Maximum Entropy Whitening Algorithm with Adaptive Neuron Slope for Blind DFE

Vladimir R. Krstić, Member, IEEE, Bojan R. Dimitrijević, and Miloš V. Bandur

Abstract — The addressed blind decision feedback equalizer (DFE) reverses the classical order of its feed-forward and feedback filters at the beginning of adaptation to achieve the best equalization of the minimum and maximum phase components of a channel transfer function. Although very effective, this blind equalization approach deals with the feedback filter mismatch at the time of its transformation from the front-end all-pole whitenner of the received signal to the decision-directed feedback equalizer placed after the feed-forward filter. To eliminate this weakness, the adaptive neuron slope is introduced instead of the fixed one into the stochastic gradient whitening algorithm based on the joint entropy maximization cost. The performance of the innovated algorithm is verified by simulating m-QAM (quadrature amplitude modulation) signals transmission over multipath channels. The algorithm with the adaptive neuron slope achieves a better whitening of the received signal spectrum, and, hence, increases the equalization successfulness.

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

I. INTRODUCTION

A channel equalizer is the adaptive filter devised to compensate for the intersymbol interference (ISI) effects at the output of a non-ideal channel. It adjusts its own coefficients according to the given optimization rule to estimate the unknown channel inverse transfer function. The classical equalizers estimate a channel with the help of the training preamble known at the receiver side while the unsupervised (blind) equalizers perform the same task relying on the available received signal and the knowledge of few signal statistical features. Making the system free from a preamble, a blind equalizer increases an effective system throughput and facilitates its operation in an environment where it is not practical or possible to use a preamble [1].

A decision feedback equalizer (DFE) combines a linear feed-forward (FFF) filter and a feedback (FBF) filter including the decision device. In comparison with the corresponding linear equalizer (LE), the DFE equalizer eliminates the post-cursor ISI with a smaller number of coefficients and without noise enhancement owing to the recursive nature of its non-linear FBF filter. The superiority of the DFE is evidenced for a wide range of channel transfer characteristics including channels with deep spectral nulls [2] and channels facing long delay spread [3]. Unfortunately, this advantage can be severely degraded by erroneously detected symbols propagating through the FBF equalizer [2]. To eliminate the error propagation effects the classical DFE equalizers periodically readjust adaptive parameters by means of a preamble while the blind equalizers rely on the resourceful cost criteria and algorithms able to eliminate the error propagation certainly emerging at the early beginning of their adaptation process.

To escape the error propagation effects, Labat et al. [4] designed the self-optimized DFE (SO-DFE) scheme optimizing both the structure and the adaptation criterion according to the equalizer convergence state. This scheme is based on the theory that the minimum mean square error (MMSE) infinitely long LE and DFE equalizers share the same components in the frequency domain which can be factored into two equalizers: the all-pole recursive amplitude equalizer and the phase equalizer compensating, respectively, for the minimum and maximum phase components of a channel transfer function [4], [5].

Later in [6] and [7], the SO-DFE scheme was innovated by the soft FBF (SFBF) equalizer which was introduced instead of the original MMSE-FBF filter to improve the estimation of a severe channel spectrum and to mitigate effects of a sudden equalizer structure-criterion switching. In this innovated SO-DFE scheme, named Soft-DFE, the SFBF equalizer exploits the information-theoretic blind deconvolution approach by Bell and Sejnowski [8] and, specifically, the equivalence between the minimization of mutual information between interfering symbols and the maximization of the joint Shannon’s entropy of equalizer outputs. It was verified by simulations that the Soft-DFE, employing the Joint Entropy Maximization (JEM) cost [9], attains remarkably better MSE convergence than the original SO-DFE in the system transmitting m-QAM (quadrature amplitude modulation), m=16,32,64, signals over deep-fade channels [6], [7].

Although very attractive, it has been noticed in [10] and later confirmed in [7] that the practical self-optimized DFE scheme, employing finite length filters, presents some limitations which are manifested as follows. The Soft-DFE deals with the position changing of the SFBF filter during the equalizer convergence process (see Fig. 1). In the blind
mode the $SFBF$ acts as a front-end all-pole amplitude equalizer and, after the equalizer structure-criterion switching, as a decision-directed feedback equalizer placed after the FFF. In the case of higher-order QAM constellations, such as, for example, 64-QAM, the equalizer structure-criterion switching is typically followed by convergence instability which, depending on the channel severity, may lead to an equalization failure. This behavior can be caused by several factors:
- The finite length of the $SFBF$ equalizer results in the coefficients mismatch which violates the assumed equivalence between infinite MMSE-LE length and MMSE-DFE equalizers [10].
- An inadequate ability of the selected neuron function of the $SFBF$ to match the variance increase of ISI probability density function (PDF) of the received signal [8].
- Too large coefficient magnitudes of filters estimating deep channel spectral nulls; the Constant Modulus (CMA) [11] and Joint Entropy Maximization (JEM) [6] algorithms include “power of three” calculations that generate large numbers risking to induce instability or overflow difficulties [12].
- A sudden structure-criterion switching of the equalizer from its blind to soft-transition operation mode.

To mitigate the above limitations of the Soft-DFE scheme, this paper proposes a new variant of the JEM whitening algorithm for the $SFBF$ equalizer. Using the whitening algorithm with an adaptive neuron slope (JEM-AS) instead of a fixed one, the $SFBF$ aims to improve ISI variance matching during the blind mode and, hence, reduce its own coefficients mismatch.

The rest of the paper is organized as follows: Section II describes the Soft-DFE equalizer operation through its three operation modes and the entropy-based algorithms of the $SFBF$ equalizer model. The proposed JEM whitening algorithm with an adaptive slope is considered in Section III. The simulation results presenting the influence of the JEM-AS on the equalizer performances are given in Section IV.

II. SOFT-DFE DESCRIPTION

A. Soft-DFE structure-criterion transformation

The Soft-DFE equalizer, given through its operation modes, 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 $x(t_i + nT/2) = x_n, i = 1, 2, \ldots$ 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 $b_n = [b_{1,n}, \ldots, b_{N,n}]$ and $c_n = [c_{1,n}, \ldots, c_{N,n}]$. The Soft-DFE operates passing 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) including 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 whitener, which is obtained by dropping the decision device in the $SFBF$, is adapted by the stochastic gradient algorithm based on the JEM cost. Operating independently of the WT, the linear equalizer TE compensates for a phase distortion (introduced by a channel+whitener combination) using the CMA-2 algorithm [11]. In the next operation mode, namely 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 active from the early beginning of equalization, and it is implemented by the classical phase locked loop [1] including some modifications [7].

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 $\text{MSE} \leq \text{TL3}$ the equalization is decided to be successful and for $\text{MSE} > \text{TL3}$ unsuccessful.

Fig. 2  Soft feedback filter: basic model.
B. SFBF equalizer: background and algorithms

The SFBF equalizer acts as a neuron-unit maximizing the joint Shannon’s entropy (JEM criterion) [9]

\[ J_{\mu}(b_n) = E \left[ \log \left( \frac{d^2(z_n)}{d\xi^2} \right) \right] \]  

(1)

assuming the basic system model presented in Fig. 2:

- 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 \( g() \) is a nonlinear monotone saturating function tending to transform the unknown PDF function of the input data into the uniform PDF in the finite range of output data symbols \( r_n = g(z_n) \).
- The previously detected symbols at the output of the SFBF are correct, i.e. \( r_{n-j} = a_{n-j}, \ j = 1, ..., N \).

Generally, the selection of neuron function in the complex domain, suitable for maximizing JEM criterion (1) through a stochastic gradient searching of complex-valued data, is a specific issue that is not particularly addressed by this paper. The interested reader should refer to [13,14] for more information. In our case of the SFBF equalizer, the neuron is defined by the complex-valued parametric function given by

\[ g(z_n, \beta) = z_n \left( 1 + \beta |z_n|^2 \right) \]  

(2)

where the parameter \( \beta \) represents the neuron’s slope. The slope \( \beta \) matches the neuron’s slope to the input PDF function in accordance with the maximum entropy deconvolution model mapping the input data corrupted by ISI into mutually independent symbols [8].

The basic stochastic (ascent) gradient JEM algorithm [14] has two heuristic variants optimizing the cost (1) through two operation modes according to the SFBF filter transformation from the decision-directed SFBF equalizer to the all-pole recursive whitener (and reverse). The JEM-D algorithm for the decision-directed SFBF is given by

\[ b_{n+1,i,j} = b_{n,i,j} - \beta_p z_n \left( 1 - \beta_p \right) z_{n-j}^* \beta_{n,i,j} \]  

(3)

and the JEM-W algorithm for the WT whitener is

\[ b_{n+1,i,j} = (1-\gamma) b_{n,i,j} - \mu_p u_n \left( 1 - \beta_p \right) u_{n-j}^* \beta_{n,i,j} \]  

(4)

where \((\mu_p, \beta_p)\) and \((\beta_w, \beta_D)\) are, respectively, the corresponding adaptation steps and neuron slopes, and the parameter \( \gamma \) the leaky factor.

It should be noted that the whitening algorithm JEM-W in (3) is the variant of the JEM-D algorithm obtained by dropping the decision device in SFBF equalizer, i.e., it is a result of linearization of the SFBF. Two algorithms share the same JEM-type nonlinear gradient term whose performances strongly depend on the slope \((\beta_w, \beta_D)\) selection.

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 that the previous decisions are correct. The leaky factor in (3) is an option which can be zero or a small positive number. It prevents the overgrowth of the whitener’s vector (Euclidian) norm \( \|b_n\| \) in the case of higher-order QAM signal constellations [7]. Acting oppositely to the entropic term \( \beta_w \|u_n\|^2 \) the coefficient leaky term decreases the risk of the equalizer instability at the critical time of its structure-criterion switching. As well known, the coefficient leaky is a 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 [15, and references therein].

The slope selection is an essential issue of the JEM algorithm implementation. The JEM-D algorithm is directed by detected symbols \( \hat{a}_n \) whose PDF is determined by the given signal constellation. This situation simplifies the selection of the slope \( \beta_D \). In [6] and [7] it is evidenced that the optimal value of the slope \( \beta_D \) depends only on the given signal constellation, and the optimal slope \( \beta_D \) can be seen as a statistical constant of the given signal and neuron function. On the other hand, the selection of the slope \( \beta_w \) is a more complex problem. The JEM-W algorithm is directed by the whitener outputs which are not ISI free because the whitener deals only with the second-order statistics of the received signal.

Thus, to minimize the mismatch between the WT coefficient vector reached during the blind mode and the coefficient vector expected by the SFBF at the beginning of the soft transition mode, we introduced an adaptive neuron slope \( \beta_{n,W} \) into the JEM-W algorithm supposing that the higher quality of power spectra estimation leads to the further mitigation of coefficient mismatch effects.

III. JEM-W ALGORITHM WITH ADAPTIVE SLOPE

The JEM-W algorithm with the adaptive neuron slope, referred to further as JEM-AS, is given by

\[ b_{n+1,i,j} = (1-\gamma) b_{n,i,j} - \mu_p u_n \left( 1 - \beta_p \right) u_{n-j}^* \beta_{n,i,j} \]  

(5)

where the term

\[ e_n = u_n \left( 1 - \beta \right) u_n^* \]  

(6)

represents an \( a \ priori \) prediction error and the whitener current output is given by

\[ u_n = x_n - b_n \]  

(7)

where \( \gamma b_n \) that here has two roles. First, as mentioned before, the coefficient leakage prevents an unconstrained growth of whitener coefficients. Second, we use its regularization capability to get a measure of the slope adaptation quality. Namely, it is used to decide whether to increase or decrease the slope \( \beta_n \) through its adaptation process.

The adaptation rule of the slope \( \beta_n \) is based on the \( a \ posteriori \) prediction errors and a greedy punish/reward heuristic [15]. It means that the whitener, at the time \( t = nT \)
but before the next updated input $x_{n+1}$, calculates the coefficient vector $\mathbf{b}_{n+1}^\text{IS}$ with ($\gamma > 0$) and $\mathbf{b}_{n+1}$ without ($\gamma = 0$) leaky and then the corresponding a posteriori outputs ($\mathbf{u}_{n}^\text{IS}$, $\mathbf{e}_{n}$) and errors ($\mathbf{e}_{n}^\text{IS}$, $\mathbf{\hat{e}}_{n}$). Finally, using the difference between modules of a posteriori errors the whitener decides whether to increase or decrease the slope $\beta_n$ according to the rule: if $|e_n^\text{IS}| < \bar{e}_n$ the slope decreases and if $|e_n^\text{IS}| > \bar{e}_n$ the slope increases. The pseudo-code, implementing the adaptation rule, is given by the following “if-else” relation

\[
\text{if } |e_n^\text{IS}| < \bar{e}_n \text{ then }
\begin{align*}
m_{n+1} &= \max(m_n - l_d, 0) \\
\text{else}
\end{align*}
\text{set } m_{n+1} = \min(m_n + l_u, M)
\text{endif (8)}
\]

and the quantized function

\[
\beta_{n+1}^\text{w} = f(m_{n+1}) = \beta_{n}^\text{w} - \alpha(m_{n+1} / M) \quad (9)
\]

which calculates the next slope $\beta_{n+1}$ using an independent variable $m_n = 0, \ldots, M$ and user-definable parameters $(M, l_d, l_u) \in \mathbb{Z}$, $(\beta_{n}^\text{w}, \alpha) \in \mathbb{R}$, $m_0$. As can be seen in (9), the slope $\beta_{n+1}$ varies in the range from $(\beta_{n}^\text{w} - \alpha)$ to $\beta_{n}^\text{w}$. For the selected set of parameters $\{\beta_{n}^\text{w} = 1.6, \alpha = 1.2, l_{d} = 5, l_{u} = 30, M = 1000, m_0 = 100\}$, Fig. 3 illustrates the neuron slope learning of the statistics of the 64-QAM signal convolved by channels Mp-(A,C,E); the Mp channel characteristics are described in the next section (see Fig.4).

The introduction of an adaptive part into the neuron slope in (9) is resulting in the complexity increase of the JEM-AS algorithm. It is the result of additional calculations of a posteriori quantities $\{u_{n}^\text{IS}, \mathbf{u}_{n}, \mathbf{e}_{n}, \mathbf{\hat{e}}_{n}, \mathbf{b}_{n}^\text{IS}, \mathbf{b}_{n}\}$ as well as the “if-else” decision rule (8) and the quantization function (9). Although relatively simple, these calculations should be taken into consideration of the entire benefits of the innovated whitener.

**IV. SIMULATION RESULTS**

The efficiency of the JEM-AS algorithm is considered through its influence on the effective equalizer performances given in terms of the MSE convergence and the equalization success index (ESI); the ESI is defined as the ratio between the number of successful equalizations and the total number of Monte Carlo runs. The JEM-AS is compared to the original JEM-W algorithm given in (3) aiming to determine the increase or decrease of the ESI index which is closely related to the presence of the error propagation effects. The performance testing environment is the system simulator including the m-QAM, m={16,64}, signal transmitter, the time-invariant multipath channel adding the white Gaussian noise and the Soft-DFE at the receiver side.

Fig. 4. Attenuation of multipath channels Mp-(A,C,E).
The MSE and ESI characteristics, presented in Figures 5-7, are obtained, respectively, for 200 and 10000 independent runs initiated with a random carrier phase taken in the range $(0, 2\pi)$. The fixed slope $\beta_w$ in recursions (3) and (9) is varied in the range $\beta_{W,16} = (4.0, 7.0)$ and $\beta_{W,64} = (1.0, 2.2)$ radians. Fig. 5a presents the influence of the adaptive slope on the MSE convergence for the 16-AM signal for two values of the slope $\beta_{W,16} = (5.0, 7.0)$ and Fig. 5b compares MSE convergence for JEM-AS and JEM-W algorithms using the same slope $\beta_{W,16} = 7.0$. The same tests are repeated with the 64-QAM signal using a different set of slopes $\beta_{W,64} = (1.6, 1.8)$, Fig 6. As can be seen in Figures 5a and 6a, the fixed part of the adaptive slope determined by equation (9) has the ability to change the MSE convergence rate with neglecting effects on the residual MSE in the steady-state. In other words, the adding of the adaptive slope does not degrade the effective MSE convergence obtained by the original JEM-W algorithm supposing the ESI index is 100%. It is verified by the results in Figures 5b and 6b that compare the MSE convergence achieved by the JEM-W and JEM-AS algorithms using the same fixed slope $\beta_{W,16} = 7.0$ and $\beta_{W,64} = 1.6$.

Fig. 5. MSE, 16-QAM, Mp-(A,C,E): a) JEM-AS, $\beta_w = (5.0, 7.0)$, b) JEM-W and JEM-AS for $\beta_w = 7.0$.

Fig. 6. MSE, 64-QAM, Mp-(A,C,E): a) JEM-AS, $\beta_w = (1.6, 1.8)$, b) JEM-W and JEM-AS for $\beta_w = 1.6$.

V. CONCLUSION

The paper proposes the JEM whitening algorithm with an adaptive neuron slope to reduce the feedback filter mismatch effects present in the Soft-DFE using the whitening algorithm with a fixed slope. Improving the received signal whitening, the algorithm with an adaptive slope effectively softens the equalizer structure-criterion switching from the blind to the soft operation mode. The efficiency of the innovated algorithm is evidenced by the increase of the equalization successfulness, and it is particularly evidenced for the 64-QAM signal and severe channels. At the same time, the MSE convergence characteristics of the equalizer are preserved practically unchanged with respect to the characteristics obtained by the JEM whitening algorithm using the fixed slope.
Fig. 7. ESI for JEM-W and JEM-AS: a) 16-QAM, Mp-C, b) 64-QAM, Mp-C, c) 64-QAM, Mp-E.

REFERENCES