A Variable Leaky Entropy-Based Whitening Algorithm for Blind Decision Feedback Equalization

Vladimir R. Krstić[1], Aco M. Stevanović[2], Borislav Lj. Odadžić[3]

[1] "Mihajlo Pupin" Institute, University of Belgrade, Volgina 15, 11060 Belgrade, Serbia

vladimir.krstic@pupin.rs

[2] Astel Project, Kraljice Natalije 38, 11000 Belgrade, Serbia

[3]Technical Faculty "Mihajlo Pupin" Zrenjanin, University of Novi Sad, Serbia

Abstract This paper addresses the joint entropy maximization algorithm constrained by the variable leaky factor (JEM-VL) aiming at mitigating the feedback filter (FBF) mismatch effects characterizing the operation of the decision feedback equalizer which, at the start of adaptation, swaps positions of feedforward and feedback filters so that the latter acts as a linear all-pole whitener of the received signal. The FBF mismatch is a result of disparity between the FBF setup achieved in the blind mode by observing channel outputs and an excepted FBF setup in the tracking mode which is driven by detected data symbols. Depending on the given signal complexity and inter-symbol interference severity, the FBF filter mismatch is typically manifested by the equalizer convergence instability or even the catastrophic error propagation effects arising at the time of the equalizer structure-criterion switching from the blind to the tracking operation mode. The constraint of superfluous coefficients of the FBF filter by means of the JEM-VL algorithm eliminates the equalizer convergence instability at the time of its switching and, consequently, increases equalization successfulness. The efficiency of the JEM-VL algorithm is verified by simulations using the 64-QAM signal.

Keywords Blind adaptive equalization, decision feedback equalizer, joint entropy maximization cost, coefficients leakage, variable leaky.

1 Introduction

The adaptive equalization of linear time dispersive channels is one of the fundamental problems in data communications. Traditionally, adaptive equalizers operate with assistance from the training signal (pilot) that ensures a quick and precise adaptation of equalizer parameters. On the other hand, the transmission of pilots is unproductive from the perspective of efficient exploitation of the available channel bandwidth [1-3]. Particularly, it becomes critical in the system environments calling for a frequent sending of pilots [3]. This spectral efficiency problem has created unsupervised (blind) equalization which has no access to a pilot and, hence, increases the effective system data throughput.

In the area of blind equalization a decision feedback equalizer structure is of a particular research interest because of its performance and implementation advantages over a liner equalizer (LE). As well-known, the decision feedback equalizer (DFE) combines linear feedforward (FFF) and feedback (FBF) filters where the latter exploits the previously detected symbols to cancel the post-cursor inter-symbol interference (ISI). Based on the assumption of correctly detected symbols, the DFE removes the post-cursor ISI more efficiently than the corresponding LE, without noise enhancement and by using a smaller number of coefficients [4]. On the other hand, the performance of DFE can be seriously degraded by erroneously detected symbols when they propagating through the decision-

feedback loop [4]. Although present in both blind and non-blind DFEs, the phenomenon of error propagation is inherently present at early beginning of a DFE blind adaptation process threatening to break it off.

Based on the recent developments, the existing solutions of DFE blind adaptation may be divided into two major approaches. In the first approach, the FFF and FBF filters of DFE are jointly optimized by means of the same cost criterion [2], [5] and in the second approach they are optimized independently of each other [6-9]. The joint optimization of FFF and FBF filters is practical from the implementation point of view, but it is characterized by a convergence state that can lead to the so-called degenerative solutions [10] when the equalizer does not recover transmitted data but generates output symbols by its own. As it is proved in [10,11], this pathology can be avoided by additional constraints depending on the given signal constellation complexity and channel severity. In the second approach, which is motivated by the idea of evading the error propagation effects, the self-optimized DFE (SO-DFE) [7] temporarily drops a hard decision device, transforms itself into the linear blind equalizer which places the FBF filter before the FFF one to open the channel signal enough and then switches itself back to the classical DFE structure performing decision-directed minimum mean square error (MMSE) equalization. This approach is based on the theory that the infinite-length MMSE-LE and MMSE-DFE equalizers share the same components [7,12] which, in the frequency domain, can be factored into the all-pole recursive amplitude equalizer and the phase equalizer which, respectively, compensates for minimum and maximum components of a channel transfer function.

The basic limitation of the SO-DFE scheme comes from its primary model that assumes the equivalence between the MMSE-LE and MMSE-DFE equalizers. This assumption is generally violated in practical implementations based on the finite-impulse response (FIR) filters; as opposed to the optimal infinite-length and finite-length FFF filters which presents the structure similarities, the optimal finite-length FBF filter is not necessarily minimum phase as its infinite-length counterpart [13, 14]. This SO-DFE structure discrepancy, denoted as a filter mismatch, is being manifested through the adaptation process of the FBF filter which, in the blind mode operates as an all-pole filter (whitener) of the received signal and then, in the tracking mode, continues adaptation as the decision-directed equalizer. Another weakness of the SO-DF scheme, which is closely related to the previous one, relates to its sudden structure-criterion switching (transition) from the blind to the tracking mode. Depending on the signal constellation complexity and channel severity, this transition can induce a bursting of signals and parameters leading the equalizer convergence towards instability and even failure.

The blind DFE solutions presented in [8,9,15] demonstrate several different algorithms and/or structure improvements that evade or mitigate limitations of the SO-DFE scheme. The DFE in [8], called Soft-DFE, introduces the new “soft” FBF (SFBF) filter optimizing the Joint Entropy Maximization (JEM) cost through two operation modes. In the blind mode, acting as an all-pole amplitude equalizer (whitener) of the received signal, the SFBF compensates more efficiently for deep spectral nulls than the corresponding whitener controlled by the extended LMS algorithm [7] and then, in the next soft transition mode, continues entropy maximization as a soft decision-directed equalizer performing in between data whitening and data recovery; effectively, the SFBF improves the received signal whitening and softens the process of the equalizer structure-criterion switching. In [9], the noise-predictive DFE scheme [16] has been proposed which evades the FBF filter mismatch by avoiding both the structure and the criterion switching. It employs two FBF filters, placed before and after the FFF filter as well as the decision rule which softly transforms the equalization between extreme linear and hard decision operating modes; the method is verified by the 64-QAM signal. In [15], the self-optimized DFE uses the zero-pole whitener instead of a commonly used all-pole whitening filter. The zero-pole whitener is viewed as the most general

whitening filter to equalize different types of channels and to approximate a long FBF filter with a smaller number of poles and zeros preventing the error propagation phenomena.

Most recently, in [17] and [18], we have innovated operation of the SFBF filter in the blind mode by introducing the coefficient leakage term [19,20] into the JEM whitening algorithm; the goal was to avoid the overgrow of the whitener's coefficients caused by the (64,128)-QAM signals which presented larger variance distributions at the channel output then the (16,32)-QAM signals. In [17] we varied both the leakage rate and the neuron slope to improve the equalizer MSE convergence performance for the 64-QAM signal, and in [18], for the fixed leakage rate, we considered the influence of the adaptive neuron slope on the convergence of the whitener for the (16, 64, 128)-QAM signals. Motivated by these results and the work in [21], which proposes a variable leaky LMS algorithm for linear filter estimation and system identification problems, in this paper we have introduced a variable leaky into the JEM whitening algorithm to eliminate undesirable effects of the SFBF filter mismatch on the Soft-DFE performance. Aiming to maintain a good balance between the MSE convergence rate and the residual MSE of the Soft-DFE, our focus in this paper is to improve the equalization successfulness index which is used as an effective measure of the equalizer robustness to both the SFBF mismatch and sudden structure-criterion switching.

The rest of the paper is organized as follows. Section 2 describes a system model and the operation of the Soft-DFE scheme through three operation modes. Section 3 recalls the backgrounds and algorithms of the SFBF and defines the difficulties of the neuron slope selection in the context of SFBF mismatch. In Section 4, the JEM-type whitening algorithm with the variable leaky term is derived. Section 5 presents simulation results that demonstrate the impact of the variable leaky JEM whitening algorithm on the effective performances of the Soft-DFE equalizer with the 64-QAM signal. The final Section 6 addresses conclusions.

2 Soft-DFE: Basics of structure-criterion adaptation

Throughout this paper, the fractionally-spaced Soft-DFE equalizer is considered as the receiver in a single carrier system presented in Fig. 1. The sequence of data symbols $\{a_n\}$, generated at the rate 1/T (T is symbol period in seconds), consists of i.i.d. complex, zero-mean, variables with a finite variance and sub-Gaussian distribution. The time-invariant channel pulse response $\{h_n\}$ represents combined effects of the transmitter filter, channel impulse response and anti-alias filter at the receiver side. The noise is real Gaussian with zero-mean and independent of the input data. In the equalizer steady-state, the signal $x(t)$ at the input of the fractionally-spaced FFF filter is sampled at the rate 2/T and odd and even samples $x(t_0 + nT - iT/2) = x_{n,i}$ are alternatively shifted to the delay lines of the corresponding FIR symbol-spaced filters defined with coefficient vectors $\mathbf{c}_i = [c_{i,1}, ..., c_{i,L}]^T$, $i = 1, 2$. Otherwise, in the equalizer blind-state when FBF acts as a front-end equalizer, samples $x_{n,i}$ are applied to the FBF filter including two FIR filters defined by vectors $\mathbf{b}_i = [b_{i,1}, ..., b_{i,N}]^T$.

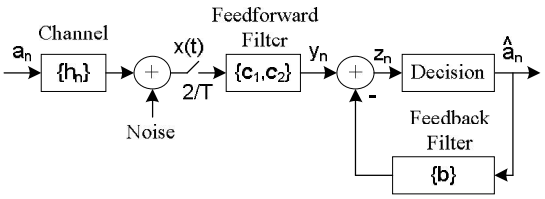
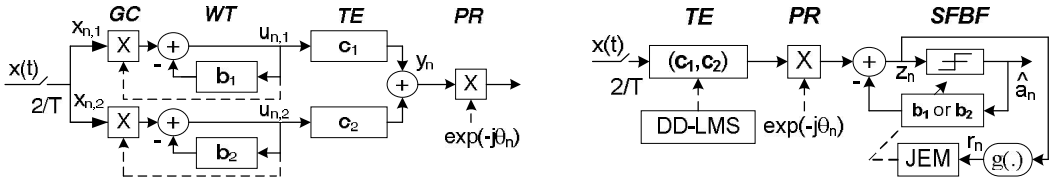


Fig. 1 Blok diagram of the system with Soft-DFE.

To achieve the optimal coefficients setting and carrier phase recovery, the Soft-DFE passes through three operation stages named: blind acquisition, soft transition and tracking. During the blind operation mode, the Soft-DFE effectively acts as a linear fractionally-spaced equalizer (FSE) which divides the equalization task between four signal transformers ordered in the following cascade: gain control (*GC*), whitener (*WT*), equalizer (*TE*) and phase rotator (*PR*), Fig. 2a. The gain control *GC* and whitener *WT* are coupled in pair where *GC* recovers the transmitted signal energy using a slightly modified single-coefficient equalization rule [7] and the JEM- whitener equalizes a non-flat channel spectrum. At the same time and independently of (*GC*+*WT*), the equalizer *TE* controlled by constant modulus algorithm (CMA) [22] compensates for a phase distortion (introduced by a channel-whitener combination). In the next soft transition mode, one of two JEM-whiteners, selected according to energy criterion, transforms itself back into the decision-directed soft feedback filter *SFBF* which continues with JEM adaptation, while the equalizer *TE* switches adaptation rule from the CMA to the decision-directed LMS (DD-LMS), Fig 2b. Effectively, during the soft transition mode, the Soft-DFE is optimized by the combined DD-(MMSE+JEM) criterion. Finally, for the signal eye opened enough, the complete Soft-DFE is being switched into the classical MMSE-DFE controlled by the DD-LMS algorithm (tracking mode).

The phase rotator *PR* is given in the form of the digital phase locked-loop which operation is modified in a way to avoid the contestation spin effects following the CMA equalization of high-order QAM signals [1]. *PR* begins with the carrier phase acquisition in the blind mode by using the reduced signal constellation including only twelve corner-symbols with the largest energy, and then, for the enough opened signal eye, continues operation with the full constellation. The carrier phase estimation is not a particular focus of this paper.



(a) **(b)**

Fig. 2 Scheme of the Soft-DFE equalizer. **a** blind acquisition, **b** soft transition mode

The Soft-DFE structure-criterion adaptation is controlled by the MSE estimator which has a task to switch both the structure and the adaptation criterion according to the *a priori* selected MSE threshold levels (TL): for TL1 the Soft-DFE switches itself from the blind to the soft transition mode, for TL2 from the soft transition to the tracking mode and the threshold TL3 switches *PR* operation between reduced and full signal operation. The MSE threshold TL3 is also used as a measure of equalization successfulness quantified by the equalization success index (ESI).

3 Soft Feedback Filter: backgrounds and JEM algorithms

In this section, we shell recall the information theoretic backgrounds of the SFBF filter (equalizer) and its practical realization in the complex domain. In contrast to the MMSE-FBF, which commonly exploits hard-decision symbol estimates to cancel the post-cursor ISI, the SFBF performs the same task employing soft-decision symbol estimates provided by a monotone function (neuron) $g(\cdot)$ that maximizes the joint Shannon's entropy (JEM) according to the blind deconvolution theory by Bell and Sejnowski [23]. As illustrated in Fig. 3, the neuron function $g(\cdot)$ in the structure of the SFBF performs mapping of inputs symbols z_n in a way that maximizes the joint entropy of output

symbols r_n and, assuming the previous output symbols are correct[1], $r_{n-j} = a_{n-j}$, $j = 1,...,N$, eliminates their interfering effect on the current channel output x_n; the corresponding input-output relation of the SFBF is given by $z_n = x_n - \mathbf{b}_n^T \mathbf{r}_n$ where $\mathbf{r}_n = [r_{n,1},...,r_{n,N}]^T$ and $\mathbf{b}_n = [b_{n,1},...,b_{n,N}]^T$ are column vectors and the current output is $r_n = g(z_n)$.

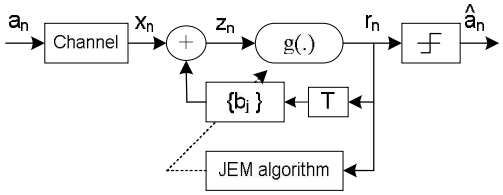


Fig. 3 Soft feedback filter: basic model.

Assuming the previous symbols are correctly detected and channel outputs are noiseless[2] the joint Shannon's entropy $H[r_{n,1},...,r_{n,N+1}]$ of the SFBF model presented in Fig. 3 is given by [24]

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

where $E\{\cdot\}$ denotes statistical expectation. It is worth to be noted that the main idea underlying the SFBF equalizer model is the observation that the maximization of the joint entropy $J_H(\mathbf{b}_n)$ in the presence of ISI (i.e., dependence) leads to the reduction of the statistical dependence between the current output $r_n = g(z_n)$ and the previous outputs (r_{n-j}, $j = 1,...,N$) and, hence, to ISI suppression. In the probability density function domain, the SFBF transforms the PDF of the input sequence into the sequence presenting the uniform distribution in a limited range of variables $\{a_n\}$.

The selection of the mapping function $g(\cdot)$ is a key point in the SFBF design. According to the theory in [23], the function $g(\cdot)$ is selected to suitably fit the expected cumulative distribution of inputs z_n in accordance with the relation $g(z) \approx \int_{-\infty}^{z} p(u)du$, $g(z) \in [0,1]$, where $p(z)$ is PDF of an input z. Consequently, the neuron's slope $\partial g(z_n)/\partial z_n$ has to match the PDF of an input z_n, i.e., $g(z_n)' \approx p(z)$. But, in practice, the selection of neuron function is not a simple task since the PDF of ISI in a channel output x_n, and thus of z_n ($z_n = x_n - \mathbf{b}_n^T \mathbf{r}_n$) is generally unknown and there is also a lack of appropriate nonlinearities. To resolve this difficulty one approach is to use parametric nonlinearities $g(z_n, \beta)$ where the parameter β can be employed as a tool to vary the "slope" of the neuron in a way to be as close as possible to the expected cumulative distribution of ISI. Generally, the optimal

[1] The hypothesis of correctness of previous detected symbols is commonly use in the DFE analysis.

[2] Although the outputs z_n are not noise free, the noiseless system model is assumed in order to simplify the derivation of the JEM algorithm. In simulations presented in the paper we have used more realistic noisy channels.

slope enabling a neuron unit to maximize the joint entropy must be inversely proportional to the variance of its input distribution.

In the following, we will present the SFBF equalizer extended to the complex domain and identify difficulties of the neuron slope selection reflecting complexity of the applied QAM signal constellations and specifics of the self-optimized Soft-DFE scheme. The complex model of the SFBF equalizer is designed for the complex-valued continuously differentiable function of complex variable z_n [8]

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

and its stochastic gradient (ascent) algorithm obtained for the cost $J(\mathbf{b}_n) = \ln|\partial r_n / \partial z_n|$ is given by

$$b_{n+1,j} = b_{n,j} - \mu z_n \left(1 - \beta |z_n|^2\right) r_{n-j}^*, \ j = 1, ..., j \tag{3}$$

where β is the real positive number which varies the neuron's slope, μ is the adaptation step-size and the operator $*$ denotes complex conjugation.

In order to meet the operation requirements of the Soft-DFE scheme, the basic model of the SFBF is modified according to the two following heuristic simplifications:

- in the blind mode, the SFBF acts as a linear all-pole filter (see *WT* in Fig. 2a) optimized by the JEM-type whitening algorithm (JEM-W) given by

$$b_{n+1,\mathrm{j}} = b_{n,j} - \mu_W u_n \left(1 - \beta_W |u_n|^2\right) u_{n-j}^*, \ j = 1, ..., N \tag{4}$$

- in the soft transition mode, based on the hypothesis of correctly detected symbols, the SFBF transforms itself into a soft decision-directed equalizer (see *SFBF* in Fig. 2b) controlled by the JEM-D algorithm given by

$$b_{n+1,j} = b_{n,j} - \mu_D z_n \left(1 - \beta_D |z_n|^2\right) \hat{a}_{n-j}^*, \tag{5}$$

where $\{\mu_W, \mu_D\}$ are adaptation step-sizes, and $\{\beta_W, \beta_D\}$ are neuron-slopes; in recursion (4) the index i of the vector $\mathbf{b}_{n,i}$ is dropped for simplicity.

Thus, based on the above structure and algorithm modifications, the *SFBF* equalizer transforms the PDF of the input sequence using two different slopes to match statistics of the input data provided that the step-sizes $\{\mu_W, \mu_D\}$ are suitably selected to guarantee stability of JEM-W and JEM-D algorithms. In the blind mode, the whitener *WT* manipulates the unknown PDF of the received signal in order to reconstruct the second-order statistic of the given signal, and the optimization of the JEM-W algorithm is achieved by varying the neuron-slope β_W. In the next soft transition mode, the *SFBF* continues to deal with the PDF of the residual post-cursor ISI acting as a higher-order statistics equalizer; in this stage, the JEM-D algorithm is driven by symbol estimates $\hat{a}_n$ of the given signal, and its optimization is achieved by the neuron-slope β_D. According to this scenario, the expected variance of ISI of the channel outputs x_n is larger than the corresponding variance of symbol estimates z_n and therefore the selected slope β_W should be smaller than the slope β_D. Besides, it is worth to be noted that the slope β_D can be seen as a statistical constant for the given signal constellation since the JEM-D algorithm (5) correlates the pseudo-error signal $z_n \left(1 - \beta_D |z_n|^2\right)$ with symbol estimates $\hat{a}_n$.

To summarize effectiveness of the SFBF we will recall the optimal neuron-slope selections $\{\beta_{W,16}=1.4, \beta_{D,16}=12\}$, $\{\beta_{W,32}=1.2, \beta_{D,32}=10\}$ and $\{\beta_{W,64}=0.2, \beta_{D,64}=2.0\}$ achieved, respectively, for signal constellations 16-QAM, 32-QAM, 64-QAM [17] which present distributions with an increasing variance, and emphasize two facts of particular importance. First, the selected slopes prove the inverse relationship between the neuron-slopes and the variance of input distributions. Second, in dealing with the neuron-slopes β_W it has been verified that the selected slopes β_W for 16- and 32-QAM signals provide a good trade-off between the MSE convergence speed, residual MMSE and the equalization successfulness. On the other hand, it is not the case for the 64-QAM signal where the equalizer performance suffers from a slow convergence rate and high residual MSE for a small β_W ($\beta_{W,64}=0.2$). This behaviour is a result of deficiency of the JEM-W algorithm which for small values of the slopes β_W has no ability to estimate deep spectral fades of channel output. In fact, for small slopes β_W the influence of the quadratic term in the prediction error $e_n(\beta_W, u_n)=u_n\left(1-\beta_W|u_n|^2\right)$ in (4) is not sufficient to provided adequate estimates. Otherwise, by increasing values of the slope $\beta_{W,64}$, the Soft-DFE presents better convergence characteristics but its equalization successfulness gets worse due to an overgrowth of the whitener coefficient norm. To eliminate described deficiency of the JEM-W algorithm with the severely corrupted 64-QAM signal, we have extended JEM-W algorithm with the coefficient leakage term [17] which constrains a coefficients overgrowth. In fact, by introducing the coefficient leakage we have open a space for using larger neuron-slopes β_W and in this way extended the operation of JEM-W algorithm toward more complex signal constellations than (16, 32)-QAM ones. The price of this improvement, however, is a more complex implementation of the JEM-W algorithm whose optimization, besides the adaptation step-size and the neuron slope, depends also on the leakage factor which determines the coefficient leakage rate.

4 JEM Whitening Algorithm with Variable Leaky

Based on the previous findings, it is obvious that the more complex signal constellations than (16, 32)-QAM require the JEM-whitener with an adaptive neuron-slope and/or some an additional signal processing more robust to larger dynamic signals. Besides, it can be a means to facilitate free parameters selections aimed at achieving the desired trade-off between the MSE convergence characteristics and the equalization successfulness.

The innovated JEM-W whitening algorithm presented in this section remains a fixed neuron-slope β_W, as in the recursion (4), and introduces a variable leaky factor to adjust the coefficient leakage rate with respect to the quality of the whitener prediction error. The motivation underlying this approach is observation that the forcing of larger neuron-slopes β_W, which provides a fast and efficient estimation of deep spectral fades at the early beginning of the received signal whitening process, must be followed by larger leakage rates aimed at constraining the coefficients overgrowth and, also, that inadequately large coefficient leaky rates can degrade the equalizer convergence performances.

In the following we will define the JEM whitener constraining superfluous coefficients by leaking them off and then derive the algorithm which performs adaptation of the coefficient leakage process. The JEM cost $J(\mathbf{b}_n)$ modified by the term $\gamma\|\mathbf{b}_B\|^2$ which penalizes a whitener coefficient vector norm overgrowth is

$$J_L(\mathbf{b}_B) = J(\mathbf{b}_B) - \gamma \|\mathbf{b}_B\|^2 \ , \ \gamma \|\mathbf{b}_B\|^2 \ll J(\mathbf{b}_B) \tag{6}$$

and the corresponding extension of the JEM-W algorithm (JEM-FL) is given by

$$b_{n+1,\mathrm{j}} = (1-\gamma) b_{n,j} - \mu_W u_n \left(1 - \beta_W |u_n|^2\right) u_{n-j}^* \ , \ j = 1, \ldots, N \tag{7}$$

where $\gamma \geq 0$ is the leaky factor which determines the coefficient leakage rate; the index B in (6) refers to the blind operation mode. As can be seen in (7), the leaky term γb_n acts in opposition and independently of the entropy-gradient term controlled by both the whitener outputs u_n and the neuron-slope β_W, and in this manner it persistently decreases the magnitude of whitener coefficients even though the whitener input is turned off and/or its outputs converge to zero. To diminish undesirable effects of this scenario, the leaky factor is selected to be as small as possible ($\gamma \|\mathbf{b}_B\|^2 \ll J_H(\mathbf{b}_B)$) which is a common practice in coefficient leaky-based filter estimation problems to trade-off bias and variance of estimates [20]. On the other hand, as it is mentioned before, using larger slopes β_W in order to improve the Soft-DFE performance with higher-order QAM signals, asks for adequately large leaky factors able to constraint superfluous coefficients. This has been the motivation to employ a adaptive leaky factor for JEM whitening algorithm which varies between zero and some predetermined positive value $\gamma_{\max}$.

In the case of the JEM-whitener, the adaptation of the leaky factor is based on the *a posteriori* error analysis and the heuristic punish/award discrete rule [21] as well as the ability of the JEM whitener to efficiently manipulate the PDF of the received signal. The *a posteriori* error analysis includes the *a posteriori* errors of the whitener predictions achieved with and without coefficient leakage. The *a posteriori* error e_n^{VL} obtained by the variable leaky JEM-VL algorithm is given by

$$\begin{aligned} &\mathbf{b}_{n+1} = \mathbf{b}_n - \gamma_n \mathbf{b}_n - \mu_W u_n (1 - \beta_W |u_n|^2) u_{n-j}^* \ , \ j = 1, \ldots, N \\ &\tilde{u}_n = x_n - \mathbf{b}_{n+1}^T \mathbf{u}_n \\ &e_n^{VL} = \tilde{u}_n (1 - \beta_W |\tilde{u}_n|^2) \end{aligned} \tag{7}$$

and the *a posteriori* error e_n^W obtained by the JEM-W ($\gamma = 0$) algorithm is given by

$$\begin{aligned} &\mathbf{b}_{n+1} = \mathbf{b}_n - \mu_W u_n (1 - \beta_W |u_n|^2) u_{n-j}^* \\ &\tilde{u}_n = x_n - \mathbf{b}_{n+1}^T \mathbf{u}_n \\ &e_n^W = \tilde{u}_n (1 - \beta_W |\tilde{u}_n|^2) \ . \end{aligned} \tag{8}$$

It should be noted in (7) and (8) that both errors are obtained using the same (current) value of the input x_n while the *a posteriori* value of the output $\tilde{u}_n$ in (7) and (8) is, respectively, obtained for the current leaky γ_n and for zero leaky ($\gamma = 0$). In the next step, the quality of errors is decided by using the punish/award heuristic given by the following if-else relation

$$\begin{aligned} &\text{if } e_n^{VL} > e_n^W \text{ then} \\ &\quad \text{set } m_{n+1} = \max(m_n - l_d, 0) \\ &\text{else} \end{aligned} \tag{9}$$

set $m_{n+1} = \min(m_n + l_u, M)$

end if

which analyses the difference $\Delta = e_n^{VL} - e_n^W$ and decides when and how much to increase or decrease the leaky. Finally, the amount of the current leaky change is calculated by the quantized exponential function

$$\gamma_n = f(m_n) = \gamma_{\max}(m_n / M) \tag{10}$$

where $m_n = 0,...,M$ is an independent variable, and $(M, l_d, l_u) \in \mathbb{Z}$, $\gamma_{\max} \in \mathbb{R}$, m_0 are user-definable parameters.

In order to gain a better understanding of the JEM-VL algorithm influence on the whitener and the equalizer behaviour, we have varied the maximum value of the leaky factor $\gamma_{\max}$, and the corresponding results of simulations are presented figures 4 and 5; the presented results are averaged over 200 Monte Carlo runs for the 64-QAM signal, the Mp-E channel (see Fig. 6) under a signal-to-nose ratio (SNR) of 30 dB, the neuron-slope $\{\beta_W = 2.0,\ \beta_D = 1.95\}$ and the set of leaky parameters $\{\gamma_{\max} = (2^{-13}, 2^{-12}, 2^{-11}),\ M = 400,\ l_d = 6,\ l_u = 40,\ m_0 = 40\}$. Fig. 4 presents the influence of the maximum leaky rate on the time variation of the **(a)** *a posteriori error* e_n^{VL}, **(b)** difference $\Delta = e_n^{VL} - e_n^W$, **(c)** variable leaky factor γ_n and **(d)** whitener vector norm $\|\mathbf{b}_B\|$. In latter figure, we have compared the influence of the JEM-FL and JEM-VL algorithms on the whitener vector norm overgrowth using the same set of leaky factors $\gamma = \gamma_{\max}$.

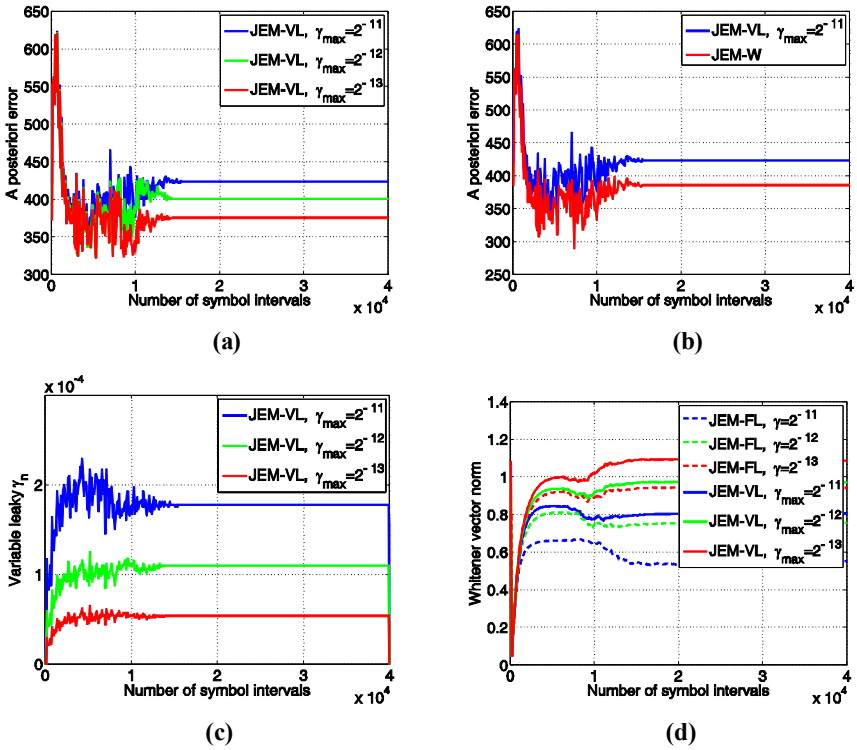


Fig. 4 Influence analysis of the maximum leaky $\gamma_{\max} = (2^{-13}, 2^{-12}, 2^{-11})$ on the time variation of (**a**) *a posteriori* error e_n^{VL}, (**b**) *a posteriori* error difference $\Delta = e_n^{VL} - e_n^W$, (**c**) leaky factor γ_n, (**d**) whitener vector norm $\|\mathbf{b}_B\|$.

Fig. 5 presents one-pass MSE convergence characteristics of the Soft-DFE obtained with JEM-FL and JEM-VL algorithms under the same conditions as in the case of Fig. 4. It is evident that the JEM-VL algorithm provides better convergence characteristics than JEM-FL for the same leaky $\gamma = \gamma_{\max}$. Also, it should be noted that the JEM-VL for $\gamma_{\max} = 2^{-11}$ and the JEM-FL for $\gamma = 2^{-13}$ demonstrate a similar influence on the equalizer convergence.

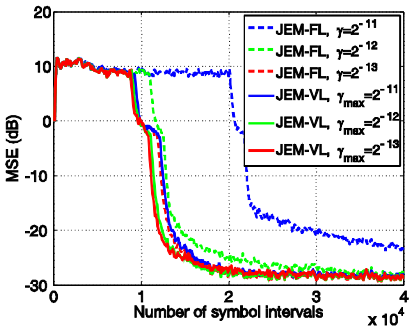


Fig. 5 One-pass MSE convergence of Soft-DFE for JEM-VL: $\{\gamma_{\max} = (2^{-11}, 2^{-12}, 2^{-13})\}$ and JEM-FL: $\{\gamma = (2^{-11}, 2^{-12}, 2^{-13})\}$ algorithms, $\beta_W = 2.0$, $\beta_D = 1.95$, Mp-E, SNR=30 dB.

Based on the previous results, two points should be stressed. First, the using too large (fixed) leaky factors with the JEM-FL algorithm is prohibited since they degrade whitener coefficient estimates and, hence, the equalizer convergence despite of fact the JEM-whitener is suspended after the equalizer transition to the soft-transition mode. Second, the JEM-VL algorithm, which varies the leaky factor in a range from zero to $\gamma_{\max}$, provides better regularization of the whitener coefficient vector norm then the JEM-FL for $\gamma = \gamma_{\max}$ and also softens the equalizer transition process from the blind to the tracking operation mode.

5 Simulation results: Soft-DFE performance

Simulation results presented in this section demonstrate influence of the SFBF equalizer, i.e. JEM-(W,FL,VL) algorithms, on the effective performance of the Soft-DFE equalizer with a particular focus on the efficiency of the JEM-VL algorithm. The performance evaluation is carried out for the 64-QAM single-carrier system with the Soft-DFE equalizer at the receiver side. The simulation results are given in the terms of the equalizer MSE convergence and the equalization success index (ESI) defined as the ratio between the number of successfully completed equalizations and the total number of Monte Carlo runs.

The 64-QAM system is characterized by the multipath (three-ray) channel model (Mp) involved in the transmitter filter with a roll-off factor of 0.12. Figure 6 depicts a normalized attenuation response of the Mp channels which are obtained for attenuation and propagation parameters selected to gradually increase an ISI level. In the given class of Mp channels, Mp-A is declared as moderate and Mp-(C, E) as severe; the SNR at the output of the channel is 30 dB. For the Soft-DFE described in Section 2, the FFF and FBF filter lengths (given in T intervals) are, respectively L=24 and N=5, and the initial values of *TE* coefficients are all zero except of the central ones $c_{1,r} = c_{2,r} = 1.0$. The adaptation step sizes for FFF [FBF] in blind, soft and tracking modes are, respectively, $\mu_{CMA} = 2^{-21}$ [$\mu_W = 2^{-22}$], $\mu_{LMS} = 2^{-20}$ [$\mu_D = 2^{-21}$] and $\mu_{LMS} = 2^{-16}$ [$\mu_{LMS} = 2^{-14}$]. The adaptation of the *GC* is controlled by two steps $\{2^{-11}, 2^{-20}\}$ selected to provide a fast power recovery of the transmitted signal and, also, to

avoid the *WT* in doing the same task since they are coupled. The neuron-slope β_W for JEM-(W,FL,VL) algorithms is used as a free parameter, user-definable parameters of the leaky adaptation rule for the JEM-VL are selected to be $\{\gamma_{\max} = 2^{-11},\ M = 400,\ l_d = 6,\ l_u = 40,\ m_0 = 40\}$ and the optimal neuron-slope for the JEM-D algorithm with the 64-QAM signal is $\beta_D = 1.95$ [19]. The carrier phase rotator *PR* switching between the reduced and the full signal constellation is controlled by the MSE threshold TL3, and the boundary separating the reduced signal space (including twelve corner-symbols) is defined by the constant modulus equals 71. The three operation modes of the Soft-DFE are controlled by the MSE threshold levels given by {TL1=8.0 dB, TL2= -1.94 dB, TL3= -3.0 dB}; the selection of thresholds is based on both the worst-case transmission scenario forcing Mp-(C, E) channels and the equalizer performances.

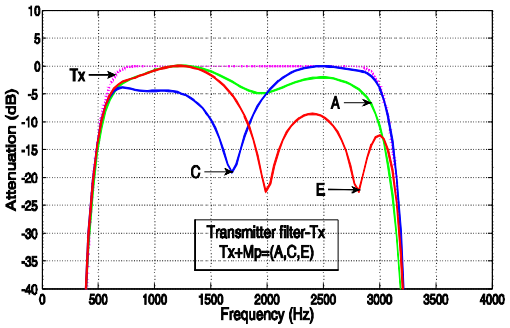
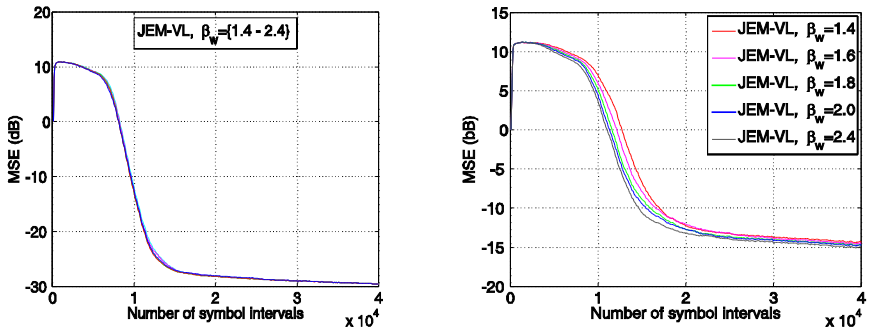
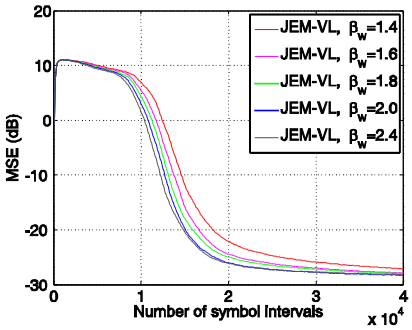


Fig. 6 Attenuation characteristics of Mp-(A, C, E) channels.

The equalizer MSE convergence curves, presented in Fig. 7, demonstrate the influence of the neuron-slope β_W on the JEM-VL performance and, hence, on the effective equalizer convergence characteristics. The selection of the maximum leak value $\gamma_{\max} = 2^{-11}$ is based on the results presented in the previous section and on the evaluation of the ESI index aiming at achieving the best trade-off between the convergence rate, residual MSE and the ESI. Evidently, for the moderate Mp-A channel the Soft-DFE presents practically the same MSE convergence characteristics for the neuron-slope range of interest $\beta_W = (1.6 - 2.4)$. On the other hand, the influence of the slope β_W on the MSE convergence is clearly demonstrated in the case of severe channels Mp-(C, E) channels. It should be noted that the desired convergence trade-off is reached for values of the slope β_W in the vicinity of the slope $\beta_D = 1.95$ which determines optimal performing of the JEM-D algorithm for the 64-QAM signal.



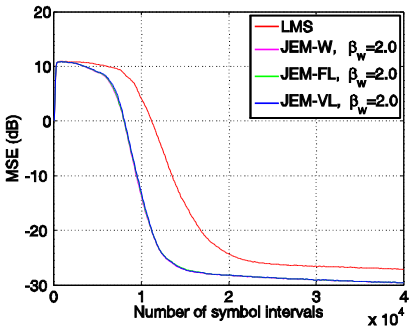
(a) **(b)**



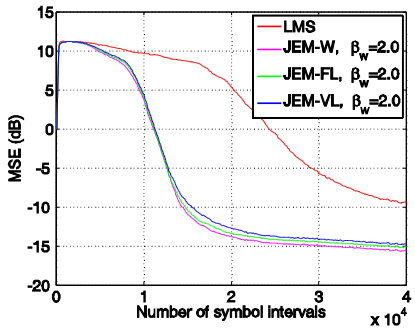
(c)

Fig. 7 MSE convergence characteristics obtained with JEM-VL: $\{\gamma_{\max} = 2^{-11}\ \beta_W = (1.4, 1.6, 1.8, 2.0, 2.4)\}$ and averaged over 200 Monte Carlo runs. **a** channel Mp-A, **b** channel Mp-C, **c** channel Mp-E

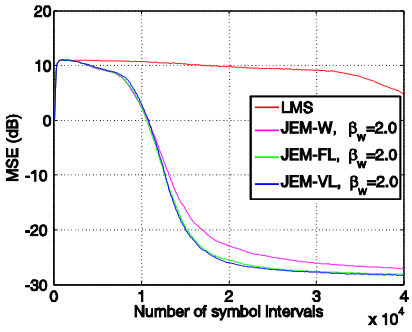
Fig. 8 presents influence of three different whitening algorithms JEM-W:($\beta_W = 2.0$), JEM-FL:($\gamma = 2^{-13}$, $\beta_W = 2.0$) and JEM-VL:($\gamma_{\max} = 2^{-11}, \beta_W = 2.0$) on the MSE convergence. In this test, we have also addressed the LMS algorithm [7] in order to stress its weakness to compensates for deep spectral nulls of the received signal, a fact that is evident in the case of the Mp-(C,E) channels; for the purpose of a correct comparison of LMS with JEM-type algorithms, the soft transition mode in the Soft-DFE scheme was temporally suspended. On the other hand, the JEM-type algorithms demonstrate a high robustness to severe ISI channels. Besides, the equalizer convergence for both the JEM-FL and the JEM-VL algorithms results in a faster convergence in comparison with the original JEM-W algorithm. If we neglect the influence of the JEM-FL and JEM-VL algorithm on the equalization successfulness, it is interesting to note that the JEM-VL for $\gamma_{\max} = 2^{-11}$ influences the equalizer convergence on a similar way as the JEM-FL with $\gamma = 2^{-13}$ does despite the fact that the maximum leaky value $\gamma_{\max}$ is four times larger than the leaky factor in the JEM-FL. It is result of a time variable leaky rate in the JEM-VL algorithm that employs the coefficient leakage only when and in an amount it is necessary. Let us remember that the JEM-FL for the fixed leaky factor $\gamma = 2^{-11}$ dramatically degrades MSE convergence, see Fig. 5.



(a)



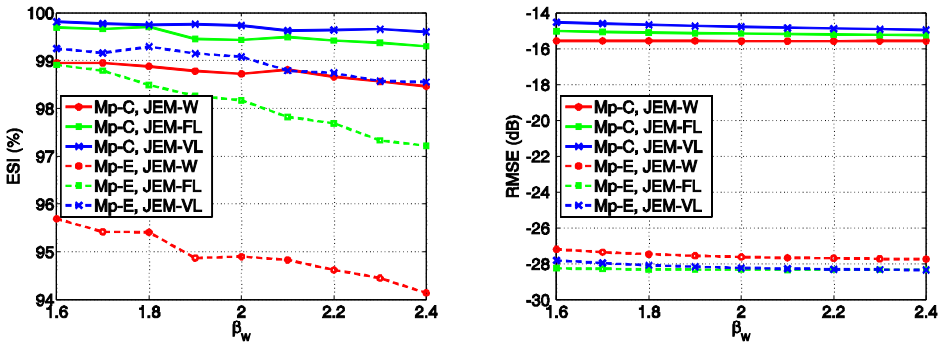
(b)



(c)

Fig. 8 MSE convergence comparison for four whitening algorithms: LMS, JEM-W ($\beta_W = 2$), JEM-FL ($\beta_W = 2.0$) and JEM-VL ($\beta_W = 2.0$). **a** channel Mp-A, **b** channel Mp-C, **c** channel Mp-E

To get a deeper insight into the JEM-(W,FL,VL) algorithms influence on the equalizer convergence behaviour at the time of the structure-criterion switching, we carried out extensive measuring of ESI index which used as a quantitative measure of the equalizer ability to successfully perform structure-criterion switching. This test is also used to get a better view of the algorithms influence on residual MSE (RMSE) which is measured in the steady-state of successfully finished equalizations. Fig. 9 presents the ESI and RMSE performance of the equalizer obtained by JEM-W, JEM-FL:($\gamma = 2^{-13}$) and JEM-VL:($\gamma_{\max} = 2^{-11}$) algorithms that vary the neuron-slope β_W in a relatively large range from 1.6 to 2.4; in the given curves each measuring point is the result of averaging over the 10000 Monte Carlo runs. The obtained results have revealed the following important findings. First, the highest ESI index is achieved by the JEM-VL algorithm for the neuron-slope β_W in the range from 1.8 to 2.0; the ESI for Mp-C and Mp-E channel is, respectively, higher than 99.7% and 99%. Also, this neuron-slope range coincides with the slope range for which the equalizer achieves the best convergence characteristics (see figures 7 and 8). Second, the achieved RMSE performance does not favour any of the applied JEM algorithms. These results indicate that the potential degradation of the RMSE introduced by the coefficient leakage (resulting from a biased estimate of whitening) is not of practical importance. Third, the JEM-W algorithm demonstrates an evident inferiority to the ESI index for both Mp-(C, E) channels. The corresponding results achieved with the moderate channel Mp-A, which are not depicted in the figures for better readability of graphs, are practically the same for all three algorithms, the ESI is 100% and the RMSE is less than -28 dB.



(a) **(b)**

Fig. 9 Soft-DFE: equalization success index and residual MSE versus the neuron-slope β_W for JEM-W, JEM-FL, JEM-VL algorithms and Mp-(C,E) channels. **a** ESI, **b** RMSE

Based on the previous analysis for the 64-QAM signal, it can be concluded that the JEM-VL algorithm with the neuron-slope selected in the range from 1.8 to 2.0 provides the best trade-off between the MSE convergence rate, residual MSE and equalization successfulness. Besides, it is interesting to note that the optimal values of the slope β_W are located in the vicinity of the β_D, i.e., $\beta_W \approx \beta_D$. This fact simplifies the JEM-VL optimization process since it can be started for the neuron slope $\beta_W \approx \beta_D$, where the slope β_D is known for the given signal constellation.

6 Conclusions

The results presented in this paper provide more evidence of importance of the received signal whitening for the Soft-DFE scheme which optimizes FFF and FBF filters independently combining CMA, JEM and MMSE criteria. The JEM-FL algorithm using a fixed leaky-factor favours larger neuron-slopes and, in this way, provides a faster equalizer convergence rates and lower residual MSE with QAM signals presenting increased variance statistics but, on the other hand, it does not guarantee desirable equalization successfulness. This insufficiency is eliminated by the variable leaky JEM-VL algorithm which exploits the whitener *a posteriori* error statistics to adapt the coefficient leaky rate. The whitener estimation improvement achieved by the adaptive coefficient leakage mitigates the feedback filter mismatch effects, which is verified by simulations presenting the increase of equalization successfulness with severe ISI channels. Besides this advantage, the using of variable leaky simplifies selection of the whitener neuron-slope value which provides the equalizer performance trade-off.

Acknowledgements This work was supported by the Ministry of Science and Technological Development of the Republic of Serbia; the project of technological development TR 32037, 2011-2016.

References

1. Treichler, J. R., Larimore, M. G., & Harp J. C. (1998). Practical Blind Demodulators for High-Order QAM Signals. *Proc. IEEE*, vol. 86, no. 10, (pp. 1907-1926).
2. Ghosh, M. (1998). Blind Decision Feedback Equalization for terrestrial television receivers. *Proc. IEEE*, vol. 86, no. 10, (pp. 1907-1926).
3. Savaux, V., Bader, F., & Palicot, J. (2016). OFDM/OQAM Blind Equalization Using CNA Approuch. *IEEE Trans. Signal Processing,* vol. 64, no. 9, (pp. 2324-2333). DOI: 10.1109/TSP.2016.2519000
4. Proakis, J. G. (1995). *Digital Communications.*3rd ed. New York: McGraw-Hill.
5. Papadias, C. B., & Paulraj, A. (1995). Decision-Feedback Equalization and Identification of Linear Channels Using Blind Algorithms of the Bussgang Type. In *Proceedings of Twenty-Ninth Asilomar Conference on Signals, Systems and Computers*, Pacific Grove, CA, USA, (pp. 335-340).
6. Rocha, C. A. F., Macchi, O., & Romano, J. M. (1994). An Adaptive Nonlinear IIR Filter foe Self-learning Equalization. In *Proceedings of International Telecommunications Symposium* – IT94, Rio de Janeiro, Brazil, (pp.184-190).
7. Labat, J., Macchi, O., & Laot, C. (1998). Adaptive Decision Feedback Equalization: Can You Skip the Training Period? *IEEE Trans. Commun.*, vol. 46, no. 7, (pp. 921-930).
8. Krstić, V. R., & Dukić, M. L. (2009). Blind DFE With Maximum-Entropy Feedback. *IEEE Signal Processing Letters*, vol. 16, no. 1, (pp. 26-29).
9. Goupil, A., & Palicot, J. (2010). An Efficient Blind Decision Feedback Equalizer. *IEEE Commun. Letters,* vol. 14, no. 5, (pp. 462-464).
10. Szczecinski, L. L., & Gei, A. (2002). Blind decision feedback equalisers, how to avoid degenerative solutions," *Signal Processing*, vol. 82, issue 11, (pp. 1675-1693).
11. Filha, J. M., Mirinda, M. D., & Silva, T. M. (2011). An efficient algorithm for Decision Feedback Blind Equalization. Revista Telecomunicacoes, Vol. 13, No. 02, (pp. 79-86).
12. Cioffi, J. M., Dudevior, G. P., Eyuboglu, M. V., & Forney, G. D. (1995). MMSE Decision-Feedback Equalizers and Coding-Part I: Equalization Results. *IEEE Trans. Commun.*,vol. 43, no. 10, (pp. 2582-2594).
13. Al-Dhahir, N., & Cioffi, J. C. (1995). MMSE Decision-Feedback Equalizers: Finite-Length Results. *IEEE Trans. Information Theory*, vol. 42, no. 4, (pp. 961-975).
14. Casas, R. A., Johnson Jr., C. R., Harp, J., & Caffee, S. (1999). On Initialization Strategies for Blind Adaptive DFEs. In *Proceedings of 1st IEEE Wireless Communications and Networking Conference*, New Orleans, LA, USA, (pp. 792-796).
15. Chang W. C., & Chuang, S.H. (2014). A Zero-Pole Whitening Filter in Adaptive Blind Decision Feedback Equalizers. In *Proceedings of IEEE Inter. Conf. on Systems, Man, Cybernetics*, Oct. 5-8, 2014, San Diego, CA, USA, (pp.3259-3264).
16. Belfiore, C. A., & Park, J. H. (1979). Decision Feedback Equalization. *Proc. IEEE,* vol. 67, no. 8, (pp.1143-1156).

17. Krstić, V. R., & Dukić, M. L. (2014). 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, http://dx.doi.org/10.1155/2014/987039.
18. Krstić, V. R. (2016). Entropy-based stochastic gradient algorithm with adaptive neuron slope for all-pole filtering. *Electronic Letters, The Institute of Engineering and Technologies*, vol. 52, Issue 5, 03 March, (pp. 397-399). DOI: 10.1049/el.2015.3052.
19. Rey, G. J., Bitmead, R. R., & Johnson Jr., C. R. (1991). The Dynamics of Bursting in Simple Adaptive Feedback Systems with Leakage. *IEEE Trans Circuit and Systems*, vol. 38, no. 5, (pp. 476-488).
20. Ljung, L., & Sjoberg, J. A. (1992). Comment on "leakage" in adaptive algorithms. *Department of Electrical Engineering*, Linkoping University. Available at: http://www.diva-portal.org.
21. Kamenetsky, M., & Widrow, B. (2004). A Variable Leaky LMS Adaptive Algorithm", in *Proc. Thirty-Eighth Asilomar Conference on Signal, Systems and Computers*, vol.1, (pp. 125-126).
22. Godard, D. N. (1980). Self-Recovering Equalization and Carrier Tracking in Two-Dimensional Data Communication Systems. *IEEE Trans.Commun.,* vol. 18, no. 11, (pp. 1867-1875).
23. Bell, A. J., & Sejnowski, T. J. (1995). An information-maximization approach to blind separation and blind deconvolution. *Neural Computation*, vol. 7, no. 6, (pp. 1129-1159).
24. Kim, Y. H., & Shamsunder, H. S. (1998). Adaptive algorithms for channel equalization with soft decision feedback. *IEEE Journal on Selected Areas in Communications,* vol. 16, no. 9, (pp. 1660-1669).