A Stochastic Whitening Algorithm with Time Variable Neuron Slope for Blind DFE

Vladimir R. Krstić, *Member, IEEE,* Bojan R. Dimitrijević, Miloš V. Banđur

***Abstract* — The self-optimized blind decision feedback equalizer (DFE), which independently optimizes feed-forward and feedback filters, deals with the feedback filter mismatch. To mitigate the filter mismatch effects, we have proposed a stochastic entropy gradient algorithm with a time variable neuron slope for all-pole filter which acts as a front-end amplitude equalizer in the blind operation mode. The algorithm performances are verified by simulations for 16- and 64-QAM (quadrature amplitude modulated) signals.**

***Keywords* — Blind decision feedback equalization, joint entropy maximization cost, adaptive neuron slope.**

I. Introduction

A channel equalizer is an adaptive filter which purpose is to compensate for the intersymbol interference (ISI) at the channel output. With respect to the equalizer parameters training method, the initial equalizer adaptation can be achieved with or without a training sequence (preamble). Making the system free from sending a training sequence, a blind equalizer increases effective system throughput and facilitates its operation in systems where it is not practical or possible to send a preamble [1].

Recently, the blind equalization based on the decision feedback filter structure has attracted attention of the researchers since the decision feedback equalizer (DFE) over-performs the corresponding linear equalizer (LE) for both channels with deep spectral nulls [2] and channels facing long delay spread [3]. The DFE divides the equalization task between two filters (equalizers), the linear feed-forward finite impulse response equalizer (FFF) and the nonlinear feedback infinite impulse response equalizer (FBF), where the latter eliminates the post-cursor ISI with a smaller number of coefficients and without noise enhancement compared to the LE. On the other hand, the main drawback of DFE is the error propagation phenomena caused by data symbols erroneously detected by the FBF [2].

To evade the error propagation effects, Labat et al. [4] designed a self-optimized DFE (SO-DFE) scheme for QAM (quadrature amplitude modulated) signals which optimize both the structure and the adaptation criterion according to the equalizer convergence state. This scheme is based on the theory that the infinitely long minimum mean square error (MMSE) LE and MMSE DFE equalizers share the same components which, in the frequency domain, can be factored into the all-pole recursive amplitude equalizer and the phase equalizer which, respectively, compensate for minimum and maximum phase components of a channel transfer function. Although very attractive, this scheme presents weakness in practical implementations based on the finite length filters. Namely, the SO-DFE deals with the FBF mismatch resulting from the change of its position during the equalizer convergence process. In the blind mode the FBF acts as a front-end all-pole amplitude equalizer and, after the equalizer structure-criterion switching, as a decision directed feedback equalizer which is placed after the FFF. Strictly, the FBF equalizer is not a minimum phase filter.

Aiming to mitigate the FBF mismatch effects, the soft FBF (*SFBF*) filter has been introduced in the SO-DFE scheme in [5]. Based on the one-neuron-unit deconvolution theory [6], the innovated SO-DFE scheme, called Soft-DFE, significantly over-performs the original solution. These results have initiated further improvement of the *SFBF* filter efficiency during the blind operation mode. In this paper we propose the stochastic gradient algorithm with the time varying neuron slope for the all-pole amplitude equalizer in the Soft-DFE scheme.

The rest of the paper is organized as follows: Section II describes Soft-DFE equalizer operation modes and *SFBF* adaptation algorithms. The proposed algorithm is described in Section III, and the simulation results presenting the influence on the equalizer performance are given in Section IV.

II. Soft-DFE Scheme

A. Structure-criterion transformation

The Soft-DFE equalizer is presented in Fig. 1. The received signal $x(t)$ is sampled at a rate that is twice bigger than the symbol rate 1/T, and then odd and even samples $x(t_0+nT-iT/2)=x_{n,i}$, $i=1,2$, are alternatively shifted to the delay lines of the corresponding filters. The FFF and FBF parts of the equalizer include four T-spaced finite impulse response filters which are defined, respectively, with coefficient vectors $\mathbf{b}_i=[b_{i,1},...,b_{i,N}]^T$ and $\mathbf{c}_i=[c_{i,1},...,c_{i,L}]^T$.

This work was supported by the Ministry of Science and Technological Development of the Republic of Serbia: the project TR 32037, 2011-2015.

Vladimir R. Krstić, Institute „Mihajlo Pupin“, Volgina 15, Beograd, Serbia (tel.: 381-64-1581684, e-mail: vladimir.krstic@pupin.rs)

Bojan Dimitrijević, Faculty of Electronic Engineering,University of Niš, Serbia (e-mail: bojan.dimitrijevic@elfak.ni.ac.rs)

Miloš V. Banđur, Faculty of Technical Sciences, University of Pristina, Kosovska Mitrovica, Serbia (e-mail: milos.bandjur@pr.ac.rs).

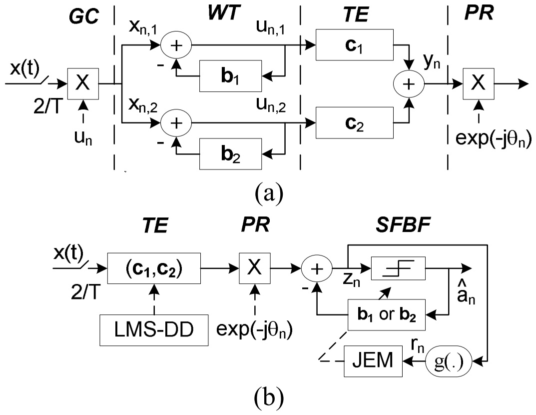


Fig. 1. Soft-DFE equalizer: (a) blind acquisition and (b) soft transition mode.

The Soft-DFE operates through three operation modes: blind acquisition, soft transition and tracking. In the blind mode the equalizer effectively acts as a linear T/2 fractionally spaced equalizer (T/2-FSE) which includes four signal transformers ordered in cascade with increasing complexity: gain control (*GC*), whitener (*WT*), equalizer (*TE*) and phase rotator (*PR*), Fig. 1a. Transformers *GC* and *WT* are coupled in a pair where *GC* recovers the transmitted signal energy and the all-pole whitener *WT* performs channel spectrum equalization. The *WT* is adapted by the stochastic gradient algorithm based on the joint entropy maximization (JEM) cost [7]. The linear equalizer *TE*, which operates independently of the *GC*+*WT*, compensates for a phase distortion (introduced by a channel+whitener combination) by using the Constant Modulus Algorithm (CMA-2) [8]. In the next operation mode, named the soft transition mode, one out of two whiteners, selected according to energy criterion, transforms itself into the *SFBF* equalizer keeping on JEM adaptation, while the equalizer *TE* switches adaptation from the CMA to the decision-directed LMS (DD-LMS), Fig 1b. Effectively, during the soft transition mode, the Soft-DFE is optimized by the combined (MSE+JEM) criterion. Finally, the *SFBF* switches itself into the classical decision feedback equalizer performing DD-LMS adaptation (tracking mode). The phase rotator *PR* is a phase locked-loop which purpose is to estimate the carrier phase and to place the signal constellation in the nominal position [1], [9].

The process of Soft-DFE adaptation is controlled by the MSE monitor that switches both the structure and the criterion for *a priori* selected MSE thresholds (TRs): for TL1 from the blind to the soft transition and for TL2 from the soft transition to the tracking mode. Besides, the threshold TL3 is introduced as a measure of equalization successfulness: for MSE $\le$ TL3 the equalization is decided to be successful and for MSE $>$ TL3 unsuccessful.

B. SFBF equalizer: backgrounds and algorithms

The *SFBF* equalizer acts as a neuron-unit maximizing the joint Shannon's entropy (JEM criterion) [7]

$$J_H(\mathbf{b}_n) = E\left\{\ln\left|\frac{\partial g(z_n)}{\partial z_n}\right|\right\} \quad (1)$$

assuming the following system model:

- Data symbols $\{a_n\}$ applied to a noiseless linear time-invariant channel are independent identically distributed zero-mean variables with finite variance and sub-Gaussian distribution,
- The neuron-unit $g(\cdot)$ is a nonlinear monotone saturating function which transforms the unknown probability density function (PDF) of the input data into the uniform PDF in the range of the output data symbols $r_n = g(z_n)$.
- The previously detected symbols at the output of the *SFBF* are correct, i.e. $r_{n-j} = a_{n-j}$, $j = 1,...,N$.

The stochastic gradient algorithms optimizing the JEM criterion are derived for the complex-valued parametric function [5]

$$g(z_n, \beta) = z_n\left(1+\beta|z_n|^2\right) \quad (2)$$

where the parameter β is used as a tool to match the neuron slope to the PDF of ISI, and they are given by two recursions

$$b_{n+1,i,j} = (1-\gamma)b_{n,i,j} - \mu_W u_{n,i}\left(1-\beta_W\left|u_{n,i}\right|^2\right)u^*_{n-j,i} \quad (3)$$

$$b_{n+1,j} = b_{n,j} - \mu_D z_n\left(1-\beta_D\left|z_n\right|^2\right)\hat{a}^*_{n-j} \quad (4)$$

where (μ_W, μ_D) and (β_W, β_D), respectively, represent the corresponding adaptation steps and neuron slopes, and γ is the leaky factor. The JEM-W algorithm in (3) performs all-pole whitening of the received signal in the blind mode and the JEM-D in (4) continues entropy maximization in the soft transition mode assuming the previous decisions are correct. The leaky factor in (3) is an option which can be zero or small positive number, and it prevents the overgrowth of the whitener's vector (Euclidian) norm $\|\mathbf{b}_n\|$ in the case of higher-order QAM signals [9]. As well known, the coefficient leaky is commonly used technique in linear regression models with different motivations: to regularize their transient behaviour, to improve stability in a finite precision implementation and to reduce different undesirable effects [10] (and references therein).

One should note that the JEM-W and JEM-D algorithms share the same JEM-type nonlinear gradient term which performing depends on input-output signals and slopes (β_W, β_D).

The slope selection is an essential issue of the JEM algorithms implementation. The JEM-D algorithm is driven by detected symbols $\hat{a}_n$ whose PDF is determined by the given signal constellation. This situation simplifies the selection of the slope β_D. The previous results in [5] and [9] proved that the optimal value of the slope β_D depends only on the given signal constellation. In other words, the optimal slope β_D can be seen as a statistical constant of the given signal. On the other hand, the selection of the slope β_W is more complex problem because it optimizes the JEM-W algorithm directed by the decorrelated outputs of the WT which are determined by the second order statistics of the given signal. In fact, the slope β_W represents only up to second order statistics of the signal. Because of that, to minimize the mismatch between two JEM algorithms and to achieve a better whitening of the received signals with time varying statistics, we have introduced the time varying slope β_W into the JEM-W algorithm.

III. JEM-W Algorithm With Adaptive Slope

Let us consider the whitening algorithm with time variable slope JEM-VS

$$b_{n+1,j} = (1-\gamma)b_{n,j} - \mu_W u_n \left(1-\beta_n |u_n|^2\right) u^*_{n-j}\ ,\ \ j=1,...,N \tag{5}$$

where the term

$$e_n = u_n\left(1-\beta_n |u_n|^2\right) \tag{6}$$

can be seen as an *a priori* prediction error and the whitener current output is given by

$$u_n = x_n - \mathbf{b}_n^T \mathbf{u}_n\ ,\ \ \mathbf{u}_n = [u_{n-1},...,u_{n-N}]^T \tag{7}$$

The subscript W in β_n (5) is dropped for simplicity.

The JEM-VS algorithm besides the adaptive slope β_n also includes the leaky term γb_n which has two purposes: First, as mentioned before, the leaky term acts in opposition to the entropic term $\beta_n|u_n|^2$ to prevent unconstrained growth of whitener coefficients. In fact, it enables the use of larger values of the slope which improves the equalizer convergence speed without the risk of instability. Second, we have used its ability to regularize the transient behaviour of the JEM algorithm as a measure in the slope adaptation rule to decide whether to increase or decrease the slope β_n.

The adaptation of the slope β_n is *a posteriori* based one. It means that the whitener, at the time $t=nT$ but before the next updated input x_{n+1}, calculates the coefficient vector $\mathbf{b}_{n+1}^{VS}$ with ($\gamma>0$) and $\mathbf{b}_{n+1}$ without ($\gamma=0$) leaky and the corresponding *a posteriori* outputs ($\tilde{u}_n^{VS}, \tilde{u}_n$) and errors ($\tilde{e}_n^{VS}$, $\tilde{e}_n$). Further, based on the difference between modules of *a posteriori* errors the whitener decides whether to increase or decrease the slope β_n according to the rule: if $|\tilde{e}_n^{VS}| \le |\tilde{e}_n|$ the slope decreases and if $|\tilde{e}_n^{VS}| > |\tilde{e}_n|$ the slope increases.

The adaptation of the slope is implemented in a similar manner as it was done in [10] by the following if-else relation

$$\begin{aligned}
&\text{if } |\tilde{e}_n^{VS}| > |\tilde{e}_n| \text{ then}\\
&\text{set } m_{n+1} = \max(m_n - l_d, 0)\\
&\text{else}\\
&\text{set } m_{n+1} = \min(m_n + l_u, M)\\
&\text{endif}
\end{aligned} \tag{8}$$

and using the quantized function

$$\beta = f(m) = \beta_W - \alpha(m_n / M) \tag{9}$$

which calculates the slope β_{n+1} given the current slope β_n, where $m_n = 0,..,M$ is an independent variable and $(M, l_d, l_u) \in \mathbb{Z}$, $(\beta_W, \alpha) \in \mathbb{R}$, m_0 are user-definable parameters. As can be seen in (9), the slope β_n varies in the range from ($\beta_W - \alpha$) to β_W.

IV. Simulation Results

The efficiency of the JEM-VS algorithm is verified by extensive simulations using the single carrier system transmitting 16- and 64-QAM signals and the Soft-DFE at the receiver side. The equalizer performances are given in the terms of MSE convergence characteristics and equalization success index (ESI) representing the ratio between the number of successful equalizations and the total number of Monte Carlo runs. Fig. 2 depicts the normalized attenuation response of three ray multipath (Mp) channels; the channel Mp-A is identified as moderate and channels Mp-(C, E) as severe. The Mp channel model is involved in the transmitter filter with a roll-off factor of 0.12. To switch operation modes, the Soft-DFE for 16- and 64-QAM signals respectively uses MSE thresholds {TL1=1.3 dB, TL2=-5.9 dB, TL3=-7.8 dB} and {TL1=8.0 dB, TL2=-1.9 dB, TL3= -4.6 dB}.

The fixed parameters selected in JEM-W and JEM-VS algorithms for 16-QAM [64-QAM] signal are given as follows: the neuron slope β_D for the JEM-D in (4) is $\beta_D = 11.4$ $[\beta_D = 1.955]$ and the parameter α in (9) is $\alpha = 1.1$, $[\alpha = 0.8]$. The rest of user-definable parameters for both algorithms and QAM signals are: $\gamma = 2^{-13}$, $l_d = 6$, $l_u = 20$, $m_0 = 40$ and $M = 1000$.

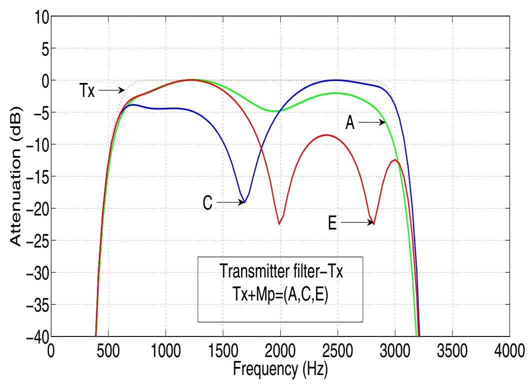
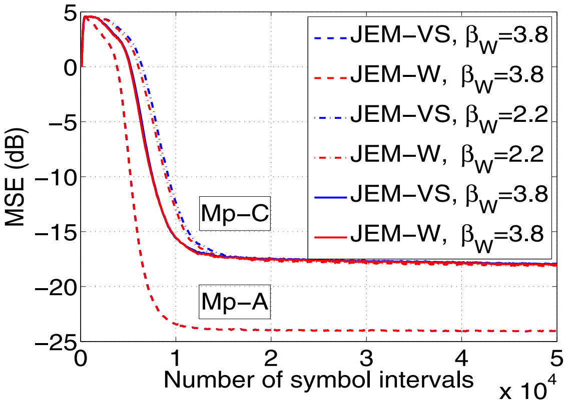
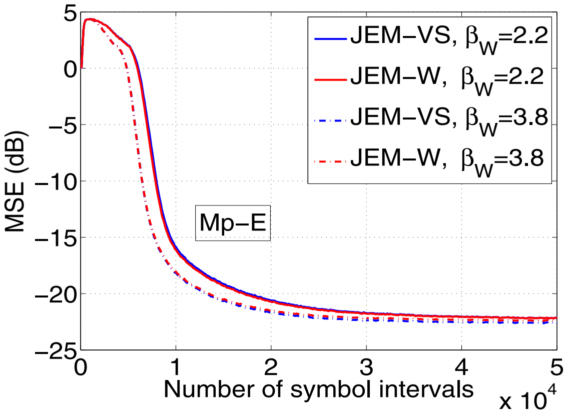


Fig. 2. Attenuation characteristics of Mp channels.

The effective MSE convergence characteristics of the equalizer are presented in Figures 3 and 4. The characteristics are given for the 16-QAM [64-QAM] signal with the signal-to-noise ratio of SNR=25 dB [SNR=30 dB], and the JEM-W and JEM-VS algorithms using the same fixed slopes $\beta_{W,16} = (2.2, 3.8)$ and $\beta_{W,64} = 1.6$ in relations (3) and (9). As can be seen in figures, despite the fact the slope β_n in JEM-VS is smaller than β_W in the JEM-W (see relation in (9))



(a)



(b)

Fig. 3. MSE convergence for JEM-W and JEM-VS with $\beta_W = (2.2, 3.8)$, 16-QAM: a) Mp-(A,C), b) Mp-E.

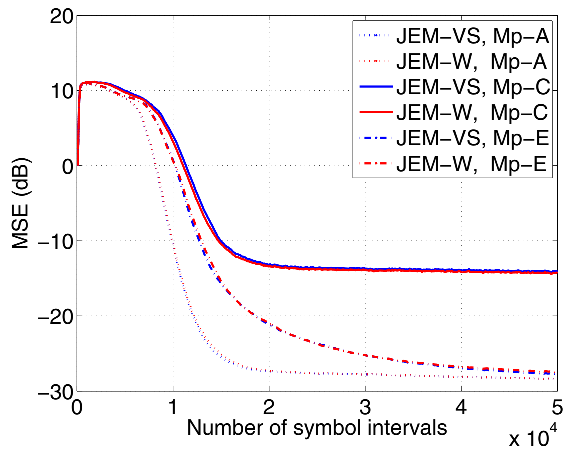


Fig. 4. MSE convergence for JEM-W and JEM-VS with $\beta_W = 1.6$, 64-QAM, Mp-(A,C,E).

the MSE convergence characteristics are practically the same for both algorithms.

In Fig. 5 we compare the influence of JEM-W and JEM-VS algorithms on the ESI index for the 64-QAM signal; the slope β_W in (3) and (9) varies in the range (1.0-2.0). For both severe channels the equalizer reaches a higher equalization success with the JEM-VS algorithm. This indicates that the more precise estimation of the

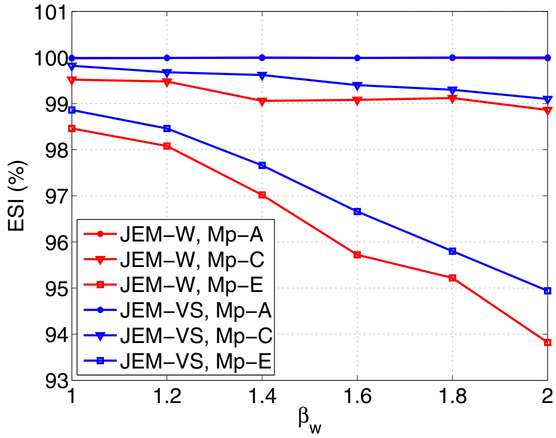


Fig. 5. ESI index for JEM-W and JEM-VS for 64-QAM.

received signal power spectrum alleviates the FBF filter mismatch effects. For 16-QAM the corresponding ESI is in the range from 99% to 100%, for both algorithms.

II. Conclusion

The JEM whitening algorithm with the time variable neuron slope is a more efficient power spectrum estimator than its counterpart with the fixed slope. The result of this improvement is the increased successful equalization of the Soft-DFE scheme. On the other hand, the cost of this benefit is covered by the increased algorithm complexity. Besides the blind DFE context presented in this paper, the JEM-VS all-pole whitener could be also applied to other types of equalizers and signal receivers.

References

[1] J. R. Treichler, M. G. Larimore and J. C. Harp "Practical Blind Demodulators for High-Order QAM Signals," *Proc. IEEE*, vol. 86, pp. 1907-1926, 1998.

[2] J. G. Proakis, *Digital Communications.* 3rd ed. New York: McGraw-Hill, 1995.

[3] M. Ghosh, "Blind Decision Feedback Equalization for terrestrial television receivers," *Proc. IEEE*, vol. 86, pp. 1907-1926, 1998.

[4] J. Labat, O. Macchi and C. Laot, "Adaptive decision feedback equalization: can you skip the training period?," *IEEE Transactions on Communications*, vol. 46, no. 7, pp. 921-930, 1998.

[5] V. R. Krstić and M. L. Dukić, "Blind DFE with maximum-entropy feedback," *IEEE Signal Processing Letters*, vol. 16, pp. 26-29, 2009.

[6] A. J. Bell and T. J. Sejnowski, "An information-maximization approach to blind separation and blind deconvolution," *Neural Computation*, vol. 7, no. 6, pp. 1129-1159), 1995.

[7] Y. H. Kim and H. S. Shamsunder, "Adaptive algorithms for channel equalization with soft decision feedback," *IEEE Journal on Selected Areas in Communications,* vol. 16, no. 9, pp. 1660-1669, 1998.

[8] D. N. Godard, "Self-recovering equalization and carrier tracking in two-dimensional data communication systems," *IEEE Transactions on Communications,* vol. 18, no. 11, pp. 1867-1875, 1980.

[9] V. R. Krstić and M. L. Dukić, „Decision Feedback Blind Equalizer with Tap-Leaky Whitening for Stable Structure-Criterion Switching," *International Journal of Digital Multimedia Broadcasting*, Vol. 2014, Article ID 987039, pages 10. Available: ,http://dx.doi.org/10.1155/2014/907039.

[10] M. Kamenetsky and B. Widrow, "A Variable Leaky LMS Adaptive Algorithm", in *Proc. Thirty-Eighth Asilomar Conference on Signal, Systems and Computers*, vol.1, pp. 125-126, 2004.