Abstract – This paper considers the blind decision feedback equalizer for quadrature amplitude modulation (QAM) systems and proposes the optimization method of parametric stochastic gradient algorithms particularly derived for its entropy-based feedback filter. The optimization method is based on the fact that the “slope” of applied parametric complex-valued nonlinearity can be adjusted to the unknown probability distribution of an input intersymbol interference to respond with a maximum entropy output. The parametric optimization is achieved for 16-, 32- and 64-QAM signals.

Keywords – Blind equalization, decision feedback equalizer, joint entropy criterion, soft feedback filter.

I. INTRODUCTION

The major drawback of the blind decision feedback equalization is the error propagation phenomenon emerging at the beginning of the detection process. To eliminate this difficulty, the blind decision feedback equalizer addressed in this paper (Soft-DFE) combines both the structure decomposition method proposed by Labat et al. [1] and the “soft” feedback filter (SFBF) [2], [3] based on the blind deconvolution theory approach by Bell-Sejnowski [4]. According to this theory, the SFBF acts as a single neuron unit which reduces the post cursor intersymbol interference (ISI) by the joint entropy maximization (JEM) criterion.

This paper addresses the optimization method for the parametric stochastic gradient algorithms which adjust SFBF’s coefficients during the initial phase of Soft-DFE convergence. For that purpose the neuron slope is varied by free parameters to respond with a maximum entropy signal with a motivation to implement the computationally efficient solution of blind DFE for 16-, 32- and 64-QAM signals. The alternative to this parametric neuron approach is the one which is based on the adaptive mapping functions which generally leads to solutions burden with a higher degree of complexity [5].

II. SOFT FEEDBACK FILTER

The basic model of SFBF is presented in Fig. 1. The data $a_n$ applied to a linear time-invariant noiseless channel is a sequence of zero-mean i.i.d. variables with a sub-Gaussian distribution and the neuron input $z_n = x_n + \mathbf{b}_n^T \mathbf{r}_n$ is a sum of channel output $x_n$ and a weighted sum of neuron outputs where the coefficients $b_j$ and outputs $r_{n-j}$, $j = 1, ..., N$, are elements of vector $\mathbf{b}_n = [b_1, ..., b_N]^T$ and $\mathbf{r}_n = [r_1, ..., r_N]^T$, respectively. According to the Bell-Sejnowski theory, the neuron $g(z)$ is a monotone increasing sigmoidal function the form of which approximates the expected cumulative distribution of inputs according to the relation $g(z) \approx \int p(u)du$ where $p(z)$ is the probability density function (pdf) of input $z$. Unfortunately, since the pdf of ISI, and hence of $z$, is generally unknown and also there is a lack of appropriate (sigmoidal) nonlinearities, it is practical to use a parametric nonlinearity $g(z, \beta)$ where the parameter $\beta$ varies the neurons slope in a way to be as close as possible to the probability distribution of ISI.

For the complex-valued SFBF model described by the nonlinearity $g(z_n, \beta) = z_n \left(1 + \beta |z_n|^2\right)$, the stochastic gradient (ascent) algorithm has been derived in [3]

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

which optimizes the JEM criterion $J_E(\mathbf{b}) = E \left[ \ln \left| \frac{\partial r_n}{\partial z_n} \right| \right]$ under the hypothesis of correctly detected symbols $r_{n-j} = a_{n-j}$, $j = 0, ..., N$, where $\mu$ is a step size and $\beta$ is a real parameter. According to this approach, the SFBF manipulates the pdf of outputs toward its uniformity using only on-line data.

III. SOFT-DFE: STRUCTURE AND ALGORITHMS

The basic model of SFBF is heuristically modified to be implemented in Soft-DFE scheme optimizing both the structure and the criterion through three operation modes: blind acquisition, soft transition and tracking. During the blind acquisition, the SFBF is transformed into a linear all-pole filter (whitener) and placed at the front end of the Soft-DFE.
to perform non-flat channel spectrum equalization, Fig. 2a. In this phase, the Soft-DFE operates as a linear \( T/2 \) fractionally-spaced equalizer \( (T \) is the symbol interval) including four signal processing circuits ordered in cascade with increasing complexity - gain control \( (G) \), whitener \( (W) \), fractionally-spaced equalizer \( (T) \) and phase-locked loop \( (P) \) - where \( W \) and \( T \) perform the most critical tasks.

The whitener \( W \), which includes two purely recursive filters with coefficient vectors \( b_i = [b_{i,1}, ..., b_{i,N}]^T \), \( i = 1, 2 \), is adjusted by the JEM-W recursion

\[
b_{i,n+1,j} = (1 - \gamma W) b_{i,n,j} - \mu_W u_{i,n} \left( 1 - \beta W \right) |u_{i,n}|^2 u_{i,n-j}^* \quad (2)
\]

where \( \mu_W \) is step-size, \( \beta_W \) is slope factor and \( \gamma_W \) is the leakage factor which determines the coefficients leakage rate. The leakage term in JEM-W is applied to restrict an unconstrained growth of whitener’s coefficients. Let us remind that the coefficient leakage method is the standard practice in fractionally-spaced equalization [6].

At the same time and independently of \( W \), the \( T \), which is defined by coefficient vectors \( c_i = [c_{i,0}, ..., c_{i,L-1}]^T \), compensates for a phase distortion introduced by a channel-whitener combination using the leakage variant of Godard’s constant modulus algorithm (CMA) [7] given by

\[
c_{n+1,k} = (1 - \gamma_G) c_{n,k} - \mu_G y_n \left| y_n \right|^2 - R c_{n-k}^* \quad (3)
\]

where \( \mu_G \) is step-size, \( \gamma_G \) is leakage factor and \( R \) is signal dependent statistical constant. According to the previous experience, the leakage terms in (2) and (3) for 16- and 32-QAM signal packets of errors for an optimally selected slope \( \beta_D \). The efficiency of SFBF is measured in two ways depending on the signal complexity. For the middle dense 16- and 32-QAM constellations, it is practical to measure the symbol error rate (SER) versus \( \beta_D \). On the other hand, for the higher dense 64-QAM, it is more convenient to measure the convergence time between thresholds MSE-TL1 and MSE-TL2 versus \( \beta_D \).

**A. Optimal \( \{b_{u}, b_{D}\} \) for 16- and 32-QAM**

In the following text the experimental data obtained via simulations are analyzed to select the optimal values of \( \{g_u, b_u, b_D\} \) parameters for 16-, 32- and 64-QAM signals.

Fig. 3 presents the kurtosis of symbols \( y_n \) originating from a 16-QAM signal which is measured at the end of blind mode for different values of \( \beta_u \) and multipath channels \( M_p \) (see Fig. 7). The corresponding curves for 32-QAM closely follow the curves for 16-QAM and are not presented here. Using these kurtosis curves as well as the corresponding MSE convergence characteristics of Soft-DFE, the optimal values of \( \beta_u \) are decided in the relatively wide range of values.
Fig. 3. Kurtosis versus $W$.

Fig. 4. Symbol error rate versus $D$ for 16- and 32-QAM.

The selection of JEM parameters for 64-QAM signal is a little more complex than for 16- and 32-QAM because it includes three parameters $\{g_W, b_W, b_D\}$. However, this complexity can be relaxed taking into consideration the fact that parameters $\{g_W, b_W\}$ and $b_D$ can be selected independently of each other in a similar manner as it is done in the case of 16- and 32-QAM signals. Fig. 5 presents kurtosis curves versus $\beta_W$ obtained for a suitably selected set of leakage factors $\{0.0, 2^{-14}, 2^{-13}, 2^{-12}\}$ for the worst case channel $\text{Mp-E}$. As can be seen, by increasing leakage from $\gamma_W = 0$ (corresponding to the basic variant of JEM-W) to $\gamma_W = 2^{-12}$, the kurtosis statistics has been improved in the sense that the strong saturation characterizing the curve $\gamma_W = 0$ disappears by increasing leakage so that the curve $\gamma = 2^{-12}$ becomes similar to the curves obtained for 16- and 32-QAM signals (see Fig. 3). It has been the motivation to estimate that the maximal kurtosis $Kur = 0.6625$ for curve $\gamma_W = 2^{-12}$ is achieved for $\beta_W = 1$, and then to decide other $\{g_W, b_W\}$ pairs which also optimize JEM-W to affect the same kurtosis value. Thus, in the case of the leakage variant of JEM-W, for each given leakage factor $\gamma_W$ the corresponding maximal slope $\beta_W$ can be selected which guarantees a successful equalization. According to this method the following $\{g_W, b_W\}$ pairs are obtained: $\{0.0, 0.1\}$, $\{2^{-14}, 0.4\}$, $\{2^{-13}, 0.6\}$, $\{2^{-12}, 1.0\}$. It should be stressed that in the case of 64-QAM only the last two pairs are of practical interest because they force an approximately linear kurtosis increase which ensures a good compromise between effective convergence speed and successful equalization.

To select the slope $b_D$ for JEM-D, the MSE transition time (MSE-TT) from MSE-TL1 to MSE-TL2 is observed for the motivating set of $b_D$ in the range from 0.5 to 4.0. Fig. 6a presents the MSE-TT in symbol intervals versus $b_D$ for $\text{Mp}$-channels. The smooth hyperbolic-like MSE-TT curves show that their unique minima converge into a relatively narrow range of $\beta_D$ from 1.75 to 2.25 indicating that the positions of their minima are practically independent of channel. It can be also shown that the optimal value of $\gamma_W$ is practically independent of $W$.

Based on the above simulation results, it is shown that the optimality of JEM-W and JEM-D can be achieved by their corresponding parameters. In contrast to the optimal values of $\beta_D$, the above maximal values of $\beta_W$ are additionally fitted to achieve the best compromise between convergence rate and equalization successfulness. Hence, the optimal pairs of parameters are decided approximately as follows $\{b_W, b_D\} = \{1.3, 1.0, 16, 12\}$, $\{0.0, 1.0, 32, 10\}$ and $\{64, 0.5\}$.

\begin{align*}
\{g_W, b_W, b_D\} & \text{ for 64-QAM} \\
\{0.8, 2, 0\} & \text{ and } \{0.8, 1, 4\} \text{ for 16-QAM and 32-QAM, respectively. For minimal values of } \beta_W \text{ the algorithm JEM-W loses ability to cope with deep spectral nulls of the received signal, and for maximal ones the switching robustness of Soft-DFE to transform itself from blind to tracking operation mode is being degraded. The last is typically characterized by the increased number of equalization failures due to error propagation.} \\
\text{The SER curves versus } \beta_D \text{ shown in Fig. 4 indicate the ability of JEM-D to minimize SER caused by error propagation. Independently of channels, these curves present minima for the same optimal value of } \beta_D \text{ that can be estimated as } \beta_{D,16} = 12 \text{ and } \beta_{D,32} = 10 \text{ for 16- and 32-QAM signals, respectively. Besides, these optimal values of } \beta_D \text{ are independent of } W, \text{ which can be seen as a part of channel.} \\
\end{align*}
V. SYSTEM SIMULATOR

The simulations are carried out for the time-invariant frequency-selective fading channels. Fig. 7 shows the attenuation characteristics of a three-ray channel model involved into the transmitter filter of system. The length of equalizer is $L=22$ and $N=6$ for $(16,32)$-QAM, and $L=25$ and $N=5$ for $64$-QAM in its $T$ and $W$ parts, respectively. The initial vectors of $T$ are with zero components except the centred (reference) ones $c_{1,r} = c_{2,r} = 1$. For $(16, 32)$-QAM, the Soft-DFE switches from blind to soft transition mode at the threshold level $\text{MSE-TL1}=1.5$ dB and then the soft transition continues during the next 2000 $T$ intervals, and for $64$-QAM the threshold levels are selected to be $\text{MSE-TL1}=7.9$ dB and $\text{M-TL2}=-2.2$ dB. The simulations are carried out under a signal-to-noise ratio of 25 and 30 dB for $(16, 32)$-QAM and $64$-QAM, respectively.

VI. CONCLUSIONS

In this paper the optimization method for the parametric recursive part of the blind Soft-DFE is presented. It is proved via simulation that the parameters of the selected complex-valued nonlinearity can be optimally adjusted for the given signal in the system with a large scale of severe ISI channels. The efficiency of the presented method is verified with 16-, 32- and $64$-QAM signals.

ACKNOWLEDGEMENT

This paper was supported by Ministry of Education, Science and Technological Development of the Republic of Serbia, Project No. TR32037 - “The robust transmission systems for corporate networks”. The authors would like to thank to anonymous reviewers for their helpful and constructive comments.

REFERENCES