An Efficient Joint Data Detection and Channel Estimation Technique for Uplink MC-CDMA Systems Based on SAGE Algorithm

Erdal Panayırcı, Hakan Doğan, Hakan A. Çırpan, Bernard H. Fleury

Abstract

This paper is concerned with joint channel estimation and data detection for uplink multicarrier code-division multiple-access (MC-CDMA) systems in the presence of frequency fading channel. The detection and estimation algorithm, implemented at the receiver, is based on a version of the expectation maximization (EM) technique, called the spece-alternating-generalized-expectation-maximization (SAGE) algorithm which is very suitable for the the multicarrier signal formats. Application of the SAGE 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 algorithm has excellent BER end estimation performance.

Index Terms: Joint data detection and channel estimation, MC-CDMA Systems, SAGE algorithm.

I. Introduction

In this paper we consider an efficient iterative algorithm based on the SAGE technique for multi-user data detection and channel estimation, jointly for uplink MC-CDMA systems in the presence of frequency selective fading channels. The SAGE algorithm is a broadly applicable approach to the iterative computation of parameters from intractable and high complexity likelihood functions. An EM approach proposed for the superimposed signals [1] is applied to the channel estimation for OFDM systems [2], [3] and compared with SAGE version in [4]. As will be seen shortly, a partial parallel interference cancelation (PIC) is incorporated into the resulting detection algorithm [5], [6]. The work is an extension of [7] in which joint data detection and channel estimation of uplink DS-CDMA systems were considered in the presence of flat Rayleigh channels. We extend 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 in case of a downlink transmission.

Notation: Vectors (matrices) are denoted by boldface lower (upper) case letters; all vectors are column vectors; $(.)^*$, $(.)^T$ 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.

II. Signal Model

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

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

This research has been conducted within the NEWCOM Network of Excellence in Wireless Communications funded through the EC 6th Framework Programme and the Research Fund of Istanbul University under Projects UDP-921/09052007,T-856/02062006. This work was also supported in part by the Turkish Scientific and Technical Research Institute (TUBITAK) under Grant 104E166.

Hakan Doğan and Hakan A. Çırpan are with the Department of Electrical and Electronics Engineering, Istanbul University, Avcilar 34850, Istanbul, Turkey, e-mail: {hdogan,hcirpan}@istanbul.edu.tr

Erdal Panayırcı is with the Department of Electronics Enginering, Kadir Has University, Cibali 34083, Istanbul, Turkey, e-mail: eepanay@khas.edu.tr

Bernard H. Fleury is with the Section Navigation and Communications, Department of Electronic Systems, Aalborg University Fredrik Bajers Vej 7A3 DK-9000 Aalborg, Denmark and Forschungszentrum Telekommunikation Wien (ftw.) Donau-City-Strasse 1 A-1220 Vienna, Austria e-mail: bfl@kom.auc.dk

where L is the number of paths in the kth users channel; g_{kl} and τ_{kl} are, respectively, the complex fading processes and the delay of the kth users lth path. It is assumed that only the second-order statistics of the fading processes are known to the receiver. Also, fading can vary from symbol to symbol but remains constant over a symbol interval.

At the receiver, the received signal is first serial-to-parallel (S/P) converted, then CP is removed and DFT is then applied to the received discrete time signal to obtain the received vector expressed as

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

where $b_k(m)$ denotes data sent by the user k within the mth symbol; $\mathbf{C}_k = \text{diag}(\mathbf{c}_k)$ with $\mathbf{c}_k = [c_{k1}, c_{k1}, ..., c_{kP}]^T$ where each chip, c_{ik}, takes values in the set $\{-\frac{1}{\sqrt{P}}, \frac{1}{\sqrt{P}}\}$ denoting the kth users spreading code ; $\mathbf{F} \in \mathbb{C}^{P\times L}$ denotes the DFT matrix with the (k, l)th element given by $\frac{1}{\sqrt{P}}e^{-j2\pi kl/P}$; 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 $\mathbf{y}$=$[\mathbf{y}^T(1), ..., \mathbf{y}^T(M)]^T$. Then the received signal model can be written as

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

and can be rewritten in more succinct form

$$\mathbf{y} = \mathbf{A}\mathbf{h} + \mathbf{w} \tag{4}$$

By using the assumption; $\mathbf{h}_k \sim N(0, \mathbf{\Sigma}_{\mathbf{h}_k})$ where $\mathbf{\Sigma}_{\mathbf{h}_k} = E[\mathbf{h}_k\mathbf{h}_k^\dagger]$, we have $\mathbf{h} \sim N(0, \mathbf{\Sigma}_{\mathbf{h}})$ with $\mathbf{\Sigma}_{\mathbf{h}} = \text{diag}[\mathbf{\Sigma}_{\mathbf{h}_1}, \cdots, \mathbf{\Sigma}_{\mathbf{h}_K}]$.

III. Joint Data Detection and Channel Estimation

The problem of interest is to derive an iterative algorithm based on the SAGE algorithm for data detection and channel estimation jointly employing the signal model given by (2). The SAGE algorithm, a generalized form of the EM algorithm [4], allows a more flexible optimization scheme and sometimes converges faster than the EM algorithm. Since the SAGE method has been studied and applied to a number of problems in communications over the years, the details of the algorithm will not be presented in this paper. The reader is suggested to read [8] for a general exposition to SAGE algorithm and [7] for its application to the estimation problem related to the work herein.

Our main objective 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 $\{\mathbf{h}_k\}$ 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 $\mathbf{b}_k = [b_k(1), b_k(2), \cdots, b_k(M)]$ of $\mathbf{b}$ indexed $k = k(i) = i$ mod 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 $\boldsymbol{\chi} = \{(\mathbf{y}_1, \mathbf{h}_1), (\mathbf{y}, \mathbf{h})\}$.

The SAGE algorithm is defined by the Expectation(E) and Maximization(M) steps as follow:

At the ith iteration the E-step computes

$$Q_k\left(\mathbf{b}_k|\mathbf{b}^{(i)}\right) = E\left\{\log p(\boldsymbol{\chi}|\mathbf{b}_k, \mathbf{b}_{\tilde{k}}^{(i)}|\mathbf{y}, \mathbf{b}^{(i)}\right\} \tag{5}$$

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

$$\mathbf{b}_k^{(i+1)} \;=\; \arg\max_{\mathbf{b}} Q_k(\mathbf{b}_k|\mathbf{b}^{(i)}) \tag{6}$$

$$\mathbf{b}_{\tilde{k}}^{(i+1)} \;=\; \mathbf{b}_{\tilde{k}}^{(i)} \tag{7}$$

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

$$\log p(\boldsymbol{\chi}|\mathbf{b}) = \log p(\mathbf{y}, \mathbf{h}|\mathbf{b}), \tag{8}$$

where

$$\log p(\mathbf{y}, \mathbf{h}|\mathbf{b}) = \log p(\mathbf{y}|\mathbf{b}, \mathbf{h}) + \log p(\mathbf{h}|\mathbf{b}). \tag{9}$$

We neglect the $\log p(\mathbf{h}|\mathbf{b})$ term in (9) since the data sequence $\mathbf{b}$ and $\mathbf{h}$ are independent of each other.

A. Implementation of the Expectation Step *(E-Step)*

From (2), the term $\log p(\mathbf{y}|\mathbf{b},\mathbf{h})$ in (9) can be expressed as

$$\log p(\mathbf{y}|\mathbf{b},\mathbf{h}) \sim \sum_{m=1}^{M} 2\Re\left\{\left(\sum_{j=1}^{K} b_j(m)\mathbf{C}_j\mathbf{F}\mathbf{h}_j\right)^{\dagger}\mathbf{y}(m)\right\} - \|\left(\sum_{j=1}^{K} b_j(m)\mathbf{C}_j\mathbf{F}\mathbf{h}_j\right)^{\dagger}\mathbf{y}(m)\|^2. \tag{10}$$

Inserting (10) in (5), we have for $Q_k(\mathbf{b}_k|\mathbf{b}^{(i)})$

$$Q_k(\mathbf{b}_k|\mathbf{b}^{(i)}) = \sum_{m=1}^{M} b_k(m)\Re\left\{(\mathbf{h}_k^{[i]})^{\dagger}\mathbf{F}^{\dagger}\mathbf{C}_k^T\mathbf{y}(m) - \sum_{j=1,j\neq k}^{K} b_j(m)(\mathbf{h}_k^{\dagger}\mathbf{F}^{\dagger}\mathbf{C}_k^T\mathbf{C}_j\mathbf{F}\mathbf{h}_j)^{[i]}\right\} \tag{11}$$

where, adopting the notation used in [7],

$$(\mathbf{h}_k)^{[i]} \triangleq E\left\{\mathbf{h}_k|\mathbf{y},\mathbf{b}^{(i)}\right\} \tag{12}$$

$$(\mathbf{h}_k^{\dagger}\mathbf{F}^{\dagger}\mathbf{C}_k^T\mathbf{C}_j\mathbf{F}\mathbf{h}_j)^{[i]} \triangleq E\left\{\mathbf{h}_k^{\dagger}\mathbf{F}^{\dagger}\mathbf{C}_k^T\mathbf{C}_j\mathbf{F}\mathbf{h}_j|\mathbf{y},\mathbf{b}^{(i)}\right\} \tag{13}$$

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

$$\begin{aligned} p(\mathbf{h}|\mathbf{y},\mathbf{b}^{(i)}) &\sim p(\mathbf{y}|\mathbf{h},\mathbf{b}^{(i)})p(\mathbf{h}) \\ &\sim \exp\left[-\frac{1}{\sigma^2}(\mathbf{y}-\mathbf{A}\mathbf{h})^{\dagger}(\mathbf{y}-\mathbf{A}\mathbf{h}) - \mathbf{h}^{\dagger}\mathbf{\Sigma}_{\mathbf{h}}^{-1}\mathbf{h}\right]. \end{aligned} \tag{14}$$

After some algebra it can be shown that

$$p(\mathbf{h}|\mathbf{y},\mathbf{b}^{(i)}) \sim N(\boldsymbol{\mu}_{\mathbf{h}}^{(i)},\mathbf{\Sigma}_{\mathbf{h}}^{(i)}) \tag{15}$$

where

$$\begin{aligned} \boldsymbol{\mu}_{\mathbf{h}}^{(i)} &= \frac{1}{\sigma^2}\mathbf{\Sigma}_{\mathbf{h}}^{(i)}\mathbf{A}^{(i)\dagger}\mathbf{y} \\ \mathbf{\Sigma}_{\mathbf{h}}^{(i)} &= \left[\mathbf{\Sigma}_{\mathbf{h}}{}^{-1} + \frac{1}{\sigma^2}(\mathbf{A}^{(i)})^{\dagger}\mathbf{A}^{(i)}\right]^{-1} \end{aligned} \tag{16}$$

and the matrix $\mathbf{A}$ is defined in (3) and (4).

The expectation in (12) can be computed as

$$(\mathbf{h}_k)^{[i]} \triangleq E\{\mathbf{h}_k|\mathbf{y},\mathbf{b}^{(i)}\} = \boldsymbol{\mu}_h^{(i)}[k]. \tag{17}$$

To calculate the 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 (13), define $\mathbf{\Psi}_j \triangleq \mathbf{C}_j\mathbf{F}$ and $\mathbf{s}_j \triangleq \mathbf{\Psi}_j\mathbf{h}_j$. It then follows that

$$\mathbf{s} = \mathbf{\Psi}\mathbf{h} = \begin{bmatrix} \mathbf{s}_1 \\ \vdots \\ \mathbf{s}_K \end{bmatrix} = \begin{bmatrix} \mathbf{\Psi}_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \mathbf{\Psi}_K \end{bmatrix} \begin{bmatrix} \mathbf{h}_1 \\ \vdots \\ \mathbf{h}_K \end{bmatrix}, \tag{18}$$

$$\mathbf{\Sigma}_s^{(i)} \triangleq E[\mathbf{s}\mathbf{s}^{\dagger}|\mathbf{y},\mathbf{b}^{(i)}] = E[\mathbf{\Psi}\mathbf{h}\mathbf{h}^{\dagger}\mathbf{\Psi}^{\dagger}|\mathbf{y},\mathbf{b}^{(i)}] = \mathbf{\Psi}\mathbf{\Sigma}_{\mathbf{h}}^{(i)}\mathbf{\Psi}^{\dagger}. \tag{19}$$

Therefore,

$$E\{\mathbf{h}_k^{\dagger}\mathbf{F}^{\dagger}\mathbf{C}_k^T\mathbf{C}_j\mathbf{F}\mathbf{h}_j|\mathbf{y},\mathbf{b}^{(i)}\} = E[\mathbf{s}^{\dagger}\mathbf{s}|\mathbf{y},\mathbf{b}^{(i)}] = \mathrm{tr}\left[\mathbf{\Sigma}_s^{(i)}[k,j] + \boldsymbol{\mu}_s^{(i)}[k]\boldsymbol{\mu}_s^{(i)}[j]^{\dagger}\right] \tag{20}$$

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

B. Implementation of the Maximization-Step (M-Step)

The M-Step can be performed using eq.(6) and the final result is as follows

$$b_k^{(i+1)}(l) = \text{sgn}\left[\Re\left\{\left[(\mathbf{h}_k^\dagger)^{[i]}\mathbf{F}^\dagger\mathbf{C}_k^T\mathbf{y}(m) - \sum_{j=1}^{K} b_j^{[i]}(m)(\mathbf{h}_k\mathbf{F}^\dagger\mathbf{C}_k^T\mathbf{C}_j\mathbf{F}\mathbf{h}_j)^{[i]}\right]\right\}\right] \tag{21}$$

where, the quantities $(\mathbf{h}_k^\dagger)^{[i]}$ and $(\mathbf{h}_k\mathbf{F}^\dagger\mathbf{C}_k^T\mathbf{C}_j\mathbf{F}\mathbf{h}_j)^{[i]}$ in (21) are given by (17)and (20), respectively. Note that for large observations frame M, tr $\left[\boldsymbol{\Sigma}_s^{(i)}[k,j]\right] \approx 0$ compared to tr $\left[\boldsymbol{\mu}_s^{[i]}[k]\boldsymbol{\mu}_s^{[i]}[j]^\dagger\right]$, [8]. That is

$$\begin{aligned}\lim_{m\to\infty}(\mathbf{h}_k\mathbf{F}^\dagger\mathbf{C}_k^T\mathbf{C}_j\mathbf{h}_j)^{[i]} &\approx \text{tr}\left[\boldsymbol{\mu}_s^{[i]}[k]\boldsymbol{\mu}_s^{[i]}[j]^\dagger\right] \\ &= \boldsymbol{\mu}_s^{[i]}[k]^\dagger\boldsymbol{\mu}_s^{[i]}[j] \equiv (\mathbf{h}_k^\dagger)^{[i]}\mathbf{h}_k^{[i]}\end{aligned} \tag{22}$$

Note that the identity on the right hand side of (22) follows from the facts that $\boldsymbol{\mu}_s^{(i)} \triangleq \boldsymbol{\Psi}\boldsymbol{\mu}_h^{(i)}$ and $\boldsymbol{\Psi}_k^\dagger\boldsymbol{\Psi}_k = \frac{1}{P}\mathbf{I}$. Discarding the second term, we finally have

$$b_k^{i+1}(m) = \text{sgn}\left[\Re\left\{\boldsymbol{\mu}_h^{(i)}[j]\boldsymbol{\Psi}_k^T\left[\mathbf{y}(m) - \sum_{j=1,j\neq k}^{K} b_j^{[i]}(m)\boldsymbol{\Psi}_j\boldsymbol{\mu}_h^{(i)}[j]\right]\right\}\right] \tag{23}$$

As a conclusion, Equation (23) can be interpreted as joint channel estimation and 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 have been decomposed into K independent optimization problems which can be resolved in a computationally feasible way. Finally we remark that this paper is an extension of the work presented by [7] to the problem of joint channel estimation and data detection for the uplink multicarrier CDMA systems operating in the presence of the frequency selective channels. Fleury and Kocian [7] investigates the same problem for the uplink Direct Sequence CDMA systems in the presence of flat fading channels.

IV. Simulations

In this section, the performance of uplink MC-CDMA systems based on the proposed receiver is analyzed for frequency selective channels. In computer simulations, we assume that all users are received with the same power level. Orthogonal Walsh sequences selected as a spreading code and the processing gain is chosen equal to the number of subcarriers $P = 16$. The number of active users are selected as $K = 16$ and each user sends its data frame composed of T preamble bits, and F data bits, over mobile fading channel. Wireless channels between mobiles antennas and receiver antenna are assumed to be complex Gaussian channel of length L with $N(0, \boldsymbol{\Sigma}_h)$.

In the receiver, the initial MMSE channel estimate is obtained by using T preamble bits while channel covariance matrix $\boldsymbol{\Sigma}_h$ is assumed to be known. Initial MMSE estimate of F data bits is computed from the observation vector $\mathbf{y}$ assuming that the channel coefficients were estimated by means of the pilot symbols. We refer to this method for obtaining $\mathbf{h}$ and $\mathbf{b}$ as the MMSE Separate Detection and Estimation (MMSE-SDE). In simulations, if the output of the the MMSE-SDE is applied to the parallel interference cancelation (PIC) receiver which is compared with the SAGE-JDE, it is referred to the Combined MMSE-PIC Receiver. Moreover, we also simulated both MMSE-SDE and Combined MMSE-PIC detectors referred as CSI-MMSE and CSI-Combined MMSE-PIC, respectively for the perfect channel state information cases.

Fig.1A compares the BER performance of the proposed SAGE-JDE approach with MMSE-SDE, Combined-MMSE-PIC, CSI-MMSE and CSI-Combined MMSE-PIC. For fair comparison we simulated Combined-MMSE-PIC, CSI-Combined MMSE-PIC and the EM-JDE receivers for three iterations. It is observed that the proposed EM-JDE outperforms MMSE-SDE, Combined-MMSE-PIC as well as CSI-MMSE and approaches the CSI-Combined MMSE-PIC.

To investigate initialization effect on the SAGE-JDE and Combined-MMSE-PIC, we also studied the BER performance of these algorithms for different preamble lengths presented in Fig.1B It is concluded that low preamble

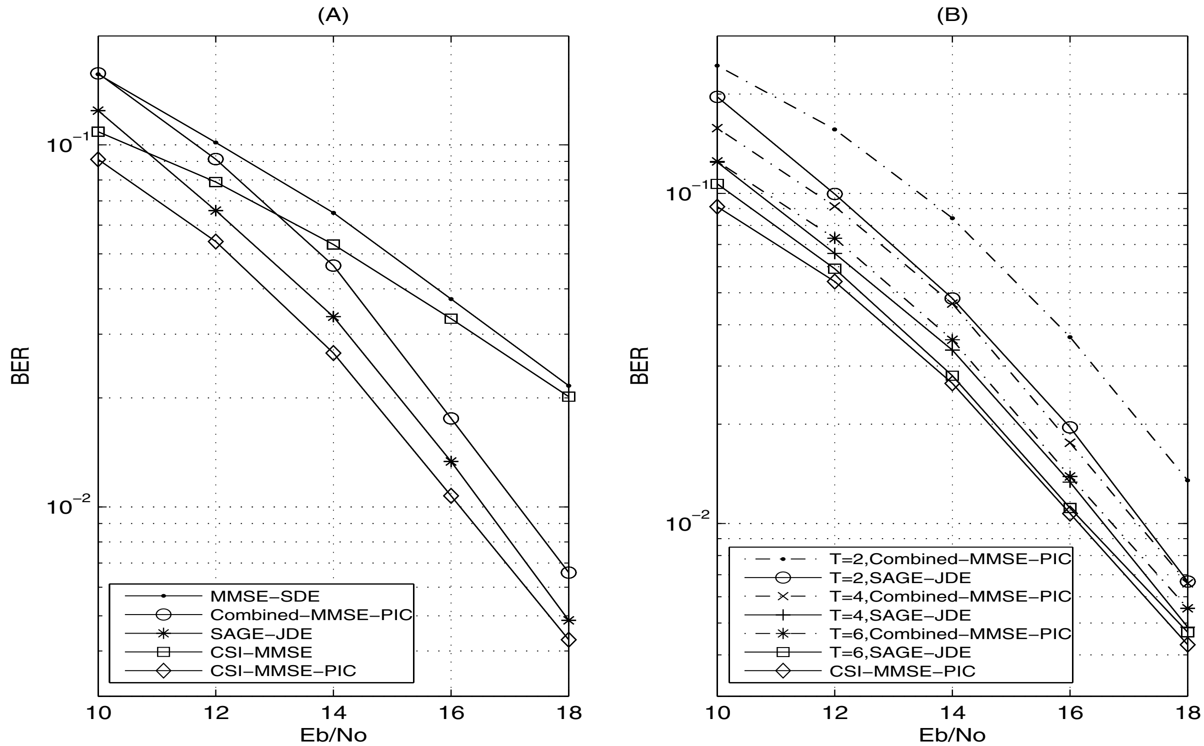

Fig. 1. (A) Behavior of the BER performance as a function of used preamble bits (F=20,T=4,L=5)(B)Behavior of the BER performance as a function of used preamble bits (F=20,L=5)

lengths provide poor initial estimates, resulting in a more BER performance degradation in the Combined-MMSE-PIC as compared to the SAGE-JDE system. On the other hand for higher preamble lengths it was observed that performance difference between Combined-MMSE-PIC and EM-JDE decreases because of considerably good initial estimates of channel coefficients.

V. Conclusions

The problem of joint data detection and channel estimation for uplink MC-CDMA systems operating in the presence of frequency selective fading channels was investigated in this work. We presented an iterative approach based on a version of the SAGE algorithm suitable for superimposed signals. 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 initiated the SAGE algorithm very effectively. Computer simulations were presented to demonstrate the effectiveness of the proposed algorithm in terms of BER performances.

References

[1] M. Feder and E. Weinstein, "Parameter Estimation of superimposed signals using the EM algorithm,"*IEEE Tran. on Acoustic, Speech and Signal Processing*, Vol. 36, pp. 477-489, April 1988.

[2] H.A Cirpan , E. Panayirci , H. Dogan, "Nondata-aided Channel Estimation for OFDM Systems with Space-Frequency Transmit Diversity"*IEEE Transactions on Vehicular Technology*, March 2006, vol. 55, Issue 2, pp. 449 - 457

[3] H. Dogan, H. A. Cirpan and Erdal Panayirci, "Iterative Channel Estimation and Decoding of Turbo Coded SFBC-OFDM Systems," *IEEE Trans. Wireless Commun.*, accepted for publication in Oct. 2006.

[4] X. Yongzhe and C.N. Georghiades, "Two EM-type channel estimation algorithms for OFDM with transmitter diversity," *IEEE Trans. on Commun.*, vol. 51, No. 1, pp. 106-115 January 2003.

[5] V. Kuhn, "Combined MMSE-PIC in coded OFDM-CDMA systems," *IEEE GLOBECOM 01*, pp. 231-235, Nov. 2001.

[6] S. Iraji, T. Sipila and J. Lilleberg, "Channel Estimation and Signal Detection for MC-CDMA in Multipath Fading Channels," *in Proc. IEEE Int. Symp. on Personal, Indoor and Mobile Radio Commun. (PIMRC)*, September 1993

[7] 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.

[8] J.A. Fessler, A.O. Hero, Space-alternating generalized expectation-maximization, Algorithm, *IEEE Trans. Signal Process*. 42 (10) pp. 2664-2677, October 1994