Adaptive Decision Feedback Equalization: Can You Skip the Training Period?

Joël Labat, Odile Macchi, *Fellow, IEEE,* and Christophe Laot

***Abstract*— This paper presents a novel unsupervised (blind) adaptive decision feedback equalizer (DFE). It can be thought of as the cascade of four devices, whose main components are a purely recursive filter $(\mathcal{R})$ and a transversal filter $(\mathcal{T})$. Its major feature is the ability to deal with severe quickly time-varying channels, unlike the conventional adaptive DFE. This result is obtained by allowing the new equalizer to modify, in a reversible way, both its structure and its adaptation according to some measure of performance such as the mean-square error (MSE). In the starting mode, $\mathcal{R}$ comes first and whitens its own output by means of a prediction principle, while $\mathcal{T}$ removes the remaining intersymbol interference (ISI) thanks to the Godard (or Shalvi–Weinstein) algorithm. In the tracking mode the equalizer becomes the classical DFE controlled by the decision-directed (DD) least-mean-square (LMS) algorithm. With the same computational complexity, the new unsupervised equalizer exhibits the same convergence speed, steady-state MSE, and bit-error rate (BER) as the trained conventional DFE, but it requires no training. It has been implemented on a digital signal processor (DSP) and tested on underwater communications signals—its performances are really convincing.**

***Index Terms*— Adaptive equalization, blind decision feedback equalization, blind deconvolution, blind equalization.**

I. Introduction

DECISION feedback equalizers (DFE's) are very often used to combat the distortion of communication channels because of their many advantages—even with severe and noisy channels, they can reach pretty good steady-state performance, e.g., a small output mean-square error (MSE), at a very much lower computational cost than other efficient techniques such as maximum-likelihood sequence estimation. Since the channels are unknown, the DFE must be implemented in an adaptive way. In a classical DFE, adaptation cannot be started without the transmission of a known training data sequence $d(k)$. This is the so-called training period. Then the proper transmission begins and the tracking of channel distortion is pursued by unsupervised[1] adaptation using the detected data $\hat{d}(k)$ in place of the true data $d(k)$. It is the decision-directed (DD) mode. Unfortunately, the channel can drastically change during the tracking period. As a consequence, a continued adaptation requires the periodic transmission of the training sequence, which decreases the effective bit rate and is not always possible. That is why unsupervised DFE's, without any training sequence, are necessary—adaptation starts as soon as the (unknown) data $d(k)$ are transmitted.

In the past decade many unsupervised adaptive algorithms have been designed for transversal equalizers [1]–[4] based on the assumption that the sequence $d(k)$ is zero-mean independent identically distributed (i.i.d.). Unfortunately, with a recursive equalizer such as a DFE, the phenomenon of error propagation restricts the use of unsupervised adaptation to the case of an initially well "open eye," corresponding to a mild channel. On the other hand, severe channels cannot be properly corrected with a transversal equalizer, especially when there is additive noise, hence the dilemma. There have already been a few approaches to solve it. In [5]–[7] the equalizer is split into a cascade of several linear filters, which involves at least one recursive filter. In particular, in [6] the following cascade has been proposed—an attenuation equalizer which is a purely recursive whitening filter $(\mathcal{R})$, a complex-valued gain control, and a phase equalizer which is an all-pass transversal/recursive filter. The complex gain control itself is decomposed into the cascade of a real gain control $(\mathcal{GC})$ and a phase rotator $(\mathcal{PR})$.

In this paper the new equalizer also appears as a four-linear-device cascade including $\mathcal{R}$, $\mathcal{GC}$, and $\mathcal{PR}$. For reasons that will be explained in the sequel, the all-pass transversal/recursive filter of [6] is replaced by a purely transversal filter $(\mathcal{T})$. The basic idea of the cascade is to split the difficult task of unsupervised equalization into several, but easier, subtasks. Furthermore, the structure itself is made adaptive and a performance index—the estimated MSE—is used to select the appropriate structure and its adaptation [8], [9]. The unsupervised starting period begins with $\mathcal{R}$ before $\mathcal{T}$. Once the eye is open, this equalizer is switched into a classical DFE simply by placing $\mathcal{R}$ after $\mathcal{T}$ and introducing the detection device into the recursive path of $\mathcal{R}$. Adaptation is then pursued in the classical DD mode. This method greatly improves the steady-state performances which reach those of the conventional trained DFE in terms of both residual MSE and bit-error rate (BER), without increasing the computational complexity. Moreover, the system easily switches back to the starting mode when a sudden change occurs in the channel. This is a very attractive feature.

Section II deals with the minimum MSE (MMSE) linear equalizer (LE) and brings a (recursive) solution to imple-

Paper approved by J. H. Winters, the Editor for Equalization of the IEEE Communications Society. Manuscript received August 6, 1996; revised July 24, 1997 and January 20, 1998.

J. Labat and C. Laot are with the Ecole Nationale Supérieure des Telcommunications de Bretagne, 29285 Brest Cédex, France (e-mail: joel.labat@enst-bretagne.fr).

O. Macchi is with the French National Center for Scientific Research, Laboratoire des Signaux et Systèmes, Ecole Supérieure d'Electricité, 91192 Gif Sur Yvette Cédex, France.

Publisher Item Identifier S 0090-6778(98)05167-8.

[1]The adjective "blind" is often used, but it has, in our mind, a pejorative connotation, so we avoid it.

ment it. Section III (resp. IV) investigates the starting (resp. tracking) period of the new equalizer. Section V (resp. VI) evidences the good behavior of the new DFE by means of computer simulations (resp. real underwater signals). In both cases the new DFE succeeds in tracking the time variations of the communication channel, whereas the conventional DFE is inefficient to deal with such impairments. Section VII presents our conclusions.

II. Recalls About the Linear MMSE Equalizer

The variance of the data $d(k)$ is denoted σ_d^2. The discrete channel $\mathcal{F}$ has a causal transfer function (TF) of order N denoted $F(z)$. The received signal $s(k)$ is corrupted by an independent additive zero-mean white noise $n(k)$ with variance σ_n^2. We recall [10] that the LE, which is optimal in the MSE sense, has TF

$$C(z) = \frac{\sigma_d^2 F^*(1/z^*)}{\sigma_d^2 F(z) F^*(1/z^*) + \sigma_n^2}. \tag{1}$$

The denominator of this expression is the power spectral density (PSD) of $s(k)$. It can be factorized as

$$D(z) = \sigma_d^2 F(z) F^*(1/z^*) + \sigma_n^2 = S G(z) G^*(1/z^*) \tag{2}$$

with S denoting a real positive number depending on σ_d^2, σ_n^2, and $F(z)$. Clearly, $D(z)$ has N pairs of roots $(z; 1/z^*)$, none of them being on the unit circle (U). So, its factorization is unique when $G(z)$ is constrained to be causal and minimum phase. If $z_i \ i = 1, 2, \cdots, N$ denotes the N roots of $D(z)$ located inside U, $G(z)$ is written

$$G(z) = \prod_{i=1}^{N} [1 - z_i z^{-1}] \quad \text{with} \quad |z_i| < 1. \tag{3}$$

As a result, the filter $\mathcal{R}$ with TF $1/G(z)$ is both causal and stable. Furthermore, when the signal $s(k)$ is filtered by $\mathcal{R}$, the output PSD has a constant value S, which means that $\mathcal{R}$ is a whitening filter—its output is an uncorrelated sequence. According to (1)–(3), the optimal LE can be viewed as the cascade of this purely recursive (all-pole) stable whitening filter $\mathcal{R}$ and a linear filter with TF

$$T(z) = [\sigma_d^2 / S] F^*(1/z^*) [G^*(1/z^*)]^{-1}. \tag{4}$$

The transfer functions $F^*(1/z^*)$ and $G^*(1/z^*)$ are anti-causal. As a consequence, the implementation of (4) requires a delay L and a transversal filter $\mathcal{T}_\mathcal{L}$, with TF

$$T_L(z) = z^{-L} \ \text{TSE}[T(z)] \tag{5}$$

where TSE stands for "truncated series expansion" of order L, in terms of the positive powers of z. Moreover, the filter $\mathcal{T}_\mathcal{L}$ can itself be split into a $\mathcal{G}C$ g, a $\mathcal{PR}$ $e^{-j\theta}$, and a transversal filter $\mathcal{T}$ as in [8] and [9]. Thus, up to a delay, the optimal LE takes the cascaded form

$$C = \mathcal{T}_\mathcal{L} \circ \mathcal{R} \qquad \mathcal{T}_\mathcal{L} = g e^{-j\theta} \mathcal{T}. \tag{6}$$

Clearly, the order of the four linear transformations $\mathcal{G}C$, $\mathcal{R}$, $\mathcal{T}$, and $\mathcal{PR}$ is irrelevant in steady state. Nevertheless, it turns out to be critical in an unsupervised mode. For reasons which will be developed later, it is recommended to plug $\mathcal{R}$ before $\mathcal{T}$, and $\mathcal{PR}$ downstream. The place of the $\mathcal{G}C$, though less critical, should be located in the vicinity of $\mathcal{R}$. So a possible form for the LE (1) corresponds to the cascade of Fig. 1. This is the architecture chosen for the starting mode and, more generally, for the difficult periods, e.g., after an abrupt channel change.

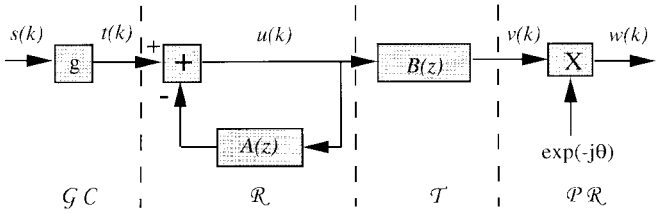


Fig. 1. Cascaded structure of the new equalizer in the starting period.

The Special Case of a Noiseless Channel $(\sigma_n^2 = 0)$*:* The TF $F(z)$ of $\mathcal{F}$ can be decomposed as

$$F(z) = f \prod_{i=1}^{N_1} (1 - z_{I,i} z^{-1}) \prod_{j=1}^{N_2} (z_{O,j}^{-1} - z^{-1}) \tag{7}$$

where $N = N_1 + N_2$, f can be complex, and the $z_{I,i}$ (resp. $z_{O,j}$) denotes the zeros of $F(z)$ inside (resp. outside) U

$$|z_{O,j}| > 1 \qquad |z_{I,i}| < 1. \tag{8}$$

With a slight loss of generality, it is assumed that $F(z)$ has no zeros on U. Then, $G(z)$ and $T(z)$ defined in (3) and (4), respectively, read

$$G(z) = \prod_{i=1}^{N_1} (1 - z_{I,i} z^{-1}) \prod_{j=1}^{N_2} (1 - (z_{O,j}^* z)^{-1}) \tag{9}$$

$$T(z) = \frac{1}{f} \prod_{j=1}^{N_2} \frac{1 - (z_{O,j}^* z)^{-1}}{z_{O,j}^{-1} - z^{-1}}. \tag{10}$$

In the abscence of noise, the MSE equalizer is a zero-forcing (ZF) device which merely inverts the channel

$$F(z) \frac{1}{G(z)} T(z) = 1. \tag{11}$$

According to (10), $T(z)$ is the TF of an all-pass filter TF. Hence, $\mathcal{R}$ compensates the channel attenuation distortion while $T(z)$ compensates the phase distortion introduced both by the channel $\mathcal{F}$ and the recursive filter $\mathcal{R}$. Unlike in [6], the filter of TF $T(z)$ is implemented by a transversal filter with TF

$$T_L(z) = \frac{z^{-L}}{f} \ \text{TSE} \left[\prod_{j=1}^{N_2} 1 - (z_{O,j}^* z)^{-1} \prod_{j=1}^{N_2} [z_{O,j}^{-1} - z^{-1}]^{-1} \right] \tag{12}$$

where L is the necessary delay, this series expansion being in terms of positive powers of z. So the MMSE LE is the cascade of a purely recursive amplitude equalizer $\mathcal{R}$ (with TF $1/G(z)$) and a phase equalizer which can itself be split into a gain control $\mathcal{G}C$, a transversal filter $\mathcal{T}$, and a phase rotator $\mathcal{PR}$.

III. Starting Period: $\mathcal{R}$ Before $\mathcal{T}$

A. Structure and Equations of the Cascade

In the starting period the equalizer is configurated as the cascade $\mathcal{GC}$, then $\mathcal{R}$, then $\mathcal{T}$, then $\mathcal{PR}$, as depicted in Fig. 1 where the notations are given for the signals at the various places in the cascade. However, the choice of the respective places for each part needs some explanation. First let us consider the noiseless channels. In this case it clearly appears that, when located before $\mathcal{T}$, $\mathcal{R}$ is an amplitute equalizer that compensates the channel attenuation distortion while $\mathcal{T}$ compensates the remaining phase distortion. If the criterion (25) is used for adaptation of $\mathcal{T}$, the uncorrelated sequence $u(k)$ is then transformed into an independent sequence $v(k)$. So the criteria are more and more demanding all along the cascade, which is a wise choice. It is obvious that permutating the order of the filters $\mathcal{R}$ and $\mathcal{T}$ would be meaningless. In addition, if the channel is minimum phase, then $\mathcal{R}$ totally removes the intersymbol interference (ISI) and, as a consequence, $\mathcal{R}$ must be located upstream. In the more general case of noisy channels the explanation is not so evident. However, if the noise level is weak, the situation is not far from the previous case. Besides, it is known [10] that the recursive part of the whitening filter is what the conventional DFE (optimal in the MMSE sense) needs in its feedback part, which is perhaps one of the main key points of the new unsupervised equalizer. Hence, it is natural to begin with whitening the sequence $s(k)$ first, i.e., to locate $\mathcal{R}$ upstream. In this case, since the sequence at the input of $\mathcal{T}$ is uncorrelated, then its correlation matrix is well conditioned, and so one can reasonably hope that the convergence speed of $\mathcal{T}$ is increased. This is the case in practice. The corresponding time equations are given below.

For $\mathcal{GC}$: In order to decrease the signal dynamical range, we shall implement it at the front edge. It is a one-coefficient real equalizer with input $s(k)$. It generates the output

$$t(k) = gs(k). \tag{13}$$

For $\mathcal{R}$: The input is $t(k)$ and the output is

$$u(k) = t(k) - \hat{t}(k) \tag{14}$$

where

$$\hat{t}(k) = \sum_{l=1}^{N} a_l u(k-l) = \boldsymbol{A}^T \boldsymbol{U}_N(k-1) \tag{15}$$

$$\boldsymbol{A} = [a_1, \cdots, a_{N-1}, a_N]^T \tag{16}$$

$$\boldsymbol{U}_N(k-1) = (u(k-1), \cdots, u(k-N))^T. \tag{17}$$

In (15) the signal $\hat{t}(k)$ is built up with past samples only. It can be viewed as a predicted value of the input $t(k)$ in such a way that $\mathcal{R}$ is an "innovator" which generates the prediction error $u(k)$ or "innovation" (14) contained in the new sample of $t(k)$. $\boldsymbol{A}$ is called the "prediction vector."

For $\mathcal{T}$: The input is $u(k)$ and the output is

$$v(k) = \sum_{l=0}^{L} b_l u(k-l) = \boldsymbol{B}^T \boldsymbol{U}_{L+1}(k) \tag{18}$$

where

$$\boldsymbol{B} = [b_0, b_1, \cdots, b_L]^T \tag{19}$$

$$\boldsymbol{U}_{L+1}(k) = (u(k), u(k-1), \cdots, u(k-L))^T. \tag{20}$$

For $\mathcal{PR}$: The input is $v(k)$ and the output is

$$w(k) = v(k) \exp(-j\theta). \tag{21}$$

Last in the cascade is the nonlinear decision device (slicer) that provides $\hat{d}(k)$, which is an estimate of $d(k-\delta)$, δ being a suitable delay.

B. Unsupervised Criteria of Optimality for the Four Devices

For $\mathcal{GC}$: It is used to fix the power level of the samples $u(k)$ at a particular value, for instance, σ_d^2

$$E\{|u(k)|^2\} = \sigma_d^2. \tag{22}$$

This equation constitutes an unsupervised criterion of optimality for the $\mathcal{GC}$.

For $\mathcal{R}$: Because $u(k)$ has a constant PSD σ_d^2, $\mathcal{R}$ is the unique stable recursive whitener of its input $t(k)$. Thus, the optimal prediction vector minimizes

$$I(\boldsymbol{A}) = E\{|u(k)|^2\} \tag{23}$$

and the sequence $u(k)$ is uncorrelated (white). This is an unsupervised criterion to optimize $\boldsymbol{A}$—although the transmitted data $d(k)$ are unknown, the recursive part $\mathcal{R}$ of the optimal MSE LE is exactly found.

For $\mathcal{T}$: There are a few unsupervised criteria, based on statistics of order higher than two, that allow (blind) deconvolution of nonminimum phase channels. For instance, the Godard criterion [1] minimizes

$$J_G(\boldsymbol{B}) = E\{[|v(k)|^p - R_p]^2\} \quad \text{with} \quad R_p = \frac{E\{|d(k)|^{2p}\}}{E\{|d(k)|^p\}} \tag{24}$$

In what follows we only investigate the classical case of $p = 2$. Then $R_2 = 1$ for the 4-quadrature amplitude modulation (QAM) scheme (resp. $R_2 = 1.32$ for the 16-QAM scheme), when considering unitary power data ($\sigma_d^2 = 1$) taking equiprobable values. Note that the criterion (24) is essentially the same as the criterion proposed by Shalvi and Weinstein [2], i.e.,

$$J_{SW}(\boldsymbol{B}) = E\{|v(k)|^4\} \quad \text{subject to} \quad E\{|v(k)|^2\} = \sigma_d^2. \tag{25}$$

The constraint in (25) can be achieved by forcing the vector $\boldsymbol{B}$ to have an unitary norm, from the moment that $u(k)$ is a white sequence of variance σ_d^2, which, unfortunately, is not true at the beginning of the equalization process. Both criteria will maximize the absolute (normalized) kurtosis of $v(k)$. Even with a noisy channel, if the additive noise is Gaussian, the optimum is reached if and only if the cascade is a ZF equalizer [2] (see Appendix A for more details). The resulting vector $\boldsymbol{B}$ might therefore be slightly different from the MSE solution which corresponds to the expansion of $T(z)$ in (4).

For $\mathcal{PR}$: To deal with the phase error brought by the channel and the demodulator, a phase rotation $e^{-j\theta}$ is required, where θ is an estimate of the phase error between the modulating and the demodulating carrier waves. To adjust θ, a possible criterion is to minimize the DDMSE

$$K(\theta) = E\{|v(k)e^{-j\theta(k-1)} - \hat{d}(k)|^2\}. \tag{26}$$

Since this criterion depends on the decisions $\hat{d}(k)$, the $\mathcal{PR}$ has to be located downstream in order not to disturb the previous stages of the equalizer. Furthermore, since the statistics of QAM or phase-shift keying (PSK) signals are unaffected by some particular rotations, it will always remain a phase indetermination in an unsupervised approach. This problem is classically solved by differential encoding.

The three last devices of the cascade have decoupled criteria in the sense that the criterion for $\mathcal{R}$ is unaffected by the parameters $\boldsymbol{B}$ and θ of the next two devices, and that the criterion for $\mathcal{T}$ is unaffected by the state θ of the $\mathcal{PR}$ which follows. This is why it is critical to plug $\mathcal{R}$ before $\mathcal{T}$ so that the difficult task of unsupervised equalization can be realized step by step— first the sequence is whitened by $\mathcal{R}$, then the remaining ISI is removed by $\mathcal{T}$.

C. Adaptation Algorithms

For $\mathcal{GC}$: It is easy to design an adaptive algorithm controlling g to satisfy (22), e.g.,

$$G(k) = G(k-1) + \mu_G[1 - |u(k)|^2] \tag{27}$$

$$g(k) = \sqrt{|G(k)|} \tag{28}$$

where $G(0) = 1$, and μ_G is a small positive step size. This is a kind of automatic gain control; see [11, ch. 1]. To implement the adaptive $\mathcal{GC}$, the gain must be in its $(k-1)$th state when the kth observation is processed

$$t(k) = g(k-1)s(k). \tag{29}$$

The same principle holds for all adaptive systems below.

For $\mathcal{R}$: There are a few algorithms available to minimize $I(\boldsymbol{A})$ in (23), e.g., the stochastic gradient algorithm whose increment is

$$\Delta\boldsymbol{A} = -(\mu_{\boldsymbol{A}}/2)\nabla_{\boldsymbol{A}}|u|^2, \qquad \mu_{\boldsymbol{A}} > 0. \tag{30}$$

Because $\mathcal{R}$ is a recursive filter, this last quantity cannot be exactly computed, but efficient approximate algorithms are available, for example,

$$\boldsymbol{A}(k) = \boldsymbol{A}(k-1) + \mu_{\boldsymbol{A}}u(k)\boldsymbol{U}_N^*(k-1) \tag{31}$$

$$u(k) = t(k) - \boldsymbol{A}(k-1)^T\boldsymbol{U}_N(k-1) \tag{32}$$

where $\boldsymbol{A}(0) = [0, 0, \cdots, 0]^T$. The superscript * stands for complex conjugate. Stability of this adaptive predictor has been investigated in detail in [11, ch. 15]. It is ensured by: 1) the presence of noise at the input which pushes the poles of $\mathcal{R}$ well inside U and 2) the so-called "self-stabilization" property of (31) which holds even in the absence of noise as proven in [11].

For $\mathcal{T}$: The stochastic gradient algorithm to minimize $J_G(\boldsymbol{B})$ in (24) is straightforward

$$\boldsymbol{B}(k) = \boldsymbol{B}(k-1) + \mu_{\boldsymbol{B}}v(k)(R_2 - |v(k)|^2)\boldsymbol{U}_{L+1}^*(k) \tag{33}$$

$$v(k) = \boldsymbol{B}^T(k-1)\boldsymbol{U}_{L+1}(k) \tag{34}$$

where $\boldsymbol{B}(0) = [0, 0, \cdots, 0, 1, 0, \cdots, 0]^T$ and $\mu_{\boldsymbol{B}}$ is a small positive step size.

The theoretical study of the criteria (24), (25) is still in progress in the open literature. However, it has been stated in [12] that, provided L is large enough and with an appropriate (center tap) initialization strategy, in both cases of Godard and Shalvi–Weinstein criteria, the equalizer $\mathcal{T}$ will converge to a stable equilibrium near a desired global minimum. Moreover, the convergence speed strongly depends on the value of the initial kurtosis ratio $K(v)/K(d)$, where $K(v)$ (resp. $K(d)$) denotes the normalized kurtosis of $v(k)$ (resp. $d(k)$). If this ratio is lower than 0.5, then the equalizer $\mathcal{T}$ will generally exhibit slower convergence toward a desired equilibrium. In the noiseless case or with sufficiently large signal-to-noise ratio (SNR) the corresponding algorithms are efficient to recover the true data $d(k)$ up to an arbitrary phase θ and an unknown delay δ. So in steady state

$$v(k) \approx d(k-\delta)\exp(j\theta). \tag{35}$$

For $\mathcal{PR}$: Classical algorithms to minimize $K(\theta)$ in (26) can be found in [14]. For instance,

$$w(k) = v(k)\exp(-j\theta(k-1)) \tag{36}$$

$$\theta(k) = \theta(k-1) + \mu_\theta\left(\varepsilon(k) + \beta\sum_{i=1}^{k}\varepsilon(i)\right) \tag{37}$$

$$\varepsilon(k) = \mathrm{Im}\{w(k)[\hat{d}(k) - w(k)]^*\} \tag{38}$$

which corresponds to a second-order phase tracking loop. In (37) and (38) μ_θ is a small positive step size, β an appropriate positive parameter, and $\theta(0) = 0$.

The theoretical analysis of the new equalizer (convergence and steady state) is simplified by the decoupled character of the criteria controlling $(g, \boldsymbol{A})$, $\boldsymbol{B}$, and θ, as pointed out in the previous subsection. Convergence of $g(k)$ and $\boldsymbol{A}(k)$ can be studied first, $\boldsymbol{B}(k)$ and $\theta(k)$ being then irrelevant. Then, once $g(k)$ and $\boldsymbol{A}(k)$ have reached their optimal values $\tilde{g}$ and $\tilde{\boldsymbol{A}}$, convergence of $\boldsymbol{B}(k)$ (see [12] for criteria (24) and (25)] can be investigated independently of $\theta(k)$. Then convergence of $\theta(k)$ can be studied with $g(k)$, $\boldsymbol{A}(k)$, and $\boldsymbol{B}(k)$ in their respective optimal states $\tilde{g}$, $\tilde{\boldsymbol{A}}$, and $\tilde{\boldsymbol{B}}$. In other words, one can investigate separately the adaptation algorithms of the pair $(g(k)\boldsymbol{A}(k))$, the vector $\boldsymbol{B}(k)$, and the angle $\theta(k)$. As for $\mathcal{GC}$, adaptation of its gain $g(k)$ remains coupled with $\mathcal{R}$ that is plugged downstream [see (22) and (23)]. (We recall that the coupling originates in the fact that the $\mathcal{GC}$ is plugged before $\mathcal{R}$ in order to reduce the signal dynamic range in the implementation). This coupling makes the theoretical investigation more intricate but, in practice, it raised no specific difficulty.

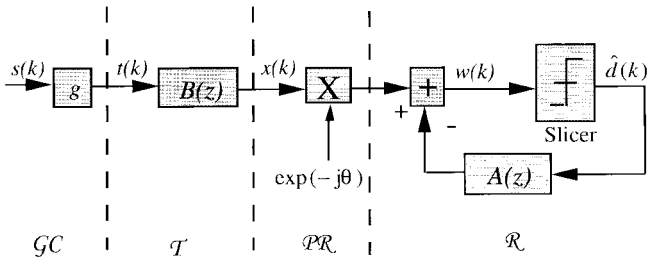


Fig. 2. Structure of the equalizer in the tracking period—classical DFE.

Computational Complexity: The equations governing a classical DFE equipped with a $\mathcal{PR}$ are given in the following subsection [see (41)–(49)]. Comparing these equations with (27)–(29), (31)–(34), and (36)–(38), it appears that the new structure in its starting mode has exactly the same computational complexity, at least when the classical DFE is preceded by an adaptive gain control (which is a conservative practice).

IV. Tracking Period: $\mathcal{T}$ Before $\mathcal{R}$

A. Structure and Switching Rule

During the first running mode, the system can be viewed as a ZF equalizer including a whitening all-pole filter whose recursive part is precisely the feedback filter of the recursive LE (see Section II) as well as the one of the conventional DFE [10]. That is why when the decisions $\hat{d}(k)$ become reliable, which is deduced from the index of performance defined below, the equalizer structure is modified into the cascade $\mathcal{GC}$, then $\mathcal{T}$, then $\mathcal{PR}$, and lastly $\mathcal{R}$, which is now fed in by the detected data, as depicted by Fig. 2. Clearly this structure is merely a DFE equipped with a front edge $\mathcal{GC}$ and a $\mathcal{PR}$ to ensure the carrier tracking. It is the structure adopted for the new equalizer in the tracking period. When modifying the relative place of $\mathcal{T}$ and $\mathcal{R}$ and feeding $\mathcal{R}$ by the past decisions in order to get the classical DFE, the system can only suffer from the fact that the transversal part $\mathcal{T}$ of the equalizer has (nearly) reached a ZF solution instead of the MSE solution. As a consequence, minimization of the DDMSE can yield a (slight) change in parameter $\boldsymbol{B}$ only. Nevertheless, with large SNR, the change is not really significant, so the modification of structure is justified. In this way the new DFE obviously yields a smaller MSE than the linear recursive ZF equalizer of Section III.

To control the running mode (starting or tracking mode), some performance measure is required, e.g., an estimation of the (true) output MSE, say M. Since the true data are unknown, they are replaced by $\hat{d}(k)$ at the slicer output and the performance measure is the DDMSE (26). This is a normalized performance measure since $d(k)$ has unit power. Besides, it can be estimated using the recursive formula

$$M_{\mathrm{DD}}(k) = \lambda M_{\mathrm{DD}}(k-1) + (1-\lambda)|\hat{d}(k) - w(k)|^2 \quad (39)$$

where λ, the so-called "forgetting factor," is taken at 0.99 in the sequel. Clearly the eye is open when M is low ($M < M_0$). In this case M and M_{DD} are essentially the same. Hence, the rule for modifying the structure

$$\begin{aligned} &M_{\mathrm{DD}}(k_0) \geq M_0 \text{ starting mode for } k > k_0 \\ &M_{\mathrm{DD}}(k_0) < M_0 \text{ tracking mode for } k > k_0. \end{aligned} \quad (40)$$

To ensure a safe transition in the tracking mode (DFE), the threshold M_0 must be reasonably small. So what we really have to do is to choose a threshold M_0 corresponding to a sufficiently low BER (typically less than 0.02) in order to avoid a pathological behavior of the DFE. But since, in the starting mode, the equalizer is ZF, the BER can be expressed as a function of the (true) MSE. Hence, the appropriate threshold M_0 can be chosen, depending on the modulation scheme. For example, in the 4-QAM case, this leads to $M_0 = 0.25$ (-6 dB). Note that, from this threshold M_0, the DDMSE becomes a good estimation of the true MSE and, consequently, a good index of the equalizer performance.

B. Optimality Criterion and Adaptation

The DDMSE criterion in (26) is applicable to optimize the parameters of this classical DFE. Then $\mathcal{GC}$ and $\mathcal{T}$ are redundant, so the coefficient $g(k)$ can be held at a fixed value g, so

$$t(k) = gs(k). \quad (41)$$

The notations are given in Fig. 2. The eight following equations are those of a classical DFE including a $\mathcal{PR}$. The output is

$$w(k) = y(k) - \boldsymbol{A}^T(k-1)\hat{\boldsymbol{D}}(k) \quad (42)$$

where

$$y(k) = [\boldsymbol{B}^T(k-1)\boldsymbol{T}(k)]\exp(-j\theta(k-1)) \quad (43)$$

$$\boldsymbol{T}(k) = [t(k), \cdots, t(k-L)]^T \quad (44)$$

$$\hat{\boldsymbol{D}}(k) = [\hat{d}(k-1), \hat{d}(k-2), \cdots, \hat{d}(k-N)]^T. \quad (45)$$

Adaptation: The simplest updating algorithm is the stochastic gradient of the criterion (26)

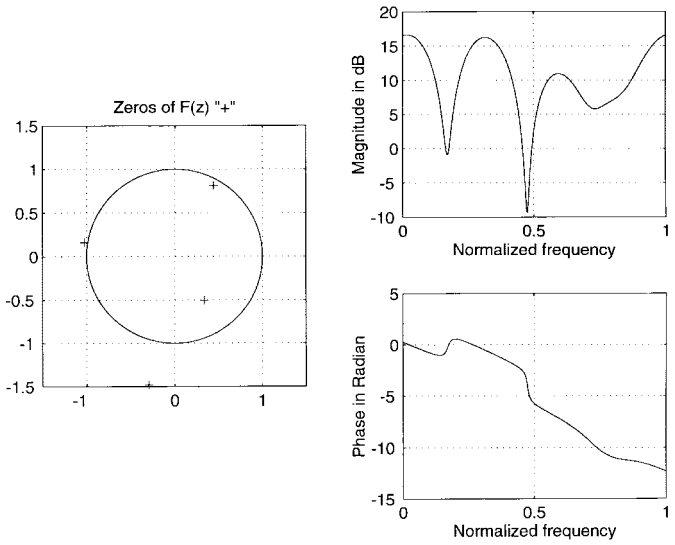
$$\boldsymbol{B}(k) = \boldsymbol{B}(k-1) + \mu_{\boldsymbol{B}}[\hat{d}(k) - w(k)]\exp(j\theta(k-1))\boldsymbol{T}^*(k) \quad (46)$$

$$\boldsymbol{A}(k) = \boldsymbol{A}(k-1) - \mu_{\boldsymbol{A}}[\hat{d}(k) - w(k)]\hat{\boldsymbol{D}}^*(k) \quad (47)$$

$$\theta(k) = \theta(k-1) + \mu_\theta\left(\varepsilon(k) + \beta\sum_{i=1}^{k}\varepsilon(i)\right) \quad (48)$$

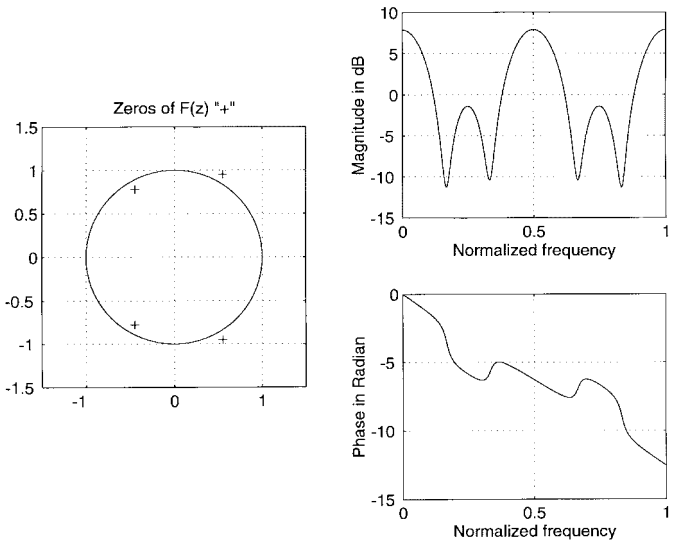
$$\varepsilon(k) = \mathrm{Im}\{y(k)[\hat{d}(k) - w(k)]^*\}. \quad (49)$$

The performance measure M_{DD} is permanently calculated along (39) and tested according to (40). If a sudden change occurs at step k_0 in the channel, the eye gets closed and M_{DD} oversteps M_0. The switchback in starting mode is straightforward. For the $\mathcal{GC}$, (27)–(29) are started with $G(k_0-1) = g^2$. For $\mathcal{R}$, (31) and (32) are started with the initial value $\boldsymbol{A}(k_0-1)$ and with $\boldsymbol{U}_N(k_0-1) = \boldsymbol{0}$. Similarly for $\mathcal{T}$, (33) and (34) are started with $\boldsymbol{B}(k_0-1)$ and with $\boldsymbol{u}_{L+1}(k_0-1) = \boldsymbol{0}$. The $\mathcal{PR}$ is started with $\theta(k_0-1)$.



(a) (b)

Fig. 3. (a) Zeros of $F_1(z)$. (b) Attenuation and phase responses.



(a) (b)

Fig. 4. (a) Zeros of $F_2(z)$. (b) Attenuation and phase responses.

V. Computer Simulations

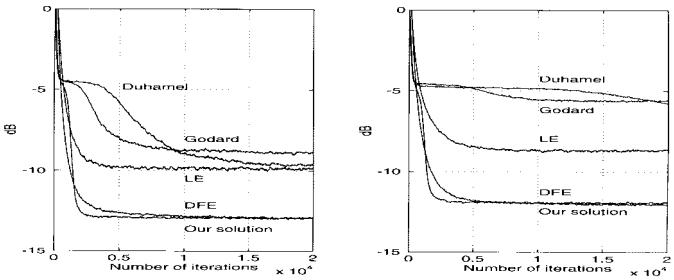
A. Stationary Channels

With such channels, the equalizer achievements can be characterized in terms of convergence speed and steady-state error. Results have been obtained via Monte Carlo simulations using 200 different runs and two severe channels. Both have nonminimal phase. Channels 1 and 2 have been proposed in [3] and [5], respectively. Their impulse responses are

$$\text{Channel 1:}\quad \boldsymbol{f}_1 = [2-0.4j, 1.5+1.8j, 1, 1.2-1.3j, 0.8+1.6j]$$
$$\text{Channel 2:}\quad \boldsymbol{f}_2 = [0.8264, -0.1653, 0.8512, 0.1636, 0.81]$$

Fig. 3(a) [resp. Fig. 4(a)] depicts the location of the TF zeros of channel 1 (resp. 2) and gives the amplitude and phase responses of these two channels. Note that both channels exhibit deep fading frequencies as well as severely nonlinear phase distortion, as a result of the zeros of $F(z)$ which are outside U and close to it. For such severe channels, most unsupervised equalizers proposed in the literature do not succeed to open the eye.

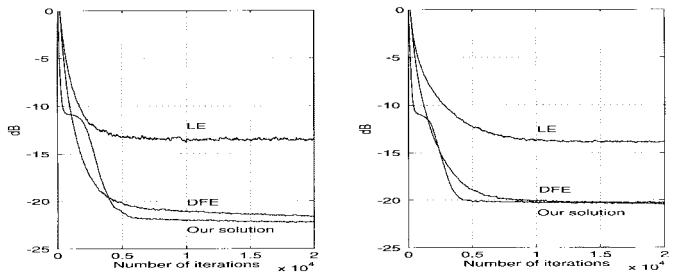
Five adaptive equalizers have been tested. Three of them are transversal LE's with 31 taps, started with a centered reference tap—the Godard unsupervised equalizer [1], the Duhamel and Hilal one [4], and the classical trained LE. The two recursive equalizers are the trained DFE and the new unsupervised DFE of this paper. Both recursive equalizers have 5 and 20 taps in their recursive and transversal parts, respectively. For the 16-QAM scheme, the Duhamel and the Godard unsupervised algorithms are totally inefficient for these two severe channels. So our system is only compared to the trained LE and DFE. To compute the DDMSE in (39), we start with $M_{\mathrm{DD}}(0) = 1$, while the switching threshold M_0 is 0.25 (-6 dB) for the 4-QAM and 0.063 (-12 dB) for the 16-QAM. It means that the system starts with the appropriate structure—$\mathcal{R}$ comes first. The equalizer input SNR for the 4- and 16-QAM schemes is 15 and 25 dB, respectively. The impulse responses are normalized to unity. The additive noise is zero mean, white, and Gaussian, whereas the channel phase-shift standing for the demodulation phase error is $\phi(n) = \pi n \Delta F T + \Phi$, T denoting the symbol duration. In the 200 trials, Φ is uniformly distributed in $[0, 2\pi]$ and the normalized frequency offset $\Delta F T$ is uniformly distributed in $[0, 1e-3]$ for 4-QAM and in $[0, 1e-4]$ for 16-QAM. Then the DDMSE is estimated in (39) for every experiment, then averaged over the 200 trials and plotted in Figs. 5 and 6 versus the number of iterations.



(a) (b)

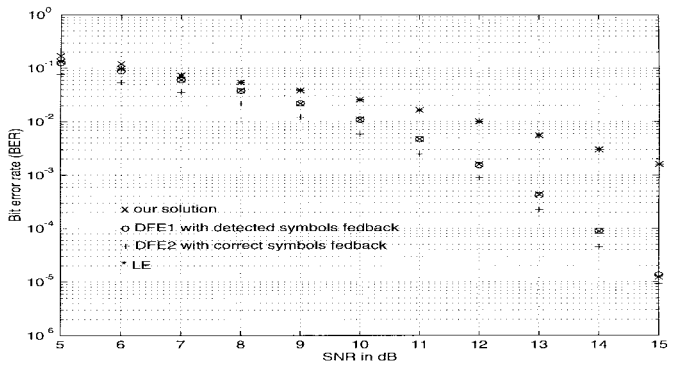
Fig. 5. Case of 4-QAM—compared performance of the new equalizer with other unsupervised equalizers (Godard and Duhamel) or trained equalizers (LE and DFE). (a) Channel 1, MSE ($\mathrm{SNR} = 15$ dB). (b) Channel 2, MSE ($\mathrm{SNR} = 15$ dB).

The striking important conclusion drawn from the simulations of Figs. 5 and 6 is that, with the same step size, our new unsupervised system achieves essentially the same steady-state MSE as the best trained equalizer, which is the DFE. In the 16-QAM case and with channel 1, we even note a 1-dB difference at iteration 20 000— it means that the final value of the MSE is reached sooner with the new equalizer. The transversal equalizers are totally outperformed (3–9 dB of output DDMSE gain) because they are not matched to such severe channels. Convergence of the new equalizer, respectively, requires less than 2000 and 5000 iterations, in the 4- and 16-QAM cases. It is very fast. The BER is plotted

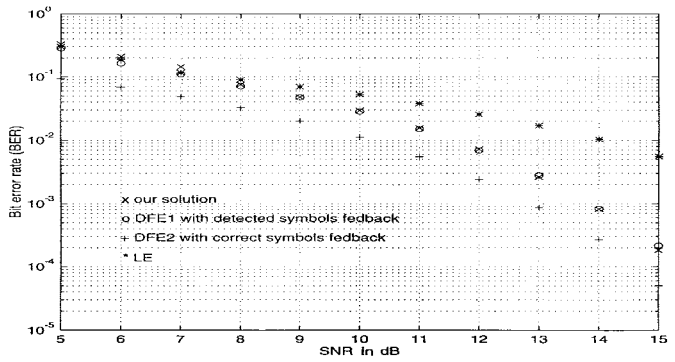


(a) (b)

Fig. 6. Case of 16-QAM—compared performance of the new unsupervised equalizer with trained LE and DFE equalizers (a) Channel 1, MSE ($\mathrm{SNR} = 25$ dB). (b) Channel 2, MSE ($\mathrm{SNR} = 25$ dB).



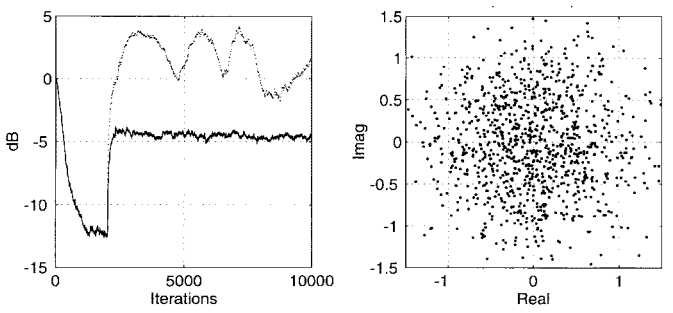
(a)



(b)

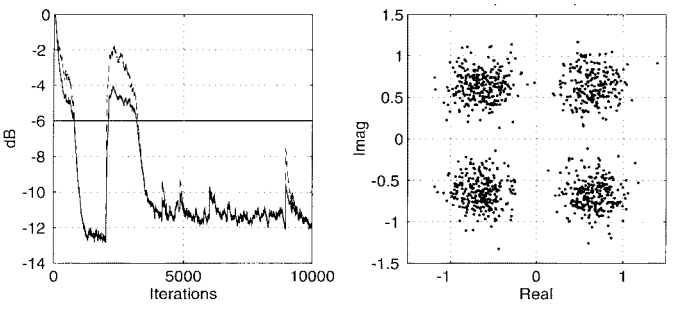
Fig. 7. (a) Compared BER—case of channel 1 (4-QAM). (b) Compared BER—case of channel 2 (4-QAM).

in Fig. 7, versus the SNR, with a 4-QAM scheme, for a trained LE $(*)$, a trained DFE $(\circ)$, and the new equalizer $(\times)$. As a performance upper bound, we have indicated the BER of a "supervised" DFE $(+)$ whose recursive path and adaptation are controlled with the true data (this DFE stands as a reference). For an SNR below 7 dB, all equalizers perform essentially the same way, that is, bad, with BER higher than 0.1. For an SNR higher than 7 dB, the trained DFE and the new one have the same achievement in BER, much better than the trained LE. They display only a slight loss with respect to the optimal supervised DFE. For instance, with channel 1 and at a BER of 10^{-4}, the loss is about 0.5 dB. Therefore, the new equalizer appears really attractive in terms of both convergence speed and steady-state performances (MSE and BER).



(a) (b)

Fig. 8. The classical trained DFE faced to a sudden and permanent change in the channel. True MSE—dotted line; DDMSE—solid line. (a) MSE. (b) Equalizer output.



(a) (b)

Fig. 9. The novel unsupervised DFE faced to a sudden and permanent change in the channel. True MSE—dotted line; DDMSE—solid line. (a) MSE. (b) Equalizer output.

B. Nonstationary Channels

With such channels one can investigate the tracking achievement of the equalizer, when facing a permanent time variation. In the simulations we begin with a fixed maximal phase channel $\mathcal{F}_o$ whose TF has one zero at $z_{O,1} = 1.1$. Then the channel becomes time-varying $\mathcal{F}(k)$ with a second mobile zero $z_2(k)$ appearing after 2000 iterations, with, for $k \geq 2000$

$$z_2(k) = \exp(j2\pi/3) + 0.1\exp(j2\pi 10^{-4}(k - 2000)). \quad (50)$$

This zero is located half-time inside and half-time outside U. Both zeros are close to U, so the time-varying channel $\mathcal{F}(k)$ is very severe. In fact, $z_2(k)$ corresponds to the classical differential Doppler effect—a new path suddenly arrives on the receiver at time $k \geq 2000$.

In the case of 4-QAM modulation, Fig. 8 (resp. 9) displays the behavior of the classical trained DFE and of the new unsupervised DFE. The DDMSE (solid line) and true MSE (dotted line) are depicted in Figs. 8(a) and 9(a), and the equalizer output diagram around iteration $k = 5000$ is shown in the Figs. 8(b) and 9(b). It appears that M_{DD} is a good estimate of the true MSE as soon as $M_{\mathrm{DD}} \leq -6$ dB. Due to training, the classical DFE begins with successfully compensating the fixed channel $\mathcal{F}_o$. However, when the channel is suddenly modified at time $k_0 = 2000$, it is unable to reopen the eye without a new training sequence. On the other hand, the new DFE takes approximately 850 iterations (in the starting mode) to correct the fixed (maximal phase) channel $\mathcal{F}_o$. Then it switches its structure in the tracking mode but at time $k_0 = 2000$, where

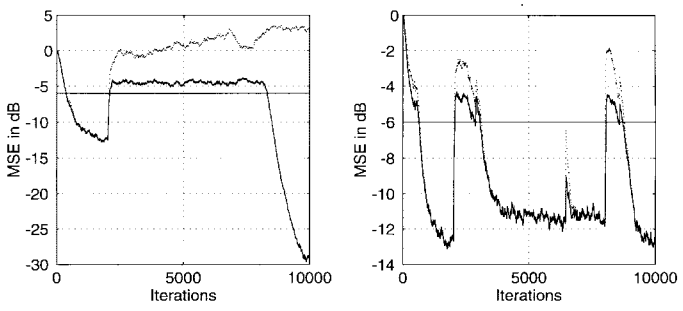


Fig. 10. Two sudden changes plus a drift in between them. True MSE—dashed line; DDMSE—solid line. (a) Trained DFE. (b) Novel unsupervised DFE.

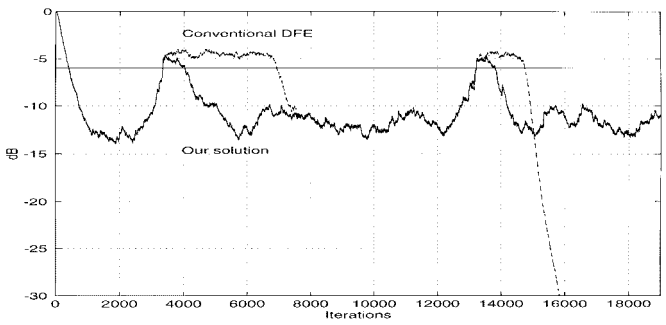


Fig. 11. MSE. Successful transmission through the shallow water channel with the novel unsupervised DFE.

the channel is suddenly modified, it immediately switches back to the starting mode. Without a new training, this DFE takes 1000 iterations to reopen the eye, then it switches again to the tracking mode and is able to properly track the permanent time variations created by the drifting zero $z_2(k)$ in (50). The diagram in Fig. 9(b) shows that, in steady state, despite the severity of this channel, most of the ISI is removed.

In order to evidence still better the ability of the new DFE to correct fast variations of the channel, we have made another sudden change at iteration $k_1 = 8000$ by canceling the zero $z_2(k)$ in (50). As shown in Fig. 10(a), this second change is well accepted by the novel DFE, which switches back to the starting mode. It takes 850 iterations to reopen the eye and then switches again to the tracking mode. On the other hand, the classical DFE [Fig. 10(b)] gets completely lost at this second change—its parameters converge toward a false equilibrium state for which the equalizer output $w(k)$ is locked onto two periodical values of the data diagram. The previous simulations are very convincing of the superiority of this novel DFE. It does work without training even with severe and fast-changing channels. Its achievement will now be illustrated in the case of a real (underwater) digital transmission.

VI. Underwater Acoustic Channels

In the field of underwater acoustic communications, digital communications systems generally use binary PSK (BPSK) and quadrature PSK (QPSK) signaling [13]. They correct multipath effects and Doppler frequency shifts by means of an adaptive trained DFE. The synchronization (carrier and timing recovery) is jointly optimized with the DFE according to some appropriate criterion. For the vertical underwater acoustic channel, this solution (DFE) yields an acceptable spectral efficiency of the order of 2 b/s/Hz. The channel impulse response is usually very short, so the equalizer only needs a small number of taps.

Unfortunately, the trained DFE is inefficient, unless it is periodically retrained, for digital communications through a shallow water channel. This latter turns out to be among the most severe channel encountered in practice and mainly suffers from two kinds of impairments. The first one is the strong channel distortion due to multipath propagation because of multiple scattering by surface and bottom. Since the delay spread may reach tens of symbol durations, the number of taps of the equalizer increases drastically, leading to a computational complexity that can severely limit the achievable data rate. This is an obvious reason to choose a recursive equalizer, regardless of stability consideration. The second strong impairment is the fast nonstationarity of the channel which originates in the relative motion between the transmitter and receiver and in the waves at the sea surface. This yields a Doppler frequency shift on the carrier and a differential Doppler effect on the various multipaths which makes the channel filter quickly changing. It also renders optimal sampling more difficult. From this point of view, the new equalizer turns out to be particularly well suited because of its ability to deal with rapidly varying channels by modifying both its optimality criteria and its structure when necessary. Two main solutions can be thought of to solve this problem. The first one is carried out by means of maximum-likelihood sequence estimation, but is not well suited here because of the impulse response length that would imply a prohibitive computational complexity. Up to now, the second solution was the periodically trained DFE.

In this section we present a shallow underwater communication system based on the new unsupervised DFE, where the modulation scheme is a 4-QAM and the bit rate is 6 kb/s.The carrier frequency is 12 kHz and the transducer bandwidth is close to 4 kHz. The experiment was carried out in the bay of Brest (France). The transmitter was located on the bottom of the sea and the transducer was kept under the boat at a depth of 5 m below the sea surface. The transducer output signal was first amplified, filtered, and then recorded on a digital audio tape during a few hours. In this way we obtained files to test the new equalizer. In addition, the whole receiver has been implemented on a digital signal processor (Motorola DSP 56 002) and its behavior observed during the few hours of recording. There are $N = 29$ (resp. $L + 1 = 20$) taps in the feedback (resp. forward) part of the DFE with the step sizes $\mu_{\boldsymbol{A}} = \mu_{\boldsymbol{B}} = 0.003$, while the step size is $\mu_G = 0.01$ for the $\mathcal{G}C$. Fig. 11 displays the evolution of the estimated MSE $M_{\mathrm{DD}}(k)$ for both the conventional trained DFE and the new unsupervised DFE. The latter switches for the first time in tracking mode after about 500 iterations, which is very fast for an unsupervised equalizer with this kind of channel. After approximatively 3300 iterations, $M_{\mathrm{DD}}(k)$ oversteps the threshold M_0. So the new DFE switches back to the starting mode, where it remains for approximately 500 iterations. In

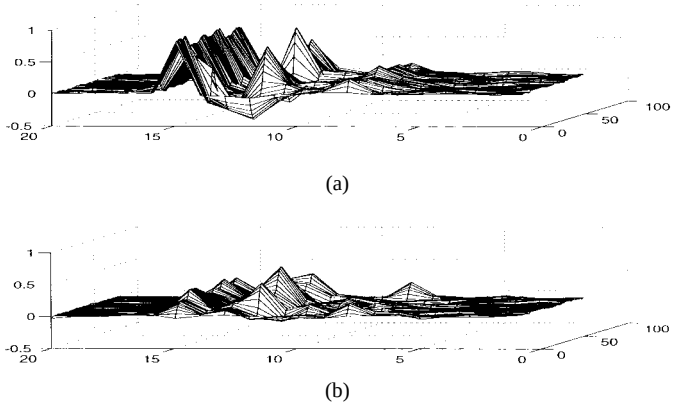


Fig. 12. Time evolution of the shallow water channel response.

this way it is able to deal with this sudden change of the channel characteristics.

On the other hand, the classical DFE cannot be started without a training sequence and so would not work with this file. To have an idea of its behavior, we have started the DFE with our structure and we have inhibited any further switchback from the tracking mode to the starting mode. This is essentially equivalent to a trained classical DFE. From iteration 3300, it gets lost and remains in this state during approximately 4000 iterations, i.e., 1.3 s in this example. Then it appears as very lucky to converge again toward the right solution and so ensures a correct transmission during 6300 iterations. Unfortunately, from iteration 13 000 until the end of the file, it remains irremediably lost—the estimated MSE is near zero because the output is locked onto a periodic sequence in the data diagram and this pathological state is stable. Such a behavior has often been observed in this channel. In such conditions a conventional DFE has to be periodically trained, whereas the new DFE remains efficient with only a few hundreds of false bits during the starting periods and a few isolated false bits in the tracking periods. Such an achievement is really convincing! Besides, the whole impulse response, including transmitter and receiver filters as well as transducer and channel, has been identified by estimating, with an adaptive algorithm, the filter $\hat{d}(k) \rightarrow s(k)\exp(-j\theta(k-1))$. The plot in Fig. 12 displays this response every 250 T and for a file duration of 6 s (76 successive responses). It shows that the channel response typically spreads over ten symbol durations, which is consistent with the fact that the receiver was close to the transmitter (50 m for the studied file). Fig. 12 evidences a first transmission path that is approximately constant and a second one that is delayed by $2T$ with time-varying amplitude (real and imaginary parts). So this real channel is not very far from the time-varying simulation model adopted in subsection V-B with the TF having a fixed zero $z_{O,1}$ and a mobile periodic zero $z_2(k)$. Even the period of 10^4T chosen in the simulations for the channel variations appears consistent with the period that can be appreciated on the basis of Fig. 12 in the real underwater experiment. Despite a delay spread that may reach 40 symbol durations (more than 10 ms) in the worst case, the new system never got lost during 2 h of transmission and with a distance from transmitter to receiver ranging from a few meters up to about 1 km. The new DFE is a fully satisfactory system in practice.

VII. Conclusion

The new unsupervised DFE introduced in this paper is based on the decomposition of the equalizer into a cascade of four devices, whose major components are a recursive whitening filter $(\mathcal{R})$ and a transversal filter $(\mathcal{T})$. The specificity is that the positions of these two devices can be switched on and back in a reversible way, according to some estimation of the output MSE. In the starting period (large MSE) the equalizer is linear—$\mathcal{R}$ comes first and whitens its own output, whereas $\mathcal{T}$ removes the remaining ISI (ZF equalizer). In the tracking mode (small MSE) $\mathcal{R}$ comes last and its loop is fed in with the detected data, so the equalizer becomes the classical DFE. This novel equalizer is very well suited for both severe stationary channels as well as quickly time-varying channels—where a conventional DFE would suffer from error propagation and bursting due to a sudden change, the new DFE switches back to its starting mode. On the other hand, it is worth emphasizing that this unsupervised DFE not only has the same steady-state performance as a usual trained DFE but has approximately the same convergence speed. Moreover, its computational complexity is not increased. As a consequence, a wide field of applications can be thought of, such as radiowave transmission, TV broadcasting [15], voice band modems, asymmetrical digital subscriber line (ADSL), and so on.

Hence, the conclusion—adaptive DFE can now skip the training period and receive a permanent digital flow without periodical retraining, even in a severely nonstationary context. Moreover, the performances can be improved by implementing $\mathcal{T}$ as a fractionally spaced equalizer during the tracking period. In addition, at the price of an increasing complexity, an recursive-least-squares (RLS) algorithm can be used to update $\mathcal{R}$ during the starting periods, and both $\mathcal{T}$ and $\mathcal{R}$ in the tracking periods.

Appendix A
Recall About the Shalvi and Weinstein Theorem [2]

If $x(k)$ denotes a zero-mean i.i.d. input sequence and $\boldsymbol{f}$ denotes the vector corresponding to the impulse response of a noiseless system (for instance channel + equalizer) with output $\hat{x}(k)$, it then comes, according to the above-mentioned theorem, that the normalized kurtosis of the output $\hat{x}(k)$ can be expressed as

$$K(\hat{x}) = K(x)\left[\frac{||\boldsymbol{f}||_4}{||\boldsymbol{f}||_2}\right]^4 \quad \text{where} \quad ||\boldsymbol{f}||_p = \left[\sum_k |f_k|^p\right]^{1/p}.$$

The main conclusions derived from this result are

$$|K(\hat{x})| \leq |K(x)|$$
$$|K(\hat{x})| = |K(x)| \quad \text{if and only if}$$
$$\boldsymbol{f} = (\cdot,\cdot,0,0,\rho e^{j\theta},0,0,\cdot,\cdot)^T.$$

This means that the maximization of $|K(\hat{x})|$ is equivalent to the deconvolution (ZF) of the system (with impulse response $\boldsymbol{f}$).

As a consequence, in the starting mode the new unsupervised equalizer is a ZF one, when using (25). Since the transversal part $\mathcal{T}$ is located downstream, this criterion will give a solution which inverts the linear cascade $\mathcal{F}$–$\mathcal{GC}$–$\mathcal{R}$. This remains valid for Gaussian channels as well. This results from the fact that, with the notations used in our paper, $K(v) = K_S(v)$, where $K_S(v)$ denotes the signal contribution kurtosis. This is because, by assumption, signal and noise are independent processes and, moreover, the noise is Gaussian.

Note that, according to [12], what is important for the convergence of the Godard and Shalvi–Weinstein algorithms is the initial (positive) kurtosis ratio

$$\frac{K(\hat{x})}{K(x)} = \left[\frac{\|\boldsymbol{f}\|_4}{\|\boldsymbol{f}\|_2}\right]^4$$

with $\boldsymbol{f}$ denoting the initial impulse response of the system (channel + equalizer), which is nothing but the impulse response of the channel when a center tap initialization strategy is used. When this ratio is greater than 0.5, then, for a small step size and a sufficient length, the equalizer will converge to a stable equilibrium near a desired global minimum [12]. Otherwise, it will generally take a longer time, subject to an appropriate initialization strategy, to converge to a desired equilibrium. For the two channels of Section V-A, the initial kurtosis conditions are not satisfied. The kurtosis ratios are 0.2375 for channel 1 and 0.3174 for channel 2. They clearly are lower than 0.5, which explains that the Godard or Shalvi–Weinstein linear equalizers take a long time to converge to a desired equilibrium. However, despite the severity of the two proposed channels, the convergence of the new equalizer is very fast.

REFERENCES

[1] D. N. Godard, "Self-recovering equalization and carrier tracking in two dimensional data communication system," *IEEE Trans. Commun.*, vol. COM-28, pp. 1867–1875, Nov. 1980.

[2] O. Shalvi and E. Weinstein, "New criteria for blind deconvolution of nonminimum phase systems (channels)," *IEEE Trans. Inform. Theory*, vol. 36, pp. 312–321, Mar. 1990.

[3] B. Porat and B. Friedlander, "Blind equalization of digital communication channels using high-order moments," *IEEE Trans. Signal Processing*, vol. 39, pp. 522–526, 1991.

[4] K. Hilal and P. Duhamel "A blind equalizer allowing soft transition between the CMA and the DD algotithm for PSK modulated signals" in *Proc. ICC*, Geneva, Switzerland, May 1993, pp. 1144–1148.

[5] O. Macchi, C. A. F Da Rocha, and J. M. T. Romano, "Egalization adaptative autodidacte par rétroprédiction et prédiction," in *14th GRETSI Symp.*, Juan-Les-Pins, France, 1993, pp. 491–494.

[6] C. A. F Da Rocha, O. Macchi, and J. M. T. Romano, "An adaptive nonlinear IIR filter for self-learning equalization," in *ITC 94*, Rio de Janeiro, Brazil, 1994, pp. 6–10.

[7] C. A. F. Da Rocha and O. Macchi,, "A novel self-learning adaptive recursive equalizer with unique optimum for QAM," in *Proc. ICASSP 94*, vol. III, Adalaïde, South Australia, Apr. 1994, pp. 481–484.

[8] J. Labat, C. Laot, and O. Macchi, "Dispositif dégalization adaptatif pour systèmes de communications numériques," French Patent 9510832, Sept. 15, 1995.

[9] J. Labat, O. Macchi, C. Laot, and N. Lesquin, "Is training of adaptive equalizers still useful?," in *Proc. Globecom'96*, vol. 2, London, U.K., Nov. 18–22, pp. 968–972.

[10] S. U. H. Qureshi, "Adaptive equalization," *Proc. IEEE*, vol. 73, pp. 1349–1387, Sept. 1985.

[11] O. Macchi "Adaptive processing," in *The LMS Approach with Applications in Transmission*. New York: Wiley, 1995.

[12] Y. Li and Z. Ding, "Convergence analysis of finite length blind adaptive equalizers," *IEEE Trans. Signal Processing*, vol. 43, pp. 2120–2129, Sept. 1995.

[13] M. Stojanovic, J. Catipovic, and J. G. Proakis, "Coherent communications over long range underwater acoustic telemetry channels," in *Acoustic Signal Processing for Ocean Exploration*, NATO ASI Series, J. M. F. Moura and I. M. G. Lourtie, Eds. Norwell, MA: Kluwer, 1993, pp. 607–612.

[14] D. D. Falconer, "Jointly adaptive equalization and carrier recovery in two-dimensional data communication systems," *Bell Syst. Tech. J.*, vol. 55, pp. 317–334, Mar. 1976.

[15] D. Mottier, M. Hélard, and J. Labat, "Self-adaptive decision feedback equalizer: Application to high-order QAM signals" in *Proc. ICC 97*, Montreal, P.Q., Canada, June 1997, pp. 1100–1104.

Joël Labat was born in Dakar, Senegal, in 1951. He received the Eng. degree from the Conservatoire National des Arts et Métiers (CNAM), Brest, France, in 1989, and the Ph.D. degree from the University of Bretagne Occidentale (UBO), Brest, France, in 1994.

He is presently with the Ecole Nationale Supérieure des Télécommunications de Bretagne (ENSTB), Brest, France, as a Maître de Conférences. His research interests include digital comunications over fading multipath channels and related problems such as adaptive equalization and synchronization.

Odile Macchi (M'75–SM'84–F'89) was born in Aurillac, France, in 1943. She graduated from the Ecole Normale Supérieure, Paris, France. She received the M.Sc. degree in both physics and mathematics in 1964 and 1965, respectively, and the teaching degree "Agrégation" from Paris University in 1966. She received the degree of Doctor of Science in 1972.

She is currently a Director of Research with the French National Center for Scientific Research, Laboratoire des Signaux et Systèmes, Ecole Supérieure d'Electricité, Gif Sur Yvette, France, in the field of adaptive communications and signal and image digital processing.

Dr. Macchi was awarded three French prizes in the field of electricity, and was nominated Distinguished Lecturer for 1994–1995 by the IEEE Signal Processing Society. In 1994 she was elected as a Member of the French "Académie des Sciences."

Christophe Laot was born in Brest, France, on March 12, 1967. He received the Eng. degree from the Ecole Francaise d'Electronique et d'Informatique (EFREI), Paris, France, in 1991, and the Ph.D. degree from the University of Rennes, Rennes, France, in 1997.

He is currently with the Digital Communications Team, Ecole Nationale Supérieure des Telcommunications de Bretagne (ENSTB), Brest, France. His professional interests include equalization, channel coding, and iterative processing.