Physical Communication 3 (2010) 87–96

Contents lists available at ScienceDirect

Physical Communication

journal homepage: www.elsevier.com/locate/phycom

Full length article

Iterative joint data detection and channel estimation for uplink MC-CDMA systems in the presence of frequency selective channels

Erdal Panayırcı [a], Hakan Doğan [b,*], Hakan A. Çırpan [b], Alexander Kocian [c], Bernard H. Fleury [d]

[a] Department of Electronics Engineering, Kadir Has University, Cibali 34083, Istanbul, Turkey

[b] Department of Electrical and Electronics Engineering, Istanbul University, Avcilar 34320, Istanbul, Turkey

[c] University of Rome Tor Vergada, Center for TeleInfrastructure(CTiF), Rome, Italy

[d] Aalborg Univeristy, Fredrik Bajers Vej 7A, A3-202, Aalborg DK-9220, Denmark

ARTICLE INFO

Keywords:
MC-CDMA Systems
EM, SAGE algorithm
Joint multiuser detection
Multichannel estimation

ABSTRACT

This paper is concerned with joint multiuser detection and multichannel estimation (JDE) for uplink multicarrier code-division multiple-access (MC-CDMA) systems in the presence of frequency selective channels. The detection and estimation, implemented at the receiver, are based on a version of the expectation maximization (EM) algorithm and the space-alternating generalized expectation–maximization (SAGE) which are very suitable for multicarrier signal formats. The EM-JDE receiver updates the data bit sequences in parallel, while the SAGE-JDE receiver reestimates them successively. The channel parameters are updated in parallel in both schemes. Application of the EM-based algorithm to the problem of iterative data detection and channel estimation leads to a receiver structure that also incorporates a partial interference cancelation. Computer simulations show that the proposed algorithms have excellent BER end estimation performance.

1. Introduction

Future communications will be driven by the need to provide more integrated high-capacity, wide-coverage services to face new challenges in meeting the ubiquity and mobility requirements of cellular systems. For the 21st century user, extensive attempts have therefore been made, and further spectacular enabling technology advances are expected, in an effort to render ubiquitous wireless connectivity a reality. One promising approach is the integration of multiple access and modulation technologies. In particular, the combination of multicarrier and code division multiple access (MC-CDMA) has been preliminarily successful because it incorporates the benefits of orthogonal frequency-division multiplexing (OFDM) and spread-spectrum presents good potentialities to make it the best technology to support broadband applications [1,2]. Moreover, MC-CDMA systems relieve the limitations on system capacity that occur in direct-sequence code division multiple access (DS-CDMA) systems.

The major advantages of MC-CDMA which lie behind its success are robustness in the case of multipath fading, a very reduced system complexity due to equalization in the frequency domain, and the capability of narrow-band interference rejection. It also has an ability to reduce users signal power during transmission using a spreading so that the user can communicate using a low-level transmitted signal, which is closer to the noise power level.

In conventional MC-CDMA systems, multiple access interference (MAI) mitigation is accomplished at the receiver using single-user or multi-user detection schemes [3]. However, even though a multiuser detection scheme is known to increase the bandwidth efficiency of the system drastically, its detection complexity grows exponentially with the number of users and the number of multipaths,

* Corresponding author.
E-mail addresses: eepanay@khas.edu.tr (E. Panayırcı), hdogan@istanbul.edu.tr (H. Doğan), hcirpan@istanbul.edu.tr (H.A. Çırpan), Alexander.Kocian@uniroma2.it (A. Kocian), bfl@kom.aau.dk (B.H. Fleury).

1874-4907/$ – see front matter Crown Copyright © 2009 Published by Elsevier B.V. All rights reserved.
doi:10.1016/j.phycom.2009.06.001

which makes its implementation unfeasible. Several sub-optimal detection techniques have been proposed in the literature such as linear multiuser detection [4] and iterative cancelation of MAI, either in a successive or parallel way in the received signal prior to data detection [5]. However, all these detection schemes require explicit knowledge of the channel parameters of the active users. A considerable amount of research has therefore been devoted to the problem of channel estimation in MC-CDMA systems. In [6], a subspace-based blind channel identification algorithm is proposed. Although blind solutions are attractive from the point of view of bandwidth efficiency, they are often computationally complex and require long data records to achieve good performance. As an alternative, data-aided channel estimation technique has been discussed in [7]. Moreover, channel acquisition and tracking in the uplink of MC-CDMA systems have also been studied in [8]. In this approach, channel tracking is pursued by means of a least mean square (LMS) algorithm while channel estimation is performed using different schemes based on a maximum likelihood (ML) criterion.

In contrast to previous approaches, several alternative algorithms can be considered to refine channel estimates through iterations. An attractive iterative technique is the expectation–maximization (EM) algorithm which has already been considered in several channel estimation scenarios [9–12]. In [11], a two-step detection, channel estimation procedure is adopted which uses the EM algorithm to estimate the channel in the first step and then uses the estimated channel to perform coherent detection in the second step. Moreover, an EM approach has been proposed for the general estimation of superimposed signals applied to the channel estimation for transmit diversity OFDM systems and was then compared with the space-alternating generalized expectation–maximization (SAGE) algorithm [13].

When dealing with the multiuser scenario, it is necessary to make excellent joint data and channel estimations for the initialization of the interference cancelation detector. The work is an extension of [14], in which joint data detection and channel estimation of uplink DS-CDMA systems were considered based on an EM algorithm in the presence of flat Rayleigh channels. We have extended their results for the uplink MC-CDMA systems with frequency selective channels. The channel estimation becomes more challenging for uplink systems since each channel between every user and the base station must be estimated rather than estimating a single channel as is the case in a downlink transmission. In this paper, we apply the EM and SAGE algorithms to the problem of joint multiuser data detection and the channel estimation (JDE) of MC-CDMA signals transmitted through frequency-selective channels. In this way, we obtain iterative methods of tractable complexity which intelligently combine the two processes of data detection and channel estimation.

Note that perfect timing synchronization between the users and the base station in MC-CDMA systems is an important issue and has to be solved before the channel estimation and the data detection processes. This is the case for all synchronous multiuser uplink systems. Timing offsets in the uplink are mainly due to the propagation delay incurred by users' signals. The timing error of each user with respect to the BS time reference can be decomposed into an integer part plus a fractional part with respect to the sampling period. As explained in [15], the fractional part can be incorporated into the channel impulse response and so is not considered in the analysis. Thus, under such circumstances, the use of a sufficiently long guard interval (in the form of a cyclic prefix) provides intrinsic protection against inter frame interference at the expense of extra overhead. Therefore, even in an imperfect timing synchronization scenario, the proposed technique in this paper for joint channel estimation and data detection will still work if timing offsets are incorporated as part of the channel impulse response. Note that quite recently, Kocian and Fleury [16] extended their earlier work to an asynchronous case which dealt with EM-based joint data detection and the channel estimation of DS-CDMA signals in the presence of quasi-static flat Rayleigh fading channels. A similar extension can be made for MC-CDMA systems with frequency selective channels if timing offsets are not incorporated as part of the channel impulse response.

The organization of this paper is as follows. The signal model of MC-CDMA systems and the channel model considered in this work are given in Section 2. The joint schemes for data detection and channel estimation based on EM and SAGE algorithms are presented in Sections 3 and 4, respectively. The performance of the algorithms proposed in the paper are assessed in Section 5 by computer simulations. Finally, the main conclusions of the paper are presented in Section 6.

Notation: Vectors (matrices) are denoted by boldface lower (upper) case letters; all vectors are column vectors; $(.)^*$, $(.)^{\mathrm{T}}$ and $(.)^H$ denote the conjugate, transpose and conjugate transpose, respectively; $\|.\|$ denotes the Frobenius norm; $\mathbf{I}_L$ denotes the $L \times L$ identity matrix; diag$\{.\}$ denotes a diagonal matrix; and finally, tr$\{.\}$ denotes the trace of a matrix.

2. Signal model

We considered a baseband MC-CDMA uplink system with P sub-carriers and K mobile users which are simultaneously active. For the kth user, each transmit symbol is modulated in the frequency domain by means of a $P \times 1$ specific spreading sequence $\mathbf{c}_k$. After transformation by a P-point IDFT and parallel-to-serial (P/S) conversion, a cyclic prefix (CP) is inserted of a length equal to at least the channel memory (L). In this work, to simplify the notation, it is assumed that the spreading factor is equivalent to the number of sub-carriers and all users have the same spreading factor. Finally, the signal is transmitted through a multipath channel with impulse response

$$g_k(t) = \sum_{l=1}^{L} g_{k,l}\delta(t - \tau_{k,l}) \tag{1}$$

where L is the number of paths in the kth user's channel; $g_{k,l}$ and $\tau_{k,l}$ are, respectively, the complex fading coefficient and the delay of lth path and P_k is the transmit power of the kth user. The fading process is assumed to be white. Its second-order statistics are known to the receiver. Note that the L-dimensional discrete channel impulse response vector $\mathbf{g}_k = [g_{k,1}, g_{k,2}, \ldots, g_{k,L}]^{\mathrm{T}}$ and the transmission power P_k can be combined as $\mathbf{h}_k = \sqrt{P_k}\mathbf{g}_k$, since they cannot be separated from each other.

In the receiver, the received signal is sampled at chip-rate and serial-to-parallel (S/P) converted. The CP is removed, and DFT is then applied to the discrete time signal to obtain the received vector expressed as

$$\mathbf{y}(m) = \sum_{k=1}^{K} b_k(m)\mathbf{C}_k\mathbf{F}\mathbf{h}_k + \mathbf{w}(m), \quad m = 1, 2, \ldots, M \quad (2)$$

where $b_k(m) \in \{+1, -1\}$ denotes binary data sent by the user k within the mth symbol time; $\mathbf{C}_k = \text{diag}(\mathbf{c}_k)$ with $\mathbf{c}_k = [c_{1k}, c_{2k}, \ldots, c_{Pk}]^{\text{T}}$ where each chip, c_{ik}, takes values in the set $\{-\frac{1}{\sqrt{P}}, \frac{1}{\sqrt{P}}\}$ denoting the kth user's spreading code ; $\mathbf{F} \in \mathbb{C}^{P\times L}$ denotes the DFT matrix with the (k, l)th element given by $\mathrm{e}^{-j2\pi kl/P}$; and $\mathbf{w}(m)$ is the $P \times 1$ zero-mean, i.i.d. Gaussian vector that models the additive noise in the P tones, with variance $\sigma^2/2$ per dimension.

Suppose M symbols are transmitted. We stack $\mathbf{y}(m)$ as $\mathbf{y} = [\mathbf{y}^{\text{T}}(1), \ldots, \mathbf{y}^{\text{T}}(M)]^{\text{T}}$. Then the received signal model can be written as

$$\mathbf{y} = \begin{bmatrix} b_1(1)\mathbf{C}_1\mathbf{F} & \cdots & b_K(1)\mathbf{C}_K\mathbf{F} \\ \vdots & \ddots & \vdots \\ b_1(M)\mathbf{C}_1\mathbf{F} & \cdots & b_K(M)\mathbf{C}_K\mathbf{F} \end{bmatrix} \begin{bmatrix} \mathbf{h}_1 \\ \vdots \\ \mathbf{h}_K \end{bmatrix} + \begin{bmatrix} \mathbf{w}(1) \\ \vdots \\ \mathbf{w}(M) \end{bmatrix} \quad (3)$$

and can be rewritten in the more succinct form

$$\mathbf{y} = \mathbf{A}\mathbf{h} + \mathbf{w} \quad (4)$$

where $\mathbf{h}_k$'s are modeled as complex Gaussian random variables with $\mathbf{h}_k \sim N(0, \boldsymbol{\Sigma}_{\mathbf{h}_k})$ and $\boldsymbol{\Sigma}_{\mathbf{h}_k} = E[\mathbf{h}_k\mathbf{h}_k^{\dagger}]$. It is then clear that $\mathbf{h} \sim N(0, \boldsymbol{\Sigma}_{\mathbf{h}})$ with $\boldsymbol{\Sigma}_{\mathbf{h}} = \text{diag}[\boldsymbol{\Sigma}_{\mathbf{h}_1}, \ldots, \boldsymbol{\Sigma}_{\mathbf{h}_K}]$. We assume that the covariance matrix $\boldsymbol{\Sigma}_{\mathbf{h}_k}$ of each user k is known, or measured by means of pilot symbols. Otherwise, a least-square estimator can be applied to estimate the channel and to measure $\boldsymbol{\Sigma}_{\mathbf{h}_k}$ as well [17]. Note that due to the orthogonality property of spreading sequences, $\mathbf{C}_k^{\text{T}}\mathbf{C}_k = \frac{1}{P}\mathbf{I}_P$.

3. Joint data detection and channel estimation with EM algorithms (EM-JDE)

Let $\mathbf{b}$ denote possibly vector-valued parameter to be estimated from some possibly vector-valued observation $\mathbf{y}$ with probability density $p(\mathbf{y}|\mathbf{b})$. The EM algorithm provides an iterative scheme to approach the ML estimate $\hat{\mathbf{b}} = \arg\max_b p(\mathbf{y}|\mathbf{b})$ in cases where a direct computation of $\hat{\mathbf{b}}$ is prohibitive. The derivation of the EM algorithm relies on the concept of a hypothetical, so-called complete unobservable data $\boldsymbol{\chi}$ which, if it could be observed, would ease the estimation of $\mathbf{b}$. The observed random variable $\mathbf{y}$ which is referred to as the incomplete data within the EM framework, is related to $\boldsymbol{\chi}$ by a mapping $\boldsymbol{\chi} \mapsto \mathbf{y}(\boldsymbol{\chi})$.

The suitable approach for applying the EM algorithm to the problem at hand is to decompose the received vector in (2) into the sum [18]

$$\mathbf{y}(m) = \sum_{k=1}^{K} \mathbf{x}_k(m), \quad m = 1, 2 \ldots, M \quad (5)$$

where

$$\mathbf{x}_k(m) = b_k(m)\mathbf{C}_k\mathbf{F}\mathbf{h}_k + \mathbf{w}_k(m). \quad (6)$$

$\mathbf{x}_k(m)$ represents the received signal component transmitted by the kth user through the channel with impulse response $\mathbf{h}_k$. The Gaussian noise vector, $\mathbf{w}_k(m)$ in (6) represents the portion of $\mathbf{w}(m)$ in the decomposition defined by $\sum_{k=1}^{K} \mathbf{w}_k(m) = \mathbf{w}(m)$, whose variance is $N_0\beta_k$. The coefficient β_k determines that part of the noise power of $\mathbf{w}(m)$ assigned to $\mathbf{x}_k(m)$, satisfying $\sum_{k=1}^{K} \beta_k = 1, 0 \le \beta_k \le 1$.

The problem now is to estimate the transmitted symbols $\mathbf{b} = \{b_k(m)\}_{k=1,m=1}^{K,M}$ and the complex channel responses $\mathbf{h}_k$ for each user, based on observed data $\mathbf{y}$. In the EM algorithm, we view the observed data $\mathbf{y}$ as the incomplete data, and define the complete data as $\boldsymbol{\chi} = \{(\mathbf{x}_1, \mathbf{h}_1), (\mathbf{x}_2, \mathbf{h}_2), \ldots, (\mathbf{x}_K, \mathbf{h}_K)\}$ where $\mathbf{x}_k = [\mathbf{x}_k(1), \ldots, \mathbf{x}_k(M)]^{\text{T}}$ for $k = 1, 2, \ldots, K$. Given the complete data set, the loglikelihood function of the parameter vector to be estimated ($\mathbf{b}$) can be expressed as

$$\log p(\boldsymbol{\chi}|\mathbf{b}) = \sum_{k=1}^{K} \log p(\mathbf{x}_k, \mathbf{h}_k|\mathbf{b}_k) \quad (7)$$

where

$$\log p(\mathbf{x}_k, \mathbf{h}_k|\mathbf{b}_k) = \log p(\mathbf{x}_k|\mathbf{b}_k, \mathbf{h}_k) + \log p(\mathbf{h}_k|\mathbf{b}_k) \quad (8)$$

and, $\mathbf{b}_k = [b_k(1), b_k(2), \ldots, b_k(M)]^{\text{T}}$. Because of the model assumptions, the second conditional pdf on the right hand side of (8) is independent of $\mathbf{b}$. It may, therefore, be discarded since in the following Maximization Step of the EM algorithm involving (7), the maximization is taken over the parameter $\mathbf{b}$, only. Moreover, neglecting those terms independent of $\mathbf{b}$, we have obtained from (6)

$$\log p(\mathbf{x}_k|\mathbf{b}_k, \mathbf{h}_k) \sim \sum_{m=1}^{M} \Re\{b_k(m)\mathbf{h}_k^{\dagger}\mathbf{F}^{\dagger}\mathbf{C}_k^{\text{T}}\mathbf{x}_k(m)\}. \quad (9)$$

Expectation Step (E-Step): The first step to implement the EM algorithm, called the *Expectation Step (E-Step)*, is to compute the average log-likelihood function, denoted by $Q(.|.)$. The conditional expectation is taken over $\boldsymbol{\chi}$ given the observation $\mathbf{y}$ and that $\mathbf{b}$ equals its estimate calculated at the ith iteration as

$$Q\left(\mathbf{b}|\mathbf{b}^{(i)}\right) = E\left\{\log p(\boldsymbol{\chi}|\mathbf{b})|\mathbf{y}, \mathbf{b}^{(i)}\right\}. \quad (10)$$

Taking into account the special form of $\log p(\boldsymbol{\chi}|\mathbf{b})$ in (7), Eq. (10) can be decomposed as

$$Q\left(\mathbf{b}|\mathbf{b}^{(i)}\right) = \sum_{k=1}^{K} Q_k(\mathbf{b}_k|\mathbf{b}^{(i)}) \quad (11)$$

where

$$Q_k(\mathbf{b}_k|\mathbf{b}^{(i)}) = E\left\{\log p(\mathbf{x}_k, \mathbf{h}_k|\mathbf{b}_k)|\mathbf{y}, \mathbf{b}^{(i)}\right\}. \quad (12)$$

Note that after discarding the second term on the right hand side of (8), due to the reasons mentioned above, (12) can be expressed as

$$Q_k(\mathbf{b}_k|\mathbf{b}^{(i)}) = E\left\{\log p(\mathbf{x}_k|\mathbf{b}_k)|\mathbf{y}, \mathbf{b}^{(i)}\right\}. \quad (13)$$

Inserting (9) in (13), we have for $Q_k(\mathbf{b}_k|\mathbf{b}^{(i)})$

$$Q_k(\mathbf{b}_k|\mathbf{b}^{(i)}) = \sum_{m=1}^{M} \Re\{b_k(m)(\mathbf{h}_k^{\dagger}\mathbf{F}^{\dagger}\mathbf{C}_k^{\text{T}}\mathbf{x}_k(m))^{(i)}\} \quad (14)$$

where, adopting the notation used in [14],

$$(\mathbf{h}_k^\dagger \mathbf{F}^\dagger \mathbf{C}_k^T \mathbf{x}_k(m))^{(i)} \triangleq E\left\{\mathbf{h}_k^\dagger \mathbf{F}^\dagger \mathbf{C}_k^T \mathbf{x}_k(m)|\mathbf{y}, \mathbf{b}^{(i)}\right\}, \tag{15}$$

the quantity $(\mathbf{h}_k^\dagger \mathbf{F}^\dagger \mathbf{C}_k^T \mathbf{x}_k(m))^{(i)}$ can be calculated by applying the conditional expectation rule as

$$= E\{\mathbf{h}_k^\dagger E(\mathbf{F}^\dagger \mathbf{C}_k^T \mathbf{x}_k(m)|\mathbf{y}, \mathbf{b}^{(i)}, \mathbf{h})|\mathbf{y}, \mathbf{b}^{(i)}\}$$
$$= E\{\mathbf{h}_k^\dagger \mathbf{F}^\dagger \mathbf{C}_k^T E(\mathbf{x}_k(m)|\mathbf{y}, \mathbf{b}^{(i)}, \mathbf{h})|\mathbf{y}, \mathbf{b}^{(i)}\}. \tag{16}$$

The conditional distribution of $\mathbf{x}_k(m)$ given $\mathbf{y}$, $\mathbf{h}$ and $\mathbf{b} = \mathbf{b}^{(i)}$ is Gaussian with the mean

$$E(\mathbf{x}_k(m)|\mathbf{y}, \mathbf{b}^{(i)}, \mathbf{h}) = b_k^{(i)}(m)\mathbf{C}_k \mathbf{F}\mathbf{h}_k + \beta_k \left(\mathbf{y}(m) - \sum_{j=1}^{K} b_j^{(i)}(m)\mathbf{C}_j \mathbf{F}\mathbf{h}_j\right) \tag{17}$$

where $b_k^{(i)}(m) \triangleq E(b_k(m)|\mathbf{y}, \mathbf{b}^{(i)}, \mathbf{h})$. Inserting (17) in (16) and using the properties $\mathbf{F}^\dagger\mathbf{F} = P\mathbf{I}_P$ and $\mathbf{C}_k^T\mathbf{C}_k = \frac{1}{P}\mathbf{I}_P$ we can rewrite (15) as

$$(\mathbf{h}_k^\dagger \mathbf{F}^\dagger \mathbf{C}_k^T \mathbf{x}_k(m))^{(i)} = (b_k(m))^{(i)} E\{\mathbf{h}_k^\dagger \mathbf{h}_k|\mathbf{y}, \mathbf{b}^{(i)}\} + \beta_k E\{\mathbf{h}_k^\dagger|\mathbf{y}, \mathbf{b}^{(i)}\}\mathbf{F}^\dagger \mathbf{C}_k^T \mathbf{y}(m) - \beta_k \sum_{j=1, j\neq k}^{K} b_j^{(i)}(m) E\{\mathbf{h}_k^\dagger \mathbf{F}^\dagger \mathbf{C}_k^T \mathbf{C}_j \mathbf{F}\mathbf{h}_j|\mathbf{y}, \mathbf{b}^{(i)}\}. \tag{18}$$

On the other hand, since $\mathbf{w} \sim N(0, \sigma^2\mathbf{I})$ and the prior pdf of $\mathbf{h}$ is chosen as $\mathbf{h} \sim N(0, \boldsymbol{\Sigma}_\mathbf{h})$, we can write the conditional pdf's of $\mathbf{h}$ given $\mathbf{y}$ and $\mathbf{b}^{(i)}$ as

$$p(\mathbf{h}|\mathbf{y}, \mathbf{b}^{(i)}) \sim p(\mathbf{y}|\mathbf{h}, \mathbf{b}^{(i)})p(\mathbf{h}) \sim \exp\left[-\frac{1}{\sigma^2}(\mathbf{y} - \mathbf{A}^{(i)}\mathbf{h})^\dagger(\mathbf{y} - \mathbf{A}^{(i)}\mathbf{h}) - \mathbf{h}^\dagger \boldsymbol{\Sigma}_\mathbf{h}^{-1}\mathbf{h}\right].$$

After some algebra it can be shown that

$$p(\mathbf{h}|\mathbf{y}, \mathbf{b}^{(i)}) \sim N(\boldsymbol{\mu}_\mathbf{h}^{(i)}, \boldsymbol{\Sigma}_\mathbf{h}^{(i)}) \tag{19}$$

where

$$\boldsymbol{\mu}_\mathbf{h}^{(i)} = \frac{1}{\sigma^2}\boldsymbol{\Sigma}_\mathbf{h}^{(i)}\mathbf{A}^{(i)\dagger}\mathbf{y}$$
$$\boldsymbol{\Sigma}_\mathbf{h}^{(i)} = \left[\boldsymbol{\Sigma}_\mathbf{h}^{-1} + \frac{1}{\sigma^2}(\mathbf{A}^{(i)})^\dagger\mathbf{A}^{(i)}\right]^{-1} \tag{20}$$

and the matrix $\mathbf{A}$ is defined in (3). Note that the complexity of computing the mean vector and covariance matrix in (20) can be determined as follows. Since $\mathbf{A}$ and $\boldsymbol{\Sigma}_\mathbf{h}^{(i)}$ are $MP \times KL$ and $KL \times KL$ dimensional matrices respectively and $\mathbf{y}$ is an $MP \times 1$ dimensional vector, it can be easily seen that the complex multiplications $KLMP + (KL)^2$ and $(KL)^2MP + (KL)^3$ are required to compute the mean vector and the covariance matrix in (19), respectively. Thus, the total number of multiplications required is $(KL)^2MP + KLMP + (KL)^2 \approx (KL)^2MP$ for $K \gg 1, L \gg 1$.

Now let us compute the terms on the right hand side of Eq. (18). Firstly, we compute $E\{\mathbf{h}_k^\dagger\mathbf{h}_k|\mathbf{y}, \mathbf{b}^{(i)}\}$ as follows. From (19) we have

$$E\{\mathbf{h}\mathbf{h}^\dagger|\mathbf{y}, \mathbf{b}^{(i)}\} = \boldsymbol{\Sigma}_\mathbf{h}{}^{(i)} + \boldsymbol{\mu}_h^{(i)}\boldsymbol{\mu}_h^{(i)\dagger}. \tag{21}$$

For the kth element we get

$$E\{\mathbf{h}_k\mathbf{h}_k^\dagger|\mathbf{y}, \mathbf{b}^{(i)}\} = \boldsymbol{\Sigma}_\mathbf{h}^{(i)}[k, k] + \boldsymbol{\mu}_h^{(i)}[k]\boldsymbol{\mu}_h^{(i)}[k]^\dagger \tag{22}$$

where $\boldsymbol{\Sigma}_\mathbf{h}[i, j]$ denotes the (i, j)th element of the matrix $\boldsymbol{\Sigma}_\mathbf{h}$. We can then calculate $E\{\mathbf{h}_k^\dagger\mathbf{h}_k|\mathbf{y}, \mathbf{b}^{(i)}\}$ from (22) as

$$(\|\mathbf{h}_k\|^2)^{(i)} \triangleq E\{\mathbf{h}_k^\dagger\mathbf{h}_k|\mathbf{y}, \mathbf{b}^{(i)}\} = \mathrm{tr}\left[\boldsymbol{\Sigma}_\mathbf{h}^{(i)}[k, k] + \boldsymbol{\mu}_h^{(i)}[k]\boldsymbol{\mu}_h^{(i)}[k]^\dagger\right]. \tag{23}$$

The second expectation in (18) can be computed as

$$(\mathbf{h}_k)^{(i)} \triangleq E\{\mathbf{h}_k|\mathbf{y}, \mathbf{b}^{(i)}\} = \boldsymbol{\mu}_h^{(i)}[k]. \tag{24}$$

Finally, to calculate the last expectation $E\{\mathbf{h}_k^\dagger \mathbf{F}^\dagger \mathbf{C}_k^T \mathbf{C}_j \mathbf{F}\mathbf{h}_j|\mathbf{y}, \mathbf{b}^{(i)}\}$ in (18), we define $\boldsymbol{\Psi}_j \triangleq \mathbf{C}_j\mathbf{F}$ and $\mathbf{s}_j \triangleq \boldsymbol{\Psi}_j\mathbf{h}_j$. It then follows that

$$\mathbf{s} = \boldsymbol{\Psi}\mathbf{h} = \begin{bmatrix}\mathbf{s}_1\\ \vdots\\ \mathbf{s}_K\end{bmatrix} = \begin{bmatrix}\boldsymbol{\Psi}_1 & 0 & 0\\ 0 & \ddots & 0\\ 0 & 0 & \boldsymbol{\Psi}_K\end{bmatrix}\begin{bmatrix}\mathbf{h}_1\\ \vdots\\ \mathbf{h}_K\end{bmatrix}, \tag{25}$$

$$\boldsymbol{\Sigma}_s^{(i)} \triangleq E[\mathbf{s}\mathbf{s}^\dagger|\mathbf{y}, \mathbf{b}^{(i)}] = E[\boldsymbol{\Psi}\mathbf{h}\mathbf{h}^\dagger\boldsymbol{\Psi}^\dagger|\mathbf{y}, \mathbf{b}^{(i)}] = \boldsymbol{\Psi}\boldsymbol{\Sigma}_\mathbf{h}^{(i)}\boldsymbol{\Psi}^\dagger. \tag{26}$$

Therefore,

$$E\{\mathbf{h}_k^\dagger \mathbf{F}^\dagger \mathbf{C}_k^T \mathbf{C}_j \mathbf{F}\mathbf{h}_j|\mathbf{y}, \mathbf{b}^{(i)}\} = E[\mathbf{s}^\dagger\mathbf{s}|\mathbf{y}, \mathbf{b}^{(i)}] = \mathrm{tr}\left[\boldsymbol{\Sigma}_s^{(i)}[k, j] + \boldsymbol{\mu}_s^{(i)}[k]\boldsymbol{\mu}_s^{(i)}[j]^\dagger\right] \tag{27}$$

where, $\boldsymbol{\mu}_s^{(i)} \triangleq \boldsymbol{\Psi}\boldsymbol{\mu}_h^{(i)}$.

Maximization-Step (M-Step): The second step in implementing the EM algorithm is the *M-Step* where the parameter $\mathbf{b}$ is updated at the $(i + 1)$th iteration according to

$$\mathbf{b}^{(i+1)} = \arg\max_{\mathbf{b}} Q(\mathbf{b}|\mathbf{b}_i) = \sum_{k=1}^{K} Q_k(\mathbf{b}_k|\mathbf{b}^{(i)}). \tag{28}$$

The M-Step can be performed by maximizing $Q_k(\mathbf{b}_k|\mathbf{b}^{(i)})$ individually in (28), as follows

$$\mathbf{b}_k^{(i+1)} = \arg\max_{\mathbf{b}_k} Q_k(\mathbf{b}_k|\mathbf{b}^i) \tag{29}$$

where from (14)

$$Q_k(\mathbf{b}_k|\mathbf{b}^{(i)}) = \sum_{m=1}^{M} b_k(m)\Re\{(\mathbf{h}_k^\dagger \mathbf{F}^\dagger \mathbf{C}_k^T \mathbf{x}_k(m))^{(i)}\}. \tag{30}$$

Moreover, when no coding is used, it follows from (30) that each component of $\mathbf{b}_k^{(i+1)}$ can be separately obtained by maximizing the corresponding summation in the right-hand expression, as follows

$$b_k^{(i+1)}(m) = \mathrm{sgn}\left[\Re\{(\mathbf{h}_k^\dagger \mathbf{F}^\dagger \mathbf{C}_k^T \mathbf{x}_k(m))^{(i)}\}\right] \tag{31}$$

where we have previously obtained that

$$(\mathbf{h}_k^\dagger \mathbf{F}^\dagger \mathbf{C}_k^T \mathbf{x}_k(m))^{(i)} = b_k^{(i)}(m)(\|\mathbf{h}_k\|^2)^{(i)} + \beta_k\left[(\mathbf{h}_k^\dagger)^{(i)}\mathbf{F}^\dagger \mathbf{C}_k^T \mathbf{y}(m) \tag{32}$$
$$- \sum_{j=1}^{K} b_j^{(i)}(m)(\mathbf{h}_k^\dagger \mathbf{F}^\dagger \mathbf{C}_k^T \mathbf{C}_j \mathbf{F}\mathbf{h}_j)^{(i)}\right]. \tag{33}$$

The quantities $(\|\mathbf{h}_k\|^2)^{(i)}$, $(\mathbf{h}_k^\dagger)^{(i)}$ and $(\mathbf{h}_k\mathbf{F}^\dagger\mathbf{C}_k^{\mathrm{T}}\mathbf{C}^j\mathbf{F}\mathbf{h}_j)^{(i)}$ in (32) are given by (23), (24) and (27), respectively. It was shown in [14] that if the length M of the observations frame is large enough, the first term on the right hand side of (27) is negligible compared to the second one. That is,

$$\lim_{m\to\infty}(\mathbf{h}_k\mathbf{F}^\dagger\mathbf{C}_k^{\mathrm{T}}\mathbf{C}_j\mathbf{F}\mathbf{h}_j)^{(i)} \approx \operatorname{tr}\left[\boldsymbol{\mu}_s^{(i)}[k]\boldsymbol{\mu}_s^{(i)}[j]^\dagger\right]$$
$$= \boldsymbol{\mu}_s^{(i)}[k]^\dagger\boldsymbol{\mu}_s^{(i)}[j] \equiv \boldsymbol{\mu}_h^{(i)}[k]^\dagger\boldsymbol{\mu}_h^{(i)}[j]. \quad (34)$$

Note that the identity on the right hand side of (34) follows from the facts that $\boldsymbol{\mu}_s^{(i)} \triangleq \boldsymbol{\Psi}\boldsymbol{\mu}_h^{(i)}$ and $\boldsymbol{\Psi}^\dagger\boldsymbol{\Psi} = \mathbf{I}_{KL}$. Discarding the first term in (27), through a slight rewrite, (32) and (33) can be simplified to

$$b_k^{i+1}(m) = \operatorname{sgn}\left[\Re\left\{b_k^{(i)}(m)\|\boldsymbol{\mu}_h^{(i)}[k]\|^2(1-\beta_k) + \beta_k(\boldsymbol{\mu}_h^{(i)}[j])^\dagger\boldsymbol{\Psi}_k^\dagger\left[\mathbf{y}(m) - \sum_{j=1,j\neq k}^{K} b_j^{(i)}(m)\boldsymbol{\Psi}_j\boldsymbol{\mu}_h^{(i)}[j]\right]\right\}\right]. \quad (35)$$

As a result, Eq. (35) can be interpreted as a joint channel estimation and a data detection with partial interference cancelation. At each iteration step during data detection, the interference-reduced signal is fed into a single user receiver consisting of a conventional coherent detector. As a result, a K-user optimization problem has been decomposed into K independent optimization problems whose resolution is computationally feasible. Finally, it should be concluded that this paper is an extension of the work [14] on the problem of joint channel estimation and data detection for uplink multicarrier CDMA systems operating in the presence of frequency-selective channels. In [14] the same problem is investigated for DC-CDMA systems in the presence of flat fading channels.

3.1. Optimal selection of β_k's

In usual parameter estimation problems in the presence of superimposed signals it has been shown that the optimal values of the coefficients β_k's are chosen as equal weights; that is $\beta_k = 1/K$ [18]. However, the equally selected weights will not be optimal when the received SNR's of each of K users are not equal to each other and if there is some correlation between the super imposed signals, as is the case under consideration here. The optimal β_k values can be determined in this case so as to minimize the bit-error probability as the number of iterations i goes to infinity. Since this is a mathematically intractable nonlinear optimization problem we will adopt a more manageable yet a suboptimal approach presented by Kocian and Fleury [14] and extend their method for the case when each user is affected by a different frequency-selective channel. As pointed out in [14], a tractable way to determine the optimal coefficients of all the users $\boldsymbol{\beta} = [\beta_1, \beta_2, \ldots, \beta_K]^{\mathrm{T}}$ is to minimize the total linear mean-squared error between the true signal components $\boldsymbol{\xi}_k(m) = b_k(m)\mathbf{C}_k\mathbf{F}\mathbf{h}_k$ and their estimated values at the ith iteration $\boldsymbol{\xi}_k^{(i)}(m) \triangleq E\left\{\mathbf{x}_k(m)|\mathbf{y}, \mathbf{b}^{(i)}\right\}$ for $k = 1, 2, \ldots, K$, after projected on $\boldsymbol{\Psi}_k \triangleq \mathbf{C}_k\mathbf{F}$. Thus

$$\boldsymbol{\beta}_{opt}^{(i)} \triangleq \arg\min_{\boldsymbol{\beta}} \sum_{k=1}^{K} E\left\{\left\|\boldsymbol{\Psi}_k^\dagger\left(\boldsymbol{\xi}_k^{(i)}(m) - \boldsymbol{\xi}_k(m)\right)\right\|^2\right\}$$

under the constraints that $\sum_{k=1}^{K}\beta_k = 1$ and $\beta_k \geq 0$. The solution for an optimal $\boldsymbol{\beta}$ is given in the following lemma.

Lemma 1. *Suppose that for each $k = 1, 2, \ldots, K$,*

$P_k^{(i)} - \lambda/2 > 0.$

The optimal β_k's are given by

$$\beta_{k,opt}^{(i)} = \frac{P_k^{(i)} - \lambda/2}{Q_k^{(i)}}, \quad 0 \leq \beta_{k,\mathrm{opt}}^{(i)} \leq 1 \quad \text{and}$$
$$\beta_{k,\mathrm{opt}}^{(i)} > 0 \quad (36)$$

where

$$P_k^{(i)} \triangleq 4\operatorname{tr}(\boldsymbol{\Sigma}_{\mathbf{h}_k})P_{b,k}^{(i)}$$
$$Q_k^{(i)} \triangleq 4\sum_{j=1}^{K}\operatorname{tr}(\boldsymbol{\Upsilon}_{kj}\boldsymbol{\Sigma}_{\mathbf{h}_j}\boldsymbol{\Upsilon}_{kj}^\dagger)P_{b,j}^{(i)}$$
$$\lambda/2 \triangleq \frac{1 - \sum_{r=1}^{K} P_r^{(i)}/Q_r^{(i)}}{\sum_{r=1}^{K} 1/Q_r^{(i)}}$$

and $P_{b,j}^{(i)} \triangleq \operatorname{Prob}\left[b_j^{(i)}(m) \neq b^{(j)}(m)\right]$, $\boldsymbol{\Upsilon}_{kj} \triangleq \mathbf{F}^\dagger\mathbf{C}_k^{\mathrm{T}}\mathbf{C}_j\mathbf{F}$.

Proof. Using the Lagrange optimization method this can be converted into an unconstraint minimization problem as follows

$$\boldsymbol{\beta}_{opt}^{(i)} = \arg\min_{\boldsymbol{\beta}} J(\boldsymbol{\beta})$$

where

$$J(\boldsymbol{\beta}) \triangleq \sum_{k=1}^{K} E\left\{\|\boldsymbol{\Psi}_k^\dagger\left(\boldsymbol{\xi}_k^{(i)}(m) - \boldsymbol{\xi}_k(m)\right)\|^2\right\} + \lambda\left(\sum_{k=1}^{K}\beta_k - 1\right). \quad (37)$$

and λ is a Lagrange coefficient. Taking the expectation with respect to $\mathbf{h}$ in (17) we have

$$\boldsymbol{\xi}_k^{(i)}(m) = b_k^{(i)}(m)\mathbf{C}_k\mathbf{F}\mathbf{h}_k^{(i)} + \beta_k\left(\mathbf{y}(m) - \sum_{j=1}^{K} b_j^{(i)}(m)\mathbf{C}_j\mathbf{F}\mathbf{h}_j^{(i)}\right). \quad (38)$$

Substituting (2) in (38), with $\mathbf{C}_k^{\mathrm{T}}\mathbf{C}_k = \frac{1}{P}\mathbf{I}_P$, assuming $\mathbf{w}(m) \approx 0$ and taking into account the fact that the channel is asymptotically known, that is $\mathbf{h}_k^{(i)} \to \mathbf{h}_k$ as $i \to +\infty$, the terms on the left hand side of (37) can be expressed as

$$\boldsymbol{\Psi}_k^\dagger\boldsymbol{\xi}_k^{(i)}(m) = b_k^{(i)}(m)\mathbf{h}_k + \beta_k\sum_{j=1}^{K}\boldsymbol{\Upsilon}_{kj}\mathbf{h}_j\left(b_j(m) - b_j^{(i)}(m)\right) \quad (39)$$

$$\boldsymbol{\Psi}_k^\dagger\boldsymbol{\xi}_k(m) = b_k(m)\mathbf{h}_k. \quad (40)$$

Note that $\boldsymbol{\Upsilon}_{kk}^{\dagger}\boldsymbol{\Upsilon}_{kk} = \mathbf{I}_L$. Substituting (39) and (40) in (37) and after some algebra yields

$$J(\boldsymbol{\beta}) = (-2\beta_k + \beta_k^2)E\left\{\|\mathbf{h}_k\|^2 \mid b_k(m) - b_k^{(i)}(m)\mid^2\right\} + \beta_k^2 \sum_{j\neq k} E\left\{\mathbf{h}_j^{\dagger}\boldsymbol{\Upsilon}_{kj}^{\dagger}\boldsymbol{\Upsilon}_{kj}\mathbf{h}_j \mid b_j(m) - b_j^{(i)}(m)\mid^2\right\} - \lambda\left(\sum_{k=1}^{K}\beta_k - 1\right). \tag{41}$$

The expectations above can be evaluated as follows.

$$E\left\{\|\mathbf{h}_k\|^2 \mid b_k(m) - b_k^{(i)}(m)\mid^2\right\} = 2\mathrm{tr}(\boldsymbol{\Sigma}_{\mathbf{h}_k})\left[1 - E\{b_k(m)b_k^{(i)}(m)\}\right] = 4\mathrm{tr}(\boldsymbol{\Sigma}_{\mathbf{h}_k})P_{b,k}^{(i)}$$

$$E\left\{\mathbf{h}_j^{\dagger}\boldsymbol{\Upsilon}_{kj}^{\dagger}\boldsymbol{\Upsilon}_{kj}\mathbf{h}_j \mid b_j(m) - b_j^{(i)}(m)\mid^2\right\} = 2\mathrm{tr}(\boldsymbol{\Upsilon}_{kj}\boldsymbol{\Sigma}_{\mathbf{h}_j}\boldsymbol{\Upsilon}_{kj}^{\dagger}) \times \left[1 - E\{b_j(m)b_j^{(i)}(m)\}\right] = 4\mathrm{tr}(\boldsymbol{\Upsilon}_{kj}\boldsymbol{\Sigma}_{\mathbf{h}_j}\boldsymbol{\Upsilon}_{kj}^{\dagger})P_{b,j}^{(i)}$$

Differentiating (41) with respect to $\beta_k, k = 1, 2, \ldots, K$, equating the resulting equations to zero and solving for β_k's and λ, the optimal solutions are obtained as in (36). By hypothesis, the right hand side of (36) is strictly positive for each k. Note that if the assumption in Lemma 1 is not satisfied, then the Lagrange maximization will yield negative or zero values for at least one of the β_k's, indicating that the maximization distribution is located on the boundary. It then becomes necessary to set some of the $\beta_k's$ equal to zero and to try to maximize $J(\boldsymbol{\beta})$ as a function of the remaining variables. In this case the Lemma 1 does not apply and the problem can be solved by a convex programming program.

The bit-error probability $P_{b,j}^{(i)}$ can be evaluated by assuming the performance of the multiuser detector is close to a single user detector performance. In this case $P_{b,j}^{(i)} \approx Q(\sqrt{2\|\mathbf{h}_j\|^2/\sigma^2})$ where $Q(.)$ is the error function defined by $Q(x) = 1/\sqrt{2\pi}\int_x^{\infty} \mathrm{e}^{-t^2/2}\mathrm{d}t$. □

4. Joint data detection and channel estimation with SAGE algorithm (SAGE-JDE)

The SAGE algorithm proposed by Fessler et al. [19] is a twofold generalization of the EM algorithm. First, rather than updating all parameters simultaneously at iteration i, only a subset $\mathbf{b}_k$ of $\mathbf{b}$ indexed by $k = k(i)$ is updated while keeping the parameters in the complement set $\mathbf{b}_{\tilde{k}} \triangleq \mathbf{b} \setminus \mathbf{b}_k$ fixed; and second, the concept of the complete data χ is extended to that of the so-called hidden data χ_k to which the incomplete data $\mathbf{y}$ is related by means of a possibly nondeterministic mapping $\boldsymbol{\chi}_k \mapsto \mathbf{y}(\boldsymbol{\chi}_k)$ exhibiting some particular property. A hidden data space would be a complete data space for $\mathbf{b}_k$ in the EM framework, if $\mathbf{b}_{\tilde{k}}$ were known [19]. The particular property of the mapping $\boldsymbol{\chi}_k \mapsto \mathbf{y}(\boldsymbol{\chi}_k)$ guarantees that the SAGE algorithm exhibits the monotonicity property as well.

The convergence rate of the SAGE algorithm is usually higher than that of the EM algorithm, because the conditional Fisher information matrix of $\boldsymbol{\chi}_k$ given $\mathbf{y}$ for each set of parameters $\mathbf{b}_k$ is likely smaller than that of the complete data χ, given $\mathbf{y}$ for the entire space $\mathbf{b}$. The SAGE algorithm, a generalized form of the EM algorithm [13], allows a more flexible optimization scheme and sometimes converges faster than the EM algorithm. Our main objective is to estimate the transmitted symbols $\mathbf{b} = \{b_k(m)\}_{k=1,m=1}^{K,M}$ for each user k, based on observed data $\mathbf{y}$. The complex channel responses $\mathbf{h} = [\mathbf{h}_1, \mathbf{h}_2, \ldots, \mathbf{h}_K]^{\mathrm{T}}$ are treated as nuisance parameters. In the SAGE algorithm, we view the observed data $\mathbf{y}$ as the incomplete data. At each iteration i, only the data sequence $\mathbf{b}_k = [b_k(1), b_k(2), \ldots, b_k(M)]$ of $\mathbf{b}$ indexed $k = k(i) = i \bmod K$ is updated while keeping the data sequences in the complement set $\mathbf{b}_{\tilde{k}}$ fixed. $\mathbf{b}_{\tilde{k}}$ is the vector obtained by canceling the components of $\mathbf{b}_k$ in $\mathbf{b}$. Then a natural choice for the so-called "hidden-data" set would be $\boldsymbol{\chi} = (\mathbf{y}, \mathbf{h})$.

The SAGE algorithm is defined by the Expectation (E) and Maximization (M) steps as follows: At the ith iteration the E-step computes

$$Q_k\left(\mathbf{b}_k|\mathbf{b}^{(i)}\right) = E\left\{\log p(\boldsymbol{\chi}|\mathbf{b}_k, \mathbf{b}_{\tilde{k}}^{(i)}|\mathbf{y}, \mathbf{b}^{(i)})\right\}. \tag{42}$$

In the M-step, only $\mathbf{b}_k$ is updated as

$$\mathbf{b}_k^{(i+1)} = \arg\max_{\mathbf{b}} Q_k(\mathbf{b}_k|\mathbf{b}^{(i)}) \qquad \mathbf{b}_{\tilde{k}}^{(i+1)} = \mathbf{b}_{\tilde{k}}^{(i)}. \tag{43}$$

Given the complete data set $\boldsymbol{\chi}$, the loglikelihood function of the parameter vector $\mathbf{b}$ to be estimated can be expressed as

$$\log p(\boldsymbol{\chi}|\mathbf{b}) = \log p(\mathbf{y}, \mathbf{h}|\mathbf{b}) = \log p(\mathbf{y}|\mathbf{h}, \mathbf{b}) + \log p(\mathbf{h}|\mathbf{b}). \tag{44}$$

As in the previous section, due to the model assumptions, the second term on the right hand side above may be discarded since it does not depend on $\mathbf{b}$. From (2), the term $\log p(\mathbf{y}|\mathbf{b}, \mathbf{h})$ in (44) can be expressed as

$$\log p(\mathbf{y}|\mathbf{b}, \mathbf{h}) \sim \sum_{m=1}^{M} 2\Re\left\{\left(\sum_{j=1}^{K} b_j(m)\mathbf{C}_j\mathbf{F}\mathbf{h}_j\right)^{\dagger}\mathbf{y}(m)\right\} - \left\|\sum_{j=1}^{K} b_j(m)\mathbf{C}_j\mathbf{F}\mathbf{h}_j\right\|^2. \tag{45}$$

Inserting (45) in (42), we have for $Q_k(\mathbf{b}_k|\mathbf{b}^{(i)})$

$$Q_k(\mathbf{b}_k|\mathbf{b}^{(i)}) = \sum_{m=1}^{M}\Re\left\{b_k(m)(\mathbf{h}_k^{(i)})^{\dagger}\mathbf{F}^{\dagger}\mathbf{C}_k^{\mathrm{T}}\mathbf{y}(m) - \sum_{j=1,j\neq k}^{K} b_j(m)(\mathbf{h}_k^{\dagger}\mathbf{F}^{\dagger}\mathbf{C}_k^{\mathrm{T}}\mathbf{C}_j\mathbf{F}\mathbf{h}_j)^{(i)}\right\} \tag{46}$$

where the quantities $\mathbf{h}_k^{(i)}$ and $(\mathbf{h}_k^{\dagger}\mathbf{F}^{\dagger}\mathbf{C}_k^{\mathrm{T}}\mathbf{C}_j\mathbf{F}\mathbf{h}_j)^{(i)}$ are defined as

$$\mathbf{h}_k^{(i)} \triangleq E(\mathbf{h}_k \mid \mathbf{y}, \mathbf{b}^{(i)})$$

$$(\mathbf{h}_k^{\dagger}\mathbf{F}^{\dagger}\mathbf{C}_k^{\mathrm{T}}\mathbf{C}_j\mathbf{F}\mathbf{h}_j)^{(i)} \triangleq E\left\{\mathbf{h}_k^{\dagger}\mathbf{F}^{\dagger}\mathbf{C}_k^{\mathrm{T}}\mathbf{C}_j\mathbf{F}\mathbf{h}_j \mid \mathbf{y}, \mathbf{b}^{(i)}\right\}.$$

Table 1
Users transmission power.

User	Linear	Logarithmic (dB)
1	1	0
2	0.9560	−0.1954
3	0.9139	−0.3909
4	0.8737	−0.5863
5	0.8353	−0.7817
6	0.7985	−0.9772
7	0.7634	−1.1726
8	0.7298	−1.3680
9	0.6977	−1.5635
10	0.6670	−1.7589
11	0.6376	−1.9543
12	0.6096	−2.1498
13	0.5827	−2.3452
14	0.5571	−2.5406
15	0.5326	−2.7361
16	0.5092	−2.9315

These quantities can be computed by (24) and (27), respectively.

The M-Step can be performed by maximizing each summand of the right-hand side expression individually in (46). After some algebra the final result is as follows.

$$b_k^{(i+1)}(m) = \mathrm{sgn}\left[\Re\left\{((\mathbf{h}_k^\dagger)^{(i)}\mathbf{F}^\dagger\mathbf{C}_k^T\mathbf{y}(m) - \sum_{j=1,j\neq k}^{K} b_j^{(i)}(m)(\mathbf{h}_k\mathbf{F}^\dagger\mathbf{C}_k^T\mathbf{C}_j\mathbf{F}\mathbf{h}_j)^{(i)}\right\}\right]. \quad (47)$$

If the observation frame length M is large enough, we can again neglect the first term in (27) and (47) can be approximately expressed as

$$b_k^{i+1}(m) = \mathrm{sgn}\left[\Re\left\{\boldsymbol{\mu}_h^{(i)}[j]\boldsymbol{\Psi}_k^T \times \left[\mathbf{y}(m) - \sum_{j=1,j\neq k}^{K} b_j^{(i)}(m)\boldsymbol{\Psi}_j\boldsymbol{\mu}_h^{(i)}[j]\right]\right\}\right]. \quad (48)$$

According to (48), the tentative decisions of the bits are used to calculate an estimate of the MAI which is likely to be increasingly reliable with iteration i.

5. Simulations

In this section, the performance of an uplink MC-CDMA system based on a proposed receiver operating over frequency-selective channels is investigated by computer simulations. In the simulations, it is assumed that all users receive different average signal powers, chosen according to the values in the Table 1. The orthogonal Walsh sequences are selected as a spreading code and the processing gain is equal to the number of subcarriers ($P = 16$). The number of users selected is $K = 16$ and each user sends a frame over the fading channel which is composed of T preamble bits, and D data bits. The BER performances of several receiver types are investigated below as a function of SNR per information bit. Wireless channels between mobiles' antennas and the receiver antenna are modeled based on a realistic channel model determined by the COST-207 project in which the Typical Urban (TU) channel model is considered to have the channel length $L = 4$ and the covariance $\boldsymbol{\Sigma}_\mathbf{h}$. The channel tap gains are given in Table 2. BPSK signal modulation format has been adopted with a bandwidth of 1.228 MHz (Qual Comm-CDMA).

Table 2
Taps power.

Delay (μs)	Linear	Logarithmic (dB)
0	0.6564	−1.8286
0.81	0.2086	−6.8072
1.62	0.0790	−11.0210
2.44	0.0560	−12.5171

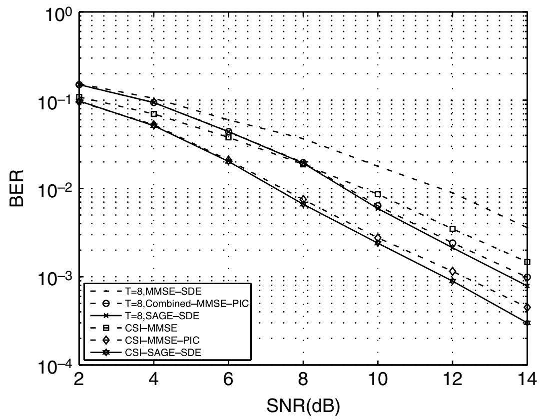

Fig. 1. BER performances of SDE receivers ($T = 8$, $D = 40$).

Traditional receivers for MC-CDMA systems are based on separate estimation and detection (SDE) methods whose performances are limited by the number of used preamble bits. Therefore we first investigate BER performances of SDE receivers for different lengths of preamble bits. In the receiver, the initial MMSE channel estimate is obtained by using T preamble bits while the channel covariance matrix $\mathbf{C}_h$ is assumed to be known. An initial MMSE estimate of D data bits is computed from the observation vector $\mathbf{y}$ while assuming the channel coefficients have already been estimated. We will refer to this method as the MMSE separate detection and estimation (MMSE-SDE) scheme. If the output of the (MMSE-SDE) is applied to a parallel interference cancelation (PIC) receiver or the SAGE receiver, the resulting receiver structures are referred to the Combined MMSE-PIC and the SAGE-SDE, respectively. There are two existing strategies on how to rank the users for SAGE receivers. The first one is that the users are sorted according to their estimated strength, so that the user with the weakest received signal is ranked first. The other one is that the users are ranked in order of decreasing strength.

In these simulations, the first sorting method is used for all SAGE simulations. Note that in [14], the first sorting method yields better performance results. Moreover, we also determined the performances of MMSE-SDE, Combined MMSE-PIC and SAGE-SDE receivers for the perfect channel state information (CSI) case are referred to

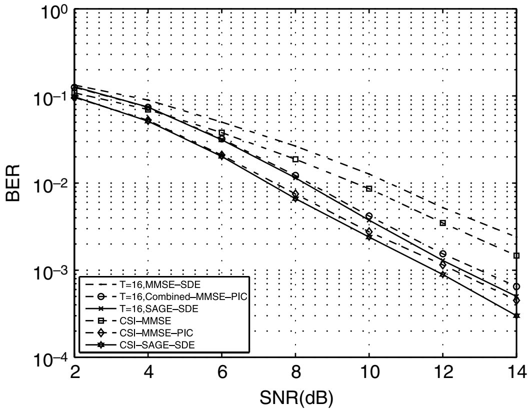

Fig. 2. BER performances of SDE receivers ($T = 16$, $D = 40$).

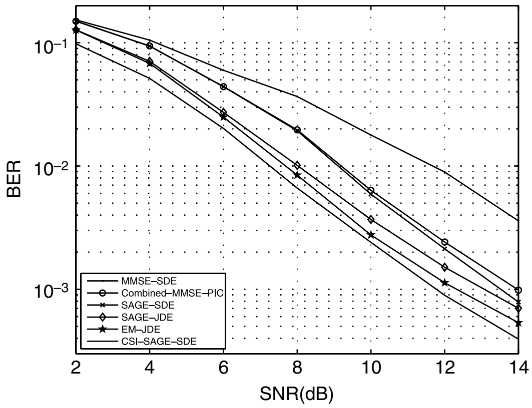

Fig. 3. BER performances of SDE and JDE receivers ($T = 8$, $D = 40$).

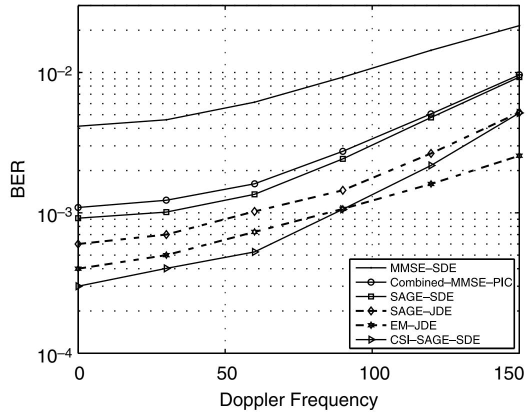

Fig. 4. BER performances of SDE and JDE receivers in the case of channel is time varying ($T = 8$, $D = 40$).

CSI-MMSE, CSI-Combined MMSE-PIC and CSI-SAGE-SDE, respectively.

Fig. 1 compares the BER performances of the MMSE-SDE, Combined-MMSE-PIC, SAGE-SDE, CSI-MMSE, CSI-Combined MMSE-PIC and the CSI-SAGE-SDE schemes as a function of SNR. For fair comparison, we simulated Combined-MMSE-PIC, SAGE-SDE, CSI-Combined MMSE-PIC and CSI-SAGE-SDE receivers employing only four stages. For fair comparison, we simulated all receivers employing only three stages. One stage corresponds to the number of iterations required to update every user's bit sequence once, i.e., one iteration in the Combined-MMSE-PIC scheme and K iterations in the SAGE-SDE scheme. It is observed that the SAGE-SDE and the Combined MMSE-PIC receivers, based on the interference cancelation, outperform the MMSE-SDE receiver. On the other hand, it is well known that the successive interference cancelation (SIC) scheme outperforms the PIC scheme when the received signals have distinctly different strengths. Therefore, the SAGE-SDE receiver performance is better than the Combined-MMSE-PIC. Note that the SAGE-SDE receiver needs more time to update the user's data since the interference is canceled successively. It is also clear that the processing time increases with the number of active users in the system.

From the simulation results in Fig. 2, we can see that as length of preamble sequence increases to 16, the SAGE-SDE and Combined MMSE-PIC receiver performances approach the CSI cases slightly. Moreover, it was shown that Combined-MMSE-PIC-CSI and SAGE-SDE-CSI gain by about 3 dB over the MMSE-SDE at BER $= 10^{-3}$. In practice, this is unfeasible because of the effective usage of bandwidth requirements. Moreover, increasing the preamble sequence will increase the SNR per information bit due to low rate. Thus, in the following, we simulated the proposed two joint-channel estimation and data detection (JDE) methods to improve the performance at shorter preamble sequence lengths.

Fig. 3 presents the simulation results where BER performances of the JDE methods are compared with that of the SDE methods. The MMSE-SDE technique has been used to initialize the EM-JDE and SAGE-JDE receivers. For a fair comparison, we simulated SDE and JDE methods employing only four stages. In Fig. 3, it is observed that the proposed JDE techniques outperform all the SDE approaches when the channel is unknown. As mentioned earlier, the SIC scheme outperforms the PIC scheme if the estimation and detection steps are implemented separately. The validity of this assertion has also been shown in [14] for joint channel estimation and data detection in the DS-CDMA systems over flat Rayleigh fading channels. Therefore, it is expected that SAGE-JDE will outperform EM-JDE, in which the channel coefficients are updated only once at every stage, rather than K times as performed in the SAGE-JDE receiver. This is due to the following reason: In the SAGE-JDE receiver, at each iteration, the bit sequence of only one of the users is updated and the other user's bit sequences are updated successively, while the channel coefficients are reestimated in parallel after completing the updating of each sequence. Therefore, at the first iteration, only the first column of the $\mathbf{A}$ matrix in Eq. (3) is updated and the channel coefficients are reestimated according to the updated $\mathbf{A}$ matrix in Eq. (20).

Consequently, the channel updating process is not as efficient as the one employed in the EM-JDE scheme. This is due to the fact that the EM-JDE receiver updates the channel after having completed the update all columns of in the **A** matrix as contrast to the SAGE-JDE technique. As a result, the performance of the SAGE-JDE scheme appears to be worse than the EM-JDE for the parameters chosen for the simulations.

In computer simulations, so far, the channel was assumed to be constant (static) regardless of any changes in the impulse response of the mobile channel (as a function of the Doppler frequency (Hz)). In Fig. 4, BER performances of the JDE and SDE receivers employing three stages are presented in the presence of different Doppler frequencies for SNR $= 14$ dB. Based on results presented in Fig. 4, we have concluded that JDE receivers are more robust against channel variations than SDE receivers. Therefore, we conclude that the JDE methods are very good candidates for operating over static as well as quasi-static channels.

Finally to make a fair comparison between the different estimation techniques in terms of computational complexity, we have the following observation. The joint estimation of the channel coefficients basically dominates the computational complexity of both EM-JDE and SAGE-JDE algorithms. For a fair comparison, the algorithms are simulated employing only four stages. One stage corresponds to one iteration in EM-based algorithms while K iteration in SAGE based algorithms are required to update every user's bit sequence. Based on the complexity analysis presented in Section 3, we have concluded that the complexity per iteration of the EM-JDE and SAGE-JDE algorithm is bounded by $O(K^2L^2MP)$ and $O(K^3L^2MP)$.

6. Conclusions

We presented two efficient iterative receiver structures of tractable complexity for the joint multiuser detection and multichannel estimation (JDE) of direct-sequence code-division multiple-access signals. The schemes result from an application of EM and SAGE algorithms, respectively. The EM-JDE receiver updates the data bit sequences in parallel, while the SAGE-JDE receiver reestimates them successively. The channel parameters are updated in parallel in both schemes. A closed form expression was derived for the data detection which incorporates the channel estimation as well as the partial interference cancelation steps in the algorithm. It was concluded that few pilot symbols were sufficient to initiate the EM-JDE and SAGE-JDE algorithms very effectively. A comparison with other previously known receiver structures was also made. These computer simulations demonstrated the effectiveness of the proposed algorithms in terms of BER performances when the channel needs to be estimated. We conclude that the EM-JDE and SAGE-JDE which smartly combine the data detection and channel estimation in multiuser systems, are robust unlike architectures where both process are implemented separately and we observed that the EM-JDE performed better than the SAGE-JDE. Finally, we have demonstrated that JDE receivers are more robust against the channel variations than SDE receivers.

Acknowledgement

This research has been conducted within the NEWCOM++ Network of Excellence in Wireless Communications funded through the EC 7th Framework Programme.

References

[1] N. Yee, J.-P. Linnarz, G. Fettweis, Multi-carrier CDMA in indoor wireless radio networks, in: Proc. IEEE Int. Symp. on Personal, Indoor and Mobile Radio Commun., PIMRC 93, September 1993, pp. 109–113.

[2] K. Fazel, L. Papke, On the performance of convolutionally-coded CDMA/OFDM for mobile communication system, in: Proc. IEEE Int. Symp. on Personal, Indoor and Mobile Radio Commun., PIMRC 93 September 1993, pp. 468–472.

[3] S. Hara, R. Prasad, Overview of Multicarrier CDMA, IEEE Commun. Magazine (1997) 126–133.

[4] Z. Xie, R.T. Short, C.K. Rushfortli, Family of suboptimum detectors for coherent multiuser communications, IEEE J. Select. Areas Commun. 8 (4) (1990) 683–690.

[5] V. Kuhn, Combined MMSE-PIC in coded OFDM-CDMA systems, IEEE GLOBECOM 01 (2001) 231–235.

[6] X. Wu, Q. Yin, A. Feng, K. Deng, Equivalently blind time-domain channel estimation for MC-CDMA systems over frequency selective fading channels in multiuser scenario, IEEE Trans. Veh. Technol. (2001) 2687–2691.

[7] P. Marques, A. Pereira, A. Gameiro, Pilot and data aided channel estimation for uplink MC-CDMA mobile systems, in: EUSIPCO, September 04–08 2005, Antalya, Turkey.

[8] L. Sanguinetti, M. Morelli, Channel acquisition and tracking for MC-CDMA uplink transmissions, IEEE Trans. Veh. Technol. 55 (3) (2006).

[9] X. Ma, H. Koyabashi, S. Schwartz, EM-Based channel estimation algorithms for OFDM, EURASIP J. Appl. Signal Process. (10) (2004) 1460–1477.

[10] H. Dogan, H.A. Çirpan, E. Panayirci, Iterative channel estimation and decoding of turbo coded SFBC-OFDM systems, IEEE Trans. Wireless Commun. 6 (7) (2007).

[11] H.A Cirpan, E. Panayirci, H. Dogan, Nondata-aided channel estimation for OFDM systems with space-frequency transmit diversity, IEEE Trans. Veh. Technol. 55 (2) (2006) 449–457.

[12] E. Panayirci, H. Dogan, H.A. Cirpan, B.H. Fleury, Joint data detection and channel estimation for uplink MC-CDMA systems over frequency selective channels, in: 6th Int. Workshop on Multi-Carrier Spread Spectrum, MC-SS 2007, 2007, Herrsching, Germany.

[13] X. Yongzhe, C.N. Georghiades, Two EM-type channel estimation algorithms for OFDM with transmitter diversity, IEEE Trans. Commun. 51 (1) (2003) 106–115.

[14] A. Kocian, B.H. Fleury, EM-based joint data detection and channel estimation of DS-CDMA signals, IEEE Trans. Commun. 51 (10) (2003) 1709–1720.

[15] M. Morelli, Timing and frequency synchronization for the uplink of an OFDMA system, IEEE Trans. Commun. 52 (2) (2004) 296–306.

[16] A. Kocian, I. Land, B.H. Fleury, Joint channel estimation, partial successive interference cancelation, and data decoding for DS-CDMA based on the SAGE algorithm, IEEE Trans. Commun. 55 (6) (2007).

[17] B. Lu, X. Wang, Y.G. Li, Iterative receivers for space-time block-coded OFDM systems in dispersive fading chanels, IEEE Trans. Wireless Commun. 1 (2) (2002) 213–225.

[18] M. Feder, E. Weinstein, Parameter Estimation of superimposed signals using the EM algorithm, IEEE Tran. Acoust. Speech, Signal Process. 36 (1988) 477–489.

[19] J.A. Fessler, A.O. Hero, Space-alternating generalized expectation-maximization algorithm, IEEE Trans. Signal Process. 42 (1994) 2664–2677.

Erdal Panayırcı received the Diploma Engineering degree in electrical engineering from the Istanbul Technical University, Istanbul, Turkey, in 1964 and the Ph.D. degree in electrical engineering and system science from Michigan State University, East Lansing, in 1970. From 1970 to 2000, he was with the Faculty of Electrical and Electronics Engineering, Istanbul Technical University, where he was a Professor and the Head of the Telecommunications Chair. He has also been a part-time Consultant to several leading companies in telecommunications in Turkey. From 1979 to 1981, he was with the Department of Computer Science, Michigan State University, as a Fulbright-Hays Fellow and a NATO Senior Scientist. Between 1983 and 1986, he served as a NATO Advisory Committee Member for the Special Panel on Sensory Systems for Robotic Control. From August 1990 to December 1991, he was a Visiting Professor at the Center for Communications and Signal Processing, New Jersey Institute of Technology, Newark, and took part in the research project on interference cancelation by array processing. Between 1998 and 2000, he was a Visiting Professor at the Department of Electrical Engineering, Texas AM University, College Station, and took part in research on developing efficient synchronization algorithms for orthogonal frequency-division multiplexing (OFDM) systems. In 2005 he was a Visiting Professor at the

Department of Electrical and Electronics Engineering, Bilkent University, Ankara, Turkey. He is currently a Professor and Department Head at the Electronics Engineering Department, Kadir Has University, Istanbul, Turkey. He is engaged in research and teaching in digital communications and wireless systems, equalization and channel estimation in multicarrier (OFDM) communication systems, and efficient modulation and coding techniques (TCM and turbo coding). Prof. Panayırcıis a member of Sigma Xi. He was the Editor for the IEEE TRANSACTIONS ON COMMUNICATIONS in the fields of synchronization and equalizations from 1995 to 1999. He is currently the Head of the Turkish Scientific Commission on Signals, Systems, and Communications of the International Union of Radio Science.

Hakan Doğan Hakan Dogan was born in Istanbul, Turkey, on 1979. He received the B.S., M.S. and and the Ph.D. degrees in electronics engineering from Istanbul University, Istanbul, Turkey, in 2001, 2003, 2007 respectively. From 2001 to 2007, he was a Research Assistant at the faculty of the Department of Electrical and Electronics Engineering, University of Istanbul, working on signal processing algorithms for wireless communication systems. In 2007, he joined the same faculty as an Assistant Professor. His general research interests cover communication theory, estimation theory, statistical signal processing, and information theory. His current research areas are focused on wireless communication concepts with specific attention to equalization and channel estimation for spread-spectrum and multicarrier (orthogonal frequency-division multiplexing) systems.

Hakan A. Çırpan (M97) received the B.S. degree from Uludag University, Bursa, Turkey, in 1989, the M.S. degree from the University of Istanbul, Istanbul, Turkey, in 1992, and the Ph.D. degree from Stevens Institute of Technology, Hoboken, NJ, in 1997, all in electrical engineering. From 1995 to 1997, he was a Research Assistant at Stevens Institute of Technology, working on signal processing algorithms for wireless communication systems. In 1997, he joined the faculty of the Department of Electrical and Electronics Engineering, University of Istanbul. His general research interests cover wireless communications, statistical signal and array processing, system identification, and estimation theory. His current research activities are focused on signal processing and communication concepts with specific attention to channel estimation and equalization algorithms for space-time coding and multicarrier (orthogonal frequency-division multiplexing) systems. Dr. Çirpan is a member of Sigma Xi. He received the Peskin Award from Stevens Institute of Technology as well as the Prof. Nazım Terzioglu Award from the Research Fund of the University of Istanbul.

Alexander Kocian was born in Vienna, Austria, on January 22, 1971. He received the Dipl. Ing. degree (with distinction) in electrical engineering from Vienna University of Technology, Vienna, Austria, in 1997. He is currently working toward the Ph.D. degree at the Department of Communication Technology, Aalborg University, Aalborg, Denmark. From 1997–1999, he was with the Spread Spectrum Team at the Communication Technology Laboratory at the Swiss Federal Institute of Technology (ETH), Zurich, Switzerland. In 1999, he joined the faculty of Aalborg University. He was a Visiting Research Scholar at the Wireless Systems Laboratory, Georgia Institute of Technology, Atlanta, in 2001. His research interests include joint data detection and channel estimation in multiple- access communication systems and characterization of multiple-input multiple-output (MIMO) channels.

Bernard H. Fleury received the diploma in electrical engineering and mathematics in 1978 and 1990 respectively, and the doctoral degree in electrical engineering in 1990 from the Swiss Federal Institute of Technology Zurich (ETHZ), Switzerland. Since 1997 Bernard H. Fleury has been with the Department of Communication Technology, Aalborg University, Denmark, where he is Professor in Digital Communications. He has also been affiliated with the Telecommunication Research Center, Vienna (ftw.) since April 2006. Bernard H. Fleury is presently Chairman of Department 2 Radio Channel Modeling for Design Optimization and Performance Assessment of Next Generation Communication Systems of the on-going FP6 network of excellence NEWCOM (Network of Excellence in Communications). During 1978–85 and 1988–92 he was Teaching Assistant and Research Assistant, respectively, at the Communication Technology Laboratory and at the Statistical Seminar at ETHZ. In 1992 he joined again the former laboratory as Senior Research Associate. In 1999 he was elected IEEE Senior Member. Bernard H. Fleurys general fields of interest cover numerous aspects within Communication Theory and Signal Processing mainly for Wireless Communications. His current areas of research include stochastic modeling and estimation of the radio channel, characterization of multiple-input multiple-output (MIMO) channels, and iterative (turbo) techniques for joint channel estimation and data detection/decoding in multi-user communication systems.