INTERNATIONAL JOURNAL OF ADAPTIVE CONTROL AND SIGNAL PROCESSING, VOL. 11, 621–640 (1997)

OPTIMAL DELAY ESTIMATION AND PERFORMANCE EVALUATION IN BLIND EQUALIZATION

M. PRANDINI, M. C. CAMPI* AND R. LEONARDI

Dipartimento di Elettronica per l'Automazione, Facoltà di Ingegneria, Università degli Studi di Brescia, Via Branze 38, I-25123 Brescia, Italy

SUMMARY

This paper deals with the problem of recovering the input signal applied to a linear time-invariant system from the measurements of its output and the *a priori* knowledge of the input statistics (blind equalization). Under the assumption of an i.i.d. non-Gaussian input sequence a new iterative procedure based on phase-sensitive high-order cumulants for adjusting the coefficients of a transversal equalizer is introduced. The main feature of the proposed technique is the automatic selection of the equalization delay so as to improve the equalization performance. A method for the *a posteriori* evaluation of the obtained accuracy in PAM systems is also introduced. It consists of the computation of an upper bound on the probability of error depending on certain moments of the equalizer output and the statistics of the channel input and therefore can be used in a blind equalization context. Based on the result of such a computation, it can be decided whether it is necessary to consider a longer equalization filter in the iterative procedure. © 1997 by John Wiley & Sons, Ltd.

Int. J. Adapt. Control Signal Process., **11**, 621–640 (1997)
No. of Figures: 8 No. of Tables: 3 No. of References: 27

Key words: blind equalization; linear filters; optimal delay estimation; error probability; accuracy evaluation

1. INTRODUCTION

Blind equalization[1] deals with the recovery of the input to an unknown system from the measurements of its output and the *a priori* knowledge of the input statistics. In a digital communication link the input is the sequence of transmitted signals, while the system represents the distortion caused by the transmission channel between the information source and the receiver. In this context, blind equalization is useful to recover a transmission after a severe distortion without having to send over the channel a training sequence (*a priori* known sequence of transmitted data); once the channel has been equalized, the blind algorithm that leads to the coefficients of the equalization filter is turned off.

The attempt to cancel the channel distortion so as to obtain a delayed version of the transmitted message is usually pursued through an adjustable FIR linear system called an equalizer, while the channel is modelled by a possibly non-minimum phase linear system.

This paper was recommended for publication by editor C.F.N. Cowan

* Correspondence to: M. C. Campi, Dipartimento di Elettronica per l'Automazione, Facoltà di Ingegneria, Università degli Studi di Brescia, Via Branze 38, I-25123 Brescia, Italy, E-mail: campi@elet.polimi.it

CCC 0890–6327/97/070621–20$17.50

Received 15 January 1996
Accepted 11 January 1997

A classical approach to blind equalization is to optimize a non-quadratic cost function of the equalizer output with a gradient search algorithm.[2–4] The cost function depends on the equalizer coefficients and its optimization leads to a matching of the probability density function of the individual recovered symbols with the probability density function of the individual input symbols. If the input is i.i.d. and non-Gaussian, the fulfillment of this condition implies that the channel has been equalized,[3] but the algorithm globally converges only under the unrealistic assumption of the use of a non-causal IIR equalizer and for particular input probability distributions.[3,5]

Recently, Shalvi and Weinstein have presented a different approach[6] which imposes no restrictions on the input distribution except for non-Gaussianity. It is based on the observation that solving the blind equalization problem requires the equalization of just a few moments of the input and output probability distributions. In practice it is sufficient to equalize the variance and the fourth-order cumulant.

To improve the approach proposed in Reference 6, in the interesting paper of Reference 7, Jelonnek and Kammeyer suggest a quality criterion, still based on the fourth-order cumulant, which can be optimized in a closed-form solution and therefore is computationally convenient. The corresponding algorithm is known under the acronym EVA (eigen vector approach). The analysis of EVA presented in Reference 8 shows that the procedure is strictly connected to Shalvi and Weinstein's approach.

A first objective of this paper is to critically analyse the EVA approach and to introduce a new improved iterative procedure still based on the closed-form solution presented in Reference 7.

A second objective consists of the design of an *a posteriori* measure to evaluate the equalizer accuracy reached by the proposed algorithm. This measure is applicable to PAM (pulse-amplitude-modulated) systems and gives an upper bound on the probability of error (i.e. the probability that an individual recovered symbol is different from the corresponding transmitted symbol), which represents a fundamental measure of performance.[9] The alternative technique based on correlation estimates on the output of the equalizer,[10] in contrast with the method developed in the present paper, is applicable only when the channel has a finite impulse response of known length.

The evaluation of the error probability is generally a difficult problem. In References 11–18 such an evaluation is addressed by assuming that the filter coefficients of the cascaded channel –equalizer are known. Typically an approximate value for the error is obtained by truncation on the impulse response[11–13] or using standard statistical inequalities.[14–18] More experimental approaches for error rate monitoring based only on the equalizer output are proposed in References 19–21. In References 19 and 20 a set of possible models for the description of the equalizer output statistics is assumed to be available *a priori*. By comparing the corresponding probability distribution with the distribution estimated from the data, a selection of the most suitable model is made and the probability of error is finally computed based on the chosen model. The main drawback of this approach rests on the fact that the output sequence is assumed to be independent, an assumption which looks unrealistic in the case of significant intersymbol interference, which may still be present if the equalizer is not performing as expected. A similar approach is considered in Reference 21 for the estimation of small probabilities of error. The fundamental hypothesis is that the output distribution belongs to a given parametrized family, so that the problem consists of the determination of a particular distribution in such a family. Again an independence assumption is made in the mathematical developments. For a review of the existing methods for error probability evaluation the reader is referred to Reference 22.

The paper is organized as follows. A brief review of the EVA algorithm is given in Section 2, while the method for the automatic search of the best delay is presented in Section 3. In Section 4, basic concepts on PAM transmission systems are introduced and a preliminary expression for the probability of error is derived. The upper bound on the probability of error is then determined in Section 5, while some simulations are provided in Section 6.

2. BRIEF REVIEW OF EVA ALGORITHM

Consider a linear, time-invariant and (possibly) non-minimum phase channel described by the transfer function

$$H(z) = \sum_{k=0}^{\infty} h(k) z^{-k}$$

The signal $d(\cdot)$ transmitted through the channel is assumed to be an i.i.d. sequence of non-Gaussian random variables with zero mean and variance σ_d^2. The received signal

$$v(t) = \sum_{k=0}^{\infty} h(k) d(t-k)$$

is processed by an equalizer with transfer function $E(z)$, which we describe as a tapped delay line with a fixed number n of parameters:

$$E(z) = \sum_{k=0}^{n-1} e(k) z^{-k}$$

The output of the equalizer with input $v(\cdot)$ is denoted by $x(\cdot)$. The EVA algorithm proposed in Reference 7 is based on the use of an additional filter, the so-called 'reference system'. It is a tapped delay line with the same number n of parameters as the equalizer:

$$F(z) = \sum_{k=0}^{n-1} f(k) z^{-k}$$

The output of the reference system with input $v(\cdot)$ is named $y(\cdot)$. The complete block diagram of the introduced systems is shown in Figure 1.

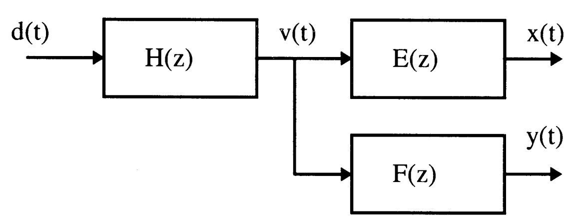

Figure 1. Basic block diagram for EVA algorithm

The EVA algorithm can be explained as follows.

Recall that the fourth-order cross-cumulant between two zero-mean stationary stochastic processes $p(\cdot)$ and $r(\cdot)$ is defined as

$$c_4^{pr}(n_1, n_2, n_3) := \mathrm{E}[p(t)r(t+n_1)p(t+n_2)r(t+n_3)] - \mathrm{E}[p(t)r(t+n_1)]\,\mathrm{E}[p(t+n_2)r(t+n_3)]$$
$$- \mathrm{E}[p(t)p(t+n_2)]\,\mathrm{E}[r(t+n_1)r(t+n_3)] - \mathrm{E}[p(t)r(t+n_3)]\,\mathrm{E}[r(t+n_1)p(t+n_2)]$$

When process $p(\cdot)$ coincides with $r(\cdot)$, we simply speak of the fourth-order cumulant of the process (see Reference 23, pp. 279–280 for a comprehensive presentation of cumulants and related properties).

Let us consider the synchronous (i.e. $n_1 = n_2 = n_3 = 0$) fourth-order cross-cumulant of processes $x(\cdot)$ and $y(\cdot)$, namely $c_4^{xy}(0,0,0) = \mathrm{E}[x^2(t)y^2(t)] - 2(\mathrm{E}[x(t)y(t)])^2 - \mathrm{E}[x^2(t)]\mathrm{E}[y^2(t)]$. By exploiting the relation $x(t) = \sum_{k=0}^{n-1} e(k)v(t-k)$, a simple computation shows that

$$c_4^{xy}(0,0,0) = \mathbf{e}^{\mathrm{T}}\,\mathbf{C}_4^{yv}\mathbf{e} \tag{1}$$

where

$$\mathbf{e} = [e(0)\;\; e(1)\;\; \ldots\;\; e(n-1)]^{\mathrm{T}}$$

is the equalizer coefficient vector and

$$\mathbf{C}_4^{yv} = \begin{bmatrix} c_4^{yv}(0,0,0) & c_4^{yv}(-1,0,0) & \cdots & c_4^{yv}(-n+1,0,0) \\ c_4^{yv}(-1,0,0) & c_4^{yv}(-1,0,-1) & \cdots & c_4^{yv}(-n+1,0,-1) \\ \vdots & \vdots & \ddots & \vdots \\ c_4^{yv}(-n+1,0,0) & c_4^{yv}(-n+1,0,-1) & \cdots & c_4^{yv}(-n+1,0,-n+1) \end{bmatrix}$$

The variance σ_x^2 of $x(\cdot)$ can be expressed as a quadratic function of $\mathbf{e}$ as well:

$$\sigma_x^2 = \mathbf{e}^{\mathrm{T}}\mathbf{R}_{vv}\mathbf{e} \tag{2}$$

with

$$\mathbf{R}_{vv} = \begin{bmatrix} r_{vv}(0) & r_{vv}(1) & \cdots & r_{vv}(n-1) \\ r_{vv}(1) & r_{vv}(0) & \cdots & r_{vv}(n-2) \\ \vdots & \vdots & \ddots & \vdots \\ r_{vv}(n-1) & r_{vv}(n-2) & \cdots & r_{vv}(0) \end{bmatrix}$$

where $r_{vv}(k)$ denotes the correlation coefficient $\mathrm{E}[v(t+k)v(t)]$.

In the EVA algorithm the reference system is supposed to be fixed in advance (later in this section, when discussing the iterative EVA procedure, we shall focus on the way in which the reference system is actually selected). If we accept that $F(z)$ is fixed (and known), a sample estimate of matrices $\mathbf{C}_4^{yv}$ in (1) and $\mathbf{R}_{vv}$ in (2) can be obtained from the data $v(\cdot)$ (i.e. the channel output) and the filtered signal $y(\cdot) = F(z)v(\cdot)$ by simply replacing expectations with empirical averages. As a matter of fact it is known that such sample estimates are consistent (see e.g. References 24 and 25). Name $\hat{\mathbf{C}}_4^{yv}$ and $\hat{\mathbf{R}}_{vv}$ such matrices. Below we first introduce the EVA algorithm and then discuss the intuitive rationale for it.

EVA algorithm

Maximize the cost function $|\hat{c}_4^{xy}(0,0,0)| = |\mathbf{e}^{\mathrm{T}}\hat{\mathbf{C}}_4^{yv}\mathbf{e}|$ subject to the constraint $\hat{\sigma}_x^2 := \mathbf{e}^{\mathrm{T}}\hat{\mathbf{R}}_{vv}\mathbf{e} = \sigma_d^2$ (recall that σ_d^2 is the variance of the transmitted message) (3)

The solution to this maximization problem is obtained as the solution to the classical generalized eigenvector problem

$$\hat{\mathbf{C}}_4^{yv}\mathbf{e} = \lambda\hat{\mathbf{R}}_{vv}\mathbf{e}$$

where λ is the generalized eigenvalue with the maximum absolute value ($|\lambda| = \max\{|\lambda_1|, \ldots, |\lambda_n|\}$), with the constraint on the norm of vector $\mathbf{e}$ provided by the relation $\mathbf{e}^{\mathrm{T}}\hat{\mathbf{R}}_{vv}\mathbf{e} = \sigma_d^2$. From this comes the name 'eigen vector approach' (EVA) of the algorithm (see Reference 7 for more details). It is worth noting that the EVA algorithm requires a computational effort lower than that required by other standard algorithms[2–6] based on the stochastic gradient method (see Reference 8 for further discussion on this point).

To understand that the problem (3) leads to coefficients $\mathbf{e}$ which equalize the channel $H(z)$, let us introduce the notation $s(k) = h(k) * e(k)$ (i.e. $s(k)$ is the impulse response of the channel –equalizer cascade) and $w(k) = h(k) * f(k)$. With this notation the left-hand sides of equations (1) and (2) can be rewritten as

$$c_4^{xy}(0, 0, 0) = c_4^{dd}(0, 0, 0) \sum_{k=0}^{\infty} w^2(k)\, s^2(k) \tag{4}$$

$$\sigma_x^2 = \sigma_d^2 \sum_{k=0}^{\infty} s^2(k) \tag{5}$$

Suppose that coefficients $\{s(k)\}$ can be selected freely in the maximization problem (this is in fact not the case, since $s(k) = h(k) * e(k)$ and $\{e(k)\}$ is formed by a finite number of coefficients). In this case, if $\{|w(k)|\}$ has a unique maximum value $|w(k_0)|$, maximizing $|c_4^{xy}(0, 0, 0)|$ subject to $\sigma_x^2 = \sigma_d^2$ obviously leads to $s(k) = \pm\,\delta(k - k_0)$, i.e. to channel equalization. In the non-ideal situation in which the number of coefficients $e(k)$ is finite, the equalization objective can only be partially achieved and the obtained solution depends on the value of coefficients $\{w(k)\}$. In this case, in order to improve the equalization result over the performance achievable with the one-step method (3), it can be convenient to resort to an iterative procedure[8] as explained below.

Iterative EVA procedure

1. Fix an initial reference system $F^{(0)}(z)$ and set $j = 0$.
2. Solve the optimization problem (3) with reference system $F^{(j)}(z)$ and denote by $E^{(j)}(z)$ the corresponding solution. Set $j = j + 1$.
3. Set $F^{(j)}(z) = E^{(j-1)}(z)$.
4. If $j \leqslant N$ (N is an *a priori* fixed integer), go to step 2.

The intuitive idea behind this iterative procedure is as follows. At the first iteration the reference system is fixed somehow arbitrarily and therefore coefficients $\{w(k)\}$ are fixed to a somewhat random value. In particular there is no reason why $\{w(k)\}$ should have a dominant coefficient. When maximizing $|c_4^{xy}(0, 0, 0)| = |c_4^{dd}(0, 0, 0)\sum_{k=0}^{\infty} w^2(k)s^2(k)|$ (see equation (4)), sequence $\{s(k)\}$ will tend to have a dominant coefficient $s(k_0)$ corresponding to the maximum value $|w(k_0)|$ of sequence $\{|w(k)|\}$. Moreover, $s(k_0)$ will be more dominant over the other coefficients to the extent by which coefficient $w(k_0)$ dominates over the other $w(k)$. In the second iteration the condition that $\{w(k)\}$ has a dominant coefficient is enforced by replacing the reference system with the equalizer computed at the first iteration. By repeating this procedure several times, it is expected that the final equalizer will exhibit improved equalization properties with respect to those obtained through a single iteration. As a matter of fact in Reference 8 it is shown that the equalization performance can be improved in just a few iterations.

As an additional remark we note that, because of the above-explained mechanism, the iterative EVA procedure has a natural tendency to maintain the position of the dominant coefficient of $\{s(k)\}$ in subsequent iterations. Thus, if k_0 is the initial position of the dominant coefficient (i.e. $|s(k_0)| > |s(k)|, \forall\, k \neq k_0$), it is probable that $s(k_0)$ will be dominant even at the end of the procedure.

In the next section we critically discuss the iterative EVA procedure and point out some of its drawbacks. A modified version is then introduced to circumvent these difficulties. This leads to an equalization algorithm in which the equalization delay in the reconstructed message is optimally selected so as to maximize the performance of the equalizer.

3. TOWARDS AUTOMATIC SEARCH FOR OPTIMAL DELAY

For convenience let us introduce the term *equalization delay* as the index of the dominant coefficient in the impulse response of the channel–equalizer cascade:

$$\text{equalization delay} = \arg\max_k \{|s(k)|\}$$

Roughly the equalization delay is the time interval which separates the reconstruction of the transmitted message at the output of the equalizer from the time it was sent over the channel.

As mentioned at the end of the previous section, the iterative EVA procedure has a natural tendency to maintain the equalization delay in subsequent iterations. As a consequence, if k_0 is the initial equalization delay, i.e. $k_0 = \arg\ \max_k |h(k) * e^{(0)}(k)|$, where $e^{(0)}(k)$ represents the equalizer coefficients at iteration 0, it is highly probable that the final equalization delay will be k_0 as well. This statement is confirmed by the simulation results in Section 6.

On the other hand, one can notice the following.

(a) The best achievable equalization accuracy for a given equalization delay $\bar{k}$ dramatically depends on the value of $\bar{k}$. In other words,

$$\min_{\{e(k)\}} \mathrm{E}[(e(k) * h(k) * d(k) - d(k - \bar{k}))^2] \tag{6}$$

significantly depends on $\bar{k}$. The intuitive reason for this can be easily understood. In expression (6) the term $e(k) * h(k) * d(k)$ represents the response of the equalizer fed by the channel output $h(k) * d(k)$. When the channel is minimum phase, sample $d(t)$ can be reconstructed from the channel output up to the current instant point t (simply by inverting the channel transfer function). On the other hand, when the channel is non-minimum phase, this is not possible (the inverse channel transfer function would be unstable!) and the reconstruction of $d(t)$ can be improved by using also future channel outputs. Therefore we see that increasing the equalization delay $\bar{k}$ may have a beneficial effect on the reconstruction capability. On the other hand, the equalization delay cannot be enlarged at will. This is because the equalizer length is finite and increasing $\bar{k}$ to be too large results in a decrease in the number of past samples of $d(t)$ used in the reconstruction. A significant example of this behaviour can be found later in Section 6 (see Figure 8).

(b) The initial equalization delay k_0 is not known. As a matter of fact, $k_0 = \arg\max_k |h(k) * e^{(0)}(k)|$ depends on the channel coefficients $\{h(k)\}$ which are unknown.

From (a) it follows that it is highly desirable to select a suitable value of the equalization delay. Point (b), however, shows that the iterative EVA procedure is not suitable in this respect. These observations prompt the need to modify the original iterative EVA algorithm so as to correctly adjust the equalization delay to reach a better equalization performance than the one obtained with the initial equalization delay.

The basic iteration of the proposed algorithm is still based on the maximization problem (3), which is an effective and computationally convenient way to search for the equalizer that reconstructs the transmitted messages with a certain delay. However, differently from the iterative EVA procedure, during a single iteration of the new algorithm, three sets of equalizer coefficients ($\mathbf{e}_0$, $\mathbf{e}_{-1}$, $\mathbf{e}_1$) are evaluated, one corresponding to a given reference system $F(z)$ and the other two corresponding to $zF(z)$ and $z^{-1}F(z)$. By means of this simple modification a joint optimization of the equalizer coefficients and of the equalization delay can be achieved.

To be specific, vectors $\mathbf{e}_0$, $\mathbf{e}_{-1}$ and $\mathbf{e}_1$ are computed as follows.

1. Vector $\mathbf{e}_0$ is computed by maximizing the absolute value of the cross-cumulant between $x(\cdot)$ and $y(\cdot)$ under the constraint $\sigma_x^2 = \sigma_d^2$.
2. Vector $\mathbf{e}_{-1}$ is computed by maximizing the absolute value of the cross-cumulant between $x(\cdot)$ and $y_{-1}(\cdot)$, where $y_{-1}(\cdot)$ is obtained as the output of the shifted reference system $F_{-1}(z) = zF(z)$ ($\{w_{-1}(k)\} = \{w(k+1)\}$) (Figure 2), under the same constraint $\sigma_x^2 = \sigma_d^2$.
3. Vector $\mathbf{e}_1$ is computed by maximizing the absolute value of the cross-cumulant between $x(\cdot)$ and $y_1(\cdot)$, where $y_1(\cdot)$ is obtained as the output of the shifted reference system $F_1(z) = z^{-1}F(z)$ ($\{w_1(k)\} = \{w(k-1)\}$) (Figure 3), under the same constraint $\sigma_x^2 = \sigma_d^2$.

To understand the effect on the equalization delay of the shifting of $y(\cdot)$, notice that if coefficients $\{s(k)\}$ could be selected freely and $\{|w(k)|\}$ had a unique maximum $|w(k_0)|$, the three maximization problems would imply perfect channel equalization with delay k_0 and shifted

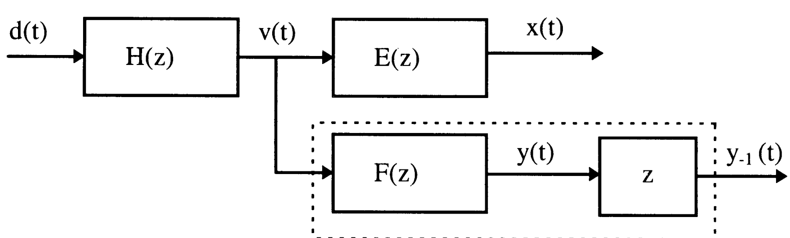

Figure 2. Block diagram to determine equalizer $\mathbf{e}_{-1}$

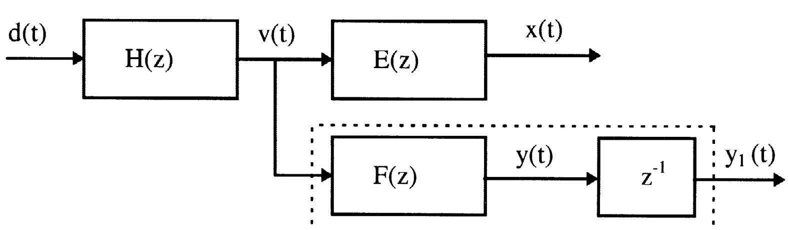

Figure 3. Block diagram to determine equalizer $\mathbf{e}_1$

delays $k_0 - 1$ and $k_0 + 1$ respectively. In the real case, where $\{s(k)\}$ cannot be selected freely because $\{e(k)\}$ is formed by a finite number of coefficients, this fact does not hold true rigorously. However, in normal applications it is still approximately correct as shown by simulation results (see Section 6).

Remark 1

Two main exceptions to the general shifting rule described above are obtained as a result of 'boundary effects'.

(a) When $k_0 = 0$, $|w_{-1}(-1)|$ is the maximum value of $\{|w_{-1}(k)|\}$ but does not influence the cross-cumulant value

$$c_4^{xy_{-1}}(0,0,0) = c_4^{dd}(0,0,0) \sum_{k=0}^{\infty} w_{-1}^2(k) s^2(k)$$

since the channel and the equalizer are both causal.

(b) When channel is an FIR system and $k_0 = q$, where q is the index of the last non-zero coefficient of the channel–equalizer cascade, $|w_1(q+1)|$ is the maximum value of $\{|w_1(k)|\}$ but does not influence the cross-cumulant value

$$c_4^{xy_1}(0,0,0) = c_4^{dd}(0,0,0) \sum_{k=0}^{q} w_1^2(k) s^2(k)$$

since $s(k) = 0$ for $k > q$.

In the presence of these two types of 'boundary effect' the equalization delay corresponding respectively to $\mathbf{e}_{-1}$ and $\mathbf{e}_1$ depends on the position of $\max_{k \neq k_0}\{|w(k)|\}$. □

In an EVA iteration the cross-cumulant absolute value between the equalizer and reference system outputs represents the performance index to determine the equalizer coefficients. At each iteration of the new algorithm we choose among $\mathbf{e}_{-1}, \mathbf{e}_0$ and $\mathbf{e}_1$ by comparing the three cross-cumulant absolute values as well.

Remark 2

Observe that the modified EVA procedure requires a computational effort per iteration which is less than three times that of the EVA procedure. In fact one has to solve three eigenvector problems and to compute a sample estimate of the three matrices $\mathbf{C}_4^{yv}$, $\mathbf{C}_4^{y_{-1}v}$ and $\mathbf{C}_4^{y_1 v}$. However, as can be easily verified, most elements of these three matrices are coincident and therefore need to be estimated only once. □

On the grounds of the above discussion we introduce the following new procedure for channel equalization.

Modified iterative EVA procedure

1. Fix an initial reference system $F^{(0)}(z)$ and set $j = 0$.
2. Solve the optimization problem (3) with reference system $F^{(j)}(z)$ and denote by $E^{(j)}(z)$ the corresponding solution.
3. Set $j = j + 1$ and $F^{(j)}(z) = E^{(j-1)}(z)$.

4. Compute each of the three sets of equalizer coefficients $\mathbf{e}_{-1}, \mathbf{e}_0$ and $\mathbf{e}_1$ by solving the optimization problem (3) with $zF^{(j)}(z), F^{(j)}(z)$ and $z^{-1}F^{(j)}(z)$ respectively as reference system.
5. Select the equalizer leading to the greatest cross-cumulant value. Denote it by $E^{(j)}(z)$.
6. If $j < N$ (N is an *a priori* fixed integer), go to step 3.

The procedure is described in the block diagram of Figure 4.

The modified iterative EVA procedure can be applied to any non-Gaussian i.i.d. input signal. In the next section we focus on the particular case of PAM (pulse-amplitude-modulated) communication. In this case a bound on the probability of error for the *a posteriori* evaluation of the equalization performance can be derived. This result can be used to design a suitable stop criterion for the modified EVA algorithm.

4. PROBABILITY OF ERROR IN PAM SYSTEMS

The typical communication system known under the acronym PAM[9] is characterized by a transmitted signal $d(\cdot)$ belonging to a finite alphabet A. We assume that $d(\cdot)$ is an i.i.d. sequence with an even number M of equiprobable values. Precisely,

$$d(t) \in A = \{-(M-1), \ -(M-3), \ldots, M-3, \ M-1\}$$

$$\Pr(d(t) = 2i - M - 1) = 1/M, \quad i = 1, 2, \ldots, M$$

In PAM systems it is common practice to add at the output of the equalizer $E(z)$ a non-linear memoryless system called a nearest-neighbour M-ary quantizer (Figure 5). It is a threshold device which is introduced to cancel the residual distortion to which the equalized signal $x(\cdot)$ can still be subject.

Specifically, the quantizer output $x_q(t)$ is the value in the alphabet A nearest to the equalized signal $x(t)$, i.e.

$$x_q(\beta) = \sum_{j=1-M/2}^{M/2-1} \operatorname{sgn}(\beta + 2j)$$

where

$$\operatorname{sgn}(\beta) = \begin{cases} 1, & \beta > 0 \\ 0, & \beta = 0 \\ -1, & \beta < 0 \end{cases}$$

Denoting by $\bar{k}$ the equalization delay, i.e. the index of the dominant coefficient in the impulse response of the channel–equalizer cascade, we say that an error occurs when $x_q(t) \neq \operatorname{sgn}(s(\bar{k}))d(t-\bar{k})$. Consequently, a meaningful measure of the overall PAM system performance is provided by the probability P_e of the event $x_q(t) \neq \operatorname{sgn}(s(\bar{k}))d(t-\bar{k})$, considering all possibilities of transmitted symbols, i.e. $d(t-\bar{k}) = 2i - M - 1$, $i = 1, 2, \ldots, M$.

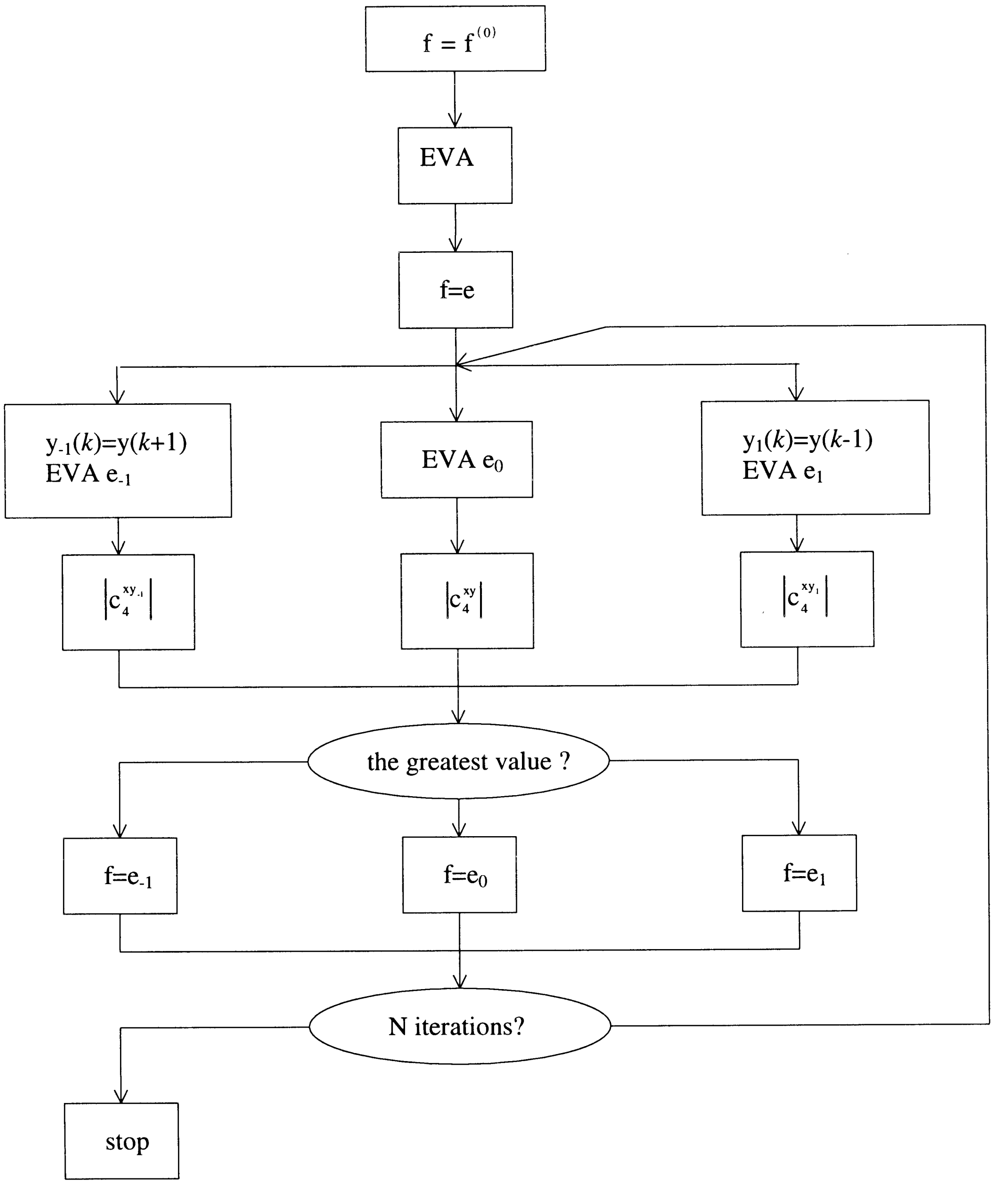

Figure 4. Block diagram of modified iterative EVA procedure

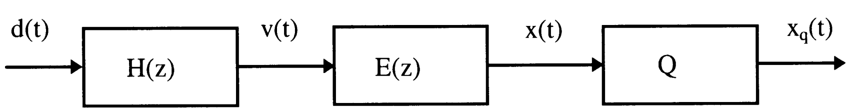

Figure 5. Complete block scheme for PAM communication system

Applying the theorem of total probability, P_e can be computed as

$$P_e = \sum_{i=1}^{M} \Pr\{x_q(t) \neq \operatorname{sgn}(s(\bar{k}))d(t-\bar{k})/d(t-\bar{k}) = 2i-M-1\} \Pr\{d(t-\bar{k}) = 2i-M-1\}$$

$$= \frac{1}{M} \sum_{i=1}^{M} \Pr\{x_q(t) \neq \operatorname{sgn}(s(\bar{k}))d(t-\bar{k})/d(t-\bar{k}) = 2i-M-1\}$$

where $\Pr\{A/B\}$ denotes the conditional probability of A given B.

In the following we assume that the channel output samples used to compute the equalizer coefficients are independent of the realization on which the equalizer performance is tested. Therefore coefficients $\{s(k)\}$ will be regarded as deterministic and not a function of $d(\cdot)$. In this case the expression of P_e can be simplified by observing that $x(t)$ is given by

$$x(t) = s(\bar{k})\, d(t-\bar{k}) + \xi(t)$$

where $\xi(t) := \sum_{k \neq \bar{k}} s(k) d(t-k)$, which denotes the *intersymbol interference*, is independent of $s(\bar{k})d(t-\bar{k})$ and symmetrically distributed. This implies that

$$\Pr\{x_q(t) \neq \operatorname{sgn}(s(\bar{k}))d(t-\bar{k})/d(t-\bar{k}) = 2i-M-1\}$$
$$= \Pr\{x_q(t) \neq \operatorname{sgn}(s(\bar{k}))d(t-\bar{k})/d(t-\bar{k}) = 2(M+1-i)-M-1\}, \quad i = 1, 2, \ldots, M/2$$

from which it follows that

$$P_e = \frac{2}{M} \sum_{i=M/2+1}^{M} \Pr\{x_q(t) \neq \operatorname{sgn}(s(\bar{k}))\, d(t-\bar{k})/d(t-\bar{k}) = 2i-M-1\}$$

We now introduce the technical assumption

$$\text{A1:} \quad (M-1)|s(\bar{k})| > M-2$$

which is satisfied if $|s(\bar{k})|$ is close enough to one. Under this hypothesis an error occurs only if $\xi(t)$ makes $x(t)$ go beyond the boundaries of the quantization interval in which $s(\bar{k})d(t-\bar{k})$ falls. P_e can then be calculated as

$$P_e = \frac{2}{M}\Bigg(\sum_{i=M/2+1}^{M-1} [\tfrac{1}{2}\Pr\{|\xi(t)| \geqslant (2i-M-1)|s(\bar{k})| - (2i-M-2)\}$$
$$+ \tfrac{1}{2}\Pr\{|\xi(t)| \geqslant (2i-M) - (2i-M-1)|s(\bar{k})|\}]$$
$$+ \tfrac{1}{2}\Pr\{|\xi(t)| \geqslant (M-1)|s(\bar{k})| - (M-2)\}\Bigg) \tag{7}$$

Equation (7) provides a rigorous expression for computing the probability of error P_e. The evaluation of the different terms in (7) would, however, require the knowledge of $s(\bar{k})$ and of the statistical properties of $\xi(t)$, quantities which are obviously not available. In the next section we first show a means to evaluate the range of possible values for $s(\bar{k})$ based on measurements. Then we determine an upper bound for P_e which only depends on $s(\bar{k})$, so that P_e can be estimated.

5. UPPER BOUND ON PROBABILITY OF ERROR

Denote by α the normalized fourth-order cumulant of the equalizer output $x(t)$, i.e.

$$\alpha := c_4^{xx}(0, 0, 0)/c_4^{dd}(0, 0, 0)$$

It is easily seen that α can be expressed in terms of the coefficients of the cascaded channel–equalizer as (from now on, all summations are computed from $k = 0$ to ∞)

$$\alpha = \sum_k s^4(k)$$

α can be estimated from a realization of the equalizer output and can thus be used to define an admissible range for $s(\bar{k})$. In fact, under the condition

$$\sum_k s^2(k) = 1$$

directly imposed by both the EVA and modified EVA procedures (see (5) and remember that the condition $\sigma_x^2 = \sigma_d^2$ is imposed in the EVA and modified EVA procedures), we have that if $\alpha = 1, s(\bar{k})$ can only have unitary modulus (and $s(k) = 0$, $\forall k \neq \bar{k}$), while if $\alpha < 1$, then the following proposition holds.

Proposition 1

Let $\{s(k)\}$ be subject to the conditions

$$\sum_k s^2(k) = 1 \tag{8}$$

$$\sum_k s^4(k) = \alpha \tag{9}$$

Then, for $\alpha > \frac{1}{2}$, $s^2(\bar{k}) := \max_k \{s^2(k)\}$ is unique and satisfies the inequality

$$\frac{1 + \sqrt{(2\alpha - 1)}}{2} \leqslant s^2(\bar{k}) \leqslant \sqrt{\alpha} \tag{10}$$

Proof. See Appendix. □

Remark 3

(a) It is easy to see that for $\alpha \leqslant \frac{1}{2}$, it may happen that the sequence $\{s^2(k)\}$ admits more than a single absolute maximum. In this case Proposition 1 is no longer valid. On the other hand, for the equalization performance to be acceptable, α must be close to one (if this is not satisfied, the dominant coefficient $|s(\bar{k})|$ cannot be close to one!). Therefore the case $\alpha \leqslant \frac{1}{2}$ is of no practical interest.

(b) As a matter of fact it is easy to show that the lower and upper bounds in (10) cannot be improved. In other words, there exist sequences $\{s(k)\}$ such that $s^2(\bar{k})$ approaches at will such bounds. □

Our goal is the determination of an upper bound on P_e based on $s(\bar{k})$, which, in view of (10), can be successively expressed as a function of α only.

We see from (7) that P_e is the sum of terms all of the form

$$\Pr(|\xi(t)| \geqslant A) \tag{11}$$

where A is a function of $s(\bar{k})$.

In order to compute the upper bound for such a probability, one can resort to one of the following three well-known inequalities (which one of these three inequalities is convenient to use depends on the actual values of $s(\bar{k})$ and α; see below):

$$\Pr(|\xi(t)| \geqslant A) \leqslant \frac{\operatorname{var}\xi(t)}{A^2} \quad \text{(Tchebycheff inequality[26])} \tag{12}$$

$$\Pr(|\xi(t)| \geqslant A) \leqslant 2\exp\left(-\frac{1}{2}\frac{A^2}{\operatorname{var}\xi(t)}\right) \quad \text{(Chernoff inequality[15])} \tag{13}$$

$$\Pr(|\xi(t)| \geqslant A) \leqslant \frac{\mathrm{E}[\xi^4(t)]}{A^4} \quad \text{(Markov inequality[27])} \tag{14}$$

The right-hand sides of equations (12)–(14) can be computed in terms of $s(\bar{k})$ and α. This is obvious for (12) and (13), since $\operatorname{var}\xi(t) = (1 - s^2(\bar{k}))\sigma_d^2$, while as far as equation (14) is concerned, note that

$$\mathrm{E}[\xi^4(t)] = \sum_{k \neq \bar{k}} s^4(k) c_4^{dd}(0,0,0) + 3(\operatorname{var}\xi(t))^2 \quad \text{and} \quad \sum_{k \neq \bar{k}} s^4(k) = \alpha - s^4(\bar{k})$$

The upper bound for the probability (11) is then determined as

$$\Pr(|\xi(t)| \geqslant A) = \min\left\{1, \frac{(1 - s^2(\bar{k}))\sigma_d^2}{A^2}, 2\exp\left(-\frac{1}{2}\frac{A^2}{(1 - s^2(\bar{k}))\sigma_d^2}\right), \frac{(\alpha - s^4(\bar{k}))c_4^{dd}(0,0,0) + 3(1 - s^2(\bar{k}))^2\sigma_d^4}{A^4}\right\} \tag{15}$$

By inserting (15) into (7) and using bounds (10), a simple but tedious computation leads to the following final upper bound on P_e:

$$P_e \leqslant \frac{1}{M}\sum_{i = M/2+1}^{M} L_i \tag{16}$$

where

$$L_i = \min\left\{1, \frac{(1 - f(\alpha))\sigma_d^2}{a_i}, \quad 2\exp\left(-\frac{1}{2}\frac{a_i}{(1 - f(\alpha))\sigma_d^2}\right), \quad \frac{(\alpha - f^2(\alpha))c_4^{dd}(0,0,0) + 3(1 - f(\alpha))^2\sigma_d^4}{a_i^2}\right\}$$
$$+ \min\left\{1, \frac{(1 - f(\alpha))\sigma_d^2}{b_i}, \quad 2\exp\left(-\frac{1}{2}\frac{b_i}{(1 - f(\alpha))\sigma_d^2}\right), \quad \frac{(\alpha - f^2(\alpha))c_4^{dd}(0,0,0) + 3(1 - f(\alpha))^2\sigma_d^4}{b_i^2}\right\}$$

for $i = M/2 + 1, \ldots, M - 1$ and

$$L_M = \min\left\{1, \frac{(1 - f(\alpha))\sigma_d^2}{a_M}, \quad 2\exp\left(-\frac{1}{2}\frac{a_M}{(1 - f(\alpha))\sigma_d^2}\right), \quad \frac{(\alpha - f^2(\alpha))c_4^{dd}(0,0,0) + 3(1 - f(\alpha))^2\sigma_d^4}{a_M^2}\right\}$$

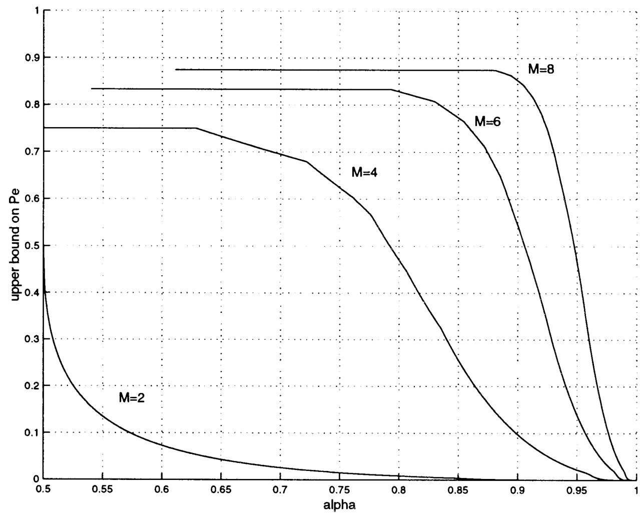

Figure 6. Upper bound on probability of error P_e

with

$$a_i = [(2i - M - 1)\sqrt{f(\alpha)} - (2i - M - 2)]^2$$

$$b_i = [(2i - M) - (2i - M - 1)\sqrt{g(\alpha)}]^2$$

$$f(\alpha) = \frac{1 + \sqrt{(2\alpha - 1)}}{2}, \qquad g(\alpha) = \sqrt{\alpha}$$

The upper bound for P_e is displayed for different values of M in Figure 6. It can be shown that for $\alpha \approx 1$ the tightest bound is obtained by means of the Chernoff inequality, which corresponds to the third element under the sign of 'min' in the expression for L_i. The Markov and Tchebycheff inequalities give better bounds for lower values of α. The joint use of the three inequalities provides a tight bound for a wide range of values of α.

Bound (16) quantifies the equalization accuracy and can be used as a measure to decide whether the equalizer performance is satisfactory. Obviously such a measure works no matter which blind algorithm is being used. On the other hand, we can note that its introduction in the modified EVA procedure does not require additional computational effort. The reason for this lies in the fact that an estimate of the fourth-order equalizer output cumulant is directly available as the final value of the performance index for the algorithm. Depending on the bound of the probability of error, a decision can be made whether it is necessary to increase the length of the equalizer. If the probability of error is high, one can enlarge the equalizer length so as to ascertain a better equalization accuracy. Moreover, in doing so, one can exploit the actual equalizer as initial reference system so as to reduce the number of iterations needed to determine the new equalizer.

6. SIMULATION EXAMPLE

In this section we present an example which illustrates the performance of the new procedure introduced in this paper. In order to make a significant comparison, we fix the same initial reference system ($F_0(z) = z^{-p}$ with $0 \leqslant p \leqslant n-1$) for the EVA and modified EVA procedures. Among other things, in the case of an i.i.d. 4-ary equiprobable input we obtain that the upper bound on the probability of error for the modified EVA algorithm solution is lower than the corresponding EVA solution, as expected.

Consider a telephone channel[9] with transfer function

$$H(z) = 0{\cdot}04 - 0{\cdot}05z^{-1} + 0{\cdot}07z^{-2} - 0{\cdot}21z^{-3} - 0{\cdot}50z^{-4} + 0{\cdot}72z^{-5} + 0{\cdot}36z^{-6} + 0{\cdot}21z^{-8} + 0{\cdot}03z^{-9} + 0{\cdot}07z^{-10}$$

fed by a 4-ary i.i.d. equiprobable signal with values in $\{-3, -1, 1, 3\}$, and a five-tap equalizer.

The performances obtained by the two procedures are shown in Figure 7, describing the evolution of the equalization delay and of the normalized cumulant α of the equalizer output for

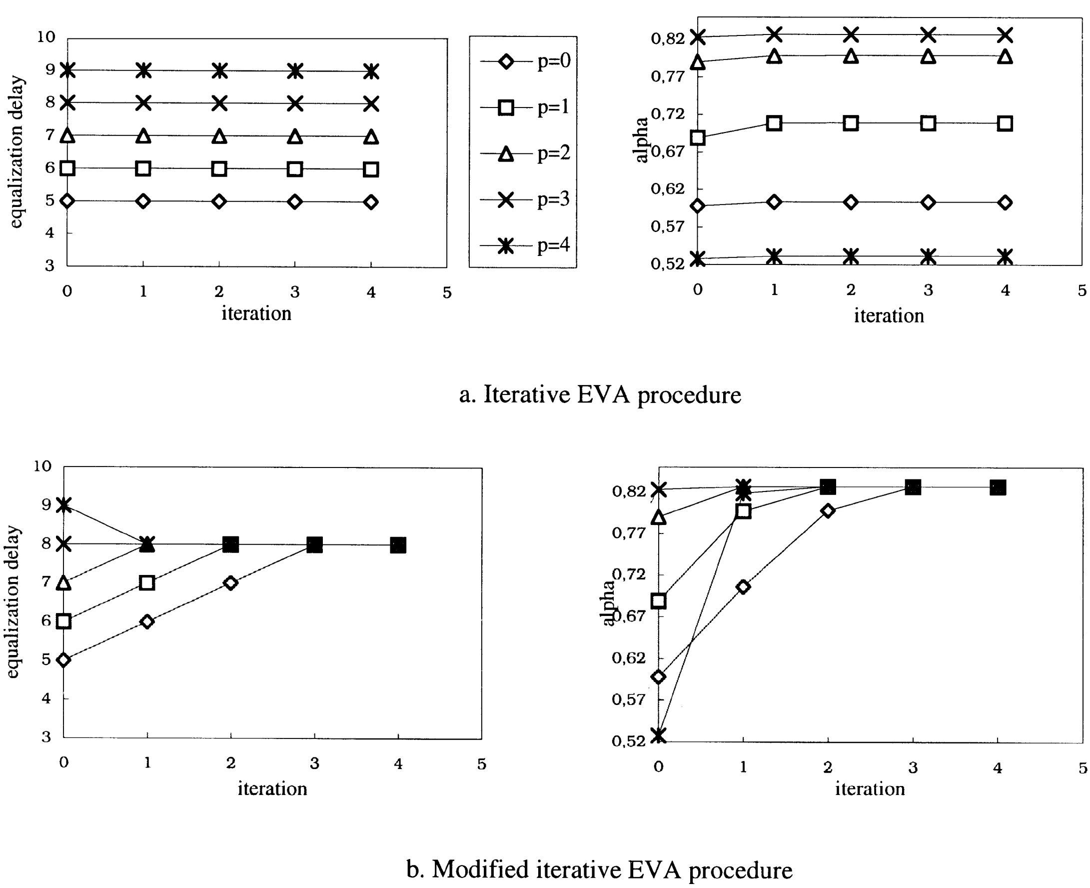

Figure 7. (a) Iterative EVA procedure. (b) Modified iterative EVA procedure

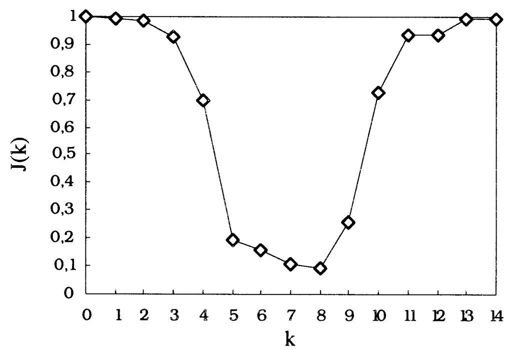

Figure 8. Performance index of mean square error equalizer

all possible values of the initial reference system delay p. In the simulations all the expected values are computed as sample averages from a realization of 1000 samples of the channel input $d(\cdot)$.

As discussed in Section 3, we find that the iterative EVA procedure tends to maintain the initial equalization delay, while the new procedure gives the same final delay regardless of the initialization. As indicated by the mean square restoration error

$$J(k) = \frac{1}{\sigma_d^2} \min_{\{e(k)\}} \mathrm{E}[(x(t) - d(t-k))^2]$$

displayed in Figure 8, the procedure presented in this paper leads to an equalization delay that corresponds to an optimal mean square error solution delay. Moreover, if we normalize the coefficients of the mean square error equalizer minimizing $J(k)$ so that the variance of the reconstructed input is equal to the input variance, we find the equalizer

$$E'(z) = 0{\cdot}1319 + 0{\cdot}2836z^{-1} + 0{\cdot}2996z^{-2} + 0{\cdot}8218z^{-3} - 0{\cdot}4664z^{-4}$$

which is quite close to the solution

$$E(z) = 0{\cdot}1396 + 0{\cdot}2809z^{-1} + 0{\cdot}2993z^{-2} + 0{\cdot}8199z^{-3} - 0{\cdot}4668z^{-4}$$

obtained by means of the proposed algorithm.

The value of bound (16) can be further used as an *a posteriori* measure to estimate the performance of the obtained equalizer and decide whether it is convenient to increase the equalizer length.

The tightness of the bound has been tested by simulation on both the modified EVA and EVA solutions determined for $p = 2$. Ten records, of 10 000 samples each, have been considered in both cases. It is shown that the bound gives results which are conservative by approximately a factor of three with respect to simulations. Discrepancies between the theoretical bound and practical experiments can be expected, since the bound in equation (16) must remain valid for any channel–equalizer cascade and not only for the one considered in the simulations.

EVA solution ($p = 2$)

The normalized fourth-order cumulant estimated from the equalizer output is $\alpha = 0{\cdot}7986$. The corresponding upper bound for the probability of error is 0·4760. Table I displays the number of errors for each simulation of 10 000 samples.

Modified EVA solution ($p = 2$)

The normalized fourth-order cumulant estimated from the equalizer output is $\alpha = 0{\cdot}8270$. The corresponding upper bound for the probability of error is 0·3572. Table II displays the number of errors for each simulation of 10 000 samples.

As a final experiment we have generated 10 independent realizations, of 1000 samples each, of the input process and determined the corresponding equalizers by the EVA and modified EVA procedures. In Table III the characteristics of the final equalizers (obtained after five iterations of the procedures) are displayed in terms of the mean value and standard deviation (SD) of the output normalized cumulant α.

Table I. Simulation results (EVA, $p = 2$, 10 000 samples/record)

Simulation	Number of errors
Record 1	1857
Record 2	1764
Record 3	1775
Record 4	1845
Record 5	1805
Record 6	1800
Record 7	1810
Record 8	1797
Record 9	1842
Record 10	1877

Table II. Simulation results (modified EVA, $p = 2$, 10 000 samples/record)

Simulation	Number of errors
Record 1	1614
Record 2	1703
Record 3	1660
Record 4	1619
Record 5	1647
Record 6	1616
Record 7	1673
Record 8	1648
Record 9	1669
Record 10	1711

Table III. Summary of EVA and modified EVA procedure solutions

	Iterative EVA procedure		Modified iterative EVA procedure	
Initialization	Equalization delay	α (mean $\pm$ SD)	Equalization delay	α (mean $\pm$ SD)
$p=0$	5	$0{\cdot}6574 \pm 0{\cdot}0399$	8	$0{\cdot}8224 \pm 0{\cdot}0294$
$p=1$	6	$0{\cdot}7253 \pm 0{\cdot}0326$	8	$0{\cdot}8224 \pm 0{\cdot}0294$
$p=2$	7	$0{\cdot}7920 \pm 0{\cdot}0261$	8	$0{\cdot}8225 \pm 0{\cdot}0293$
$p=3$	8	$0{\cdot}8225 \pm 0{\cdot}0293$	8	$0{\cdot}8225 \pm 0{\cdot}0293$
$p=4$	9	$0{\cdot}5774 \pm 0{\cdot}0488$	8	$0{\cdot}8225 \pm 0{\cdot}0293$

7. CONCLUSIONS

We have presented an iterative procedure for blind equalization based on fourth-order statistics, which automatically searches for the optimal equalization delay and at the same time tries to determine the best equalizer for the final delay. Therefore the algorithm copes with the problem of the selection of a suitable value for the equalization delay. This problem is particularly relevant because of the strong dependence of the equalization accuracy on this delay. We have introduced an upper bound on the probability of error in PAM systems. Since this bound only depends on the equalizer output measurements and on the input statistics, it is suitable for the *a posteriori* evaluation of the blind equalizer performance. It was identified that for high values of the normalized output fourth-order cumulant α such a bound is better obtained with the use of the Chernoff inequality. On the other hand, the Tchebycheff and Markov inequalities provide a tighter bound when α gets small.

Simulation results are included. They show that the proposed procedure identifies an equalizer similar to the best achievable one in the mean square sense and that the bound on the probability of error can provide an estimate of the reached equalization accuracy.

APPENDIX: PROOF OF PROPOSITION 1

The following equation is easily derived from (8):

$$\left(\sum_k s^2(k)\right)^2 = \sum_k s^4(k) + \sum_k s^2(k) \sum_{j \neq k} s^2(j) = 1$$

which can be simplified by (9), leading to

$$\sum_k s^2(k) \sum_{j \neq k} s^2(j) = 1 - \alpha$$

By extracting the term $s^2(\bar{k})$, it can be rewritten as

$$2s^2(\bar{k}) \sum_{j \neq \bar{k}} s^2(j) + \sum_{k \neq \bar{k}} s^2(k) \sum_{j \neq k,\bar{k}} s^2(j) = 1 - \alpha \tag{17}$$

Obviously

$$\sum_{j \neq k, \bar{k}} s^2(j) \geqslant 1 - 2s^2(\bar{k}), \quad \forall k \tag{18}$$

Thus (17) and (18) entail the inequality

$$2s^2(\bar{k}) \sum_{j \neq \bar{k}} s^2(j) + \sum_{k \neq \bar{k}} s^2(k)(1 - 2s^2(\bar{k})) \leqslant 1 - \alpha$$

which can be reduced to

$$s^2(\bar{k}) \geqslant \alpha \tag{19}$$

If $\alpha > \frac{1}{2}$, it is then evident that no other coefficient but $s(\bar{k})$ has its square greater than α and consequently $s^2(\bar{k})$ is the only maximum value of the sequence $\{s^2(k)\}$.

From equation (17) we also have, by discarding the second term after the ' + ' sign,

$$2s^2(\bar{k})(1 - s^2(\bar{k})) \leqslant 1 - \alpha$$

which leads to

$$s^2(\bar{k}) \leqslant \frac{1 - \sqrt{(2\alpha - 1)}}{2} \quad \text{or} \quad s^2(\bar{k}) \geqslant \frac{1 + \sqrt{(2\alpha - 1)}}{2} \tag{20}$$

Since (19) and (20) have to be satisfied simultaneously, $s^2(\bar{k}) \leqslant [1 - \sqrt{(2\alpha - 1)}]/2$ in (20) must be discarded, so that we finally obtain

$$s^2(\bar{k}) \geqslant \frac{1 + \sqrt{(2\alpha - 1)}}{2}$$

which is the lower bound in equation (10).

The upper bound immediately follows from condition (9).

REFERENCES

1. Haykin, S. (ed.), *Blind Deconvolution*, Prentice-Hall, Englewood Cliffs, NJ, 1994.
2. Sato, Y., 'A method of self-recovering equalization for multilevel-amplitude modulation systems', *IEEE Trans. Commun.*, **COM-23,** 679–682 (1975).
3. Benveniste, A., M. Goursat and G. Ruget, 'Robust identification of a nonminimum phase system: blind adjustment of a linear equalizer in data communications', *IEEE Trans. Automatic Control*, **AC-25,** 385–399 (1980).
4. Godard, D.N., 'Self-recovering equalization and carrier tracking in two-dimensional data communication systems', *IEEE Trans. Commun.*, **COM-28,** 1867–1875 (1980).
5. Ding, Z., C. R. Johnson Jr. and R. A. Kennedy, 'Non-global convergence of blind recursive identifiers based on gradient descent of continuous cost functions', *Proc. 29th IEEE Conf. on Decision and Control*, Honolulu, HI, December 1990, IEEE, New York, 1990, pp. 225–230.
6. Shalvi, O. and E. Weinstein, 'New criteria for blind deconvolution of nonminimum phase systems (channels)', *IEEE Trans. Info. Theory*, **IT-36,** 312–321 (1990).
7. Jellonek, B. and K. D. Kammeyer, 'A closed form solution to blind equalization', *Signal Process.*, **36,** 251–259 (1994).
8. Jellonek, B. and K. D. Kammeyer, 'Eigenvector algorithm for blind equalization', *Proc. IEEE Signal Processing Workshop on HOS*, Lake Tahoe, CA, June 1993, IEEE, New York, 1993, pp. 19–23.
9. Proakis, J. G., *Digital Communications*, 3rd edn, McGraw-Hill, New York, 1995.
10. Kennedy, R. A., G. Pulford, B. D. O. Anderson and R. R. Bitmead, 'When has a decision-directed equalizer converged?', *IEEE Trans. Commun.*, **COM-37,** 879–884 (1989).
11. Ho, E. H. and Y. S. Yeh, 'A new approach for evaluating the error probability in the presence of intersymbol interference and additive Gaussian noise', *Bell Syst. Tech. J.*, **49,** 2249–2266 (1970).

12. Shimbo, O. and M. I. Celebiler, The probability of error due to intersymbol interference and Gaussian noise in digital communication systems', *IEEE Trans. Commun. Technol.*, **COM-19,** 113–119 (1971).
13. Benedetto, S., G. de Vincentiis and A. Luvison, 'Error probability in the presence of intersymbol interferece and additive noise for mutilevel digital signals', *IEEE Trans. Commun.*, **COM-21,** 181–190 (1973).
14. Saltzberg, B. R., 'Intersymbol interference error bounds with application to ideal bandlimited signaling', *IEEE Trans. Info. Theory*, **IT-14,** 563–568 (1968).
15. Lugannani, R., 'Intersymbol interference and probability of error in digital systems', *IEEE Trans. Info. Theory*, **IT-15,** 682–688 (1969).
16. Prabhu, V. K., 'Some considerations of error bounds in digital systems', *Bell Syst. Tech. J.*, **50,** 3127–3151 (1971).
17. Matthews, J. W., 'Sharp error bounds for intersymbol interference', *IEEE Trans. Info. Theory*, **IT-19,** 440–447 (1973).
18. Jenq, Y. C., B. Liu and J. B. Thomas, 'Probability of error in PAM systems with intersymbol interference and additive noise', *IEEE Trans. Info. Theory*, **IT-23,** 575–582 (1977).
19. Scholz, J. B., S. C. Cook and T. C. Giles, 'A scheme for high performance real-time BER measurement', *IEEE Trans. Commun.*, **COM-40,** 1574–1576 (1992).
20. Scholz, J. B. and B. R. Davis, 'Error probability estimator structures based on analysis of the receiver decision variable', *IEEE Trans. Commun.*, **COM-43,** 2311–2315 (1995).
21. Weinstein, S. B., 'Estimation of small probabilities by linearization of the tail of a probability distribution function', *IEEE Trans. Commun. Technol.*, **COM-19,** 1149–1155 (1971).
22. Newcombe, E. A. and S. Pasupathy, 'Error rate monitoring for digital communications', *Proc. IEEE*, **70,** 805–828 (1982).
23. Mendel, J. M., 'Tutorial on higher-order statistics (spectra) in signal processing and system theory: theoretical results and some applications', *Proc. IEEE*, **79,** 278–305 (1991).
24. Giannakis, G. B. and M. K. Tsatanis, 'A unifying maximum-likelihood view of cumulants and polyspectral measures for non-gaussian signal classification and estimation', *IEEE Trans. Info. Theory*, **IT-38,** 386–406 (1992).
25. Ljung, L. and T. Söderström, *Theory and Practice of Recursive Identification*, MIT Press, Cambridge, MA, 1983.
26. Papoulis, A., *Probability, Random Variables and Stochastic Processes*, McGraw-Hill, New York, 1984.
27. Shiryayev, A.N., *Probability*, Springer, New York, 1984.

