EM-Based Joint Channel Estimation and Data Detection for MIMO-CDMA Systems

Ayman Assra, Walaa Hamouda, *Senior Member, IEEE*, and Amr Youssef, *Senior Member, IEEE*

***Abstract*—In this paper, we present an iterative joint channel-estimation and data-detection technique for multiple-input–multiple-output (MIMO) code-division multiple-access (CDMA) systems over Rayleigh fading channels. The proposed receiver performs channel estimation and data detection using the expectation–maximization (EM) algorithm. We derive a closed-form expression for the optimized weight coefficients of the EM algorithm, which is shown to provide large performance improvement relative to the conventional equal-weight EM-based signal decomposition. Our results show that the receiver can achieve near-optimum performance with modest complexity using very few training symbols. Furthermore, our simulation results confirm that the proposed receiver is near–far resistant and offers fast convergence in severe near–far scenarios.**

***Index Terms*—Code-division multiple-access (CDMA) systems, Cramér–Rao lower bound (CRLB), expectation–maximization, joint detection and estimation techniques, maximum likelihood (ML) solution, multiple-input–multiple-output (MIMO) systems, space–time spreading.**

I. Introduction

SUPPORTING reliable high-data-rate application is one of the main requirements of next-generation wireless communication systems. In this quest, multiple-input–multiple-output (MIMO) technology has proven to offer considerable throughput improvement without additional power consumption or bandwidth expansion [1]. Within the context of MIMO, space–time (ST) coding techniques provide spatial diversity with lower complexity as in ST block coding [2] or both spatial diversity and coding gain as in ST trellis coding [3]. The former is a generalization of the Alamouti's dual-transmit diversity scheme [4] to multiple transmit antennas, whereas the latter is a generalization of trellis coding to multiple transmit and receive antennas. With code-division multiple access (CDMA) being one of the generic multiple-access schemes in second- and third-generation wireless systems [5], the integration of CDMA with MIMO techniques has become an active area of research. One of the first ST systems that implement ST spreading (STS) in CDMA was introduced in [6]. The proposed STS system in [6] is shown to achieve full spatial diversity and maintain high spectral efficiency without wasting system resources [7]. Given these advantages, it has become a strong candidate for next-generation wireless networks, where it has been considered as part of the IS-2000 wideband CDMA standard [8].

In MIMO-CDMA systems, perfect channel knowledge is essential for efficient detection [9]. Compared with single-input–single-output (SISO) CDMA systems, channel estimation in MIMO systems becomes even more challenging as the number of simultaneous transmissions and interference levels increases. Whereas the majority of the works in MIMO systems assume perfect channel estimation, relatively few researchers have investigated the effect of channel-estimation errors and the possible estimation techniques (e.g., [9] and [10] and references therein). Commonly used channel estimation techniques either reduce the effective data rate as in training-based channel estimation [10], [11] or increase the computational complexity of the system as in blind-based channel-estimation techniques [12]–[14]. In the former, channel estimation is performed by periodically inserting known training bits among the data frames, whereas in the latter, channel estimation is implemented by exploiting the statistical characteristics of the transmitted signals.

Recently, there has been an increasing interest in iterative joint channel-estimation and data-detection algorithms [15], [16], which take the advantage of the detected data symbols in the channel-estimation process. These iterative receivers have shown enhanced performance with reasonable convergence rates using very short training sequences. Among the iterative techniques, expectation–maximization (EM) [17] has been considered in iterative receivers for its attractive features such as iteratively attaining the maximum-likelihood (ML) solution with reduced complexity [18]. Based on the observed data (i.e., the incomplete data), the ML-based estimation has high computational complexity, particularly for a large number of unknown parameters. On the other hand, the EM algorithm defines the so-called complete data set, which enjoys flexible computation of the ML estimates of unknown parameters. These complete data are related to the incomplete data through some, possibly random, mapping.

In the past, an extensive effort has been focused on employing the EM algorithm in joint detection and estimation (JDE) techniques. Considering SISO systems, Georghiades and Han [19] proposed a JDE receiver based on the EM algorithm in time-variant Rayleigh flat-fading channels. Borran and Nasiri-Kenari [20] applied the EM algorithm for synchronous CDMA systems in additive white Gaussian noise channels. Motivated

Manuscript received June 15, 2009; revised September 16, 2009 and November 17, 2009. First published December 18, 2009; current version published March 19, 2010. The review of this paper was coordinated by Prof. H.-F. Lu.

A. Assra and W. Hamouda are with the Department of Electrical and Computer Engineering, Concordia University, Montreal, QC H3G 1M8, Canada (e-mail: a_assra@ece.concordia.ca; hamouda@ece.concordia.ca).

A. Youssef is with the Concordia Institute for Information Systems Engineering, Concordia University, Montreal, QC H3G 1M8, Canada (e-mail: youssef@ciise.concordia.ca).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TVT.2009.2038925

by the EM algorithm related to the problem of parameter estimation of superimposed signals [21], Kocian and Fleury [22] investigated the application of the EM algorithm in CDMA systems over flat-fading channels. An iterative JDE receiver for direct-sequence CDMA in frequency-selective fading channels has also been proposed in [23]. Recently, there has been an interest in applying the EM algorithm in MIMO systems. For example, Cozzo and Hughes [24] have proposed a JDE technique based on the EM algorithm in flat-fading channels with multiple antennas at both the transmitter and the receiver. So and Cheng [16] have also proposed an iterative receiver for an ST trellis-coded system in frequency-selective fading channels, where channel estimation and data detection are performed using the EM algorithm.

In MIMO-CDMA systems, the channel estimation poses a challenging problem, as the performance of the system becomes sensitive to channel-estimation errors and multiple-access interference (MAI), which lead to diversity loss. Given the advantages of the EM algorithm as opposed to training-based and blind channel-estimation techniques, in this paper, we propose a JDE technique based on the EM algorithm for MIMO-CDMA systems, assuming synchronous transmission over flat-fading channels. We show the benefits of the iterative EM algorithm in terms of its convergence to the optimum ML with lower complexity and resistance to near–far effects. The proposed JDE receiver structure is derived, where we show that it can bring an optimum balance between the single-user (SU) matched filter detector and the parallel interference cancellation (PIC)-based detector. This balance is optimized through the weight coefficient parameter that is updated at each EM iteration. In the sequel, we derive the optimum weight based on the MMSE criterion. With this optimum balancing weight, the proposed JDE is shown to offer large improvement gains relative to the conventional equal-weight balancing, particularly in near–far situations. Our results show that the proposed receiver not only offers fast convergence, with a few number of iterations and very few training symbols, but also achieves full system diversity. We also assess the channel estimator in terms of the Cramér–Rao lower bound (CRLB). In particular, we prove that the estimator is asymptotically efficient with respect to the SNR, i.e., it converges to the CRLB at a high SNR.

The rest of this paper is organized as follows. The system model is described in the following section. In Section III, the EM-based ST receiver is discussed. In Section IV, we derive a closed-form expression for the optimized weight coefficients and the initialization of the EM receiver. The Cramér–Rao bound on the channel estimates is presented in Section V. Simulation results and discussions are then presented in Section VI. Finally, conclusions are drawn in Section VII.

II. System Model

Throughout this paper, we consider a transmit diversity scheme with $N = 2$ transmit antennas at the mobile user and M multiple receive antennas at the base station. We also consider the original STS scheme proposed in [6] for a synchronous K-user system over a slow flat-fading channel. The channel coefficients are, therefore, considered fixed within a block of L codewords, where each codeword has a period of $T_s = 2T_b$, and T_b denotes the bit period. The received complex low-pass equivalent signal at the mth receive antenna is given by

$$\begin{aligned} r^m(t) = \sum_{l=1}^{L}\sum_{k=1}^{K} h_{1m}^k \left[b_{k1}[l]c_{k1}(t-lT_s)+b_{k2}[l]c_{k2}(t-lT_s)\right] \\ + h_{2m}^k \left[b_{k2}[l]c_{k1}(t-lT_s) - b_{k1}[l]c_{k2}(t-lT_s)\right] + n^m(t) \quad (1) \end{aligned}$$

where $b_{k1}[l]$ and $b_{k2}[l]$ are the odd and even data streams of the kth user within the lth codeword interval. The codes $c_{k1}(t)$ and $c_{k2}(t)$ represent the kth user's spreading sequences with processing gain $2N_c$, where $N_c = T_b/T_c$ is the number of chips per bit, and T_c is the chip duration. Based on [6], $c_{k1}(t)$ and $c_{k2}(t)$ are assumed to be orthogonal. However, the effect of cross correlation among different users on the overall system performance is considered. In (1), h_{im}^k, $i = 1, 2$ is the attenuation coefficient corresponding to the kth user from the ith transmit antenna to the mth receive antenna, where $h_{im}^k = \sqrt{E_k/2}\alpha_{im}^k$, α_{im}^k is the corresponding fading channel coefficient, and E_k is the kth user transmit energy. These attenuation coefficients are modeled as independent complex Gaussian random variables with zero mean and variance σ_k^2, where $\sigma_k^2 = (E_k/2)\sigma_h^2$, and $\sigma_h^2 = 1$. The noise $n^m(t)$ is Gaussian with zero mean and variance N_o. At the receiver side, the received signal is passed through a bank of matched filters, where the received signal is correlated with the spreading codes assigned to the K users. Let $\mathbf{z}^m(l)$ represent the output vector of the matched filter bank at the mth receive antenna during the lth codeword period. Then, we have

$$\mathbf{z}^m(l) = \mathbf{R}\mathbf{B}(l)\mathbf{h}^m + \mathbf{n}^m(l) \quad (2)$$

where $\mathbf{R} = \mathbf{C}^H\mathbf{C}$ is the cross-correlation matrix, $\mathbf{C}$ is a $2N_c \times 2K$ matrix that consists of the users' code sequences defined as

$$\mathbf{C} = [\mathbf{C}_1, \mathbf{C}_2, \ldots, \mathbf{C}_K]$$

with $\mathbf{C}_k = [\mathbf{c}_{k1}, \mathbf{c}_{k2}]$, and H denotes conjugate transpose. In (2), $\mathbf{B}(l)$ represents the user data matrix within the lth period, which is defined as

$$\mathbf{B}(l) = \mathrm{diag}\left\{\mathbf{B}_1(l), \mathbf{B}_2(l), \ldots, \mathbf{B}_K(l)\right\}$$

where

$$\mathbf{B}_k(l) = \begin{bmatrix} b_{k1}[l] & b_{k2}[l] \\ b_{k2}[l] & -b_{k1}[l] \end{bmatrix}, \quad k=1,\ldots,K \quad l=1,\ldots,L$$

and $\mathbf{h}^m$ is a $(2K \times 1)$ channel-coefficient vector defined as

$$\mathbf{h}^m = \left[\mathbf{h}_1^{mT}, \mathbf{h}_2^{mT}, \ldots, \mathbf{h}_K^{mT}\right]^T$$

where $\mathbf{h}_k^m = [h_{1m}^k, h_{2m}^k]^T$, and T denotes transpose. The channel vector $\mathbf{h}^m$ is Gaussian-distributed with zero mean and covariance matrix $\mathbf{\Sigma}_{hh} = \mathrm{diag}\{\sigma_1^2, \sigma_1^2, \sigma_2^2, \sigma_2^2, \ldots, \sigma_K^2, \sigma_K^2\}$. The noise vector $\mathbf{n}^m(l)$ includes the noise samples at the output of the mth matched filter bank during the lth code period and has a Gaussian distribution with zero mean and covariance matrix $N_o\mathbf{R}$. We assume that $\{\mathbf{C}_k\}$, $k = 1, \ldots, K$ are selected under the condition that $\mathbf{R}$ is positive definite. Then, $\mathbf{R}$ can be factorized using Cholesky decomposition to $\mathbf{R} = \mathbf{F}^H\mathbf{F}$, where $\mathbf{F}$ is a unique lower triangular invertible matrix [25]. Multiplying $\mathbf{z}^m(l)$ by $(\mathbf{F}^H)^{-1}$, we obtain

$$\mathbf{y}^m(l) = \mathbf{F}\mathbf{B}(l)\mathbf{h}^m + \mathbf{n}_w^m(l) \tag{3}$$

where $\mathbf{n}_w^m(l)$ is a complex Gaussian vector with zero mean and covariance matrix $N_o\mathbf{I}_{2K}$, and $\mathbf{I}_{2K}$ is an identity matrix of dimension $2K$. Note that both $\mathbf{z}^m(l)$ and $\mathbf{y}^m(l)$ have the same information about the transmitted data. Due to the whitening noise property of (3), our subsequent analysis will be based on $\mathbf{y}^m(l)$.

III. Expectation–Maximization-Based ST Receiver

Our subsequent analysis is based on the approach proposed in [21] for the estimation problem of superimposed signals. Using this approach, the observed data are decomposed into their signal components. Then, the parameters of each signal component are separately and iteratively estimated using the EM algorithm. Accordingly, we decompose the whitening filter output $\mathbf{y}^m(l)$ into a sum of K statistically independent components, i.e.,

$$\mathbf{y}^m(l) = \sum_{k=1}^{K} \mathbf{g}_k^m(l) \tag{4}$$

where $\mathbf{g}_k^m(l) = \mathbf{F}_k\mathbf{B}_k(l)\mathbf{h}_k^m + \mathbf{n}_{wk}^m(l)$, $\mathbf{F}_k$ is a $2K \times 2$ vector, including the two columns corresponding to the kth user in the matrix $\mathbf{F}$, $\mathbf{n}_{wk}^m(l)$ is a complex Gaussian vector with zero mean and covariance matrix $\beta_k^m N_o\mathbf{I}_{2K}$, and β_k^m is a nonnegative value satisfying the constraint $\sum_{k=1}^{K}\beta_k^m = 1$. Our goal is to obtain the users' data estimates using the EM algorithm. First, we define the EM algorithm parameters.

1) The *observed data* $\mathbf{y}$, which include the outputs of the M whitening matched filters within a frame of L codes, are given by

$$\mathbf{y} = \left[\mathbf{y}^{1T}, \mathbf{y}^{2T}, \ldots, \mathbf{y}^{MT}\right]^T$$

where

$$\mathbf{y}^m = \left[\mathbf{y}^m(1)^T, \mathbf{y}^m(2)^T, \ldots, \mathbf{y}^m(L)^T\right]^T, \quad m = 1, \ldots, M.$$

2) The *parameters to be estimated* $\mathbf{b}$ include the transmitted data bits from the K users within the frame period, i.e.,

$$\mathbf{b} = \left[\mathbf{b}_1^T, \mathbf{b}_2^T, \ldots, \mathbf{b}_K^T\right]^T$$

where

$$\mathbf{b}_k = \left[\mathbf{b}_k(1)^T, \mathbf{b}_k(2)^T, \ldots, \mathbf{b}_k(L)^T\right]^T, \quad k = 1, \ldots, K$$

and $\mathbf{b}_k(l) = [b_{k1}[l], b_{k2}[l]]^T, l = 1, \ldots, L$.

3) *Complete data* $\mathbf{G}$: We employ the complete data definition in [22], where the unknown channel coefficient vectors are included as a part of the complete data as follows:

$$\mathbf{G} = \left[\mathbf{G}^{1T}, \mathbf{G}^{2T}, \ldots, \mathbf{G}^{MT}\right]^T$$

where

$$\mathbf{G}^m = \left[\left(\mathbf{g}_1^m, \mathbf{h}_1^m\right), \left(\mathbf{g}_2^m, \mathbf{h}_2^m\right), \ldots, \left(\mathbf{g}_K^m, \mathbf{h}_K^m\right)\right]$$
$$m = 1, \ldots, M$$
$$\mathbf{g}_k^m = \left[\mathbf{g}_k^m(1), \mathbf{g}_k^m(2), \ldots, \mathbf{g}_k^m(L)\right], \quad k = 1, \ldots, K.$$

Since the components of $\mathbf{G}$ given $\mathbf{b}$ are statistically independent, the complete log-likelihood function is given by

$$\Phi(\mathbf{G}|\mathbf{b}) = \sum_{m=1}^{M}\sum_{k=1}^{K}\Phi\left(\mathbf{g}_k^m, \mathbf{h}_k^m|\mathbf{b}_k\right) \tag{5}$$

where

$$\Phi\left(\mathbf{g}_k^m, \mathbf{h}_k^m|\mathbf{b}_k\right) = \Phi\left(\mathbf{g}_k^m|\mathbf{h}_k^m, \mathbf{b}_k\right) + \Phi\left(\mathbf{h}_k^m|\mathbf{b}_k\right). \tag{6}$$

The second summand in (6) is neglected as it is independent of $\mathbf{b}$. Therefore, (6) is reduced to

$$\Phi\left(\mathbf{g}_k^m, \mathbf{h}_k^m|\mathbf{b}_k\right) \propto \sum_{l=1}^{L}\left(\mathbf{g}_k^m(l) - \mathbf{F}_k\mathbf{B}_k(l)\mathbf{h}_k^m\right)^H \times \left(\mathbf{g}_k^m(l) - \mathbf{F}_k\mathbf{B}_k(l)\mathbf{h}_k^m\right). \tag{7}$$

By neglecting the terms in (7), which are independent of $\mathbf{b}$, the conditional likelihood in (7) can be simplified to

$$\Phi\left(\mathbf{g}_k^m, \mathbf{h}_k^m|\mathbf{b}_k\right) \propto \sum_{l=1}^{L} Re\left\{\mathbf{h}_k^{m^H}\mathbf{B}_k(l)\mathbf{F}_k^H\mathbf{g}_k^m(l)\right\} \tag{8}$$

where $Re\{\cdot\}$ denotes the real part of the argument. Notice that, due to the orthogonality assumption between the two codes assigned to each user [6], $\mathbf{F}_k^H\mathbf{F}_k$ is reduced to an identity matrix of dimension 2, i.e., $\mathbf{I}_2$, and consequently, $\mathbf{B}_k(l)^H\mathbf{F}_k^H\mathbf{F}_k\mathbf{B}_k(l) = 2\mathbf{I}_2$. At the ith iteration, the E-step of the EM algorithm is implemented by taking the expectation of the complete log-likelihood function defined in (5) with respect to the observed data vector $\mathbf{y}$ and the current EM data estimates $\mathbf{b}^i$, i.e.,

$$\mathcal{Q}(\mathbf{b}|\mathbf{b}^i) = \sum_{k=1}^{K}\mathcal{Q}_k(\mathbf{b}_k|\mathbf{b}^i) \tag{9}$$

where

$$\mathcal{Q}_k(\mathbf{b}_k|\mathbf{b}^i) = \sum_{m=1}^{M} E\left[\Phi\left(\mathbf{g}_k^m, \mathbf{h}_k^m|\mathbf{b}_k\right)|\mathbf{y}, \mathbf{b}^i\right] \tag{10}$$

and E represents the expectation. From (8), the expectation of the individual log-likelihood function is reduced to

$$\mathcal{Q}_k(\mathbf{b}_k|\mathbf{b}^i) = \sum_{m=1}^{M}\sum_{l=1}^{L} Re\left\{E\left[\mathbf{h}_k^{m^H}\mathbf{B}_k(l)\mathbf{F}_k^H\mathbf{g}_k^m(l)|\mathbf{y}, \mathbf{b}^i\right]\right\}. \tag{11}$$

To find the joint conditional expectation in (11), we evaluate $E[\mathbf{g}_k^m(l)|\mathbf{y}, \mathbf{b}^i, \mathbf{h}]$, where $\mathbf{h} = [\mathbf{h}^1, \mathbf{h}^2, \ldots, \mathbf{h}^M]$. Then, the subsequent expression is used to find $E[f(\mathbf{h}_k^m)|\mathbf{y}, \mathbf{b}^i]$, where $f(\mathbf{h}_k^m)$ denotes the resultant function in $\mathbf{h}_k^m$. By noting that the conditional probability density function $P(\mathbf{g}_k^m(l)|\mathbf{y}, \mathbf{b}^i, \mathbf{h})$ is Gaussian with mean [21]

$$E\left[\mathbf{g}_k^m(l)|\mathbf{y}, \mathbf{b}^i, \mathbf{h}\right] = \mathbf{F}_k \mathbf{B}_k(l)^i \mathbf{h}_k^m + \beta_k^m \times \left(\mathbf{y}^m(l) - \sum_{j=1}^{K} \mathbf{F}_j \mathbf{B}_j(l)^i \mathbf{h}_j^m\right) \quad (12)$$

the conditional expectation of the likelihood function in (11), after some algebraic manipulations, can be expressed as (see Appendix A)

$$\mathcal{Q}_k(\mathbf{b}_k|\mathbf{b}^i) = \sum_{l=1}^{L} \mathcal{Q}_k\left(\mathbf{b}_k(l)|\mathbf{b}^i\right) \quad (13)$$

where

$$\begin{aligned}\mathcal{Q}_k(\mathbf{b}_k(l)|\mathbf{b}^i) &= \sum_{m=1}^{M} Re\Bigg\{(1-\beta_k^m)\Big(a_{11}(l)^i|\left(h_{1m}^k\right)^i|^2 \\ &\quad + a_{12}(l)^i\left(h_{1m}^k\right)^{i^*}\left(h_{2m}^k\right)^i \\ &\quad + a_{21}(l)^i\left(h_{2m}^k\right)^{i^*}\left(h_{1m}^k\right)^i \\ &\quad + a_{22}(l)^i|\left(h_{2m}^k\right)^i|^2\Big) \\ &\quad + \beta_k^m \left(\mathbf{h}_k^m\right)^{i^H}\mathbf{B}_k(l) \\ &\quad \times \left(\mathbf{z}_k^m(l) - \sum_{j=1, j\neq k}^{K} \mathbf{R}_{kj}\mathbf{B}_j(l)^i\left(\mathbf{h}_j^m\right)^i\right)\Bigg\}.\end{aligned} \quad (14)$$

$a_{qv}(l)^i$, q, $v \in \{1, 2\}$, are defined in terms of the current and next data estimates (see Appendix A), and $\mathbf{R}_{kj}$ is the 2×2 cross-correlation matrix corresponding to the two spreading codes assigned to users k and j. In (14), the conditional expectation of the attenuation coefficients given $\mathbf{y}$ and $\mathbf{b}^i$ is given by

$$\left(h_{qm}^k\right)^i = E[\mathbf{h}^m|\mathbf{y}, \mathbf{b}^i]_{2(k-1)+q} = \left[(\mathbf{h}^m)^i\right]_{2(k-1)+q} \quad (15)$$

$$\left(h_{qm}^{k*} h_{vm}^j q\right)^i = \left(h_{qm}^k\right)^{i^*}\left(h_{vm}^j\right)^i + \left(\mathbf{\Omega}_{hh}^i\right)_{2(j-1)+v, 2(k-1)+q} \quad (16)$$

where $q, v \in \{1, 2\}$, $k, j \in \{1, \ldots, K\}$, and

$$\mathbf{\Omega}_{hh}^i = E\left[\left(\mathbf{h}^m - (\mathbf{h}^m)^i\right)\left(\mathbf{h}^m - (\mathbf{h}^m)^i\right)^H|\mathbf{y}, \mathbf{b}^i\right]. \quad (17)$$

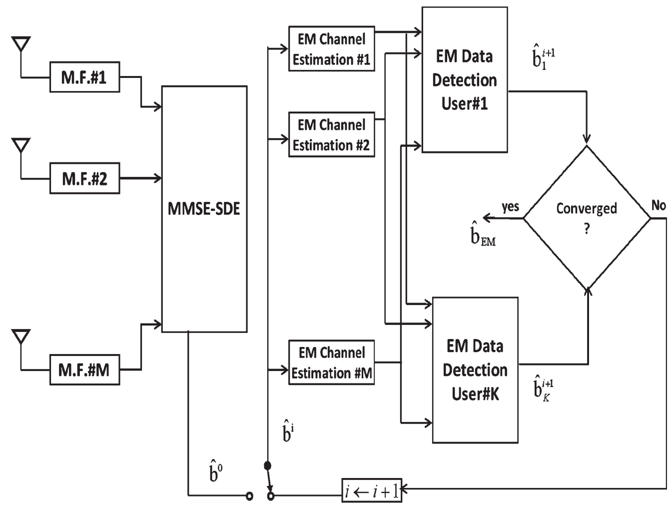

Fig. 1. ST JDE receiver based on the EM algorithm.

In Appendix B, we prove that the conditional distribution of the channel vector $\mathbf{h}^m$ given $\mathbf{y}$ and $\mathbf{b}^i$ is Gaussian with mean

$$(\mathbf{h}^m)^i = \left(\sum_{l=1}^{L} \mathbf{B}(l)^i \mathbf{R} \mathbf{B}(l)^i + N_o \mathbf{\Sigma}_{hh}^{-1}\right)^{-1} \sum_{l=1}^{L} \mathbf{B}(l)^i \mathbf{z}^m(l) \quad (18)$$

and covariance

$$\mathbf{\Omega}_{hh}^i = \left(N_o^{-1} \sum_{l=1}^{L} \mathbf{B}(l)^i \mathbf{R} \mathbf{B}(l)^i + \mathbf{\Sigma}_{hh}^{-1}\right)^{-1}. \quad (19)$$

From (9), the M-step of the EM algorithm is performed by maximizing the individual likelihood functions $\mathcal{Q}_k(\mathbf{b}_k|\mathbf{b}^i)$, $k = 1, \ldots, K$ as

$$\mathbf{b}_k^{i+1} = \arg\max_{\mathbf{b}_k} \mathcal{Q}_k(\mathbf{b}_k|\mathbf{b}^i). \quad (20)$$

Furthermore, as the codewords of each user are statistically independent, each component of $\mathbf{b}_k^{i+1}$ can be separately obtained by maximizing the corresponding summand in (13), i.e.,

$$\mathbf{b}_k^{i+1}(l) = \arg \max_{\{b_{k1}(l), b_{k2}(l)\}} \mathcal{Q}_k\left(\mathbf{b}_k(l)|\mathbf{b}^i\right). \quad (21)$$

Considering binary phase-shift keying transmission, the maximization is performed over four possibilities for the kth user data bits $b_{k1}(l)$ and $b_{k2}(l)$, i.e., $\{(1, 1), (1, -1), (-1, 1), (-1, -1)\}$. From the likelihood function in (14), we notice that the EM-based ST receiver can be interpreted as follows: The channel coefficients of the K users are estimated based on the observed data $\mathbf{y}$ and the previous data estimates $\mathbf{b}^i$. The data bits of each user are then detected from (14) based on the balancing weight β_k^m between the ST parallel interference cancellation receiver and the ST SU coherent detector. We can also notice that by using the EM algorithm, the K-user optimization problem is converted into K parallel SU optimization problems leading to low computational complexity. A block diagram of the proposed ST EM-JDE receiver is shown in Fig. 1.

IV. Expectation–Maximization Optimized Weights and Initialization

Here, we derive the optimized weights of the EM algorithm to ensure optimum performance. Also, the conditions on the EM initialization are derived and discussed in detail.

A. Optimized Weights (β_k^m)

As discussed before, the decoupling of the received superimposed signal involved in the EM algorithm heavily depends on the choice of the balancing weight at each iteration β_k^m. Here, we obtain an optimized weight that, as will be shown later, brings the performance of the EM receiver close to the SU bound.

In (14), we notice that $Q_k(\mathbf{b}_k(l)|\mathbf{b}^i)$ is a sum of M statistically independent terms given $\mathbf{b}$ and its EM estimate $\mathbf{b}^i$, which are related to the M receive antennas. Since the spatial channels corresponding to the links between transmit and receive antennas are considered independent, β_k^m can be separately optimized. In this case, we choose the weight coefficients to minimize the linear MSE between the true signal vector $\mathbf{d}_k^m(l) = \mathbf{F}_k\mathbf{B}_k(l)\mathbf{h}_k^m$ and its estimate $(\mathbf{d}_k^m(l))^i = E[\mathbf{g}_k^m(l)|\mathbf{y},\mathbf{b}^i]$ as

$$\beta_k^m = \arg\min_{\beta_k^m} E\left[\left\|(\mathbf{d}_k^m(l))^i - \mathbf{d}_k^m(l)\right\|^2\right] \tag{22}$$

where $\|\cdot\|$ denotes the vector norm. Taking the expectation of (12) with respect to $\mathbf{h}^m$, we have

$$(\mathbf{d}_k^m(l))^i = \mathbf{F}_k\mathbf{B}_k(l)^i\,(\mathbf{h}_k^m)^i + \beta_k^m \times\left(\mathbf{y}^m(l) - \sum_{j=1}^{K}\mathbf{F}_j\mathbf{B}_j(l)^i(\mathbf{h}_j)^i\right). \tag{23}$$

To simplify our analysis, we assume that $\mathbf{n}_w = 0$ and $L \to \infty$. As proved in Appendix C, according to these asymptotic assumptions, the estimate $(\mathbf{h}^m)^i$ is consistent, and it can be considered known at the receiver. It follows that

$$(\mathbf{d}_k^m(l))^i = \mathbf{F}_k\mathbf{B}_k(l)^i\mathbf{h}_k^m + \beta_k^m \times\left(\sum_{j=1}^{K}\mathbf{F}_j\left(\mathbf{B}_j(l) - \mathbf{B}_j(l)^i\right)\mathbf{h}_j^m\right). \tag{24}$$

Substituting both $\mathbf{d}_k^m(l)$ and $(\mathbf{d}_k^m(l))^i$ in (22) and by removing the terms that are independent of β_k^m, we have

$$\beta_k^m = \arg\min_{\beta_k^m}\Bigg\{-2\beta_k^m Re\Bigg\{E\Bigg[\mathbf{h}_k^{m^H}\left(\mathbf{B}_k(l) - \mathbf{B}_k(l)^i\right) \times\sum_{j=1}^{K}\mathbf{R}_{kj}\left(\mathbf{B}_j(l) - \mathbf{B}_j(l)^i\right)\mathbf{h}_j^m\Bigg]\Bigg\} + (\beta_k^m)^2 \times E\left[\left\|\sum_{j=1}^{K}\mathbf{F}_j\left(\mathbf{B}_j(l) - \mathbf{B}_j(l)^i\right)\mathbf{h}_j^m\right\|^2\right]\Bigg\}. \tag{25}$$

Since

$$E\left[h_{qm}^{k^*}h_{vm}^{j}\right] = \begin{cases}\sigma_k^2, & j = k, q = v\\ 0, & \text{otherwise}\end{cases}$$

where $j,\ k \in \{1,\ldots,K\}$ and $q,\ v \in \{1,2\}$, the optimized weight in (25) reduces to

$$\beta_k^m = \arg\min_{\beta_k^m}\Bigg\{-4\beta_k^m\sigma_k^2 E\left[\left(\left(b_{k1}(l) - b_{k1}(l)^i\right)^2 + \left(b_{k2}(l) - b_{k2}(l)^i\right)^2\right)\right] + 2(\beta_k^m)^2 \times\sum_{j=1}^{K}\sigma_j^2 E\left[\left(\left(b_{j1}(l) - b_{j1}(l)^i\right)^2 + \left(b_{j2}(l) - b_{j2}(l)^i\right)^2\right)\right]\Bigg\}. \tag{26}$$

We notice that

$$\begin{aligned}E\left[\left(b_{jq}(l) - b_{jq}(l)^i\right)^2\right] &= 2\left(1 - E\left[b_{jq}(l)b_{jq}(l)^i\right]\right)\\ &= 4P_{e_j}^{m,i}, j = 1,\ldots,k, q = 1,2\end{aligned} \tag{27}$$

where the probability of error $P_{e_j}^{m,i} = P(b_{j1}(l) \neq b_{j1}(l)^i) = P(b_{j2}(l) \neq b_{j2}(l)^i)$. Using (27), (26) can be expressed as

$$\beta_k^m = \arg\min_{\beta_k^m}\left\{16\left((\beta_k^m)^2 - 2\beta_k^m\right)\sigma_k^2 P_{e_k}^{m,i} + 16(\beta_k^m)^2\sum_{j=1,j\neq k}^{K}\sigma_j^2 P_{e_j}^{m,i}\right\}. \tag{28}$$

By differentiating (28) with respect to β_k^m, we obtain the MMSE optimum balancing weight at the ith EM iteration as

$$(\beta_k^m)^i = \frac{\sigma_k^2 P_{e_k}^{m,i}}{\sum_{j=1}^{K}\sigma_j^2 P_{e_j}^{m,i}}. \tag{29}$$

Suppose that the performance of the EM-based ST receiver with $M = 1$ receive antenna will converge to the SU bound with a known channel, which is given by

$$P_{e,\mathrm{SU}} = Q\left(\sqrt{\frac{2\left(\left|h_{1m}^k\right|^2 + \left|h_{2m}^k\right|^2\right)}{N_o}}\right)$$

where $Q(x) = 1/\sqrt{2\pi}\int_x^\infty e^{-x^2/2}dx$.

Substituting with the EM channel estimates defined in (18) in the SU bound, i.e., $P_{e,\mathrm{SU}}$, we obtain an approximation for $P_{e_k}^{m,i}$, i.e.,

$$P_{e_k}^{m,i} \approx Q\left(\sqrt{\frac{2\left(\left|\left(h_{1m}^k\right)^i\right|^2+\left|\left(h_{2m}^k\right)^i\right|^2\right)}{N_o}}\right).$$

The importance of the optimized weight coefficients arises from the fact that it determines the best balance between the SU matched filter detector and the ST PIC-based detector. In the literature, the partial PIC has proven to be near–far resistant, where it achieves performance that is close to the ML detector [26].

B. EM Initialization

Since the EM algorithm is sensitive to the initialization of the parameters to be estimated [27], as well as due to the high computational complexity of joint estimation and detection in MIMO systems, we assume that our proposed EM-based ST receiver is initialized by reliable estimates. This will guarantee that the performance of our proposed receiver converges to the global maximum of the likelihood function with fast convergence.

We assume that our proposed receiver is initialized using the ST MMSE separate detection and estimation (ST MMSE-SDE) technique. The MMSE-SDE receiver was first proposed in [22] for SISO systems. Here, we extend this work to MIMO-CDMA systems as follows. To estimate the channel, we assume that each user transmits p training codewords known at the receiver, i.e., $2p$ bits. Let $\mathbf{z}_p^m$ include the output of the mth matched filter bank within a frame of p codewords. Based on $\mathbf{z}_p^m$, we can estimate the channel vector at each receive antenna $\mathbf{h}_{\mathrm{MMSE}}^m$ [28]. Then, following the same procedure, the MMSE data estimate is obtained based on $\mathbf{z}^M$ while assuming $\mathbf{h}=\mathbf{h}_{\mathrm{MMSE}}$ [29], where $\mathbf{z}^M$ represents the output of the M matched filter banks within a frame of L codewords.

V. Cramér–Rao Lower Bound on Channel Estimates

In the search for minimum-variance unbiased estimators, the CRLB is commonly used in estimation theory to assess the accuracy of the estimator in terms of its error variance. In Appendix C, we prove that the estimate is unbiased. To formulate the CRLB, the received signal $r^m(t)$ in (1) is sampled at the chip rate $R_c = 1/T_c$, and then, we collect $2N_c$ samples of the received signal corresponding to the T_s transmission period. Using vector notation, we have

$$\mathbf{r}^m(l) = \mathbf{C}\mathbf{B}(l)\mathbf{h}^m + \mathbf{N}^m(l), \quad l=1,\ldots,L;\ m=1,\ldots,M. \tag{30}$$

Let

$$\mathbf{r} = [\mathbf{r}(1), \mathbf{r}(2), \ldots, \mathbf{r}(L)] \tag{31}$$

where

$$\mathbf{r}(l) = \left[\mathbf{r}^1(l), \mathbf{r}^2(l), \ldots, \mathbf{r}^M(l)\right], \quad l=1,\ldots,L. \tag{32}$$

Then, the log-likelihood function

$$\Phi(\mathbf{r}|\mathbf{h},\mathbf{b}) \propto -\frac{1}{N_o}\left(\sum_{l=1}^{L}\sum_{m=1}^{M}\left(\mathbf{r}^m(l)-\mathbf{C}\mathbf{B}(l)\mathbf{h}^m\right)^H \times \left(\mathbf{r}^m(l)-\mathbf{C}\mathbf{B}(l)\mathbf{h}^m\right)\right). \tag{33}$$

Neglecting the terms that are independent of the channel vector $\mathbf{h}$, we have

$$\Phi(\mathbf{r}|\mathbf{h},\mathbf{b}) = -\frac{1}{N_o}\left(\sum_{l=1}^{L}\sum_{m=1}^{M} -2Re\left\{\mathbf{h}^{m^H}\mathbf{B}(l)\mathbf{z}^m(l)\right\} + \mathbf{h}^{m^H}\mathbf{B}(l)^H\mathbf{R}\mathbf{B}(l)\mathbf{h}^m\right). \tag{34}$$

Focusing on the channel estimates, we assume that the data vector $\mathbf{b}$ is known *a priori* or has been correctly detected. The CRLB provides a lower bound on the MSE of the channel estimates as follows:

$$E\left\{\left|\left(h_{qm}^k\right)^i - h_{qm}^k\right|^2\right\} \geq -\frac{\left\{\frac{\partial}{\partial h_{qm}^k}E\left[\left(h_{qm}^k\right)^i\right]\right\}^2}{E\left\{\frac{\partial^2}{\partial h_{qm}^{k\,2}}\Phi(\mathbf{r}|\mathbf{h})\right\}} \quad q=1,2;\ k=1,\ldots,K. \tag{35}$$

To evaluate $(\partial/\partial h_{qm}^k)E[(h_{qm}^k)^i]$, we replace $\mathbf{B}(l)^i$ by $\mathbf{B}(l)$ and substitute (2) in (18), yielding

$$E\left[(\mathbf{h}^m)^i\right] = \boldsymbol{\mathcal{F}}\mathbf{h}^m \tag{36}$$

where $\boldsymbol{\mathcal{F}}{=}E[(\sum_{l=1}^{L}\mathbf{B}(l)\mathbf{R}\mathbf{B}(l){+}N_o\boldsymbol{\Sigma}_{hh}^{-1})^{-1}\sum_{l=1}^{L}\mathbf{B}(l)\mathbf{R}\mathbf{B}(l)]$. Let us consider the asymptotic case when the average SNR of each user increases, i.e., $(\sigma_k^2/N_o)\to\infty$ for $k=1,\ldots,K$. In this case, $\boldsymbol{\mathcal{F}}$ becomes a unitary matrix, and

$$E\left[(\mathbf{h}^m)^i\right] = \mathbf{h}^m \tag{37}$$

proving that the EM estimates are asymptotically unbiased. Consequently, (35) is reduced to

$$E\left\{\left|\left(h_{qm}^k\right)^i - h_{qm}^k\right|^2\right\} \geq -\frac{1}{E\left\{\frac{\partial^2}{\partial h_{qm}^{k\,2}}\Phi(\mathbf{r}|\mathbf{h})\right\}}. \tag{38}$$

It is convenient to split the channel estimates into their real and imaginary components, i.e., $h_{qm}^k = h_{qm,r}^k + jh_{qm,i}^k$. Then, the CRLB can be computed as

$$E\left\{\left|\left(h_{qm}^k\right)^i - h_{qm}^k\right|^2\right\} \geq \frac{-1}{E\left\{\frac{\partial^2}{\partial h_{qm,r}^{k\,2}}\Phi(\mathbf{r}|\mathbf{h})\right\}} + \frac{-1}{E\left\{\frac{\partial^2}{\partial h_{qm,i}^{k\,2}}\Phi(\mathbf{r}|\mathbf{h})\right\}}. \tag{39}$$

By computing the second derivatives of (34) with respect to $h_{qm,r}^k$ and $h_{qm,i}^k$ and expressing the channel variables as

$h_{qm}^k = \sqrt{E_k/2}\alpha_{qm}^k$, we obtain the MSE of the EM channel estimates as follows:

$$E\left\{\left|\left(\alpha_{qm}^k\right)^i - \alpha_{qm}^k\right|^2\right\} \geq \frac{N_o}{E_k L}. \tag{40}$$

In (40), it is noted that the power of the estimation error is inversely proportional to both the SNR and the length of the observation window.

VI. Simulation Results

Here, we examine the bit-error-rate (BER) performance of the proposed EM-JDE receiver in MIMO-CDMA systems. In all cases, we consider a system with two transmit and $M = 1, 2$ receive antennas. We also consider an uplink synchronous transmission of a data block of 20 codewords $(L = 20)$ over a flat-fading channel. Without loss of generality, we consider a five-user system with the first user as the desired one. A training codeword of two training bits is used for the initialization of the EM receiver. We design the cross-correlation matrix to satisfy the standard form of STS systems [6], [30]. Therefore, $\mathbf{R}$ is represented by

$$\mathbf{R} = \begin{bmatrix} \mathbf{R}_{11} & \dots & \mathbf{R}_{1K} \\ \vdots & \dots & \vdots \\ \mathbf{R}_{K1} & \dots & \mathbf{R}_{KK} \end{bmatrix}$$

where

$$\mathbf{R}_{kj} = \begin{bmatrix} \rho^{kj} & 0 \\ 0 & \rho^{kj} \end{bmatrix}, k, j = 1, \dots, K$$

$$\rho^{kj} = \begin{cases} 1 & j = k \\ \rho = 0.3 & j \neq k. \end{cases}$$

Without loss of generality, throughout all our simulation results, we consider the average BER of the first user BER_1. As reference, we include the BER performance of the ST MMSE-SDE receiver with perfect channel estimation. Fig. 2 presents the BER performance of the proposed ST EM-JDE and ST MMSE-SDE receivers using 2×1 antenna configuration. As a reference, the results are compared with the BER performance of the SU system assuming an unknown channel. To show that our proposed receiver attains the full system diversity, we compare our results with a maximal ratio combiner (MRC) with the same number of diversity branches assuming perfect channel estimation. The MRC represents an optimal combiner for a receive diversity scheme of one transmit and multiple receive antennas [31]. Considering the STS system with two transmit antennas and one receive antenna, it seems that the EM-based ST receiver attains the full system diversity, i.e., same as an MRC with two diversity branches. The same conclusion also follows from Fig. 3.

In Figs. 4 and 5, we examine the near–far effect property of the proposed receiver for $M = 1, 2$ receive antennas, respectively. We fix the received SNR of the first user γ_1 at 16 and 8 dB for $M = 1, 2$, respectively, while the interfering users have equal SNRs relative to γ_1, varying from -10 to 60 dB. We also compare the performance of the ST EM-JDE receiver considering optimum β_k^m values (29) and equal β_k^m values

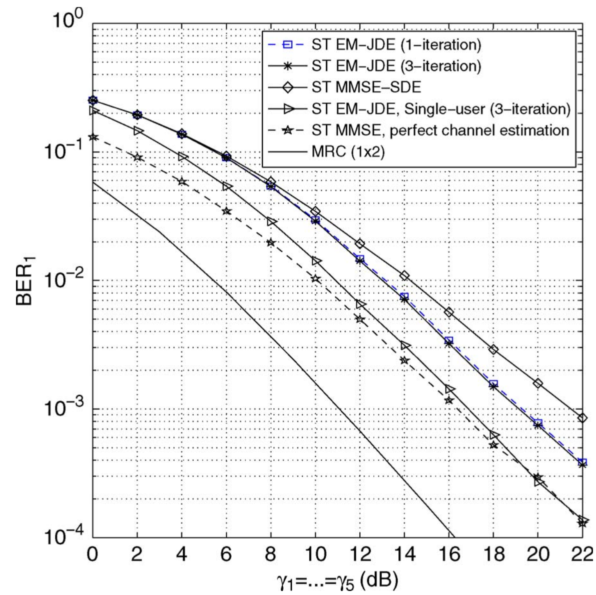

Fig. 2. BER performance of the first user considering the ST EM-JDE receiver with two transmit antennas and one receive antenna. The channel coefficients are assumed unknown at the receiver $(L = 20, p = 1, \text{and } \rho = 0.3)$.

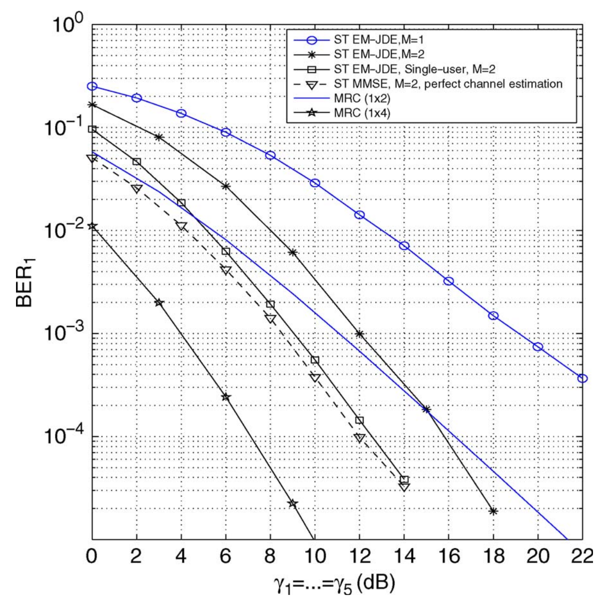

Fig. 3. BER performance of the first user considering the ST EM-JDE receiver with $M = 2$ receive antennas, $L = 20$, $p = 1$, $\rho = 0.3$, and three iterations.

$(\beta_k^m = 1/K)$. The results show that the EM receiver with optimum β_k^m is near–far resistant. Also, when the interference level is high, a reliable estimate of the MAI is obtained, and, consequently, the MAI removal is performed efficiently. As the level of MAI decreases, we note that the EM receiver reduces to the SU matched filter, where the performance converges to the SU bound. On the other hand, the performance of the MMSE-SDE receiver degrades due to noise enhancement. We can notice the effect of β_k^m on the performance of the ST EM-JDE receiver. That is, compared with the case of equal β_k^m, the optimum weights β_k^m achieve the best balance between the ST SU coherent detector and the ST parallel interference cancellation receiver (14). One should note that with the relatively high SNR (i.e., $\gamma_1 = 16$ dB) in Fig. 4, the limiting factor on the

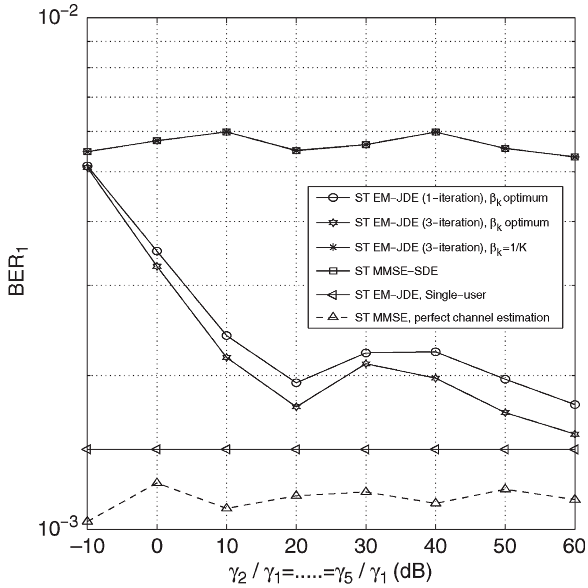

Fig. 4. BER behavior of the first user as a function of the MAI level with $M = 1$ receive antenna, $L = 20$, $p = 1$, $\rho = 0.3$, and $\gamma_1 = 16$ dB.

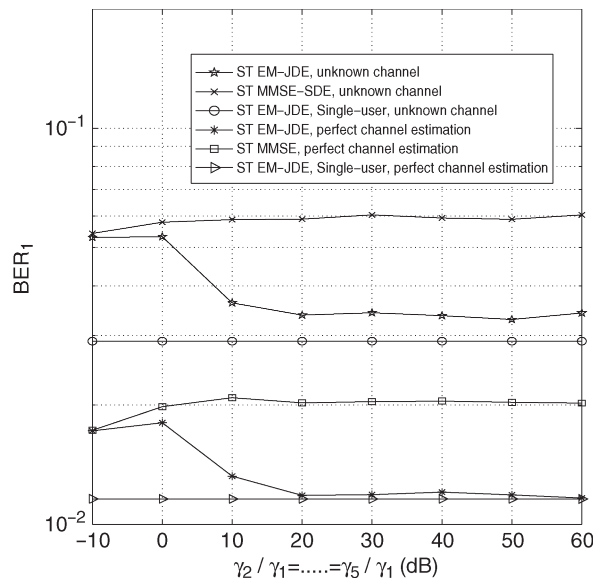

Fig. 6. Effect of the channel-estimation error on the BER performance of the first user considering the EM-JDE receiver as a function of MAI energy ($L = 20$, $p = 1$, $\rho = 0.3$, $\gamma_1 = 8$ dB, and three iterations).

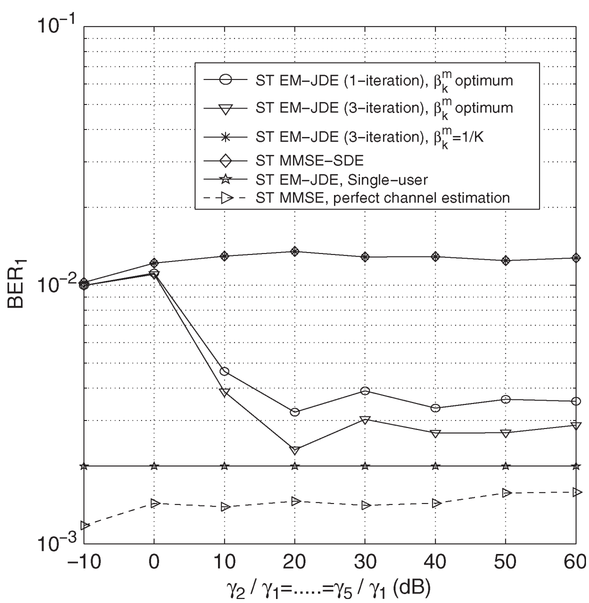

Fig. 5. BER behavior of the first user as a function of the MAI level for $M = 2$ receive antennas, $L = 20$, $p = 1$, $\rho = 0.3$, and $\gamma_1 = 8$ dB.

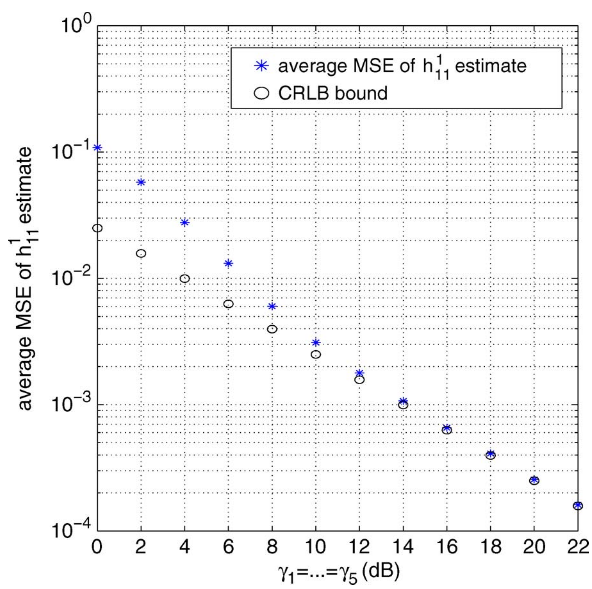

Fig. 7. MSE of channel estimates in the MIMO-CDMA system with $M = 2$ receive antennas, $L = 20$, $p = 1$, $\rho = 0.3$, and two iterations.

system performance is mostly due to the MAI contribution. At severe near–far scenarios (i.e., large MAI levels), the proposed receiver in this case is shown to achieve performance that is close to that of the ST MMSE with perfect channel estimation.

It is important to mention that the derived results can be well extended to the case of a frequency-selective fading channel. Considering an $N \times M$ MIMO system in a frequency-selective channel with W-resolvable paths, the expected full system diversity in this case is NMW. Furthermore, the analysis for a K-user asynchronous system with a transmitted frame length of L per user can be viewed as an equivalent synchronous system with LK users [32].

In Fig. 6, we compare the performance of the proposed receiver considering two cases at the receiver: perfect channel state information (CSI) and an unknown channel. Considering the perfect CSI case, the EM-based receiver achieves an excellent match with the corresponding SU bound at extremely high MAI levels. On the other hand, for the case of an unknown channel, there is some performance loss due to channel-estimation errors.

Finally, in Fig. 7, we assess the accuracy and the asymptotic performance of the channel estimates based on the EM algorithm for a system with $M = 2$ receive antennas. In this figure, we simulate the MSE of the channel estimate $\widehat{h}_{11}^1$, which is averaged over 10^5 channel realizations at different SNRs. The results show that the channel estimates are asymptotically efficient, where the average MSE of h_{11}^1 estimate converges to the CRLB at a high SNR, confirming our analytical results presented in Section V.

VII. Conclusion

We have developed an iterative JDE receiver based on the EM algorithm for STS systems. Using Monte Carlo simulations, we have examined the performance of our proposed receiver in flat-fading channels. It has been shown that with a few training bits, the receiver can achieve performance that is close to the SU bound in a few iterations. We have also shown that the proposed receiver achieves the full system diversity through accurate channel estimates.

Appendix A
Estimation of $\mathcal{Q}_k(\mathbf{b}_k|\mathbf{b}^i)$

Using some mathematical manipulations, the conditional expectation of the likelihood function in (11) can be expressed as

$$\mathcal{Q}_k(\mathbf{b}_k|\mathbf{b}^i) = \sum_{m=1}^{M}\sum_{l=1}^{L} Re\Big\{E\Big[A_{k1}^m(l)^i + \beta_k^m\left(A_{k2}^m(l)^i - A_{k3}^m(l)^i\right)\Big]\Big\} \quad \text{(A1)}$$

where

$$A_{k1}^m(l)^i = E\left[\mathbf{h}_k^{m^H}\mathbf{B}_k(l)\mathbf{F}_k^H\mathbf{F}_k\mathbf{B}_k(l)^i\mathbf{h}_k^m|\mathbf{y},\mathbf{b}^i\right] \quad \text{(A2)}$$

$$A_{k2}^m(l)^i = E\left[\mathbf{h}_k^{m^H}\mathbf{B}_k(l)\mathbf{F}_k^H\mathbf{y}^m(l)|\mathbf{y},\mathbf{b}^i\right] \quad \text{(A3)}$$

$$A_{k3}^m(l)^i = \sum_{j=1}^{K} E\left[\mathbf{h}_k^{m^H}\mathbf{B}_k(l)\mathbf{F}_k^H\mathbf{F}_j\mathbf{B}_j(l)^i\mathbf{h}_j^m|\mathbf{y},\mathbf{b}^i\right]. \quad \text{(A4)}$$

Starting with $A_{k1}^m(l)^i$ and after some mathematical manipulations, (A2) can be represented as

$$A_{k1}^m(l)^i = a_{11}(l)^i\left(\left|h_{1m}^k\right|^2\right)^i + a_{12}(l)^i\left(h_{1m}^{k*}h_{2m}^k\right)^i + a_{21}(l)^i\left(h_{2m}^{k*}h_{1m}^k\right)^i + a_{22}(l)^i\left(\left|h_{2m}^k\right|^2\right)^i \quad \text{(A5)}$$

where

$$\begin{aligned} a_{11}(l)^i &= b_{k1}(l)b_{k1}(l)^i + b_{k2}(l)b_{k2}(l)^i \\ a_{12}(l)^i &= b_{k1}(l)b_{k2}(l)^i - b_{k2}(l)b_{k1}(l)^i \\ a_{21}(l)^i &= b_{k2}(l)b_{k1}(l)^i - b_{k1}(l)b_{k2}(l)^i \\ a_{22}(l)^i &= b_{k1}(l)b_{k1}(l)^i + b_{k2}(l)b_{k2}(l)^i. \end{aligned}$$

$(h_{qm}^{k^*}h_{vm}^j)^i = E[h_{qm}^{k^*}h_{vm}^j|\mathbf{y},\mathbf{b}^i]$, $q, v \in \{1,2\}$, $j, k \in \{1,\ldots,K\}$, and * denotes the complex conjugate. From (A3), $A_{k2}^m(l)^i$ is given by

$$\begin{aligned} A_{k2}^m(l)^i &= E\left[\mathbf{h}_k^{m^H}\mathbf{B}_k(l)\mathbf{z}_k^m(l)|\mathbf{y},\mathbf{b}^i\right] \\ &= \left(h_{1m}^k\right)^{i^*}\left(b_{k1}(l)z_{k1}^m(l) + b_{k2}(l)z_{k2}^m(l)\right) \\ &\quad + \left(h_{2m}^k\right)^{i^*}\left(b_{k2}(l)z_{k1}^m(l) - b_{k1}(l)z_{k2}^m(l)\right) \end{aligned} \quad \text{(A6)}$$

where $(h_{qm}^k)^i = E[h_{qm}^k|\mathbf{y},\mathbf{b}^i]$, and $\mathbf{z}_k^m(l) = [z_{k1}^m(l), z_{k2}^m(l)]^T = \mathbf{F}_k^H\mathbf{y}^m(l)$ is a (2×1) vector consisting of the matched filter outputs corresponding to the kth user at the lth received period. In a similar manner, one can express $A_{k3}^m(l)^i$ in (A4) as

$$A_{k3}^m(l)^i = \sum_{j=1}^{K} d_{11}^{kj}(l)^i\left(h_{1m}^{k*}h_{1m}^j\right)^i + d_{12}^{kj}(l)^i\left(h_{1m}^{k*}h_{2m}^j\right)^i + d_{21}^{kj}(l)^i\left(h_{2m}^{k*}h_{1m}^j\right)^i + d_{22}^{kj}(l)^i\left(h_{2m}^{k*}h_{2m}^j\right)^i \quad \text{(A7)}$$

where

$$\begin{aligned} d_{11}^{kj}(l)^i &= \left(\rho_{11}^{kj}b_{k1}(l) + \rho_{21}^{kj}b_{k2}(l)\right)b_{j1}(l)^i + \left(\rho_{12}^{kj}b_{k1}(l) + \rho_{22}^{kj}b_{k2}(l)\right)b_{j2}(l)^i \\ d_{12}^{kj}(l)^i &= \left(\rho_{11}^{kj}b_{k1}(l) + \rho_{21}^{kj}b_{k2}(l)\right)b_{j2}(l)^i - \left(\rho_{12}^{kj}b_{k1}(l) + \rho_{22}^{kj}b_{k2}(l)\right)b_{j1}(l)^i \\ d_{21}^{kj}(l)^i &= \left(\rho_{11}^{kj}b_{k2}(l) - \rho_{21}^{kj}b_{k1}(l)\right)b_{j1}(l)^i + \left(\rho_{12}^{kj}b_{k2}(l) - \rho_{22}^{kj}b_{k1}(l)\right)b_{j2}(l)^i \\ d_{22}^{kj}(l)^i &= \left(\rho_{11}^{kj}b_{k2}(l) - \rho_{21}^{kj}b_{k1}(l)\right)b_{j2}(l)^i - \left(\rho_{12}^{kj}b_{k2}(l) - \rho_{22}^{kj}b_{k1}(l)\right)b_{j1}(l)^i \end{aligned}$$

and ρ_{qv}^{kj} is the cross-correlation value between c_{kq} and c_{jv}. According to [6], and based on the synchronous transmission model, $\rho_{11}^{kj} = \rho_{22}^{kj}$, and $\rho_{12}^{kj} = \rho_{21}^{kj} = 0$. Substituting (A5)–(A7) into (A1), the conditional likelihood function corresponding to user k is given by

$$\begin{aligned} \mathcal{Q}_k(\mathbf{b}_k|\mathbf{b}^i) &= \sum_{m=1}^{M}\sum_{l=1}^{L} Re \\ &\times\Big\{a_{11}(l)^i\left(\left|h_{1m}^k\right|^2\right)^i + a_{12}(l)^i\left(h_{1m}^{k*}h_{2m}^k\right)^i \\ &\quad + a_{21}(l)^i\left(h_{2m}^{k*}h_{1m}^k\right)^i + a_{22}(l)^i\left(\left|h_{2m}^k\right|^2\right)^i \\ &\quad + \beta_k^m\Big(\left(h_{1m}^k\right)^{i^*}\times\left(b_{k1}(l)z_{k1}^m(l) + b_{k2}(l)z_{k2}^m(l)\right) \\ &\quad + \left(h_{2m}^k\right)^{i^*}\left(b_{k2}(l)z_{k1}^m(l) - b_{k1}(l)z_{k2}^m(l)\right) \\ &\quad - \sum_{j=1}^{K} d_{11}^{kj}(l)^i\left(h_{1m}^{k*}h_{1m}^j\right)^i \\ &\quad + d_{12}^{kj}(l)^i\left(h_{1m}^{k*}h_{2m}^j\right)^i + d_{21}^{kj}(l)^i\left(h_{2m}^{k*}h_{1m}^j\right)^i \\ &\quad + d_{22}^{kj}(l)^i\left(h_{2m}^{k*}h_{2m}^j\right)^i\Big)\Big\}. \end{aligned} \quad \text{(A8)}$$

Considering a large frame L, the second summand in the right-hand side of (16) can be neglected, and (A8) is reduced to

$$\mathcal{Q}_k(\mathbf{b}_k|\mathbf{b}^i) = \sum_{l=1}^{L}\mathcal{Q}_k\left(\mathbf{b}_k(l)|\mathbf{b}^i\right) \quad \text{(A9)}$$

where

$$\begin{aligned}\mathcal{Q}_k\left(\mathbf{b}_k(l)|\mathbf{b}^i\right) &= \sum_{m=1}^{M} Re\Bigg\{ (1-\beta_k^m)\left(a_{11}(l)^i\left|\left(h_{1m}^k\right)^i\right|^2 + a_{12}(l)^i\left(h_{1m}^k\right)^{i^*}\left(h_{2m}^k\right)^i + a_{21}(l)^i\left(h_{2m}^k\right)^{i^*}\left(h_{1m}^k\right)^i + a_{22}(l)^i\left|\left(h_{2m}^k\right)^i\right|^2\right) \\ &\quad + \beta_k^m\left(\mathbf{h}_k^m\right)^{i^H}\mathbf{B}_k(l) \times \left(\mathbf{z}_k^m(l) - \sum_{j=1,j\neq k}^{K}\mathbf{R}_{kj}\mathbf{B}_j(l)^i\left(\mathbf{h}_j^m\right)^i\right)\Bigg\}. \end{aligned} \quad \text{(A10)}$$

Appendix B
Distribution of $\mathbf{h}^m|\mathbf{y}, \mathbf{b}^i$

Let

$$\mathbf{n}_w = \left[\mathbf{n}_w^1, \mathbf{n}_w^2, \ldots, \mathbf{n}_w^M\right]$$

where

$$\mathbf{n}_w^m = \left[\mathbf{n}_w^m(1), \mathbf{n}_w^m(2), \ldots, \mathbf{n}_w^m(L)\right].$$

Then, the distribution of $\mathbf{y}|\mathbf{h}, \mathbf{b}^i$ is given by

$$P(\mathbf{y}|\mathbf{h}, \mathbf{b}^i) = \prod_{m=1}^{M} P(\mathbf{y}^m|\mathbf{h}^m, \mathbf{b}^i) \quad \text{(B1)}$$

where

$$\begin{aligned}P(\mathbf{y}^m|\mathbf{h}^m, \mathbf{b}^i) &= \frac{1}{\pi^{2LK}(\det \mathbf{\Sigma}_{wl})^L} \times e^{\frac{-1}{N_o}\left(\sum_{l=1}^{L}\left(\mathbf{y}^m(l)-\mathbf{FB}(l)^i\mathbf{h}^m\right)^H\left(\mathbf{y}^m(l)-\mathbf{FB}(l)^i\mathbf{h}^m\right)\right)} \\ &= \frac{1}{\pi^{2LK}(\det \mathbf{\Sigma}_{wl})^L} e^{\frac{-1}{N_o}\left\{\sum_{l=1}^{L}\mathbf{y}^m(l)^H\mathbf{y}^m(l)\right\}} \\ &\quad \times e^{\frac{-1}{N_o}\left(-\left\{\sum_{l=1}^{L}\mathbf{y}^m(l)^H\mathbf{FB}(l)^i\right\}\mathbf{h}^m - \mathbf{h}^{m^H}\left\{\sum_{l=1}^{L}\mathbf{B}(l)^i\mathbf{F}^H\mathbf{y}^m(l)\right\}\right)} \\ &\quad \times e^{\frac{-1}{N_o}\left(\mathbf{h}^{m^H}\left\{\sum_{l=1}^{L}\mathbf{B}(l)^i\mathbf{RB}(l)^i\right\}\mathbf{h}^m\right)}. \end{aligned} \quad \text{(B2)}$$

det denotes the matrix determinant, and $\mathbf{\Sigma}_{wl} = N_o\mathbf{I}_{2K}$. The channel vector $\mathbf{h}$ has a multivariate Gaussian distribution, which is defined as

$$P(\mathbf{h}) = \prod_{m=1}^{M} P(\mathbf{h}^m) \quad \text{(B3)}$$

where

$$P(\mathbf{h}^m) = \frac{1}{\pi^{2K}\det \mathbf{\Sigma}_{hh}} e^{-\mathbf{h}^{m^H}\mathbf{\Sigma}_{hh}^{-1}\mathbf{h}^m}. \quad \text{(B4)}$$

Using (B1) and (B3), the joint distribution

$$P(\mathbf{y}, \mathbf{h}|\mathbf{b}^i) = \prod_{m=1}^{M} P(\mathbf{y}^m, \mathbf{h}^m|\mathbf{b}^i) \quad \text{(B5)}$$

where we have (B6), shown at the bottom of the page. In (B5), we can notice that the pairs $(\mathbf{y}^m, \mathbf{h}^m|\mathbf{b}^i)$, $m = 1, \ldots, M$, are mutually independent. Therefore, we will subsequently focus on the individual joint distribution of each pair $P(\mathbf{y}^m, \mathbf{h}^m|\mathbf{b}^i)$. The exponential term, including the second and third terms in the right-hand side of (B6), can be reformed to take the form of the exponential term of the standard multivariate Gaussian distribution by multiplying (B6) with $e^{-\mathbf{m}_h^{m^H}\mathbf{\Sigma}_h^{-1}\mathbf{m}_h^m}/e^{-\mathbf{m}_h^{m^H}\mathbf{\Sigma}_h^{-1}\mathbf{m}_h^m}$, where

$$\mathbf{\Sigma}_h = \left(N_o^{-1}\sum_{l=1}^{L}\mathbf{B}(l)^i\mathbf{RB}(l)^i + \mathbf{\Sigma}_{hh}^{-1}\right)^{-1} \quad \text{(B7)}$$

$$\mathbf{m}_h^m = \left(\sum_{l=1}^{L}\mathbf{B}(l)^i\mathbf{RB}(l)^i + N_o\mathbf{\Sigma}_{hh}^{-1}\right)^{-1}\sum_{l=1}^{L}\mathbf{B}(l)^i\mathbf{z}^m(l) \quad \text{(B8)}$$

and $\mathbf{z}^m(l) = F^H\mathbf{y}^m(l)$. By integrating (B6) with respect to $\mathbf{h}^m$, we obtain

$$P(\mathbf{y}^m|\mathbf{b}^i) = \frac{\det \mathbf{\Sigma}_h}{\pi^{2K(L+1)}(\det \mathbf{\Sigma}_{wl})^L\det \mathbf{\Sigma}_{hh}} \times e^{\frac{-1}{N_o}\left\{\sum_{l=1}^{L}\mathbf{y}^m(l)^H\mathbf{y}(l)\right\} + \mathbf{m}_h^{m^H}\mathbf{\Sigma}_h^{-1}\mathbf{m}_h^m}. \quad \text{(B9)}$$

Finally, using (B6) and (B9), we can estimate the conditional distribution of $(\mathbf{h}^m|\mathbf{y}^m, \mathbf{b}^i)$ as

$$P(\mathbf{h}^m|\mathbf{y}^m, \mathbf{b}^i) = \frac{P(\mathbf{h}^m, \mathbf{y}^m|\mathbf{b}^i)}{P(\mathbf{y}^m|\mathbf{b}^i)} = \frac{1}{\pi^{2K}\det \mathbf{\Sigma}_h} e^{-\left(\mathbf{h}^m-\mathbf{m}_h^m\right)^H\mathbf{\Sigma}_h^{-1}\left(\left(\mathbf{h}^m-\mathbf{m}_h^m\right)\right.} \quad \text{(B10)}$$

$$\begin{aligned}P(\mathbf{y}^m, \mathbf{h}^m|\mathbf{b}^i) &= \frac{e^{\frac{-1}{N_o}\left\{\sum_{l=1}^{L}\mathbf{y}^m(l)^H\mathbf{y}^m(l)\right\}}}{\pi^{2K(L+1)}(\det \mathbf{\Sigma}_{wl})^L\det \mathbf{\Sigma}_{hh}} \\ &\quad \times e^{\left\{N_o^{-1}\left\{\sum_{l=1}^{L}\mathbf{y}^m(l)^H\mathbf{FB}(l)^i\right\}\mathbf{h}^m + N_o^{-1}\mathbf{h}^{m^H}\left\{\sum_{l=1}^{L}\mathbf{B}(l)^i\mathbf{F}^H\mathbf{y}^m(l)\right\}\right\}} \\ &\quad \times e^{-\mathbf{h}^{m^H}\left\{No^{-1}\sum_{l=1}^{L}\mathbf{B}(l)^i\mathbf{RB}(l)^i + \mathbf{\Sigma}_{hh}^{-1}\right\}\mathbf{h}^m} \end{aligned} \quad \text{(B6)}$$

which represents a Gaussian distribution with mean

$$E[\mathbf{h}^m|\mathbf{y}^m,\mathbf{b}^i] = \mathbf{m}_h^m = \left(\sum_{l=1}^{L}\mathbf{B}(l)^i\mathbf{R}\mathbf{B}(l)^i + N_o\mathbf{\Sigma}_{hh}^{-1}\right)^{-1} \times \sum_{l=1}^{L}\mathbf{B}(l)^i\mathbf{z}^m(l) \quad \text{(B11)}$$

and covariance

$$E\left[\left(\mathbf{h}^m - (\mathbf{h}^m)^i\right)\left(\mathbf{h}^m - (\mathbf{h}^m)^i\right)^H|\mathbf{y}^m,\mathbf{b}^i\right] = \left(N_o^{-1}\sum_{l=1}^{L}\mathbf{B}(l)^i\mathbf{R}\mathbf{B}(l)^i + \mathbf{\Sigma}_{hh}^{-1}\right)^{-1} = \mathbf{\Sigma}_h. \quad \text{(B12)}$$

APPENDIX C
CONSISTENCY OF CHANNEL ESTIMATES

In this part, we show that the channel estimate $(\mathbf{h}^m)^i$ is consistent. First, we prove that the estimate $(\mathbf{h}^m)^i$ is asymptotically unbiased. In other words

$$\lim_{L\to\infty} E\left[(\mathbf{h}^m)^i|\mathbf{h},\mathbf{b}^i\right] = \mathbf{h}^m. \quad \text{(C1)}$$

We start by the expectation of (18) given $\mathbf{h}^m$ and $\mathbf{b}^i$. We have

$$E\left[(\mathbf{h}^m)^i|\mathbf{h}^m,\mathbf{b}^i\right] = \left(\sum_{l=1}^{L}\mathbf{B}(l)^i\mathbf{R}\mathbf{B}(l)^i + N_o\mathbf{\Sigma}_{hh}^{-1}\right)^{-1} \times \sum_{l=1}^{L}\mathbf{B}(l)^i E\left[\mathbf{z}^m(l)|\mathbf{h},\mathbf{b}^i\right]. \quad \text{(C2)}$$

Substituting (2) in (C2) and inserting the factor L^{-1}, we obtain

$$E\left[(\mathbf{h}^m)^i|\mathbf{h}^m,\mathbf{b}^i\right] = \left(L^{-1}\sum_{l=1}^{L}\mathbf{B}(l)^i\mathbf{R}\mathbf{B}(l)^i + L^{-1}N_o\mathbf{\Sigma}_{hh}^{-1}\right)^{-1} \times L^{-1}\sum_{l=1}^{L}\mathbf{B}(l)^i\mathbf{R}\mathbf{B}(l)^i\mathbf{h}^m. \quad \text{(C3)}$$

Invoking the strong law of large numbers, $L^{-1}\sum_{l=1}^{L}\mathbf{B}(l)^i\mathbf{R}\mathbf{B}(l)^i$ converges to $E[\mathbf{B}(l)^i\mathbf{R}\mathbf{B}(l)^i] = 2\mathbf{I}_{2K}$ as $L\to\infty$. Thus, (C1) follows. Finally, we show that the error covariance matrix of the channel estimates Ω_{hh}^i in (19) converges to $\mathbf{0}$ as $L\to\infty$, i.e.,

$$\lim_{L\to\infty}\Omega_{hh}^i = \mathbf{0}. \quad \text{(C4)}$$

Starting from (19), we have

$$L\mathbf{\Omega}_{hh}^i = \left(L^{-1}N_o^{-1}\sum_{l=1}^{L}\mathbf{B}(l)^i\mathbf{R}\mathbf{B}(l)^i + L^{-1}\mathbf{\Sigma}_{hh}^{-1}\right)^{-1}. \quad \text{(C5)}$$

Employing the same argument as above, it is easy to show that $L\mathbf{\Omega}_{hh}^i$ converges to $(N_o/2)\mathbf{I}_{2K}$ as $L\to\infty$. Consequently, $\mathbf{\Omega}_{hh}^i$ converges to $\mathbf{0}$ with rate $1/L$, hence, proving (C4).

REFERENCES

[1] H. Bolckei, D. Gesbert, and A. Paulraj, "On the capacity of OFDM-based spatial multiplexing systems," *IEEE Trans. Commun.*, vol. 50, no. 2, pp. 225–234, Feb. 2002.

[2] V. Tarokh, H. Jafarkhani, and A. R. Calderbank, "Space–time block codes from orthogonal designs," *IEEE Trans. Inf. Theory*, vol. 45, no. 5, pp. 1456–1467, Jul. 1999.

[3] V. Tarokh, N. Seshadri, and A. R. Calderbank, "Space–time codes for high data rate wireless communication: Performance criterion and code construction," *IEEE Trans. Inf. Theory*, vol. 44, no. 2, pp. 744–765, Mar. 1998.

[4] S. M. Alamouti, "A simple transmit diversity technique for wireless communications," *IEEE J. Sel. Areas Commun.*, vol. 16, no. 8, pp. 1451–1458, Oct. 1998.

[5] M. J. Juntti and M. Latva-aho, "Multiuser receivers for CDMA systems in Rayleigh fading channels," *IEEE Trans. Veh. Technol.*, vol. 49, no. 3, pp. 885–899, May 2000.

[6] B. Hochwald, T. Marzetta, and C. Papadias, "A transmitter diversity scheme for wideband CDMA systems based on space–time spreading," *IEEE J. Sel. Areas Commun.*, vol. 19, no. 1, pp. 1451–1458, Jan. 2001.

[7] R. A. Soni and R. M. Buehre, "On the performance of open-loop transmit diversity techniques for IS-2000 systems: Comparative study," *IEEE Trans. Wireless Commun.*, vol. 3, no. 5, pp. 1602–1615, Sep. 2004.

[8] 3rd Generation Partnership Project 2, "Physical Layer Standard for cdma2000 Spread Spectrum Systems Release D," Feb. 2004.

[9] X. Gao, B. Jiang, X. You, Z. Pan, and Y. Xue, "Efficient channel estimation for MIMO single-carrier block transmission with dual cyclic time slot structure," *IEEE Trans. Commun.*, vol. 55, no. 11, pp. 2210–2223, Nov. 2007.

[10] G. V. V. Sharma and A. Chockalingam, "Performance analysis of maximum-likelihood multiuser detection in space–time-coded CDMA with imperfect channel estimation," *IEEE Trans. Veh. Technol.*, vol. 55, no. 6, pp. 1824–1837, Nov. 2006.

[11] L. Chong and L. Milstein, "The effects of channel-estimation errors on a space–time spreading CDMA system with dual transmit and dual receive diversity," *IEEE Trans. Commun.*, vol. 52, no. 7, pp. 1145–1151, Jul. 2004.

[12] E. de Carvalho and D. T. M. Slock, "Maximum-likelihood blind FIR multi-channel estimation with Gaussian prior for the symbols," in *Proc. IEEE Int. Conf. Acoust. Speech, Signal Process.*, Apr. 1997, vol. 5, pp. 3593–3596.

[13] E. de Carvalho and D. T. M. Slock, "Cramér–Rao bounds for semi blind, blind and training sequence based channel estimation," in *Proc. 1st IEEE Signal Process. Workshop Signal Process. Adv. Wireless Commun.*, Apr. 1997, vol. 1, pp. 129–132.

[14] D. Gore, S. Sandhu, and J. Paulraj, "Blind channel identification and projection receiver determination for multicode and multirate situations in DS-CDMA systems," in *Proc. IEEE ICC*, Apr./May 2002, vol. 3, pp. 1949–1953.

[15] Y. Huang and J. A. Ritcey, "Joint iterative channel estimation and decoding for bit-interleaved coded modulation over correlated fading channels," *IEEE Trans. Wireless Commun.*, vol. 4, no. 5, pp. 2549–2558, Sep. 2005.

[16] D. K. C. So and R. S. Cheng, "Iterative EM receiver for space–time coded systems in MIMO frequency-selective fading channels with channel gain and order estimation," *IEEE Trans. Wireless Commun.*, vol. 3, no. 6, pp. 1928–1935, Nov. 2004.

[17] T. K. Moon, "The expectation–maximization algorithm," *IEEE Signal Process. Mag.*, vol. 13, no. 6, pp. 45–59, Nov. 1996.

[18] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via EM algorithm," *J. R. Stat. Soc.*, vol. 39, no. 1, pp. 1–38, Jan. 1977.

[19] C. N. Georghiades and J. C. Han, "Sequence estimation in the presence of random parameters via the EM algorithm," *IEEE Trans. Commun.*, vol. 45, no. 3, pp. 300–308, Mar. 1997.

[20] M. J. Borran and M. Nasiri-Kenari, "An efficient detection technique for synchronous CDMA communication systems based on the expectation maximization algorithm," *IEEE Trans. Veh. Technol.*, vol. 49, no. 5, pp. 1663–1668, Sep. 2000.

[21] M. Feder and E. Weinstein, "Parameter estimation of superimposed signals using the EM algorithm," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 36, no. 4, pp. 477–489, Apr. 1988.

[22] A. Kocian and B. H. Fleury, "EM-based joint data detection and channel estimation of DS-CDMA signals," *IEEE Trans. Commun.*, vol. 51, no. 10, pp. 1709–1720, Oct. 2003.

[23] S.-H. Wu, U. Mitra, and C.-C. J. Kuo, “Iterative joint channel estimation and multiuser detection for DS-CDMA in frequency-selective fading channels,” *IEEE Trans. Signal Process.*, vol. 56, no. 7, pp. 3261–3277, Jul. 2008.
[24] C. Cozzo and B. Hughes, “Joint channel estimation and data detection in space–time communications,” *IEEE Trans. Commun.*, vol. 51, no. 8, pp. 1266–1270, Aug. 2003.
[25] Y.-P. Cheng, K.-Y. Zhang, and Z. Xu, *Matrix Theory*. Xi’an, China: Northwestern Polytechnical Univ. Press, 2002.
[26] D. Divsalar, M. K. Simon, and D. Raphaeli, “Improved parallel interference cancellation for CDMA,” *IEEE Trans. Commun.*, vol. 46, no. 2, pp. 258–268, Feb. 1998.
[27] J. Ylioinas and M. Juntti, “Iterative joint detection, decoding, and channel estimation in turbo-coded MIMO-OFDM,” *IEEE Trans. Veh. Technol.*, vol. 58, no. 4, pp. 1784–1796, May 2009.
[28] A. Klein, G. K. Kaleh, and P. W. Baier, “Zero forcing and minimum mean-square error equalization of multiuser detection in code-division multiple-access channels,” *IEEE Trans. Veh. Technol.*, vol. 45, no. 2, pp. 276–287, May 1996.
[29] H. V. Poor, *An Introduction to Signal Detection and Estimation*. New York: Springer-Verlag, 1994.
[30] L.-L. Yang, “MIMO-assisted space-code-division multiple-access: Linear detectors and performance over multipath fading channels,” *IEEE J. Sel. Areas Commun.*, vol. 24, no. 1, pp. 121–131, Jan. 2006.
[31] J. G. Proakis, *Digital Communications*. New York: McGraw-Hill, 2000.
[32] S. Verdu, *Multiuser Detection*. Cambridge, U.K.: Cambridge Univ. Press, 1998.

Walaa Hamouda (S’96–M’02–SM’06) received the M.A.Sc. and Ph.D. degrees in electrical and computer engineering from Queen’s University, Kingston, ON, Canada, in 1998 and 2002, respectively.

Since July 2002, he has been with the Department of Electrical and Computer Engineering, Concordia University, Montreal, QC, Canada, where he is currently an Associate Professor. In June 2006, he was appointed as a Concordia University Research Chair in Communications and Networking. His current research interests include wireless networks, multiple-input–multiple-output space–time processing, multiuser communications, cross-layer design, and source and channel coding.

Dr. Hamouda has served or is currently serving on the technical program committee of several IEEE conferences, including the 2006–2009 IEEE International Conference on Communications (ICC), the 2007 IEEE Wireless Communications and Networking Conference, the 2006–2008 IEEE Global Telecommunications Conference, and the 2001–2003, Fall 2005, and Spring 2008 IEEE Vehicular Technology Conference (VTC). He served as the Technical Cochair of many conferences and symposiums, including the Ad-hoc, Sensor, and Mesh Networking Symposium of the 2010 ICC, the 25th Queen’s Biennial Symposium on Communications, the 2006 4th Arab Computer Society/IEEE International Conference on Computer Systems and Applications: Ad Hoc and Sensor Networks Workshop, and the 2005 Signal Processing Symposium of the IEEE International Conference of Wireless Networks, Communications, and Mobile Computing. He served as the Track Chair for the Radio Access Techniques of the Fall 2006 IEEE VTC. From September 2005 to November 2008, he was the Chair of the IEEE Montreal Chapter in Communications and Information Theory. He has received of many awards, including the Best Paper Award [WNS] at the 2009 ICC, Dresden, Germany, and the IEEE Canada Certificate of Appreciation in 2007 and 2008. He serves as an Associate Editor for the IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY, IEEE COMMUNICATIONS LETTERS, *Wiley International Journal of Communication Systems*, and the *Journal of Computer Systems, Networks, and Communications*.

Ayman Assra received the B.Sc. degree in electrical engineering from Alexandria University, Alexandria, Egypt, in 2000 and the M.Sc. degree from the Arab Academy for Science and Technology and Maritime Transport, Alexandria, in 2004. He is currently working toward the Ph.D. degree in electrical and computer engineering at Concordia University, Montreal, QC, Canada.

Since May 2005, he has been a Research Assistant with Concordia University. His current research interests include space–time processing and multiple-input–multiple-output communications.

Amr Youssef (SM’06) received the B.Sc. and M.Sc. degrees from Cairo University, Cairo, Egypt, in 1990 and 1993, respectively, and the Ph.D. degree from Queens University, Kingston, ON, Canada, in 1997.

He worked for Nortel Networks; the Center for Applied Cryptographic Research, University of Waterloo, Waterloo, ON; IBM; and Cairo University. In 2004, he joined Concordia Institute for Information Systems Engineering, Concordia University, Montreal, QC, Canada, as an Associate Professor. His main research interests include the areas of sequence design, cryptology, and network security.