A New Approach Based on *"Soft Statistics"* to the Nonlinear Blind-Deconvolution of Unknown Data Channels

Enzo Baccarelli and Stefano Galli, *Member, IEEE*

***Abstract*—In this paper, we present a new nonlinear receiver for the blind deconvolution of intersymbol interference (ISI) impaired data. The proposed receiver achieves fast identification of an unknown transmission channel using *only one* channel estimator and requiring the computation of *only the second-order conditional statistics* of the *baud-rate* sampled received signal and the knowledge of the *transmitted constellation*. The main novelty of the proposed approach is that the receiver accomplishes fast channel-identification by using *soft-statistics*. In particular, the receiver consists of a symbol-by-symbol *maximum a posteriori* (SbS-MAP) detector that feeds a *nonlinear Kalman-like channel estimator* with the soft statistics constituted by the *a posteriori probabilities (APPs)* of the state sequence of the ISI channel. Several numerical results confirm that the proposed blind detector achieves the identification of nonminimum phase channels with deep spectral notches within 300 symbols, even at low signal-to-noise ratios (SNRs). Furthermore, an attractive feature of the proposed blind channel estimator is that it directly estimates the discrete-time impulse response of the unknown channel so that, in principle, any equalization technique for known channels may be performed after channel identification has been achieved.**

***Index Terms*—Blind equalization, channel estimation, GSM test channels, MAP detectors, nonlinear MMSE estimation.**

I. Introduction

THE ultimate goal of blind-equalization algorithms is to detect ISI-impaired signals transmitted over unknown channels without resorting to known training sequences. For this purpose, a pioneering self-recovering adaptive equalizer was proposed by Sato [1] more than 20 years ago. As is well known, Sato's approach requires the utilization of nonconvex cost functions different from the traditional quadratic ones usually employed for trained equalizers. The work of Sato has been further developed by, among others, Godard [2], Benveniste and Goursat [3], Picchi and Prati [4], Shalvi and Weinstein [5], and, more recently, Tugnait and Gummadavelli [12]. The approaches followed by these authors are based on similar principles and require the computation of suitably defined *higher order* statistics of the received signal. The main drawback suffered by the resulting blind equalizers is typically a *slow* convergence to the unknown channel-impulse response. In fact, it has been experienced that several thousand observation samples are generally necessary to achieve channel identification [2], [6], [12]. Moreover, the utilization of nonconvex cost functions may lead to estimates of the channel affected by high residual mean square errors (MSEs) [6], [12].

To bypass these drawbacks, a blind channel equalizer based on second-order statistics has been proposed in [6]. This receiver directly exploits the cyclostationarity property of the observation process through a suitable *oversampling* of the received signal. The contribution in [6] has represented a novel approach to the problem of blind deconvolution and has stimulated further research (see, for example, [8] and references therein).

An alternative approach to blind identification of linear multisensor-based systems has been pursued in [26]. More precisely, an enhanced version of the expectation-maximization (EM) algorithm has been developed for blind estimation of a desired signal transmitted over an unknown ISI channel when both spectrally shaped cochannel random interference and observation noise are present. Interestingly, the version in [26] of the EM algorithm works also when the spectral parameters of the interfering signal and noise are *a priori* unknown to the multisensor receiver. In addition, it achieves blind signal estimation in a sequential fashion via an adaptive Kalman estimator.

In some recent contributions [7], [13]–[16], joint channel-estimation and data-detection techniques have been proposed for blind equalization of channels with severe ISI at low and moderate SNRs. The blind receivers presented in the cited contributions share a common structure and in essence consist of a data-detector aided by *a bank* of channel estimators. In particular, in [13] and [14] a maximum likelihood sequence estimator (MLSE) is employed for data-detection, whereas in [7], and [15], various forms of reduced complexity SbS-MAP detectors are used for the same purpose. Furthermore, in the above contributions, channel-estimation is accomplished via a *bank* of (at least) $S^{(L-1)}$ estimators[1] whose output channel-estimates are conditioned on the states of the trellis of the ISI channel [7, Sect. III], [13, Fig 1], [14, Sect. 4], [15, Fig. 2]. Although the resulting channel-estimates are, indeed, nearly optimal in a minimum mean square error (MMSE) sense, the computational load requested to implement the bank of channel-estimators *increases in an exponential fashion* with the (*a priori* expected) length L of the unknown channel impulse response. From this point of view, blind deconvolution techniques based on a *single* channel estimator would be more appealing in practical applications.

Manuscript received February 16, 2000; revised March 20, 2001. The associate editor coordinating the review of this paper and approving it for publication was Prof. Michail K Tsatsanis.

E. Baccarelli is with the INFO-COM Department, University of Rome "La Sapienza," Rome, Italy

S. Galli is with the Telcordia Technologies, Inc., Morristown, NJ 07960 USA.

Publisher Item Identifier S 1053-587X(01)05192-3.

[1] S is the size of the employed data constellation, and L is the *a priori* expected length of the unknown channel impulse-response in multiples of the signaling period.

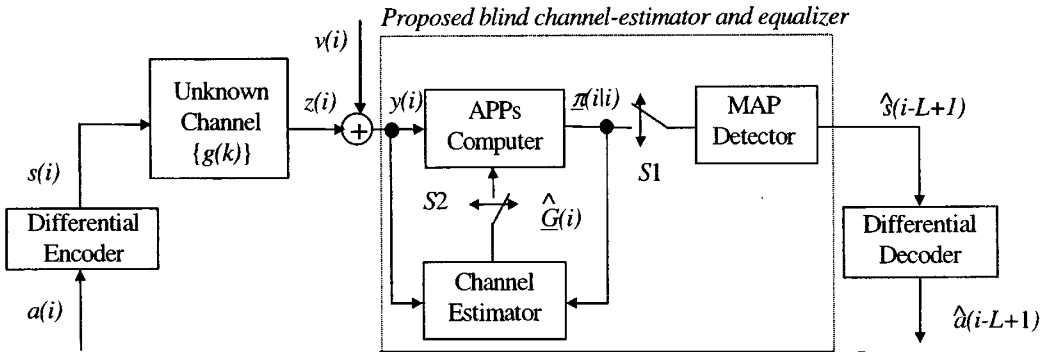

Fig. 1. Discrete-time lowpass version of the considered ISI-impaired communication system. Switch S1 is open, and switch S2 is shut during the SOM and the QHOM. Switch S1 is shut and switch S2 is open in the data detection operating mode (i.e., after having achieved channel identification at the end of the QHOM).

A somewhat different approach to the joint blind-deconvolution and data-detection has been recently proposed in [17], where the MLSE or the SbS-ML detectors are replaced by an improved blind version of the extended Kalman filter (EKF) equalizer formerly presented in [18]. In [17], an approximate MMSE-type joint estimate of the transmitted data and unknown channel impulse response is carried out via the so-called extended-Bayesian Filter (EBF) [17, Sect. II]. However, in the derivation of the EBF, a Gaussian approximation is introduced on some suitably defined conditional distributions of the detected data [17, Sect. II], and then, an *approximate* combined nonlinear MMSE estimate of the data and the channel is updated by performing (on a step-by-step basis) some sort of linearization on the observation equation [17, Sect. II]

In the present contribution, we present a novel blind equalizer constituted of an MMSE-type *single* channel estimator fed by the *"soft-statistics"* supplied by an SbS-MAP detector. The resulting blind detector is *nonlinear* and *only requires* the computation of the conditional second-order statistics of the observations and the knowledge of the *transmitted constellation*. Moreover, it requires neither oversampling of the received signal, nor the restriction of Gaussian approximation on the conditional distributions of the transmitted data, nor any form of linearization of the observation equation. More precisely, during the channel identification period, the SbS-ML detector embedded in the proposed blind-receiver feeds a *single nonlinear second-order Kalman-like channel estimator* with two kinds of *soft information*: at first, the *soft statistics* constituted by the APPs of the state of the Markov chain of the ISI channel and, afterwards, the nonlinear MMSE estimates of the transmitted symbols. For deriving the proposed blind-equalizer, we have resorted to the analytical tool constituted by the so called "Martingale difference (MD) processes" theory [10], which has been experienced to provide a theoretic framework suitable for developing optimal (and quasioptimal) nonlinear MMSE recursive finite-dimensional channel-estimators of practical interest.

Extensive simulations of the presented blind equalizer have shown the feasibility of identification of *both* channels with deep spectral notches and nonminimum phase channels *within 150–350 observation samples, even at low and moderate SNRs*. Moreover, the knowledge of the exact length of the impulse response is *not requested* by the proposed equalizer since it is sufficient that the impulse response sought by the channel-estimator is longer than the actual one.

The remainder of the paper is organized as follows. After modeling the considered communication system in Section II, we develop the relationships requested for implementing the proposed blind equalizer in Section III. Computational aspects and implementation consideration are covered in Section IV, whereas in Section V, we present several numerical results and comparisons to test the performance of the proposed blind equalizer in terms of convergence-rate, residual MSE in the channel-estimate and resulting bit error rate (BER). Some conclusive remarks are reported in the final Section VI.

II. Model of the Considered Transmission-System

The baud-rate sampled baseband equivalent data transmission system considered here is sketched in Fig. 1, where the discrete-time ISI-channel accounts for the combined effects of transmitting filter, noisy time-dispersive analog (unknown) waveform channel, front-end receiving filter, and symbol-rate sampler. After differential encoding[2] and quadrature amplitude modulation (QAM) of the i.i.d. zero-mean source stream $\{a(i)\}$, the resulting S-ary sequence $\{s(k) \in A \equiv \{1, \ldots, s_S\} \subset \boldsymbol{C}\}$ is transmitted over a linear channel with unknown time-invariant discrete-time impulse-response $\{g(k)\, 0 \le k, \le L-1\}$. Thus, the ISI corrupted noisy random sequence observed at the input of the blind detector is given by (see Fig. 1)[3]

$$y(i) = \sum_{k=0}^{L-1} g(k)s(i-k) + v(i) \equiv \underline{G}^{\mathrm{T}}\underline{x}(i) + v(i)$$
$$\equiv z(i) + v(i), \quad i \ge 1 \tag{1}$$

where $\underline{G} \equiv [g(0) \ldots g(L-1)]^{\mathrm{T}} \in \boldsymbol{C}^L$ is the (unknown) L-variate impulse-response vector, $\underline{x}(i) \equiv [s(i) \ldots s(i-L+1)]^{\mathrm{T}}$ is the channel-state vector, and $\{v(i)\}$ is a complex zero-mean Gaussian noise sequence whose uncorrelated components share

[2]As is well known, differential encoding of the source stream constitutes an effective means to combat ambiguity phenomena arising from the bilinear form of the observation model in (1) in the unknown channel $\{g(k)\}$ and unknown transmitted symbols $\{s(k)\}$

[3]From now on, vectors are denoted by underlined letters, whereas matrices are in bold characters. Furthermore, diag$\{\alpha_1, \ldots, \alpha_p\}$ denotes a $P \times P$ diagonal matrix with the elements $\{\alpha_1, \ldots \alpha_p\}$ disposed along the main diagonal, whereas $\underline{1}_m$ indicates the column vector with m unit elements.

a common variance equal to $(N_o/2)$. The L-variate random sequence $\{\underline{x}(i) \in A^L \equiv \{\underline{\xi}_1, \ldots, \underline{\xi}_N\} \subset \boldsymbol{C}^L\}$ is a first-order Markov chain generally referred to as the "state sequence" of the ISI channel [11]; it may assume $N \equiv S^L$ distinct values $\{\underline{\xi}_i, 1 \leq i \leq N\}$, which correspond to the L-long ordered subsequences of constellation symbols

$$\underline{\xi}_j \equiv \left[s_1^{(j)} \ldots s_{\mathrm{L}}^{(j)}\right]^{\mathrm{T}}, \quad 1 \leq j \leq N \tag{2}$$

where $s_i^{(j)} \in A$ indicates the ith component of the jth outcome $\underline{\xi}_j$ of the channel-state $x(i)$. From a statistical point of view, $\{\underline{x}(i)\}$ is an homogeneous Markov chain and can be described by the usual corresponding $N \times N$ state-transition probability matrix Φ (see, for example, [11]). Since this chain evolves as the state of a right-shift register of length L, the elements $\{\phi(l, m)\}$ of the state-transition probability matrix Φ are defined according to the (usual) relationship

$$\begin{aligned}\phi(l,m) \equiv& P\left(\underline{X}(n+1) = \underline{\xi}_1 | \underline{X}(n) = \underline{\xi}_m\right)\\ =& \begin{cases} \frac{1}{S}, & \text{if } \underline{\xi}_1 \text{ is an allowable continuation of } \underline{\xi}_m \\ 0, & \text{otherwise.}\end{cases}\end{aligned} \tag{3}$$

Now, having denoted by $y_1^i = \{y(1), \ldots y(i)\},\ i \geq 1$ the observed sequence in (1) available at the receiving side until step i, an SbS-MAP detector with decision-delay D equal to the memory length $(L-1)$ of the transmission channel[4] generates the hard-decided data sequence $\{\hat{s}_{\mathrm{MAP}}(i-L+1|i) \in A,\ i \geq 1\}$ according to the usual maximum *a posteriori* (MAP) decision rule

$$\begin{aligned}\hat{s}_{\mathrm{MAP}}(i-L+1|i) =& s_r \text{ if } P(s(i-L+1) = s_r | y_1^i)\\ &> P(s(i-L+1) = s_j | \underline{y}_1^i), \quad \forall j \neq r\end{aligned} \tag{4}$$

where $\{P(s(i-L+1) = s_r | \underline{y}_1^i), 1 \leq r \leq S\}$ is the set of the APPs of the random variable $s(i-L+1)$ conditioned on the received sequence $\underline{y}_1^i$. Now, a direct application of the "total probability theorem" allows us to relate the APPs of (4) to those pertaining to the Markov chain $\{x(i)\}$ through the expression [20]

$$\begin{aligned}P\left(s(i-L+1) = s_m | \underline{y}_1^i\right) =& \sum_{\underline{\xi}_j \in A(r,m)} P\left(\underline{X}(i) = \underline{\xi}_j | \underline{y}_1^i\right)\\ 1 \leq& m \leq S, \quad i \geq 1\end{aligned} \tag{5}$$

where $A(m) \equiv \{\underline{\xi}_j \in A^L : s_{\mathrm{L}-1}^{(j)} = s_m\}, 1 \leq m \leq S$ is the subset of A^L constituted by the outcomes $\{\underline{\xi}_j\}$ with the $(L-1)$th element equal to the constellation symbol s_m. Therefore, the computation of the APP set $\{P(\underline{X}(i) = \underline{\xi}_j | \underline{y}_1^i), 1 \leq j \leq N\}$ is sufficient to deliver the symbol decision in (4).

In the next section, we derive the nonlinear MMSE estimator that employs soft statistics for channel identification. For this purpose, it is useful to introduce the $L \times N$ mapping matrix $\mathbf{M} \equiv [\underline{\xi}_1, \ldots, \underline{\xi}_N]$, whose N columns are constituted by the previously defined outcomes $\{\underline{\xi}_i, 1 \leq i \leq N\}$ allowable to the channel state $\underline{X}(i)$.

[4]Several analytical and simulation results support the conclusion that a decision delay D that is equal to the memory length of the channel is, in general, adequate to obtain reliable SbS-MAP decisions [11], [19, Sect. V, VI], [20, eq.(6)] and following text]. Therefore, in the following, only the case $D = L - 1$ is explicitly developed.

III. Proposed Soft-Statistic Based Blind Channel Estimator

During the channel identification period, the task of a blind equalizer is to achieve a reliable estimate of the unknown impulse response vector $\underline{G}$ of the transmission channel in (1). In the proposed detector of Fig. 1, this task is pursued by cascading a preliminary *soft operating mode* (SOM) with a final *quasi-hard operating mode* (QHOM). During the SOM, a coarse estimate of the unknown impulse response is generated on the basis of the APPs of the state $\{\underline{x}(i)\}$ of the ISI channel, and afterwards, this channel estimate is refined during the QHOM by using a nonlinear MMSE estimate of the transmitted symbols $\{s(i)\}$. These operative modes are detailed in the next two subsections.

A. Soft Operating Mode (SOM)

In this operating mode, the channel identification is pursued via a nonlinear MMSE estimation of the unknown impulse-response vector $\underline{G}$ that, when *a priori* information on the channel is not available, can be effectively modeled as a zero-mean Gaussian random vector. Thus, the nonlinear MMSE estimation of $\underline{G}$ can be obtained by exploiting the properties of the so-called "Martingale difference" (MD) sequences (see [10] for a good introduction to the basic concepts of the MD theory). In fact, after being indicated as $\underline{\hat{G}}(i) \equiv E\{\underline{G}|\underline{y}_1^{i-1}\}$, $\underline{\hat{G}}(i-1) \equiv E\{\underline{G}|\underline{y}_1^{i-1}\}$, and $\hat{y}(i|i-1) \equiv E\{y(i)|\underline{y}_1^{i-1}\}$, the MMSE nonlinear estimates of $\underline{G}$ and $y(i)$ on the basis of the observed sequences $\underline{y}_1^i$ and $\underline{y}_1^{i-1}$, respectively, we note that the two related series $\{\underline{\mu}(i) \equiv \underline{\hat{G}}(i) - \underline{\hat{G}}(i-1)\}$ and $\{\Theta(i) \equiv y(i) - \hat{y}(i|i-1)\}$ are MD sequences (in the sense of [10, Sect.VII]) with respect to the observations $\{\underline{y}_1^i, i \geq 1\}$. Therefore, an application of the so-called *MD representation theorem* of [10, Sect.VII] allows us to write

$$\underline{\mu}(i) = \underline{K}_G(i)\Theta(i) \tag{6}$$

where the observation-dependent time-varying filter gain $\underline{K}_G(i)$ is defined by the expression [10, eq.(107)]

$$\begin{aligned}\underline{K}_G(i) \equiv& \frac{E\left\{\Theta^*(i)\underline{\mu}(i)|\underline{y}_1^i\right\}}{E\left\{\|\Theta(i)\|^2 | \underline{y}_1^i\right\}}\\ \equiv& \frac{E\left\{[y(i) - \hat{y}(i|i-1)]^* \left[\underline{\hat{G}}(i) - \underline{\hat{G}}(i-1)\right] | \underline{y}_1^i\right\}}{E\left\{\|y(i) - \hat{y}(i|i-1)\|^2 | \underline{y}_1^i\right\}}.\end{aligned} \tag{7}$$

Therefore, directly from the above definitions of $\{\mu(i)\}$ and $\{\Theta(i)\}$, (6) can be rewritten in the form of a Kalman-like filter as

$$\underline{\hat{G}}(i) = \underline{\hat{G}}(i-1) + \underline{K}_G(i)\left[y(i) - \hat{y}(i|i-1)\right]. \tag{8}$$

Now, from the whiteness property of the noise process $\{v(i)\}$ of (1) and the independence between $\underline{G}$ and $X(i)$, the following relationship arises for the updating of the one-step MMSE predictor present in (8):

$$\hat{y}(i|i-1) = \underline{\hat{G}}^{\mathrm{T}}(i-1)\underline{\hat{x}}(i|i-1) \tag{9}$$

where, in turn, the MMSE prediction of the channel-state $\underline{\hat{x}}(i|i-1) \equiv E\{\underline{x}(i)|\underline{y}_1^{i-1}\}$ can be computed as

$$\begin{aligned}\underline{\hat{x}}(i|i-1) \equiv& E\left\{\underline{x}(i)|\underline{y}_1^{i-1}\right\}\\ \equiv& \sum_{j=1}^{N}\underline{\xi}_j P\left(\underline{X}(i)=\underline{\xi}_j|\underline{y}_1^{i-1}\right) = \mathbf{M}\underline{\pi}(i|i-1)\end{aligned} \tag{10}$$

where

$$\underline{\pi}(i|i-1) \equiv [P(\underline{X}(i)=\underline{\xi}_1|\underline{y}_1^{i-1})\dots P(\underline{X}(i)=\underline{\xi}_N|\underline{y}_1^{i-1})]^{\mathrm{T}}$$

is the N-variate APP vector of the channel state $\underline{X}(i)$ computed on the basis of the observed sequence $\underline{y}_1^{i-1}$ available until step $(i-1)$.

As far as the evaluation of the filtering gain sequence $\{\underline{K}_G(i), i \ge 1\}$ in (7) is concerned, the analysis in [10] leads to the general conclusion that, in principle, this gain sequence is not predictable (in the sense of [10]) with respect to the observation stream $\{\underline{y}_1^i, i \ge 1\}$, and therefore, the resulting channel estimator (8) is *not recursive*; therefore, its computational complexity is no longer linear in the length of the observed record. In order to retain the recursive structure of the estimator (8), we must assume that $\{\underline{K}_G(i), i \ge 1\}$ is *predictable* with respect to $\{\underline{y}_1^i, i \ge 1\}$, and this requires replacing $\underline{y}_1^i$ with $\underline{y}_1^{i-1}$ in the conditional expectation of (7). As a consequence, the developed channel estimator may be considered to be the best one within the class of the *second-order MMSE recursive* filters. In practice, the simulation results of Section IV confirm the actual effectiveness of the introduced assumption. Therefore, after replacing $\underline{y}_1^i$ with $\underline{y}_1^{i-1}$ in (7), the numerator can be evaluated as

$$\begin{aligned}E\left\{\Theta^*(i)\,\underline{\mu}(i)|\underline{y}_1^{i-1}\right\} \equiv& E\left\{\underline{\mu}(i)y^*(i)|\underline{y}_1^{i-1}\right\}\\ &- E\left\{\underline{\mu}(i)\hat{y}^*(i|i-1)|\underline{y}_1^{i-1}\right\}\ {}^{(a)}\\ =& E\left\{\underline{\mu}(i)y^*(i)|\underline{y}_1^{i-1}\right\}\end{aligned} \tag{11}$$

where, in (a), we have exploited the property that the random variable $\underline{\mu}(i)$ is not predictable (in the sense of [10, Sect.VI]) with respect to $\underline{y}_1^{i-1}$. Now, by introducing the observation model in (1) and then exploiting the definition of $\underline{\mu}(i)$, the conditional expectation in (11) can be further developed as

$$\begin{aligned}&E\left\{\underline{\mu}(i)y^*(i)|\underline{y}_1^{i-1}\right\}\\ &\equiv E\left\{\underline{\hat{G}}(i)y^*(i)|\underline{y}_1^{i-1}\right\} - E\left\{\underline{\hat{G}}(i-1)y^*(i)|\underline{y}_1^{i-1}\right\}\ {}^{(b)}\\ &= E\left\{\underline{\hat{G}}(i)y^*(i)|\underline{y}_1^{i-1}\right\} - \underline{\hat{G}}(i-1)\underline{\hat{G}}^{\mathrm{H}}(i-1)\underline{\hat{x}}^*(i|i-1)^{(c)}\\ &= E\left\{\underline{GG}^{\mathrm{H}}|\underline{y}_1^{i-1}\right\}\underline{\hat{x}}^*(i|i-1)\\ &\quad - \underline{\hat{G}}(i-1)\underline{\hat{G}}^{\mathrm{H}}(i-1)\underline{\hat{x}}^*(i|i-1)\\ &= \mathbf{Cov_G}(i-1)\underline{\hat{x}}^*(i|i-1)\ {}^{(d)}\\ &= \mathbf{Cov_G}(i-1)\mathbf{M}^*\underline{\pi}(i|i-1)\end{aligned} \tag{12}$$

where

$$\mathbf{Cov_G}(i-1) \equiv E\{[\underline{G}-\underline{\hat{G}}(i-1)][\underline{G}-\underline{\hat{G}}(i-1)]^{\mathrm{H}}|\underline{y}_1^{i-1}\}$$

denotes the error covariance matrix in the estimation of the channel impulse-vector $\underline{G}$ on the basis of observations available until step $(i-1)$. The identities (b) and (c) in (12) follow from an exploitation of the observation model (1), whereas (d) directly arises from (10). Furthermore, in the derivation of (c), we have also exploited the so-called "*smoothing properties*" of nested conditional expectations of random variables (see [27, eqs.(7.23) and (7.24)]), which allows us to write

$$\begin{aligned}E\left\{\underline{\hat{G}}(i)y^*(i)|\underline{y}_1^{i-1}\right\} \equiv& E\left\{E\left\{\underline{G}\,y^*(i)|\underline{y}_1^{i}\right\}|\underline{y}_1^{i-1}\right\}\\ \equiv& E\left\{\underline{G}\,y^*(i)|\underline{y}_1^{i-1}\right\}.\end{aligned}$$

As far as the denominator of (7) is concerned, from the observation model of (1) and the independence of $\underline{G}$ on $\{\underline{x}(i)\}$, we obtain the following chain of relationships for the conditional error covariance of the MMSE estimate $\hat{y}(i|i-1)$ of the ith observation $y(i)$:

$$\begin{aligned}&E\left\{\|\Theta(i)\|^2|\underline{y}_1^{i-1}\right\}\\ &= N_o + E\left\{\left\|\underline{G}^{\mathrm{T}}\underline{x}(i)\right\|^2|\underline{y}_1^{i-1}\right\}\\ &\quad - \|\hat{y}(i|i-1)\|^2\\ &= N_o + E\left\{\underline{G}^{\mathrm{T}}E\left\{\underline{x}(i)\underline{x}^{\mathrm{H}}(i)|\underline{y}_1^{i-1}\right\}\underline{G}^*|\underline{y}_1^{i-1}\right\}\\ &\quad - \|\hat{y}(i|i-1)\|^2\ {}^{(e)}\\ &= N_o + E\Bigg\{\text{Trace}\left\{\left(E\left\{\underline{x}(i)\underline{x}^{\mathrm{H}}(i)|\underline{y}_1^{i-1}\right\}\right)^{\mathrm{T}}\underline{GG}^{\mathrm{H}}\right\}\\ &\qquad\times|\underline{y}_1^{i-1}\Bigg\} - \|\hat{y}(i|-1)\|^2\ {}^{(f)}\\ &= N_o + \text{Trace}\Bigg\{\left(E\left\{\underline{x}(i)\underline{x}^{\mathrm{H}}(i)|\underline{y}_1^{i-1}\right\}\right)^{\mathrm{T}}\\ &\qquad\times E\left\{\underline{GG}^{\mathrm{H}}|\underline{y}_1^{i-1}\right\}\Bigg\} - \|\hat{y}(i|i-1)\|^2\end{aligned} \tag{13}$$

where (e) and (f) in (13) arise from the basic properties of the Trace operator. Therefore, after introducing the channel-state conditional error covariance matrix $\mathbf{Cov_X}(i|i-1) \equiv E\{[\underline{x}(i)-\underline{\hat{x}}(i|i-1)][\underline{x}(i)-\underline{\hat{x}}(i|i-1)]^{\mathrm{H}}|\underline{y}_1^{i-1}\}$, a direct exploitation of the two following relationships:

$$\begin{aligned}\mathbf{Cov_X}(i|i-1) =& E\left\{\underline{x}(i)\underline{x}^{\mathrm{H}}(i)|\underline{y}_1^{i-1}\right\}\\ &- \underline{\hat{x}}(i|i-1)\underline{\hat{x}}^{\mathrm{H}}(i|i-1)\end{aligned} \tag{14}$$

$$\mathbf{Cov_G}(i-1) = E\left\{\underline{GG}^{\mathrm{H}}|\underline{y}_1^{i-1}\right\} - \underline{\hat{G}}(i-1)\underline{\hat{G}}^{\mathrm{H}}(i-1) \tag{15}$$

allows us to rewrite (13) as

$$\begin{aligned}&E\left\{\|\Theta(i)\|^2|\underline{y}_1^{i-1}\right\}\\ &\equiv N_o + \text{Trace}\left\{\mathbf{Cov_X^*}((i|i-1)\mathbf{Cov_G}(i-1)\right\}\\ &\quad + \underline{\hat{G}}^{\mathrm{T}}(i-1)\mathbf{Cov_X}(i|i-1)\underline{\hat{G}}^*((i-1)\\ &\quad + \underline{\hat{x}}^{\mathrm{H}}(i|i-1)\mathbf{Cov_G^*}(i-1)\underline{\hat{x}}(i|i-1).\end{aligned} \tag{16}$$

TABLE I
ORDERED LIST OF THE RECURSIONS REQUESTED BY THE PROPOSED BLIND EQUALIZER

	Ordered list of recursions
SOM stage	Eqs. (20), (10), (9), (17), (16), (12), (7), (8), (22), (21), (19), (18).
QHOM stage	Eqs. (20), (9), (10), (23), (12), (7), (8), (22), (21), (19), (18).
Data Detection stage	Eqs. (20), (22), (21), (5), (4).

Although it is easy to show that the evaluation of the error covariance matrix in (14) can be performed on the basis of the above-defined APP vector $\pi(i|i-1)$ as

$$\mathbf{Cov_X}(i|i-1) = \mathbf{M}\left[\text{diag}\left\{\underline{\pi}(i|i-1)\right\} - \underline{\pi}(i|i-1)\underline{\pi}^{\mathrm{T}}(i|i-1)\right]\mathbf{M}^{\mathrm{H}} \quad (17)$$

the updating of the error covariance matrix of the channel-estimate $\mathbf{Cov_G}(i)$ is not as straightforward. However, in the Appendix, it is pointed out that the recursive updating of $\mathbf{Cov_G}(i)$ can be effectively carried out via the following approximate relationship:

$$\mathbf{Cov_G}(i) \cong \left[\mathbf{I}_{\mathrm{L}\times\mathrm{L}} - \underline{K}_G(i)\hat{\underline{x}}^{\mathrm{T}}(i|i)\right]\mathbf{Cov_G}(i-1) \quad (18)$$

where, similar to (10), the nonlinear MMSE filtered estimate $\hat{\underline{x}}(i|i)$ of the channel-state $x(i)$ on the basis of $\underline{y}_1^i$ can be computed from the corresponding APP vector $\underline{\pi}(i|i) \equiv [P(\underline{X}(i)=\underline{\xi}_1|\underline{y}_1^i)\dots P(\underline{X}(i)=\underline{\xi}_N|\underline{y}_1^i)]^{\mathrm{T}}$ as

$$\hat{\underline{x}}(i|i) = \mathbf{M}\underline{\pi}(i|i). \quad (19)$$

The last step concerns the recursive updating of the two APP vectors $\underline{\pi}(i|i-1)$ and $\underline{\pi}(i|i)$ requested for the evaluation of (10) and (19). As far as the updating of $\underline{\pi}(i|i)$ from $\underline{\pi}(i|i-1)$ is concerned, a straightforward application of the Bayes' rule and the total probability theorem followed by the exploitation of the Markovian property of the channel-state sequence $\{\underline{x}(i)\}$ leads to the usual relationship [10], [11], [19]

$$\underline{\pi}(i|i-1) = \mathbf{\Phi}\underline{\pi}(i-1|i-1). \quad (20)$$

It is also easy to prove that a reiterated application of the Bayes' rule gives rise to the following expression for the computation of $\underline{\pi}(i|i)$ from $\underline{\pi}(i|i-1)$ [11], [19]:

$$\underline{\pi}(i|i) = \frac{[\mathbf{D}(i)\underline{\pi}(i|i-1)]}{\lfloor(\underline{1}_N)^{\mathrm{T}}\mathbf{D}(i)\underline{\pi}(i|i-1)\rfloor} \quad (21)$$

where, as a consequence of the Gaussianity property of the channel-noise sequence $\{v(i)\}$ in (1), the $N\times N$ observation-dependent diagonal matrix $\mathbf{D}(i)$ is composed by N exponential terms as follows:

$$\mathbf{D}(i) \equiv \text{diag}\left\{\exp\left\{-\frac{\left\|y(i)-\hat{\underline{G}}^{\mathrm{T}}(i)\underline{\xi}_1\right\|^2}{N_o}\right\} \dots \exp\left\{-\frac{\left\|y(i)-\hat{\underline{G}}^{\mathrm{T}}(i)\underline{\xi}_N\right\|^2}{N_o}\right\}\right\}. \quad (22)$$

In summary, from (5)–(12), it can be seen that during the SOM the channel estimator utilizes the soft-statistics constituted by the APP sequence $\{\pi(i|i)\}$ of the states $\{\underline{x}(i)\}$ of the ISI channel for achieving a coarse channel identification. Furthermore, (7) points out that during the SOM the gain vector $\underline{K}_G(i)$ of the channel estimator is *time varying and observation dependent* so that the resulting channel estimator is *time-variant and nonlinear*. For the sake of clarity, the ordered set of relationships to be updated by the proposed blind equalizer during the SOM stage is listed in Table I.

It is worth pointing out that during the SOM stage, the updating of the APP vector $\underline{\pi}(i|i)$ in (21) is, by fact, approximate since it is based on the available estimate $\hat{\underline{G}}(i)$ of the true impulse response $\underline{G}$. However, when $\hat{\underline{G}}(i)$ approaches $\underline{G}$, the computation of $\underline{\pi}(i|i)$ in (21) becomes exact.

As far as the initialization of the presented blind equalizer is concerned, an effective means to reflect the lack of any *a priori* knowledge on the channel impulse response $\underline{G}$ is to pose $\underline{\pi}(0|0) \equiv 1/N[1\dots 1]^{\mathrm{T}}$ and $\mathbf{Cov_G}(0) \equiv (1/L)\mathbf{I}_{\mathrm{L}\times\mathrm{L}}$. Moreover, the simulations have shown that

$$\hat{\underline{G}}(0) \equiv [1\ 0\ \dots\ 0]^{\mathrm{T}}$$

is a good choice as an initial estimate of the unknown vector $\underline{G}$.

Finally, according to a usual taxonomy [9], [10], we remark that the channel estimator in (8) constitutes *an MMSE-type nonlinear second-order recursive channel estimator*. From the outset, we can also conclude that the channel estimates $\{\hat{\underline{G}}(i)\}$ in (8) represent *the best* (in an MMSE sense) available estimates for the considered transmission system in (1) when *only one recursive second-order channel estimator* is employed at the receiving side for achieving blind channel equalization.

B. *Quasi-Hard Operating Mode (QHOM)*

After achieving a coarse estimate of the channel impulse response $\underline{G}$, a conventional blind equalizer would switch to a decision-driven operating mode to refine channel identification and then speed up convergence on the basis of hard-detected data [hard operating mode (HOM)]. Therefore, in the HOM, the data output by the detector are assumed *error-free* and are then fed to the channel estimator as in the conventional trained (i.e., non-blind) adaptive equalizers [2]–[5]. However, at low and moderate SNRs, the decisions supplied by the detector at the end of the SOM are, in general, not reliable, and this may cause a *degradation* of the channel estimate $\hat{\underline{G}}$ achieved at the end of the SOM [14].

We have ascertained, via extensive computer simulations, that an effective means to *soften* these effects and, at the same time, *speed up* the convergence rate of the proposed blind receiver simply consists of *forcing to zero* the conditional covariance error matrix $\mathbf{Cov_x}(i|i-1)$ in (17) of the channel state estimates. Therefore, in the QHOM, the updating of the conditional error variance $E\{\|\Theta(i)\|^2|\underline{y}_1^{i-1}\}$ of the observation estimates in (16) is carried out by using the following simplified relationship:

$$E\left\{\|\Theta(i)\|^2\,|\underline{y}_1^{i-1}\right\} = N_o + \hat{\underline{x}}^{\mathrm{H}}(i|i-1)\mathbf{Cov}^*_{\mathbf{G}}(i-1)\hat{\underline{x}}(i|i-1). \quad (23)$$

It is worth pointing out that forcing to zero $\mathbf{Cov_x}(i|i-1)$ is equivalent to considering the MMSE estimates of the channel state error-free during the QHOM. Although it is not exact, this assumption is *less stringent* than assuming that the hard decisions in (4) are themselves error free. By doing so, we have experienced that the proposed blind equalizer is able to *soften* the effects of wrong hard decisions, especially at low and moderate SNRs, and the simulation results of Section V directly support this conclusion. In Table I, the ordered set of equations to be updated by the presented blind equalizer during the QHOM stage is reported.

IV. Computational Aspects and Implementation Considerations

As pointed out in Section I, the proposed blind equalizer of Fig. 1 achieves channel identification by using *only one* channel estimator that exploits *the second order* soft statistics of the *baud-rate sampled* received signal.

The proposed blind equalizer is essentially composed of two processing units that run in parallel during the SOM and the QHOM stages and are fed by the baud-rate sampled received signal (see Fig. 1 with the switch S2 shut and S1 open). More precisely, the task of the processing unit labeled as "APPs computer" in Fig. 1 is to calculate the APPs $\{\underline{\pi}(i|i)\}$ of the state of the ISI channel, and this task is accomplished via the recursive updating of (20)–(22). Therefore, since the size N of the APP vector $\underline{\pi}(i|i)$ to be updated is $N = S^L$, according to [19, Tab. I], we conclude that, in principle, the computational load (per S-ary transmitted symbol) sustained by the APPs computing unit of Fig. 1 is of the order of $O(S^L)$, i.e., it increases in an exponential way with the (expected) channel length L. However, for large L, the approach recently presented in [21] can be effectively pursued to reduce the complexity requested by the updating of and (21), whereas the quasioptimal versions of the SbS-MAP equalizer proposed in [19] can be also implemented to avoid a direct computation of the N exponential terms in (22).

As far as the processing unit labeled as "channel estimator" in Fig. 1 is concerned, its task is to compute the channel estimate $\underline{G}(i)$ in (8) via the recursive updating of (7), (9), and (18). Now, an examination of (8) shows that this channel estimator constitutes an enhanced version of a conventional Kalman-type tracker. In fact, the only difference between the channel estimator in (8) and the conventional ones present in the usual adaptive equalizers [18], [22], [23] is constituted by the type of statistics (*soft* for the proposed estimator and *hard* for the conventional ones) supplied to the channel-estimator by the corresponding detectors present in the receivers. Therefore, the computational complexity really requested for the updating of (8) deeply depends on the specific algorithm actually used to implement (8). For example, it is known that the computational complexity requested by the usual "square root Kalman filter" is $O(L^2)$ [22], which can be lowered to $O(L)$ by resorting to "fast Kalman" algorithms (see [23] and references therein).

On the basis of the above considerations, we can conclude that the implementation of the proposed blind equalizer of Fig. 1 requires a computational effort that is of the same order of the conventional adaptive (trained) equalizers described, for example, in [18], [22], and [23], and based on Viterbi-like detectors supported by recursive-least-squares (RLS) or Kalman-type channel trackers.

V. Simulation Results and Performance Comparisons

In Section V-A, we present simulation results that confirm the fast identification capability and good steady-state BERs of the proposed receiver in Fig. 1 over time-invariant channels. In Section V-B, we compare the performance achieved by the proposed receiver with other known blind receivers. Finally, in Section V-C, we test the robustness of the proposed receiver over time-variant Rayleigh fading channels.

In the simulations presented in Section V-A and Section V-B, we have considered three time-invariant channels with the following impulse responses: $\underline{G}_A \equiv [0.707\ 0\ -0.707]^{\mathrm{T}}$, $\underline{G}_B \equiv [0.55\ 0.33\ 0.77]^{\mathrm{T}}$, and $\underline{G}_C \equiv [0.408\ 0.816\ 0.408]$. Channel $\underline{G}_A$ exhibits *two spectral nulls* at low and high frequencies, and channel $\underline{G}_B$ is a *nonminimum phase* channel with a deep notch in the middle of the bandwidth, whereas channel $\underline{G}_C$ has *spectral nulls* at high frequencies. Moreover, the randomly time-varying radio channel composed by three equal-power uncorrelated scattering (US) Rayleigh-faded taps has been considered in the simulation reported in Section V-C.

A. Performance of the Proposed Blind Equalizer Over Static Channels

The obtained behavior of the MSEs of the channel estimates for the first two channels $\underline{G}_A$ and $\underline{G}_B$ are shown in Figs. 2, 3. All simulated trials have been carried out using 100 observation samples during the SOM and 400 observation samples during the QHOM. The results were also averaged over 5000 independent trials. As Figs. 2 and 3 point out, the SOM appears mandatory to achieve a preliminary rough estimation of the unknown channel impulse responses. Moreover, during the QHOM, the channel estimation improves rapidly in terms of residual MSE. The improvement arising at low and moderate SNRs from the utilization in the QHOM of the MMSE estimates of the channel states in place of their corresponding hard decisions can be also appreciated from Figs. 2 and 3. In fact, for SNRs up to 13 dBs, the residual MSE of the channel estimates is noticeably lower in the QHOM than in the HOM. However, at higher SNRs, the HOM obviously approaches the same performances of the QHOM.

The trajectory of the estimated channel taps for channel G_C during the SOM is shown in Fig. 4 versus the number of observed samples at SNRs of 10 and 20 dB. It can be seen that only 100–150 samples are generally requested to achieve a preliminary good identification of the channel.

The performance of the proposed blind receiver in terms of BER versus SNR is shown in Figs. 5 and 6 for a binary shift keying (BPSK) modulation format. As a benchmark, the performance of the theoretically optimum minimum error probability SbS-MAP equalizer [11] for *the ideal case of known channel impulse response* is also reported for the same figures. For the simulations in Figs. 5 and 6, 100 observation samples in the SOM and 300 in the QHOM/HOM have been employed. After the QHOM/HOM, the channel estimation procedure is stopped,

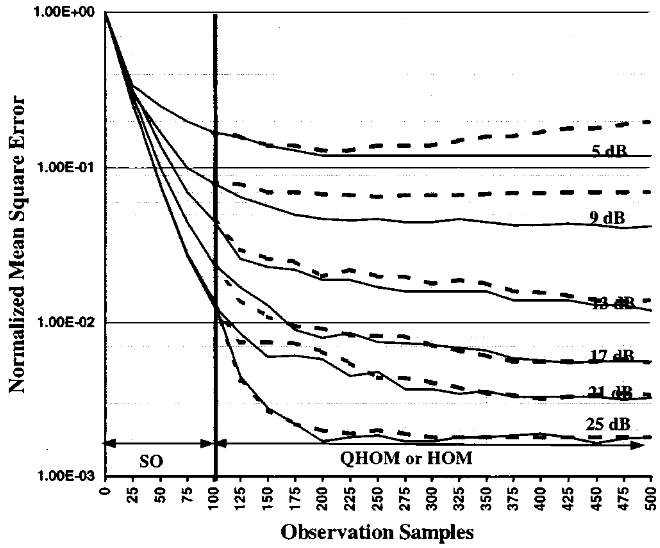

Fig. 2. Behavior of the average MSE of the estimate of $\underline{G}_A$ versus the number of observation samples for different values of the SNR E_b/N_o. The reported curves refer to a differentially encoded (DE) and BPSK modulated (DE-BPSK) data stream and are averaged over 5000 independent trials.

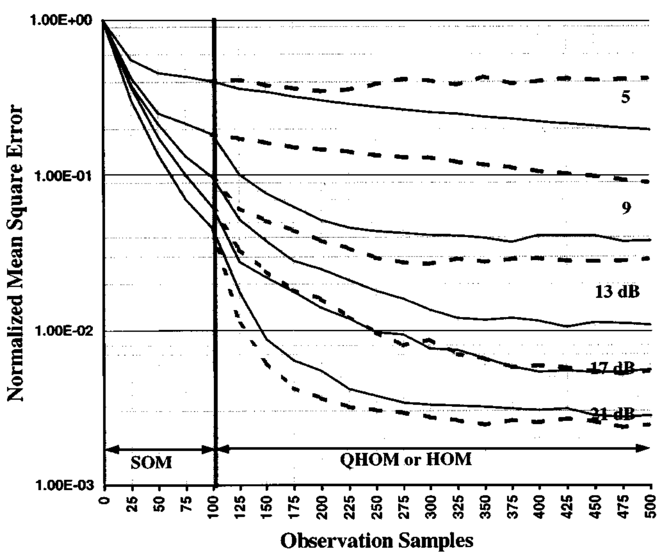

Fig. 3. Same as in Fig. 2 for channel $\underline{G}_B$.

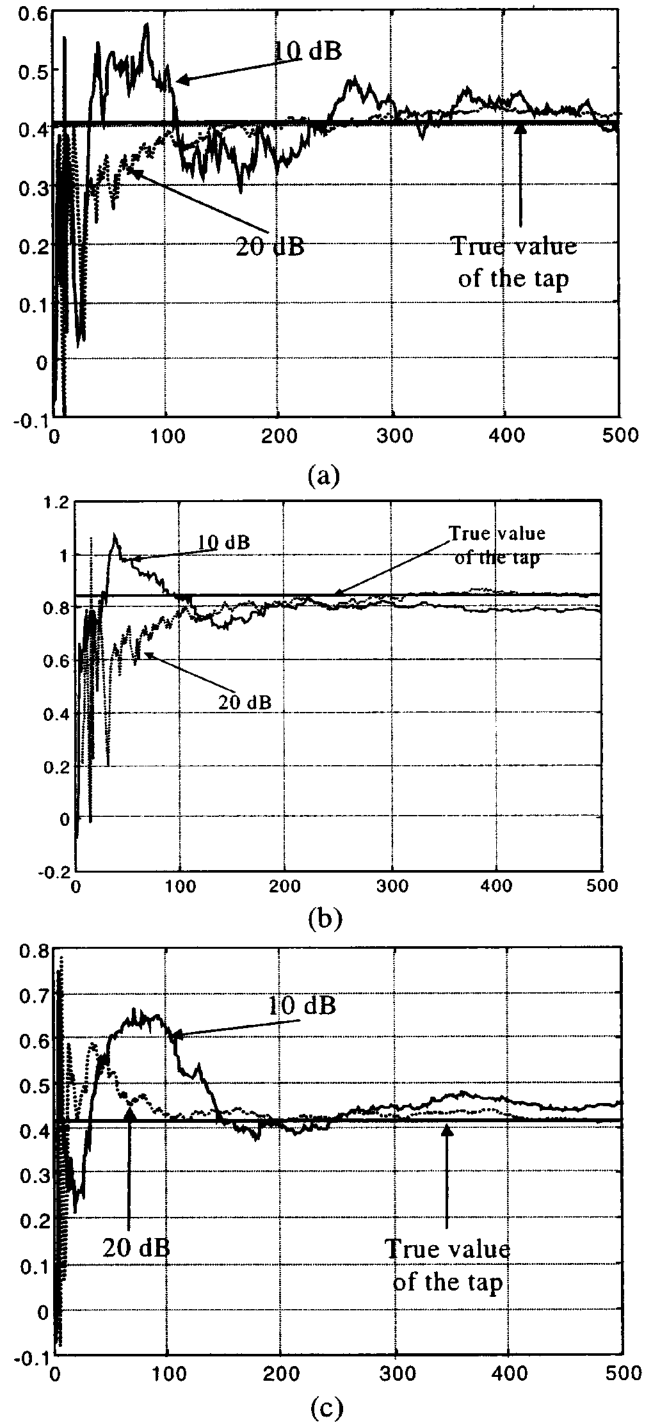

Fig. 4. Trajectory of the estimates of the taps of the channel G_C during the SOM as a function of the number of the observed samples for different values of the SNR (in decibels). DE-BPSK modulation has been adopted in the simulations. Plots (a), (b), and (c) refer to the channel taps $G_C(0)$, $G_C(1)$, and $G_C(2)$, respectively.

and data recovery is performed by the conventional SbS-MAP equalizer of [11] with a decision delay D set to $L-1$. By referring to the block diagram of Fig. 1, after channel-identification is achieved, the switch S2 is shut so that the receiver accomplishes data detection via the recursive updating of (20)–(22) with $\hat{\underline{G}}(i) \equiv \hat{\underline{G}}$ fixed (i.e., *i-invariant*) and set to the value assumed at the end of the QHOM stage. Therefore, at each step i, the $(L-1)$-delayed MAP hard decision $\hat{s}(i-L+1)$ of the transmitted symbol $s(i-L+1)$ is computed from the corresponding APP vector $\pi(i|i)$ in (21) via an application of the relationships in (4) and (5). This last task is carried out in the data-detection operating mode by the block labeled "MAP detector" (see Table I for a list of the recursions performed during the data-detection operating mode).

As Figs. 5 and 6 confirm, the blind detector of Fig. 1 is able to obtain BER performances that are *very close* to those of the *ideal case* of known channels, even at low and moderate SNRs. Furthermore, as it may be expected, Figs. 5 and 6 point out that the BER degradation induced by the residual MSE present in the final channel estimates generated at the end of the QHOM stage is more evident at higher SNRs, and this is due to the enhanced sensitivity of the SbS-MAP detector to small errors that are possibly present in the final channel estimates at high SNRs.

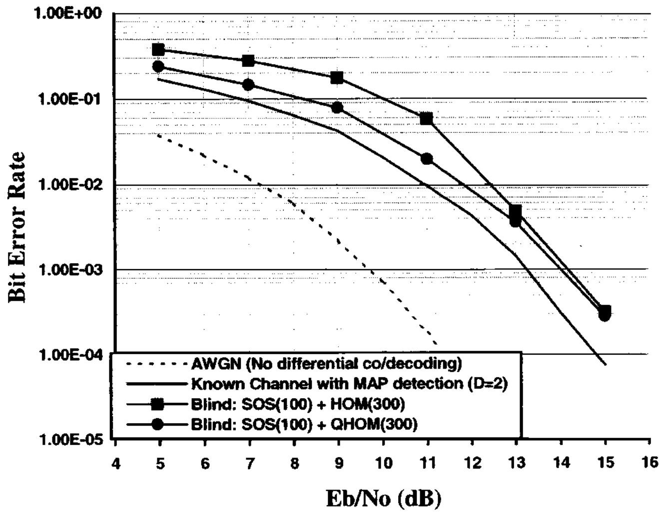

Fig. 5. Steady-state BER for channel $\underline{G}_A$ versus E_b/N_o. The simulations refer to a DE-BPSK-modulated data stream and are averaged over 5000 independent trials.

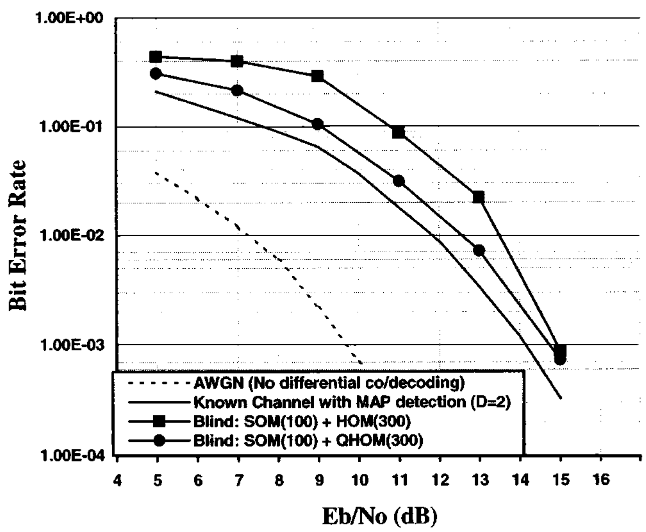

Fig. 6. Same as in Fig. 5 for channel $\underline{G}_B$.

B. Performance Comparison with Other Blind Equalizers

The steady-state BER performance of the proposed blind detector of Fig. 1 on the above cited $\underline{G}_C$ channel is shown in Fig. 7 together with the corresponding performances of Sato's equalizer (SATE) [1, eq.(2)], the constant modulus equalizer (CME) of [25 ,eqs.(11) and (12)], the Bayesian MAP equalizer (BAMAPE) of [15], and the extended Kalman filter equalizer (EKFE) of [18]. In particular, the simulated SATE and CME are constituted by FIR filters with 17 equally spaced taps, whereas, according to [15, Sect.V], the implemented BMAPE presents a bank of eight least mean square (LMS) channel estimators. All the simulation results reported in Fig. 7 refer to DE-BPSK modulated data, and they are averaged over 5000 independent trials. The first 300 symbols of each transmitted record have been used for blind channel estimation. As a benchmark, the performance of the SbS-MAP equalizer [11] for the ideal case of known channel impulse response has been also plotted for $D = 2$.

An examination of Fig. 7 shows that the performance plots of the simulated SATE, CME, and EKFE are nearly flat, and therefore, their actual utilization over channels with deep spectral nulls does not appear attractive. On the contrary, both the proposed detector of Fig. 1 and the BMAPE of [15] give rise to BER curves that are (quasi) exponential in the SNR and approach that of the ideal case of known channel-impulse response. Although the BMAPE appears to perform slightly better than the proposed one, it is worth pointing out that the blind equalizer of Fig. 1 utilizes *only one* channel estimator, whereas the simulated BMAPE utilizes a *bank of eight LMS* channel estimators.

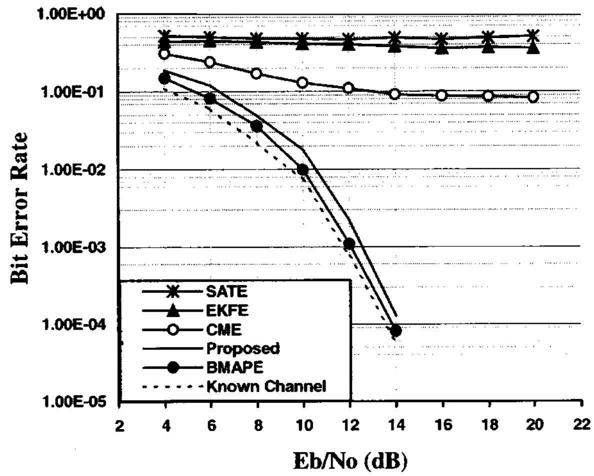

Fig. 7. Steady-state BERs for channel G_C versus E_b/N_o for the proposed blind detector of Fig. 1 and the SATE, EKFE, BMAPE, and CME ones. The performance of the MAP equalizer [11] with $D = 2$ for the ideal case of known channel impulse is also shown.

C. Performance Over Randomly Time-Variant GSM-Like Channels

The robustness of the proposed blind detector has been tested on some time-varying multipath channels affected by Rayleigh fading. In particular, the radio channel composed by three [i.e., $L = 3$ in (1)] equal-power, US, symbol-spaced, Rayleigh-faded taps has been considered. The simulated links can be considered to be representative of micro and macro cellular land-mobile radio links in urban areas [28]. As it is known, in these environments, the Doppler power density spectrum $S(f)$ describing the randomly time-variant channel fluctuations is well modeled by Clarke's formula [28]

$$S(f) = \begin{cases} \frac{1}{\sqrt{1-\left(\frac{f}{B_D}\right)^2}}, & |f| \le B_D \\ 0, & |f| \ge B_D \end{cases} \qquad (24)$$

where the Doppler-spread $B_D(Hz)$ of the considered channel is related to the mobile speed v and carrier frequency f_c as $B_D = (vf_c)/c$ (c is the speed of light). In the carried out simulations of Fig. 9, values of the product Doppler spread $\times$ signaling period $B_D T_S$ of $0, 10^{-4}$, and $5 \cdot 10^{-4}$ have been considered. These values correspond, in the GSM environment, to speeds of 0, 30, and 150 Km/h, respectively.

As far as the simulated packet structure is concerned, we recall that 26 of the 142 bits constituting the so-called "normal" burst of the actual GSM standard [see Fig. 8(a)] are, at the present, devoted to the training of the receiver. The use of training sequences leaves the GSM system with an overhead of $26/116 = 22.4\%$, which, in principle, could be used for other purposes (e.g., source and/or

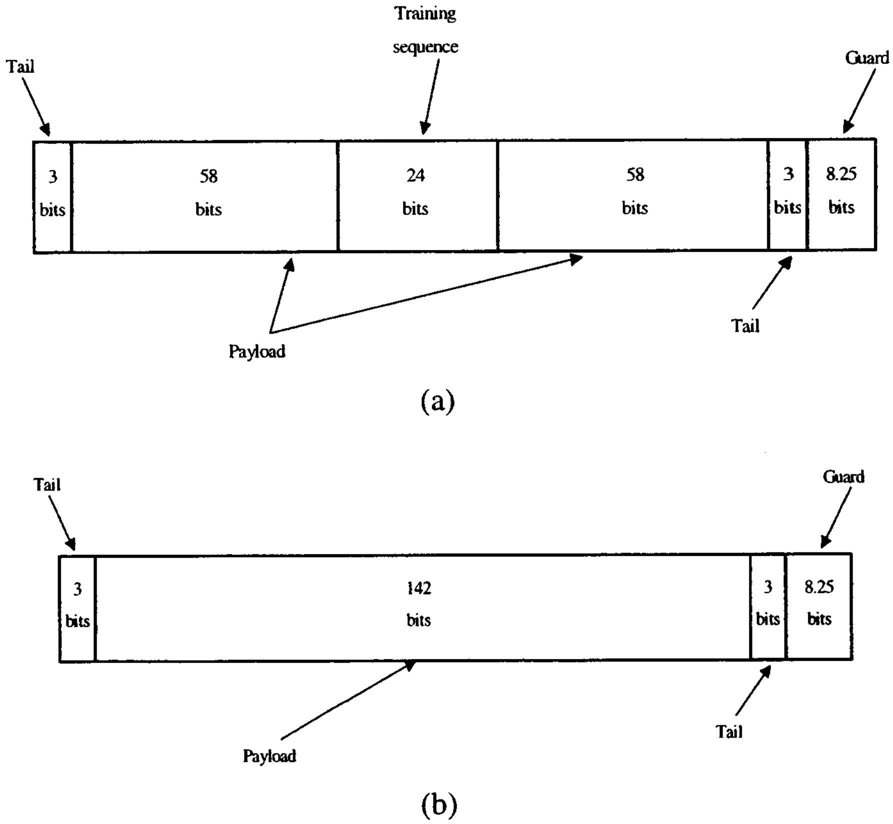

Fig. 8. (a) Normal burst for GSM-type trained (i.e., not blind) application. (b) Simulated GSM-type "blind" burst.

channel coding) if blind channel estimation were used for the detection of the transmitted packets. Therefore, in the simulations carried out in Fig. 9, we have considered the GSM "blind" burst described in Fig. 8(b), where the payload sequence comprehends all 142 data bits present in a GSM burst.

The simulated blind packet-detector operates as described in the following. At first, the 142 data bits of Fig. 8(b) are only used for the estimation of the channel impulse response $\underline{G}$ via an activation of the SOM/QHOM stages described in Section III-A and B. Afterwards, the resulting channel estimate $\hat{\underline{G}}(142)$ obtained at the end of the burst of Fig. 8(b) is used to detect the overall 142-long payload packet via the activation of the data-detection operating mode. In this last operative stage, we have utilized an SbS-MAP equalizer for data-detection with a decision delay D set to the memory of the considered channel, i.e., $D = 2$ (see the last row of Table I). In this way, the total delay embedded in the simulated blind packet detector is limited up to two packets periods (i.e., about 1.5 ms for the GSM standard). In general, when the channel is static (i.e., vanishing $B_D T_S$), the vector $\hat{\underline{G}}(142)$ obtained at the end of the SOM/QHOM stages represents an estimate of the true time-invariant impulse response $\underline{G}$ exhibited by the channel over the corresponding packet period, whereas in the case of a (slowly) time-variant link (i.e., $B_D T_S$ greater than zero), the obtained $\hat{\underline{G}}(142)$ can be interpreted as an estimate of the "average value" assumed by the channel-impulse response over the transmission packet.

The average BERs obtained for the described blind receiver with DE-BPSK modulated data-burst are reported in Fig. 9 for the previously described Rayleigh faded channels. Each BER value has been averaged over 100 000 independent transmitted packets impaired by different channel-realizations.[5]

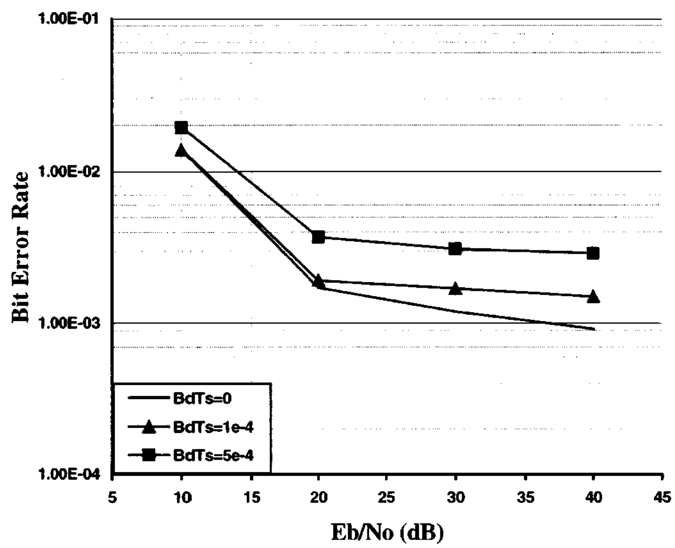

Fig. 9. Average BERs versus E_b/N_o for the three-tap Rayleigh faded channel of Section V-C with DE-BPSK modulation. Values of $B_D T_S$, 0, 10^{-4}, and $5 \cdot 10^{-4}$ are considered.

Since typical GSM-like voice applications may tolerate BERs up to 10^{-2} for the uncoded equalized links [28], an examination of the performance plots of Fig. 9 supports the conclusion that

[5]It is indeed necessary to generate many channel realizations because of the nonergodic nature of slowly time-varying channels as those here considered in the simulations we have carried out.

the proposed detector of Fig. 1 may be a possible candidate for blind GSM-like applications for values of the product $B_D T_S$ up to 10^{-3}.

VI. Conclusions

The reported simulations support the conclusion that the proposed blind equalizer achieves reliable estimates of unknown channel impulse responses within 150–300 observation samples. Although the presented blind detector employs *only one* channel estimator to achieve channel identification, the numerical results reported in Section V confirm that it exhibits appealing convergence rates and steady-state performances, even on nonminimum phase channels. Moreover, the time-variant and nonlinear nature of the channel estimator of (8) allows the receiver to achieve fast identification of slowly varying channels with *deep spectral notches*. In principle, this feature could make the proposed blind channel estimator suitable for application on slowly time-variant channels such as the land mobile cellular ones of urban and suburban areas, and the results of Section V-C support this conclusion.

Another attractive feature of the proposed blind channel estimator is that it *directly estimates* the (equivalent discrete-time) impulse response of the channel, which is output by the channel estimator of Fig. 1 at the end of the QHOM stage. In this way, the use of an SbS-MAP equalizer in the data-detection operating mode is *not strictly* mandatory, and in principle, any equalization technique (e.g., MLSE implemented via the Viterbi algorithm) can be used after achieving channel identification. However, since the proposed blind channel estimator *needs* the APPs of the states of the ISI channel to perform channel identification during the SOM and the QHOM, the utilization of an SbS-MAP equalizer to detect the transmission stream after the SOM/QHOM stages exhibits an appealing tradeoff in terms of implementation complexity versus achievable performances.

Appendix

Derivation of (18)

As far as the computation of the covariance matrix of the channel estimation error $\mathbf{Cov}_{\mathbf{G}}(i)$ is concerned, we note that due to the non-Gaussian distribution of the observations process in (1), its recursive updating from $\mathbf{Cov}_{\mathbf{G}}(i-1)$ should require, in principle, the computation of the *overall* higher order moments of the (non-Gaussian) distribution of $\underline{G}$ conditioned on $\underline{y}_1^i$. Since it would give rise to a *useless infinite-dimensional* channel estimator, to bypass this problem, we introduce a simplifying assumption that is quite common in the Bayesian approach to non-Gaussian and/or nonlinear MMSE filtering (see, for example, [7], [17], and references therein). More precisely, we assume that the third-order cumulants of the distribution of $\underline{G}$ conditioned on the observations $\underline{y}_1^i$ vanish. In fact, this assumption allows us to update $\mathbf{Cov}_{\mathbf{G}}(i)$ on the basis of $\mathbf{Cov}_{\mathbf{G}}(i-1)$ *only* via *the same relationship* employed by the standard Kalman filter for the estimation of Gauss–Markov sequences impaired by additive Gaussian noise. Therefore, under the above-stated simplifying assumption, the derivation of (18) from (8) is standard and can be carried out by following the same steps reported, for example, in [9, Ch. 7, Sect. 3] and [26, Sect. 3, eq. (60)] for the conventional Kalman filter.

References

[1] Y. Sato, "A method of self-recovering equalization for multilevel amplitude-modulation," *IEEE Trans. Commun.*, vol. COMM-23, pp. 679–682, June 1975.

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

[3] A. Benveniste and M. Goursat, "Blind equalizers," *IEEE Trans. Commun.*, vol. COMM-32, pp. 871–883, Nov. 1980.

[4] G. Picchi and G. Prati, "Blind equalization and carrier recovery using a stop-and-go decision directed algorithm," *IEEE Trans. Commun.*, vol. COMM-35, pp. 877–887, Sept. 1980.

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

[6] L. Tong, G. Xu, and T. Kailath, "Blind identification and equalization based on second-order statistics: A time domain approach," *IEEE Trans. Inform. Theory*, vol. 40, pp. 340–349, Mar. 1994.

[7] G. Lee, S.B. Gelfand, and M. P. Fitz, "Bayesan decision feedback techniques for deconvolution," *IEEE J. Select. Areas Commun.*, vol. 13, pp. 155–165, Jan. 1995.

[8] E. Moulines, P. Duhamel, J. F. Cardoso, and S. Mayrargue, "Subspace methods for blind identification of multichannel FIR filters," *IEEE Trans. Signal Processing*, vol. 43, pp. 516–525, Feb. 1995.

[9] A. Jazwinski, *Stochastic Processes and Filtering Theory*. New York: Academic, 1970.

[10] A. Segall, "Stochastic processes in estimation theory," *IEEE Trans. Inform. Theory*, vol. IT-22, pp. 275–286, May 1976.

[11] K. Abend and B. D. Fritchman, "Statistical detection for communication channels with intersymbol interference," *Proc. IEEE*, vol. 58, pp. 779–785, May 1970.

[12] J. K. Tugnait and U. Gummadavelli, "Blind equalization and channel estimation with partial response input signals," *IEEE Trans. Commun.*, vol. 45, pp. 1025–1031, Sept. 1997.

[13] R. A. Iltis, "A Bayesian maximum likelihood sequence estimation algorithm for a priori unknown channel and symbol timing," *IEEE J. Select. Areas Commun.*, vol. 10, pp. 579–588, Apr. 1992.

[14] N. Seshadri, "Joint data and channel estimation using blind trellis search techniques," *IEEE Trans. Commun.*, vol. 42, pp. 1000–1011, Feb./Mar./Apr. 1994.

[15] R. A. Iltis, J. J. Shunk, and K. Giridhar, "Bayesian algorithms for blind equalization using adaptive filtering," *IEEE Trans. Commun.*, vol. 42, pp. 1017–1032, Feb./Mar./Apr. 1994.

[16] M. Gosh and C. L. Weber, "Maximum likelihood blind equalization," *Opt. Eng.*, vol. 31, pp. 1224–1228, June 1992.

[17] G. K. Lee, S. B. Gelfald, and M. P. Fitz, "Bayesian techniques for blind deconvolution," *IEEE Trans. Commun.*, vol. 44, pp. 826–835, July 1996.

[18] R. E. Lawrence and H. Kaufman, "The Kalman filter for the equaliztion of digital communication channel," *IEEE Trans. Commun.*, vol. COMM-19, pp. 1137–1141, Dec. 1971.

[19] Y. Li, B. Vucetic, and Y. Sato, "Optimum soft-output detection for channels with intersymbol interference," *IEEE Trans. Inform. Theory*, vol. 41, pp. 704–713, May 1995.

[20] E. Baccarelli, R. Cusani, and S. Galli, "Novel analytical performance bounds for symbol-by-symbol decoding of digital data impaired by ISI and AWGN," *IEEE Trans. Inform. Theory*, vol. IT-14, pp. 618–626, Mar. 1997.

[21] E. Baccarelli, A. Fasano, S. Galli, and A. Zucchi, "A novel reduced-complexity MAP equalizer using soft-statistics for decision-feedback ISI cancellation," in *Proc. IEEE Globecom Conf.*, Rio de Janeiro, Brazil, Dec. 5–12, 1999.

[22] F. M. Hsu, "Square root Kalman filtering for high speed data received over fading dispersive HF channels," *IEEE Trans. Inform. Theory*, vol. IT-28, pp. 753–763, Sept. 1982.

[23] A. P. Clark and R. Harun, "Assessment of Kalman filter channel-estimators for HF radio links," *Proc. Inst. Elect. Eng. F*, vol. 133, no. 6, pp. 513–521, Oct. 1986.

[24] F. M. Gaston and G. W. Irwin, "Systolic Kalman filtering theory: An overview," *Proc. Inst. Elect. Eng. D*, vol. 137, no. 4, pp. 235–244, July 1990.

[25] J. R. Treichler and B. G. Agee, "A new approach to multipath correction of constant modulus signals," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-31, pp. 459–471, Apr. 1983.
[26] E. Weinstein, A. V. Oppeenheim, M. Feder, and J. R. Buck, "Iterative and sequential algorithm for multisensor signal enhancement," *IEEE Trans. Signal Processing*, vol. 42, pp. 846–859, Apr. 1994.
[27] E. Wong and B. Hajek, *Stochastic Processes in Engineering Systems*. Berlin, Germany: Springer-Verlag, 1994.
[28] T. S. Rappaport, *Wireless Communications: Principles and Practice*. Englewood Cliffs, NJ: Prentice-Hall, 1996.

Enzo Baccarelli was born in Todi, Italy, in 1962. He received the Laurea degree in electronic engineering and the Ph.D. degree in communication theory from the University of Roma "La Sapienza," Rome, Italy, in 1989 and 1993, respectively.

In 1995, he was with the Post-Doctorate course at the INFO-COM Department, the University of Rome, where he has been a Researcher since 1996 and where he is currently an Associate Professor of signal theory and wireless communications. His main research activities are in the areas of random processes and information theory, with applications to radio-mobile digital communications and coding. He is author of about 30 papers in international journals and 40 conferences proceedings on these topics. He holds a U.S. patent in adaptive equalization of radio channels.

Stefano Galli (M'98) was born in Florence, Italy, in 1966. He received the Laurea (M.S.) degree in electronic engineering and the Dottorato di Ricerca (Ph.D.) degree in information theory and communications from the University of Rome "La Sapienza," Rome, Italy, in 1994 and 1998, respectively.

His first research studies were on coherent optical communications. However, while he was a Ph.D. candidate, his interests moved to digital communications in time-variant environments, with particular emphasis on adaptive channel equalization. After completing the Ph.D. degree, he continued as a Teaching Assistant in Signal Theory at the INFO-COM Department of the University of Rome. At the same time, he began to work as a free consultant for Italian telecommunications companies. He also worked on the Trans-European Trunked Radio (TETRA) project, in particular, on the definition of the line dispatcher functionalities. In October 1998, he joined Bellcore (now Telcordia Technologies, an SAIC company), Morristown, NJ, as a Research Scientist in the Broadband Access and Premises Internetworking Department. His main research efforts are devoted to the problem of automatic loop qualification and, more recently, to the analysis and performance assessment of wireless home networks and power line carriers. His research interests also include detection and estimation, channel equalization, personal wireless communications, xDSL systems, and crosstalk modeling.