Hindawi Publishing Corporation
EURASIP Journal on Wireless Communications and Networking
Volume 2006, Article ID 24132, Pages 1–13
DOI 10.1155/WCN/2006/24132

Multipass Channel Estimation and Joint Multiuser Detection and Equalization for MIMO Long-Code DS/CDMA Systems

Stefano Buzzi

DAEIMI, Università degli Studi di Cassino, Via G. Di Biasio 43, 03043 Cassino, Italy

Received 8 April 2005; Revised 16 October 2005; Accepted 28 November 2005

Recommended for Publication by Wolfgang Gerstacker

The problem of joint channel estimation, equalization, and multiuser detection for a multiantenna DS/CDMA system operating over a frequency-selective fading channel and adopting long aperiodic spreading codes is considered in this paper. First of all, we present several channel estimation and multiuser data detection schemes suited for multiantenna long-code DS/CDMA systems. Then, a multipass strategy, wherein the data detection and the channel estimation procedures exchange information in a recursive fashion, is introduced and analyzed for the proposed scenario. Remarkably, this strategy provides, at the price of some attendant computational complexity increase, excellent performance even when very short training sequences are transmitted, and thus couples together the conflicting advantages of both trained and blind systems, that is, good performance and no wasted bandwidth, respectively. Space-time coded systems are also considered, and it is shown that the multipass strategy provides excellent results for such systems also. Likewise, it is also shown that excellent performance is achieved also when each user adopts the same spreading code for all of its transmit antennas. The validity of the proposed procedure is corroborated by both simulation results and analytical findings. In particular, it is shown that adopting the multipass strategy results in a remarkable reduction of the channel estimation mean-square error and of the optimal length of the training sequence.

1. INTRODUCTION

Direct-sequence code-division multiple-access (DS/CDMA) techniques are of considerable interest, since they are among the basic technologies for the realization of the air interface of current and future wireless networks [1]. One of the salient features of the emerging CDMA-based wireless networks standards is the adoption of long (aperiodic) spreading codes. Even though the use of long codes ensures that all the users achieve “on the average” the same performance in a frequency-flat channel with perfect power control, it destroys the bit-interval cyclostationarity properties of the CDMA signals and thus renders ineffective many of the advanced signal processing techniques that have been developed for blind multiuser detection and adaptive channel estimation in short-code CDMA systems [2, 3]. The design of intelligent signal processing techniques for DS/CDMA systems with aperiodic spreading codes is thus a challenging research topic.

While most contributions in this area (e.g., [4–7]) either propose detection and estimation algorithms with heavy computational complexity or rely on prior knowledge of the propagation delay of the user of interest and of the interfering signature waveforms, in [8] a least-squares channel estimation procedure has been introduced. This procedure relies on the transmission of known pilot symbols, and it may be implemented with a computational complexity which is quadratic in the processing gain, and is suited for both the uplink and the downlink. The channel estimation procedures presented in [8] have been also used in [9], wherein a recursive algorithm, based on an iterative exchange of information between the data detector and the channel estimator, is proposed in order to improve the system performance.

So far, the problem of devising effective channel estimation algorithms for long-code CDMA systems has mainly focused on the case that the transmitter and the receiver are equipped with a single antenna, and indeed all the papers so far cited refer to this scenario. On the other hand, of late there has been a growing interest in the design and the analysis of communication systems employing multiple transmit and receive antennas, also known as multiple-input multiple-output (MIMO) systems. Indeed, recent results from information theory have shown that in a rich scattering environment the capacity of multiantenna communication systems grows with a law approximately linear in the minimum between the number of transmit and receive antennas [10–12]. In general, the use of multiple antennas has a complicated impact on the performance of a wireless communication system. The use of multiple antennas at the transmitter, in uncoded systems, permits attaining a wireless communication link with large spectral efficiency. Otherwise stated, having N_t antennas at the transmitter and a serial to parallel converter with N_t outputs, permits transmitting a given symbol stream with a bandwidth N_t times smaller than the one required by a system using the same modulation format and having only one transmit antenna. In a space-time coded system, instead, part of the increase in the spectral efficiency may be sacrificed for transmit diversity, which provides increased performance and resistance to fading. The use of multiple antennas at the receiver, instead, provides a receiver diversity advantage, since the receiving multiple antennas provide multiple independently faded replicas of the transmitted signals. This helps to improve performance and, also, to separate the symbols transmitted by different antennas through suitable signal processing techniques. Several multiantenna communication architectures have been thus proposed, formerly for single-user systems [13, 14], and then for multiuser systems [15, 16]. In particular, the paper [16] develops subspace-based blind adaptive multiuser detectors for short-code DS/CDMA systems with transceivers equipped with multiple antennas. Since these subspace-based techniques rely on the symbol-interval cyclostationarity of the observed data, they are not applicable to CDMA systems employing long codes.

Following on this track, in this paper we consider the problem of channel estimation for multiantenna DS/CDMA systems employing long codes. The contributions of this paper can be summarized as follows.

(i) We extend the iterative channel estimation and data detection procedure in [9] to the case where each user is equipped with multiple transmit and receive antennas. It is thus shown that the iterative strategy permits achieving, at the price of little attendant computational complexity increase, excellent performance in MIMO systems also for very short lengths of the training sequence.

(ii) With regard to the problem of channel estimation, we extend the least-squares channel estimation procedures of [8] to the case of a multiantenna transceiver.

(iii) We provide a theoretical performance analysis of the proposed iterative strategy, which leads to a closed-form formula relating the mean square channel estimation error at each iteration with the error probability achieved by the data detector at the previous iteration.

(iv) We show that the proposed iterative strategy provides excellent results also in the case that each user is assigned only one spreading code, which is thus used to spread the data symbols on all of its transmit antennas.

(v) The problem of how to set the length of the training sequence is considered. Indeed, this length should be chosen as a compromise between the conflicting requirements of achieving a reliable channel estimate and of not reducing too much the system throughput, that is, the fraction of information bits in each data packet. A cost function is thus introduced, whose minimization can be used to set the optimal training length. Through our analysis, it is thus shown that the proposed multipass strategy permits achieving, in the region of interest of moderately small and small error probabilities, lower values of the cost function with lower values of the training sequence length (i.e., with a larger throughput).

(vi) It is shown that the proposed iterative channel estimation and data detection scheme can be extended to orthogonal space-time coded system with moderate efforts. In particular, simulation results for the Alamouti space-time code [17] are provided.

The rest of this paper is organized as follows. Section 2 contains the signal model for the considered multiuser long-code CDMA MIMO system. Sections 3 and 4 are devoted to the synthesis of multiuser MIMO channel estimation techniques and of MIMO multiuser detection algorithms, respectively. In Section 5 the basic idea of the multipass strategy is presented along with a thorough performance analysis showing its merits. In particular, we present both theoretical findings and numerical simulation results, demonstrating the accuracy of the theoretical analysis. In Section 6 space-time coded long-code CDMA systems are briefly examined, and it is shown that the proposed multipass approach can be applied to such systems too with excellent performance. Finally, concluding remarks are given in Section 7.

2. MULTIANTENNA DS/CDMA SIGNAL MODEL

Consider an asynchronous DS/CDMA system with K active users employing long (aperiodic) codes. We assume that every transceiver is equipped with N_t transmit antennas and N_r receive antennas;[1] the information stream of the kth user (at rate R) is demultiplexed into N_t information substreams at rate $R_b = R/N_t$; each substream is independently transmitted by only one transmit antenna.[2] Denote by $T_b = 1/R_b$ the symbol interval, by N the processing gain, by $T_c = T_b/N$ the chip interval, and assume that $\{\beta_{k,p}^{(n),n_t}\}_{n=0}^{N-1}$ is the kth user spreading sequence in the pth symbol interval for the n_tth transmit antenna.[3] Denoting by $u_{T_c}(t)$ a unit-height rectangular pulse supported in $[0, T_c]$, the signature waveform transmitted by the n_tth transmit antenna of the kth user in the pth symbol interval is written as[4]

$$s_{k,p}^{n_t}(t) = \sum_{n=0}^{N-1} \beta_{k,p}^{(n),n_t} u_{T_c}(t - nT_c). \tag{1}$$

In keeping with [4–8], we consider the case of slow frequency-selective fading channels, and denote by $c_k^{n_t,n_r}(t)$ the impulse response of the channel linking the kth user n_tth

[1] It is usually assumed that $N_r \geq N_t$; however, this hypothesis is not needed here.

[2] A block scheme of the considered system is depicted in Figure 1.

[3] Note that, for the moment, we are assuming that each user is assigned N_t different spreading codes, one for each transmit antenna.

[4] For the sake of simplicity, we consider here the use of rectangular chip pulses. Note, however, that all of the subsequent derivations can be extended in a straightforward manner to bandlimited chip pulses.

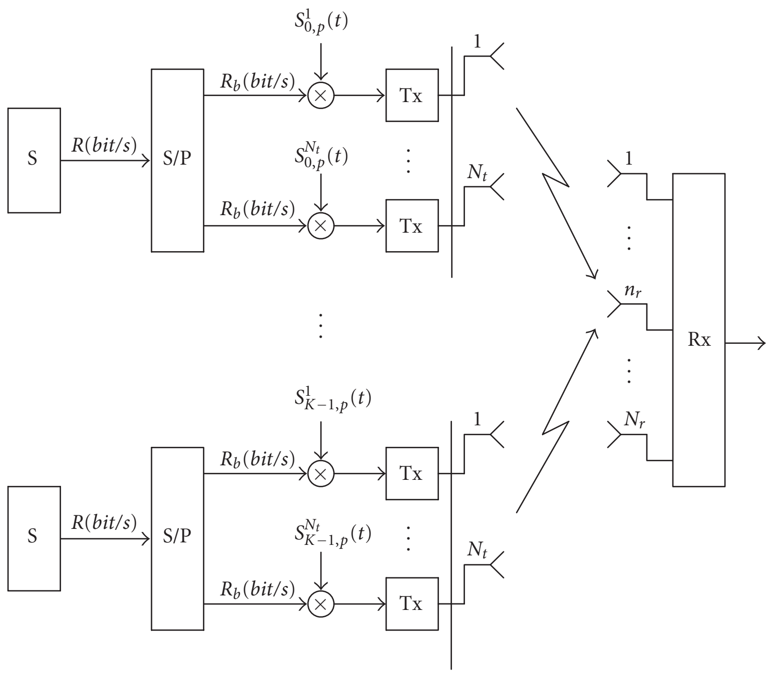

Figure 1: A multiantenna multiuser communication system.

transmit antenna with the n_rth receive antenna; it is here assumed that the scattering environment is "rich" and that the antenna elements are sufficiently spaced so that the channel impulse responses are independent for all k, n_t, n_r. Based on the above assumptions, the complex envelope of the signal observed at the n_rth receive antenna is written as

$$r_{n_r}(t) = \sum_{p=1}^{B} \sum_{k=0}^{K-1} \sum_{n_t=0}^{N_t} A_k b_k^{n_t}(p) s_{k,p}^{n_t}(t - \tau_k - pT_b) * c_k^{n_t,n_r}(t) + w_{n_r}(t). \quad (2)$$

In the above equation, $*$ denotes convolution, B is the length of the data frame measured in symbol intervals, $b_k^{n_t}(p) \in \{+1, -1\}$ is the symbol transmitted in the pth signaling interval on the n_tth transmit antenna of the user k (note that we are here considering BPSK modulation), A_k and τ_k are the amplitude and timing offset of the kth user. We also assume that the multipath delay spread T_m is such that $\tau_k + T_m < T_b$. Finally, $w_{n_r}(t)$ is the additive thermal noise that we model as a white complex zero-mean Gaussian random process with power spectral density (PSD) $2\mathcal{N}_0$; we also have $E[w_{n_r}(t)w_{n_r'}^*(u)] = 0$, for $n_r \neq n_r'$. Now, the signal (2) can be cast in a form such that it resembles the signal model of a synchronous DS/CDMA system with no fading. Indeed, letting

$$h_{k,p}^{n_t,n_r}(t) = A_k s_{k,p}^{n_t}(t - \tau_k) * c_k^{n_t,n_r}(t) = \sum_{n=0}^{N-1} \beta_{k,p}^{(n),n_t} \underbrace{A_k u_{T_c}(t - \tau_k - nT_c) * c_k^{n_t,n_r}(t)}_{g_k^{n_t,n_r}(t-nT_c)}, \quad (3)$$

the unknown channel impulse response and timing offset are shoved in the waveform $g_k^{n_t,n_r}(\cdot)$, which is supported on the interval $[\tau_k, \tau_k + T_c + T_m] \subseteq [0, T_b + T_c]$; on the other hand, note also that $h_{k,p}^{n_t,n_r}(t)$ is supported on $[\tau_k, \tau_k + T_b + T_m] \subseteq [0, 2T_b]$. Based on (3), it is seen that (2) can be thus written as

$$r_{n_r}(t) = \sum_{p=1}^{B} \sum_{k=0}^{K-1} \sum_{n_t=1}^{N_t} b_k^{n_t}(p) h_{k,p}^{n_t,n_r}(t - pT_b) + w_{n_r}(t). \quad (4)$$

Now, the received signal is converted to discrete time at a rate of M samples per chip interval according to the following projection:

$$r_{n_r}(\ell) = \sqrt{\frac{M}{T_c}} \int_{\ell T_c/M}^{(\ell+1)T_c/M} r_{n_r}(t)dt. \quad (5)$$

Stacking in the vector $\mathbf{r}_{n_r}(p)$ the NM samples arising from the discretization of the received signal as observed in the pth signaling interval $[pT_b, (p+1)T_b]$, it can be shown that $\mathbf{r}_{n_r}(p)$ can be expressed as

$$\mathbf{r}_{n_r}(p) = \sum_{k=0}^{K-1} \sum_{n_t=1}^{N_t} \left[b_k^{n_t}(p-1)\mathbf{h}_{k,p-1}^{l,n_t,n_r} + b_k^{n_t}(p)\mathbf{h}_{k,p}^{u,n_t,n_r} \right] + \mathbf{w}_{n_r}(p), \quad (6)$$

where it is assumed that $b_k^{n_t}(0) = 0$, for all $k = 0, \ldots, K-1$ and for all $n_t = 1, \ldots, N_t$. In the above equation, $\mathbf{h}_{k,p-1}^{l,n_t,n_r}$ and $\mathbf{h}_{k,p-1}^{u,n_t,n_r}$ denote the discretized contributions of the waveforms $\mathbf{h}_{k,p-1}^{n_t,n_r}(t - (p-1)T_b)$ and $\mathbf{h}_{k,p}^{n_t,n_r}(t - pT_b)$, respectively, to the interval $[pT_b, (p+1)T_b]$, while $\mathbf{w}_{n_r}(p)$ is the vector of the thermal noise projections, which are independent zero-mean complex Gaussian random variates with variance $2\mathcal{N}_0$.

Now, note that upon defining the projections

$$g_k^{n_t,n_r}(\ell) = \sqrt{\frac{M}{T_c}} \int_{\ell T_c/M}^{(\ell+1)T_c/M} g_k^{n_t,n_r}(t)dt, \tag{7}$$

we have that $g_k^{n_t,n_r}(\ell) = 0$ for $\ell \notin \{0, 1, \ldots, (N+1)M-1\}$, whereby we can stack in the $(N+1)M$-dimensional vector $\mathbf{g}_k^{n_t,n_r}$ the nonzero projections of the waveform $g_k^{n_t,n_r}(t)$, that is,

$$\mathbf{g}_k^{n_t,n_r} = \left[g_k^{n_t,n_r}(0), \ldots, g_k^{n_t,n_r}((N+1)M-1) \right]^T. \tag{8}$$

Moreover, denote by $\mathbf{C}_{k,p}^{n_t}$ the following $2NM \times (N+1)M$-dimensional matrix, containing properly shifted versions of the kth user spreading code adopted in the pth symbol interval on the n_tth transmit antenna:

$$\mathbf{C}_{k,p}^{n_t} = \begin{bmatrix} \beta_{k,p}^{(0),n_t} & 0 & \cdots & \cdots & 0 \\ \beta_{k,p}^{(1),n_t} & \beta_{k,p}^{(0),n_t} & 0 & \cdots & 0 \\ \vdots & \beta_{k,p}^{(1),n_t} & \ddots & \cdots & 0 \\ \beta_{k,p}^{(N-1),n_t} & \vdots & \ddots & \ddots & 0 \\ 0 & \beta_{k,p}^{(N-1),n_t} & \ddots & \ddots & \beta_{k,p}^{(0),n_t} \\ \vdots & 0 & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & \beta_{k,p}^{(N-1),n_t} \end{bmatrix} \otimes \mathbf{I}_M, \tag{9}$$

with $\otimes$ denoting the Kronecker product and $\mathbf{I}_M$ the identity matrix of order M. The matrix $\mathbf{C}_{k,p}^{n_t}$ can be partitioned into two $NM \times (N+1)M$-dimensional matrices that we denote by $\mathbf{C}_{k,p}^{u,n_t}$ and $\mathbf{C}_{k,p}^{l,n_t}$, that is,

$$\mathbf{C}_{k,p}^{n_t} = \begin{bmatrix} \mathbf{C}_{k,p}^{u,n_t} \\ \mathbf{C}_{k,p}^{l,n_t} \end{bmatrix}. \tag{10}$$

Based on the above notation, it can be shown that the relations $\mathbf{h}_{k,p-1}^{l,n_t,n_r} = \mathbf{C}_{k,p-1}^{l,n_t}\mathbf{g}_k^{n_t,n_r}$ and $\mathbf{h}_{k,p}^{u,n_t,n_r} = \mathbf{C}_{k,p}^{u,n_t}\mathbf{g}_k^{n_t,n_r}$ hold, whereby the vector $\mathbf{r}_{n_r}(p)$ in (6) can be cast in the following form:

$$\mathbf{r}_{n_r}(p) = \sum_{k=0}^{K-1}\sum_{n_t=1}^{N_t} \Big[b_k^{n_t}(p-1)\mathbf{C}_{k,p-1}^{l,n_t} + b_k^{n_t}(p)\mathbf{C}_{k,p}^{u,n_t} \Big] \mathbf{g}_k^{n_t,n_r} + \mathbf{w}_{n_r}(p). \tag{11}$$

The above representation, which extends to the multiple-antenna scenario the one developed in [8] for single-antenna systems, is extremely powerful; indeed, from (11) it is seen that even though aperiodic long codes changing at each symbol interval are adopted, and even though the propagation delay and the channel impulse response are not known, the discrete-time signatures may be deemed as the product of a time-varying, but *known*, matrix, containing properly shifted versions of the spreading codes, times an unknown, but *time-invariant*, vector, which carries information on the channel impulse response and timing offset. Now, based on the representation in (11), our actual goal is to provide an estimation algorithm for the channel vectors $\mathbf{g}_k^{n_t,n_r}$.

3. MULTIUSER MIMO CHANNEL ESTIMATION

We consider the case that the channel vectors of all the active users are to be estimated, based on the knowledge of their spreading codes and relying on the transmission of known pilot symbols; this is a typical situation in the uplink of cellular CDMA systems.

First of all, we have to develop a compact representation for the discrete-time signals received on all the N_r receive antennas. To this end, let $\overline{\mathbf{C}}_{k,p}^{n_t} = [\mathbf{C}_{k,p-1}^{l,n_t} \ \ \mathbf{C}_{k,p}^{u,n_t}]$ be an $NM \times 2(N+1)M$-dimensional matrix, and

$$\overline{\mathbf{B}}_k^{n_t}(p) = \begin{bmatrix} b_k^{n_t}(p-1) \\ b_k^{n_t}(p) \end{bmatrix} \otimes \mathbf{I}_{(N+1)M} \tag{12}$$

a $2(N+1)M \times (N+1)M$-dimensional matrix. We thus have that (11) can be expressed as

$$\mathbf{r}_{n_r}(p) = \sum_{k=0}^{K-1}\sum_{n_t=1}^{N_t} \overline{\mathbf{C}}_{k,p}^{n_t}\overline{\mathbf{B}}_k^{n_t}(p)\mathbf{g}_k^{n_t,n_r} + \mathbf{w}_{n_r}(p). \tag{13}$$

Letting now $\mathbf{A}_{k,p} = [\overline{\mathbf{C}}_{k,p}^{1}\overline{\mathbf{B}}_k^{1}(p), \ldots, \overline{\mathbf{C}}_{k,p}^{N_t}\overline{\mathbf{B}}_k^{N_t}(p)]$ be an $NM \times N_t(N+1)M$-dimensional matrix and[5] letting $\widetilde{\mathbf{g}}_k^{n_r} = [\mathbf{g}_k^{1,n_r T} \ \ \mathbf{g}_k^{2,n_r T} \ \cdots \ \mathbf{g}_k^{N_t,n_r T}]^T$ be a column vector of length $(N+1)MN_t$, the summation over the index n_t may be shoved in the following matrix notation:

$$\mathbf{r}_{n_r}(p) = \sum_{k=0}^{K-1} \mathbf{A}_{k,p}\widetilde{\mathbf{g}}_k^{n_r} + \mathbf{w}_{n_r}(p). \tag{14}$$

The above representation holds for all $n_r = 1, \ldots, N_r$; stacking the vectors $\mathbf{r}_{n_r}(p)$ in an N_rNM-dimensional vector, say $\mathbf{r}(p)$, we have

$$\mathbf{r}(p) = \begin{bmatrix} \mathbf{r}_1(p) \\ \mathbf{r}_2(p) \\ \vdots \\ \mathbf{r}_{N_r}(p) \end{bmatrix} = \sum_{k=0}^{K-1} \begin{bmatrix} \mathbf{A}_{k,p}\widetilde{\mathbf{g}}_k^{1} \\ \mathbf{A}_{k,p}\widetilde{\mathbf{g}}_k^{2} \\ \vdots \\ \mathbf{A}_{k,p}\widetilde{\mathbf{g}}_k^{N_r} \end{bmatrix} + \begin{bmatrix} \mathbf{w}_1(p) \\ \mathbf{w}_2(p) \\ \vdots \\ \mathbf{w}_{N_r}(p) \end{bmatrix}. \tag{15}$$

Upon defining the $N_rNM \times N_tN_r(N+1)M$ block diagonal matrix $\mathbf{X}_{k,p} = \text{Diag}\Big\{ \underbrace{\mathbf{A}_{k,p}, \ldots, \mathbf{A}_{k,p}}_{N_r} \Big\}$ and the $N_rN_t(N+1)M$-dimensional vector $\widetilde{\mathbf{g}}_k = [\widetilde{\mathbf{g}}_k^{1T} \ \ \widetilde{\mathbf{g}}_k^{2T} \ \cdots \ \widetilde{\mathbf{g}}_k^{N_r T}]^T$, (15) can be finally written through the following compact notation:

$$\mathbf{r}(p) = \sum_{k=0}^{K-1} \mathbf{X}_{k,p}\widetilde{\mathbf{g}}_k + \mathbf{w}(p). \tag{16}$$

[5] $(\cdot)^T$ denotes transpose.

Finally, letting $\mathbf{F}_p = [\mathbf{X}_{0,p}, \mathbf{X}_{1,p}, \ldots, \mathbf{X}_{K-1,p}]$ be an $N_rNM \times KN_tN_r(N+1)M$-dimensional matrix and

$$\mathbf{q} = \begin{bmatrix} \widetilde{\mathbf{g}}_0^T & \widetilde{\mathbf{g}}_1^T & \cdots & \widetilde{\mathbf{g}}_{K-1}^T \end{bmatrix}^T \tag{17}$$

a $KN_tN_r(N+1)M$-dimensional vector, containing all the unknown quantities for all the active users, the observable $\mathbf{r}(p)$ in (16) can be expressed as

$$\mathbf{r}(p) = \mathbf{F}_p\mathbf{q} + \mathbf{w}(p). \tag{18}$$

Given the above representation, performing pilot-aided centralized channel estimation amounts to estimating the unknown vector $\mathbf{q}$ based on the knowledge of the matrices $\mathbf{F}_1, \ldots, \mathbf{F}_T$, with T denoting the number of signaling intervals devoted to the training phase. Accordingly, assuming that the receiver has an initial uncertainty on the delays $\tau_0, \ldots, \tau_{K-1}$ equal to $[-T_b/2, T_b/2]^K$, an estimate, say $\widehat{\mathbf{q}}(n)$, of the vector $\mathbf{q}$, available after observation of training symbols for n symbol intervals, is obtained by solving the problem

$$\widehat{\mathbf{q}}(n) = \arg\min_{\widehat{\mathbf{q}}} \sum_{p=1}^{n} \frac{1}{n} \left|\left|\mathbf{r}(p) - \mathbf{F}_p\widehat{\mathbf{q}}\right|\right|^2. \tag{19}$$

It is readily seen that solving the above problem requires that

$$n > \frac{KN_t(N+1)}{N}, \tag{20}$$

that is, there is a minimum number of symbol intervals that have to be devoted to training in order to enable the least-squares channel estimation. Equation (20) is a necessary condition for the existence of the inverse of the matrix $\sum_{p=1}^{n} \mathbf{F}_p^H, \mathbf{F}_p$. If (20) holds, the solution to (19) under mild conditions can be written as

$$\widehat{\mathbf{q}}(n) = \left(\sum_{p=1}^{n} \mathbf{F}_p^H\mathbf{F}_p\right)^{-1} \cdot \left(\sum_{p=1}^{n} \mathbf{F}_p^H\mathbf{r}(p)\right). \tag{21}$$

It is worth pointing out that, given the signal model (18), the least-squares solution (21) does coincide with the maximum-likelihood estimate of the vector $\mathbf{q}$; moreover, since there is a linear relationship between the thermal-noise-free observables and the vector $\mathbf{q}$, (21) coincides also with the minimum variance unbiased estimator (MVUE). With regard to the computational complexity, given the sparse nature of the matrix $\mathbf{F}_p^H\mathbf{F}_p$, it is easy to show that the solution (21) entails an $\mathcal{O}((KN_tM)^3(N+1)^3)$ computational complexity. Likewise, it can be also shown that processing separately the signals observed on each receive antenna does not yield any performance loss. Moreover, a lower complexity estimation rule can be obtained by resorting to the stochastic gradient recursive update, which yields

$$\begin{aligned} \widehat{\mathbf{q}}(n) = &\left[\mathbf{I}_{KN_tN_r(N+1)M} - \mu \sum_{p=1}^{n} \frac{1}{n}\mathbf{F}_p^H\mathbf{F}_p\right] \\ &\cdot \widehat{\mathbf{q}}(n-1) + \mu \sum_{p=1}^{n} \frac{1}{n}\mathbf{F}_p^H\mathbf{r}(p). \end{aligned} \tag{22}$$

Computational complexity is now reduced to $\mathcal{O}((KN_tNM)^2)$.

4. MIMO MULTIUSER DETECTION

In the following we extend some multiuser detection strategies to multiantenna DS/CDMA systems employing aperiodic spreading codes. It is assumed that channel estimation has been first accomplished, so that the receiver has knowledge of the estimates of the vectors $\mathbf{g}_k^{n_t,n_r}$. Note that, in order to detect the symbols $b_k^{n_t}(p)$, for all k and for all n_t, it is convenient to consider the discrete-time samples of the received signal corresponding to the interval $\mathcal{I}_p = [pT_b, (p+2)T_b]$, since, due to the assumption that $\tau_k + T_m < T_b$, the contribution of these bits falls entirely within $\mathcal{I}_p$. It is easy to show that the discrete-time version of the signal received on the n_rth receive antenna in the interval $\mathcal{I}_p$ is expressed through the following $2NM$-dimensional vector:

$$\begin{aligned} \mathbf{r}_2^{n_r}(p) = &\sum_{k=0}^{K-1}\sum_{n_t=1}^{N_t} \Big(b_k^{n_t}(p-1)\widetilde{\mathbf{C}}_{k,p-1}^{l,n_t} + b_k^{n_t}(p)\mathbf{C}_{k,p}^{n_t} \\ &+ b_k^{n_t}(p+1)\widetilde{\mathbf{C}}_{k,p+1}^{u,n_t}\Big)\mathbf{g}_k^{n_t,n_r} + \mathbf{w}_2^{n_r}(p). \end{aligned} \tag{23}$$

In the above equation, $\mathbf{w}_2^{n_r}(p)$ is the thermal noise contribution, while $\widetilde{\mathbf{C}}_{k,p-1}^{l,n_t}$ and $\widetilde{\mathbf{C}}_{k,p+1}^{u,n_t}$ are $2NM \times (N+1)M$-dimensional matrices defined as

$$\widetilde{\mathbf{C}}_{k,p-1}^{l,n_t} = \begin{bmatrix} \mathbf{C}_{k,p-1}^{l,n_t} \\ \mathbf{O}_{NM,(N+1)M} \end{bmatrix}, \qquad \widetilde{\mathbf{C}}_{k,p+1}^{u,n_t} = \begin{bmatrix} \mathbf{O}_{NM,(N+1)M} \\ \mathbf{C}_{k,p+1}^{u,n_t} \end{bmatrix}. \tag{24}$$

Upon defining the matrices

$$\begin{aligned} \widetilde{\mathbf{D}}_{k,p-1}^{l,n_t} = \mathbf{I}_{N_r} \otimes \widetilde{\mathbf{C}}_{k,p-1}^{l,n_t}, \qquad \mathbf{D}_{k,p}^{n_t} = \mathbf{I}_{N_r} \otimes \mathbf{C}_{k,p}^{n_t}, \\ \widetilde{\mathbf{D}}_{k,p+1}^{u,n_t} = \mathbf{I}_{N_r} \otimes \widetilde{\mathbf{C}}_{k,p+1}^{u,n_t}, \end{aligned} \tag{25}$$

and the $2N_rNM$-dimensional vectors

$$\widetilde{\mathbf{g}}_k^{n_t} = \begin{bmatrix} \mathbf{g}_k^{n_t,1} \\ \mathbf{g}_k^{n_t,2} \\ \vdots \\ \mathbf{g}_k^{n_t,N_r} \end{bmatrix}, \qquad \mathbf{w}_2(p) = \begin{bmatrix} \mathbf{w}_2^1(p) \\ \mathbf{w}_2^2(p) \\ \vdots \\ \mathbf{w}_2^{N_r}(p) \end{bmatrix}, \tag{26}$$

the $2N_rNM$-dimensional vector $\mathbf{r}_2(p)$, obtained by stacking the $2NM$-dimensional vectors $\mathbf{r}_2^1(p), \ldots, \mathbf{r}_2^{N_r}(p)$ vectors, is written as

$$\begin{aligned} \mathbf{r}_2(p) = &\sum_{k=0}^{K-1}\sum_{n_t=1}^{N_t} \Big(b_k^{n_t}(p-1)\widetilde{\mathbf{D}}_{k,p-1}^{l,n_t} + b_k^{n_t}(p)\mathbf{D}_{k,p}^{n_t} \\ &+ b_k^{n_t}(p+1)\widetilde{\mathbf{D}}_{k,p+1}^{u,n_t}\Big)\widetilde{\mathbf{g}}_k^{n_t} + \mathbf{w}_2(p). \end{aligned} \tag{27}$$

Based on (27), it is now easy to extend multiuser detection strategies to MIMO DS/CDMA systems.

4.1. The linear MMSE receiver

In order to detect the bit $b_h^{n_t}(p)$, transmitted by the n_tth transmit antenna of the hth user, the linear MMSE receiver implements the following rule:

$$\widehat{b}_h^{n_t}(p) = \operatorname{sgn}\left\{\Re\left[\left(\mathbf{D}_{h,p}^{n_t}\widetilde{\mathbf{g}}_h^{n_t}\right)^H \times \left(\mathbf{H}(p)\mathbf{H}(p)^H + 2\mathcal{N}_0\mathbf{I}_{2NMN_r}\right)^{-1}\mathbf{r}_2(p)\right]\right\}, \tag{28}$$

wherein $\operatorname{sgn}(\cdot)$ and $\Re(\cdot)$ denote the signum function and real part, respectively, and $\mathbf{H}(p)$ is a $2NN_tN_rM \times 3K$-dimensional matrix containing on its columns the discrete-time windowed signatures $\widetilde{\mathbf{D}}_{k,p-1}^{l,n_t}\widetilde{\mathbf{g}}_k^{n_t}$, $\mathbf{D}_{k,p}^{n_t}\widetilde{\mathbf{g}}_k^{n_t}$, and $\widetilde{\mathbf{D}}_{k,p+1}^{u,n_t}\widetilde{\mathbf{g}}_k^{n_t}$, for all $k = 0,\dots,K-1$, $n_t = 1,\dots,N_t$. It is worth noting that, due to the use of aperiodic spreading codes, the matrix $\mathbf{H}(p)$ depends on the temporal index p, and implementing the MMSE decision rule (28) requires a matrix inversion at each bit interval. Note that, in a CDMA system using short codes, the matrix $\mathbf{H}(p)$ is generally constant over several symbol intervals, since its variability depends on the channel impulse response variations only.

4.2. Iterative MMSE: serial interference cancellation

Since the real-time matrix inversion required by (28) may be prohibitive in some applications, it is convenient to resort to lower-complexity detection structures. To this end, note that an approximate MMSE receiver can be implemented through the use of iterative techniques. Indeed, upon letting $\mathbf{R}_{\mathbf{r}_2\mathbf{r}_2}(p) = \mathbf{H}(p)\mathbf{H}(p)^H + 2\mathcal{N}_0\mathbf{I}_{2NMN_r}$, it is easily seen that the test statistic in (28) can be written as $\widetilde{\mathbf{g}}_h^{n_tH}\mathbf{D}_{h,p}^{n_tH}\mathbf{y}(p)$, wherein $\mathbf{y}(p)$ is the solution to the following linear system:

$$\mathbf{R}_{\mathbf{r}_2\mathbf{r}_2}(p)\mathbf{y}(p) = \mathbf{r}_2(p). \tag{29}$$

As a consequence, the Gauss-Seidel iterative procedure can be used to solve the above system and to avoid the real-time matrix inversion [18]. In particular, upon letting

$$\mathbf{R}_{\mathbf{r}_2\mathbf{r}_2}(p) = \mathbf{R}_{\mathbf{r}_2\mathbf{r}_2}^U(p) + \mathbf{R}_{\mathbf{r}_2\mathbf{r}_2}^L(p) + \mathbf{R}_{\mathbf{r}_2\mathbf{r}_2}^D(p), \tag{30}$$

with $\mathbf{R}_{\mathbf{r}_2\mathbf{r}_2}^U(p)$, $\mathbf{R}_{\mathbf{r}_2\mathbf{r}_2}^L(p)$, and $\mathbf{R}_{\mathbf{r}_2\mathbf{r}_2}^D(p)$ the upper-triangular, lower-triangular, and diagonal parts of $\mathbf{R}_{\mathbf{r}_2\mathbf{r}_2}(p)$, the output of the iterative algorithm at the ℓth iteration is written as

$$\begin{aligned}\mathbf{y}^{(\ell)}(p) = &-\left(\mathbf{R}_{\mathbf{r}_2\mathbf{r}_2}^D(p) + \mathbf{R}_{\mathbf{r}_2\mathbf{r}_2}^L(p)\right)^{-1}\mathbf{R}_{\mathbf{r}_2\mathbf{r}_2}^U(p)\mathbf{y}^{(\ell-1)}(p) \\ &+ \left(\mathbf{R}_{\mathbf{r}_2\mathbf{r}_2}^D(p) + \mathbf{R}_{\mathbf{r}_2\mathbf{r}_2}^L(p)\right)^{-1}\mathbf{r}_2(p),\end{aligned} \tag{31}$$

and the estimate of the bit $b_h^{n_t}(p)$ at the ℓth iteration is written as

$$\widehat{b}_h^{(\ell),n_t}(p) = \operatorname{sgn}\left\{\Re\left[\widetilde{\mathbf{g}}_h^{n_tH}(\cdot)\mathbf{D}_{h,p}^{n_tH}\mathbf{y}^{(\ell)}(p)\right]\right\}. \tag{32}$$

Some remarks are in order on the detection rule (32). First, note that, since $\mathbf{R}_{\mathbf{r}_2\mathbf{r}_2}(p)$ is positive definite, the iterative procedure is guaranteed to converge to the MMSE multiuser receiver regardless of the starting point $\mathbf{y}^{(0)}$. Moreover, note that each iteration of the Gauss-Seidel algorithm has a quadratic, rather than cubic, computational complexity in the processing gain. Finally, note that applying the iterative Gauss-Seidel procedure is equivalent, from a multiuser detection point of view, to the adoption of a linear serial interference cancellation (SIC) receiver.

4.3. MMSE-like multiuser BLAST detection

Another possible detection strategy for multiantenna DS/CDMA systems is to devise a receiver that suppresses the multiple-access interference according to an MMSE criterion, and that then decodes the data from the transmit antennas of the user of interest through a nulling and cancellation receiver, also known as BLAST [13]. To be more precise, let us denote by $\mathbf{b}_h(p) = [b_h^1(p),\dots,b_h^{N_t}(p)]^T$ the N_t-dimensional vector containing the hth user symbols transmitted in the pth signaling interval, and by $\mathbf{H}_h(p)$ the $2N_rNM \times N_t$-dimensional matrix $\mathbf{H}_h(p) = [\mathbf{D}_{h,p}^1\widetilde{\mathbf{g}}_h^1,\dots,\mathbf{D}_{h,p}^{N_t}\widetilde{\mathbf{g}}_h^{N_t}]$. It is easy to show that the vector $\mathbf{r}_2(p)$ can be written as

$$\mathbf{r}_2(p) = \mathbf{H}_h(p)\mathbf{b}_h(p) + \mathbf{z}_h(p). \tag{33}$$

In (33) we have isolated the contribution from the vector $\mathbf{b}_h(p)$ of interest, while $\mathbf{z}_h(p)$ is the overall interference, which is made of the superposition of multiple-access interference, intersymbol interference and thermal noise. In order to suppress this interference term, the vector $\mathbf{r}_2(p)$ is processed according to the rule $\mathbf{y}_h(p) = \mathbf{D}_h^H(p)\mathbf{r}_2(p)$, wherein the matrix $\mathbf{D}_h(p)$ is $N_t \times 2N_rNM$-dimensional and solves the following constrained optimization problem:

$$\begin{aligned}\mathbf{D}_h(p) = \arg\min_{\mathbf{X}\in\mathcal{C}^{N_t\times 2N_rNM}} E\left\{\left\|\mathbf{X}^H\mathbf{r}_2(p)\right\|^2\right\},\\ \text{subject to } \mathbf{D}_h^H(p)\mathbf{H}_h(p) = \mathbf{I}_{N_t}.\end{aligned} \tag{34}$$

Applying standard Lagrangian techniques, it is easily shown that the matrix $\mathbf{D}_h(p)$ is written as

$$\mathbf{D}_h(p) = \mathbf{R}_{\mathbf{r}_2\mathbf{r}_2}^{-1}(p)\mathbf{H}_h(p)\left(\mathbf{H}_h^H(p)\mathbf{R}_{\mathbf{r}_2\mathbf{r}_2}^{-1}(p)\mathbf{H}_h(p)\right)^{-1}. \tag{35}$$

Now, assuming that the overall interference has been suppressed by the filter $\mathbf{D}_h(p)$, the N_t-dimensional vector $\mathbf{y}_h(p)$ can be approximately written as $\mathbf{y}_h(p) \approx \mathbf{b}(p) + \mathbf{D}_h^H(p)\mathbf{w}_2(p)$, that is, as the superposition of the symbols of interest and of a nonwhite Gaussian vector with covariance matrix $2\mathcal{N}_0\mathbf{D}_h^H(p)\mathbf{D}_h(p)$. Letting $\mathbf{U}_h(p)\mathbf{\Lambda}_h(p)\mathbf{U}_h^H(p)$ be the eigendecomposition of the matrix $\mathbf{D}_h^H(p)\mathbf{D}_h(p)$, the vector $\mathbf{y}_h(p)$ can be whitened through the following processing $\mathbf{y}_{h,w}(p) = \mathbf{\Lambda}_h^{-1/2}(p)\mathbf{U}_h^H(p)\mathbf{y}_h(p)$. Now, since the vector $\mathbf{y}_{h,w}(p)$ is the superposition of the useful term $\mathbf{\Lambda}_h^{-1/2}(p)\mathbf{U}_h^H(p)\mathbf{b}_h(p)$ and of additive white thermal noise, the nulling and cancellation receiver proposed in [13] can be applied in order to detect the entries of the symbol vector $\mathbf{b}_h(p)$. For the sake of brevity, we omit here further details on this receiver, since they can be easily found in the literature.

5. THE MULTIPASS STRATEGY

The multipass strategy that is explored in this paper is based on the following idea. Once the $(B-T)N_t$ data symbols for

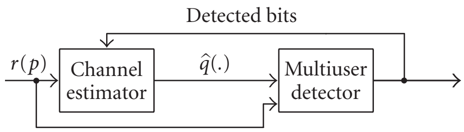

Figure 2: Block-scheme representation of the multipass strategy.

each user have been detected, they can be fed back, along with the training bits, to the channel estimation algorithm that can treat them as a fictitious training sequence of length KN_tB. Based on the knowledge of such fictitious training symbols, a new channel estimate can be thus computed. Obviously, intuition suggests that if the data symbols have been detected with a sufficiently low error probability, the new channel estimate will be much more reliable than the previous one, and, accordingly, feeding this channel estimate to the data detector will provide an even lower data error probability. If, instead, the data symbols have been detected with a large error probability, we expect that the new channel estimate will be worse than the previous one and an error propagation process may arise. Luckily enough, both analytical findings and numerical results, to be illustrated in the remainder of the paper, will confirm the following two remarkable features of the multipass strategy: (a) under many scenarios of relevant interest the proposed iterative approach is convenient even when the data symbols are detected with an error probability which is about 10^{-1}; and (b) few iterations (i.e., 2-3) between the data detector and the channel estimator are sufficient to provide huge performance gains with respect to the case that no multipass strategy is employed. A block scheme of the multipass estimator/detector is depicted in Figure 2. Obviously, any channel estimation and data detection algorithm illustrated in the previous section can be used as building blocks of the scheme in the figure.

5.1. Performance analysis

Before illustrating some numerical results, we provide a theoretical analysis and derive an approximate closed-form formula for the channel estimation mean square error (CEMSE) at a given iteration of the multipass strategy.

To begin with, we first consider the initial stage of the multipass strategy analyzing the CEMSE when only the known TN_t training bits are used for channel estimation. Substituting (16) into (21), with T in place of n, and letting $\mathbf{Q}_T = \sum_{p=1}^{T} \mathbf{F}_p^H \mathbf{F}_p$, it is easily seen that[6]

$$\widehat{\mathbf{q}}(T) = \mathbf{q} + \mathbf{Q}_T^{-1}\left(\sum_{p=1}^{T} \mathbf{F}_p^H \mathbf{w}(p)\right), \tag{36}$$

whereby we can claim that the channel estimator $\widehat{\mathbf{q}}(T)$ is unbiased and the corresponding CEMSE is given by

$$E\left[\|\widehat{\mathbf{q}}(T) - \mathbf{q}\|^2\right] = 2\mathcal{N}_0 \operatorname{trace}\left(\mathbf{Q}_T^{-1}\right). \tag{37}$$

The CEMSE can be also given a more informative approximate expression. Indeed, since in a long-code CDMA system the spreading codes are well modeled as realizations of a sequence of independent equally likely binary variates, substituting the time average of the matrices $\mathbf{F}_p^H \mathbf{F}_p$ with a statistical expectation, the following approximate formula for the CEMSE at the initial stage of the multipass strategy can be found:

$$E\left[\|\widehat{\mathbf{q}}(T) - \mathbf{q}\|^2\right] \approx 2\mathcal{N}_0 \frac{(N+1)MKN_tN_r}{NT}. \tag{38}$$

It is thus seen that, as expected, the CEMSE is a decreasing function of the number of signaling intervals devoted to training.

Let us now consider the more interesting situation that the entire frame of duration BT_b is fed back to the channel estimator. Equation (21) is now written as

$$\widehat{\mathbf{q}}(B) = \left(\sum_{p=1}^{B} \widehat{\mathbf{F}}_p^H \widehat{\mathbf{F}}_p\right)^{-1} \cdot \left(\sum_{p=1}^{B} \widehat{\mathbf{F}}_p^H \mathbf{r}(p)\right), \tag{39}$$

wherein the matrix $\widehat{\mathbf{F}}_p$ contains, for $p > T$ the detected bits, say $\widehat{b}_k^{n_t}(p)$, in lieu of the true information symbols. Letting $\widehat{\mathbf{Q}}_B = \sum_{p=1}^{B} \widehat{\mathbf{F}}_p^H \widehat{\mathbf{F}}_p$, and substituting (16) into (39), we have

$$\widehat{\mathbf{q}}(B) = \mathbf{q} + \widehat{\mathbf{Q}}_B^{-1}\left\{\sum_{p=1}^{B} \widehat{\mathbf{F}}_p^H \left[(\widehat{\mathbf{F}}_p - \mathbf{F}_p)\mathbf{q} + \mathbf{w}(p)\right]\right\}. \tag{40}$$

From the above equation, it is seen that the iterative strategy makes the channel estimate no longer unbiased. In order to come up with a closed-form formula for the CEMSE, we make the assumption of considering the statistics of the detected bits $\widehat{b}_k^{n_t}(\cdot)$ independent of the additive thermal noise, and, also, approximate the computation of fourth-order moments in terms of products of second-order moments.[7] First of all, we have to consider the term

$$E\left[\sum_{p=1}^{B} (\mathbf{F}_p - \widehat{\mathbf{F}}_p)^H \widehat{\mathbf{F}}_p\right] = E\left[\sum_{p=T+1}^{B} (\mathbf{F}_p - \widehat{\mathbf{F}}_p)^H \widehat{\mathbf{F}}_p\right]. \tag{41}$$

In order to give an informative expression to the right-hand side of (41), we note that the matrix $(\mathbf{A}_{k,p} - \widehat{\mathbf{A}}_{k,p})^H \widehat{\mathbf{A}}_{k,p}$ can be approximated as block diagonal; moreover, it can be also shown that

$$\overline{\mathbf{C}}_{k,p}^{n_t H} \overline{\mathbf{C}}_{k,p}^{n_t} \approx \operatorname{Diag}(0, 1, \ldots, N-1, N, N, N-1, \ldots, 1, 0) \otimes \mathbf{I}_M. \tag{42}$$

[6] Note that in this scenario, the considered least-squares estimator coincides with the maximum-likelihood channel estimate.

[7] Note that this last assumption is quite customary in the analysis of adaptive algorithms. Moreover, numerical results will show that both these assumptions have a negligible effect on the accuracy of the derived formulas.

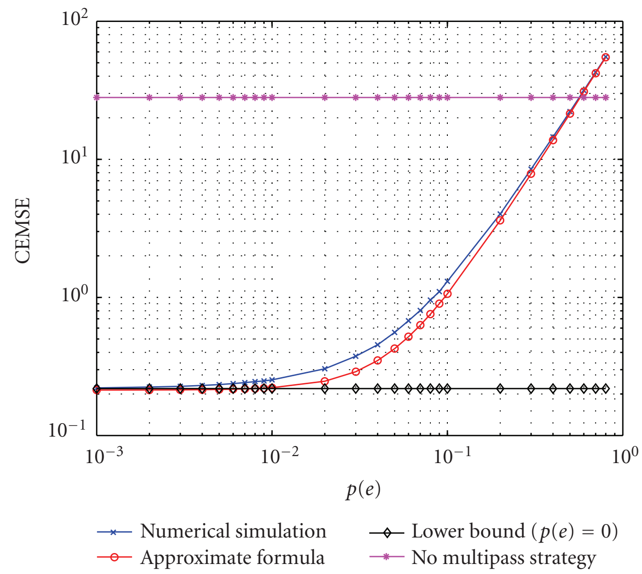

FIGURE 3: CEMSE achieved by the channel estimator versus the error probability $p(e)$ of the data detector at the previous iteration. SNR = 10 dB, $N_t = 2$, $N_r = 2$, $K = 5$, $B = 400$, $T = 15$, $N = 15$, $M = 2$.

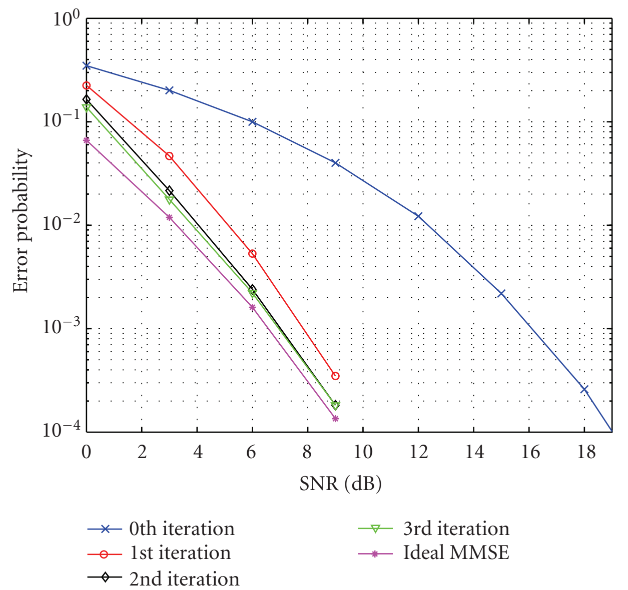

FIGURE 4: Error probability for the linear MMSE receiver versus the SNR. $N_t = 2$, $N_r = 2$, $K = 5$, $B = 400$, $T = 15$, $N = 15$, $M = 2$.

Upon denoting by $p(e)$ the bit error probability achieved by the data detector at the previous iteration, we also have that

$$E\left[\left(b_k^{n_t}(p) - \widehat{b}_k^{n_t}(p)\right)\widehat{b}_k^{n_t}(p)\right] = -2p(e). \tag{43}$$

Based on the above relations, some lengthy algebraic manipulations, not reported here for the sake of brevity, lead to

$$E\left[\sum_{p=T+1}^{B}\left(\mathbf{F}_p - \widehat{\mathbf{F}}_p\right)^H \widehat{\mathbf{F}}_p\right] \approx 2p(e)(B-T)N\mathbf{I}_{(N+1)MKN_tN_r},$$
$$E\left[\widehat{\mathbf{Q}}_B^{-1}\right] \approx \frac{1}{BN}\mathbf{I}_{(N+1)MKN_tN_r}. \tag{44}$$

Exploiting (44), it can be finally shown that the CEMSE achieved by the channel estimator exploiting information bits detected with a bit error probability $p(e)$ can be finally expressed as

$$E\left[\|\widehat{\mathbf{q}}(B) - \mathbf{q}\|^2\right] \approx 4p(e)^2\frac{(B-T)^2}{B^2}\|\mathbf{q}\|^2 + \frac{2\mathcal{N}_0}{BN}K(N+1)MN_tN_r. \tag{45}$$

It is worth noting that relation (45) is extremely powerful, in that it provides a simple expression of the CEMSE as a function of the fundamental system parameters such as the number of users, the processing gain, the number of transmit and receive antennas, and, obviously, the error probability achieved at the previous iteration. In Figure 3 we plot the approximate relation (45) versus the error probability $p(e)$; for comparison purposes, we also report the results of computer simulations, as well as the CEMSE that would be achieved in the case that all the information bits are detected with no error and the CEMSE (38) corresponding to the situation that no multipass strategy is adopted. A Rayleigh-distributed 3-path channel model has been considered; the considered system parameters are reported in the caption of the figure. The computer simulation results have been obtained by repeating 10^5 times the following procedure. First, N_tB bits are randomly generated and used to generate the discrete-time vectors $\mathbf{r}(1), \ldots, \mathbf{r}(B)$; then, random errors with probability $p(e)$ are introduced on the $N_t(B-T)$ information bits and, after that, these errored bits are fed to the channel estimator, that uses them, along with the N_tT actual training bits, to perform the channel estimate; based on the output of the channel estimator the CEMSE can be computed. Interestingly, it is seen that the experimental results are in excellent agreement with the approximate relation (45). Moreover, on one hand, it is seen that even large values of the error probability (close to 0.5) lead to a reduction of the CEMSE. On the other hand, results show that the case that $p(e) \leq 10^{-2}$ (Note that such values of the error probability may be obtained, even with an initially large CEMSE, by properly increasing the signal-to-noise ratio) permits achieving the same CEMSE that would be achieved in the ideal situation that the whole transmitted packet is known and exploited for channel estimation.

5.2. *Numerical bit error rate results*

The results of Figure 3 have shown that the performance of the channel estimation scheme can have a large benefit from the use of the multipass strategy; accordingly, it is reasonably expected that the CEMSE reduction leads to a considerable reduction in the bit error rate also. Indeed, this intuition is confirmed by the results of some computer

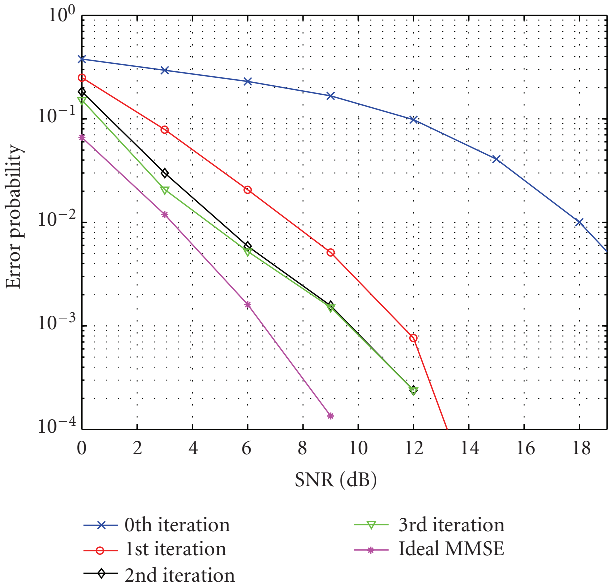

FIGURE 5: Error probability for the iterative MMSE receiver versus the SNR. $N_t = 2$, $N_r = 2$, $K = 5$, $B = 400$, $T = 15$, $N = 15$, $M = 2$.

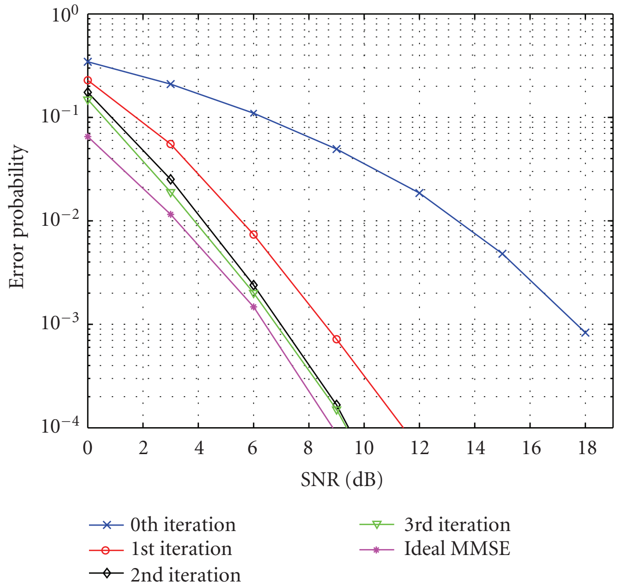

FIGURE 6: Error probability for the linear MMSE receiver versus the SNR. Each user is assigned one signature waveform, which is thus used to spread data symbols on all its transmit antennas. $N_t = 2$, $N_r = 2$, $K = 5$, $B = 400$, $T = 15$.

simulations. In Figures 4 and 5 we thus report the error probability versus the signal-to-noise ratio for the linear MMSE receiver (Section 4.1) and for the iterative MMSE receiver (Section 4.2). The considered system parameters are reported in the caption of the figures. Again we consider a Rayleigh-distributed 3-path channel model. The curves labeled as “0 iteration” correspond to the situation that no multipass strategy has been employed, while the remaining curves show the error probability after some iterations. Moreover, for comparison purposes, we also report the error probability of the ideal linear MMSE receiver, which assumes perfect knowledge of the channel vectors. It is clearly seen that the multipass strategy permits achieving a performance gain of about 10 dB, and, as regards linear MMSE detection, performs pretty close to the ideal MMSE receiver which has a perfect knowledge of the channel. As expected, it is seen that the iterative MMSE receiver performance is worse than that of the linear MMSE receiver, but, however, also for the iterative receiver the multipass strategy leads to a remarkable performance improvement. The results of Figures 4 and 5 refer to the situation that each user is assigned N_t different spreading codes, that is, one for each transmit antenna. However, the proposed channel estimation and data detection scheme does work also when just one spreading code is assigned to each user, and is used to spread the information symbols on all the transmit antennas. In Figures 6 and 7 we thus report the performance of the linear MMSE receiver (Section 4.1) and of the iterative MMSE receiver (Section 4.2), respectively, in the “same signatures” scenario. It is seen that the performance is practically coincident with the one reported in Figures 4 and 5. This is a remarkable feature of the proposed strategy. Indeed, the use of different spreading codes may lead to a spreading code shortage and, eventually, to a drastic reduction in the number of users, thus implying that the ability to use just one spreading code per user in multiantenna systems is a fundamental requisite.

Overall, results show that the multipass strategy is an effective strategy to achieve excellent performance levels with very short training sequences. Otherwise stated, the multipass strategy retains the advantages of both trained and blind systems, that is, excellent performance levels and close-to-one throughput.

5.3. *Setting of the optimal training length*

A general question in the design of wireless communication systems is how to set the amount of time devoted to training. In principle, the length of the training phase is to be chosen as a compromise between the conflicting requirements of achieving a reliable channel estimate and of not reducing too much the system throughput. As a consequence, a possible reasonable optimization strategy is to choose the training length T as the one that minimizes the following objective function:

$$\gamma(T) = \frac{E\left[||\widehat{\mathbf{q}}(\cdot) - \mathbf{q}||^2 \right]}{(B-T)/B} \tag{46}$$

which is the CEMSE-to-throughput ratio. Based on the approximate expressions (38) and (45), it is easily seen that the objective function $\gamma(T)$ is expressed as

$$\gamma(T) = 2\mathcal{N}_0 \frac{B(N+1)MKN_tN_r}{NT(B-T)}, \tag{47}$$

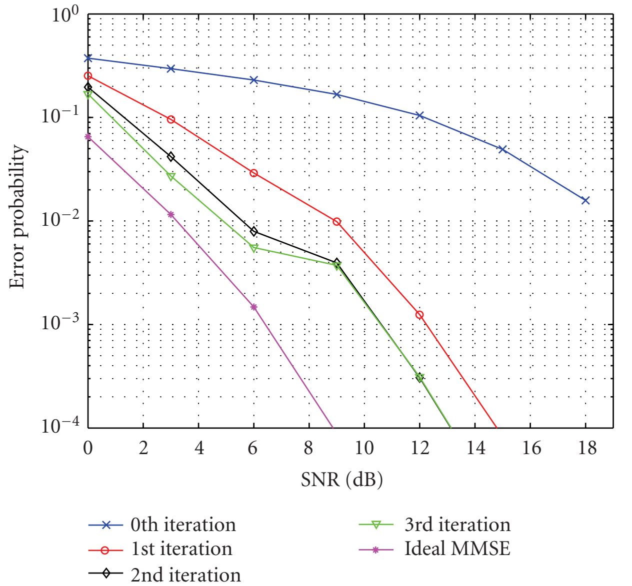

FIGURE 7: Error probability for the iterative MMSE receiver versus the SNR. Each user is assigned one signature waveform, which is thus used to spread data symbols on all its transmit antennas. $N_t = 2$, $N_r = 2$, $K = 5$, $B = 400$, $T = 15$.

for the case that the multipass strategy is not implemented, and

$$\gamma(T) = 4p(e)^2 \frac{B-T}{B} \|\mathbf{q}\|^2 + \frac{2\mathcal{N}_0}{(B-T)N} K(N+1)MN_tN_r, \quad (48)$$

when the multipass strategy is adopted. Through elementary calculus it is seen that (47) is minimum for $T = B/2$, that is, when no multipass strategy is adopted half of the time should be spent in training. As to (48), unfortunately its minimization is not trivial, since the error probability $p(e)$ depends in a complicated way on the training length T. In order, however, to be able to assess the impact of the multipass strategy on the objective function $\gamma(T)$, in Figure 8(a) we report the minimum (with respect to the training sequence length T) of $\gamma(T)$ versus the error probability $p(e)$, for both the cases that the multipass strategy is adopted and that the multipass strategy is not adopted. In this latter situation, obviously, the minimum of $\gamma(T)$ is independent of $p(e)$ and is thus represented by a straight horizontal line in the figure. In Figure 8(b), instead, the minimizer of $\gamma(T)$ is represented versus $p(e)$, for the multipass and the conventional strategy. Also in this case the considered system parameters are reported in the figures caption. Interestingly, it is seen that, in the considered scenario, it suffices to have $p(e) \leq 4 \cdot 10^{-2}$ for the multipass strategy to outperform the conventional strategy (note that such small values for the error probability can be achieved, for small training lengths, by properly increasing the signal-to-noise ratio). In particular, it is thus seen that in the region of interest of low error probabilities the multipass strategy permits achieving a smaller value of the cost function $\gamma(T)$. Moreover, and mostly important, it is seen from the lower plot that for moderately low error probabilities the optimal training sequence length coincides with its minimum value (see (20)), thus implying that, as already pointed out, the system achieves at the same time a throughput close to that of blind systems and a performance close to that of systems adopting long training sequences. The experimental results thus confirm again the huge performance gains that are granted by the use of this recursive approach.

5.4. *Multiple antennas versus single-antenna systems*

As already commented, it is well known that multiple-antenna systems are capable of achieving much better performance than single-antenna systems. It should be however noted that most studies available in the literature compare single-antenna and multiple-antenna systems under the assumption that perfect channel knowledge is available at the receiver. In practice, however, the channel realizations are to be estimated at the receiver, and, since the task of channel estimation is much more challenging for multiple-antenna systems, the question thus arises to understand whether multiple-antenna systems are still so advantageous when no channel state information is assumed. This issue has been recently tackled in the literature (see for instance [19, 20]), and it has been shown that, even if channel estimation is explicitly accounted for, using multiple-antenna systems brings performance improvements with respect to the use of a single-antenna system. A thorough theoretical discussion of this issue is well beyond the scope of this paper. However, in the following we present some simulation results that compare a single-antenna system with multiple-antenna systems. In particular, putting a constraint on the available bandwidth and on the data rate of the information stream to be conveyed, we have compared a single-antenna system employing 8 PSK modulation with a multiple-antenna system with $N_t = 3$, $N_r = 1$, and with $N_t = N_r = 2$, and with BPSK modulation. The results are shown in Figure 9, wherein the performance of the ideal MMSE receiver (i.e., assuming a known channel) for the 8 PSK single-antenna system is reported versus the performance of the proposed multipass strategy for the systems with $N_t = 3$, $N_r = 1$, and with $N_t = N_r = 3$. While the system with just one receive antenna performs slightly worse than the 8 PSK single-antenna system, it is seen that, despite the challenge of estimating 6 different channels for each user, the system equipped with 3 antennas at the transmitter and 2 at the receiver well outperforms by many dBs the single-antenna system. Other simulations, whose results are not reported here for the sake of brevity, have shown that such a performance gain is even larger when the number of receive antennas is increased. In agreement with the findings of [19, 20], it is thus experimentally shown that the use of multiple antennas is beneficial despite the increased number of channel impulse responses that are to be estimated.

6. SPACE-TIME CODED CDMA SYSTEMS

In this section we show how the proposed iterative strategy can be extended to space-time coded systems. In particular,

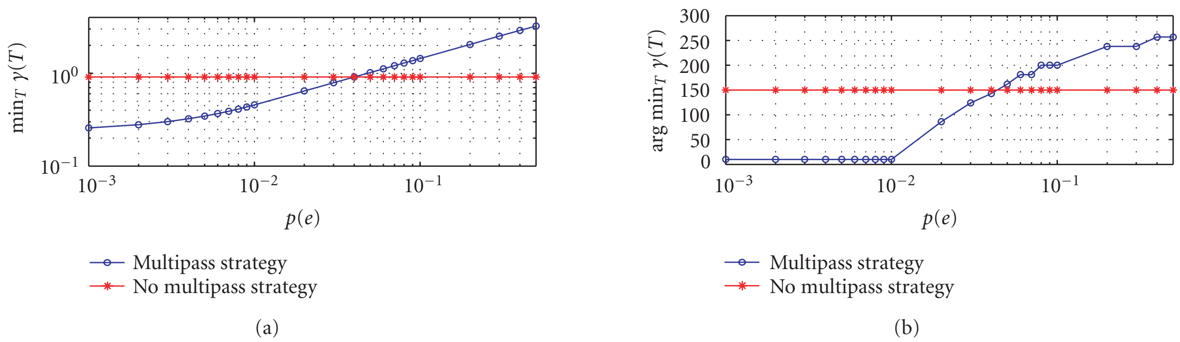

FIGURE 8: (a) Minimum of $\gamma(T)$ versus $p(e)$ for the multipass and conventional strategy. (b) Minimizer of $\gamma(T)$ versus $p(e)$ for the multipass and conventional strategy. It is seen that, for reasonable values of $p(e)$, the multipass strategy achieves lower values of the cost-function $\gamma(\cdot)$ and, also, a larger throughput. SNR = 6 dB, $N_t = 2$, $N_r = 2$, $K = 4$, $B = 300$, $N = 15$, $M = 2$.

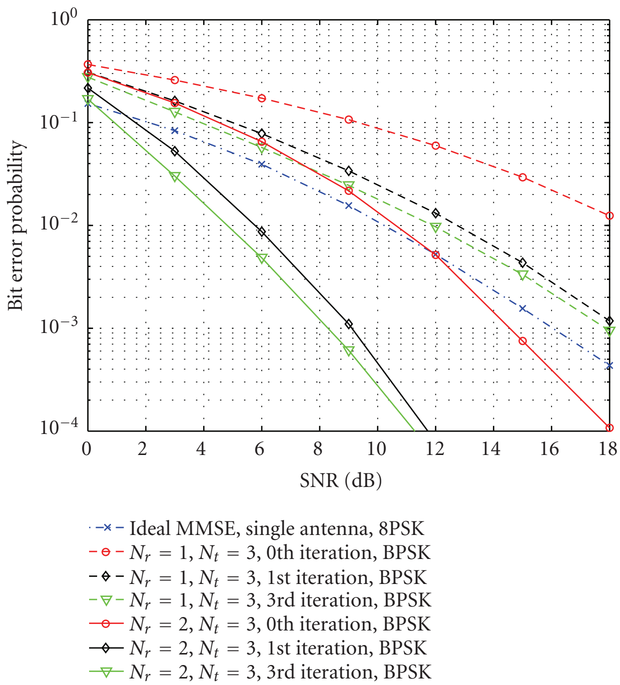

FIGURE 9: Error probability for a single-antenna system with 8 PSK modulation versus BPSK multiple-antenna systems with $N_t = 3$ and $N_r = 1, 3$. Linear MMSE Receiver, $K = 4$, $B = 200$, $T = 20$.

we focus on the popular Alamouti space-time coding scheme [17]; note, however, that our approach can be extended with moderate efforts to any orthogonal space-time code. In its basic configuration, the Alamouti scheme requires $N_t = 2$ transmit antennas and $N_r = 1$ receive antenna. Denoting by $\cdots b_k(p) b_k(p+1) b_k(p+2) \cdots$ the information stream of the kth user, in the $(2p)$th signaling interval, the symbols $b_k(2p)$ and $b_k(2p+1)$ are sent by the first and second transmit antennas of the kth user, respectively, while, in the subsequent $(2p+1)$th signaling interval, the symbols $-b_k^*(2p+1)$ and $-b_k^*(2p)$ are transmitted, each on a separate antenna (see also [17] for further details). As regards the issue of channel estimation, the structure impressed by the Alamouti scheme on the transmitted symbols has no effect on the previously outlined channel estimation procedure, whereby we just dwell on the issue of data decoding. It is easily shown that the discrete-time signal observed in the interval $\mathcal{I}_{2p}$ is given by the $2NM$-dimensional vector

$$\begin{aligned} \mathbf{r}_2(2p) = \sum_{k=0}^{K-1} \Big[&- b_k^*(2p-1)\widetilde{\mathbf{D}}_{k,2p-1}^{l,1} \\ &+ b_k(2p)\mathbf{D}_{k,2p}^{1} - b_k^*(2p+1)\widetilde{\mathbf{D}}_{k,2p+1}^{u,1}\Big]\widetilde{\mathbf{g}}_k^1 \\ &+ \Big[b_k^*(2p-2)\widetilde{\mathbf{D}}_{k,2p-1}^{l,2} + b_k(2p+1)\mathbf{D}_{k,2p}^{2} \\ &+ b_k^*(2p)\widetilde{\mathbf{D}}_{k,2p+1}^{u,2}\Big]\widetilde{\mathbf{g}}_k^2 + \mathbf{w}_2(2p), \end{aligned} \tag{49}$$

while the vector $\mathbf{r}_2(2p+1)$, observed in the interval $\mathcal{I}_{2p+1}$, is written as

$$\begin{aligned} \mathbf{r}_2(2p+1) = \sum_{k=0}^{K-1} \Big[& b_k(2p)\widetilde{\mathbf{D}}_{k,2p}^{l,1} - b_k^*(2p+1)\mathbf{D}_{k,2p+1}^{1} \\ &+ b_k^*(2p+2)\widetilde{\mathbf{D}}_{k,2p+2}^{u,1}\Big]\widetilde{\mathbf{g}}_k^1 \\ &+ \Big[b_k^*(2p+1)\widetilde{\mathbf{D}}_{k,2p}^{l,2} + b_k^*(2p)\mathbf{D}_{k,2p+1}^{2} \\ &+ b_k(2p+3)\widetilde{\mathbf{D}}_{k,2p+2}^{u,2}\Big]\widetilde{\mathbf{g}}_k^2 + \mathbf{w}_2(2p+1). \end{aligned} \tag{50}$$

Consider now the $4NM$-dimensional data vector $\mathbf{y}(p) = [\mathbf{r}_2^T(2p) \;\; \mathbf{r}_2^H(2p+1)]$. Since we have assumed that the information symbols are real, it is easily shown that the vector

$\mathbf{y}(p)$ can be written as

$$
\begin{aligned}
\mathbf{y}(p) = \sum_{k=0}^{K-1} & b_k(2p-2)\begin{bmatrix} \widetilde{\mathbf{D}}^{l,2}_{k,2p-1}\widetilde{\mathbf{g}}^2_k \\ \mathbf{0} \end{bmatrix} \\
& - b_k(2p-1)\begin{bmatrix} \widetilde{\mathbf{D}}^{l,1}_{k,2p-1}\widetilde{\mathbf{g}}^1_k \\ \mathbf{0} \end{bmatrix} \\
& + b_k(2p)\begin{bmatrix} \mathbf{D}^{1}_{k,2p}\widetilde{\mathbf{g}}^1_k + \widetilde{\mathbf{D}}^{u,2}_{k,2p+1}\widetilde{\mathbf{g}}^2_k \\ \left(\widetilde{\mathbf{D}}^{l,1}_{k,2p}\widetilde{\mathbf{g}}^1_k + \mathbf{D}^{2}_{k,2p+1}\widetilde{\mathbf{g}}^2_k\right)^* \end{bmatrix} \\
& - b_k(2p+1)\begin{bmatrix} \widetilde{\mathbf{D}}^{u,1}_{k,2p+1}\widetilde{\mathbf{g}}^1_k - \mathbf{D}^{2}_{k,2p}\widetilde{\mathbf{g}}^2_k \\ \left(\mathbf{D}^{1}_{k,2p+1}\widetilde{\mathbf{g}}^1_k - \mathbf{D}^{l,2}_{k,2p}\widetilde{\mathbf{g}}^2_k\right)^* \end{bmatrix} \\
& + b_k(2p+2)\begin{bmatrix} \mathbf{0} \\ \widetilde{\mathbf{D}}^{u,1}_{k,2p+2}\widetilde{\mathbf{g}}^1_k \end{bmatrix}^* \\
& + b_k(2p+3)\begin{bmatrix} \mathbf{0} \\ \widetilde{\mathbf{D}}^{u,2}_{k,2p+2}\widetilde{\mathbf{g}}^2_k \end{bmatrix}^* + \begin{bmatrix} \mathbf{w}_2(2p) \\ \mathbf{w}_2(2p+1)^* \end{bmatrix}.
\end{aligned}
\tag{51}
$$

Now, note that the above equation (51) has a structure which is similar to that of (27). As a consequence, the same steps that have led to the previous derivation of the MIMO multiuser detectors processing (27) can be followed in order to obtain MIMO multiuser receivers that process the data (27) in order to provide estimates of the symbols $b_k(2p)$ and $b_k(2p+1)$ for all $k = 0, \ldots, K-1$. Due to lack of space, we do not provide here more details, and just report some simulation results showing that the multipass strategy is very effective also for space-time coded multiuser systems. In Figure 10 we show the error probability of the MMSE receiver in the situation that the Alamouti space-time code and the multipass strategy are employed. The considered system parameters are reported in the figure's caption. It is seen that also in this case the multipass strategy achieves a huge performance gain in just one iteration, while, after three iterations and at an error probability of 10^{-4}, the system is less than 1 dB far from the ideal MMSE receiver that assumes perfect knowledge of the channel vectors.

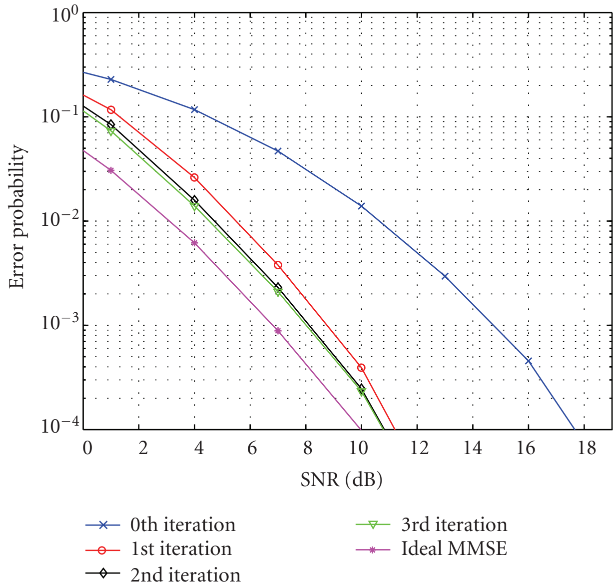

FIGURE 10: Error probability for the linear MMSE receiver versus the SNR. Alamouti space-time coded system. A Rayleigh-distributed 3-path channel model has been considered. The processing gain N is 15 and the oversampling factor M is 2. $N_t = 2$, $N_r = 1$, $K = 6$, $B = 200$, $T = 20$.

7. CONCLUSIONS

In this paper the issue of joint channel estimation and multiuser detection for long-code MIMO DS/CDMA systems operating on frequency-selective channels has been considered. Extending the results reported in [9] with regard to a single-antenna system, we have proposed to use a multipass strategy wherein the channel estimator and the data detector recursively exchange information in order to improve the system performance. It is seen that this simple strategy yields huge and impressive performance gains even when the training length is very small. In particular, we have shown, through both theoretical considerations and simulation results, that the proposed multipass strategy achieves a reduction in the CEMSE, in the system error probability, and in the optimal length of the training phase. It is worth pointing out that the multipass strategy has a general validity, and can be applied to improve performance in many wireless communication systems. As an example, the proposed strategy can be used also to perform timing and frequency synchronization. An interesting generalization of the multipass strategy is also the consideration of convolutionally coded systems. Indeed, it is expected that designing ad hoc iterative strategies wherein the data detector, the code decoder, and the channel estimator exchange soft information in a turbo-like fashion may lead to very remarkable performance gains. These issues form the object of current research.

REFERENCES

[1] E. Dahlman, P. Beming, J. Knutsson, F. Ovesjo, M. Persson, and C. Roobol, "WCDMA-the radio interface for future mobile multimedia communications," *IEEE Transactions on Vehicular Technology*, vol. 47, no. 4, pp. 1105–1118, 1998.

[2] U. Madhow, "Blind adaptive interference suppression for direct-sequence CDMA," *Proceedings of the IEEE*, vol. 86, no. 10, pp. 2049–2069, 1998.

[3] X. Wang and H. V. Poor, "Blind multiuser detection: a subspace approach," *IEEE Transactions on Information Theory*, vol. 44, no. 2, pp. 677–690, 1998.

[4] A. J. Weiss and B. Friedlander, "Channel estimation for DS-CDMA downlink with aperiodic spreading codes," *IEEE Transactions on Communications*, vol. 47, no. 10, pp. 1561–1569, 1999.

[5] Z. D. Xu and M. K. Tsatsanis, "Blind channel estimation for long code multiuser CDMA systems," *IEEE Transactions on Signal Processing*, vol. 48, no. 4, pp. 988–1001, 2000.

[6] C. J. Escudero, U. Mitra, and D. T. M. Slock, "A Toeplitz displacement method for blind multipath estimation for long code DS/CDMA signals," *IEEE Transactions on Signal Processing*, vol. 49, no. 3, pp. 654–665, 2001.

[7] V. Tripathi, A. Mantravadi, and V. V. Veeravalli, "Channel acquisition for wideband CDMA signals," *IEEE Journal on Selected Areas in Communications*, vol. 18, no. 8, pp. 1483–1494, 2000, Special issue on Wideband CDMA.

[8] S. Buzzi and H. V. Poor, "Channel estimation and multiuser detection in long-code DS/CDMA systems," *IEEE Journal on Selected Areas in Communications*, vol. 19, no. 8, pp. 1476–1487, 2001.

[9] S. Buzzi and H. V. Poor, "A multipass approach to joint data and channel estimation in long-code CDMA systems," *IEEE Transactions on Wireless Communications*, vol. 3, no. 2, pp. 612–626, 2004.

[10] I. E. Telatar, "Capacity of multi-antenna Gaussian channels," *European Transactions on Telecommunications*, vol. 10, no. 6, pp. 585–595, 1999.

[11] G. J. Foschini and M. J. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," *Wireless Personal Communications*, vol. 6, no. 3, pp. 311–335, 1998.

[12] T. L. Marzetta and B. M. Hochwald, "Capacity of a mobile multiple-antenna communication link in Rayleigh flat fading," *IEEE Transactions on Information Theory*, vol. 45, no. 1, pp. 139–157, 1999.

[13] G. J. Foschini, "Layered space-time architecture for wireless communication in a fading environment when using multi-element antennas," *Bell Labs Technical Journal*, vol. 1, no. 2, pp. 41–59, 1996.

[14] G. J. Foschini, G. D. Golden, R. A. Valenzuela, and P. W. Wolniansky, "Simplified processing for high spectral efficiency wireless communication employing multi-element arrays," *IEEE Journal on Selected Areas in Communications*, vol. 17, no. 11, pp. 1841–1852, 1999.

[15] B. M. Hochwald, T. L. Marzetta, and C. B. Papadias, "A transmitter diversity scheme for wideband CDMA systems based on space-time spreading," *IEEE Journal on Selected Areas in Communications*, vol. 19, no. 1, pp. 48–60, 2001.

[16] D. Reynolds, X. Wang, and H. V. Poor, "Blind adaptive space-time multiuser detection with multiple transmitter and receiver antennas," *IEEE Transactions on Signal Processing*, vol. 50, no. 6, pp. 1261–1276, 2002.

[17] S. M. Alamouti, "A simple transmit diversity technique for wireless communications," *IEEE Journal on Selected Areas in Communications*, vol. 16, no. 8, pp. 1451–1458, 1998.

[18] G. H. Golub and C. F. Van Loan, *Matrix Computations*, The Johns Hopkins University Press, Baltimore, Md, USA, 3rd edition, 1996.

[19] L. Mailaender, "Capacity bound for linear MIMO systems with channel estimation," in *Proceedings of IEEE 6th Workshop on Signal Processing Advances in Wireless Communications (SPAWC '05)*, pp. 1048–1052, New York, NY, USA, June 2005.

[20] T. Yoo and A. Goldsmith, "Capacity of fading MIMO channels with channel estimation error," in *Proceedings of IEEE International Conference on Communications (ICC '04)*, vol. 2, pp. 808–813, Paris, France, June 2004.

Stefano Buzzi was born in Piano di Sorrento, Italy, on December 10, 1970. He received with honors the Dr. Eng. degree in 1994, and the Ph.D. degree in electronic engineering and computer science in 1999, both from the University of Naples "Federico II." In 1996 he spent six months at Centro Studi e Laboratori Telecomunicazioni (CSELT), Turin, Italy, while from November 1999 through December 2001 he has spent eight months at the Department of Electrical Engineering, Princeton University, as a Visiting Research Fellow. He is currently an Associate Professor at the University of Cassino, Italy. His current research and study interests lie in the area of statistical signal processing, with emphasis on signal detection in non-Gaussian noise and multiple-access communications. He was awarded by the AEI (Associazione Elettrotecnica ed Elettronica Italiana) the "G. Oglietti" scholarship in 1996, and was the recipient of a NATO/CNR advanced fellowship in 1999 and of a CNR short-term mobility grant in 2000 and 2001. He is currently serving as an Associate Editor for the IEEE Signal Processing Letters, the IEEE Communications Letters, and the Journal of Communications and Networks.