A Two-stage CMA-Based Receiver for Blind Joint Equalization and Multiuser Detection in High Data-Rate DS-CDMA Systems

Giacinto Gelli, *Member, IEEE,* Luigi Paura, and Francesco Verde

***Abstract*— The paper deals with the problem of blind mitigation of intersymbol interference (ISI) as well as multiple-access interference (MAI) in asynchronous high data-rate direct-sequence code-division multiple-access (DS-CDMA) systems. A blind adaptive multiuser receiver based on the constant-modulus algorithm (CMA) is proposed, which demodulates each desired user by exploiting only the knowledge of its spreading code, without requiring estimation of the users's channels and timings. In order to overcome the CMA interference capture problem, which arises in a multiuser scenario, a two-stage adaptive receiver is adopted: in the first stage, partial MAI and ISI suppression is blindly achieved by exploiting the desired user signature structure properties; in the second stage, the residual MAI and the ISI are removed by using the CMA, and the information symbols of the desired user are reliably recovered. Theoretical analysis and simulation results show that the first stage is an effective blind adaptive strategy which allows the CMA detector in the second stage to lock on the desired-user symbol, at a particular delay. The proposed blind receiver achieves a significant performance gain in comparison with existing blind methods.**

***Index Terms*— Adaptive multiuser detection, blind equalization, code-division multiple-access (DS-CDMA), constant modulus algorithm.**

I. INTRODUCTION

THE need of a reliable wireless transmission technology for delivering advanced broadband multimedia services, with data rates of the order of tens of megabits per second or even higher, has driven a significant bulk of research towards design and optimization of high data-rate direct-sequence code-division (DS-CDMA) systems, which offer a high degree of service and operator flexibility, with a manageable implementation complexity. For low-rate DS-CDMA systems that are essentially oriented to voice communications, the multipath nature of the wireless channel can be neglected and, thus, the major source of performance degradation is represented by the multiple-access interference (MAI), which arises from mutual interference among all the (non-orthogonal) active users of the system. However, when CDMA technology is employed for high data-rate wireless networks envisioned to support multimedia applications, the channel multipath spread might exceed the symbol period, giving rise to intersymbol interference (ISI) and, hence, the system performance becomes severely limited by both ISI and MAI. In such context, conventional *single-user* detection techniques (e.g., matched filter or RAKE receiver), which are already highly susceptible to the *near-far* problem in low-rate DS-CDMA systems, cannot longer be used due to the long-delay multipath [1]. On the other hand, optimal *multiuser* detection [2], i.e., joint maximum-likelihood (ML) detection of all the active users of the system, provides a significant performance gain over single-user receivers; unfortunately, already in the low-rate context, the ML solution exhibits an exponential complexity in the number of users. Moreover, the implementation of the ML detector is much more impractical in high data-rate DS-CDMA (namely, in the presence of ISI) than in low-rate ones since, besides knowledge of all active users' spreading codes, amplitudes and timings, it requires explicit knowledge of their channel-impulse responses. Among the linear suboptimal multiuser receivers [2], the minimum mean-square error (MMSE) detector [3] offers high near-far resistance, without sharing the prohibitive computational complexity of the ML detector, and has the important property that a single user can be detected without detecting all other users [4], that is, without requiring knowledge of the interfering users' codes, channel-impulse responses and timings. In the context of low-rate DS-CDMA systems, the MMSE multiuser detector can be implemented via *blind* (i.e., no training sequence) adaptation, requiring the knowledge of desired user's synchronization and spreading code [5] or requiring only the knowledge of the desired user's spreading code [6]. When the multipath nature of the channel is no longer negligible, besides knowledge of the desired user's synchronization and spreading code, the implementation of the MMSE multiuser receiver still requires training sequences of symbols for the desired transmission [7]. However, since the periodic transmission of training sequences significantly reduces the spectrum efficiency, the derivation of blind multiuser receivers with MMSE performance, which eliminate the use of training sequences, has been gaining recently a significant attention in the literature. To this end, it has been observed (see, e.g., [8]) that, after sampling the received signal at the chip-rate, the problem of multiuser detection in multipath channels turns out to be equivalent to joint equalization and separation of multiple signals in a finite-impulse-response (FIR) multiple-input (the number of users) multiple-output (the processing gain) (MIMO) channel. Based on this observation, several blind multiuser receivers

Manuscript received xxx; revised xxx and xxx

The authors are with the Dipartimento di Ingegneria Elettronica e delle Telecomunicazioni, Università degli Studi di Napoli Federico II, via Claudio 21, I-80125 Napoli, Italy (e-mail gelli@unina.it, paura@unina.it, f.verde@unina.it)

Publisher Item Identifier S 0000-0000(00)00000-3.

were proposed by integrating ideas from the area of blind source separation and equalization for MIMO systems. In [9], [10] a multiuser generalization of the Godard algorithm [11], also known as the constant modulus algorithm (CMA) [12], was proposed, where all the active users are separated in the presence of ISI, requiring knowledge of users' codes and timings. However, when employed in a fully blind setting, the multiuser CMA suffers of an intrinsic detection ambiguity: the detector is capable of perfectly recovering (in the absence of noise) all the active users of the system up to (unknown) permutations in the ordering of the signals (see the discussion in [10]). Moreover, this receiver can be adopted only in the uplink (i.e., mobile to base station) of a DS-CDMA network, since its synthesis requires knowledge of the codes and timings of all the active users of the system, which, in its turn, may lead to an unsustainable implementation complexity [5]. In this respect, it should be noted that, without a good initialization guess, the estimation of the multiuser cost function proposed in [9], [10] requires more data than its conventional (i.e., single-user) counterpart.

Some attention has been focused recently on a different family of blind multiuser detectors, which do not require knowledge of the spreading codes and timings of the interfering users. These detectors are well suited for the DS-CDMA downlink (i.e., base station to mobile) as well as for the uplink, where all the users can be detected in parallel by using a bank of receivers with near-far resistance capabilities. The methods proposed in literature can be roughly classified as *indirect* and *direct* ones: the former perform a channel identification first, and then the estimated channel is used to jointly suppress ISI and MAI; the latter extract detector parameters by the received data, without performing any explicit channel identification. Several indirect methods, which exploit second-order statistical properties of the received vector data (i.e., the data correlation matrix), borrow concepts from subspace-based blind channel identification, first proposed in the context of single-user channel equalization [13]. Perhaps the first contribution in this area is the paper of Torlak and Xu [14], where a subspace-based approach for channel identification is proposed for asynchronous CDMA systems, under the assumption that the channel multipath spread T_m is much smaller than the symbol period T_s and, hence, ISI would span only a fraction of the symbol period. The obtained channel estimate is then plugged into a linear MMSE multiuser receiver to suppress the MAI. Similar approaches have been proposed in [8], [15], where in addition no restriction about the channel length is made; in particular, the solution of [8] is similar in spirit to that of [14], whereas the method of [15], unlike [8], [14], performs the subspace-based channel identification by resorting to linear prediction and does not necessarily require knowledge of the desired-user synchronization. Recent developments indicate that direct methods may lead to simpler ISI-resilient multiuser algorithms, which are more tolerant to channel estimation errors. Belonging to this class are the methods of [16] and [17]: in the former, array processing concepts are exploited to generalize the minimum-output energy (MOE) [5], [6] receiver to the multipath case, under the assumption that $T_m < T_s$ and the desired-user delay is known; in the latter, an array of sensors is employed at the receiver, and the subspace approach is used to directly estimate from data the weight vector of the MMSE receiver, without restrictions on the channel length and without requiring desired-user synchronization.

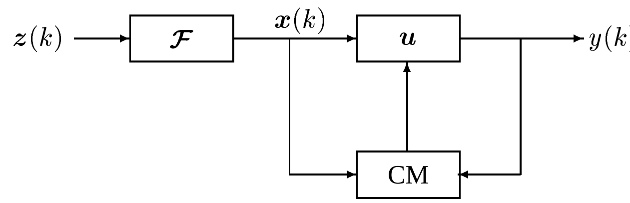


Fig. 1. The proposed two-stage blind multiuser receiver.

Unlike the previous techniques, which use second-order statistics (SOS) of the received data, the direct methods proposed in [18], [19] resort to higher-order statistics (HOS), which have been used for several years in single-user blind channel equalization since they require much milder identifiability conditions than SOS-based techniques. Since HOS-based functionals are multimodal, in a multiuser scenario they are prone to the interference capture problem, i.e., they might extract symbols of an undesired user. To overcome this drawback, the methods of [18], [19] resort to *constrained* optimization of fourth-order functionals, where the constraint is aimed at forcing the extraction of the desired user signal, and is implemented starting from the data correlation matrix (i.e., using SOS properties) and assuming only the knowledge of the desired-user code. It should be noted that these techniques are based on a clever integration of SOS (in the constraint implementation) and HOS (in the cost function) properties, and do not require desired-user synchronization. The performance results reported in [18], [19] show that these constrained HOS-based techniques significantly outperform the SOS-based ones, like those proposed in [8], [14], [17]. Despite their performance advantage, however, they are not suited to real-time implementation, since they are not entirely adaptive and, even worse, are characterized by a high computational complexity. In fact, the constraint implementation is based on singular value decomposition (SVD), whose complexity increases, as well known, cubically with the dimensions of the matrix involved, whereas optimization of the HOS cost function is made iteratively and in batch-mode. To overcome these limitations, we propose a blind direct technique for joint equalization and multiuser detection in high data-rate DS-CDMA systems, by extending some of the concepts originally proposed in [20], which have proven fruitful in the area of single-user channel equalization and interference suppression. In particular, the proposed receiver (see Fig. 1) consists of two stages: the former performs pre-filtering of the received signal, in order to mitigate MAI as well as ISI, and is based on a blind constrained optimization exploiting the SOS properties of the received signal and a suitable parameterization of the channel matrix, which allows one to impose a deterministic constraint enhancing the desired-user signal contribution; the latter exploits HOS properties to recover the desired signal, by resorting, similarly to [19], to the Godard cost function. The implementation of the first stage, which does not require any

SVD but only a matrix inversion, can easily be made adaptive by means of a recursion similar to the recursive least square (RLS) algorithm [7], whose complexity per iteration increases only quadratically with the dimension of the matrix involved. Moreover, the second stage can be adaptively implemented by the CMA, which is based on the stochastic gradient algorithm. It should be noted that the proposed method, similarly to [18], [19], does not need any assumption on the channel memory length and does not necessarily require desired-user synchronization.

The outline of the paper is the following: the system model is introduced in Section II, whereas Section III describes the proposed receiver, theoretically analyzes the disturbance (ISI-plus-MAI) suppression capabilities of the first stage and discusses the adaptive implementation of the overall receiver; finally, Section IV provides a performance comparison with existing methods, under ideal conditions, and reports extensive computer simulation results in a realistic operative environment, including a severe near-far and a nonstationary interference scenario.

II. System model

In the rest of the paper, we will use the following notations. Upper- and lower-case bold letters denote matrices and vectors, respectively; the superscripts $*$, $\dagger$, T and H denote conjugate, pseudo-inverse, transpose and conjugate transpose operations, respectively; $\mathbf{0}_n$, $\boldsymbol{O}_{n\times m}$ and $\boldsymbol{I}_n$ denote the n-column zero vector, the $n\times m$ zero and $n\times n$ identity matrices, respectively; $\|\cdot\|$, $\text{trace}(\cdot)$ and $\text{rank}(\cdot)$ denote the Frobenius norm, the trace and the rank of a matrix; $\text{null}(\boldsymbol{A})$, $\text{range}(\boldsymbol{A})$ and $\text{range}^{\perp}(\boldsymbol{A})$ denote for any $m\times n$ matrix $\boldsymbol{A}$ the null space of $\boldsymbol{A}$, the column space of $\boldsymbol{A}$ and its orthogonal complement in $\mathbb{C}^m$, respectively; the subscript c stands for continuous-time signal and $E[\cdot]$ denotes statistical averaging; $\lceil\cdot\rceil$ denotes ceiling-integer and $\text{rect}_{T_c}(t)$ is the unit-amplitude rectangular pulse of duration T_c, i.e., $\text{rect}_{T_c}(t)=1$, for $0\le t\le T_c$, and zero elsewhere.

Let us consider a short-code DS-CDMA network with J users, symbol period T_s, chip interval T_c and spreading factor N. The complex envelope of the received signal can be written as

$$r_c(t)=\sum_{j=1}^{J}\sum_{q=-\infty}^{+\infty} s_j(q)\, g_{c,j}(t-qT_c-\tau_j)+w_c(t)\,, \quad (1)$$

where $g_{c,j}(t)$ and $\tau_j\in[0,T_s)$ are the *composite* channel impulse response (including transmitting filter, physical channel, and receiving filter) and the transmission delay of the jth user, respectively; $w_c(t)$ is the additive noise at the output of the receiving filter; and, finally, the discrete-time transmitted signal of the jth user

$$s_j(q)=\sum_{i=-\infty}^{+\infty} b_j(i)\, c_j(q-iN) \quad (2)$$

is expressed in terms of the multirate convolution between the complex information symbol sequence $\{b_j(i)\}_{i\in\mathbb{Z}}$ and the spreading code $\{c_j(i)\}_{i\in\mathbb{Z}}$, which satisfies the condition $c_j(i)\equiv 0$ for $i\notin\{0,1,\ldots,N-1\}$. By substituting (2) in (1), the received signal can be equivalently written as

$$r_c(t)=\sum_{j=1}^{J}\sum_{i=-\infty}^{+\infty} b_j(i)\, h_{c,j}(t-iT_s-\tau_j)+w_c(t)\,, \quad (3)$$

where we have defined

$$h_{c,j}(t)\triangleq\sum_{q=0}^{N-1} c_j(q)\, g_{c,j}(t-qT_c)\,. \quad (4)$$

Similarly to [14]–[17] and [18], [19], we model the physical channel as a linear time-invariant discrete multipath channel: hence, denoting with $\psi_c(t)$ the convolution between the impulse responses of the transmitting and receiving filter, the composite channel of the jth user is

$$g_{c,j}(t)=\sum_{m=1}^{M_j}\alpha_{m,j}\,\psi_c(t-\tau_{m,j})\,, \quad (5)$$

where M_j is the number of paths of the physical channel, and $\alpha_{m,j}$ and $\tau_{m,j}$ are the complex gain and the propagation time of the mth path, respectively. Without loss of generality, we will assume that $\tau_{1,j}=0$ and $0<\tau_{2,j}<\tau_{3,j}<\ldots<\tau_{M_j,j}$, for each user: hence, the transmission delay τ_j in (1) represents also the delay of the first path of the jth user, relative to the beginning of each symbol interval. Such a delay τ_j is not restricted to be an integer multiple of T_c, i.e., it can be decomposed as $\tau_j=d_jT_c+\beta_j$, with $d_j\in\{0,1,\ldots,N-1\}$ and $\beta_j\in[0,T_c)$. By sampling the received signal (3) at the chip rate, we obtain the discrete-time signal

$$\begin{aligned} r(n) &\triangleq r_c(nT_c)\\ &=\sum_{j=1}^{J}\sum_{i=-\infty}^{+\infty} b_j(i)\, h_j(n-iN-d_j)+w(n)\,, \end{aligned} \quad (6)$$

where $w(n)\triangleq w_c(nT_c)$ and

$$h_j(n)\triangleq\sum_{q=0}^{N-1} c_j(q)\, g_j(n-q)\,, \quad (7)$$

with the discrete-time channel[1] $g_j(n)\triangleq g_{c,j}(nT_c-\beta_j)$ accounting also for the fractional part β_j of the transmission delay τ_j. The following customary assumptions will be considered throughout the paper: A1) the information symbols $b_j(i)$ are mutually independent zero-mean and independent identically-distributed (iid) sequences, with equal variance σ_b^2; A2) the additive noise $w_c(t)$ is a zero-mean wide-sense stationary (WSS) complex process, which is independent of the sequences $b_j(i)$, for $j=1,2,\ldots,J$.

Under assumptions A1 and A2, the discrete-time received signal $r(n)$ turns out to be *second-order cyclostationary* with period equal to the processing gain N, that is, its statistical autocorrelation function $R_{rr}(n,m)\triangleq E[r(n)\,r^*(n-m)]$ is

[1]For the sake of notation simplicity, we do not explicitly indicate the dependence of $g_j(n)$ on β_j.

periodic in n with period N. Therefore, a suitable representation for $r(n)$ is in terms of its *polyphase decomposition* with respect to N:

$$r^{(\ell)}(k) \triangleq r(kN+\ell) = \sum_{j=1}^{J} \sum_{i=-\infty}^{+\infty} b_j(i)\, h_j^{(\ell)}(k-i) + w^{(\ell)}(k)\,, \tag{8}$$

for $\ell = 0, 1, \ldots, N-1$, where we have defined the *phases* $h_j^{(\ell)}(k) \triangleq h_j(kN+\ell-d_j)$ and $w^{(\ell)}(k) \triangleq w(kN+\ell)$. Observe that, under assumptions A1 and A2, the polyphase components $r^{(\ell)}(k)$ turn out to be jointly WSS. Finally, a more compact MIMO vector model can be obtained by collecting the N different polyphase components of $r(n)$ in the N-column vector $\boldsymbol{r}(k) \triangleq [\, r^{(0)}(k),\, r^{(1)}(k),\, \ldots,\, r^{(N-1)}(k)\,]^T$, obtaining thus

$$\boldsymbol{r}(k) = \sum_{j=1}^{J} \sum_{i=-\infty}^{+\infty} b_j(i)\, \boldsymbol{h}_j(k-i) + \boldsymbol{w}(k)\,, \tag{9}$$

where we have introduced the N-column vectors $\boldsymbol{h}_j(k) \triangleq [\, h_j^{(0)}(k),\, h_j^{(1)}(k),\, \ldots,\, h_j^{(N-1)}(k)\,]^T$ and $\boldsymbol{w}(k) \triangleq [\, w^{(0)}(k),\, w^{(1)}(k),\, \ldots,\, w^{(N-1)}(k)\,]^T$. Following common practice in single-user equalization literature, we assume that the composite channel $g_{c,j}(t)$ of the jth user spans $L_j \in \mathbb{N} - \{0\}$ symbol intervals, i.e., $g_{c,j}(t) \equiv 0$ for $t \notin [\,0, L_j T_s)$, which implies that the (discrete-time) channel $g_j(n)$ is a causal finite-impulse response (FIR) filter of order less than or equal to $NL_j + 1$ chips, i.e., $g_j(n) \equiv 0$ for $n \notin \{\,0, 1, \ldots, NL_j\,\}$. Under this assumption, it results from (7) that $h_j(n) \equiv 0$ for $n \notin \{\,0, 1, \ldots, NL_j + N - 1\,\}$, which in its turn implies that $\boldsymbol{h}_j(k) \equiv \boldsymbol{0}_N$, for $k \notin \{0, 1, \ldots, L_j + 1\}$. Hence, the MIMO model (9) can be ultimately written as

$$\boldsymbol{r}(k) = \sum_{j=1}^{J} \sum_{i=0}^{L_j+1} \boldsymbol{h}_j(i)\, b_j(k-i) + \boldsymbol{w}(k)\,. \tag{10}$$

Note that, similarly to [15], [18], [19], the memory $L_j T_s$ of the composite channel can be greater than the symbol period T_s and, consequentially, the length of the FIR channel $g_j(n)$ might exceed the processing gain N; this typically happens in high-speed (data rates of 1 to 10 Mbps) DS-CDMA wireless local area networks, which are oriented to multimedia communications [21].

III. The Proposed Blind Multiuser Receiver

Without loss of generality, our aim is to demodulate the first user ($j = 1$), which will be referred to as the *desired user*. In order to compensate for the multipath-induced ISI, the proposed detection algorithm estimates the symbol $b_1(k-d)$, with d denoting a suitable *equalization delay*, on the basis of the (NL_e)-column vector $\boldsymbol{z}(k) \triangleq [\,\boldsymbol{r}^T(k), \boldsymbol{r}^T(k-1), \ldots, \boldsymbol{r}^T(k-L_e+1)\,]^T$, which collects L_e consecutive symbols (i.e., NL_e chips). A linear receiver for detecting the desired user's information symbol $b_1(k-d)$ has the form $y(k) = \boldsymbol{f}^H \boldsymbol{z}(k)$, where $\boldsymbol{f}$ is the (NL_e)-column vector collecting the weights of the receiver; then, the receiver output $y(k)$ is quantized to the nearest (in Euclidean distance) information symbol to form the estimate $\widehat{b}_1(k-d)$. Accounting for (10), the vector $\boldsymbol{z}(k)$ can be conveniently written as

$$\boldsymbol{z}(k) = \sum_{j=1}^{J} \boldsymbol{\mathcal{H}}_j\, \boldsymbol{b}_j(k) + \boldsymbol{v}(k)\,, \tag{11}$$

where we have defined the $(NL_e) \times (L_j + L_e + 1)$ block Toeplitz matrix

$$\boldsymbol{\mathcal{H}}_j \triangleq \begin{bmatrix} \boldsymbol{h}_j(0) & \ldots & \boldsymbol{h}_j(L_j+1) & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{h}_j(0) & \ldots & \boldsymbol{h}_j(L_j+1) & \boldsymbol{0} \\ \vdots & & & & \vdots \\ \vdots & & & & \vdots \\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{h}_j(0) & \ldots & \boldsymbol{h}_j(L_j+1) \end{bmatrix}, \tag{12}$$

with $\boldsymbol{b}_j(k) \triangleq [\, b_j(k),\, b_j(k-1),\, \ldots,\, b_j(k-L_j-L_e)\,]^T$ and $\boldsymbol{v}(k) \triangleq [\,\boldsymbol{w}^T(k),\, \boldsymbol{w}^T(k-1),\, \ldots,\, \boldsymbol{w}^T(k-L_e+1)\,]^T$. For any $d \in \{0, 1, \ldots, L_1 + L_e\}$, denoting with $\widetilde{\boldsymbol{b}}_1(k)$ the vector including all elements in $\boldsymbol{b}_1(k)$ except for the $(d+1)$th element $b_1(k-d)$, and with $\widetilde{\boldsymbol{\mathcal{H}}}_1$ the matrix including all the columns in $\boldsymbol{\mathcal{H}}_1$ except for the $(d+1)$th column $\boldsymbol{h}_{1,d}$, equation (11) can be re-expressed as

$$\boldsymbol{z}(k) = \boldsymbol{h}_{1,d}\; b_1(k-d) + \boldsymbol{\jmath}(k)\,, \tag{13}$$

where $\boldsymbol{\jmath}(k) \triangleq \boldsymbol{\mathcal{H}}\, \boldsymbol{b}(k) + \boldsymbol{v}(k)$ represents the overall disturbance (ISI-plus-MAI and noise) vector, with $\boldsymbol{\mathcal{H}} \triangleq [\,\widetilde{\boldsymbol{\mathcal{H}}}_1,\, \boldsymbol{\mathcal{H}}_2,\, \ldots,\, \boldsymbol{\mathcal{H}}_J\,]$ denoting the *composite disturbance matrix* and $\boldsymbol{b}(k) \triangleq [\,\widetilde{\boldsymbol{b}}_1^T(k),\, \boldsymbol{b}_2^T(k),\, \ldots,\, \boldsymbol{b}_J^T(k)]^T$ collecting all the $P \triangleq \sum_{j=1}^{J} L_j + J(L_e+1) - 1$ undesired symbols. Finally, accounting for (13), the linear receiver output can be rewritten explicitly as

$$y(k) = \boldsymbol{f}^H\, \boldsymbol{h}_{1,d}\; b_1(k-d) + \boldsymbol{f}^H \boldsymbol{\jmath}(k)\,. \tag{14}$$

The *composite signature* $\boldsymbol{h}_{1,d}$ depends on the user's spreading code $\{c_1(i)\}_{i \in \mathbb{Z}}$ and its associated transmission delay τ_1 and channel impulse response $\{g_1(n)\}_{n=0}^{NL_1}$. Thus, since in practice the channel and the transmission delay are unknown, the composite signature $\boldsymbol{h}_{1,d}$ is *unknown* at the receiver side.

In order to design a blind receiver which is joint multipath-resilient and near-far resistant, without requiring any training sequences, since the desired symbol constellation typically possesses the constant modulus property (e.g., BPSK or QPSK), the weight vector $\boldsymbol{f}$ can be chosen so as to minimize the Godard or constant modulus (CM) cost function, which is a well-known HOS-based criterion adopted in single-user blind equalization [22]. However, in the considered multiuser scenario, minimization of the CM cost function does not necessarily assure extraction of the desired symbol $b_1(k-d)$, but can instead lead to extraction of any of the symbols of the interfering users [23]. This phenomenon, known as the *capture effect*, is a direct consequence of the fact that, in the absence of noise, the CM cost function exhibits multiple global minima in the receiver parameter space, each corresponding to a particular symbol of a particular user. Sufficient conditions on the signal-to-interference-plus-noise ratio (SINR) assuring extraction of the desired symbol $b_1(k-d)$ have been singled out in [24]. Since in practice the CM cost function is minimized by resorting to gradient-descent algorithms, such conditions are employed in [24] to devise suitable initialization schemes avoiding the capture effect; however, they require some *a priori* knowledge about the desired-user channel and, moreover, are difficult to employ in nonstationary environments, where interferers may suddenly appear or disappear. To overcome these drawbacks, we factorize the overall weight vector as $\boldsymbol{f} = \boldsymbol{\mathcal{F}}\,\boldsymbol{u}$, where $\boldsymbol{\mathcal{F}}$ is a $NL_e \times K_1$ filtering matrix (with K_1 depending on the adopted optimization criterion, see Section III.A for details) that is designed blindly to mitigate the disturbance contribution (ISI-plus-MAI and noise) in order to avoid the capture effect, whereas the K_1-column weight vector $\boldsymbol{u}$ is determined according to the CM criterion. Such a factorization leads naturally to the two-stage architecture depicted in Fig. 1.

A. Design of the first stage

The output of the first stage $\boldsymbol{x}(k)$ is a linear transformation of the input vector $\boldsymbol{z}(k)$, that is,

$$\boldsymbol{x}(k) = \boldsymbol{\mathcal{F}}^H\, \boldsymbol{z}(k)\ . \qquad (15)$$

By substituting (13) in (15), we obtain a simple vector model for the output of the first stage

$$\boldsymbol{x}(k) = \boldsymbol{\mathcal{F}}^H\, \boldsymbol{h}_{1,d}\; b_1(k-d) + \boldsymbol{\mathcal{F}}^H\, \boldsymbol{\jmath}(k)\ . \qquad (16)$$

To solve for $\boldsymbol{\mathcal{F}}$ in a blind manner, we exploit the structure of the unknown composite signature $\boldsymbol{h}_{1,d}$. To this end, let us define the following $N(L_1+1)$-column vector

$$\boldsymbol{g}_1 = [\,\underbrace{0,\,\ldots,\,0}_{d_1}\ \underbrace{g_1(0),\,\ldots,\,g_1(NL_1)}_{NL_1+1}\ \underbrace{0,\,\ldots,\,0}_{(N-d_1-1)}\,]^T\ , \qquad (17)$$

which collects all the coefficients of the FIR channel $g_1(n)$ and accounts also for the unknown delay d_1. In the sequel, to globally characterize the columns $\boldsymbol{h}_{1,d}$ of the channel matrix $\boldsymbol{\mathcal{H}}_1$, we will assume that the length L_e of the observation interval is chosen not smaller than the memory of $\boldsymbol{h}_1(k)$, i.e., $L_e \geq L_1 + 2$. It is important to remark that this assumption *only* requires an upper bound $(L_1)_{\max}$ (rather than the exact knowledge) of the desired-channel order L_1. This is a reasonable assumption in practice since: (i) in general, depending on the transmitted signal parameters (carrier frequency and bandwidth) and application (indoor or outdoor), the maximum channel multipath spread is known; (ii) in high data-rate CDMA networks the multipath energy typically spans only a few symbol periods [21], i.e., L_1 is in the range $\{1, 2\}$. It is shown in Appendix I that, under the assumption $L_e \geq L_1 + 2$, the vector $\boldsymbol{h}_{1,d}$ can be linearly parameterized, over the entire range of the equalization delay values $d \in \{0, 1, \ldots, L_1 + L_e\}$, as $\boldsymbol{h}_{1,d} = \boldsymbol{\mathcal{D}}_{1,d}\, \boldsymbol{g}_{1,d}$, where $\boldsymbol{\mathcal{D}}_{1,d}$ is a full column rank *known* parameterization matrix, which depends on the code of the desired user, and the vector $\boldsymbol{g}_{1,d}$ is equal to $\boldsymbol{g}_1$, for $d \in \{L_1, L_1+1, \ldots, L_e-1\}$ (referred to as parameterization 2 in Appendix I), whereas it collects only some elements of $\boldsymbol{g}_1$, for $d \in \{0, 1, \ldots, L_1-1\}$ or $d \in \{L_e, L_e+1, \ldots, L_e+L_1\}$ (referred to as parameterization 1 and 3, respectively, in Appendix I). It should be remarked that d is a design parameter; hence, since parameterization 1 and 3 do not allow exploitation of the entire desired-signal channel energy and might thus lead to a significant SINR degradation at the output of the first stage, we will restrict our attention only to parameterization 2, that is, we consider

$$\boldsymbol{h}_{1,d} = \boldsymbol{\mathcal{D}}_{1,d}\, \boldsymbol{g}_1\ , \qquad (18)$$

for $d \in \{L_1, L_1+1, \ldots, L_e-1\}$, where $\boldsymbol{\mathcal{D}}_{1,d}$ has dimensions $(NL_e) \times N(L_1+1)$ (see Appendix I for its detailed expression). Observe that some elements of the unknown vector $\boldsymbol{g}_1$ might be zero depending on the value of $d_1 \in \{0, 1, \ldots, N-1\}$, and hence parameterization (18) can be explicitly rewritten as

$$\boldsymbol{h}_{1,d} = \boldsymbol{\mathcal{D}}_{1,d}\, \boldsymbol{\Upsilon}_{d_1}\, \widetilde{\boldsymbol{g}}_1\ , \qquad (19)$$

for $d \in \{L_1, L_1+1, \ldots, L_e-1\}$, where $\boldsymbol{\Upsilon}_{d_1} \triangleq [\boldsymbol{O}^T_{d_1 \times (NL_1+1)}, \boldsymbol{I}^T_{(NL_1+1)}, \boldsymbol{O}^T_{(N-d_1-1)\times(NL_1+1)}]^T$ and $\widetilde{\boldsymbol{g}}_1 \triangleq [g_1(0), g_1(1), \ldots, g_1(NL_1)]^T$. Substitution of (19) in (16) yields

$$\boldsymbol{x}(k) = \boldsymbol{\mathcal{F}}^H\, \boldsymbol{\mathcal{D}}_{1,d}\, \boldsymbol{\Upsilon}_{d_1}\; \widetilde{\boldsymbol{g}}_1\; b_1(k-d) + \boldsymbol{\mathcal{F}}^H\, \boldsymbol{\jmath}(k)\ , \qquad (20)$$

from which it results that, if the delay d_1 is known or estimated, by constraining the filtering matrix to satisfy $\boldsymbol{\mathcal{F}}^H\, \boldsymbol{\mathcal{D}}_{1,d}\, \boldsymbol{\Upsilon}_{d_1} = \boldsymbol{I}_{(NL_1+1)}$, the contribution of the desired symbol $b_1(k-d)$ is passed to the output of the first stage with response $\widetilde{\boldsymbol{g}}_1$. Thus, a sensible criterion is to choose $\boldsymbol{\mathcal{F}}$ so as to minimize the output power $E[\|\boldsymbol{x}(k)\|^2]$ of the first stage subject to the previous constraint, which allows one to preserve the desired symbol while minimizing the contribution of $\boldsymbol{\jmath}(k)$. On the other hand, if the delay d_1 is unknown, the matrix $\boldsymbol{\Upsilon}_{d_1}$ is also unknown and, thus, the filtering matrix $\boldsymbol{\mathcal{F}}$ can be chosen so as to minimize the output power of the first stage subject to the constraint $\boldsymbol{\mathcal{F}}^H\, \boldsymbol{\mathcal{D}}_{1,d} = \boldsymbol{I}_{N(L_1+1)}$. Note that, in the latter case, it is not required that the proposed receiver is synchronized to the desired transmission; in this case, the contribution of the desired symbol $b_1(k-d)$ is passed to the output of the first stage with response $\boldsymbol{g}_1$ instead of $\widetilde{\boldsymbol{g}}_1$. For the sake of synthesis, we will unify the treatment for the cases of known/unknown delay, by considering in the sequel the following constrained optimization problem:

$$\min_{\boldsymbol{\mathcal{F}}} E[\|\boldsymbol{x}(k)\|^2] \quad \text{subject to} \quad \boldsymbol{\mathcal{F}}^H\, \boldsymbol{\mathcal{Q}}_{1,d} = \boldsymbol{I}_{K_1}\ , \qquad (21)$$

where the $(NL_e) \times K_1$ *composite* full column rank parameterization matrix $\boldsymbol{\mathcal{Q}}_{1,d}$ is given by $\boldsymbol{\mathcal{Q}}_{1,d} = \boldsymbol{\mathcal{D}}_{1,d}\, \boldsymbol{\Upsilon}_{d_1}$, with $K_1 = NL_1 + 1$, in the case of *known* transmission delay d_1, whereas it is equal to $\boldsymbol{\mathcal{D}}_{1,d}$, with $K_1 = N(L_1+1)$, when the delay d_1 is *unknown*. As we will discuss later more in depth, since in the former case K_1 is smaller, the filtering matrix is designed starting from a constrained optimization problem with fewer constraints than in the latter one, thus the knowledge of the transmission delay is obviously expected to improve the disturbance suppression capabilities of the first stage. From a mathematical point of view, it should be observed that, since the constraint matrix $\boldsymbol{\mathcal{Q}}_{1,d}$ is full column rank, the system $\boldsymbol{\mathcal{F}}^H\, \boldsymbol{\mathcal{Q}}_{1,d} = \boldsymbol{I}_{K_1}$ admits an infinite number of solutions; thus, the optimization criterion (21) can be reinterpreted as follows: among the infinite solutions of the system $\boldsymbol{\mathcal{F}}^H \boldsymbol{\mathcal{Q}}_{1,d} = \boldsymbol{I}_{K_1}$, choose the one that minimizes the output power $E[\|\boldsymbol{x}(k)\|^2]$. In this respect, the synthesis of the first stage can be viewed as a generalization of the MOE approach to the multidimensional case . However, unlike [5], [6], we observe that the composite signature $\boldsymbol{h}_{1,d}$ is partially known and has quite a rich structure; by exploiting this structure, we impose multiple linear equality constraints which *completely* preserve the desired symbol $b_1(k-d)$ in *any* operative situation and *regardless* the (possibly) unknown transmission delay τ_1 and the unknown multipath channel $\{g_1(n)\}_{n=0}^{NL_1}$.

Accounting for (15) and using the trace properties, the constrained optimization problem (21) can be reformulated as follows

$$\min_{\boldsymbol{\mathcal{F}}} \operatorname{trace}[\boldsymbol{\mathcal{F}}^H \boldsymbol{R}_{\boldsymbol{zz}} \boldsymbol{\mathcal{F}}] \quad \text{subject to} \quad \boldsymbol{\mathcal{F}}^H \boldsymbol{\mathcal{Q}}_{1,d} = \boldsymbol{I}_{K_1} , \tag{22}$$

where $\boldsymbol{R}_{\boldsymbol{zz}} \triangleq E[\boldsymbol{z}(k)\, \boldsymbol{z}^H(k)]$ is the $NL_e \times NL_e$ statistical correlation matrix of $\boldsymbol{z}(k)$. The closed-form solution of the optimization problem (22), derived in [20], is

$$\boldsymbol{\mathcal{F}}_{\text{opt}} = \boldsymbol{R}_{\boldsymbol{zz}}^{-1} \boldsymbol{\mathcal{Q}}_{1,d} \,(\, \boldsymbol{\mathcal{Q}}_{1,d}^H \, \boldsymbol{R}_{\boldsymbol{zz}}^{-1} \, \boldsymbol{\mathcal{Q}}_{1,d}\,)^{-1} \, . \tag{23}$$

When batch algorithms are employed to estimate $\boldsymbol{\mathcal{F}}_{\text{opt}}$, the computational complexity of the first stage, which is dominated by the two (distinct) matrix inversions in (23), is of order $\mathcal{O}(N^3 L_e^3) + \mathcal{O}(K_1^3) \sim \mathcal{O}(N^3 L_e^3)$ (recall that $L_e \geq L_1 + 2$). Such a complexity can be reduced by reformulating the constrained optimization problem (22) as an unconstrained one, by resorting to an extension of the generalized sidelobe canceller decomposition [25], which was originally proposed in the array processing context and subsequently applied in [5] to derive the canonical representation of a linear multiuser detector. The starting point is to recognize that, since each column of the unknown matrix $\boldsymbol{\mathcal{F}}$ can be canonically decomposed in two orthogonal components belonging to the K_1-dimensional subspace $\operatorname{range}(\boldsymbol{\mathcal{Q}}_{1,d})$ and to the $(N L_e - K_1)$-dimensional subspace $\operatorname{range}^{\perp}(\boldsymbol{\mathcal{Q}}_{1,d})$, respectively, the entire matrix $\boldsymbol{\mathcal{F}}$ can be conveniently decomposed as $\boldsymbol{\mathcal{F}} = \boldsymbol{\mathcal{F}}^{(0)} - \boldsymbol{\mathcal{F}}^{(1)}$, where the columns of $\boldsymbol{\mathcal{F}}^{(0)}$ span the subspace $\operatorname{range}(\boldsymbol{\mathcal{Q}}_{1,d})$, i.e., $\operatorname{range}(\boldsymbol{\mathcal{F}}^{(0)}) = \operatorname{range}(\boldsymbol{\mathcal{Q}}_{1,d})$, whereas the column space of $\boldsymbol{\mathcal{F}}^{(1)}$ lies in the subspace $\operatorname{range}^{\perp}(\boldsymbol{\mathcal{Q}}_{1,d})$, i.e., $\operatorname{range}(\boldsymbol{\mathcal{F}}^{(1)}) \subseteq \operatorname{range}^{\perp}(\boldsymbol{\mathcal{Q}}_{1,d})$; moreover, since $\operatorname{range}^{\perp}(\boldsymbol{\mathcal{Q}}_{1,d})$ is an $(N L_e - K_1)$-dimensional subspace, the matrix $\boldsymbol{\mathcal{F}}^{(1)}$ can be written as $\boldsymbol{\mathcal{F}}^{(1)} = \boldsymbol{\Pi}_{1,d}\, \boldsymbol{\mathcal{F}}^{(a)}$, where the columns of the $(NL_e) \times (NL_e - K_1)$ matrix $\boldsymbol{\Pi}_{1,d}$ are an orthonormal basis for the subspace $\operatorname{range}^{\perp}(\boldsymbol{\mathcal{Q}}_{1,d})$, i.e., the matrix $\boldsymbol{\Pi}_{1,d}$ obeys the relation $\boldsymbol{\Pi}_{1,d}^H\, \boldsymbol{\mathcal{Q}}_{1,d} = \boldsymbol{O}_{(NL_e - K_1)\times K_1}$. By substitution, we can finally express $\boldsymbol{\mathcal{F}}$ in the following form:

$$\boldsymbol{\mathcal{F}} = \boldsymbol{\mathcal{F}}^{(0)} - \boldsymbol{\Pi}_{1,d}\, \boldsymbol{\mathcal{F}}^{(a)} \, . \tag{24}$$

It should be noted that the matrix $\boldsymbol{\Pi}_{1,d}$ can be obtained from the known parameterization matrix $\boldsymbol{\mathcal{Q}}_{1,d}$ by using any one of several orthogonalization procedures, such as the SVD or QR decomposition [26] and hence we assume, without loss of generality, that $\boldsymbol{\Pi}_{1,d}$ is unitary, i.e., $\boldsymbol{\Pi}_{1,d}^H\, \boldsymbol{\Pi}_{1,d} = \boldsymbol{I}_{(NL_e - K_1)}$. The advantage of decomposition (24) is that $\boldsymbol{\mathcal{F}}$ satisfies the constraint in (22) *for any choice* of the matrix $\boldsymbol{\mathcal{F}}^{(a)}$, provided that $[\boldsymbol{\mathcal{F}}^{(0)}]^H \boldsymbol{\mathcal{Q}}_{1,d} = \boldsymbol{I}_{K_1}$; recalling that $\operatorname{range}(\boldsymbol{\mathcal{F}}^{(0)}) = \operatorname{range}(\boldsymbol{\mathcal{Q}}_{1,d})$, the latter condition is uniquely satisfied by

$$\boldsymbol{\mathcal{F}}_{\text{opt}}^{(0)} = \boldsymbol{\mathcal{Q}}_{1,d}\, (\boldsymbol{\mathcal{Q}}_{1,d}^H\, \boldsymbol{\mathcal{Q}}_{1,d})^{-1} \, , \tag{25}$$

where the matrix $(\boldsymbol{\mathcal{Q}}_{1,d}^H\, \boldsymbol{\mathcal{Q}}_{1,d})$ turns out to be invertible due to the fact that $\boldsymbol{\mathcal{Q}}_{1,d}$ has full column rank; the matrix $\boldsymbol{\mathcal{F}}_{\text{opt}}^{(0)}$ may be recognized as the pseudo inverse [27] of $\boldsymbol{\mathcal{Q}}_{1,d}^H$. At this point, by substituting (24) in (22) and accounting for (25), the constrained minimization in (22) is turned into an unconstrained one

$$\min_{\boldsymbol{\mathcal{F}}^{(a)}} \operatorname{trace}\{[\, \boldsymbol{\mathcal{F}}_{\text{opt}}^{(0)} - \boldsymbol{\Pi}_{1,d}\, \boldsymbol{\mathcal{F}}^{(a)}\,]^H\, \boldsymbol{R}_{\boldsymbol{zz}}\, [\, \boldsymbol{\mathcal{F}}_{\text{opt}}^{(0)} - \boldsymbol{\Pi}_{1,d}\, \boldsymbol{\mathcal{F}}^{(a)}\,]\} \, , \tag{26}$$

where only $\boldsymbol{\mathcal{F}}^{(a)}$ must be determined. With some manipulations and using the properties of the trace operator, the solution of (26) can be expressed as (see also [5])

$$\boldsymbol{\mathcal{F}}_{\text{opt}}^{(a)} = (\, \boldsymbol{\Pi}_{1,d}^H\, \boldsymbol{R}_{\boldsymbol{zz}}\, \boldsymbol{\Pi}_{1,d}\,)^{-1}\, \boldsymbol{\Pi}_{1,d}^H\, \boldsymbol{R}_{\boldsymbol{zz}}\, \boldsymbol{\mathcal{F}}_{\text{opt}}^{(0)} \, . \tag{27}$$

One implementation advantage of this alternative formulation is that the optimal filtering matrix in (23) has been decomposed into the *non-data-dependent* component $\boldsymbol{\mathcal{F}}_{\text{opt}}^{(0)}$, which can be pre-computed off-line, plus the $(NL_e - K_1) \times K_1$ *data-dependent* component $\boldsymbol{\mathcal{F}}_{\text{opt}}^{(a)}$, which must be estimated from data; hence, if one resorts to batch algorithms, the computational complexity of the first stage, compared with (23), is now dominated by *one* matrix inversion in (27), which moreover is of reduced complexity $\mathcal{O}[(NL_e - K_1)^3]$. We will show in Section III.D that the matrix $\boldsymbol{\mathcal{F}}_{\text{opt}}^{(a)}$ in (27) can also be estimated by means of a simple and effective recursion, similar to the well-known RLS algorithm, with a complexity per iteration of order only $\mathcal{O}[(NL_e - K_1)^2]$.

B. Design of the second stage

In order to obtain an accurate estimate of the desired symbol $b_1(k-d)$ from the output $\boldsymbol{x}(k)$ of the first stage, one can resort to minimization of the CM cost function

$$J_{\text{CM}}(\boldsymbol{u}) \triangleq E[(\gamma - |\boldsymbol{u}^H\, \boldsymbol{x}(k)|^2)^2] \, , \tag{28}$$

where $\boldsymbol{u}$ is the K_1-column weight vector of the second stage and $\gamma \triangleq E\{|b_1(k)|^4\}/\sigma_b^2$ is the dispersion coefficient of the desired symbol sequence $b_1(k)$. A closed-form expression for the solution of (28) is not available and, thus, the gradient descent method [7] estimates the optimal weight vector $\boldsymbol{u}$ by resorting to the simple recursive relation

$$\boldsymbol{u}(k+1) = \boldsymbol{u}(k) - \mu\, \boldsymbol{\nabla} J_{\text{CM}}[\boldsymbol{u}(k)] \, , \tag{29}$$

where $\boldsymbol{\nabla}[\cdot]$ represents the *gradient vector* operator [7] and μ is the step-size of the algorithm. The gradient vector $\boldsymbol{\nabla} J_{\text{CM}}[\boldsymbol{u}(k)]$ can be expressed in terms of second- and fourth-order statistics of vector $\boldsymbol{x}(k)$, and can be estimated from the received data either in batch-mode (see [23] for details) or in adaptive-mode, by resorting to the stochastic gradient descent (SGD) algorithm [7], as discussed in Section III.D. Finally, it should be observed that, as most blind techniques (see, e.g., [8]–[10], [14]–[19] and [28]), the desired symbol $b_1(k-d)$ is recovered up to an arbitrary complex factor, which introduces symbol constellation rotation/scaling. In practice, the phase ambiguity (i.e., rotation) can be blindly solved by resorting to differential coding/encoding modulation techniques [1], whereas the amplitude ambiguity (i.e., scaling) can be blindly compensated by employing an automatic gain control device at the receiver.

C. Interference cancellation analysis and discussion

In this section we analyze the disturbance suppression capabilities of the first stage and discuss some interesting properties of the overall two-stage receiver. To this aim, let us consider the canonical decomposition of $\boldsymbol{\mathcal{F}}_{\text{opt}}$ [see equations (24), (25) and (27)]:

$$\boldsymbol{\mathcal{F}}_{\text{opt}} = \boldsymbol{\mathcal{F}}_{\text{opt}}^{(0)} - \boldsymbol{\Pi}_{1,d}\,\boldsymbol{\mathcal{F}}_{\text{opt}}^{(a)} = \left[\boldsymbol{I}_{NL_e} - \boldsymbol{\Pi}_{1,d}\,(\boldsymbol{\Pi}_{1,d}^H\,\boldsymbol{R}_{\boldsymbol{zz}}\,\boldsymbol{\Pi}_{1,d})^{-1}\boldsymbol{\Pi}_{1,d}^H\,\boldsymbol{R}_{\boldsymbol{zz}}\right]\boldsymbol{\mathcal{F}}_{\text{opt}}^{(0)} \,. \tag{30}$$

In order to simplify the analysis, we assume hereinafter that the noise $w(n)$ is white with variance σ_w^2. Under this hypothesis, by invoking assumptions A1-A2 and accounting for (13) and (19), the correlation matrix $\boldsymbol{R}_{\boldsymbol{zz}}$ of the input vector $\boldsymbol{z}(k)$ can be written as

$$\boldsymbol{R}_{\boldsymbol{zz}} = \sigma_b^2\,\boldsymbol{\mathcal{Q}}_{1,d}\,\boldsymbol{\varrho}_1\,\boldsymbol{\varrho}_1^H\,\boldsymbol{\mathcal{Q}}_{1,d}^H + \sigma_b^2\,\boldsymbol{\mathcal{H}}\,\boldsymbol{\mathcal{H}}^H + \sigma_w^2\,\boldsymbol{I}_{NL_e}\,, \tag{31}$$

where, according to the definition of $\boldsymbol{\mathcal{Q}}_{1,d}$, the K_1-column vector $\boldsymbol{\varrho}_1$ is given by $\boldsymbol{\varrho}_1 = \boldsymbol{\Upsilon}_{d_1}\,\widetilde{\boldsymbol{g}}_1$ in case of unknown transmission delay d_1, whereas it is equal to $\widetilde{\boldsymbol{g}}_1$ when the delay d_1 is known. Substituting (31) in (30) and recalling that $[\boldsymbol{\mathcal{F}}_{\text{opt}}^{(0)}]^H\,\boldsymbol{\mathcal{Q}}_{1,d} = \boldsymbol{I}_{K_1}$ and $\boldsymbol{\Pi}_{1,d}^H\,\boldsymbol{\mathcal{Q}}_{1,d} = \boldsymbol{O}_{(NL_e-K_1)\times K_1}$, it can be easily seen that

$$\boldsymbol{\mathcal{F}}_{\text{opt}} = \Big[\boldsymbol{I}_{NL_e} - \sigma_b^2\,\boldsymbol{\Pi}_{1,d}\left(\sigma_b^2\,\boldsymbol{\Pi}_{1,d}^H\,\boldsymbol{\mathcal{H}}\,\boldsymbol{\mathcal{H}}^H\,\boldsymbol{\Pi}_{1,d} + \sigma_w^2\,\boldsymbol{I}_{NL_e-K_1}\right)^{-1}\boldsymbol{\Pi}_{1,d}^H\,\boldsymbol{\mathcal{H}}\,\boldsymbol{\mathcal{H}}^H\Big]\boldsymbol{\mathcal{F}}_{\text{opt}}^{(0)} \,. \tag{32}$$

Since the performance of high data-rate DS-CDMA networks is mainly limited by ISI and MAI, we derive the analytical expression of the matrix $\boldsymbol{\mathcal{F}}_{\text{opt}}$ in the high SNR region, i.e., as σ_w^2/σ_b^2 approaches to zero.

Proposition 1: In the high SNR region, the optimal filtering matrix $\boldsymbol{\mathcal{F}}_{\text{opt}}$ assumes the expression

$$\overline{\boldsymbol{\mathcal{F}}}_{\text{opt}} \triangleq \lim_{\sigma_w^2/\sigma_b^2\to 0}\boldsymbol{\mathcal{F}}_{\text{opt}} = \left[\boldsymbol{I}_{NL_e} - \boldsymbol{\Pi}_{1,d}\,(\boldsymbol{\mathcal{H}}^H\boldsymbol{\Pi}_{1,d})^{\dagger}\boldsymbol{\mathcal{H}}^H\right]\boldsymbol{\mathcal{F}}_{\text{opt}}^{(0)}. \tag{33}$$

Proof: See Appendix II. ■

Accounting for (13), (15) and (33) and using the pseudo inverse's properties [27], the noiseless output of the first stage can be written as

$$\overline{\boldsymbol{x}}(k) = \overline{\boldsymbol{\mathcal{F}}}_{\text{opt}}^H\,\boldsymbol{z}(k) = \boldsymbol{\varrho}_1\,b_1(k-d) + [\boldsymbol{\mathcal{F}}_{\text{opt}}^{(0)}]^H\,\boldsymbol{\mathcal{H}}\,\boldsymbol{\mathcal{X}}\,\boldsymbol{b}(k)\,, \tag{34}$$

where the $P \times P$ matrix $\boldsymbol{\mathcal{X}} \triangleq \boldsymbol{I}_P - (\boldsymbol{\Pi}_{1,d}^H\,\boldsymbol{\mathcal{H}})^{\dagger}(\boldsymbol{\Pi}_{1,d}^H\,\boldsymbol{\mathcal{H}})$ represents the *orthogonal projector* [27] onto the null space of $\boldsymbol{\Pi}_{1,d}^H\,\boldsymbol{\mathcal{H}}$ and, thus, $\text{range}(\boldsymbol{\mathcal{X}}) = \text{null}(\boldsymbol{\Pi}_{1,d}^H\,\boldsymbol{\mathcal{H}})$. From (34), it appears that *complete* cancellation of the disturbance is obtained if and only if (*iff*) the *zero-forcing condition* $[\boldsymbol{\mathcal{F}}_{\text{opt}}^{(0)}]^H\,\boldsymbol{\mathcal{H}}\,\boldsymbol{\mathcal{X}} = \boldsymbol{O}_{K_1\times P}$ holds. When this condition is satisfied, the proposed first stage behaves as a *blind zero-forcing* detector [8], whose synthesis does not require knowledge of the composite disturbance matrix $\boldsymbol{\mathcal{H}}$ and, unlike [8], does not involve any SVD of the correlation matrix $\boldsymbol{R}_{\boldsymbol{zz}}$, lending itself to a simple and effective adaptive implementation (see Section III.D). Accounting for (25) and recalling that $\boldsymbol{\mathcal{Q}}_{1,d}$ is a full column rank matrix, the previous condition can be equivalently rewritten as $\boldsymbol{\mathcal{Q}}_{1,d}^H\,\boldsymbol{\mathcal{H}}\,\boldsymbol{\mathcal{X}} = \boldsymbol{O}_{K_1\times P}$. This condition holds *iff* the columns of $\boldsymbol{\mathcal{X}}$ belong to the null space[2] of $\boldsymbol{\mathcal{Q}}_{1,d}^H\,\boldsymbol{\mathcal{H}}$; however, since the columns of $\boldsymbol{\mathcal{X}}$ span the subspace $\text{null}(\boldsymbol{\Pi}_{1,d}^H\,\boldsymbol{\mathcal{H}})$, the disturbance contribution is nullified under the equivalent condition $\text{null}(\boldsymbol{\Pi}_{1,d}^H\,\boldsymbol{\mathcal{H}}) \subseteq \text{null}(\boldsymbol{\mathcal{Q}}_{1,d}^H\,\boldsymbol{\mathcal{H}})$.

Lemma 1: Since the matrices $\boldsymbol{\Pi}_{1,d}$ and $\boldsymbol{\mathcal{Q}}_{1,d}$ are orthogonal and full-column rank, one has

$$\text{null}(\boldsymbol{\Pi}_{1,d}^H\,\boldsymbol{\mathcal{H}}) \cap \text{null}(\boldsymbol{\mathcal{Q}}_{1,d}^H\,\boldsymbol{\mathcal{H}}) = \text{null}(\boldsymbol{\mathcal{H}})\,. \tag{35}$$

Proof: See Appendix III. ■

Based on this lemma, we infer that the zero-forcing condition is satisfied *iff* $\text{null}(\boldsymbol{\Pi}_{1,d}^H\,\boldsymbol{\mathcal{H}}) = \text{null}(\boldsymbol{\mathcal{H}})$ which, in its turn, is satisfied *iff* the null space of $\boldsymbol{\Pi}_{1,d}^H$ and column space of $\boldsymbol{\mathcal{H}}$ intersect only trivially, that is $\text{null}(\boldsymbol{\Pi}_{1,d}^H) \cap \text{range}(\boldsymbol{\mathcal{H}}) = \{\boldsymbol{0}_{NL_e}\}$. At this point, observing that $\text{null}(\boldsymbol{\Pi}_{1,d}^H) = \text{range}^{\perp}(\boldsymbol{\Pi}_{1,d})$ and recalling that $\text{range}(\boldsymbol{\Pi}_{1,d}) = \text{range}^{\perp}(\boldsymbol{\mathcal{Q}}_{1,d})$, it is easily proved the following basic result.

Theorem 1: In the absence of noise, perfect disturbance cancellation can be achieved *iff*

$$\text{range}(\boldsymbol{\mathcal{Q}}_{1,d}) \cap \text{range}(\boldsymbol{\mathcal{H}}) = \{\boldsymbol{0}_{NL_e}\}\,, \tag{36}$$

that is, the subspaces $\text{range}(\boldsymbol{\mathcal{Q}}_{1,d})$ and $\text{range}(\boldsymbol{\mathcal{H}})$ are *nonoverlapping* or *disjoint*.

An interesting observation about the above theorem is that, for (36) to hold, it suffices that the columns of $\boldsymbol{\mathcal{Q}}_{1,d}$ and $\boldsymbol{\mathcal{H}}$ are linearly independent so that the composite matrix $[\boldsymbol{\mathcal{Q}}_{1,d}, \boldsymbol{\mathcal{H}}]$ has full column rank, which is the usual identifiability condition assumed in subspace-based CDMA channel identification techniques [8], [14], [16], [28]. Moreover, it is worthwhile to note that, according to [24], it is not required that the first stage perfectly suppresses the disturbance in order to avoid the capture effect but only that, at the output of the first stage [see (34)], the disturbance power $\mathcal{P}_{\text{dist}} = \sigma_b^2\,\text{trace}(\boldsymbol{\mathcal{X}}\,\boldsymbol{\mathcal{H}}^H\,\boldsymbol{\mathcal{F}}_{\text{opt}}\,\boldsymbol{\mathcal{F}}_{\text{opt}}^H\,\boldsymbol{\mathcal{H}}\,\boldsymbol{\mathcal{X}})$ is "sufficiently" lower than the desired-symbol power $\mathcal{P}_1 = \sigma_b^2\,\|\boldsymbol{\varrho}_1\|^2$. Thus, unlike [8], [14], [16], [28], the proposed two-stage receiver can also work when condition (36) is violated. In this respect, an interesting research issue is to delineate the "minimal" conditions under which the disturbance capture is avoided in the second stage.

Some remarks are now in order about the overall receiver:

Remark 1. Since, as it is apparent from (24), the first stage must only adapt the component of $\boldsymbol{\mathcal{F}}_{\text{opt}}$ that lies in $\text{range}^{\perp}(\boldsymbol{\mathcal{Q}}_{1,d})$, the number of *available degrees of freedom* for disturbance suppression is equal to the dimension $NL_e - K_1$ of $\text{range}^{\perp}(\boldsymbol{\mathcal{Q}}_{1,d})$. The condition $\text{null}(\boldsymbol{\Pi}_{1,d}^H\,\boldsymbol{\mathcal{H}}) = \text{null}(\boldsymbol{\mathcal{H}})$

[2]Since in practice $K_1 < P$, the $K_1 \times P$ matrix $\boldsymbol{\mathcal{Q}}_{1,d}^H\,\boldsymbol{\mathcal{H}}$ is "wide" and, thus, its rank is smaller than or equal to K_1, implying that $\boldsymbol{\mathcal{Q}}_{1,d}^H\,\boldsymbol{\mathcal{H}}$ possesses a nontrivial null space whose dimensionality is not inferior to $P - K_1$.

is also equivalent to rank($\mathbf{\Pi}_{1,d}^H \boldsymbol{\mathcal{H}}$) = rank($\boldsymbol{\mathcal{H}}$), which, by using straightforward rank inequalities, necessitates that the $(NL_e - K_1) \times P$ matrix $\mathbf{\Pi}_{1,d}^H \boldsymbol{\mathcal{H}}$ be "tall", i.e., the number $NL_e - K_1$ of available degrees of freedom must be greater than or equal to P. Since P depends on J, this necessary condition provides an upper bound for the number of users that can be zero-forced by the proposed first stage. For the sake of simplicity, let us assume that the channel orders of all users are equal, i.e., $L_j = L$ for $j \in \{1, 2, \ldots, J\}$, then the system capacity is limited by

$$J \leq \left\lceil \frac{NL_e - K_1 + 1}{L_e + L + 1} \right\rceil . \tag{37}$$

When the delay d_1 is known, the available degrees of freedom are $N(L_e - L_1) - 1$, whereas, if the delay d_1 is unknown, they are only $N(L_e - L_1 - 1)$, i.e., $N - 1$ less. Hence, the knowledge or accurate estimation of the delay d_1 can improve the system capacity (37), as confirmed by the simulation results reported in Section IV. The problem of blind estimation of the transmission delay has been addressed in several recent works, e.g., [6] and [14], [15]. Although the solutions proposed in [14], [15] can be easily adapted to our framework, they require a batch SVD of the correlation matrix $\boldsymbol{R}_{\boldsymbol{zz}}$ and, hence, are computationally too expensive for adaptive applications. More interestingly, the blind adaptive technique for joint synchronization and demodulation proposed in [6] resorts to the MOE approach, avoiding an SVD. However, since the method of [6] is addressed essentially to low-rate CDMA systems, its generalization to our algorithm is not straightforward and is the focus of our current research. Observe that, besides knowledge of the transmission delay, the increase of the system capacity (37) can also be obtained: (i) by using multiple receivers (*antenna array*) [14]; (ii) by (temporally) *oversampling* the received signal $r_c(t)$ [13]; (iii) by resorting to *widely linear* transformations of the input vector $\boldsymbol{z}(k)$ [29].

Remark 2. As to the knowledge of the desired-channel memory L_1, it should be stressed that the proposed first stage works even when L_1 is overestimated, i.e., when L_1 is replaced with $(L_1)_{\max}$. However, in this case, the disturbance suppression capability of the first stage decreases as the difference $(L_1)_{\max} - L_1$ increases: this is simply due to the fact that overestimating L_1 reduces the number $NL_e - K_1$ of available degrees of freedom for the disturbance suppression, since the number $(K_1)_{\max}$ of constraints used to preserve the desired symbol is greater than the minimal one K_1.

Remark 3. In the presence of correlated noise, unlike several methods proposed in literature, e.g., [14]-[16] and [18], [19], the proposed method does not require any *a priori* knowledge on the statistical correlation matrix $\boldsymbol{R}_{\boldsymbol{vv}} \triangleq E[\boldsymbol{v}(k)\, \boldsymbol{v}^H(k)]$ of the noise vector $\boldsymbol{v}(k)$.

Remark 4. It has been shown [30] that the proposed receiver assures a significant performance gain in comparison with existing methods even though the channel memory is smaller then one symbol period: this happens typically in DS-CDMA systems transmitting at data rates of 1.5 Mbps indoors and 50 Kbps outdoors [28].

Remark 5. For a synchronous (i.e., $\tau_j = 0$ for each user) DS-CDMA system operating over a single-path channel [i.e., $M_j = 1$ in (5)], each user contributes one signature vector during a single symbol interval (i.e., there is no ISI) and its signature $\boldsymbol{h}_j(0)$ [see (10)] is given by $\boldsymbol{h}_j(0) = \alpha_{1,j}\, [c_j(0), c_j(1), \ldots, c_j(N-1)]^T$ (see, e.g., [5]). In such an *ideal* scenario, using an observation interval of one symbol (i.e., $L_e = 1$), the proposed two-stage receiver reduces to a single-stage one, where the overall weight vector $\boldsymbol{f}$ is chosen blindly so as to minimize the output power $E[|y(k)|^2]$ subject to the constraint $\boldsymbol{f}^H [c_j(0), c_j(1), \ldots, c_j(N-1)]^T = 1$.

D. Adaptive implementation of the two-stage receiver

In this section, we will show that the proposed receiver admits a simple and computationally-efficient adaptive implementation, where the optimal unconstrained component $\boldsymbol{\mathcal{F}}_{\text{opt}}^{(a)}$ and the CM weight vector $\boldsymbol{u}$ are recursively estimated from the received data. Let us firstly consider estimation of the matrix $\boldsymbol{\mathcal{F}}_{\text{opt}}^{(a)}$, given by (27). Borrowing concepts from the well-known RLS algorithm [7], it is shown in Appendix IV that the recursion for estimating $\boldsymbol{\mathcal{F}}_{\text{opt}}^{(a)}$ is

$$\boldsymbol{\mathcal{F}}_{\text{opt}}^{(a)}(k+1) = \boldsymbol{\mathcal{F}}_{\text{opt}}^{(a)}(k) + \boldsymbol{g}(k+1) \left\{ \boldsymbol{p}(k+1) - \left[\boldsymbol{\mathcal{F}}_{\text{opt}}^{(a)}(k) \right]^H \boldsymbol{q}(k+1) \right\}^H , \tag{38}$$

where $\boldsymbol{p}(k+1) \triangleq [\boldsymbol{\mathcal{F}}_{\text{opt}}^{(0)}]^H \boldsymbol{z}(k+1)$, $\boldsymbol{q}(k+1) \triangleq \mathbf{\Pi}_{1,d}^H \boldsymbol{z}(k+1)$, and the *overall gain* vector

$$\boldsymbol{g}(k+1) \triangleq \frac{\boldsymbol{P}(k)\, \boldsymbol{q}(k+1)}{\lambda + \boldsymbol{q}^H(k+1)\, \boldsymbol{P}(k)\, \boldsymbol{q}(k+1)} , \tag{39}$$

with $\boldsymbol{P}(k)$ and $\lambda \in (0, 1]$ denoting the estimate, at iteration k, of the matrix $(\mathbf{\Pi}_{1,d}^H\, \boldsymbol{R}_{\boldsymbol{zz}}\, \mathbf{\Pi}_{1,d})^{-1}$ and the forgetting factor of the recursive algorithm, respectively. According to the usual initialization strategy for the RLS algorithm, we set $\boldsymbol{P}(0) = \delta^{-1}\, \boldsymbol{I}_{(NL_e - K_1)}$ and $\boldsymbol{\mathcal{F}}_{\text{opt}}^{(a)}(0) = \boldsymbol{O}_{(NL_e - K_1) \times K_1}$, where δ is a small positive constant. With regard to the adaptive implementation of the second stage, since it is based on the CM cost function, we can resort to the simple SGD algorithm, which leads to the following recursive equation

$$\boldsymbol{u}(k+1) = \boldsymbol{u}(k) + \mu\, y^*(k)\, (\gamma - |y(k)|^2)\, \boldsymbol{x}(k) , \tag{40}$$

where $y(k) \triangleq \boldsymbol{u}(k)^H\, \boldsymbol{x}(k)$ denotes the output of the second stage at iteration k. The algorithm (40) is commonly referred to as the CMA (see [22]).

A final remark is now in order about the overall computational load of the proposed adaptive receiver. The recursive equation (38) requires a computational complexity per iteration of order only $\mathcal{O}[(NL_e - K_1)^2]$, whereas, since $\boldsymbol{u}$ is a K_1-column vector, the computational complexity per iteration of the updating equation (40) is of order only $\mathcal{O}(K_1)$. Thus, the overall computational complexity per iteration of the two-stage receiver is of order $\mathcal{O}[(NL_e - K_1)^2] + \mathcal{O}(K_1) \sim \mathcal{O}[(NL_e - K_1)^2]$ [in practice, K_1 is significantly smaller than $(NL_e - K_1)^2$]. In Section IV, we provide simulation comparisons with the CMA-based approach of Tugnait and Li [19]. The latter approach and the related methods [15],

[18] can be fundamentally decomposed in two steps: in the former, the SVD is required to estimate a certain subspace associated with the desired user's code sequence; in latter, an iterative batch-mode minimization of an HOS cost function is performed. The overall computational complexity of the first step is thus of order $\mathcal{O}(N^3 L_e^3)$, whereas the computational requirements per iteration of the highly-optimized second step are difficult to quantify since they are problem-dependent, as recognized by the same authors (see, e.g., [19]); however, based on the results of computer simulations reported in [15], [18], [19], it seems that the computational load per iteration of the second step may be comparable with the complexity of the first one.

IV. Simulation results

In this section, we present the performance analysis (carried out by Monte Carlo computer simulations) of the proposed algorithm, together with a comparison with the methods proposed in [19]. In particular, we consider two different versions of the proposed receiver: in the first one, the optimal filtering matrix $\boldsymbol{\mathcal{F}}$ is designed by assuming the knowledge of the desired user's transmission delay d_1; in the second one, instead, the matrix $\boldsymbol{\mathcal{F}}$ is designed without requiring knowledge of the desired-user synchronization. Moreover, since the approaches proposed in [19] are batch-mode algorithms, we have initially chosen to implement our methods also in their batch versions, and assuming, moreover, perfect estimation of the correlation matrices involved in the algorithms. This choice allows us not only to carry out a meaningful comparison with the method of [19], but also to accurately determine the ultimate performance penalty with respect to the *(ideal) non-blind* maximum-SINR receiver, which has exact knowledge of the desired-channel impulse response. Soon thereafter, we also provide a detailed performance analysis of the adaptive implementation of our method.

In all the experiments, the following common simulation setting is adopted. The asynchronous DS-CDMA network employs random binary code sequences of length $N = 8$, with equiprobable values ± 1, and QPSK symbol modulation, which implies that the dispersion coefficient to be used in the CM cost function (28) is $\gamma = 1$. The chip sequence $s_j(q)$ is modulated using a unit-energy rectangular pulse $T_c^{-1/2}\,\mathrm{rect}_{T_c}(t)$ and the receiver filter is matched to the transmitting pulse, i.e., its impulse response is $T_c^{-1/2}\,\mathrm{rect}_{T_c}(-t)$. In this case, the function $\psi_c(t)$ in (5) is a triangular pulse of width $2T_c$, i.e., $\psi_c(t) = 1 - |t|/T_c$, for $|t| \leq T_c$, and zero elsewhere. The transmission delays τ_j are uniformly distributed over one symbol interval, for $j = 1, 2, \ldots, J$. The multipath channel for each user, modeled similarly to [15], [18], [19], has four paths [i.e., $M_j = 4$ in (5)], where the first path $(m = 1)$ is assumed to be deterministic with amplitude $\alpha_{1,j} = 1$ and propagation time $\tau_{1,j} = 0$; the remaining path gains $\alpha_{m,j}$, for $m = 2, 3, 4$, are modeled as mutually independent complex circular Gaussian zero-mean random variables, with standard deviation 0.3, whereas the corresponding propagation times $\tau_{m,j}$ are modeled as mutually independent random variables uniformly distributed over L_j symbol intervals, for $j = 1, 2, \ldots, J$. The additive noise $w(n)$ is modeled as a complex circular zero-mean white Gaussian process, with variance σ_w^2, and the signal-to-noise ratio (SNR) of the desired user at the detector input is defined, according to (10), as

$$\mathrm{SNR} \triangleq \sigma_b^2 \, \frac{\sum_{i=0}^{L_1+1} \|\boldsymbol{h}_1(i)\|^2}{E[\,\|\boldsymbol{w}(k)\|^2\,]} \,. \tag{41}$$

We consider a severe near-far scenario: in all the experiments the path gains of each user channel are adjusted so that each interfering user is 10 dB stronger than the user of interest $(j = 1)$. As performance measure, we resort to the SINR at the output of the considered receivers, which is defined, according to (14), as

$$\mathrm{SINR} \triangleq \sigma_b^2 \, \frac{|\boldsymbol{f}^H \boldsymbol{h}_{1,d}|^2}{\boldsymbol{f}^H \boldsymbol{R}_{\boldsymbol{\jmath}\boldsymbol{\jmath}} \boldsymbol{f}} \,, \tag{42}$$

where $\boldsymbol{R}_{\boldsymbol{\jmath}\boldsymbol{\jmath}} \triangleq E[\boldsymbol{\jmath}(k)\,\boldsymbol{\jmath}^H(k)]$ is the $NL_e \times NL_e$ statistical correlation matrix of $\boldsymbol{\jmath}(k)$. This is a widespread measure of performance for MMSE and maximum-SINR detectors (see [5], for instance), since the output of these receivers is approximately Gaussian distributed [31] and, hence, the SINR values can be directly related to those of symbol error probability. Moreover, the SINR is particularly useful in a blind setting, since it is insensitive to complex scaling of the estimated symbol. After estimating the receiver weights on the basis of the given data record, an indipendent record of 3000 symbols is considered to evaluate the SINR at the output of the considered receivers. The results are obtained by carrying out 100 independent trials, with the propagation delays, the channels and the code sequences for each user randomly generated and then fixed over all Monte Carlo runs.

In the first part of this section, the performances of the proposed receivers are firstly compared with those of the non-blind maximum-SINR detector (referred to as max-SINR in the plots) and with those of the blind receivers proposed in [19]. More precisely, two different detectors are proposed in [19]: the former (referred to as CA-CMA-1) is based on a batch-mode iterative minimization of the CM cost function, where the CM weight vector is constrained to belong to a (data-dependent) subspace associated with the desired user's code sequence; the latter (referred to as CA-CMA-2) resorts to an unconstrained iterative minimization of the CM cost function, initialized by using the CA-CMA-1 algorithm. The non-blind maximum-SINR receiver and the first stage of the proposed approaches [see equation (27)] are evaluated using the theoretical correlation matrix $\boldsymbol{R}_{\boldsymbol{z}\boldsymbol{z}}$ of the received vector $\boldsymbol{z}(k)$, which is given by (31), whereas, for the methods of [19], the desired subspace is extracted by employing its (exact) theoretical dimension and using the noise-free[3] correlation matrix $\boldsymbol{R}_{\boldsymbol{z}\boldsymbol{z}} - \sigma_w^2\,\boldsymbol{I}_{NL_e}$. To carry out a fair comparison, the iterative minimization (29) in the second stage is obtained in batch-mode as suggested in [19], over 300 symbols. However, unlike [19], in the known-delay case, the CM weight vector $\boldsymbol{u}$ is initialized with a single one in the first component, i.e., $\boldsymbol{u}(0) = [\,1, 0, \ldots, 0\,]^T$, whereas, in the unknown-delay case,

[3]It is recognized in [19] that the presence of additive noise in the data reduces the effectiveness of the proposed constraint.

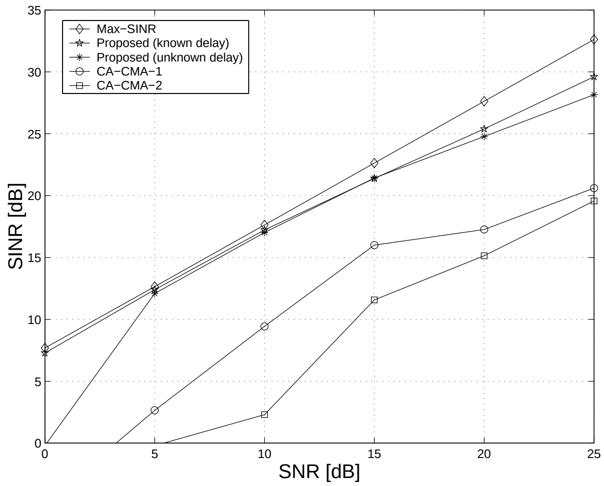


Fig. 2. SINR versus SNR (first experiment, theoretical performances, short multipath channel, three active users).

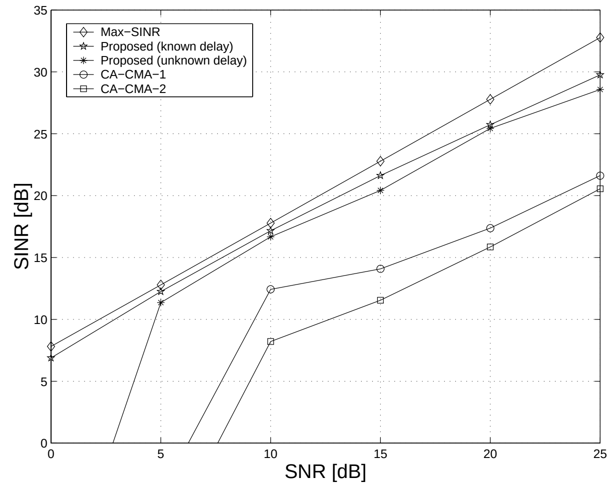


Fig. 4. SINR versus SNR (third experiment, theoretical performances, long multipath channel, three active users).

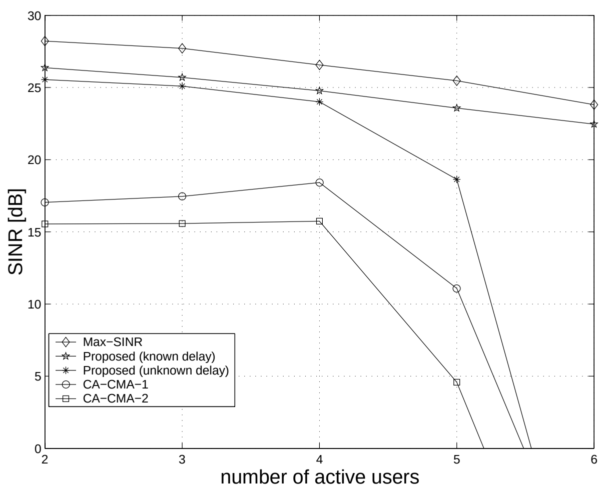


Fig. 3. SINR versus number of active users (second experiment, theoretical performances, short multipath channel, SNR$= 20$ dB).

the CM weight vector $\boldsymbol{u}$ is initialized with a single one in the Nth component, i.e., $\boldsymbol{u}(0) = [0, \ldots, 0, 1, 0, \ldots, 0]^T$.

In the first experiment, we consider a "short" multipath channel, which spans only one symbol interval, i.e., $L_j = 1$ for each user. We report in Fig. 2 the values of SINR obtained with the different receivers as a function of SNR ranging from 0 to 25 dB. The number of active users is $J = 3$ and, for all the considered receivers, we have fixed $L_e = 5$ and the equalization delay[4] at $d = 3$. The results show that the proposed receivers significantly outperform the detectors of [19] for all values of SNR. Moreover, for values of SNR exceeding 5 dB, the lack of knowledge of the desired-user synchronization does not appreciably affect the performances of the two-stage detectors: in fact, the SINR curves of both proposed receivers are remarkably close to those of the ideal max-SINR detector. The second experiment investigates the impact of the considered receivers on system capacity, for transmission over the same short multipath channel. We report in Fig. 3 the values of SINR as a function of the number of active users ranging from 2 to 6 (recall that the processing gain is $N = 8$). The SNR is set to 20 dB and, for all the considered receivers, we have fixed $L_e = 8$ and the equalization delay at $d = 5$. It can be observed that both the proposed receivers also assure a substantial performance gain with respect to CA-CMA-1 and CA-CMA-2 detectors in highly loaded systems. In particular, the proposed receiver (known delay) is able to handle up to six users, whereas the performance of the proposed receiver (unknown delay) is within a few dB from the max-SINR limit only in moderately loaded networks, i.e., for a number of users less or equal than four. It should be noted that an estimate of the transmission delay d_1 is only needed when the system loading is high, that is, the number of users J approaches the processing gain N. The third experiment is aimed at showing that the proposed methods are also effective in the presence of long-delay multipath, i.e., when the channel multipath spread exceeds the symbol duration. In particular, we consider a situation where the channel spans two symbol intervals, i.e., $L_j = 2$ for each user. Fig. 4 reports the values of SINR obtained with the different receivers as a function of SNR ranging from 0 to 25 dB. The number of active users is $J = 3$ and, for all the considered receivers, we have fixed $L_e = 8$ and the equalization delay at $d = 5$. The results show that the performances of both the proposed receivers are similar to those reported in Fig. 2, albeit a larger value of L_e is required in this case to obtain satisfactory results.

In the rest of this section, we consider the adaptive version of the proposed algorithms. It should be stressed that, since the methods of [19] do not admit an adaptive version, they will not be considered anymore. As performance bound, we report in all the simulations the (non-blind) max-SINR receiver, which is now implemented by using the sample-correlation matrix $\widehat{\boldsymbol{R}}_{\boldsymbol{zz}}$ (rather than the exact one $\boldsymbol{R}_{\boldsymbol{zz}}$) of the received vector $\boldsymbol{z}(k)$. However, it is well-known [5] that in this case there exists a significant performance gap between the estimated maximum-SINR receiver and the true one, for high values of

[4]Note that, to avoid the capture of an interfering user, the algorithms in [19] explicitly require that $d \geq 2L_1 + 1$.

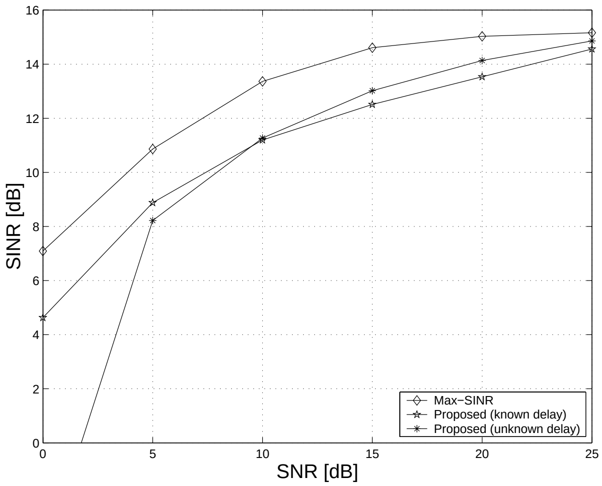


Fig. 5. SINR versus SNR (fourth experiment, adaptive implementation, short multipath channel, three active users).

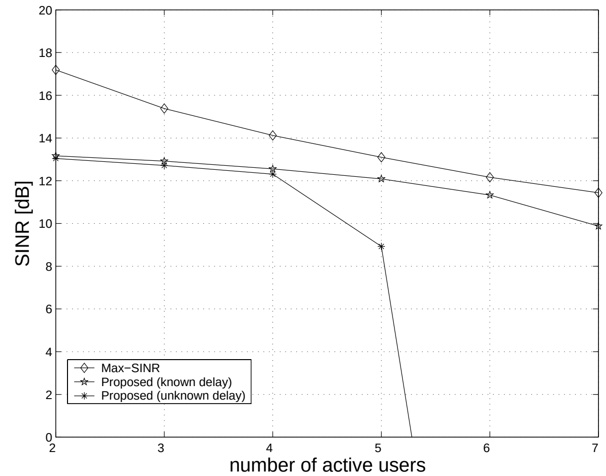


Fig. 6. SINR versus number of active users (fifth experiment, adaptive implementation, short multipath channel, SNR$=20$ dB).

SNR; to overcome this performance saturation effect, in the subsequent experiments we simulate the max-SINR receiver by using its subspace-based implementation [8]. For both the proposed algorithms, we chose $\lambda = \delta = 1$ in the first stage, whereas, in the second stage, the CMA recursive equation (40) is initialized as in the previous experiments and the step-size μ is continuously adjusted, in order to achieve fast convergence without compromising stability. More specifically, we set $\mu(k) = 0.1\,\mu_{\max}(k)$, where, according to [32], $\mu_{\max}(k)$ is the maximum value of the step-size that assures CMA stability at iteration k, and can be evaluated in real-time, since it depends only on the output $y(k)$ of the receiver and the dispersion coefficient γ.

In the fourth experiment, we consider the same simulation setting described in the first experiment (short multipath channel), with the exception that the considered receivers are estimated by using 625 symbols. We report in Fig. 5 the values of SINR as a function of SNR ranging from 0 to 25 dB. Results show that the proposed receivers provide good performances also when adaptively implemented. Moreover, for high values of SNR, their performances are close to those of the non-blind (subspace-based) max-SINR receiver. To further corroborate the satisfactory performances of the proposed adaptive receivers, we report in Fig. 6 (fifth experiment) the values of SINR as a function of the number of active users ranging from 2 to 7. The simulation setting is the same considered in the second experiment (short multipath channel) and the sample-size is 1000 symbols. The results confirm those reported in Fig. 3; in particular, it should be noted that the proposed receiver (known delay) can successfully work also with seven users in a highly-loaded system. Results, not reported here, show that the proposed receivers exhibit a remarkable robustness also when the power of the interfering users varies by several orders of magnitude, i.e., the proposed receivers achieve near-far resistance. In the sixth experiment, we consider the severe multipath environment (long multipath channel) and simulation setting used in the third experiment, with a sample-size of 1000 symbols. The results show that

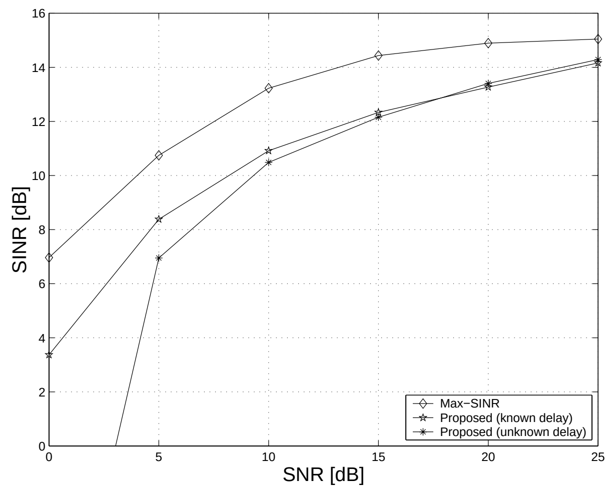


Fig. 7. SINR versus SNR (sixth experiment, adaptive implementation, long multipath channel, three active users).

the performances of both the proposed adaptive receivers are similar to those reported in Fig. 5, albeit a larger value of L_e is required in this case to perform satisfactorily. Finally, we also investigated the tracking performance of the proposed algorithm (unknown delay), i.e., its capability to operate in a situation where there is a drastic change in the interference environment. More specifically, we consider the following situation: during the first 500 iterations, there are three active users; at iteration 501 one additional user appears; at iteration 1001 one user vanishes. The details regarding SNR, receivers's length, equalization delays and sample-size are as in the fifth experiment. It should be noted that the receivers of [19] cannot work in this scenario since they require a fresh subspace extraction from data at each change. We report in Fig. 8 the values of SINR as a function of the number k of iterations, evaluated as in [5]: the proposed receiver is able to operate satisfactorily in this nonstationary environment, exhibiting a better tracking behavior when an user disappears than when a new user enters. Finally, results not reported here show that reducing the memory of the pre-filtering stage (i.e., by setting

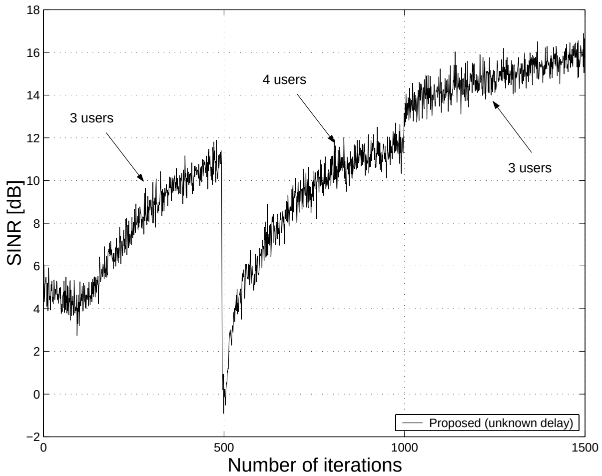


Fig. 8. SINR versus number of iterations for varying-interference scenario (seventh experiment, adaptive implementation, short multipath channel, SNR= 20 dB).

$\lambda < 1$) does not significantly affect the tracking capabilities of the overall receiver.

V. CONCLUSIONS

In this paper, we have proposed a new blind (batch or adaptive) receiver for joint equalization and multiuser detection in DS-CDMA systems. The proposed equalizer is composed of two stages: in the first stage, a linear transformation is performed on the received data, in order to blindly preserve the desired symbol while minimizing the contribution of disturbances (ISI-plus-MAI and noise); in the second stage, the CMA is used to extract the desired symbol. The proposed receiver can be adaptively implemented with a reasonable computational burden and requires only the knowledge of the desired user's spreading code in moderately loaded networks, whereas, in highly loaded networks, it can take additional advantage from the knowledge of the desired-user synchronization. Our current research is aimed at investigating the theoretical conditions assuring satisfactory disturbance cancellation in the first stage, in order to avoid the capture effect in the second stage. Moreover, the feasibility of suppressing deterministically the in-cell users in uplink environments, by using the knowledge of their spreading sequences (the so called *group-blind* multiuser detection), is currently under investigation.

APPENDIX I
PARAMETERIZATION OF THE COMPOSITE SIGNATURE $\boldsymbol{h}_{1,d}$

In this appendix, we show how the columns $\boldsymbol{h}_{1,d}$ of the channel matrix $\boldsymbol{\mathcal{H}}_1$ can be linearly parameterized over the entire range of the equalization delays $d \in \{0, 1, \ldots, L_1 + L_e\}$. In order to help streamline this section, we define the vector/matrix operator $\mathcal{T}_{n,m}[\cdot]$, which associates to any column vector $\boldsymbol{a}$ the column vector formed by extracting its components in the range (n, m), whereas it associates with any matrix $\boldsymbol{A}$ the matrix formed by extracting its columns in the range (n, m). Let us define the $N \times 1$ desired code vector $\boldsymbol{c}_1 \triangleq [\, c_1(N-1),\, c_1(N-2),\, \ldots,\, c_1(0)\,]$ and the following $N \times N(L_1+1)$ matrices constructed from $\boldsymbol{c}_1$:

$$\boldsymbol{\mathcal{C}}_1^{(0)} \triangleq \begin{bmatrix} c_1(0) & \boldsymbol{0}^T_{N(L_1+1)-1} \\ \{\mathcal{T}_{N-1,N}(\boldsymbol{c}_1)\}^T & \boldsymbol{0}^T_{N(L_1+1)-2} \\ \vdots & \vdots \\ \{\mathcal{T}_{1,N}(\boldsymbol{c}_1)\}^T & \boldsymbol{0}^T_{NL_1} \end{bmatrix}, \qquad (43)$$

$$\boldsymbol{\mathcal{C}}_1^{(L_1+1)} \triangleq \begin{bmatrix} \boldsymbol{0}^T_{NL_1+1} & \{\mathcal{T}_{1,N-1}(\boldsymbol{c}_1)\}^T \\ \boldsymbol{0}^T_{NL_1+2} & \{\mathcal{T}_{1,N-2}(\boldsymbol{c}_1)\}^T \\ \vdots & \vdots \\ \boldsymbol{0}^T_{NL_1} & \boldsymbol{0}^T_{N} \end{bmatrix}, \qquad (44)$$

and, for $m \in \{1, \ldots, L_1\}$,

$$\boldsymbol{\mathcal{C}}_1^{(m)} \triangleq \begin{bmatrix} \boldsymbol{0}^T_{N(m-1)+1} & \boldsymbol{c}_1^T & \boldsymbol{0}^T_{N(L_1-m+1)-1} \\ \boldsymbol{0}^T_{N(m-1)+2} & \boldsymbol{c}_1^T & \boldsymbol{0}^T_{N(L_1-m+1)-2} \\ \vdots & \vdots & \vdots \\ \boldsymbol{0}^T_{Nm} & \boldsymbol{c}_1^T & \boldsymbol{0}^T_{N(L_1-m)} \end{bmatrix}. \qquad (45)$$

The vector $\boldsymbol{h}_{1,d}$ can be linearly parameterized as $\boldsymbol{h}_{1,d} = \widetilde{\boldsymbol{\mathcal{D}}}_{1,d}\, \boldsymbol{g}_1$, where the expression of the parameterization matrix $\widetilde{\boldsymbol{\mathcal{D}}}_{1,d}$ depends on the value of d. Specifically, if $d \in \{0, 1, \ldots, L_1+1\}$, the vector $\boldsymbol{h}_{1,d}$ is given by

$$\boldsymbol{h}_{1,d} = [\, \underbrace{\boldsymbol{h}_1^T(d),\, \ldots,\, \boldsymbol{h}_1^T(0)}_{N(d+1)},\ \underbrace{0,\, \ldots,\, 0}_{N(L_e-d-1)}\,]^T \qquad (46)$$

and, hence, recalling the definition of $\boldsymbol{h}_1(k)$ and accounting for (7), the matrix $\widetilde{\boldsymbol{\mathcal{D}}}_{1,d}$ is given by

$$\widetilde{\boldsymbol{\mathcal{D}}}_{1,d} = \left[(\boldsymbol{\mathcal{C}}_1^{(d)})^T, \ldots, (\boldsymbol{\mathcal{C}}_1^{(0)})^T, \boldsymbol{O}^T_{N(L_e-d-1)\times N(L_1+1)} \right]^T. \qquad (47)$$

Furthermore, if $d \in \{L_1+2, L_1+3, \ldots, L_e-1\}$, the vector $\boldsymbol{h}_{1,d}$ is given by

$$\boldsymbol{h}_{1,d} = [\, \underbrace{0,\, \ldots,\, 0}_{N(d-L_1-1)},\ \underbrace{\boldsymbol{h}_1^T(L_1+1),\, \ldots,\, \boldsymbol{h}_1^T(0)}_{N(L_1+2)},\ \underbrace{0,\, \ldots,\, 0}_{N(L_e-d-1)}\,]^T \qquad (48)$$

and, hence, recalling the definition of $\boldsymbol{h}_1(k)$ and accounting for (7), the matrix $\widetilde{\boldsymbol{\mathcal{D}}}_{1,d}$ is given by

$$\widetilde{\boldsymbol{\mathcal{D}}}_{1,d} = \Big[\boldsymbol{O}^T_{N(d-L_1-1)\times N(L_1+1)}, (\boldsymbol{\mathcal{C}}_1^{(L_1+1)})^T, \ldots \\ \ldots, (\boldsymbol{\mathcal{C}}_1^{(0)})^T, \boldsymbol{O}^T_{N(L_e-d-1)\times N(L_1+1)} \Big]^T. \qquad (49)$$

Finally, if $\{L_e, L_e+1, \ldots, L_1+L_e\}$, the vector $\boldsymbol{h}_{1,d}$ is given by

$$\boldsymbol{h}_{1,d} = [\, \underbrace{0,\, \ldots,\, 0}_{N(d-L_1-1)}\ \underbrace{\boldsymbol{h}_1^T(L_1+1),\, \ldots,\, \boldsymbol{h}_1^T(d-L_e+1)}_{N(L_1+L_e-d+1)}\,]^T \qquad (50)$$

and, hence, recalling the definition of $\boldsymbol{h}_1(k)$ and accounting for (7), the matrix $\widetilde{\boldsymbol{\mathcal{D}}}_{1,d}$ is given by

$$\widetilde{\boldsymbol{\mathcal{D}}}_{1,d} = \Big[\boldsymbol{O}^T_{N(d-L_1-1)\times N(L_1+1)}, (\boldsymbol{\mathcal{C}}_1^{(L_1+1)})^T, \ldots \\ \ldots, (\boldsymbol{\mathcal{C}}_1^{(d-L_e+1)})^T \Big]^T. \qquad (51)$$

As a final remark, it is worthwhile to note that, for $d \in \{L_1, L_1+1, \ldots, L_e-1\}$, the matrix $\widetilde{\boldsymbol{\mathcal{D}}}_{1,d}$ turns out to be full column rank, whereas, for $d \in \{0, 1, \ldots, L_1-1\}$ or $d \in \{L_e, L_e+1, \ldots, L_e+L_1\}$, it is rank deficient, i.e., $\text{rank}(\widetilde{\boldsymbol{\mathcal{D}}}_{1,d}) < N(L_1+1)$. In the latter case, the rank reduction is simply due to the fact that, for $d \in \{0, 1, \ldots, L_1-1\}$, all the elements of the last $N(L_1-d)$ columns of $\widetilde{\boldsymbol{\mathcal{D}}}_{1,d}$ are zero and, for $d \in \{L_e, L_e+1, \ldots, L_e+L_1\}$, all the elements of the first $N(d-L_e)+1$ columns of $\widetilde{\boldsymbol{\mathcal{D}}}_{1,d}$ are zero. This problem can be easily circumvented by observing that, for $d \in \{0, 1, \ldots, L_1-1\}$ or $d \in \{L_e, L_e+1, \ldots, L_e+L_1\}$, the signature $\boldsymbol{h}_{1,d}$ can be linearly parameterized as $\boldsymbol{h}_{1,d} = \boldsymbol{\mathcal{D}}_{1,d}\, \boldsymbol{g}_{1,d}$ [see (18)], where the matrix $\boldsymbol{\mathcal{D}}_{1,d}$ is obtained from $\widetilde{\boldsymbol{\mathcal{D}}}_{1,d}$ by deleting the null columns and the vector $\boldsymbol{g}_{1,d}$ is accordingly obtained from $\boldsymbol{g}_1$. More precisely, the expression of $\boldsymbol{\mathcal{D}}_{1,d}$ and $\boldsymbol{g}_{1,d}$, to be used in (18), is given by one of the following parameterizations: for $d \in \{0, 1, \ldots, L_1-1\}$, parameterization (18) is given by $\boldsymbol{\mathcal{D}}_{1,d} = \mathcal{T}_{1,N(d+1)}(\widetilde{\boldsymbol{\mathcal{D}}}_{1,d})$ and $\boldsymbol{g}_{1,d} = \mathcal{T}_{1,N(d+1)}(\boldsymbol{g}_1)$ (*parameterization* 1); for $d \in \{L_1, L_1+1, \ldots, L_e-1\}$, parameterization (18) is directly given by $\boldsymbol{\mathcal{D}}_{1,d} = \widetilde{\boldsymbol{\mathcal{D}}}_{1,d}$ and $\boldsymbol{g}_{1,d} = \boldsymbol{g}_1$ (*parameterization* 2); for $d \in \{L_e, L_e+1, \ldots, L_e+L_1\}$, parameterization (18) is given by $\boldsymbol{\mathcal{D}}_{1,d} = \mathcal{T}_{N(d-L_e)+2, N(L_1+1)}(\widetilde{\boldsymbol{\mathcal{D}}}_{1,d})$ and $\boldsymbol{g}_{1,d} = \mathcal{T}_{N(d-L_e)+2, N(L_1+1)}(\boldsymbol{g}_1)$ (*parameterization* 3).

APPENDIX II
PROOF OF PROPOSITION 1

Let r denote the rank of the $P \times (NL_e - K_1)$ matrix $\boldsymbol{\mathcal{H}}^H \boldsymbol{\Pi}_{1,d}$, the following singular value decomposition (SVD) holds (see [26])

$$\boldsymbol{\mathcal{H}}^H \boldsymbol{\Pi}_{1,d} = \boldsymbol{U}\, \boldsymbol{\Lambda}\, \boldsymbol{V}^H \,, \tag{52}$$

with

$$\boldsymbol{\Lambda} = \begin{bmatrix} \boldsymbol{\Lambda}_r & \boldsymbol{O}_{r \times (NL_e - K_1 - r)} \\ \boldsymbol{O}_{(P-r)\times r} & \boldsymbol{O}_{(P-r)\times(NL_e-K_1-r)} \end{bmatrix} \tag{53}$$

where $\boldsymbol{\Lambda}_r = \text{diag}[\lambda_1, \lambda_2, \ldots, \lambda_r]$ is the $r \times r$ diagonal matrix containing the nonzero singular value of $\boldsymbol{\mathcal{H}}^H \boldsymbol{\Pi}_{1,d}$, whereas $\boldsymbol{U}$ and $\boldsymbol{V}$ are unitary matrices of dimensions $P \times P$ and $(NL_e - K_1) \times (NL_e - K_1)$, respectively. Accounting for (53) and applying the matrix inversion lemma [7], we get after rearrangement

$$\left(\sigma_b^2\, \boldsymbol{\Pi}_{1,d}^H\, \boldsymbol{\mathcal{H}}\, \boldsymbol{\mathcal{H}}^H\, \boldsymbol{\Pi}_{1,d} + \sigma_w^2\, \boldsymbol{I}_{NL_e-K_1}\right)^{-1} = \sigma_w^{-2}\, \boldsymbol{I}_{NL_e-K_1} - \boldsymbol{V}\, \widetilde{\boldsymbol{\Sigma}}\, \boldsymbol{V}^H \,, \tag{54}$$

where we have defined the $(NL_e - K_1) \times (NL_e - K_1)$ diagonal matrix

$$\widetilde{\boldsymbol{\Sigma}} \triangleq \sigma_w^{-4}\, \sigma_b^2\, \boldsymbol{\Lambda}^T \left(\boldsymbol{I}_P + \sigma_w^{-2}\, \sigma_b^2\, \boldsymbol{\Lambda}\, \boldsymbol{\Lambda}^T\right)^{-1} \boldsymbol{\Lambda} = \begin{bmatrix} \widetilde{\boldsymbol{\Sigma}}_r & \boldsymbol{O}_{r \times (NL_e-K_1-r)} \\ \boldsymbol{O}_{(NL_e-K_1-r)\times r} & \boldsymbol{O}_{(NL_e-K_1-r)\times(NL_e-K_1-r)} \end{bmatrix}, \tag{55}$$

with $\widetilde{\boldsymbol{\Sigma}}_r = \sigma_w^{-4}\, \sigma_b^2\, \boldsymbol{\Lambda}_r \left(\boldsymbol{I}_r + \sigma_w^{-2}\, \sigma_b^2\, \boldsymbol{\Lambda}_r^2\right)^{-1} \boldsymbol{\Lambda}_r$. At this point, by substituting (53) and (54) in (32), it follows after some algebra

$$\boldsymbol{\mathcal{F}}_{\text{opt}} = \left[\boldsymbol{I}_{NL_e} - \boldsymbol{\Pi}_{1,d}\, (\boldsymbol{V}\, \boldsymbol{\Sigma}\, \boldsymbol{U}^H)\, \boldsymbol{\mathcal{H}}^H\right] \boldsymbol{\mathcal{F}}_{\text{opt}}^{(0)} \,, \tag{56}$$

where, accounting for (55), we have defined the $(NL_e - K_1) \times P$ matrix

$$\boldsymbol{\Sigma} \triangleq \sigma_w^{-2}\, \sigma_b^2\, \boldsymbol{\Lambda}^T - \sigma_b^2\, \widetilde{\boldsymbol{\Sigma}}\, \boldsymbol{\Lambda}^T = \begin{bmatrix} \boldsymbol{\Sigma}_r & \boldsymbol{O}_{r\times(P-r)} \\ \boldsymbol{O}_{(NL_e-K_1-r)\times r} & \boldsymbol{O}_{(NL_e-K_1-r)\times(P-r)} \end{bmatrix} \,, \tag{57}$$

with $\boldsymbol{\Sigma}_r = \sigma_w^{-2}\, \sigma_b^2\, \boldsymbol{\Lambda}_r - \sigma_w^{-4}\, \sigma_b^4\, \boldsymbol{\Lambda}_r \left(\boldsymbol{I}_r + \sigma_w^{-2}\, \sigma_b^2\, \boldsymbol{\Lambda}_r^2\right)^{-1} \boldsymbol{\Lambda}_r^2$. It is easily seen that the ith diagonal entry $[\boldsymbol{\Sigma}_r]_i$ of the matrix $\boldsymbol{\Sigma}_r$ is given by

$$[\boldsymbol{\Sigma}_r]_i = \frac{\sigma_b^2\, \lambda_i}{\sigma_w^2 + \sigma_b^2\, \lambda_i^2} \,, \quad \text{for} \quad i \in \{1, 2, \ldots, r\} \,. \tag{58}$$

Based on (58), we can see that, as σ_w^2/σ_b^2 approaches zero, the diagonal matrix $\boldsymbol{\Sigma}_r$ converges to $\boldsymbol{\Lambda}_r^{-1}$ and, hence, $\boldsymbol{V}\, \boldsymbol{\Sigma}\, \boldsymbol{U}^H$ in (56) converges to the pseudo inverse of $\boldsymbol{\mathcal{H}}^H\, \boldsymbol{\Pi}_{1,d}$ [27], that is, $\boldsymbol{V}\, \boldsymbol{\Sigma}\, \boldsymbol{U}^H \to (\boldsymbol{\mathcal{H}}^H\, \boldsymbol{\Pi}_{1,d})^\dagger$. Therefore, in the high SNR region, the optimal filtering matrix (30) assumes the expression given by (33). ■

APPENDIX III
PROOF OF LEMMA 1

Since $\text{null}(\boldsymbol{\Pi}_{1,d}^H\, \boldsymbol{\mathcal{H}}) \supseteq \text{null}(\boldsymbol{\mathcal{H}})$ and $\text{null}(\boldsymbol{\mathcal{Q}}_{1,d}^H\, \boldsymbol{\mathcal{H}}) \supseteq \text{null}(\boldsymbol{\mathcal{H}})$, the intersection between the null spaces of $\boldsymbol{\Pi}_{1,d}^H\, \boldsymbol{\mathcal{H}}$ and $\boldsymbol{\mathcal{Q}}_{1,d}^H\, \boldsymbol{\mathcal{H}}$ can be written as $\text{null}(\boldsymbol{\Pi}_{1,d}^H\, \boldsymbol{\mathcal{H}}) \cap \text{null}(\boldsymbol{\mathcal{Q}}_{1,d}^H\, \boldsymbol{\mathcal{H}}) = \text{null}(\boldsymbol{\mathcal{H}}) \cup \mathcal{A}$, where we have defined the subset $\mathcal{A} \triangleq \{\boldsymbol{\zeta} \neq \boldsymbol{0}_{NL_e}$ lying in $\text{range}(\boldsymbol{\mathcal{H}}) \,|\, \boldsymbol{\Pi}_{1,d}^H\, \boldsymbol{\zeta} = \boldsymbol{0}_{NL_e-K_1}$ and $\boldsymbol{\mathcal{Q}}_{1,d}^H\, \boldsymbol{\zeta} = \boldsymbol{0}_{K_1}\}$. However, the subset $\mathcal{A}$ is empty: in fact, since the matrices $\boldsymbol{\Pi}_{1,d}$ and $\boldsymbol{\mathcal{Q}}_{1,d}$ are orthogonal, i.e., $\boldsymbol{\Pi}_{1,d}^H\, \boldsymbol{\mathcal{Q}}_{1,d} = \boldsymbol{O}_{(NL_e-K_1)\times K_1}$, the subspaces $\text{range}(\boldsymbol{\Pi}_{1,d})$ and $\text{range}(\boldsymbol{\mathcal{Q}}_{1,d})$ are disjoint, i.e., $\text{range}(\boldsymbol{\Pi}_{1,d}) \cap \text{range}(\boldsymbol{\mathcal{Q}}_{1,d}) = \{\boldsymbol{0}_{NL_e}\}$; consequently, the composite $NL_e \times NL_e$ matrix $[\boldsymbol{\mathcal{Q}}_{1,d}, \boldsymbol{\Pi}_{1,d}]$ is full rank and, hence, the only vector $\boldsymbol{\zeta} \in \text{range}(\boldsymbol{\mathcal{H}})$ satisfying the system $[\boldsymbol{\mathcal{Q}}_{1,d}, \boldsymbol{\Pi}_{1,d}]^H\, \boldsymbol{\zeta} = \boldsymbol{0}_{NL_e}$ is the zero vector. ■

APPENDIX IV
UPDATING EQUATION FOR $\boldsymbol{\mathcal{F}}_{\text{opt}}^{(a)}(k)$

We seek the recursive equation for updating the matrix

$$\boldsymbol{\mathcal{F}}_{\text{opt}}^{(a)}(k+1) = [\,\boldsymbol{\Pi}_{1,d}^H\; \boldsymbol{R}_{\boldsymbol{zz}}(k+1)\, \boldsymbol{\Pi}_{1,d}\,]^{-1} \times \boldsymbol{\Pi}_{1,d}^H\; \boldsymbol{R}_{\boldsymbol{zz}}(k+1)\, \boldsymbol{\mathcal{F}}_{\text{opt}}^{(0)} \,, \tag{59}$$

where the updating equation of the correlation matrix $\boldsymbol{R}_{\boldsymbol{zz}}$ is given by (see [7])

$$\boldsymbol{R}_{\boldsymbol{zz}}(k+1) = \lambda\, \boldsymbol{R}_{\boldsymbol{zz}}(k) + \boldsymbol{z}(k+1)\, \boldsymbol{z}^H(k+1) \,, \tag{60}$$

with $\lambda \in (0, 1]$ denoting the forgetting factor. By using (60), it is easily seen that

$$[\,\boldsymbol{\Pi}_{1,d}^H\; \boldsymbol{R}_{\boldsymbol{zz}}(k+1)\, \boldsymbol{\Pi}_{1,d}\,]^{-1} = \left\{\lambda\, \boldsymbol{\Pi}_{1,d}^H\; \boldsymbol{R}_{\boldsymbol{zz}}(k)\, \boldsymbol{\Pi}_{1,d} + \boldsymbol{q}(k+1)\, \boldsymbol{q}^H(k+1)\right\}^{-1} \,, \tag{61}$$

where $\boldsymbol{q}(k+1) \triangleq \boldsymbol{\Pi}_{1,d}^H \, \boldsymbol{z}(k+1)$. Denoting with $\boldsymbol{P}(k) \triangleq [\,\boldsymbol{\Pi}_{1,d}^H \, \boldsymbol{R}_{\boldsymbol{zz}}(k) \, \boldsymbol{\Pi}_{1,d}\,]^{-1}$ and applying the matrix inversion lemma to (61), one obtains after rearrangement

$$\boldsymbol{P}(k+1) = \lambda^{-1} \left[\boldsymbol{P}(k) - \boldsymbol{g}(k+1) \, \boldsymbol{q}^H(k+1) \, \boldsymbol{P}(k) \right] , \quad (62)$$

where the overall gain vector $\boldsymbol{g}(k+1)$ has already been defined in (39). In order to obtain the equation for updating $\boldsymbol{\mathcal{F}}_{1,\mathrm{opt}}^{(a)}(k)$, we substitute (60) and (62) in (59), obtaining thus

$$\begin{aligned}\boldsymbol{\mathcal{F}}_{\mathrm{opt}}^{(a)}(k+1) &= \lambda \, \boldsymbol{P}(k+1) \, \boldsymbol{\Pi}_{1,d}^H \, \boldsymbol{R}_{\boldsymbol{zz}}(k) \, \boldsymbol{\mathcal{F}}_{\mathrm{opt}}^{(0)} \\ &+ \boldsymbol{P}(k+1) \, \boldsymbol{\Pi}_{1,d}^H \, \boldsymbol{z}(k) \, \boldsymbol{p}^H(k+1) \\ &= \boldsymbol{\mathcal{F}}_{\mathrm{opt}}^{(a)}(k) + \lambda^{-1} \, \boldsymbol{P}(k) \, \boldsymbol{q}(k+1) \, \boldsymbol{p}^H(k+1) \\ &- \boldsymbol{g}(k+1) \, \boldsymbol{q}^H(k+1) \, \boldsymbol{\mathcal{F}}_{\mathrm{opt}}^{(a)}(k) \\ &- \lambda^{-1} \, \boldsymbol{g}(k+1) \, \boldsymbol{q}^H(k+1) \, \boldsymbol{P}(k) \\ &\times \boldsymbol{q}(k+1) \, \boldsymbol{p}^H(k+1) \, , \end{aligned} \quad (63)$$

where $\boldsymbol{p}(k+1) \triangleq [\,\boldsymbol{\mathcal{F}}_{\mathrm{opt}}^{(0)}]^H \boldsymbol{z}(k+1)$. Substituting (39) in (63) and taking the lowest common multiple between the last three terms, one obtains after some algebra

$$\begin{aligned}\boldsymbol{\mathcal{F}}_{\mathrm{opt}}^{(a)}(k+1) &= \boldsymbol{\mathcal{F}}_{\mathrm{opt}}^{(a)}(k) \\ &+ \left[\frac{\boldsymbol{P}(k) \, \boldsymbol{q}(k+1)}{\lambda + \boldsymbol{q}^H(k+1) \, \boldsymbol{P}(k) \, \boldsymbol{q}(k+1)} \right] \\ &\times \left\{ \boldsymbol{p}(k+1) - \left[\boldsymbol{\mathcal{F}}_{\mathrm{opt}}^{(a)}(k) \right]^H \boldsymbol{q}(k+1) \right\}^H , \end{aligned} \quad (64)$$

which, by reintroducing the overall gain vector (39), yields the updating equation for $\boldsymbol{\mathcal{F}}_{\mathrm{opt}}^{(a)}(k)$.

ACKNOWLEDGMENTS

The authors would like to thank the anonymous reviewers for their constructive comments and suggestions, which helped improve the quality of the manuscript.

REFERENCES

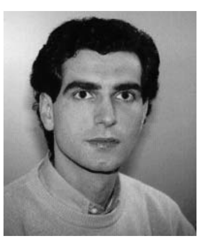
[1] J.G. Proakis, *Digital Communications*. New York: McGraw-Hill, 2001.
[2] S. Verdù, *Multiuser detection*. Cambridge: Cambridge University Press, 1998.
[3] U. Madhow and M.L. Honig, "MMSE interference suppression for direct-sequence spread-spectrum CDMA," *IEEE Trans. Commun.*, vol. 42, pp. 3178–3188, Dec. 1994.
[4] M. Honig and M.K. Tsatsanis, "Multiuser CDMA receivers," *IEEE Signal Processing Mag.*, pp. 49–61, May 2000.
[5] M. Honig, U. Madhow, and S. Verdù, "Blind adaptive multiuser detection," *IEEE Trans. Inform. Theory*, vol. 41, pp. 944–960, July 1995.
[6] U. Madhow, "Blind adaptive interference suppression for the near-far resistant acquisition and demodulation of DS-CDMA signals," *IEEE Trans. Signal Processing*, vol. 45, pp. 124–136, Jan. 1997.
[7] S. Haykin, *Adaptive Filter Theory*. New York: Prentice Hall, 1996.
[8] 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.
[9] C.B. Papadias and A.J. Paulraj, "A constant modulus algorithm for multiuser signal separation in presence of delay spread using antenna arrays," *IEEE Signal Processing Letters*, vol. 4, pp. 178–181, June 1997.
[10] C.B. Papadias and H. Huang, "Linear space-time multiuser detection for multipath CDMA channels," *IEEE J. Select. Areas Commun.*, vol. 19, pp. 254–265, Feb. 2001.
[11] D.N. Godard, "Self-recovering equalization and carrier tracking in two-dimensional data communication systems," *IEEE Trans. Commun.*, vol. 28, pp. 1867–1875, Nov. 1980.
[12] J.R. Treichler and B.G. Agee, "A new approach to multipath correction of costant modulus signals," *IEEE Trans. Acoust. Speech, Signal Processing*, vol. 31, no. 4, pp. 349–372, Apr. 1983.
[13] E. Moulines, P. Duhamel, J. Cardoso, and S. Mayrargue, "Subspace methods for the blind identification of multichannel FIR filters," *IEEE Trans. Signal Processing*, vol. 45, pp. 516–525, Feb. 1995.
[14] M. Torlak and G. Xu, "Blind multiuser channel estimation in asynchronous CDMA systems," *IEEE Trans. Signal Processing*, vol. 45, pp. 137–147, Jan. 1997.
[15] J.K. Tugnait and T. Li, "A multistep linear prediction approach to blind asynchronous CDMA channel estimation and equalization," *IEEE J. Select. Areas Commun.*, vol. 19, pp. 1090–1102, June 2001.
[16] M.K. Tsatsanis and Z.D. Xu, "Performance analysis of minimum variance CDMA receivers," *IEEE Trans. Signal Processing*, vol. 46, pp. 3014–3022, Nov. 1998.
[17] D. Gesbert, J. Sorelius, P. Stoica, and A. Paulraj, "Blind multiuser MMSE detection of CDMA signals in ISI channels," *IEEE Commun. Lett.*, vol. 3, pp. 233–235, Aug. 1999.
[18] J.K. Tugnait and T. Li, "Blind detection of asynchronous CDMA signals in multipath channels using code-constrained inverse filter criterion," *IEEE Trans. Signal Proc.*, vol. 49, pp. 1300–1309, July 2001.
[19] J.K. Tugnait and T. Li, "Blind asynchronous multiuser CDMA receivers for ISI channels using code-aided CMA," *IEEE J. Select. Areas Commun.*, vol. 19, pp. 1520–1530, Aug. 2001.
[20] G. Gelli and F. Verde, "Two-stage interference-resistant adaptive periodically time-varying CMA blind equalization," *IEEE Trans. Signal Processing*, vol. 50, pp. 662–672, Mar. 2002.
[21] J.E. Smee and S.C.S. Schwartz, "Adaptive feedforward/feedback architectures for multiuser detection in high data rate wireless CDMA networks," *IEEE Trans. Commun.*, vol. 48, pp. 996–1011, June 2000.
[22] C.R. Johnson, P. Schniter, T.J. Endres, J.D. Behm, D.R. Brown, and R.A. Casas. "Blind equalization using the constant modulus criterion: a review," *Proc. IEEE*, vol. 86, pp. 1927–1950, Oct. 1998.
[23] J.K. Tugnait, "Blind spatio-temporal equalization and impulse response estimation for MIMO channels using a Godard cost function," *IEEE Trans. Signal Processing*, vol. 45, pp. 268–271, Jan. 1997.
[24] P. Schniter and C.R. Johnson Jr., "Sufficient conditions for the local convergence of constant modulus algorithms," *IEEE Trans. Signal Processing*, vol. 48, pp. 2785–2796, Oct. 2000.
[25] L.J. Griffiths and C.W. Jim, "An alternative approach to linearly constrained adaptive beamforming," *IEEE Trans. Antennas and Propagat.*, vol. 30, pp. 27–34, Jan. 1982.
[26] R.A. Horn and C.R. Johnson, *Matrix Analysis*, Cambridge: Cambridge University Press, 1990.
[27] A. Björck, *Numerical Methods for Least Squares Problems*, Philadelphia: Siam, 1996.
[28] 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.
[29] G. Gelli, L. Paura and A.R.P. Ragozini, "Blind widely linear multiuser detection," *IEEE Commun. Lett.*, vol. 4, pp. 187–189, June 2000.
[30] G. Gelli and F. Verde, "Two-stage CMA blind multiuser detection for DS-CDMA systems in multipath fading," in *Proc. of 2nd IEEE-EURASIP International Symposium on Image and Signal Processing and Analysis*, Pula, Croatia, pp. 622–627, 2001.
[31] H.V. Poor and S. Verdù, "Probability of error in MMSE multiuser detection," *IEEE Trans. Inform. Theory*, vol. 43, pp. 868–871, May 1997.
[32] M. Rupp and S.C. Douglas, "A posteriori analysis of adaptive blind equalizers," in *Proc. 32nd Asilomar Conf. Signals, Syst. Comput.*, Pacific Grove, CA, pp. 369–373, 1998.

Giacinto Gelli (M'02) was born in Napoli, Italy, on July 29, 1964. He received the Dr. Eng. degree *summa cum laude* in electronic engineering in 1990, and the Ph.D. degree in computer science and electronic engineering in 1994, both from the University of Napoli Federico II.

From 1994 to 1998, he was an Assistant Professor with the Department of Information Engineering, Second University of Napoli. Since 1998, he has been an Associate Professor of Telecommunications with the Department of Electronic and Telecommunication Engineering, University of Napoli Federico II. His research interests are in the fields of statistical signal processing, array processing, image processing, and mobile communications, with current emphasis on code-division multiple-access systems and multicarrier modulation.

Luigi Paura was born in Napoli, Italy, on February 20, 1950. He received the Dr. Eng. degree *summa cum laude* in electronic engineering in 1974 from the University of Napoli Federico II.

From 1979 to 1984 he was with the Department of Electronic and Telecommunication Engineering, University of Napoli, Italy, first as an Assistant Professor and then as an Associate Professor. Since 1994, he has been a Full Professor of Telecommunications: first, with the Department of Mathematics, University of Lecce, Italy; then, with the Department of Information Engineering, Second University of Napoli; and, finally, from 1998 he has been with the Department of Electronic and Telecommunication Engineering, University of Napoli Federico II. He also held teaching positions at the University of Salerno, Italy, at the University of Sannio, Italy, and the University Parthenope, Napoli, Italy. In 1985-86 and 1991 he was a Visiting Researcher at the Signal and Image Processing Laboratory, University of California, Davis. At the present time, his research activities are mainly concerned with statistical signal processing, digital communication systems and medium access control in wireless networks.



Francesco Verde was born in Santa Maria Capua Vetere, Italy, on June 12, 1974. He received the Dr. Eng. degree *summa cum laude* in electronic engineering in 1998 from the Second University of Napoli, and the Ph.D. degree in information engineering in 2002, from the University of Napoli Federico II.

Since 2002, he has been an Assistant Professor with the Department of Electronic and Telecommunication Engineering, University of Napoli Federico II. His research activities lie in the area of statistical signal processing, digital communications, and communication systems. In particular, his current interests are focused on cyclostationarity-based techniques for blind identification, equalization and interference suppression for narrowband modulation systems, code-division multiple-access systems and multicarrier modulation systems.