A Maximum Likelihood Approach to Blind Multiuser Interference Cancellation

Mónica F. Bugallo, *Student Member, IEEE*, Joaquín Míguez, *Member, IEEE*, and Luis Castedo, *Associate Member, IEEE*

***Abstract*—This paper addresses the problem of blind multiple access interference (MAI) and inter-symbol interference (ISI) suppression in direct sequence code division multiple access (DS CDMA) systems. A novel approach to obtain the coefficients of a linear receiver using the maximum likelihood (ML) principle is proposed. The method is blind because it only exploits the statistical features of the transmitted symbols and Gaussian noise in the channel. We demonstrate that an adequate linear constraint on these coefficients ensures that the desired user is extracted and the resulting linearly constrained maximum likelihood linear (LCMLL) receiver can be efficiently implemented using the iterative space alternating generalized expectation–maximization (SAGE) algorithm. In order to take advantage of the diversity inherent to multipath channels, we also introduce a blind rake multiuser receiver that proceeds in two steps. First, soft estimates of the desired user transmitted symbols are obtained from each propagation path using a bank of appropiate LCMLL receivers. Afterwards, these estimates are adequately combined to enhance the signal-to-interference-and-noise ratio (SINR). Computer simulations show that the proposed blind algorithms for multiuser detection are near–far resistant and attain convergence using small blocks of data, thus outperforming existing linearly constrained minimum variance (LCMV) blind receivers.**

***Index Terms*—Blind receivers, CDMA, interference suppression, maximum likelihood, multiuser detection, rake receiver.**

I. Introduction

CODE division multiple access (CDMA) is the multiple access technique to be used in the next generation of mobile communication systems because it provides a higher spectral efficiency and a superior flexibility in the radio interface [1]–[3]. In CDMA, different users simultaneously transmit over the same bandwidth, and each user-signal modulates an unique spreading code or signature waveform. The capacity of current practical CDMA systems, however, is limited by the multiple access interference (MAI) caused by code nonorthogonality due to diverse phenomena such as asynchronous transmission, multipath propagation, or limited bandwidth. Moreover, the presence of inter-symbol interference (ISI) due to the time-dispersive nature of wireless channels is often neglected in low rate CDMA systems, but it becomes a major problem in wideband CDMA.

Different techniques have been proposed to suppress MAI as well as ISI using linear filtering. Decorrelating receivers [4] require a perfect knowledge of the received user codes, which are likely to be distorted by the unknown channel, and they suffer from noise amplification problems. Conventional linear minimum mean square error (MMSE) receivers [4], [5] overcome both drawbacks through the use of training sequences, but such sequences are not available in many applications. Therefore, alternative blind implementations are preferred [6]–[8]. Several blind schemes based on the linearly constrained minimum variance (LCMV) criterion have been proposed. The LCMV receivers described in [6] and [9] require a very precise knowledge of the desired user code and timing that is not likely to be available in practice. This limitation is overcome with the solution proposed in [7], which only requires the transmitted (i.e., nondistorted) spreading code to be known. Nevertheless, all LCMV multiuser receivers exhibit a very low convergence rate, especially at moderate and high signal-to-noise ratio (SNR) values [10], that restricts their practical applicability. Subspace techniques with somehow faster convergence rate have also been suggested [11]–[13], but their high computational complexity and their poor performance in the low SNR region are important disadvantages in real applications.

In this paper, we introduce a new blind approach to linear multiuser interference cancellation that exploits the statistical features of the desired user signal taking into account the additive white Gaussian noise (AWGN) in the channel. The maximum likelihood (ML) principle is used to estimate the coefficients of the linear multiuser receiver that supresses both MAI and ISI in time-dispersive multipath channels. Since the proposed ML linear (MLL) receiver exploits the statistical characterization of the received information-bearing signals, and this is the same both for the desired user and the interfering ones, the receiver may capture an interference instead of the user of interest. We show, however, that a linear constraint on the receiver coefficients is enough to guarantee that the resulting detector extracts the desired user symbols. Since a closed-form solution for the proposed linearly constrained (LC) MLL receiver does not exist, we also suggest an efficient iterative implementation based on the expectation-maximization (EM) algorithm [14]–[17] that provides very fast convergence.

The LCMLL multiuser receiver presents, however, an important disadvantage because it is unable to exploit the diversity inherent to multipath channels. The linear constraint avoids the capture problem by ensuring that the desired user signal arriving through one particular propagation path is never cancelled. With

Manuscript received February 16, 2000; revised February 20, 2001. This work was supported by FEDER funds under Grant 1FD97-0082 and the Xunta de Galicia under Grant PGIDT00PXI10504PR. The associate editor coordinating the review of this paper and approving it for publication was Dr. Athina Petropulu.

The authors are with Departamento de Electrónica e Sistemas, Universidade da Coruña, Facultade de Informática, Coruña, Spain (e-mail: monica@des.fi.udc.es; miguez@des.fi.udc.es; luis@des.fi.udc.es).

Publisher Item Identifier S 1053-587X(01)03878-8.

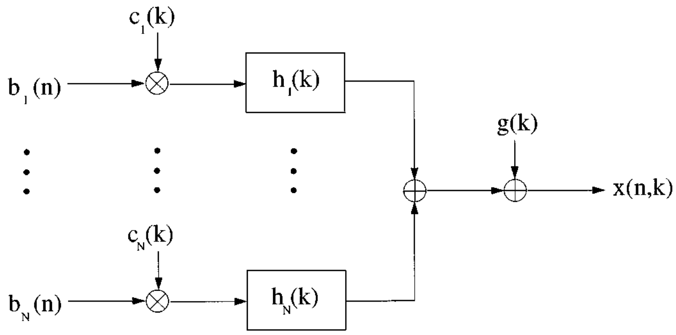

Fig. 1. Baseband discrete-time equivalent model of a DS CDMA system with time dispersive channels.

this approach, the other desired user components due to alternative paths are dealt with as interferences, and they are suppressed instead of recombined to enhance the signal to interference and noise ratio (SINR). Thus, the resulting LC MLL receiver exhibits a clearly suboptimum performance. To overcome this limitation, we introduce a blind rake multiuser receiver [4], [18] that proceeds in two steps. First, soft estimates of the desired user-transmitted symbols are obtained for each propagation path using a bank of appropiate LCMLL receivers. Second, these soft estimates are suitably recombined to enhance the SINR. The weight vector for this recombination is also estimated according to the ML criterion.

The remainder of this paper is organized as follows. The next section introduces the baseband discrete-time equivalent signal model of an asynchronous CDMA communication system with time-dispersive channels. In Section III, we introduce the LCMLL multiuser receiver. Section IV describes the iterative EM-based algorithm used to compute the filter coefficients. In Section V, the implementation of the blind rake receiver based on the ML principle is addressed. Finally, Section VI presents some illustrative computer simulation results, and Section VII is devoted to the conclusions.

II. Signal Model

Let us consider a baseband direct-sequence (DS) CDMA system with N users and time dispersive channels whose discrete-time equivalent model is shown in Fig. 1. When the ith user transmits a sequence of statistically independent complex symbols $b_i(n)$, it modulates a unique spreading code waveform $c_i(t)$. Each channel use consists of the transmission of a sequence of K symbols, and thus, the signal transmitted by the ith user is given by

$$z_i(t) = \sum_{n=0}^{K-1} b_i(n) c_i(t - nT - \tau_i) \quad (1)$$

where T is the symbol period, which is assumed to be equal to the code waveform duration, and $0 \leq \tau_i < T$ is the ith user unknown delay. The overall received signal for the ith user is

$$x_i(t) = \sum_{n=0}^{K-1} z_i(t) * h_i(t) \quad (2)$$

where $*$ denotes convolution, and $h_i(t)$ is the continuous-time channel response between the i-th transmitter and the multiuser demodulator.

The ith spreading code $c_i(t)$ can be decomposed into a sequence of L binary *chips* that modulate a pulse waveform $\alpha(t)$ of duration T_c, i.e.,

$$c_i(t) = \sum_{j=0}^{L-1} c_i(j) \alpha(t - jT_c) \quad (3)$$

where T_c is the chip period ($L = T/T_c$). Therefore, we can substitute (1) and (3) into (2) to yield

$$\begin{aligned} x_i(t) &= \sum_{n=0}^{K-1} b_i(n) \sum_{j=0}^{L-1} c_i(j) \alpha(t - nT - jT_c - \tau_i) * h_i(t) \\ &= \sum_{n=0}^{K-1} b_i(n) \sum_{j=0}^{L-1} c_i(j) h_i^{\alpha}(t - nT - jT_c) \end{aligned} \quad (4)$$

where $h_i^{\alpha}(t) = \alpha(t - \tau_i) * h_i(t)$ is the equivalent channel response obtained when the pulse $\alpha(t - \tau_i)$ is transmitted through the channel $h_i(t)$. Note that $h_i^{\alpha}(t)$ accounts not only for the continuous-time channel response but for the relative time delays of the different users (this "equivalent channel" approach is rather common; see, for instance, [7] and [11]) as well. The resulting signal is passed through a chip-matched filter followed by a chip rate sampler. The obtained output for the ith user, in the jth chip period, during the nth symbol period is

$$\begin{aligned} x_i(n,j) &= \int_{nT+jT_c}^{nT+(j+1)T_c} x_i(t) \alpha(t - nT - jT_c)\, dt \\ &= \sum_{k=0}^{K-1} b_i(k) \sum_{l=0}^{L-1} c_i(l) \int_0^{T_c} h_i^{\alpha}(t + (n-k)T \\ &\quad + (j-l)T_c) \alpha(t)\, dt. \end{aligned} \quad (5)$$

If the equivalent channel $h_i^{\alpha}(t)$ is $(P-1)T_c$ long, i.e., it is zero outside of the interval $[0, (P-1)T_c]$, the nth transmitted symbol $b_i(n)$ interferes with $b_i(n-1), \ldots, b_i(n-m+1)$, where $m = \lceil (L+P-1)/L \rceil$ is the channel memory size, and (5) can be simplified as

$$\begin{aligned} x_i(n,j) &= \sum_{k=n-m+1}^{n} b_i(k) \sum_{l=0}^{L-1} c_i(l) \int_0^{T_c} h_i^{\alpha}(t + (n-k)T \\ &\quad + (j-l)T_c) \alpha(t)\, dt \\ &= \sum_{r=0}^{m-1} b_i(n-r) \sum_{l=0}^{L-1} c_i(l) \\ &\quad \times \int_0^{T_c} h_i^{\alpha}(t + rT + (j-l)T_c) \alpha(t)\, dt \\ &= \sum_{r=0}^{m-1} b_i(n-r) \sum_{l=0}^{L-1} c_i(l) h_i(rL + j - l) \\ &= \sum_{r=0}^{m-1} b_i(n-r) d_i(rL + j) \end{aligned} \quad (6)$$

where $h_i(p) = \int_0^{T_c} h_i^{\alpha}(t + pT_c) \alpha(t)\, dt, p = 0, \ldots, P-1$ is the discrete-time equivalent channel response, and the sequence

$d_i(k) = c_i(k) * h_i(k) = \sum_{p=0}^{P-1} h_i(p)c_i(k-p)$ has length $L+P-1$ and will be termed *received code*. Using (6), we can write the overall received jth sample during the nth symbol period

$$x(n,j) = x(nL+j) = \sum_{i=1}^{N}\sum_{r=0}^{m-1} d_i(r,j)b_i(n-r) + g(n,j) \quad j=0,\ldots,L-1 \quad (7)$$

where $d_i(r,j) = d_i(rL+j)$, and $g(n,j) = g(nL+j)$ is the jth component of the AWGN sequence.[1]

Using vector notation, the $L \times 1$ vector given by the observations in (7) can be written as

$$\mathbf{x}(n) = \sum_{i=1}^{N} \mathbf{D}_i \mathbf{b}_i(n) + \mathbf{g}(n) \quad (8)$$

where $\mathbf{D}_i = [\mathbf{d}_i(0),\ldots,\mathbf{d}_i(m-1)]$ is the $L \times m$ received code matrix for the ith user, which is composed of the column vectors $\mathbf{d}_i(r) = [d_i(r,0),\ldots,d_i(r,L-1)]^T$; $\mathbf{b}_i(n) = [b_i(n),\ldots,b_i(n-m+1)]^T$ is the $m \times 1$ vector of symbols contributed by the ith user to the nth observation vector, and $\mathbf{g}(n) = [g(n,0),\ldots,g(n,L-1)]^T$ is a vector of independent and identically distributed (i.i.d.) complex Gaussian variables with zero mean and covariance matrix $E[\mathbf{g}(n)\mathbf{g}^H(n)] = \sigma_g^2 \mathbf{I}_L$.

The linear multiuser receiver consists of a finite impulse response (FIR) filter $\mathbf{w} = [w_1,\ldots,w_L]^T$ followed by a threshold detector as shown in Fig. 2. The soft estimate corresponding to the nth symbol period can be written as

$$y(n) = \mathbf{w}^H \mathbf{x}(n) \quad (9)$$

where the superindex ^{H} denotes Hermitian transposition.

III. Selection of the Receiver Coefficients

In this section, we derive a novel statistical approach to select the receiver coefficients in order to obtain MAI and ISI free estimates of the desired user symbols. The selection criterion is based on the fact that, when the MAI and the ISI are totally suppressed, the symbol soft estimate y consists of just two components: the desired user symbol b_1 and an additive Gaussian noise term g_f. Indeed, let $\mathbf{w}_*$ denote the optimum value of the filter coefficients that eliminate the MAI and the ISI. Then, we can write

$$y = \mathbf{w}_*^H \mathbf{x} = A_1 b_1 + g_f \quad (10)$$

where

- b_1 desired user symbol;
- A_1 unknown complex amplitude that depends on both the channel vector $\mathbf{h}_1$ and $\mathbf{w}_*$;
- g_f complex Gaussian random variable with zero mean and variance $\sigma_f^2 = \sigma_g^2 \mathbf{w}_*^H \mathbf{w}_*$.

Although the filtered noise variance σ_f^2 clearly depends on $\mathbf{w}_*$, we will assume in the sequel that it is *a priori* known, and therefore, it is dealt with as a constant.[2] In Appendix A, it is demonstrated that the probability density function (pdf) of $y = A_1 b + g_f$ is given by

$$f_{y;\mathbf{w}_*,A_1}(y) = f_{A_1 b + g_f}(y) = \frac{1}{\pi\sigma_f^2} E_b\left[e^{-\frac{|y-A_1 b|^2}{\sigma_f^2}}\right] \quad (11)$$

where $E_b[\cdot]$ denotes statistical expectation with respect to (w.r.t.) the desired user-transmitted symbols.

In digital communications the transmitted symbols are usually modeled as discrete i.i.d. random variables with known pdf and finite alphabet. Therefore, the statistical expectation in (11) reduces to a simple summation. Moreover, the soft estimates obtained with the optimum filter $\mathbf{w}_*$ can also be considered as i.i.d. random variables, and when a block of K observation vectors is available, the joint pdf of the resulting frame of estimates $\mathbf{y} = [y(0),\ldots,y(K-1)]^T$ is

$$f_{\mathbf{y};\mathbf{w}_*,A_1}(\mathbf{y}) = \prod_{n=0}^{K-1} f_{y;\mathbf{w}_*,A_1}(y(n)) = \left(\frac{1}{\pi\sigma_f^2}\right)^K \prod_{n=0}^{K-1} E_b\left[e^{-\frac{|y(n)-A_1 b|^2}{\sigma_f^2}}\right]. \quad (12)$$

Note that the pdf of $\mathbf{y}$ given by (12) depends on the unknown parameters $\mathbf{w}_*$ and A_1, which are given by

$$[\hat{\mathbf{w}}, \hat{A}_1] = \arg\max_{\mathbf{w}_*, A_1} \{\mathcal{L}(\mathbf{w}_*, A_1)\} \quad (13)$$

where

$$\mathcal{L}(\mathbf{w}_*, A_1) = \log \prod_{n=0}^{K-1} E_b\left[e^{-\frac{|y(n)-A_1 b|^2}{\sigma_f^2}}\right] = \sum_{n=0}^{K-1} \log E_b\left[e^{-\frac{|y(n)-A_1 b|^2}{\sigma_f^2}}\right] \quad (14)$$

is the log-likelihood of $[\mathbf{w}_*, A_1]$ w.r.t. the block of soft estimates $\mathbf{y} = [y(0),\ldots,y(K-1)]$.

Unfortunately, the log-likelihood $\mathcal{L}(\mathbf{w}_*, A_1)$ is a nonquadratic function that presents several local maxima. In particular, the solutions to problem (13) guarantee that the soft estimates $y(n)$ have a pdf close to $f_{A_1 b + g_f}(\cdot)$, but this is not enough to ensure that the desired user is extracted. Since

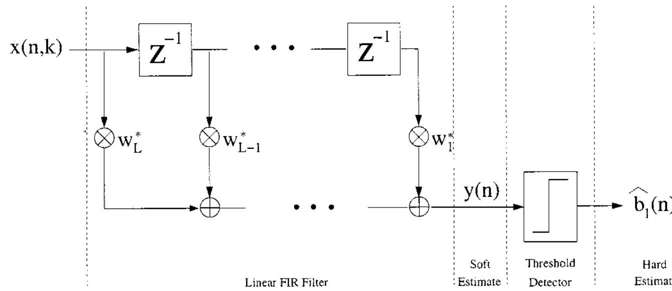

Fig. 2. Linear multiuser receiver.

[1]The Gaussian noise sequence $g(n,j)$ is white if the chip waveform $\alpha(t)$ is chosen according to the zero ISI criterion [19].

[2]Nevertheless, the computer simulation results in Section VI show that this is not an important parameter, and large deviations in the selection of σ_f^2 do not lead to a significant performance degradation.

in CDMA all users transmit symbols with the same modulation format, the pdf of the ith interference at the receiver is $f_{A_i b+g_f}(\cdot)$, which only differs from the target pdf $f_{A_1 b+g_f}(\cdot)$ in the unknown complex amplitude $A_1 \neq A_i$. Therefore, solving the optimization problem (13) may lead to the *capture* of an interference. In order to avoid this limitation, we propose to set an adequate linear constraint on the coefficient vector $\mathbf{w}$ that prevents the capture of a nondesired user.

Let us consider the factorization of the received code

$$\mathbf{d}_i(r) = \begin{bmatrix} d_i(r,0) \\ \vdots \\ d_i(r,L-1) \end{bmatrix} = \mathbf{C}_i(r)\mathbf{h}_i \tag{15}$$

where $\mathbf{h}_i = [h_i(0),\ldots,h_i(P-1)]^T$ is the $P \times 1$ vector containing the channel components for the ith user, and

$$\mathbf{C}_i(r) = \begin{bmatrix} c_i(rL) & \cdots & c_i(rL-P+1) \\ \cdots & \ddots & \cdots \\ c_i((r+1)L-1) & \cdots & c_i((r+1)L-P) \end{bmatrix} \tag{16}$$

is an $L \times P$ matrix whose columns are length L segments of the ith user transmitted code. Using this decomposition, the soft estimate $y(n)$ can be written as

$$y(n) = \mathbf{w}^H\mathbf{C}_1(0)\mathbf{h}_1 b_1(n) + \sum_{i=1}^{m-1}\mathbf{w}^H\mathbf{d}_1(i)b_1(n-i) + \sum_{i=2}^{N}\mathbf{w}^H\mathbf{D}_i\mathbf{b}_i + \mathbf{w}^H\mathbf{g}(n). \tag{17}$$

In order to prevent the desired signal component $\mathbf{w}^H\mathbf{d}_1(0)b_1(n)$ from being cancelled or attenuated when selecting the filter coefficients, vector $\mathbf{w}$ can be constrained to verify

$$\mathbf{w}^H\mathbf{C}_1(0)\mathbf{h}_1 = A_1. \tag{18}$$

It is apparent that $\mathbf{h}_1$ is unknown, but the above condition holds as long as $\mathbf{w}^H\mathbf{C}_1(0) = \mathbf{u}_d^T = [0\cdots 0 \underbrace{1}_{d} 0\cdots 0]$ and $|h_1(d)| > 0$ since

$$\mathbf{w}^H\mathbf{C}_1(0)\mathbf{h}_1 = \mathbf{u}_d^T\mathbf{h}_1 = h_1(d) = A_1. \tag{19}$$

Therefore, we propose to select the coefficient vector and the amplitude parameter estimates as the solution to the linearly constrained problem

$$[\hat{\mathbf{w}}, \hat{A}_1] = \arg\max_{\mathbf{w}_*, A_1}\{\mathcal{L}(\mathbf{w}_*, A_1)\} \quad \text{subject to } \mathbf{w}_*^H\mathbf{C}_1(0) = \mathbf{u}_d^T. \tag{20}$$

Notice that the constraint in (20) is always feasible if $P < L$, and it guarantees that an interference is not captured as long as $|h_1(d)|$ is non-negligible, as shown in Appendix B. The multiuser receiver built using vector $\hat{\mathbf{w}}$ is the LCMLL detector for user 1.

It is important to remark that, rigorously speaking, criterion (20) is inherently unrealizable because it relies on the hypothesis that the soft estimates $y(n)$ have the *desired* pdf (11). It is apparent that this assumption does not hold in practice because both $\mathbf{w}_*$ and A_1 are unknown. Nevertheless, the computer simulations in Section VI illustrate that the criterion is still valid. The explanation of this is twofold. On the one hand, criterion (20) is equivalent to a partial minimization of the Kullback–Leibler distance (KLD) between the actual pdf of $y(n)$, $f_y(\cdot)$ and the target pdf $f_{y;\mathbf{w}_*,A_1}(\cdot)$ [20]. Indeed, the KLD between both pdf can be written as

$$\begin{aligned}\mathrm{KLD}(f_y \parallel f_{y;\mathbf{w}_*,A_1}) &= E_y\left[\log\frac{f_y}{f_{y;\mathbf{w}_*,A_1}}\right] \\ &= E_y[\log f_y] - E_y[\log f_{y;\mathbf{w}_*,A_1}]\end{aligned} \tag{21}$$

and the second term in (21) can be estimated from $y(0),\ldots,y(K-1)$ as

$$-\frac{1}{K}\sum_{n=0}^{K-1}\log f_{y;\mathbf{w}_*,A_1}(y(n)) \tag{22}$$

which is, except for a scale factor, the negative of the log-likelihood $\mathcal{L}(\mathbf{w}_*, A_1)$ in (13). On the other hand, the analysis presented in Appendix C shows that the LCMLL multiuser receiver $\hat{\mathbf{w}}$, which is obtained as the solution to problem (20), is closely related to the linear MMSE detector subject to the same linear constraint. Analytical results concerning the large sample (asymptotic) properties of $\hat{\mathbf{w}}$ would also be desirable, but they exceed the scope of the present paper and remain for future work.

IV. Iterative Implementation

Unfortunately, it is not possible to find a closed-form solution to problem (20), and therefore, some optimization algorithm must be used to obtain the parameter estimates $[\hat{\mathbf{w}}, \hat{A}_1]$. In order to find an iterative rule that adequately computes $\hat{\mathbf{w}}$ and $\hat{A}_1$, we will first convert problem (20) into an unconstrained form. This can be done using the generalized sidelobe canceller (GSC) decomposition [21]

$$\mathbf{w}_* = \mathbf{w}_q - \mathbf{B}\mathbf{w}_u \tag{23}$$

where

- $\mathbf{w}_q$ quiescent vector;
- $\mathbf{B}$ blocking matrix;
- $\mathbf{w}_u$ unconstrained part of $\mathbf{w}_*$.

Both $\mathbf{w}_q$ and $\mathbf{B}$ are completely determined by the constraint: The quiescent vector belongs to the subspace defined by the constraint, i.e., it is a solution of the overdetermined linear system $\mathbf{w}_q^H\mathbf{C}_1(0) = \mathbf{u}_d^T$, and $\mathbf{B}$ is an $L \times (L-P)$ matrix that spans the null column subspace of $\mathbf{C}_1(0)$, i.e., $\mathbf{B}^H\mathbf{C}_1(0) = \mathbf{0}$. As a consequence, (20) is equivalent to

$$\hat{\Theta} = \arg\max_{\Theta}\left\{\mathcal{L}(\Theta) = \sum_{n=0}^{K-1}\log E_b\left[e^{-\frac{|y(n)-A_1 b|^2}{\sigma_f^2}}\right]\right\} \tag{24}$$

where $\Theta = [\mathbf{w}_u, A_1]$ and the dimension of $\mathbf{w}_u$ is $(L-P) \times 1$.

The next step is to compute Θ using the EM algorithm [14]–[16], [22] that provides an iterative procedure to perform ML estimation when direct maximization of the likelihood is not feasible. The EM approach postulates the existence of some missing (unobserved) data that, if known, would aid in the estimation problem. The algorithm consists of a two-step iteration: Use the incomplete (observed) data and the current parameter estimates to compute sufficient statistics of the complete data (E-step), and re-estimate the parameters using the computed complete data sufficient statistics (M-step). The sequence of estimates thus obtained exhibits the desirable property of being monotonically nondecreasing in likelihood.

In our problem, the incomplete-data set is given by the soft estimates $y(n), n = 0, \ldots, K-1$, whereas the complete-data set is given by the *extended* vectors $\mathbf{y}_e(n) = [y(n)\ b_1(n)]^T, n = 0, \ldots, K-1$. Let us build the complete-data block $\mathbf{Y}_e = [\mathbf{y}_e(0), \ldots, \mathbf{y}_e(K-1)]$ with joint pdf $f_{\mathbf{Y}_e;\Theta}(\cdot)$. It is easy to decompose $f_{\mathbf{Y}_e;\Theta}(\cdot)$ as

$$f_{\mathbf{Y}_e;\Theta}(\mathbf{Y}_e) = f_{\mathbf{Y}_e\,|\,\mathbf{y};\Theta}(\mathbf{Y}_e) f_{\mathbf{y};\Theta}(\mathbf{y}) \tag{25}$$

and taking logarithms and conditional expectations on both sides of (25), we arrive at the relationship

$$\log f_{\mathbf{y};\Theta}(\mathbf{y}) = U(\Theta, \hat{\Theta}_{i,i}) - V(\Theta, \hat{\Theta}_{i,i}) \tag{26}$$

where

$$U(\Theta, \hat{\Theta}_{i,i}) = E_{\mathbf{Y}_e\,|\,\mathbf{y};\hat{\Theta}_{i,i}}[\log(f_{\mathbf{Y}_e;\Theta}(\mathbf{Y}_e))] \tag{27}$$

and

$$V(\Theta, \hat{\Theta}_{i,i}) = E_{\mathbf{Y}_e\,|\,\mathbf{y};\hat{\Theta}_{i,i}}[\log(f_{\mathbf{Y}_e\,|\,\mathbf{y};\Theta}(\mathbf{Y}_e))]. \tag{28}$$

$\hat{\Theta}$ and $\hat{\Theta}_{i,j}$ denote $[\hat{\mathbf{w}}_u, \hat{A}_1]$, and $[\hat{\mathbf{w}}_u(i), \hat{A}_1(j)]$, respectively. An application of Jensen's inequality shows that [14]

$$V(\Theta, \hat{\Theta}_{i,i}) \le V(\hat{\Theta}_{i,i}, \hat{\Theta}_{i,i}) \tag{29}$$

for any value of Θ, and as a consequence, the sequence of estimates

$$\hat{\Theta}_{i+1,i+1} = \arg\max_{\Theta} U(\Theta, \hat{\Theta}_{i,i}) \tag{30}$$

is clearly nondecreasing in likelihood. Substituting

$$f_{\mathbf{Y}_e;\Theta}(\mathbf{Y}_e) = \prod_{n=0}^{K-1} \frac{1}{\pi\sigma_f^2} e^{-\frac{|y(n)-A_1 b_1(n)|^2}{\sigma_f^2}} f_b(b_1(n)) \tag{31}$$

(see Appendix A) into (30) and neglecting constant terms leads to the single iterative rule

$$\hat{\Theta}_{i+1,i+1} = \arg\min_{\Theta}\{\mathcal{Q}(\hat{\Theta}_{i,i})\} \tag{32}$$

where

$$\mathcal{Q}(\hat{\Theta}_{i,i}) = \sum_{n=0}^{K-1} E_{\mathbf{y}_e(n)\,|\,y(n);\hat{\Theta}_{i,i}}[|y(n) - \hat{A}_1(i) b_1(n)|^2] \tag{33}$$

that comprises the E and M steps of the EM algorithm.

Nevertheless, analytically solving (32) w.r.t. the joint parameter vector $[\mathbf{w}_u, A_1]$ is rather involved. The space alternating generalized EM (SAGE) algorithm [23] is a suitable modification of the conventional EM approach that consists of succesively maximizing function $U(\cdot,\cdot)$ w.r.t. different parameter subsets [15], [23]. In our case, it is straightforward to find separate updating rules for $\hat{\mathbf{w}}_u(i)$ and $\hat{A}_1(i)$

$$\hat{\mathbf{w}}_u(i+1) = \arg\min_{\mathbf{w}_u}\{\mathcal{Q}(\hat{\Theta}_{i,i})\} \tag{34}$$

$$\hat{A}_1(i+1) = \arg\min_{A_1}\{\mathcal{Q}(\hat{\Theta}_{i+1,i})\} \tag{35}$$

where we have neglected all terms that are constant w.r.t. $\mathbf{w}_u$ and A_1. The optimization problems (34) and (35) have closed-form solutions, and the sequence of estimates provided by the SAGE algorithm turns out to be

$$\hat{\mathbf{w}}_u(i+1) = \left(\mathbf{B}^H \sum_{n=0}^{K-1} \mathbf{x}(n)\mathbf{x}^H(n)\mathbf{B}\right)^{-1} \times \mathbf{B}^H \sum_{n=0}^{K-1} \mathbf{x}(n) e^*(n) \tag{36}$$

$$\hat{A}_1(i+1) = \left(\sum_{n=0}^{K-1} E_{b_1(n)\,|\,y(n);\hat{\Theta}_{i+1,i}}[|b_1(n)|^2]\right)^{-1} \times \left(\sum_{n=0}^{K-1} E_{b_1(n)\,|\,y(n);\hat{\Theta}_{i+1,i}}[b_1^*(n)] y(n)\right) \tag{37}$$

where

$$e^*(n) = \mathbf{x}^H(n)\mathbf{w}_q - \hat{A}_1^*(i) E_{b_1(n)\,|\,y(n);\hat{\Theta}_{i,i}}[b_1^*(n)] \tag{38}$$

and we have also used the fact that the only random part in $\mathbf{y}_e(n)\,|\,y(n)$ is $b_1(n)$. The conditioned expectations $E_{b_1(n)\,|\,y(n);\hat{\Theta}_{i,j}}[\cdot]$ in (36) and (37) are calculated in terms of $f_{y;\hat{\Theta}_{i,j}}(\cdot), f_{y\,|\,b;\hat{\Theta}_{i,j}}(\cdot)$ and $f_b(\cdot)$ using the Bayes theorem. Note that $f_{y\,|\,b;\hat{\Theta}_{i,j}}(\cdot)$ is obtained from $f_{y;\hat{\Theta}_{i,j}}(\cdot)$ by simply dropping the expectation. As a result, the following relationship is obtained:

$$E_{b_1(n)\,|\,y(n);\hat{\Theta}_{i,j}}[\Phi(b_1(n))] = \frac{E_{b_1(n)}\left[e^{-\frac{|y(n)-\hat{A}_1(i)b_1(n)|^2}{\sigma_f^2}} \Phi(b_1(n))\right]}{E_{b_1(n)}\left[e^{-\frac{|y(n)-\hat{A}_1(i)b_1(n)|^2}{\sigma_f^2}}\right]} \tag{39}$$

where $\Phi(\cdot)$ is an arbitrary function of $b_1(n)$.

At first glance, it may seem that the SAGE algorithm given by (36) and (37) is computationally very demanding due to the need to obtain an inverse matrix in (36). This is not the case in practice, however. Since we are only interested in updating the unconstrained vector, it is not necessary to explicitly carry out the matrix inversion. Vector $\hat{\mathbf{w}}_u(i+1)$ can be obtained by solving a system of linear equations, and it is well known that there are several fast and numerically stable methods to accomplish this task [24]–[26]. Therefore, our approach is computationally less complex than subspace techniques, which usually require carrying out an eigendecomposition of the observations autocorrelation matrix. Although currently feasible due to the advances in VLSI technology, the computation of eigenvalues and eigenvectors [25], [26] is clearly more demanding than the iteration of algorithm (36) and (37).

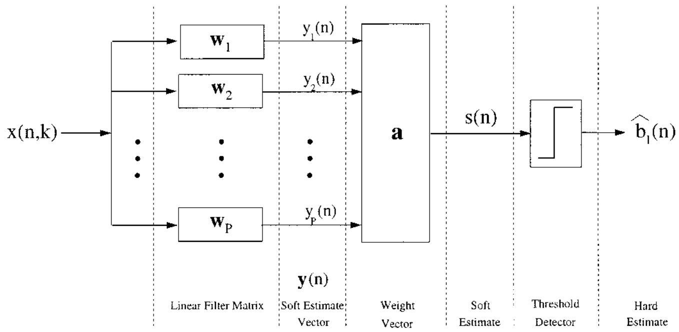

Fig. 3. Rake receiver.

V. Rake Receiver

The LCMLL receiver described so far is a valid solution to the multiuser detection problem because it provides estimates of the desired user transmitted symbols with significantly reduced MAI and ISI. Unfortunately, it is a suboptimum approach because it fails to exploit the inherent temporal diversity of multipath channels. Indeed, the linear constraint, as defined in (19), avoids the capture problem at the expense of cancelling all the desired user components except the one received through the dth propagation path. Therefore, not all the desired user received energy is fully exploited. Clearly, a more adequate choice of the linear constraint that circumvents this drawback is

$$\mathbf{w}^H\mathbf{C}_1(0) = \mathbf{u}^T \tag{40}$$

where vector $\mathbf{u}$ is selected in order to maximize the scalar magnitude

$$|\mathbf{u}^T\mathbf{h}_1| = \left|\sum_{i=1}^{P} u_i h_1(i)\right|. \tag{41}$$

However, this constraint can only be established if the channel vector $\mathbf{h}_1$ is known, which is not the case in the context of blind detection.

As an alternative approach, we propose the implementation of the blind rake detector [4], [18], which is shown in Fig. 3. It consists of a bank of P LCMLL receivers: one for each propagation path. The ith receiver provides a soft estimate $y_i(n) = \mathbf{w}_i^H\mathbf{x}(n)$ of the desired user-transmitted symbol $b_1(n)$, using the linear constraint corresponding to the ith path, i.e., $\mathbf{w}^H\mathbf{C}_1(0) = \mathbf{u}_i^T$. Afterwards, these estimates are linearly combined to yield the improved symbol estimate

$$s(n) = \sum_{i=1}^{P} a_i^* y_i(n) = \mathbf{a}^H\mathbf{y}(n) \tag{42}$$

where $\mathbf{y}(n) = [y_1(n), \ldots, y_P(n)]^T = \mathbf{W}^H\mathbf{x}(n)$ is the soft-estimate vector $\mathbf{W} = [\mathbf{w}_1, \ldots, \mathbf{w}_P]$, and $\mathbf{a} = [a_1, \ldots, a_P]^T$ is an adequately chosen weight vector. There are several criteria that may lead to a proper selection of vector $\mathbf{a}$ (e.g., MMSE and maximum SINR), but they require knowledge of either the transmitted symbols or the channel coefficients, which are not available. Notice that the ith LCMLL receiver in the bank also provides an estimate of the desired user channel gain for the corresponding path, i.e., the complex amplitude $\hat{A}_1$ is actually an ambiguous phase estimate $\hat{h}_1(i)$ of the ith channel coefficient $h_1(i)$. We can then build an estimate of the channel vector $\hat{\mathbf{h}}_1 = [\hat{h}_1(0), \ldots, \hat{h}_1(P-1)]^T$, but this is not useful at all in calculating the vector $\mathbf{a}$ because each coefficient $\hat{h}_1(i)$ has a different unknown phase rotation.

Following the same reasoning that led to the development of the LCMLL receiver, we propose to select the weight vector $\mathbf{a}$ according to the ML criterion, i.e.,

$$\hat{\mathbf{a}} = \arg\max_{\mathbf{a}} \left\{ \log \prod_{n=0}^{K-1} f_{s;\mathbf{a}_*}(s(n)) \right\} \tag{43}$$

where $f_{s;\mathbf{a}_*}(\cdot)$ is the pdf of the symbol estimate $s(n)$ when the optimum weight vector $\mathbf{a}_*$ is used. Analogously to Section III, we assume that

$$s(n) = \mathbf{a}_*^H\mathbf{y}(n) = b_1(n) + g_a(n) \tag{44}$$

where $g_a(n) = \mathbf{a}^H\mathbf{W}^H\mathbf{g}(n)$ is a Gaussian noise scalar component[3] with zero mean and variance σ_a^2, and

$$f_{s;\mathbf{a}_*}(s(n)) = \frac{1}{\pi\sigma_a^2} E_b\left[e^{-\frac{|s(n)-b|^2}{\sigma_a^2}}\right]. \tag{45}$$

Substituting (45) into (43), we arrive at the equivalent optimization problem

$$\hat{\mathbf{a}} = \arg\max_{\mathbf{a}}\{\mathcal{L}(\mathbf{a})\} \tag{46}$$

where

$$\begin{aligned}\mathcal{L}(\mathbf{a}) &= \log \prod_{n=0}^{K-1} E_b\left[e^{-\frac{|s(n)-b|^2}{\sigma_a^2}}\right] \\ &= \sum_{n=0}^{K-1} \log E_b\left[e^{-\frac{|s(n)-b|^2}{\sigma_a^2}}\right]\end{aligned} \tag{47}$$

which can be solved using the EM algorithm. Following a reasoning analogous to the one explained in Section IV and considering the incomplete data set $\{s(0), \ldots, s(K-1)\}$ and the complete data set $\{[s(0)\; b_1(0)]^T, \ldots, [s(K-1)\; b_1(K-1)]^T\}$, it is straightforward to obtain the iterative updating rule

$$\begin{aligned}\hat{\mathbf{a}}(i+1) &= \left(\sum_{n=0}^{K-1} \mathbf{y}(n)\mathbf{y}^H(n)\right)^{-1} \\ &\times \left(\sum_{n=0}^{K-1} \mathbf{y}(n) E_{b_1(n)\,|\,s(n);\hat{\mathbf{a}}(i)}[b_1^*(n)]\right).\end{aligned} \tag{48}$$

The conditioned expectation in the above equation can be calculated via the Bayes theorem to yield

$$E_{b_1(n)\,|\,s(n);\hat{\mathbf{a}}(i)}[\Phi(b_1(n))] = \frac{E_{b_1(n)}\left[e^{-\frac{|s(n)-b_1(n)|^2}{\sigma_a^2}}\Phi(b_1(n))\right]}{E_{b_1(n)}\left[e^{-\frac{|s(n)-b_1(n)|^2}{\sigma_a^2}}\right]} \tag{49}$$

where $\Phi(\cdot)$ is an arbitrary funtion of $b_1(n)$.

[3]Furthermore, note that $g_a(n)$ is a zero-mean white process. Indeed, since $E[g(n)g^*(n)] = \sigma_g^2\mathbf{I}_L$, it is straightforward to show that $E[g_a(n)g_a^*(n+k)] = \delta(k)\sigma_g^2\mathbf{a}^H\mathbf{W}^H\mathbf{W}\mathbf{a}$, where $\delta(\cdot)$ is Kronecker's delta function.

VI. Computer Simulations

Finally, we present computer simulations that illustrate the validity of our approach. We have considered an asynchronous DS CDMA communication system with N users transmitting QPSK symbols and length $L = 16$ random binary spreading codes. The length of the discrete-time equivalent channel response for each user is $P = 4$. Recall from Section I that the discrete-time channel coefficients account for the continuous-time channel response as well as for the relative delays of all users and the transmitter and receiver terminal filters. Symbols are transmitted in blocks of length K, and the channel coefficients are assumed to vary slowly enough so that they remain constant for the duration of the block. Unless something different is stated, the simulation results presented in this section have been averaged over 150 randomly generated sets of multiaccess channels. To obtain these sets, we have considered a Rayleigh channel model where each channel coefficient $h_i(p)$ $(i = 1, \ldots, N, p = 0, \ldots, P-1)$ is modeled as a complex random variable with statistically independent real and imaginary parts, where both of them are Gaussian with zero mean and standard deviation $\sigma_h = 0.3$. In order to estimate the symbol error rate (SER), we have simulated the demodulation of 10 000 length K independent data blocks for each different channel. The SAGE algorithm used to estimate the parameters of the LCMLL receivers is initialized with $\hat{\mathbf{w}}_u(0) = \mathbf{0}$ and $\hat{A}_1(0) = 1$, whereas the EM algorithm used to compute the weight vector in the blind rake receiver is always initialized with $\hat{\mathbf{a}}(0) = [1/P, \ldots, 1/P]^T$.

Fig. 4(a) plots the SER attained by the proposed LCMLL receiver for several values of the input SNR, which is defined as

$$\text{SNR} = 10 \log_{10} \frac{E[|b_1|^2]}{\sigma_g^2} \tag{50}$$

when the linear constraint is set to protect the desired user component corresponding to the propagation path with the highest gain. The number of system users is $N = 7$, the number of observation vectors available to estimate the receiver coefficients is $K = 100$, and the value of the filtered noise variance required to run the SAGE algorithm is roughly approximated by the channel noise variance, i.e., we use an estimate $\hat{\sigma}_f^2 = \sigma_g^2$ instead of the true value of σ_f^2. It is apparent that the proposed algorithm performs close to the theoretical LCMV receiver constructed with the same linear constraint as the proposed receiver and perfect knowledge of the channel vectors $\mathbf{h}_1, \ldots, \mathbf{h}_N$. In this figure, we have also plotted the SER achieved by the LCMV receiver constructed using an estimation of the autocorrelation matrix $\hat{\mathbf{R}}_x = (1/K) \sum_{n=0}^{K} \mathbf{x}(n)\mathbf{x}^H(n)$, as should be done in practice. It can be seen that the performance of the practical LCMV receiver is considerably worse than the theoretical one because 100 observations are not enough to obtain an adequate estimation of the 16×16 true autocorrelation matrix $\mathbf{R}_x = E_{\mathbf{x}(n)}[\mathbf{x}(n)\mathbf{x}^H(n)]$. We have repeated the previous experiments for $K = 300$ observation vectors and plotted the resulting curve in Fig. 4(b). It can be seen that the LCMLL receiver performance matches the theoretical limit, and the practical LCMV receiver also approaches this limit, but its convergence is still poorer for the medium to high SNR region.

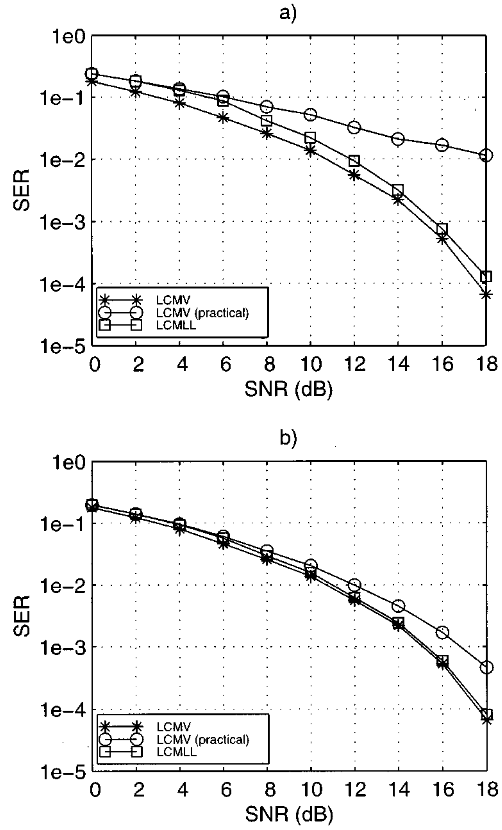

Fig. 4. SER for several values of the SNR in a time-dispersive asynchronous DS CDMA system with $N = 7$ users, length $L = 16$ random binary spreading codes, and length $P = 4$ discrete channels. (a) $K = 100$. (b) $K = 300$.

We have also verified the robustness of the proposed LCMLL receiver in near-far enviroments. Let us define the signal-to-interference ratio (SIR) of the desired user w.r.t. the jth interference as

$$\text{SIR}_j = 10 \log_{10} \frac{E[|b_1|^2]}{E[|b_j|^2]}. \tag{51}$$

First, we have chosen the value of $E[|b_j|^2]$ so that $\text{SIR}_j = -5$ dB $\forall j$. The resulting SER curves for the LCMLL receiver, the theoretical LCMV receiver, and the practical LCMV receiver are plotted in Fig. 5(a) for $N = 7$ and $K = 100$. No degradation in performance is observed, and the proposed receiver still approaches the theoretical limit. The near-far resistance property of the LCMLL detector is clearly illustrated in Fig. 5(b), where only a very slight performance loss is appreciated when $\text{SIR}_j = -10$ dB $\forall j$, and again, the theoretical performance limit is practically matched.

Another important measure of the receiver performance is the SER achieved for different system loads. Fig. 6 plots the SER for several values of the number of users N when the block size is $K = 100$ and SNR $= 12$ dB. The resulting curve shows that the performance degradation of the LCMLL receiver with increasing system load is the same one suffered by the theoretical LCMV receiver, whereas the practical LCMV receiver performance is considerably worse.

Fig. 7 illustrates the fast convergence speed of the SAGE algorithm. With the same simulation parameters as in Fig. 4(a) and

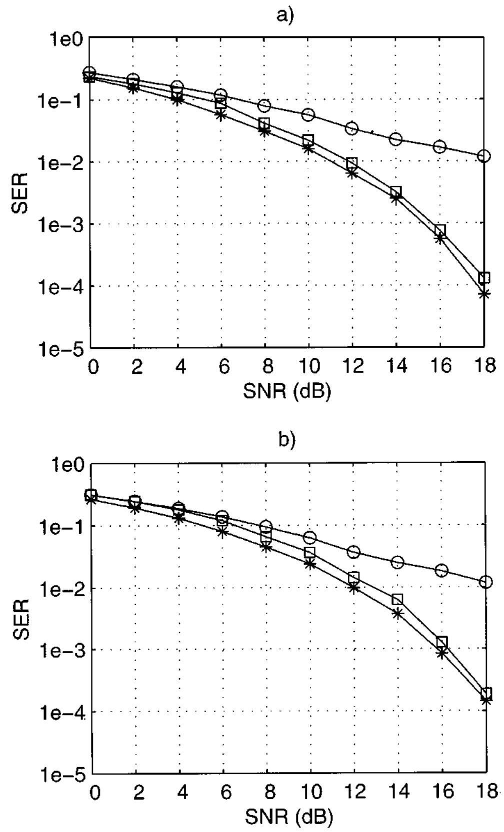

Fig. 5. SER for several values of the SNR in a time-dispersive asynchronous DS CDMA system with $N = 7$ users, length $L = 16$ random binary spreading codes, length $P = 4$ discrete channels, and block size $K = 100$. (a) $\text{SIR}_j = -5$ dB. (b) $\text{SIR}_j = -10$ dB.

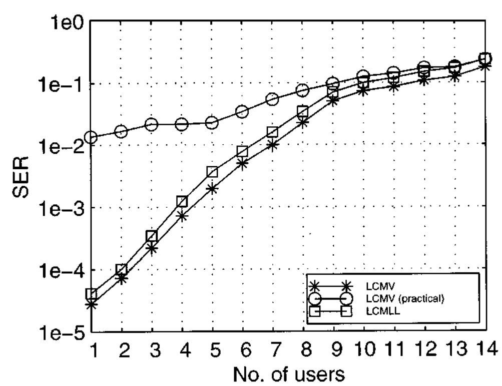

Fig. 6. SER for several values of the number of users in a time-dispersive asynchronous DS CDMA system with length $L = 16$ random binary spreading codes, length $P = 4$ discrete channels, block size $K = 100$, and SNR = 12 dB.

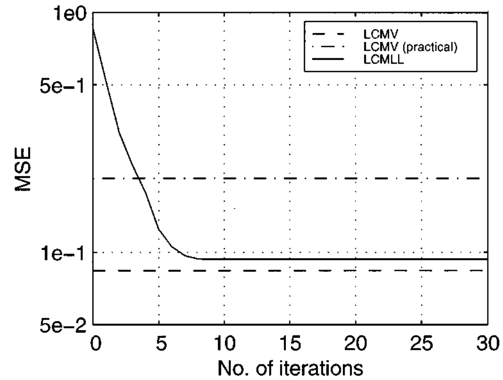

Fig. 7. MSE versus the number of iterations of the SAGE algorithm that obtains the LCMLL receiver coefficients in a DS CDMA system with $N = 7$ users, length $L = 16$ random binary spreading codes, length $P = 4$ discrete channels, observation block size $K = 100$, and SNR = 12 dB.

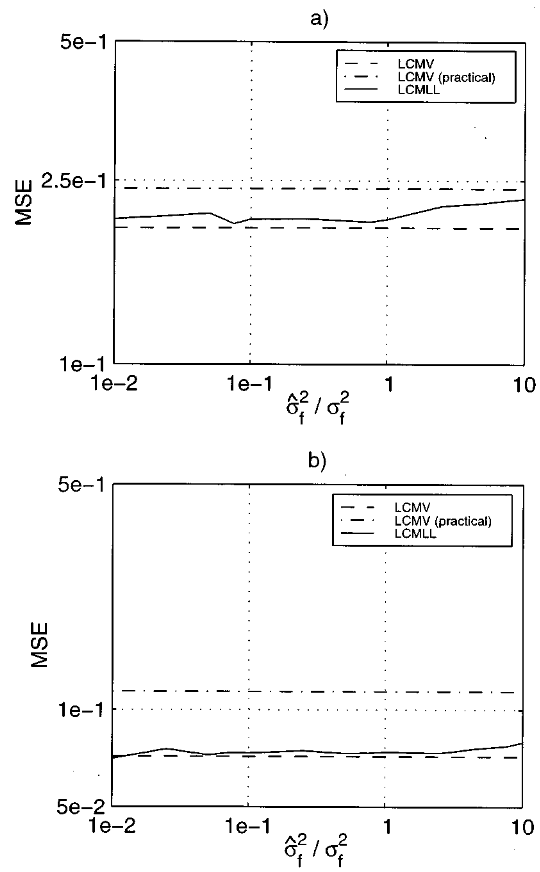

Fig. 8. MSE versus $\hat{\sigma}_f^2/\sigma_f^2$ in a time-dispersive asynchronous DS CDMA system with $N = 7$ users, length $L = 16$ random binary spreading codes, length $P = 4$ discrete channels, and block size $K = 100$. (a) SNR = 6 dB. (b) SNR = 12 dB.

a fixed SNR value of 12 dB, we have plotted the mean square error (MSE) at the receiver output as a function of the number of iterations of algorithm in (36) and (37). It is apparent that very few iterations are enough to obtain the receiver filter coefficients. This may be an important advantage when time or computational load constraints have to be fullfilled.

In order to verify the robustness of the LCMLL receiver to mismatches in the selection of the tentative filtered noise variance $\hat{\sigma}_f^2$, we have measured the MSE that is attained (in a system with $N = 7$ users and block size $K = 100$) w.r.t. the ratio $\hat{\sigma}_f^2/\sigma_f^2$, where σ_f^2 is the true value of the filtered noise variance when the optimum filter $\mathbf{w}_*$ is employed. The results can be observed in Figs. 8(a) and (b) for SNR values of 6 dB and 12 dB, respectively. It can be clearly seen that the MSE hardly varies, even when the deviation in the selected variance is very large.

Finally, we present some computer simulations that illustrate the performance of the blind rake receiver. Fig. 9 shows the performance improvement that can be achieved when using the rake receiver instead of a plain LCMLL receiver. We have considered

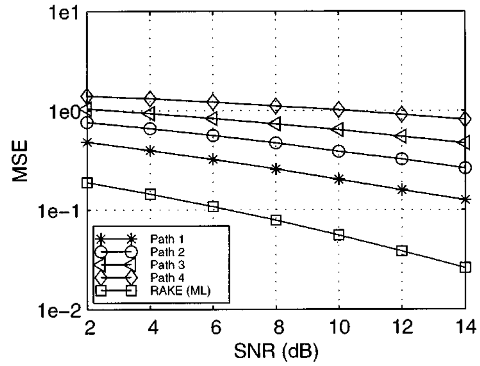

Fig. 9. MSE for several values of the SNR in a time-dispersive asynchronous DS CDMA system with $N = 10$ users, length $L = 16$ random binary spreading codes, length $P = 4$ discrete channels, and block size $K = 300$.

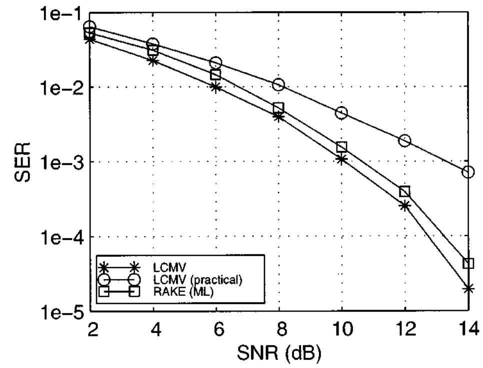

Fig. 10. SER for several values of the SNR in a time-dispersive asynchronous DS CDMA system with $N = 10$ users, length $L = 16$ random binary spreading codes, length $P = 4$ discrete channels, and $K = 300$.

a system with $N = 10$ users and block size $K = 300$ symbols. The curve labeled *Path 1* represents the MSE of the LCMLL receiver that extracts the desired user signal arriving through the strongest path, whereas *Paths 2, 3,* and *4* correspond to the desired user signal extracted from each one of the remaining paths in decreasing power order. When the soft estimates from the bank of LCMLL receivers are linearly combined using the proposed blind rake solution, a considerable reduction in the MSE is obtained, as shown by the curve labeled *RAKE (ML)*.

Fig. 10 shows the SER achieved by the blind rake receiver for several values of SNR when the number of users in the system is $N = 10$ and the block size is $K = 300$. The theoretical performance limit of this receiver is given by the linear LCMV detector

$$\mathbf{w}_{\text{LCMV}} = \arg\min_{\mathbf{w}} \{E_x[|\mathbf{w}^H \mathbf{x}|^2]\} \quad \text{subject to } \mathbf{w}^H \mathbf{C}_1(0)\mathbf{h}_1 = 1 \tag{52}$$

where the linear constraint requires the desired user channel to be known.[4] The solution of problem (52) includes the observation autocorrelation matrix $\mathbf{R}_x$, which depends on the received codes of all the system users. This knowledge is not usually available, so we distinguish the theoretical LCMV detector (curve labeled *LCMV*) and the practical implementation where $\mathbf{R}_x$ should be estimated from the available observations [curve labeled *LCMV (practical)*]. It is apparent that the proposed rake multiuser receiver practically matches the SER of the theoretical LCMV receiver and clearly outperforms the practical LCMV detector.

We have also evaluated the convergence speed of the EM algorithm to compute the weight vector $\hat{\mathbf{a}}$ in the blind rake receiver. Fig. 11 shows that convergence is achieved in less than 30 iterations when considering a system with $N = 10$ users, block size $K = 300$, and SNR $= 12$ dB.

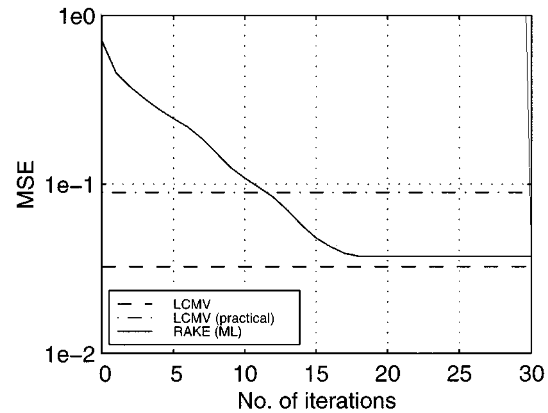

Fig. 11. MSE versus the number of iterations of the EM algorithm that obtains the weight vector $\hat{\mathbf{a}}$. System parameters: $N = 10$ users, length $L = 16$ random binary spreading codes, length $P = 4$ discrete channels, block size $K = 300$, and SNR $= 12$ dB.

VII. Conclusion

We have introduced a new blind approach to linear interference cancellation in DS CDMA that relies on the ML criterion to estimate the coefficients of a linear FIR filter that suppresses both MAI and ISI. The method is blind because it does not require the transmission of training sequences, but in turn, it exploits the knowledge of the pdf of the transmitted symbols and noise. Since the statistical characterization of all user signals is the same, a linear constraint has to be set on the receiver coefficients to ensure that the desired user is extracted. As a result, a linearly constrained maximum likelihood linear (LCMLL) multiuser receiver is obtained that can be efficiently implemented using the iterative SAGE algorithm.

The LC imposes an important limitation on the performace of the MLL receiver because it does not allow to exploit the temporal diversity inherent to multipath channels. To circumvent this drawback, we have introduced a blind rake multiuser receiver that proceeds in two steps. First, soft estimates of the desired user-transmitted symbols are obtained from each propagation path using a bank of appropiate LCMLL receivers, and second, these soft estimates are suitably combined to increase the SINR. The weight vector for this linear combination is also estimated according to the ML criterion.

Computer simulations show that the proposed blind multiuser receivers exhibit considerable near–far resistance and attain convergence using small blocks of observations ($100 \leq K \leq 300$), thus outperforming existing blind LCMV receivers.

[4]Notice that the proposed blind rake receiver does not have such knowledge.

APPENDIX A
DERIVATION OF $f_{A_1 b+g_f}(\cdot)$ AND $f_{\mathbf{Y}_e;\Theta}(\cdot)$

A. Derivation of $f_{A_1 b+g_f}(\cdot)$

Let us assume that the N system users employ the same modulation format with i.i.d. and equiprobable symbols. Thus, an arbitrary symbol $b_i(n)$ belongs to the finite alphabet $\mathcal{B} = \{b^{(1)}, \ldots, b^{(M)}\}$, where $\log_2 M$ is the number of bits per symbol, and its pdf is

$$f_b(b_i(n)) = \frac{1}{M}\sum_{l=1}^{M} \delta\left(b_i(n) - b^{(l)}\right) \tag{53}$$

where $\delta(\cdot)$ is Kronecker's delta function. Obviously, the pdf of the rescaled symbol $A_i b_i(n)$ is

$$f_{A_i b}(A_i b_i(n)) = \frac{1}{M}\sum_{l=1}^{M} \delta\left(A_1 b_i(n) - A_1 b^{(l)}\right). \tag{54}$$

The pdf of the noisy rescaled symbols $A_1 b_i(n)+g_f(n)$ is simply the convolution of $f_{A_1 b}(\cdot)$ and the Gaussian pdf $f_{g_f}(\cdot)$, i.e.,

$$\begin{aligned} f_{A_1 b+g_f}(z) &= f_{A_1 b}(z) * f_{g_f}(z) \\ &= \frac{1}{M}\sum_{l=1}^{M} f_{g_f}\left(z - A_1 b^{(l)}\right) \\ &= \frac{1}{M}\sum_{l=1}^{M} \frac{1}{\pi\sigma_f^2} e^{-\frac{|z-A_1 b^{(l)}|^2}{\sigma_f^2}} \\ &= \frac{1}{\pi\sigma_f^2} E_b e^{-\frac{|z-A_1 b|^2}{\sigma_f^2}}. \end{aligned} \tag{55}$$

B. Derivation of $f_{\mathbf{Y}_e;\Theta}(\cdot)$

When Θ is adequately chosen (i.e., $y(n) = \mathbf{w}_*^H \mathbf{x}(n) = A_1 b_1(n) + g_f(n)$) the extended vector $\mathbf{y}_e(n) = \left[\begin{smallmatrix} y(n) \\ b_1(n)\end{smallmatrix}\right]$ is easily obtained through a linear invertible transformation of the extended symbol vector $\mathbf{b}_e(n) = \left[\begin{smallmatrix} g_f(n) \\ b_1(n)\end{smallmatrix}\right]$ as

$$\mathbf{y}_e(n) = t_{A_1}(\mathbf{b}_e(n)) = \begin{bmatrix} A_1 b_1(n) + g_f(n) \\ b_1(n)\end{bmatrix}. \tag{56}$$

It is well known that the pdf's of $\mathbf{y}_e(n)$ and $\mathbf{b}_e(n)$ are related by [27]

$$f_{\mathbf{y}_e;\mathbf{w}_*}(\mathbf{y}_e(n)) = \frac{f_{\mathbf{b}_e}\left(t_{A_1}^{-1}(\mathbf{y}_e(n))\right)}{|J_t|} \tag{57}$$

where J_t is the Jacobian of the transformation, and $|\cdot|$ denotes absolute value. It is straightforward to show that

$$J_t = \det\begin{bmatrix} \frac{\partial y(n)}{\partial g_f} & \frac{\partial b_1(n)}{\partial g_f} \\ \frac{\partial y(n)}{\partial b_1(n)} & \frac{\partial b_1(n)}{\partial b_1(n)}\end{bmatrix} = \begin{vmatrix} 1 & 0 \\ A_1 & 1\end{vmatrix} = 1 \tag{58}$$

and, assuming $g_f(n)$ is statistically independent of $b_1(n)$

$$\begin{aligned} f_{\mathbf{y}_e;\Theta}(\mathbf{y}_e(n)) &= f_{\mathbf{b}_e}\left(t_{A_1}^{-1}(\mathbf{y}_e(n))\right) \\ &= f_b(b_1(n)) \cdot f_{g_f}(y(n) - A_1 b_1(n)) \\ &= f_b(b_1(n)) \cdot \frac{1}{\pi\sigma_f^2} e^{-\frac{|y(n)-A_1 b_1(n)|^2}{\sigma_f^2}}. \end{aligned} \tag{59}$$

Since the transmitted symbols are i.i.d., the joint pdf of the K extended vectors $\mathbf{Y}_e = [\mathbf{y}_e(0), \ldots, \mathbf{y}_e(K-1)]$ can be written as

$$f_{\mathbf{Y}_e;\Theta}(\mathbf{Y}_e) = \prod_{n=0}^{K-1} \frac{1}{\pi\sigma_f^2} e^{-\frac{|y(n)-A_1 b_1(n)|^2}{\sigma_f^2}} f_b(b_1(n)). \tag{60}$$

APPENDIX B
CAPTURE PROBLEM

In this Appendix, we show that if the soft estimates pdf matches the target pdf, i.e.,

$$f_{y;\mathbf{w}_*,A_1}(y(n)) = f_{A_1 b+g_f}(y(n)) \tag{61}$$

where $g_f(n) = \mathbf{w}_*^H g(n)$, the receiver necessarily extracts the desired user and not an interferent one. When the filter coefficients are subject to the constraint

$$\mathbf{w}_*^H \mathbf{C}_1(0) = \mathbf{u}_d^T \tag{62}$$

and $h_1(d) = A_1$ with non-negligible $|h_1(d)|$, the resulting soft estimates can be written as

$$\begin{aligned} y(n) &= \mathbf{w}_*^H \mathbf{d}_1(0) b_1(n) + \sum_{i=1}^{m-1} \mathbf{w}_*^H \mathbf{d}_1(i) b_1(n-i) \\ &\quad + \sum_{k=2}^{N}\sum_{i=0}^{m-1} \mathbf{w}_*^H \mathbf{d}_k(i) b_k(n-i) + \mathbf{w}_*^H \mathbf{g}(n) \\ &= A_1 b_1(n) + \sum_{i=1}^{m-1} A_1^{(i)} b_1(n-i) \\ &\quad + \sum_{k=2}^{N}\sum_{i=0}^{m-1} A_k^{(i)} b_k(n-i) + \mathbf{w}_*^H \mathbf{g}(n). \end{aligned} \tag{63}$$

Since the symbols transmitted by the N users are i.i.d. discrete random variables, the soft estimates pdf is

$$\begin{aligned} & f_{y;\mathbf{w}_*,A_1}(y(n)) \\ &\quad = f_{A_1 b}(y(n)) * f_{A_1^{(1)} b}(y(n)) \\ &\qquad * \cdots * f_{A_1^{(m-1)} b}(y(n)) * f_{A_2^{(0)} b}(y(n)) \\ &\qquad * \cdots * f_{A_2^{(m-1)} b}(y(n)) * \cdots * f_{A_N^{(0)} b}(y(n)) \\ &\qquad * \cdots * f_{A_N^{(m-1)} b}(y(n)) * f_{g_f}(y(n)). \end{aligned} \tag{64}$$

Since

$$f_{A_i^{(j)} b}(y(n)) = \frac{1}{M}\sum_{l_{i,j}=1}^{M} \delta\left(y(n) - A_i^{(j)} b^{(l_{i,j})}\right) \tag{65}$$

it follows that

$$f_{y;\mathbf{w}_*,A_1}(y(n)) = f_{A_1 b}(y(n)) * \rho(n) * f_{g_f}(y(n)) \tag{66}$$

where

$$\rho(n) = \frac{1}{M^{Nm-1}} \left[\sum_{l_{1,1}=1}^{M} \cdots \sum_{l_{N,m-1}=1}^{M} \delta \left(y(n) - A_N^{(m-1)} b^{(l_{N,m-1})} - \cdots - A_1^{(1)} b^{(l_{1,1})} \right) \right] \quad (67)$$

and it is clear that

$$f_{y;\mathbf{w}_*,A_1}(y(n)) = f_{A_1 b}(y(n)) * f_{g_f}(y(n)) = f_{A_1 b + g_f}(y(n)) \quad (68)$$

if, and only if

$$\rho(n) = \delta(y(n)) \quad (69)$$

which is equivalent to $A_i^{(j)} = 0 \, \forall i, j$.

Appendix C
Relationship Between the LCMLL and the Linear MMSE Receivers

Let us consider the linear MMSE multiuser receiver subject to the same linear constraint in problem (20), i.e.,

$$\mathbf{w}_{\text{MMSE}} = \arg\min_{\mathbf{w}} \left\{ E_{\mathbf{x}(n)}[|\mathbf{w}^H \mathbf{x}(n) - b_1(n)|^2] \right\} \quad \text{subject to } \mathbf{w}^H \mathbf{C}_1(0) = \mathbf{u}_d^T. \quad (70)$$

Applying the GSC decomposition, it is straightforward to show that the solution to the above problem is

$$\mathbf{w}_{\text{MMSE}} = \mathbf{w}_q - \mathbf{B}\mathbf{w}_{u,\text{MMSE}} \quad (71)$$

$$\mathbf{w}_{u,\text{MMSE}} = (\mathbf{B}^H \mathbf{R}_x \mathbf{B})^{-1} \mathbf{B}^H (\mathbf{R}_x \mathbf{w}_q - \mathbf{p}) \quad (72)$$

where $\mathbf{R}_x = E_{\mathbf{x}(n)}[\mathbf{x}(n)\mathbf{x}^H(n)]$, and $\mathbf{p} = E_{\mathbf{x}(n)}[\mathbf{x}(n) b_1^*(n)]$.

In this Appendix, we will show the close relationship between the LCMLL receiver and the MMSE solution given by (72). Toward this aim, let us characterize the local maxima of the log-likelihood function

$$\mathcal{L}(\mathbf{w}_u, A_1) = \sum_{n=0}^{K-1} \log E_b \left[e^{-\frac{|y(n) - A_1 b|^2}{\sigma_f^2}} \right] \quad (73)$$

w.r.t. the unconstrained vector $\mathbf{w}_u$. The stationary points of $\mathcal{L}(\mathbf{w}_u)$ are found by calculating the gradient $\nabla_{\mathbf{w}_u} \mathcal{L}$ and equalling it to zero as

$$\nabla_{\mathbf{w}_u} \mathcal{L} = \sum_{n=0}^{K-1} \mathbf{B}^H \mathbf{x}(n) E_{b_1(n)\,|\,y(n);\mathbf{w}_u} [y^*(n) - A_1 b_1^*(n)] = 0. \quad (74)$$

Taking into account that $y(n) = \mathbf{w}^H \mathbf{x}(n)$ and the GSC decomposition $\mathbf{w} = \mathbf{w}_q - \mathbf{B}\mathbf{w}_u$, the previous equation can be elaborated to yield

$$\sum_{n=0}^{K-1} \mathbf{B}^H \mathbf{x}(n)\mathbf{x}^H(n) \mathbf{B}\mathbf{w}_u = \sum_{n=0}^{K-1} \mathbf{B}^H \mathbf{x}(n) e^*(n) \quad (75)$$

where

$$e^*(n) = \mathbf{x}^H(n)\mathbf{w}_q - A_1^* E_{b_1(n)\,|\,y(n);\mathbf{w}_u}[b_1^*(n)] \quad (76)$$

and the conditioned expectation is the nonlinear mean-squared estimate of $b_1^*(n)$ [28]. Solving for $\mathbf{w}_u$, we arrive at

$$\mathbf{w}_u = (\mathbf{B}^H \hat{\mathbf{R}}_x \mathbf{B})^{-1} \mathbf{B}^H (\hat{\mathbf{R}}_x \mathbf{w}_q - A_1^* \hat{\mathbf{p}}) \quad (77)$$

where

$$\hat{\mathbf{R}}_x = \sum_{n=0}^{K-1} \mathbf{x}(n)\mathbf{x}^H(n) \quad (78)$$

is the empirical autocorrelation matrix, and

$$\hat{\mathbf{p}} = \sum_{n=0}^{K-1} \mathbf{x}(n) E_{b_1(n)\,|\,y(n)}[b_1^*(n)] \quad (79)$$

is an empirical cross-correlation vector where the transmitted symbols are substituted by their mean-squared estimates. Except for the scale factor A_1^*, it is apparent that (77) converges to the MMSE solution (72) when the block size K is large enough.

Notice that (77) is not a useful result from a practical point of view since it does not provide a closed-form solution for $\mathbf{w}_u$. This unconstrained vector must be known in order to compute the mean-squared estimates of the symbols. The SAGE algorithm proposed in this paper is actually an iterative method to numerically approximate solution (77).

References

[1] B. Chen and P. Willet, “The theoretical advantage of CDMA over FDMA in a Gaussian MAC,” *IEEE Trans. Inform. Theory*, vol. 45, pp. 2046–2053, Sept. 1999.

[2] E. Dahlman, B. Gudmundson, and M. Nilsson, “UMTS/IMT-2000 based on wideband CDMA,” *IEEE Commun. Mag.*, pp. 70–80, Sept. 1998.

[3] F. Adachi, M. Sawahashi, and H. Suda, “Wideband CDMA for next-generation mobile communications systems,” *IEEE Commun. Mag.*, pp. 56–59, Sept. 1998.

[4] S. Verdú, *Multiuser Detection*, Cambridge, U.K.: Cambridge Univ. Press, 1998.

[5] U. Madhow and M. L. Honig, “MMSE interference suppression for direct-sequence spread-spectrum CDMA,” *IEEE Trans. Commun.*, vol. 42, pp. 3178–3188, Dec. 1994.

[6] M. L. Honig, U. Madhow, and S. Verdú, “Blind adaptive multiuser detection,” *IEEE Trans. Inform. Theory*, vol. 41, pp. 944–960, July 1995.

[7] M. K. Tsatsanis, “Inverse filtering criteria for CDMA systems,” *IEEE Trans. Signal Processing*, vol. 45, pp. 102–112, Jan. 1997.

[8] J. Míguez and L. Castedo, “A constant modulus blind adaptive receiver for multiuser interference suppression,” *Signal Process.*, vol. 71, no. 1, pp. 15–27, Nov. 1998.

[9] U. Madhow, “Blind adaptive interference suppression for the near-far resistant acquisition and demodulation of direct-sequence CDMA signals,” *IEEE Trans. Signal Processing*, vol. 45, pp. 124–136, Jan. 1997.

[10] D. D. Feldman and L. J. Griffiths, “A projection approach for robust adaptive beamforming,” *IEEE Trans. Signal Processing*, vol. 42, pp. 867–876, Apr. 1994.

[11] X. Wang and H. V. Poor, "Blind equalization and multiuser detection in dispersive CDMA channels," *IEEE Trans. Commun.*, vol. 46, pp. 91–103, Jan. 1998.

[12] ——, "Blind multiuser detection: A subspace approach," *IEEE Trans. Inform. Theory*, vol. 44, pp. 677–690, Mar. 1998.

[13] S. E. Bensley and B. Aazhang, "Subspace-based channel estimation for code division multiple access communication systems," *IEEE Trans. Commun.*, vol. 44, pp. 1009–1020, Aug. 1996.

[14] G. J. McLachlan and T. Krishnan, *The EM Algorithm and Extensions*. New York: Wiley, 1997.

[15] H. Zamiri-Jafarian and S. Pasupathy, "Adaptive MLSDE using the EM algorithm," *IEEE Trans. Commun.*, vol. 47, pp. 1181–1193, Aug. 1999.

[16] ——, "EM-based recursive estimation of channel parameters," *IEEE Trans. Commun.*, vol. 47, pp. 1297–1302, Sept. 1999.

[17] L. Tong and S. Perreau, "Multichannel blind identification: From subspace to maximum likelihood methods," *IEEE Trans. Signal Processing*, vol. 45, pp. 1343–1348, May 1998.

[18] A. M. Sayeed, A. Sendonaris, and B. Aazhang, "Multiuser detection in fast-fading multipath environments," *IEEE J. Select. Areas Commun.*, vol. 16, pp. 1691–1701, Dec. 1998.

[19] J. G. Proakis, *Digital Communications*, Singapore: McGraw-Hill, 1995.

[20] J. S. Álvarez and G. Vázquez-Grau, "Statistical reference criteria for adaptive signal processing in digital communications," *IEEE Trans. Signal Processing*, vol. 45, pp. 14–31, Jan. 1997.

[21] L. J. Griffiths and C. W. Jim, "An alternative approach to linearly constrained adaptive beamforming," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-30, pp. 2737–2746, Nov. 1982.

[22] L. B. Nelson and V. Poor, "Iterative multiuser receivers for CDMA channels: An EM-based approach," *IEEE Trans. Commun.*, vol. 44, pp. 1700–1710, Dec. 1996.

[23] J. A. Fessler and A. O. Hero, "Space-alternating generalized expectation-maximization algorithm," *IEEE Trans. Signal Processing*, vol. 42, pp. 2664–2677, Oct. 1994.

[24] B. F. Plybon, *An Introduction to Applied Numerical Analysis*. Boston, MA: PWS-Kent, 1992.

[25] G. W. Stewart, *Matrix Algorithms*. Philadelphia, PA: SIAM, 1998.

[26] J. R. Rice, *Matrix Computations ans Mathematical Software*. New York: McGraw-Hill, 1981.

[27] F. Kai-Tai and Z. Yao-Ting, *Generalized Multivariate Analysis* Beijing, China, 1990.

[28] C. W. Therrien, *Discrete Random Signals and Statistical Signal Processing*. Englewood Cliffs, NJ: Prentice-Hall, 1992.

Mónica F. Bugallo (S'98) was born in Ferrol, Spain, in 1975. She received the M.Sc. degree in computer engineering from the University of A Coruña, Coruña, Spain, in 1998.

From 1998 to 2000, she participated in a research project on xDSL supported by the European Union. She is currently holding a scholarship from the local government of Galicia, Spain (Xunta de Galicia). Her research interests lie in the area of statistical signal processing focused on interference suppression and filtering in multiuser communication systems.

Joaquín Míguez (M'01) was born in Ferrol, Spain, in 1974. He received the M.Sc. and Ph.D. degrees in computer engineering from the University of A Coruña, Coruña, Spain, in 1997 and 2000, respectively. From 1999 to 2000, he held a grant from the Xunta de Galicia to pursue the Ph.D. degree at the Department of Electronics and Systems, University of A Coruña.

Since late 2000, he has been an Assistant Professor with the same departament. His research interests are in the area of signal processing for communications, including multiuser detection, space-time coding, and adaptive filtering.

Luis Castedo (A'95) was born in Santiago de Compostela, Spain, in 1966. He received the Ingeniero de Telecomunicacion and Dr.Ing. de Telecomunicación degrees, both from the Universidad Politécnica de Madrid (UPM), Madrid, Spain, in 1990 and 1993, respectively.

From 1990 to 1994, he was with the Departamento de Señales, Sistemas, y Radiocomunicación at UPM, where he worked on array processing applied to digital communications. In 1994, he joined the Departamento de Electrónica y Sistemas, Universidad de A Coruña, Coruña, Spain, where he is currently Associate Professor and teaches courses in signal processing, digital communications, and linear control systems. His research interests include blind adaptive filtering and signal processing methods for space and code diversity exploitation in communication systems.