM are considerably degraded. Also, it should be noted that the Soft-DFE is not inferior in comparison with the "trained" Soft-DFE.

VI. CONCLUSIONS

The purpose of this contribution is to discuss in some detail the behavior of the self-optimized DFE based on the complex-valued neuron of the Bell-Sejnowski class in order to find a relationship between the smoothing parameter of the chosen activation function and the error propagation effects and, hence, the convergence characteristics. The presented simulation results have proved the ability of the new algorithm to mitigate the error propagation effects for a wide range of values of the smoothing parameter. In particular, it is possible to determine the range of values of β that provides the best performance of the algorithm for a known signal constellation. Consequently, for the selected β, the proposed DD-CJEM algorithm is of the same low computation complexity as the CMA.

REFERENCES

[1] A. J. Bell, T. J. Sejnowski, "An information maximization approach to blind separation and blind deconvolution," *Neural Computation,* vol 7, 1996, pp. 1129-1159.

[2] S. Fiori, "Notes on Bell-Sejnowski PDF-Matching Neuron," Neural Computation Vol.14, No.12, Dec. 2002, pp. 2847-2855.

[3] Y. H. Kim, S. Shamsunder, "Adaptive Algorithms for Channel Equalization with Soft Decision Feedback," IEEE JSAC. vol. 16, Dec. 1998, pp. 1660-1669.

[4] V. R. Krstic, Z. Petrovic, "Decision Feedback Blind Equalizer with Maximum Entropy," EUROCON 2005, Belgrade, November, 2005.

[5] V. R. Krstic, Z. Petrovic, "Complex-Valued Maximum Joint Entropy Algorithm for Blind Decision Feedback Equalizer, "8th International Conference on Telecommunications in Modern Satellite, Cable and Broadcasting Services-TELSIKS 2007, Serbia, Niš, September 26-28, 2007, pp. 601-604.

[6] J. Labat, O. Macchi, and C. Laot, "Adaptive Decision Feedback Equalization: Can You Skip the Training Period?," IEEE Trans. Commun.., July 1998, pp.921-930.

[7] D. Godard "Self-recovering Equalization and Carrier Tracking in Two-Dimensional Data Communication Systems," IEEE Trans. Commun.., Nov. 1980, pp. 1867-1875.

[8] Y. Li and Z. Ding, "Global Convergence of Fractionally Spaced Godard (CMA) Adaptive Equalizers," IEEE Trans. Signal Processing, Apr. 1996, pp. 818-826.

[9] J.-T. Liu, S. B. Gelfand, "Optimized Decision-Feedback Equalization for Convolutional Coding With Reduced Delay," IEEE Trans. Commun., Nov., 2005, pp.1859-1866.

[10] G. Ananthaswamy, D. L. Goeckel, "A Fast-Acquiring Blind Predictive DFE," IEEE Trnas. Commun., Oct. 2002, pp.1557-1560.

[11] S. Hykin, Adaptive Filter Theory, 4th edn Prentice-Hall, 2002.

[12] O. Macchi, Adaptive processing, The LMS approach with Applications in Transmission, New York: John Wiley & Sons, 1995.