
Grzegorz Haza, Ryszard Makowski

Signal Theory Section, Wroclaw University of Technology
Wybrzeze Wyspiańskiego 27, 50-370 Wroclaw, Poland
email: grzegorz.haza@pwr.wroc.pl, ryszard.makowski@pwr.wroc.pl

1. THE ALGORITHM DESCRIPTION

The Multi Modulus Algorithm (MMA) for blind equalization has been proposed in [1]. It is based on the popular and widely used algorithm presented by Godard in [2]. The MMA cost function penalizes both, real and imaginary components of the output signal and thus it is capable to remove a residual phase offset. However, when the MMA is used for the high order QAM, the mean squared error in the steady state at the equalizer output may be too high and, as a consequence, the scatter diagram eye will not become opened. Joint blind equalization techniques can be used to avoid this problem. At the beginning, the blind equalization technique, which can cope with high level of the input signal mean squared error induced by the channel impulse response, is used. If the MSE is reduced (for a particular constellation size) to an acceptable level, another equalization technique which guarantees lower steady state MSE at the same step size is used in order to proceed further corrections. The joint blind equalization with the Constant Modulus Algorithm (CMA) at the first stage, and the Decision Directed (DD) algorithm at the second stage is shown in [3].

The Shell Partitioned Constant Modulus Algorithm (ShCMA), presented in [4], is the blind equalization technique, which can be used at the second stage instead of the DD algorithm. This solution is based on the observation that M-QAM output power can be classified into a few sections where each of them has its own radius assigned, instead of one global radius used in the conventional CMA. On the contrary to the DD algorithm, the ShCMA does not require any phase offset correction before the equalization is performed. In [4] it is shown, that the results of the convergence for the CMA+ShCMA joint blind equalizer are comparable to the CMA-DD for the constellation size up to the 64-QAM.

In this paper, the Shell Partitioned Multi Modulus Algorithm (ShMMA) is presented, and then used at the second stage of equalization instead of the DD or ShCMA. The equalizer processes a signal, which is assumed to be sampled coherently to the symbol timing clock. The signal is described as \( x(n) = \sum b(n-m)h(n-m) \cdot e^{j\phi(n)} + v(n) \), where \( b \) is a symbol chosen from the complex alphabet \( a = \{a_1, \ldots, a_M\} \) of the M-QAM, \( h \) is an impulse response of the non-minimum phase communication channel, without zeros on the unit circle in the Z-domain; the phase offset \( \phi \) is assumed to be constant in time; \( v \) is an additive white Gaussian noise.

1.1 Multi-Modulus Algorithm

The cost function for the Multi-Modulus Algorithm is described by the following equation:

\[
\Psi_1(y(n)) = E\{[\hat{y}_R(n) - R^2_k] + [\hat{y}_I(n) - R^2_k]^2\} \tag{1}
\]

Where \( y_R(n), y_I(n) \) are the real and imaginary part of the equalizer output sample and \( R^2_k, R^2_j \) represent radii for the real and imaginary part, respectively:

\[
R^2_k = \frac{E\{a_k^2\}}{E\{a_k^2\}} \quad R^2_j = \frac{E\{a_j^2\}}{E\{a_j^2\}} \tag{2}
\]

The stochastic gradient, based on the equation (1) is described by:

\[
\Psi_1(y(n)) = y_R(n)[\hat{y}_R(n) - R^2_k] + y_I(n)[\hat{y}_I(n) - R^2_k] \tag{3}
\]

1.2 The Shell Partitioned Multi Modulus Algorithm

If the mean squared error is reduced to the level, at which the eye of the scatter diagram opens, \( y_R(n) \) and \( y_I(n) \) have Gaussian distribution [5] with the mean value \( \mu_{R,j,m} \) over \( m \)th constellation point \( a_m = a_{R,m} + ja_{I,m} \). If \( y_R(n) \) and \( y_I(n) \) are i.i.d., they can be treated as two \( \sqrt{M} \)PAM. For the MMA cost function (1), only the absolute values \( |y_R(n)| \) and \( |y_I(n)| \) are used in the process of detection. Because of this, subsets \( \{G_{R,k}\}_{k=1}^M \) and \( \{G_{I,l}\}_{l=1}^M \) with related decision boundaries \( \Delta_{G_{R,k}} \) and \( \Delta_{G_{I,l}} \) respectively are applied. Classifying each constellation point \( a_R \) and \( a_I \) to the particular \( G_{R,k} \) and \( G_{I,l} \), the following set of radii can be computed:

\[
R^2_{R,k} = E\{[a_R \in G_{R,k}]^2\} \quad R^2_{I,j} = E\{[a_I \in G_{I,j}]^2\} \tag{4}
\]

Each radius \( R^2_{R,k} \) and \( R^2_{I,j} \) is related to the subset \( S_{R,k} \) and \( S_{I,k} \). To determine decision boundaries \( < s_{R,k-1}, s_{R,k} \) and \( < s_{I,k-1}, s_{I,k} \) for the subsets \( \{S_{R,k}\}_{k=1}^M \) and \( \{S_{I,l}\}_{l=1}^M \), the ML criterion is used [8]. When the eye of the scatter diagram is opened, the random variables \( \hat{y}_R(n) \) and \( \hat{y}_I(n) \) in the MMA cost function (1) are described by the non-central Chi Squared PDF \( \chi^2(1) \) [6],[7]. The exemplary probability density functions \( p(\hat{y}_R|S_{R,k})P(S_{R,k}) \) and \( p(\hat{y}_I|S_{I,k})P(S_{I,k}) \), used to compute the decision boundaries for the shell \( S_{R,k} \) for \( k = 6 \) for the 256-QAM are presented in Figure 1.

The cost function for the MMA with the shell partitioning is described by the following equation:

\[
\Psi_2(y(n)) = E\{[\hat{y}_R(n) - R^2_{R,k}]^2 + [\hat{y}_I(n) - R^2_{I,j}]^2\} \tag{5}
\]
The stochastic gradient based on (5) is described by the equation:

\[ J^\text{MIN}(n) = \lambda \text{MSE}(n) + (1 - \lambda \text{MSE}) |\hat{b}(n) - y(n)|^2 \]

Applying the adaptive estimation technique to (7), with an exponentially decaying window configured by the forgetting factor \( \lambda M \), we get:

\[ J(n) = \lambda M J(n - 1) + (1 - \lambda M) |\hat{b}(n) - y(n)|^2 \]

However, if the distortion introduced by the channel and an additional Gaussian noise closes the eye of the scatter diagram, the technique of the MSE estimation (8) fails. In this situation, the PDFs of the adjacent constellation points overlap each other, and degradation of the MSE estimation, according to the true MSE value, appears. This can thwart the proper detection of switching and, in fact, the equalization procedure itself.

When \( y \) is a random variable with Gaussian PDF \( p(y) \) for each constellation point with the variance \( \sigma^2 \), the entire constellation is expressed by the joint PDF \( p(y) = \sum_{i=1}^{M} p(y_i) \) assuming that each symbol is equally probable, i.e., \( p(y_i) = \frac{1}{M} \). Each constellation point \( y_i \) has its own subset \( D_i \) and the variance within it is \( \sigma^2_i = \text{E}[(y_i - \mu_i)^2] = J_i \). Let \( \sigma^2_0 = \frac{1}{M} \sum_{i=1}^{M} \sigma^2_i \) be a measure of the average \( \sigma^2 \) among all the subsets. The measure \( \sigma^2_0 = f(\sigma^2) \), calculated for all constellation points, is drawn by the solid line in Figure 2. It can be seen, that \( \sigma^2_0 \) strongly deviates from the true value \( \sigma^2 \) when the border (depending on the distance between the adjacent constellation points) is crossed.

This drawback can be omitted by the use of only the edge subsets \( E = \{D_1, D_M\} \) of the constellation in purpose, to estimate the Edge MSE:

\[ J_E(n) = \text{E}(|\hat{b}(n) - y(n)|^2 | y(n) \in E) \]

Results of the \( \sigma^2_0 \) calculation only for the edge subsets are presented in Figure 2 by the dashed lines. It can be seen that, using only the edge subsets, the variance \( \sigma^2_0 \) is closer to the true variance \( \sigma^2 \) compared to the case, when all subsets were used.

Thus, the MSE adaptive estimation technique is based on the edges of the probability density function is calculated according to:

\[ J_E(n) = \begin{cases} \lambda_E J_E(n - 1) + (1 - \lambda_E) |\hat{b}(n) - y(n)|^2, & y(n) \in E \\ J_E(n - 1), & y(n) \notin E \end{cases} \]

### 2. Switching Procedure

#### 2.1 Mean Squared Error Estimation

The presented joint blind equalizer has to employ some measure of the output error, which can be applied in order to switch from the MMA to the ShMMA. This can be done by using the MSE of the equalized samples \( y(n) \) with respect to the detected data \( \hat{b}(n) \):

\[ J(n) = \text{E}[(\hat{b}(n) - y(n))^2] \]

Applying the adaptive estimation technique to (7), with an exponentially decaying window configured by the forgetting factor \( \lambda M \), we get:

\[ J(n) = \lambda M J(n - 1) + (1 - \lambda M) |\hat{b}(n) - y(n)|^2 \]

### 2.2 Switching procedure

Usually in joint blind equalizers, hard switching procedure is used. In the proposed equalizer, a soft technique, based on the EMSE estimation (10), is applied. It is obtained by the following switching function:

\[ \gamma(n) = T \left( -\log J_E(n) + \log J_{\text{start}} \right) / \log J_{\text{start}} - \log J_{\text{stop}} \]

Where \( J_{\text{start}} \) and \( J_{\text{stop}} \) are switching start and stop limits, respectively, while \( T(x) \) is the following saturation function:

\[ T(n) = \begin{cases} 0, & x \leq 0 \\ x, & 0 < x \leq 1 \\ 1, & x > 1 \end{cases} \]

By using the switching function (11) for (3) and (6), a new gradient can be created and described by the given equation:

\[ \Psi(n) = (1 - \gamma(n)) \Psi_1 + \gamma(n) \Psi_2(n) \]

From (13) it follows that, when the EMSE estimation value is smaller than the \( J_{\text{start}} \), only the MMA equalizer is used. When this border is crossed, the MMA participation in the process of equalization is decreasing, whereas the ShMMA participation is rising until the EMSE estimator reaches \( J_{\text{stop}} \). Then the ShMMA takes the entire control of the equalizer and the MMA remains disabled. This equation is applied for the procedure of the filter coefficients update (24).

### 3. THE J-ORTHOGONAL FILTER

When an uncorrelated data sequence is transmitted over the communication channel with a given impulse response, the signal becomes correlated. In gradient-based techniques of seeking the settings for the equalizer coefficients, that minimize some cost function, the rate of convergence strongly depends on the eigen value spread \( \lambda_{\text{MAX}} / \lambda_{\text{MIN}} \) of the received signal correlation matrix (9).
Thus, if the channel imposes high amplitude distortion, poor performance of the gradient-based algorithms can be expected. In order to avoid this problem, the Gram–Schmidt orthogonalization procedure can be used. In literature, a lot of attention is paid to the adaptive lattice algorithms and their orthogonalizing properties. Considering the forward and backward linear prediction errors, and computing the partial correlation (PARCOR) coefficients, an orthogonal basis can be calculated [10],[11],[12].

In [10], the gradient lattice predictor is used to calculate an uncorrelated basis for the adaptive channel equalizer and evaluate its superior performance, in comparison with the conventional tapped-delay filtering. In [11], the application of the least squares lattice predictor algorithm in adaptive filtering, derived from the arithmetic approach is described. In [12], the normalized recursive least squares ladder-filter, derived geometrically, is presented. It is also pointed, that the normalized algorithm [12] has better performance, comparing to its unnormalized version [11].

The orthogonalization procedure with the use of the adaptive lattice predictor is also investigated in blind channel equalization [13]. In this paper, the technique presented in [12] is used to perform the orthogonalization procedure, and the gradient-based technique [10] is applied for updating the coefficients values.

3.1 Adaptive orthogonalization of the basis

The structure of the lattice adaptive blind channel equalization is presented in Figure 3. According to [12], if \( n = 0 \) the filter initializations are given by:

\[
R(0) = \sigma + |x(0)|^2
\]

\[
e_0(0) = r_0(0) = x(0)R(0)^{-\frac{1}{2}}
\]

where \( \sigma \) is a small positive value which prevents from (a possible) division by zero while \( e_0(0) \) and \( r_0(0) \) are the forward and backward predictor errors initializations.

In further steps, i.e., for \( n = 1, 2, \ldots \), the filter initializations are:

\[
R(n) = \lambda R(n-1) + |x(n)|^2
\]

\[
e_n(n) = r_n(n) = x(n)R(n)^{-\frac{1}{2}}
\]

where \( \lambda \) is the so-called forgetting factor, associated with exponential windowing of the filter input time-series samples.

The filter algorithm can be summarized as follows: Let

\[
E_m(n) = (1 - |e_m(n)|^2)^{\frac{1}{2}}
\]

\[
R_m(n) = (1 - |r_m(n)|^2)^{\frac{1}{2}}
\]

\[
N_m(n) = (1 - |\rho_m(n)|^2)^{-\frac{1}{2}}
\]

Then the PARCOR coefficients values, for each lattice-section \( \theta_m(n) \) for \( m = 1, \ldots, N \) are adaptively updated as follows:

\[
\rho_{m+1}(n) = \rho_{m+1}(n-1)E_{m}(n)R_{m}(n-1) + e_{m}(n)r_{m}(n-1)
\]

(21)

The forward and backward predictor error samples are computed with the following recurrence relations:

\[
e_{m+1}(n) = \frac{e_{m}(n) - \rho_{m+1}(n)r_{m}(n-1)}{R_{m}(n-1)N_{m+1}(n)}
\]

(22)

\[
r_{m+1}(n) = \frac{r_{m}(n-1) - \rho_{m+1}(n)e_{m}(n)}{E_{m}(n)N_{m+1}(n)^{\frac{1}{2}}}
\]

(23)

The forgetting factor should be set close to 1 (e.g. \( \lambda = 0.9999 \)) in aim to keep small variance of estimated PARCOR coefficients. However, if the channel statistics change quickly at some time-instant, then \( \lambda \) must be set to the lower value to enable quick PARCOR coefficients modification. In [14], the algorithm allowing for on-line modification of the forgetting factor value is presented. This technique may also be used to reset the equalization algorithm.

3.2 The equalizer gradient weights update

In [12] it is shown, that the set of the backward predictor error vectors \( r_m(n) \) where \( m = 0, \ldots, N \) form the orthogonal basis. In the presented equalizer, vector \( r(n) = [r_0(n) \ldots r_N(n)] \) is used to produce the output sample \( y(n) = w(n)r^T(n) \), where \( w(n) = [w_0(n) \ldots w_N(n)] \) is the filter coefficients vector. Each coefficient \( m = 0, \ldots, N \) is calculated by the following formula:

\[
w_m(n) = w_m(n-1) + g_m^2(n)\Psi(m)(n)r_m^*(n)
\]

(24)

Variable step size is calculated for \( m = 0, \ldots, N \) at each iteration by:

\[
g_m^2(n) = (1 - \alpha)g_m^2(n-1) + \beta_m^2(n)\]

(25)

4. SIMULATION RESULTS

The presented adaptation algorithm for the equalizer with soft switching was tested for the orthogonal and tapped–delay filter respectively. Also tapped–delay MMA, CMA and CMA+DD were simulated to provide compassion with the proposed solution. The simulations were performed for the 256 and 1024 QAM signals distorted by the ISI, with signal to noise ratio \( SNR = 45dB \). The

Figure 3: Blind MMA+ShMMA equalizer with the orthogonal basis

Figure 4: The scatter diagram of the equalizer output \( y(n) \) for the 1024-QAM, 10k symbols are plotted

\[
\rho_{m+1}(n) = \rho_{m+1}(n-1)E_{m}(n)R_{m}(n-1) + e_{m}(n)r_{m}(n-1)
\]

(21)
input data sequence was additionally distorted by the phase offset \( \phi = \pi/6 \). The FIR-type channel with the impulse response \( h = [0, 1, 0.2, 0, 4, 0.5, 1, 0.5, 0.4, 0.2, 0.1] \) was chosen. This is the non-minimum phase channel without zeros on the unit circle in the Z-domain. The eigenvalue spread of the correlation matrix is about 19/6B, which causes bad initial conditions for the algorithm.

In the lattice MMA+ShMMA equalizer, the forgetting factor of the orthogonal filter was \( \lambda = 0.9999 \) and the stochastic gradient initial adaptation step was \( \mu_m^{-2}(0) = 0.001 \) with \( \alpha = 0.003 \), where \( m = 0, \ldots, N \). For the tapped–delay MMA+ShMMA equalizer, the stochastic gradient adaptation step was \( \mu = 0.002 \). The stochastic gradient step size for the MMA and CMA was selected as \( \mu_{\text{MMA}} = 2 \cdot 10^{-3} \) and \( \mu_{\text{CMA}} = 17 \cdot 10^{-34} \) respectively. In case of the CMA+DD the stochastic gradient step size was chosen to be \( \mu_{\text{CMA+DD}} = 10^{-3} \), so the steady state MSE was small enough in order to perform hard switching to the decision directed algorithm without diverging the adaptation procedure.

All the equalizers were of the same length \( L = 20 \) and the initial filter coefficients were \( w(0) = [0 \ldots 0 \ 1 \ldots 0] \). The Edge MSE adaptive estimation was conducted with the forgetting factor \( \lambda_E = 0.98 \). Soft switching procedure for both 256 and 1024-QAM was configured to \( J_{\text{start}} = 10^{-1.8} \) and \( J_{\text{stop}} = 10^{-3.3} \). The Mean Squared Error for the lattice and tapped–delay equalizer is depicted in Figure 5. Scatter diagram for the 1024-QAM is presented in Figure 4. Real part of the received signal after equalization for the 256-QAM is shown in Figure 6.

5. CONCLUSIONS

In this paper, joint blind equalization procedure based on the MMA and ShMMA for the high order QAM has been presented. To provide a good measure of the equalizer output distortion, the edges of the constellation have been used in the process of the MSE estimation. Soft switching procedure of the equalization technique has been used to avoid a possibility of the convergence problems which can occur, when the hard switching is used for the high order QAM. Simulation results show that the joint blind equalizer MMA+ShMMA is able to cope with additional, constant phase offset. By using the J-orthogonal filter, the rate of convergence has been improved, comparing to the solution based on the tapped–delay filter.

REFERENCES