On-Line Blind Multichannel Equalization Based on Mutually Referenced Filters

David Gesbert, Pierre Duhamel, *Senior Member, IEEE*, and Sylvie Mayrargue

***Abstract*—This paper presents a novel approach to the blind linear equalization of possibly nonminimum phase and time-varying communication channels. In the context of channel diversity, we introduce the concept of *mutually referenced equalizers* (MRE's) in which several filters are considered, the outputs of which act as training signals for each other. A corresponding (constrained) multidimensional *mean-square error* (MSE) cost function is derived, the minimization of which is shown to be a necessary and sufficient condition for equalization. The links with a standard linear prediction problem are demonstrated. The proposed technique exhibits properties of important practical concern:**

1) **The proposed algorithm is globally convergent.**
2) **Simple closed-form solutions exist for the MRE's, but the MRE's also lend themselves readily to adaptive implementation. In particular, the recursive least-squares (RLS) algorithm can be used to offer optimal convergence rate.**
3) **The MRE method provides a solution for all equalization delays, which results in robustness properties with respect to SNR and ill-defined channel lengths.**

I. Introduction

TRANSMISSION over high-speed digital communication channels is subject to inter-symbol interference (ISI) inducing channel distortion, which has to be compensated for by an equalization device. In the context of digital radio-communications, the problem is very challenging due to the length of the ISI that stems from the data transmission at high rates in a multipath propagation environment.

Traditional equalization schemes rely on the periodic transmission of training sequences, which are known from the receiver and are used to acquire and update either the channel or the equalizer coefficients. This strategy, however, results in a significant reduction of the effective communication rate. The need for very high bit rates that arises in the field of digital communications makes very attractive the methods that do not resort to such training sequences (blind equalization).

Manuscript received March 28, 1996; revised March 28, 1997. The associate editor coordinating the review of this paper and approving it for publication was Prof. Georgios B. Giannakis.

D. Gesbert was with Department DMR/ATH, France Telecom-CNET, Issy-Les-Moulineaux, France. He is now with the Information Systems Laboratory, Stanford University, Stanford, CA 94305 USA (e-mail: gesbert@rascals.stanford.edu).

P. Duhamel is with the Signal Department, Telecom Paris, Paris, France (e-mail: duhamel@sig.enst.fr).

S. Mayrargue is with the Department DMR/ATH, France Telecom-CNET, Issy-Les-Moulineaux, France (e-mail: mayrargue@issy.cnet.fr).

Publisher Item Identifier S 1053-587X(97)06445-3.

Conventional blind equalization (BE) methods implicitly assume a symbol-rate monochannel transmission model and, therefore, are based on the use (explicit or not) of higher order statistics (HOS) of the received signals. Hence, adaptive techniques are typically designed as stochastic gradient descent schemes based on the minimization of various nonquadratic nonunimodal error functions [1]–[6], resulting in a possible slow/ill-convergence [7]. Recently, the use of fractionally spaced (FS) and/or multisensor (MS) reception arrays was recognized as a powerful means to solve the convergence-related problems of BE algorithms for two reasons. First, some of the HOS-based techniques such as the constant modulus algorithm (CMA) are found to be globally convergent in the FS or MS context [27], [28], and second, blind equalization is shown to be feasible, assuming FS or MS arrays, based on the sole second-order statistics of the received signals [8], [9].

Building on the results found in [8] and [9], several second-order time-domain methods were developed that generally rely on one (or more) subspace decomposition of either the received data matrix (“deterministic methods” [14], [15], [24]) or the data correlation matrix (“stochastic methods” [9], [11], [12]). For frequency-domain approaches, see [8], [10], and [13]. All these methods assume FS or MS arrays but can essentially be referred to as *multichannel* (MC) equalization methods, building on the concept of channel diversity, for which the transmitted source is seen through more than one symbol-rate linear filter. The MC second-order methods above offer an interesting alternative to the conventional HOS-based BE algorithms. However, they suffer from several drawbacks.

1) Most are computationally intensive and/or are unsuited to an adaptive implementation,
2) They only provide channel coefficient estimates and need to be further linked with an equalization device (although this does not apply to [15] and [24]).
3) They require exact knowledge of the channel order, although this condition is never met in practice.

However, the second-order methods mentioned above are not the only possible ones, and other second-order equalization techniques that do not rely on subspace decompositions include [16], [17], [19], and [21]. In this paper, we propose a novel approach for the *direct* BE problem (i.e., without channel preidentification) in the context of channel diversity. The proposed method builds on the concept of the mutually referenced equalizers (MRE's), which were first introduced in [23], in which a set of K ($K > 1$) filters are considered, the outputs of which act as training signals for each other.[1] A multidimensional mean-square error MRE criterion for blind MC equalization is derived, and certain minimization procedures are discussed. The obtained algorithm is shown to meet several conditions of important practical concern:

1) The MRE criterion is unimodal; hence, the algorithm exhibits global convergence.
2) The MRE criterion is a mean-square error. Full flexibility is gained for the implementation. In contrast, flexibility of design is difficult to obtain with subspace decomposition methods as well as with adaptive BE methods relying on a high-order nonlinear cost (this includes the Godard's algorithms).
3) The method directly provides channel inverses with all possible delays, hence, yielding robustness w.r.t. the noise amplification problems that may be related to a specific delay (typically the minimum and maximum delays yield poor results if the channel has coefficients that taper off at the ends).
4) Finally, the MRE method shows empirical robustness in the presence of channel length mismatch.

Following a short review of multichannel background in Section II, we investigate the theoretical properties of the MRE criterion in the noise-free case and show that its minimization is necessary and sufficient for linear MC equalization (Section III). We show that the MRE criterion is a multidimensional MSE that has to be constrained in order to avoid the convergence toward undesirable solutions. The optimization under different types of constraints leads to various algebraic solutions for the MRE, which are indicated in Section IV. Quadratic and linear constraints are considered. The choice of an optimal constraint with respect to the performance/implementation issue is discussed.

In Section V, we elaborate on this result by demonstrating that under a linear constraint and under mild assumptions about the transmitted symbols, the proposed method can be put in the form of a certain linear prediction problem. A standard RLS implementation of the method is proposed that offers an optimal convergence rate.

In Section VI, the MRE method is evaluated in a multipath propagation environment with signals impinging on a circular antenna array. Batch simulations for various sample and array sizes and signal-to-noise ratios (SNR's) show the consistency of the proposed criterion and illustrate the good behavior of the MR equalizers in a realistic context (noisy channels with unwell-defined orders). The behavior of the adaptive MRE method is also investigated. We show that the MRE based on the RLS algorithm reaches convergence in 100 iterations under certain conditions. Throughout the paper, the following notations are adopted:

X^t transpose of X;
X^+ conjugate transpose of X;
$|X|$ L_2-norm of X;
X^* complex conjugate of X;
I_K identity matrix of size K;
0 all-zero matrix;
$E()$ statistical expectation.

[1] A similar idea was later and independently obtained by Giannakis [22].

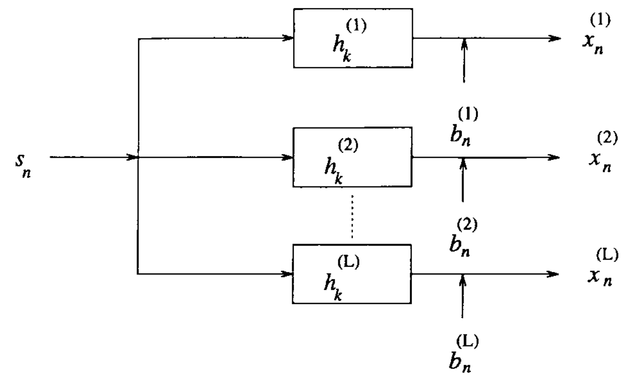

Fig. 1. Baud-rate digital multichannel model.

II. Multichannel Equalization

A. Background

Consider the continuous-time signal observed at the output of a noisy PAM/QAM communication channel

$$x(t) = \sum_{k=-\infty}^{+\infty} s_k h(t-kT) + b(t) \qquad (1)$$

where $\{s_k\}$, $\{b(t)\}$, and $\{h(t)\}$, respectively, denote the transmitted symbols sequence with rate $1/T$, the additive noise (assumed to be uncorrelated with $\{s_k\}$), and the baseband equivalent channel, including the equipment filtering, modulation, and demodulation effects. The transmitted sequence is not required to be white. Channel diversity may be represented as

$$x_n^{(i)} = \sum_{k=-\infty}^{+\infty} s_k h_{n-k}^{(i)} + b_n^{(i)}, \quad i = 1 \cdots L \qquad (2)$$

where

$h^{(i)}$ T-sampled impulse response of the ith channel;
$x_n^{(i)}$ baud-rate signal measured at the output of $h^{(i)}$;
$b_n^{(i)}$ corresponding noise sequence;
L number of channels ($L > 1$).

Note that the *same* sequence $\{s_k\}$ is observed through the different filters $h^{(1)}$, $h^{(2)}$, $\cdots$, $h^{(L)}$.

In the context of multisensor arrays, $h^{(i)}$ represents the link between the transmitter and the ith sensor. In the fractionally spaced scenario, the channels $h^{(i)}$ correspond to sampled versions (at rate T) of the same continuous-time channel $h(t)$, with various sampling phases $(i-1)T/L$, $i = 1 \cdots L$ (see [9] for details). The multichannel formalism is also easily extended to the case of oversampled multisensor arrays. The baud-rate multichannel setup is displayed in Fig. 1. Throughout the paper, we assume that the ISI is causal with length M ($M \geq 0$).

B. Vector Representation

Assume that N measurements are performed per transmitted symbol on each channel. The LN-long data vector at time n, which is denoted by $X_n = [x_n^{(1)}, \cdots, x_{n-N+1}^{(1)}, \cdots, x_n^{(L)}, \cdots, x_{n-N+1}^{(L)}]^t$, satisfies the linear equation

$$X_n = \mathcal{H} S_n + B_n \tag{3}$$

where $B_n = [b_n^{(1)}, \cdots, b_{n-N+1}^{(1)}, \cdots, b_n^{(L)}, \cdots, b_{n-N+1}^{(L)}]^t$ is the corresponding noise vector at time n, and S_n contains the set of symbols involved in the output measurements. Since the channels have length $M+1$, we have $S_n = (s_n, s_{n-1}, \cdots, s_{n-N-M+1})^t$, and $\mathcal{H}$ is the channel convolution matrix

$$\mathcal{H} = \overset{K=M+N}{\overleftrightarrow{}}\begin{pmatrix} h_0^{(1)} & \cdots & h_M^{(1)} & 0 & \cdots & 0 \\ \vdots & \cdots & \ddots & \cdots & \ddots & 0 \\ 0 & \cdots & 0 & h_0^{(1)} & \cdots & h_M^{(1)} \\ \vdots & \cdots & \vdots & \cdots & \cdots & \vdots \\ h_0^{(L)} & \cdots & h_M^{(L)} & 0 & \cdots & 0 \\ \vdots & \cdots & \ddots & \cdots & \ddots & 0 \\ 0 & \cdots & 0 & h_0^{(L)} & \cdots & h_M^{(L)} \end{pmatrix} \updownarrow L \times N.$$

Let $K = M + N$ denote the number of columns in $\mathcal{H}$. $\mathcal{H}$ is a Sylvester matrix built from L vertically stacked Toeplitz blocks, each one of size $N \times K$. The condition of blind equalizability of the multichannel model (3) is also the condition under which $\mathcal{H}$ is left invertible. Therefore, we will assume throughout this paper that

H1) *$\mathcal{H}$ has full column rank K.*

Note that H1) requires $LN \geq K$ so that $\mathcal{H}$ is vertical, or square, at least. Let $h^{(i)}(z^{-1}) = h_0^{(i)} + \cdots + h_M^{(i)} z^{-M}$ be the z transform of the ith channel impulse response. When $N \geq M$, H1) can also be reinterpreted as [13]: *There should not exist any z_0 such that $h^{(i)}(z_0^{-1}) = 0$ for all i.* In practice, the performances of second-order methods critically depend on *how close* the zeros of the various channels are. Practical concerns include the possibility of 1) channels having quasizeros in 0, i.e., very small coefficients $h_M^{(i)}$, $i = 1, L$ and 2) channels having quasizeros in ∞, i.e., very small coefficients $h_0^{(i)}$, $i = 1, L$. Both cases correspond to a likely situation in which the multichannel order is not well defined or is overestimated. This problem is discussed in more detail in Section VI.

C. Linear Equalization

Here, we focus on the problem of blind adaptive symbol recovery using linear multichannel filters. Each channel is filtered by an N-tap filter, and the L resulting outputs are added to form a symbol estimate. Let the filter coefficients be listed in the $LN \times 1$ complex-valued vector v. At time n, we have

$$v^+ X_n = \rho \hat{s}_{n-d} \tag{4}$$

where d and ρ are, respectively, arbitrary constant time delay and gain. A main advantage of the multichannel structure over the classical one ($L = 1$) lies in the availability of linear equalizers achieving zero ISI, i.e., zero-forcing (ZF) equalizers, so that the symbol recovery can be perfect in the absence of noise ($\hat{s}_n = s_n$). Algebraically, these ZF equalizers are provided by any left inverse of the channel convolution matrix $\mathcal{H}$ since

$$v^+ X_n = \rho s_{n-d} \tag{5}$$

also reads

$$v^+ \mathcal{H} = \rho(0, \cdots, 0, 1, 0, \cdots, 0), \quad \text{for } d = 0 \cdots K-1 \tag{6}$$

in the noise-free case under a sufficient excitation condition on the transmitted sequence. That the transmitted sequence can be linearly recovered, up to a finite set of time delays, ranging from 0 to $K-1$ symbol durations is also clear from (6). Hence, channel diversity induces delay diversity. Note also that undesirable *blocking* solutions that suppress the desired signal with $\rho = 0$ may appear when $\mathcal{H}^+$ has a nontrivial null space ($LN > K$ case). Then, for each possible delay, various equalizers can theoretically be computed, lying in a subspace of dimension $1 + dim[null(\mathcal{H}^+)]$.

III. The Mutually Referenced Equalizers

A. Noise-Free Case

In the previous section, we have noted the existence of independent linear equalizers associated with different reconstruction delays. We attempt now to show how this delay diversity can be exploited to permit the blind determination of the full set of channel inverses.

The simple idea of the mutually referenced equalizers evolves as follows: Consider v_i (resp. v_{i+1}), which is a multichannel equalizer satisfying (5), in the noise-free case, with $d = i$ (resp. $d = i+1$). In the following, v_i is referred to as an *i-delay equalizer*. Then, an obvious assertion is

$$v_i^+ X_n = v_{i+1}^+ X_{n+1}$$

where the outputs of the i-delay and the $(i+1)$-delay equalizers can be said to be "referenced" to each other up to a symbol duration delay. Other similar relations can be found, involving other equalizers and delays. We have, in general

$$v_k^+ X_n = v_i^+ X_{n+i-k}, \qquad \text{for } i, k = 0 \cdots K-1, \text{ and } k > i \tag{7}$$

where the v_k are also defined as k-delay equalizers. As can be seen from (7), a redundant set of $K(K-1)/2$ simple relationships may serve as necessary conditions for K filters $v_0, v_1, \cdots, v_{K-1}$ to be ideal ZF multichannel equalizers in the noiseless case. Surprisingly, we can demonstrate the *sufficiency* of such conditions.

Lemma 3.1: Assume K is known. Let $v_0, v_1, \cdots, v_{K-1}$ be $LN \times 1$ complex-valued vectors. Assume the independent relationships $v_i^+ X_n = v_{i+1}^+ X_{n+1}$ for all n and $i = 0 \cdots K-2$ are satisfied. Rewrite the vectors in matrix form as $V = (v_0, v_1, \cdots, v_{K-1})$. Then, we have[2]

i) $V^+ \mathcal{H} = \alpha I_K$ for some complex number α or equivalently

ii) $v_i^+ X_n = \alpha s_{n-i}$, $i = 0 \cdots K-1$.

[2] A similar result can be obtained by using conditions on v_0 and v_{K-1} only.

Proof: Let $\tilde{S}_n = (s_{n+1}, S_n^t)^t$. The conditions on the equalizers $\{v_i\}$ are easily rewritten as

$$(I_{K-1}, \mathbf{o})V^+X_n = (\mathbf{o}, I_{K-1})V^+X_{n+1}$$

then, from (3)

$$(I_{K-1}, \mathbf{o})V^+\mathcal{H}S_n = (\mathbf{o}, I_{K-1})V^+\mathcal{H}S_{n+1}$$
$$(I_{K-1}, \mathbf{o})V^+\mathcal{H}(\mathbf{o}, I_K)\tilde{S}_n = (\mathbf{o}, I_{K-1})V^+\mathcal{H}(I_K, \mathbf{o})\tilde{S}_n. \quad (8)$$

Provided $\{s_n\}$ is persistently exciting of order at least $K+1$, (8) yields

$$(I_{K-1}, \mathbf{o})V^+\mathcal{H}(\mathbf{o}, I_K) = (\mathbf{o}, I_{K-1})V^+\mathcal{H}(I_K, \mathbf{o}) \quad (9)$$

which reads equivalently

$$\begin{pmatrix} 0 & W_{11} & \cdots & W_{1K} \\ \vdots & \vdots & & \vdots \\ 0 & W_{(K-1)1} & \cdots & W_{(K-1)K} \end{pmatrix} = \begin{pmatrix} W_{21} & \cdots & W_{2K} & 0 \\ \vdots & & \vdots & \vdots \\ W_{K1} & \cdots & W_{KK} & 0 \end{pmatrix}$$

where W_{ij} stands for the (i, j) term of $V^+\mathcal{H}$. It is now easily found that $W = \alpha I_K$ for some complex α. Step ii) follows immediately from (3). □

Comments: By Lemma 3.1, *all* solutions $v_0, v_1, \cdots, v_{K-1}$ to our problem in (7) are ideal ZF multichannel equalizers with delay $0, 1, \cdots, K-1$, respectively. Hence, the MRE criterion provides the full set of channel FIR inverses. Note that as usual in a blind context, a gain factor α remains undetermined. This problem is usually solved by resorting to differential modulation in order to cope with the phase shift, whereas an automatic gain control device may be used to compensate for the amplitude.

B. Noisy Case

In the presence of channel additive noise, perfect symbol recovery is an impossible task. However, (7) may be transformed into a quadratic criterion measuring the distance to equality in all these equations. Lemma 3.1 then suggests a simple MSE-like equalization criterion

minimize

$$J(V) = E|E_n|^2$$

with

$$E_n = (I_{K-1}, 0V^+X_n - (0, I_{K-1})V^+X_{n+1}$$

under constraint

$$\mathcal{C}(V) = 0$$

where V is the $LN \times K$ equalizers matrix to be updated, and $\mathcal{C}(V) = 0$ is a nontriviality constraint. Throughout the paper, this criterion is referred to as the mutually referenced equalizers (MRE) criterion since it involves several linear equalizers with different delays and measures the closeness between their properly delayed output signals. Fig. 2 displays the MRE equalization structure when $K = 5$. It is seen that the output of each filter serves as a training signal to update another filter. The fact that $J(V)$ is merely a K-dimensional minimum mean-square error (MMSE) cost constitutes one desirable feature of a BE criterion. From the results in 3.1, a minimum of the MRE criterion is attained, in the noise-free case, if and only if $J(V) = 0$, which in turn leads to

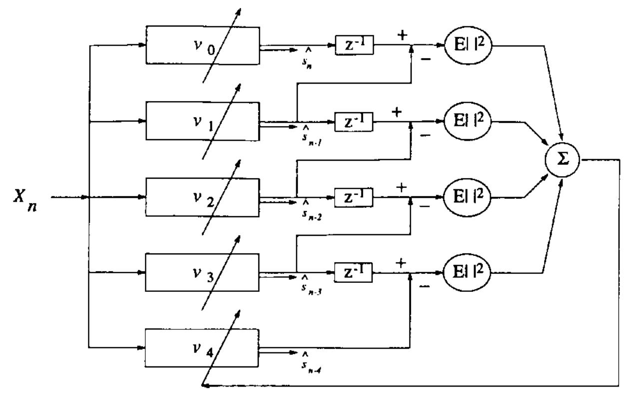

Fig. 2. MRE setup for $K = 5$.

$$V^+\mathcal{H} = \alpha I_K. \quad (10)$$

This shows that any column of V can be chosen to equalize the system with a predetermined delay. In a practical situation, however, the channel additive noise modifies the performance surface of the MRE function in such a way that (10) is no longer strictly valid. Hence, the mutually referenced equalizers differ from ZF equalizers in the presence of noise. The performances of the noised MRE are checked by simulations in this paper.

C. Choice of a Constraint

The main role of the constraint is to disqualify undesirable minima such as $V = 0$, as well as other nonzero blocking matrices V (i.e., $V^+\mathcal{H} = \mathbf{o}$) in order to achieve true equalization. In addition, the choice of $\mathcal{C}(V)$ should preserve the convexity property so that the constrained criterion exhibits no spurious local minimum. Last but not least, this choice should keep the minimization procedure as simple as possible. Typical choices include linear and quadratic constraints:

C1) $\mathcal{C}(V) = |V| - 1$ (quadratic).
C2) $\mathcal{C}(V) = \text{trace}\,(U^+V) - 1$, where U is an arbitrary $LN \times K$ matrix (linear).

Note that C1) and C2) cannot theoretically prevent the MRE from converging toward undesirable blocking matrices when the noise level is close to zero. The reason is that we can generally find a minimizing matrix V in the null space of $\mathcal{H}^+$ satisfying $|V| = 1$ or trace$(U^+V) = 1$. Poor choices of U are then characterized by $U \in \text{null}\,(\mathcal{H}^+)$ (which is unlikely to occur since null$\,(\mathcal{H}^+)$ has a much smaller dimension than the number of variables in the equalization problem). In such a situation, we may think of other choices for a constraint such as

C3) $\mathcal{C}(V) = E|V^+X_n|^2 - 1$ (keeps the equalizers output power to a constant).

However, we notice that blocking matrices move away from the criterion global minimum in the noisy situations; thus, C1) and C2) can always be used in practice. Since the quadratic constraint preserves all generality in the solution to the BE problem, it may appear to be more natural than the linear constraint, which is subject to possible poor choices of U. In fact, we note in the next sections that the linear constraint offers both more flexibility in the implementation and better conditioned solutions within the class of channels simulated.

D. Variations on the Criterion

A set of conditions that fully characterize the solutions to the multichannel equalization problem have been written in (7). A nonredundant subset of these relationships has been exploited in the derivation of a basic stochastic criterion (MRE). Several other variations may also be considered, which will not be detailed here due to the lack of space. These include the following:

- **Adding memory** in the criterion using the whole set of possible relationships in order to introduce more robustness to noise. This amounts to referencing each equalizer output to all the remaining ones, with proper delays, whereas the primitive version of the criterion only involves "neighbors."
- **Weighting** the contributions of the filters associated with different delays in the criterion, since an equalizer performance generally strongly depends on its delay in the noisy context.
- **Using a deterministic version** of the MRE criterion when, for instance, only short data records are available. This version is easily obtained since the MRE approach does not rely on specific assumptions concerning the input statistics.

IV. Asymptotic Solutions of the MRE Criterion

Section III introduced the MRE criterion as a means to solve the blind equalization problem. Here, we focus on the algebraic expressions for the criterion minimizers. Although adaptive equalization is our main goal, closed-form expressions for finite-sample and asymptotic solutions are interesting tools for investigating the intrinsic abilities of the proposed method.

The MRE criterion was first presented as a matrix optimization problem. However, a formalism using vectors only greatly simplifies algebraic developments. Recall that each column (among K) of matrix V consists of an equalizer $V = (v_0, v_1, \cdots, v_{K-1}$. Now, consider $\mathcal{V} = (v_0^t, v_1^t, \cdots, v_{K-1}^t)^t$, all the elements of V being strung out in a long vector of size LNK. In the following lemma, a simple expression of the MRE error function is provided in terms of $\mathcal{V}$. The proof is provided in Appendix A.

Lemma 4.1: Let $R_0 = E(X_n X_n^+)$ and $R_1 = E(X_{n+1} X_n^+)$ be, respectively, the zero- and first-order covariance matrices of the observed signals. Let $\mathcal{R}$ denote the $LNK \times LNK$ matrix given by

$$\mathcal{R} = \begin{pmatrix} R_0 & -R_1^+ & \mathbf{o} & & \mathbf{o} \\ -R_1 & 2R_0 & \ddots & & \\ \mathbf{o} & \ddots & \ddots & & \mathbf{o} \\ & & & 2R_0 & -R_1^+ \\ \mathbf{o} & & \mathbf{o} & -R_1 & R_0 \end{pmatrix}.$$

Then, the unconstrained MRE cost function is also found to be

$$J(V) = \mathcal{J}(\mathcal{V}) = \mathcal{V}^+ \mathcal{R} \mathcal{V}.$$

A. Quadratic Constraint Case

The MRE criterion under unit-norm constraint (denoted by C1–MRE) corresponds to a Rayleigh quotient minimization problem, involving the Hermitian positive matrix $\mathcal{R}$. The unique *stable* minimum of the C1-MRE criterion is readily found by extracting the smallest eigenvector of $\mathcal{R}$.

B. Linear Constraint Case

Based on the previous results, the minimization of the linearly constrained (C2-MRE) criterion is equivalent to that of $\mathcal{J}(\mathcal{V}) = \mathcal{V}^+ \mathcal{R} \mathcal{V}$ subject to $\mathcal{U}^+ \mathcal{V} = 1$, where $\mathcal{U}$ is the column-wise version of matrix U. The solution to this problem is well known as

$$\mathcal{V}_{opt} = \frac{\mathcal{R}^{-1} \mathcal{U}}{\mathcal{U}^+ \mathcal{R}^{-1} \mathcal{U}} \tag{11}$$

which requires solving a linear system instead of extracting an eigenvector.

C. Practical Implementation

1) Batch Implementations: A practical batch implementation of the mutually referenced equalizers goes as follows:

1) Compute finite sample estimates of R_0 and R_1 using standard time-domain averaging expressions.
2) Form an estimate of the matrix $\mathcal{R}$ based on the formula shown in Lemma 4.1. According to the chosen method, extract its smallest eigenvector (C1-MRE method) or solve the linear system (11) (C2-MRE method). Denote $\mathcal{V}_{opt}$ to be the obtained solution.
3) Reshape $\mathcal{V}_{opt}$ into the $LN \times K$ matrix V_{opt}, and choose one column of V_{opt} as a linear combiner for equalization. Avoid choosing the 0-delay and $(K-1)$-delay equalizers, which generally provide very poor noise enhancement properties when the channel ends are "small" or when the multichannel order is not well defined.

Note that the use of the output power constraint C3) would only slightly modify the calculations into the extraction of a smallest *generalized* eigenvector.

2) Adaptive Implementations (Initial Formulation): Different schemes can be used to implement the adaptive mutually referenced equalizers with the criterion of Section III-B. Constrained gradient techniques give the lowest computational cost and are thus attractive. Here, we have the expressions for the constrained LMS updates in both the quadratic and linear constraint cases. Remark that we directly update the matrix of equalizers V, rather than its column-wise version $\mathcal{V}$. Let μ be a small stepsize and V_n be the equalizer matrix estimate at time n

$$\begin{aligned} E_n &= (I_{K-1}, \mathbf{o}) V_n^+ X_n - (\mathbf{o}, I_{K-1}) V_n^+ X_{n+1} \\ \frac{\partial J}{\partial V^+}(V_n) &= X_n E_n^+ (I_{K-1}, \mathbf{o}) - X_{n+1} E_n^+ (\mathbf{o}, I_{K-1}) \\ \tilde{V}_{n+1} &= V_n - \mu \frac{\partial J}{\partial V^+}(V_n) \\ V_{n+1} &= \frac{\tilde{V}_{n+1}}{|\tilde{V}_{n+1}|} \text{ (unit-norm constraint)} \\ V_{n+1}(i,j) &= \tilde{V}_{n+1}(i,j) \text{ and } V_{n+1}(1,1) = 1 \text{ (linear constraint).} \end{aligned}$$

Comments: The C1-MRE method is implemented using a simple scaling to maintain unit-norm in V_n. The reason for this simplification is that scaling coincides with an orthogonal projection on the surface of the constraint (the unit-sphere). The C2–MRE method is shown here with $\mathcal{U} = (1, 0, \cdots, 0)^t$ but can readily be extended to arbitrary $\mathcal{U}$.

As we simultaneously track $K > 1$ filters in the MRE criterion, the overall complexity is in $O(LNK)$ instead of $O(LN)$ for a classical gradient-based equalization algorithm. As shown in the simulations, the convergence of such gradient algorithms is rather slow due to 1) the large number of parameters to be updated and 2) the lack of persistent excitation in the data vector X_n, which is structural in multichannel equalization methods. In fact, since the signal model in (3) is low rank, the correlation matrix R_0 is exactly singular in the noise-free case. As a remedy, other types of optimization techniques may be considered that show better performances in terms of convergence speed. In particular, since the computation of the mutually referenced equalizers involves that of some minimal eigenvector in the quadratic constraint case, several existing eigenvector tracking methods can be used to perform the blind equalization with the MRE criterion (see, for example, [29]). The conjugate gradient algorithm (CGA) used in [20] was extended to the case of complex signals/systems and was particularly suitable [23]. The CGA induces a noticeable gain in the convergence rate because it is much less sensitive w.r.t. the input signal statistics. Recall that the CGA theoretically converges in a finite number of steps when used to minimize a fixed quadratic cost [30]. The computational cost of the CGA in $O(LNK)^2$ is, however, prohibitive and cannot easily be reduced. In the next section, we show that more flexibility can be gained in the implementation through a convenient reformulation of the MRE criterion into a monodimensional linear prediction problem.

V. Links with Linear Prediction

Flexibility in the adaptive implementation of our method is typically possible if the matrix underlying the quadratic criterion ($\mathcal{R}$) can be put in the form of a mathematical expectation of a rank-one form. We provide such an expression in this section. While doing so, we show that the mutually referenced equalizers under the linear constraint can be computed using standard linear prediction techniques.

Lemma 5.1: Assume white noise and that H2) $\exists m$ such that $\forall n$, the symbols s_n and s_{n+m} are uncorrelated (the transmitted sequence has finite memory). Let $D \geq K + m$ be a time-shifting parameter, and let Y_n denote the modified data vector: $Y_n = (X_n^t, -X_{n+1}^t + X_{n+D}^t, \cdots, -X_{n+(K-3)D+1}^t + X_{n+(K-2)D}^t, -X_{n+(K-2)D+1}^t)^t$. Then, Y_n is a $LNK \times 1$ stationary process with covariance matrix $E(Y_n Y_n^+) = \mathcal{R}$.

A proof for Lemma 5.1 is provided in Appendix B. Based on this result, the vector-wise C2-MRE criterion now reads

$$\min \mathcal{J}(\mathcal{V}) = \mathcal{V}^+ \mathcal{R} \mathcal{V} = E|\mathcal{V}^+ Y_n|^2, \text{ under the constraint } \mathcal{U}^+\mathcal{V} = 1. \tag{12}$$

An unconstrained C2-MRE criterion is easily derived from (12). For the sake of simplicity, we choose again $\mathcal{U} = (1, 0, \cdots, 0)^t$, and we adopt the notations

$$Y_n \stackrel{\text{def}}{=} (x_n^{(1)}, Z_n^t)^t, \quad \mathcal{V} \stackrel{\text{def}}{=} (1, -\mathcal{W}^t)^t. \tag{13}$$

Now, the $(LNK-1)$-tap filter $\mathcal{W}$ can be updated to minimize the prediction variance

$$\mathcal{J}(\mathcal{W}) = E|x_n^{(1)} - \mathcal{W}^+ Z_n|^2. \tag{14}$$

In this new formulation, the mutually referenced equalizers are obtained by predicting a channel output from the signals incoming on all channels. Note the difference from the linear prediction problem of [16], [18], and [19]. Here, the present sample on all but the first channels are also used in the prediction. In [16] and [18], the minimum prediction error is found to be $h_0^{(1)} s_n$, and the practical equalization performance critically depends on $h_0^{(1)}$. In contrast, the MRE prediction error in (14) turns out to be zero in the absence of noise, independent of the channel characteristics. Note also that H2) is not a very restrictive assumption in the sense that a correlation between symbols within any finite neighborhood may be tolerated. Only an extra processing delay of $(K-2)D$ is the price to be paid for exploiting this formulation in terms of the modified data vectorY_n.

A. Adaptive Implementations (New Formulation)

The formulation of the C2–MRE algorithm above offers great flexibility for a practical implementation since several known algorithms can be used to solve the prediction problem (14) adaptively. Among those, an optimal convergence rate can be obtained using a standard RLS algorithm

$$e_n = x_n^{(1)} - \mathcal{W}_{n-1}^+ Z_n, \quad G_n = \frac{\mathcal{S}_{n-1} Z_n}{\gamma + Z_n^+ \mathcal{S}_{n-1} Z_n}$$
$$\mathcal{S}_n = \gamma^{-1}(I - G_n Z_n^+)\mathcal{S}_{n-1}, \quad \mathcal{W}_n = \mathcal{W}_{n-1} + G_n e_n^*$$

where Z_n is the modified data vector defined as in (13), and where the vector $\mathcal{W}_n$ contains all the equalizers coefficients but one at time n (the missing coefficient being forced to one). γ is a forgetting factor ($\gamma \leq 1$). $\mathcal{S}_n$ corresponds to the inverse of the correlation matrix $E(Z_n Z_n^+)$ estimate. Like for the CGA, the initial complexity of an RLS implementation in $O(LNK)^2$ seems prohibitive at the first glance. However, recent advances in the field have permitted the development of fast stabilized *multichannel* RLS versions that exploit displacement structures in the data vectors. In the context of our work, data vector Z_n consists of LK subblocks with a full displacement structure. This allows the complexity to be cut to roughly $O(LNK)$ operations [25].

VI. Simulations

Here, we evaluate the performances of the MRE method in a digital wireless communications situation at 900 MHz with L sensors distributed on an uniform circular array. The propagation channel is generated based on the model of Clarke

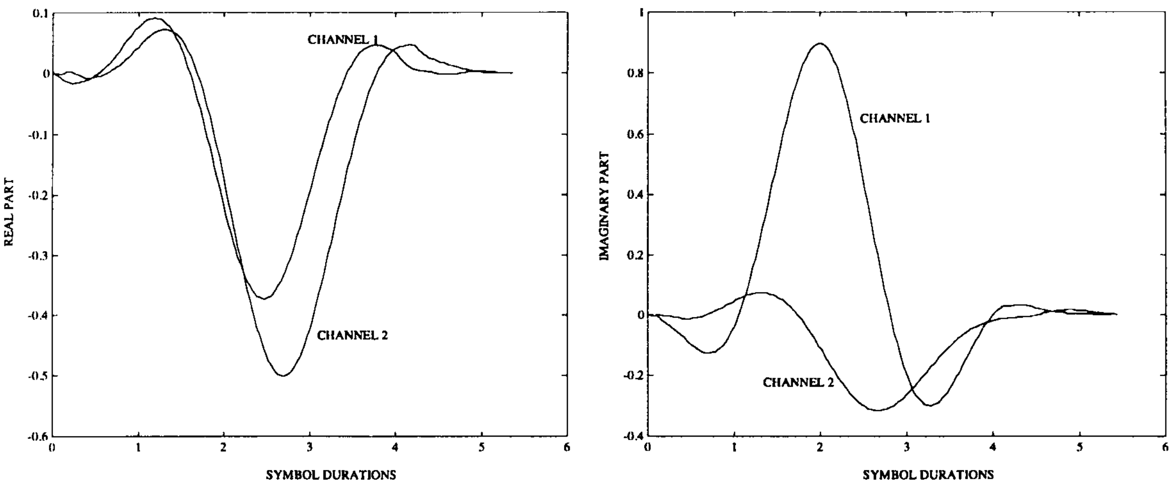

Fig. 3. Channels impulses responses.

TABLE I
"TU" CHANNELS PATH PROFILE

path	1	2	3	4	5	6	7	8	9	10	11	12
delay (μs)	0	0.2	0.4	0.6	0.8	1.2	1.4	1.8	2.4	3.0	3.2	5
attenuation (dB)	-4	-3	0	-2	-3	-5	-7	-5	-6	-9	-11	-10

[32]. At sensor i, the multipath channel is obtained through

$$k^{(i)}(t) = \sum_{p=1}^{nb_{paths}} A_p \delta(t-\tau_p) \sum_{n=1}^{nb_{rays}} \frac{1}{\sqrt{nb_{rays}}} e^{j(\phi_{n,p}+\Delta\phi_{n,p}^{i})} \tag{15}$$

where path p has amplitude A_p and delay τ_p. $nb_{rays} = 20$ is the number of rays impinging on the sensor within each path. $\phi_{n,p}$ is an i.i.d. uniform process in $[0; 2\pi]$. Assuming planar wavefronts, $\Delta\phi_{n,p}^{i} = 2\pi d/\lambda \cos(\theta_{n,p} - (i-1)2\pi/L)$ is the propagation delay of a ray with random angular incidence $\theta_{n,p}$ from the array origin to sensor i. d denotes the array radius and λ the wavelength (≈ 33 cm). The symbols are unit-variance white QPSK with duration 3.7 μs. Root–Nyquist filters are used for pulse shaping and reception filters (0.5 rolloff). The equalizer order is $N = 5$. The overall continuous-time channels are sampled at the baud rate and are normalized to have unit gain. White Gaussian noise is added with variance σ_b^2 chosen according to SNR. ($\text{SNR} = -10\log(\sigma_b^2)$ since the channels have unit gain).

In the first set of simulations, we work on a single realization of the channel model above in the case of two sensors spaced by a half wavelength. The path profile is drawn from the typical urban ("TU") model recommended by COST [33] (see Table I). The channel impulse responses are plotted in Fig. 3 and are seen to span over more than five symbols but taper off at the ends. The theoretical channel length (noiseless case) is $M + 1 = 5$. However, the practical length is not well defined and typically depends on the noise level.

Figs. 4–6 illustrate the intrinsic behavior of the MRE criterion based on the batch implementations described in Section IV-C1). The linear constraint is implemented with $\mathcal{U} = (1, 0, \cdots, 0)^t$ as this choice provided satisfying results over a wide range of channels/signals.

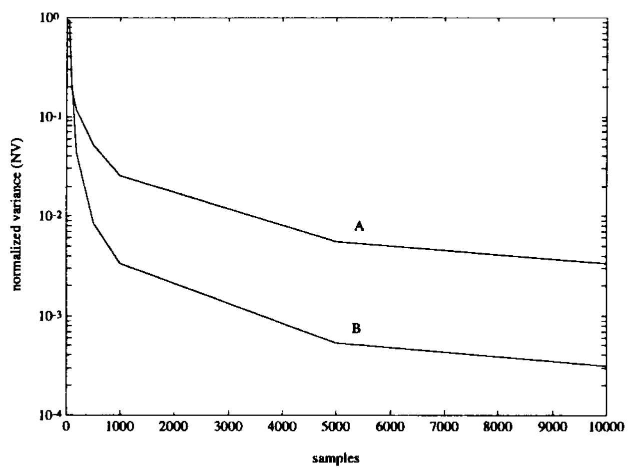

Fig. 4. Normalized variance of the MRE estimates at 10 dB output SNR. A is the unit-norm constraint, and B is the linear constraint.

Fig. 4 checks the consistency of the MRE method. A Monte Carlo experiment on 100 independent trials is conducted to compute the normalized variance (NV) of the estimated matrix of MRE, which is denoted $V(S)$, for various sample sizes S. We define $NV = < |V(S) - V|^2 > |V|^2$, where V is the true MRE matrix obtained with the procedure Section IV-C1, assuming exact statistics. $<>$ denotes the average over the Monte Carlo runs. The noise level is 10 dB SNR, and the simulations are run with $K = 6$ mutually referenced equalizers corresponding to a practical channel length of $\hat{M} + 1 = 2$ (corresponding to the number of coefficients above the noise

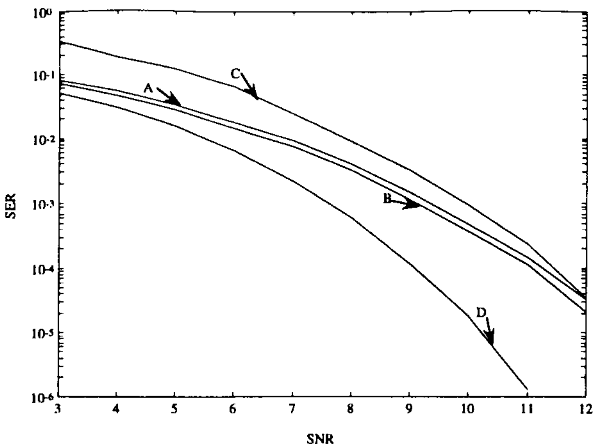

Fig. 5. BER achieved by batch blind methods, assuming the practical channel length ($\hat{M}+1=2$). A is the C1-MRE (unit-norm constraint), B is the C2-MRE (linear constraint), C is the subspace method (MUSIC-like), and D is the reference provided by the nonblind MMSE equalizer.

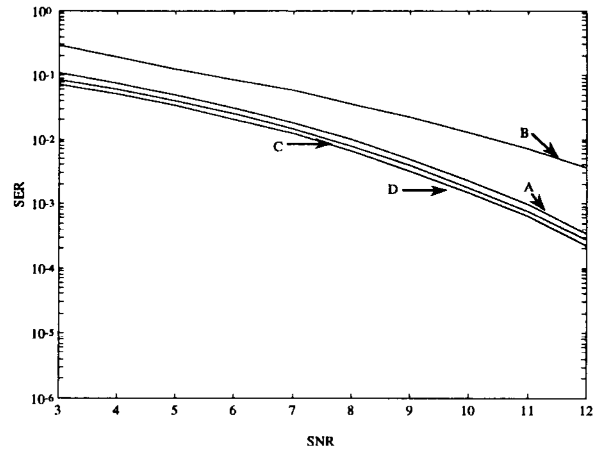

Fig. 6. BER achieved by batch blind methods in the large sample limit, assuming the theoretical channel length ($\hat{M}+1=5$). A is the C2-MRE method, B is the subspace method (MUSIC-like), C is the C2-MRE method with overestimated channel length ($\hat{M}+1=10$), and D is the C2-MRE method with largely overestimated channel length ($\hat{M}+1=15$).

level). The linear problem in Section IV-B seems better conditioned than the minimal eigenvector problem in Section IV-A. This was also observed for most channel realizations with the model used above. Note that Fig. 4 does not represent the convergence rate of the *equalization* method. In fact, in the noiseless case, a full subspace of solutions (hence, with large NV) would be satisfactory.

Fig. 5 plots the symbol error rate (SER) after linear equalization and decision for a range of SNR. We compare the MRE method with another second-order based technique, namely, the subspace identification plus equalization technique of [11]. We use 20 000 samples. In this experiment, the estimated channel length equals $\hat{M}+1=2$ in this range of SNR. This "practical" channel length provided the best results for both blind methods even if the subspace method cannot be statistically consistent in this case. For reference, we provide the SER achieved by the optimal MMSE equalizer given by $R_0^{-1}E(X_n s_{n-d}^*)$. Note that in all experiments/methods, we choose to pick the result provided by the best equalization delay. This delay was found to be the same regardless of the equalization technique. The MRE criterion performs well under severe SNR conditions due to the MSE-like form of the cost function. Perfect equalization is achieved in the absence of noise. The linear constraint provides the best results. For high SNR, the MRE and subspace methods are found to yield similar performances.

Fig. 6 investigates the behavior of the batch MRE (with linear constraint) and subspace methods in the case the algorithms are implemented with various overestimated channel lengths (i.e., greater than the practical length). For $\hat{M}=4$ (i.e., the theoretical length), both blind techniques yield degraded performances compared with the practical length case. This is not surprising since small channel coefficients are difficult to distinguish or estimate in a noisy context. The MRE method seems to show more robustness toward this problem than the subspace method, which is known to require well-defined noise/signal subspace dimensions. The MRE criterion even shows robustness toward a large overestimation of the channel length ($\hat{M}=9$, $\hat{M}=14$), whereas the subspace can no longer be used in this case. The observed robustness can be intuitively explained: 1) The method solves for all possible delays and is not directly dependent on "small channel ends." 2) Overestimating the ISI length amounts to trying to solve an overdetermined set of conditions in (7) since the set of delays for channel inverses is not sufficient for these conditions to be satisfied. Of course, some of the MRE (typically the "extremal" MRE) are unable to invert the channel. Due to the particular form of the criterion, each of the MRE's is predominantly determined by its neighbors, hence, extremal MRE are only seen as side effects by the middle MRE. The larger $\hat{M}$ gets (the larger the number of equalizers is), the more side effects vanish at least seen from middle MRE. This explains why some of the filters become unaffected by the overdetermination and may still be used to equalize the channel.

We now focus on the adaptive equalization problem. The channel is unchanged. The practical channel length is assumed. Convergence is shown in terms of the mean-square error between the equalized and the transmitted data. Note that MSE is used here for the sake of simplicity: Unlike the SER, the MSE has closed-form expressions and does not need to be estimated. Since blind equalization is achieved up to an arbitrary gain ρ, we define for a d-delay equalizer v

$$\mathrm{MSE}\,(v)=E\left|\frac{v^{+}X_n}{\rho}-s_{n-d}\right|^2.$$

Fig. 7 shows the C1-MRE (quadratic constraint) and C2-MRE (linear constraint) methods implemented with the LMS algorithm described in Section IV-C2. The stepsize is 0.01. The SNR is 10 dB. Slow convergence can be noticed, due to the large number of parameters to be updated ($LNK=60$) and to ill-conditioning effects. A better convergence rate is, however, obtained with the linear constraint.

Fig. 8 shows the learning curves obtained with the RLS implementation of the proposed algorithm based on the prediction scheme described in Section V-A. Soft-start initialization is used with $\mathcal{S}_0=100\ I$, and $\gamma=0.99$. The time shift

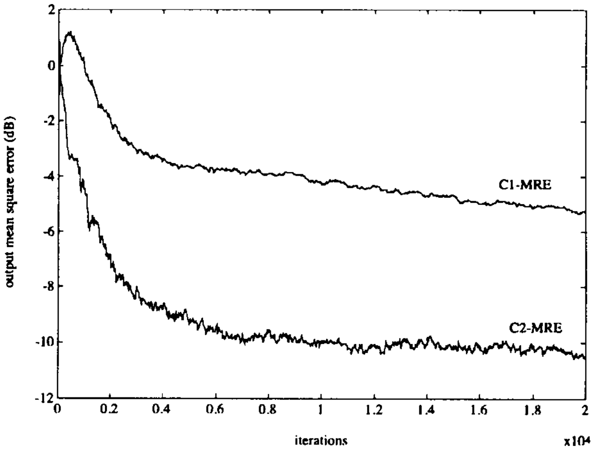

Fig. 7. Learning curves of the LMS-MRE method (10 dB SNR).

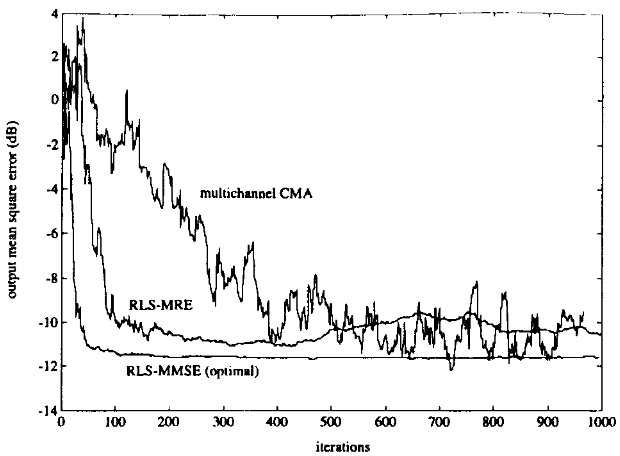

Fig. 8. Learning curves of the RLS-MRE method (10 dB SNR).

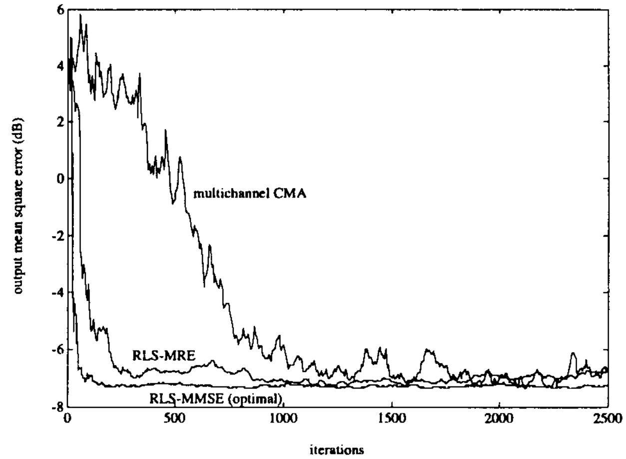

Fig. 9. Learning curves of RLS-MRE method (5 dB SNR).

TABLE II
"EQx" EQUALIZATION TEST PATH PROFILE

path	1	2	3	4	5	6
delay (μs)	0.0	2.2	4.4	6.6	8.8	11.0
attenuation (dB)	0	0	0	0	0	0

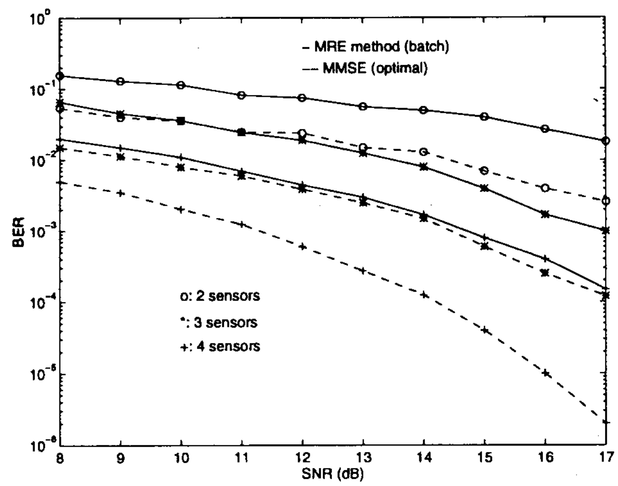

Fig. 10. Results versus the number of sensors and SNR.

parameter is $D = 10$. The SNR is 10 dB. For a comparison, the learning curves obtained with a HOS-based technique (namely, the multichannel CMA [26], [28]) and with the optimal RLS-based MMSE equalizer are also plotted. Both blind algorithms are tuned to offer comparable residual error. Fig. 9 shows a similar experiment under more severe conditions (SNR $= 5$ dB). A high convergence rate is achieved with the RLS (≈ 100 iterations at 10 dB SNR), which is close to the convergence rate of the nonblind RLS algorithm. The CMA is faster than the LMS-based MRE algorithm; however, it is outperformed by the RLS–MRE because it is essentially gradient-based and subject to ill-conditioning effects. In contrast, the convergence rate of the RLS–MRE algorithm is only bounded by the convergence rate of second-order moments.

Finally, we study the impact of channel diversity on the performances of the proposed method. Only batch implementation in the large sample limit (20 000 samples) is considered here. The path profile is the Equalization Test ("EQx") model for highly dispersive channels, which is shown in Table II. Monte Carlo simulations are conducted on the basis of 100 independent channel/signals trials.

Fig. 10 provides the results obtained with two, three, or four sensors, spaced by a half-wavelength, for a range of SNR. The results of the MRE method are plotted along with those of the optimal MMSE equalizer. Again, the best delay is picked in both methods. As we would expect, two sensors does not provide a sufficient amount of channel diversity for highly dispersive channels, unless we increase the equalizer's length. Increasing the number of sensors consistently provides better results, mainly due to better noise averaging. An extra sensor is approximately required for the performances of the blind MRE method to catch up with those of the optimal MMSE equalizer.

VII. CONCLUSION

This paper has presented a potential solution to the problem of blind adaptive equalization of multiple FIR filters, which has applicability in the context of wireless communications equipped with channel diversity. The proposed method improves on the conventional blind adaptive equalization techniques mainly in that the cost function used for equalization is a simple (constrained or not) mean-square error. This ensures global convergence and, more importantly, provides flexibility for a practical implementation since many known adaptive filtering techniques, including a standard RLS, can be used to implement the method. The initial computational load is high; however, due to the particular form of the criterion, standard fast adaptive filtering techniques are applicable here to reduce complexity. Furthermore, the complexity/performance tradeoff can be optimized by resorting to affine projection algorithms. The algorithms were tested and validated in a wireless communications context with various amounts of channel diversity. The proposed criterion seems to show some robustness w.r.t. the crucial problem of channels having not well defined or even unknown lengths. A utilization of the proposed criterion in the context of blind multiuser deconvolution seems also possible and is currently under investigation [31].

Appendix A
Proof of Lemma 4.1

The proof is easily conducted using the Kronecker product, as this formalism allows the conversion of matrix equations into vector equations. First, the error vector E_n in the criterion is rewritten in terms of the vector $\mathcal{V}$

$$\begin{aligned} E_n &= (I_{K-1}, \mathbf{o})V^+X_n - (\mathbf{o}, I_{K-1})V^+X_{n+1}, \\ E_n^+ &= X_n^+V(I_{K-1}, \mathbf{o})^+ - X_{n+1}^+V(\mathbf{o}, I_{K-1})^+, \\ E_n^* &= [(I_{K-1}, \mathbf{o}) \otimes X_n^+ - (\mathbf{o}, I_{K-1}) \otimes X_{n+1}^+]\mathcal{V}. \end{aligned} \quad (16)$$

Denote by U_n the large expression in parentheses on the right-hand side of (16). Then, we have

$$\begin{aligned} |E_n|^2 &= \mathcal{V}^+(U_n^+U_n)\mathcal{V}, \\ \mathcal{J}(\mathcal{V}) &= \mathcal{V}^+E(U_n^+U_n)\mathcal{V}. \end{aligned}$$

Straightforward manipulations lead to this expression for $\mathcal{R} \stackrel{\text{def}}{=} E(U_n^+U_n)$

$$\mathcal{R} = \text{diag}(1, 2, \cdots, 2, 1) \otimes R_0 - \begin{pmatrix} \mathbf{o} & \mathbf{o} \\ I_{K-1} & \mathbf{o} \end{pmatrix} \otimes R_1 - \begin{pmatrix} \mathbf{o} & I_{K-1} \\ \mathbf{o} & \mathbf{o} \end{pmatrix} \otimes R_1^+$$

where diag (q) stands for the diagonal matrix with diagonal q. This completes the proof. □

Appendix B
Proof of Lemma 5.1

Assuming white additive noise and the m-independence of the transmitted symbols, data vectors X_{i+1} and X_{i+D} are decorrelated for all i as soon as $D \geq K+m$ since they convey no mutual information, up to the second-order (recall (3) for this result). Consequently, the computation of $E(Y_nY_n^+)$, with $Y_n = (X_n^t, -X_{n+1}^t + X_{n+D}^t, \cdots, -X_{n+(K-3)D+1}^t + X_{n+(K-2)D}^t, -X_{n+(K-2)D+1}^t)^t$, only involves correlation terms of the form $E(X_iX_i^+)$, $E(X_iX_{i+1}^+)$, and $E(X_{i+1}X_i^+)$. Thus, one can easily check that $E(Y_nY_n^+) = \mathcal{R}$. □

Acknowledgment

Thanks are due to F.-X. Bergot (Telecom Paris) for his help with the simulations.

References

[1] A. Benveniste, M. Goursat, and G. Ruget, “Robust identification of a nonminimum phase system: Blind adjustment of a linear equalizer in data communications,” *IEEE Trans. Automat. Contr*, vol. 25, pp. 679–682, June 1980.

[2] Y. Sato, “A method of self-recovering equalization for multi-level amplitude modulation,” *IEEE Trans Commun.*, vol. COMM-6, pp. 679–682, June 1975.

[3] D. N. Godard, “Self-recovering equalization and carrier tracking in two-dimensional data communications systems,” *IEEE Trans Commun.*, vol. COMM-28, pp. 1867–1875, 1980.

[4] 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. 349–372, Apr. 1983.

[5] O. Macchi and E. Eweda, “Convergence analysis of self-adaptive equalizers,” *IEEE Trans. Inform. Theory*, vol. IT-30, pp. 161–176, Mar. 1984.

[6] 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.

[7] Z. Ding, R. A. Kennedy, B. D. O. Anderson, and C. R. Johnson, “Ill-convergence of Godard blind equalizers in data communication systems,” *IEEE Trans. Commun.*, vol. 39, pp. 1313–1329, Sept. 1991.

[8] W. A. Gardner, “A new method of channel identification,” *IEEE Trans. Commun.*, vol. 39, pp. 813–817, June 1991.

[9] 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.

[10] L. Tong, G. Xu, B. Hassibi, and T. Kailath, “Blind channel identification based on second-order statistics: A frequency-domain approach,” *IEEE Trans. Inform. Theory*, vol. 41, pp. 329–334, Jan. 1995.

[11] E. Moulines, P. Duhamel, J. F. Cardoso, and S. Mayrargue, “Subspace methods for the blind identification of multichannel FIR filters,” *IEEE Trans. Signal Processing*, vol. 43, pp. 516–526, Feb. 1995.

[12] K. Abed Meraim, P. Loubaton, and E. Moulines, “A subspace algorithm for certain blind identification problems,” to be published.

[13] Y. Li and Z. Ding, “Blind channel identification based on second-order cyclostationary statistics,” in *Proc. ICASSP*, 1993, vol. 4, pp. 81–84.

[14] H. Liu, G. Xu, and L. Tong, “A deterministic approach to blind identification of multichannel FIR systems,” in *Proc. ICASSP*, 1994, vol. 4, pp. 581–584.

[15] H. Liu and G. Xu, “A deterministic approach to blind symbol estimation,” *IEEE Signal Processing Lett.*, vol. 1, Dec. 1994.

[16] D. Slock, “Blind fractionally-spaced equalization, perfect reconstruction filter-banks and multichannel linear prediction,” in *Proc. ICASSP*, 1994, vol. 4, pp. 585–588.

[17] D. Slock and C. B. Papadias, “Further results on blind identification and equalization of multiple FIR channels,” in *Proc. ICASSP*, 1995, pp. 1964–1967.

[18] K. Abed Meraim *et al.*, “Prediction error methods for time-domain blind identification of multichannel FIR filters,” in *Proc. ICASSP*, 1995.

[19] D. Gesbert and P. Duhamel, “Robust blind channel identification and equalization based on multistep predictors,” in *Proc. ICASSP*, Apr. 1997.

[20] H. Chen, T. Sarkar, S. A. Dianat, and J. Brule, “Adaptive spectral estimation by the conjugate gradient method,” *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-34, Apr. 1986.

[21] G. Giannakis and S. Halford “Blind fractionally-spaced equalization of noisy FIR channels: Adaptive and optimal solutions,” in *Proc. ICASSP*, 1995, pp. 1972–1975.

[22] G. Giannakis “Blind equalization of time-varying channels: A deterministic multichannel approach,” in *Proc. 8th IEEE Signal Processing Workshop SSAP*, 1996.

[23] D. Gesbert, P. Duhamel, and S. Mayrargue, “Subspace-based adaptive algorithms for the blind equalization of multichannel FIR filters,” in *Proc. EUSIPCO*, 1994.

[24] A. J. Van der Veen, S. Talwar, and A. Paulraj, “Blind estimation of multiple digital signals transmitted over FIR channels,” *IEEE Signal Processing Lett.*, vol. 2, pp. 99–102, 1995.

[25] D. T. M. Slock, L. Chisci, H. Lev-Ari, and T. Kailath, “Modular and numerically stable fast transversal filters for multichannel and multiexperiments RLS,” *IEEE Trans. Signal Processing*, vol. 40, pp. 784–802, Apr. 1992.

[26] S. Mayrargue, "A blind spatio-temporal equalizer for a radio-mobile channel using the constant modulus algorithm (CMA)," in *Proc. ICASSP,* 1994, vol. 3, pp. 317–319.
[27] I. Fijalkow, F. Lopez de Victoria, and C. R. Johnson, Jr., "Adaptive fractionally-spaced blind equalization," in *Proc. IEEE DSP Workshop,* 1994.
[28] Y. Li and Z. Ding, "Global convergence of fractionally-spaced Godard (CMA) adaptive equalizers," *IEEE Trans. Signal Processing,* vol. 44, pp. 818–826, Apr. 1996.
[29] P. Comon and G. H. Golub, "Tracking a few extreme singular values and vectors in signal processing," *Proc. IEEE,* vol. 78, Aug. 1990.
[30] G. H. Golub and C. F. Van Loan, *Matrix Computations.* Baltimore, MD: Johns Hopkins Univ. Press, 1983.
[31] A. Mansour, C. Jutten, and P. Loubaton, "Subspace method for blind separation of sources and for a convolutive mixture model," in *Proc. 8th Euro. Signal Processing Conf. (EUSIPCO),* 1996.
[32] R. H. Clarke, "A statistical theory of mobile radio reception," *Bell Syst. Tech. J.,* vol. 47, pp. 987–1000, 1968.
[33] COST recommendations 05.05 for propagation models in the GSM system.

David Gesbert was born in Mantes-La-Jolie, France, in 1969. He received the M.S. degree in electrical engineering from the National Institute for Telecommunications (INT), Evry, France, in 1993 and received the Ph.D. degree from Ecole Nationale Supérieure des Télécommunications (Télécom Paris), Paris, France, in 1997.

From 1993 to 1997, he was with France Telecom-CNET, where he was involved in the development and study of blind algorithms for digital radio communications systems. Since April 1997, he has been a Visiting Scholar at the Smart Antenna Research Group, Information Systems Laboratory, Stanford University, Stanford, CA, where he is the recipient of a French Army's DGA/DRET postdoctoral scholarship. His main interests are in the areas of signal processing for digital communications, with emphasis on blind system identification and equalization, smart antennas, and adaptive filtering.

Pierre Duhamel (SM'87), for photograph and biography, see this issue, page 2202.

Sylvie Mayrargue entered the Ecole Normale Superieure de Jeunes Filles in 1973 and received the Master's degree in mathematics in 1975 and the DEA degree in the agregation of mathematics in 1976. She received the doctoral degree from the Engineering School of Télécom Paris, Paris, France, in 1978.

She joined the National Center in Telecommunications (CNET) French Telecom Research Center, Issy les Moulineaux, France, in 1978. Her first interests were in radio wave propagation. In 1985, she joined a signal processing group, where her research activities were first concerned with spectrum estimation. Her current interests are in digital radio communications, adaptive filtering, equalization, and adaptive antennas.