A Monte Carlo Implementation of the SAGE Algorithm for Joint Soft-Multiuser Decoding, Channel Parameter Estimation, and Code Acquisition

Alexander Kocian, *Member, IEEE*, Erdal Panayırci, *Fellow, IEEE*, H. Vincent Poor, *Fellow, IEEE*, and Marina Ruggieri, *Senior Member, IEEE*

***Abstract*—This paper presents an iterative scheme for joint timing acquisition, multi-channel parameter estimation, and multiuser soft-data decoding. As an example, an asynchronous convolutionally coded direct-sequence code-division multiple-access system is considered. The proposed receiver is derived within the space-alternating generalized expectation-maximization framework, implying that convergence in likelihood is guaranteed under appropriate conditions in contrast to many other iterative receiver architectures. The proposed receiver iterates between joint posterior data estimation, interference cancellation, and single-user channel estimation and timing acquisition. A Markov Chain Monte Carlo technique, namely Gibbs sampling, is employed to compute the *a posteriori* probabilities of data symbols in a computationally efficient way. Computer simulations in flat Rayleigh fading show that the proposed algorithm is able to handle high system loads unlike many other iterative receivers.**

***Index Terms*—Expectation maximization algorithms, multiaccess communication, Monte Carlo mthods.**

I. Introduction

IN practical wireless communication systems, signals are affected by physical phenomena such as time-varying channels, frequency-selective fading, multiple access interference, non-Gaussian noise, and loss of synchronization. In a parametric model, the received signal is represented as a function of unknown complex channel coefficients and transmission delays which, if perfectly known at the receiver, would improve the quality of data symbol estimation by the receiver.

For situations in which the channel parameters are known to the receiver, Giallorenzi and Wilson derived in [1] the maximum-likelihood (ML) data sequence decoder for a convolutionally coded multiuser system. Its computational complexity, however, grows exponentially with the number of users and constraint length, so that a real-time implementation in a digital signal processor is impossible even for small numbers of users and small constraint lengths. For code-division multiple-access (CDMA) systems, Alexander *et al.* [2] view the concatenation of direct-sequence spreading with the asynchronous multiple-access channel as a special form of a convolutional code. Following the "turbo principle" [3], the resulting inner code (due to the spreading sequence) and the outer code (due to the channel code) can be decoded in an iterative fashion by exchanging *a posteriori* probability (APP) information about the data symbols. The overall bit-error performance of the receiver has been shown to be near-optimum for heavily loaded systems i.e., with loads larger than one, at a computational complexity that makes its implementation feasible. In Wang and Poor [4], a soft-input soft-output (SISO) multiuser estimator and a bank of single-user (SU) channel decoders exchange extrinsic (EXT) information about the code symbols. For direct-sequence (DS) CDMA with random spreading, the latter choice is optimum in the large-system limit [5]. For systems with finite numbers of users and finite interleaver sizes there are cases in which the SU channel needs APP values about the code symbols [6].

In turn, the ML or the maximum *a posteriori* probability (MAP) delay estimator turns out to be a delay-locked loop (DLL), provided the data symbols, carrier frequency, and phase are known [7].

When the channel, symbol timing, and data are unknown, an often used approach is to iterate among timing acquisition, channel estimation, interference cancellation and single user decoding [8]–[10]. Convergence of the overall receiver cannot be guaranteed as the parameters are estimated in a heuristic fashion. An optimal receiver that jointly estimates the nuisance parameters and the data symbols of all users at polynomial computational complexity is sought.

The expectation maximization (EM) and space alternating generalized EM (SAGE) algorithms are iterative methods that approximate the maximum likelihood solution. Under certain mild conditions, convergence in likelihood is guaranteed [11], [12]. This approach was first applied to multiuser detection in [13], and later generalized in other works (e.g., [14]). For synchronized reception, the EM and SAGE receivers in [15] iterate among multichannel estimation, interference cancellation, and (hard) data decoding. Given the channel coefficients and treating the unknown transmission delays as

Manuscript received March 18, 2010; accepted July 05, 2010. Date of publication July 29, 2010; date of current version October 13, 2010. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Huaiyu Dai. This work was supported in part by the U.S. National Science Foundation under Grant CNS-09-05398. The material in the paper was presented in part at the IEEE International Workshop on Signal Processing Advances in Wireless Communications (SPAWC), Perugia, Italy, June 2009.

A. Kocian and M. Ruggieri are with the Center for TeleInFrastructure (CTiF)-Italy, University of Rome "Tor Vergata," Rome, Italy (e-mail: Alexander.Kocian@uniroma2.it; ruggieri@uniroma2.it).

E. Panayirci was with the Princeton University, Princeton, NJ 08544 USA. He is now with the Kadir Has University, Cibali, 34083, Istanbul, Turkey (e-mail: eepanay@khas.edu.tr).

H. V. Poor is with the Department of Electrical Engineering, Princeton University, Princeton, NJ 08544 USA (e-mail: poor@princeton.edu).

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/TSP.2010.2062181

nuisance parameters, the EM receiver in [16] iterates between synchronization and maximum-likelihood sequence detection (MLSD) such as the Viterbi algorithm [17]. For the estimation problem at hand, the delay estimation problem is nonlinear. The delay-estimation problem becomes multi-linear when a multi-resolution expansion is applied to the received signal instead [18]. In contrast, when the roles of estimation parameter and nuisance parameters are interchanged, the EM receiver in [19] iterates between single-user *a posteriori* data decoding, such as the Bahl–Cocke–Jelinek–Raviv (BCJR) algorithm [20], ML channel coefficient estimation and timing acquisition. Straightforward extension of [18] to multiuser transmission, however, results in an algorithm consuming a nonpolynomial number of operations in the number of users. To overcome this problem, Iltis *et al.* consider in [21] the product of (uncoded) data and channel coefficients as estimation parameters and the users' transmission delays as nuisance parameters. Monte Carlo simulations in Rayleigh fading for slightly loaded systems, i.e., with loads less than one, indicate that the proposed EM receiver is robust to channel estimation errors.

This paper presents a novel algorithm for joint multiuser soft-decoding, multichannel estimation, and synchronization based on the SAGE framework at polynomial computational complexity. To meet this challenge, we adopt Bayesian Monte Carlo methodologies in the SAGE framework. Here, an efficient Markov chain Monte Carlo (MCMC) technique [22] called *Gibbs sampling* [23] is used to compute the expected log-APPs of all the users' data symbols exactly in an adaptive fashion. Direct computation of these APPs involves a nonpolynomial number of floating point operations (FLOPs). In contrast to other MCMC techniques, such as the Metropolis–Hastings algorithm, the Gibbs sampler does not require calibration, its convergence rate is higher in the low SNR-regime, and the acceptance probability is always one. In the limiting regime, these Monte Carlo EM/SAGE algorithms and their standard counterparts have the same convergence properties [24].

The rest of the paper is organized as follows. Section II presents the system model. As an example we consider DS-CDMA transmission. Section III provides background information on the SAGE and the Monte Carlo SAGE algorithms and the derivation of the latter for joint estimation of coded data, channel coefficients, and transmission delays. Implementation issues such as initialization are discussed in Section III-E. The performance of the proposed scheme is analyzed in Section IV. We investigate the convergence rate, derive a modified Cramér–Rao bound for the estimated parameters, and give numerical examples.

Notation: In the following, the superscripts $(\cdot)^\star$ and $(\cdot)^\dagger$ denote conjugate and conjugate transpose, respectively, of the argument; $\Re(\cdot)$ and $\Im(\cdot)$ denote the real and imaginary parts of a complex argument, respectively. Column vectors and matrices are represented by boldface lowercase and uppercase letters, respectively; $\mathrm{col}\{\cdot\}$ denotes a column vector with the elements in the argument as its entries. The symbol $\mathbf{0}_{m\times n}$ denotes the $m \times n$-dimensional all-zero matrix; whereas $\boldsymbol{I}_n$ represents the identity matrix of size $n \times n$. Finally, $\delta(\cdot)$ is an indicator function that takes the value 1 if its argument is true and the value 0 otherwise.

II. System Description

We consider an asynchronous convolutionally coded DS-CDMA system with K active users. The information bit sequence $\{b_k[u]\} \in \{0,1\}^{L_b}$ of user k is encoded with a user-independent encoder of rate R, mapped into the data sequence $\{d_k[n]\} \in \{-1,+1\}^{L_d}$, $L_d = L_b/R$, $L_b/R \bmod L_d = 0$, fed into the user specific symbol interleaver $\mathbf{\Pi}_k$ with interleaving depth L_Π, and multiplexed with L_p random preamble symbols. Hence, the block length equals $L = L_p + L_d$ symbols. Each interleaved symbol $d_k[\ell]$ is then modulated with a time-varying random signature waveform of duration T_d, such that each symbol consists of N_c chips with duration $T_c = T_d/N_c$ where N_c is an integer, and transmitted over a quasi-static flat block fading channel. Notice that the average number of information bits per code symbol is $L_b/L = R(L - L_p)/L$. The received signal is the noisy sum of all users' contributions, delayed by the propagation delays $\tau_k \in]0, T_d[$, where the subscript k denotes the label of the kth user. After down-converting the received signal to baseband and passing it through an integrate-and-dump filter with integration time $T_s = T_c/Q$, $Q \in \mathbb{Q}^+$, $QN_c(L+1)$ samples over an observation frame of L symbols are stacked into a signal column-vector $\boldsymbol{r}$. Assuming that sampling is chip-synchronous without knowledge of the individual transmission delays, the vector $\boldsymbol{r}$ can be expressed as

$$\boldsymbol{r} = \boldsymbol{S}(\boldsymbol{\tau})\mathbf{A}\mathbf{d} + \boldsymbol{w}. \tag{1}$$

In this expression the matrix $\boldsymbol{S}(\boldsymbol{\tau}) \in \mathbb{C}^{QN_c(L+1)\times LK}$ contains the signature sequences of all the users:

$$\boldsymbol{S}(\boldsymbol{\tau}) = [\boldsymbol{S}_1(\tau_1),\ \boldsymbol{S}_2(\tau_2), \ldots,\ \boldsymbol{S}_K(\tau_K)]$$

where $\boldsymbol{S}_k(\tau_k) \in \mathbb{C}^{QN_c(L+1)\times L}$ has the form

$$\boldsymbol{S}_k(\tau_k) = \begin{bmatrix} | & | & & | \\ \boldsymbol{S}_k(1,\tau_k) & \boldsymbol{S}_k(2,\tau_k) & \cdots & \boldsymbol{S}_k(L,\tau_k) \\ | & | & & | \end{bmatrix}$$

and the spreading code vector $\boldsymbol{S}_k(\ell,\tau_k) \in \mathbb{C}^{QN_c(L+1)\times 1}$ is given by

$$\boldsymbol{S}_k(\ell,\tau_k) = \begin{bmatrix} \mathbf{0}_{QN_c(\ell-1)+\tau_k\times 1} \\ | \\ \boldsymbol{s}_k(\ell,\tau_k) \\ | \\ \mathbf{0}_{QN_c(L-\ell+1)-\tau_k\times 1} \end{bmatrix}.$$

The vector $\boldsymbol{s}_k(\ell,\tau_k) \in \mathbb{C}^{QN_c}$ contains the sampled spreading code of user k in signaling interval ℓ with energy $\mathcal{E}_k = \boldsymbol{s}_k^\dagger(\ell,\tau_k)\boldsymbol{s}_k(\ell,\tau_k) = 1$. The diagonal channel matrix $\boldsymbol{A} \in \mathbb{C}^{LK\times LK}$ in (1) is given by $\boldsymbol{A} = \mathrm{diag}\{\boldsymbol{A}_1, \ldots, \boldsymbol{A}_K\}$ with $\boldsymbol{A}_k = a_k\mathbf{I}_L$. The kth user's channel coefficient a_k is assumed to be a circularly symmetric complex Gaussian random variable with zero mean and variance σ_k^2, and its transmission delay is assumed to be uniformly distributed. The code symbol vector $\boldsymbol{d} \in \mathbb{C}^{LK}$ takes the form $\boldsymbol{d} = \mathrm{col}\{\boldsymbol{d}_1, \ldots, \boldsymbol{d}_K\}$, where the vector $\boldsymbol{d}_k \in \mathbb{C}^L$ contains the kth user's code symbols, i.e., $\boldsymbol{d}_k = \mathrm{col}\{d_k[1], \ldots, d_k[L]\}$ with $d_k[\ell] \in \{-1,+1\}$ denoting the symbol transmitted by the kth user during the ℓth signalling interval. Finally, the column vector $\boldsymbol{w} \in \mathbb{C}^{QN_c(L+1)}$ contains complex, circularly symmetric white Gaussian noise having covariance matrix $N_0\mathbf{I}_{QN_c(L+1)}$. We assume that the vectors $\boldsymbol{a} \triangleq \mathrm{col}\{a_1, a_2, \ldots, a_K\}$, $\boldsymbol{\tau} \triangleq \mathrm{col}\{\tau_1, \tau_2, \ldots, \tau_K\}$, $\boldsymbol{d}$ and $\boldsymbol{w}$ and their components are independent. The receiver does not know the data sequences, the (complex) channel coefficients, or the transmission delays.

III. Monte Carlo SAGE Joint Parameter Estimation

A. Motivation

The task is to obtain a model-based estimate of an (unobserved) parameter vector $\boldsymbol{\theta}$ given some observed vector $\boldsymbol{\mathcal{R}}$ with joint probability density function (pdf) $p(\boldsymbol{\theta}, \boldsymbol{\mathcal{R}})$. In this case, the MAP point estimate of $\boldsymbol{\theta}$ yields

$$\boldsymbol{\theta}_{\mathrm{MAP}} = \arg\max_{\boldsymbol{\theta}} p(\boldsymbol{\theta}, \boldsymbol{\mathcal{R}}).$$

When the model includes latent variables, it is often impossible to determine the MAP-estimate in closed form. We therefore resort to suboptimal methods.

B. The SAGE Algorithm

In previous applications, the SAGE algorithm has been extensively used to iteratively approximate the maximum-likelihood (cf. [25]) and MAP (cf. [26]) estimates of a *continuous-valued* parameter vector $\boldsymbol{\theta}$ with respect to observations $\boldsymbol{\mathcal{R}}$. In the SAGE algorithm, only the parameter components in a subset $\boldsymbol{\theta}_{\mathrm{S}}$ of $\boldsymbol{\theta}$, indexed by S, are updated. The remaining parameter components $\boldsymbol{\theta}_{\bar{\mathrm{S}}}$ of $\boldsymbol{\theta}$, indexed by the complement $\bar{\mathrm{S}}$ of S, are not re-estimated. The SAGE algorithm postulates the existence of some hypothetical *discrete-valued* data $\boldsymbol{\mathcal{D}}$ that would aid in the estimation of $\boldsymbol{\theta}$ but is not part of $\boldsymbol{\mathcal{R}}$. The data $\boldsymbol{\chi}_{\mathrm{S}} = \{\boldsymbol{\mathcal{R}}, \boldsymbol{\mathcal{D}}\}$ is said to be *admissible hidden* with respect to $\boldsymbol{\theta}_{\mathrm{S}}$ [25]. Given the current ith estimate $\boldsymbol{\theta}$, the so-called expectation (E)-step computes the Q-function

$$Q_{\mathrm{S}}(\boldsymbol{\theta}_{\mathrm{S}}; \boldsymbol{\theta}^{[i]}) \triangleq \sum_{\boldsymbol{\mathcal{D}}} \log p\left(\boldsymbol{\mathcal{R}}, \boldsymbol{\mathcal{D}} \mid \boldsymbol{\theta}_{\mathrm{S}}, \boldsymbol{\theta}_{\bar{\mathrm{S}}}^{[i]}\right) p\left(\boldsymbol{\mathcal{D}} \mid \boldsymbol{\mathcal{R}}, \theta^{[i]}\right). \tag{2}$$

Then, the maximization (M)-step seeks to find the $(i+1)$st estimate

$$\boldsymbol{\theta}_{\mathrm{S}}^{[i+1]} = \arg\max_{\boldsymbol{\theta}_{\mathrm{S}}} \left\{ Q_{\mathrm{S}}(\boldsymbol{\theta}_{\mathrm{S}}; \boldsymbol{\theta}^{[i]}) + \log p(\boldsymbol{\theta}_{\mathrm{S}}) \right\}.$$

The objective function is nondecreasing at each iteration. Notice that the computational complexity of the E-step in (2) is still NP-hard due to the discrete nature of $\boldsymbol{\mathcal{D}}$ unless the problem has special structure that alleviates this complexity.

C. The Monte Carlo SAGE Algorithm

To make the computation of the expectation in (2) feasible, we propose to use the MCMC technique to obtain the so-called Monte Carlo SAGE algorithm (cf. [24]). MCMC is a statistical technique that allows generation of a large number M of ergodic pseudo-random samples $\left\{\boldsymbol{\mathcal{D}}^{[i,m]}\right\}_{m=1}^{M}$ from the current approximation to the conditional pdf $p(\boldsymbol{\mathcal{D}} \mid \boldsymbol{\mathcal{R}}, \boldsymbol{\theta}^{[i]})$ at SAGE iteration i. These samples are used to approximate the expectation in (2) by a sample-mean.

Widely used MCMC algorithms are the Metropolis-Hastings algorithm [27], [28] and Gibbs sampling [29]. Here, we consider only the latter, mainly because of its faster convergence rate [29]. Suppose that the dimension of $\boldsymbol{\mathcal{D}}$ is $\dim\{\boldsymbol{\mathcal{D}}\} = KM$. Having initialized $\boldsymbol{\mathcal{D}}^{[0,0]}$ randomly, the Gibbs sampler iterates the following loop:

for $m = -M_b + 1, \ldots, M$
  for $k = 1, 2, \ldots, K$
    Draw sample $\mathcal{D}_k^{[i,m]}$ from
    $p(\mathcal{D}_k | \mathcal{D}_1^{[i,m]}, \ldots, \mathcal{D}_{k-1}^{[i,m]}, \mathcal{D}_{k+1}^{[i,m-1]}, \ldots, \mathcal{D}_K^{[i,m-1]}, \boldsymbol{\mathcal{R}}, \boldsymbol{\theta}^{[i]})$;
    Compute $\delta\left(\mathcal{D}_k^{[i,m]} = +1 \mid \boldsymbol{\mathcal{R}}, \boldsymbol{\theta}^{[i]}\right)$;
  end
end

The first "burn-in" samples M_b are not taken into account. From the strong law of large numbers, it follows under mild conditions on the kth entry of $\boldsymbol{\mathcal{D}}$ that [30]

$$\lim_{M\to\infty} \frac{1}{M} \sum_{m=1}^{M} \delta\left(\mathcal{D}_k^{[i,m]} = 1 \mid \boldsymbol{\mathcal{R}}, \boldsymbol{\theta}^{[i]}\right) \to p(\mathcal{D}_k = 1 \mid \boldsymbol{\mathcal{R}}, \boldsymbol{\theta}^{[i]})$$

with probability one. Following this approach, the Monte Carlo E-step yields

$$Q_{\mathrm{S}}(\boldsymbol{\theta}_{\mathrm{S}}; \boldsymbol{\theta}^{[i]}) \approx \frac{1}{M} \sum_{m=1}^{M} \log p\left(\boldsymbol{\mathcal{R}}, \boldsymbol{\mathcal{D}}^{[i,m]} \mid \boldsymbol{\theta}_{\mathrm{S}}, \boldsymbol{\theta}_{\bar{\mathrm{S}}}^{[i]}\right).$$

D. Receiver Design

We now return to the specific model of (1). To obtain a SAGE-based receiver architecture that iterates between soft-data and channel estimation, one might choose the parameter vector to be $\boldsymbol{\theta} = \{\Re(a_1), \ldots, \Re(a_K), \Im(a_1), \ldots, \Im(a_K),\ \tau_1, \ldots, \tau_K\}$. At iteration i, the parameters for user $k = k[i] = i \bmod K + 1$ are updated. For the observation $\boldsymbol{\mathcal{R}} = \boldsymbol{r}$, it can easily be shown that $\chi_k = \{\boldsymbol{r},\ \boldsymbol{d}\}$ is admissible hidden for $\{a_k,\ \tau_k\}$.

We start with the log-likelihood function

$$\log p(\boldsymbol{r}, \boldsymbol{d} \mid \boldsymbol{a}, \boldsymbol{\tau}) = \log p(\boldsymbol{r} \mid \boldsymbol{d}, \boldsymbol{a}, \boldsymbol{\tau}) + \log p(\boldsymbol{d}). \tag{3}$$

From (1), it follows that

$$\log p(\boldsymbol{r} \mid \boldsymbol{d}, \boldsymbol{a}, \boldsymbol{\tau}) \propto \Re\{\boldsymbol{r}^{\dagger} \boldsymbol{S}(\boldsymbol{\tau}) \boldsymbol{A} \boldsymbol{d}\} - \frac{1}{2} \boldsymbol{\mu}(\boldsymbol{\theta}, \boldsymbol{d})^{\dagger} \boldsymbol{\mu}(\boldsymbol{\theta}, \boldsymbol{d}) \tag{4}$$

where $\boldsymbol{\mu}(\boldsymbol{\theta}, \boldsymbol{d}) \triangleq \sum_{k=1}^{K} \sum_{\ell=1}^{L} \boldsymbol{S}_k(\ell, \tau_k) a_k d_k[\ell]$.

1) The Q-Function: Substituting (4) into (3) yields after some algebraic manipulations for the E-step in (2):

$$Q_k(\boldsymbol{\theta}_k; \boldsymbol{\theta}^{[i]}) = \frac{2}{N_0} \sum_{\ell=1}^{L} \Re\left\{a_k^{\star} \lambda(\ell, \tau_k)\right\} - \frac{L}{N_0} |a_k|^2$$

with the branch definition

$$\lambda(\ell, \tau_k) \triangleq \boldsymbol{S}_k^{\dagger}(\ell, \tau_k) \left(\tilde{d}_k^{[i]}[\ell]\, \boldsymbol{r} - \sum_{k' \neq k} \mathcal{I}_{k'}^{[i]}(\ell, \tau_{k'}) \right); \tag{5}$$

the interference term

$$\mathcal{I}_{k'}^{[i]}(\ell,\tau_{k'}) \triangleq a_{k'}^{[i]} \sum_{\ell'=\ell-1}^{\ell+1} \left(\boldsymbol{S}_{k'}(\ell',\tau_{k'}^{[i]}) \left(d_k[\widetilde{\ell]d_{k'}}[\ell'] \right)^{[i]} \right)$$

the soft-data symbols

$$\tilde{d}_k^{[i]}[\ell] \triangleq \sum_{\mu\in\{-1,+1\}} \mu P\left(d_k[\ell]=\mu \mid \boldsymbol{r},\boldsymbol{\tau}^{[i]},\boldsymbol{a}^{[i]}\right) \quad (6)$$

and the soft-product of two data symbols

$$\left(d_k[\widetilde{\ell]d_{k'}}[\ell']\right)^{[i]} \triangleq \sum_{\mu}\sum_{\nu} \mu\nu P\left(d_k[\ell]=\mu, d_{k'}[\ell']=\nu \mid \boldsymbol{r},\boldsymbol{\tau}^{[i]},\boldsymbol{a}^{[i]}\right), \quad (7)$$

$k' \neq k$, where the lag ℓ' is within the range $\ell' \in \{\ell-1,\ell,\ell+1\}$.

2) Monte-Carlo Implementation of the Symbol a posteriori Probabilities: For simplicity of notation, let $\boldsymbol{d}_{\bar{k}}$ denote the vector containing the code symbols of all users but those of user k i.e., $\boldsymbol{d}_{\bar{k}} \triangleq \boldsymbol{d}\backslash\boldsymbol{d}_k$. The APPs of $d_k[\ell]$ in (6) can be evaluated as

$$\begin{aligned} P&\left(d_k[\ell]=\mu \mid \boldsymbol{r},\boldsymbol{\tau}^{[i]},\boldsymbol{a}^{[i]}\right) \\ &= \sum_{\boldsymbol{d}_{\bar{k}}} P\left(d_k[\ell]=\mu \mid \boldsymbol{d}_{\bar{k}},\boldsymbol{r},\boldsymbol{\tau}^{[i]},\boldsymbol{a}^{[i]}\right)\, P\left(\boldsymbol{d}_{\bar{k}} \mid \boldsymbol{r},\boldsymbol{\tau}^{[i]},\boldsymbol{a}^{[i]}\right) \\ &\approx \frac{1}{M}\sum_{m=1}^{M} \delta\left(d_k[\ell]=\mu \mid \boldsymbol{d}_{\bar{k}}^{[i,m]},\boldsymbol{r},\boldsymbol{\tau}^{[i]},\boldsymbol{a}^{[i]}\right), \end{aligned} \quad (8)$$

$\ell = 1,\ldots,L$. Incorporating the code's trellis, the conditional probability $P(d_k[\ell]=\mu \mid \boldsymbol{d}_{\bar{k}}^{[i,m]},\boldsymbol{r},\boldsymbol{\tau}^{[i]},\boldsymbol{a}^{[i]})$ may be written as

$$P\left(d_k[\ell]=\mu \mid \boldsymbol{d}_{\bar{k}}^{[i,m]},\boldsymbol{r},\boldsymbol{\tau}^{[i]},\boldsymbol{a}^{[i]}\right) = C\sum_{\mathcal{S}_\ell^\mu} \alpha_{s'}[u-1]\gamma_{s',s}[u]\beta_s[u]. \quad (9)$$

The normalization constant C is resolved from the boundary condition $\sum_{d_k[\ell]\in\{-1,+1\}} P\left(d_k[\ell] \mid \boldsymbol{d}_{\bar{k}}^{[i,m]},\boldsymbol{r},\boldsymbol{\tau}^{[i]},\boldsymbol{a}^{[i]}\right) = 1$. Suppose, the kth user's symbol interleaver implements the permutation function $\boldsymbol{\Pi}_k : \mathcal{P}_k \rightarrow \mathcal{P}_k$, with $\mathcal{P}_k = \{1,2,\ldots,L_\Pi\}$. Conversely, the kth user's symbol de-interleaver implements the reverse function $\boldsymbol{\Pi}_k^{-1}$, with $\boldsymbol{\Pi}_k^{-1}(\boldsymbol{\Pi}_k(n)) = n, \forall n \in L_\Pi$. Then, from (4), the logarithm of the transition probability $\gamma_{s',s}[u]$ in (9) from state s' at time instant $u-1$ to state s at time instant u is given by

$$\ln\gamma_{s',s}[u] \propto \sum_{n=R^{-1}u}^{R^{-1}(u+1)-1} \frac{2}{N_0} d_k[\boldsymbol{\Pi}_k(n)] \times\Re\left\{ (\boldsymbol{g}_k[\boldsymbol{\Pi}_k(n)]^{[i]})^\dagger \boldsymbol{r} - \sum_{j\neq k} \mathcal{J}_{k,j}^{[m]}[\boldsymbol{\Pi}_j(n)] \right\}$$

with the interference term

$$\mathcal{J}_{k,j}^{[m]}[n] \triangleq (\boldsymbol{g}_k[n]^{[i]})^\dagger \sum_{\Delta n=-1}^{+1} \boldsymbol{g}_j[n+\Delta n]^{[i]} d_j[n+\Delta n]^{[i,m]} \quad (10)$$

at sample index m. For notational simplicity, we have defined $\boldsymbol{g}_k[n]^{[i]} \triangleq a_k^{[i]}\boldsymbol{S}_k^\dagger(n,\tau_k^{[i]})$. The "forward" and "backward" probabilities in (9) are given by $\alpha_{s'}[u] = \sum_{s'}\alpha_{s'}[u-1]\gamma_{s',s}[u]$ and $\beta_{s'}[u-1] = \sum_s \beta_s[u]\gamma_{s',s}[u]$, respectively, with boundary conditions [20]

$$\alpha_s[0] = \begin{cases} 1, & s=0 \\ 0, & s\neq 0 \end{cases}, \quad \beta_s[L_b-1] = \begin{cases} 1, & s=0 \\ 0, & s\neq 0. \end{cases}$$

Moreover, $\mathcal{S}_\ell^\mu$ in (9) is the set of state pairs (s',s) in the trellis such that $d_k[\ell]=\mu$. Notice that the algorithm in (9) computes APP values for the *coded* data symbols while the BCJR algorithm in [20] provides APP values for the *uncoded* information bits.

Having computed the *sequence* of probabilities $P(d_k[\ell] = \pm\mu \mid \boldsymbol{d}_{\bar{k}}^{[i,m]},\boldsymbol{r},\boldsymbol{\tau}^{[i]},\boldsymbol{a}^{[i]})$, $\ell=1,\ldots,L$, the Gibbs sampler draws the vector $\boldsymbol{d}_k^{[i,m]}$, computes the corresponding indicator functions in (8) at epoch m for the sequence $\boldsymbol{d}_k$, and sets $m \rightarrow m+1$.

3) Monte-Carlo Implementation of the Joint Symbol a posteriori Probabilities: Direct computation of the joint APPs in (7) for the interleaved symbols $d_k[\ell]$ and $d_{k'}[\ell']$ is infeasible, as the receiver does not have access to the state information at joint signaling interval $n = \boldsymbol{\Pi}_k^{-1}(\ell)$ and $n' = \boldsymbol{\Pi}_{k'}^{-1}(\ell')$.

To obtain a guess of the joint APPs though, (7) is first expanded as

$$P\left(d_k[\ell]=\mu, d_{k'}[\ell']=\mu' \mid \boldsymbol{r},\boldsymbol{\tau}^{[i]},\boldsymbol{a}^{[i]}\right) = \begin{cases} P\left(d_k[\ell]=\mu \mid d_{k'}[\ell']=\mu',\boldsymbol{r},\boldsymbol{\tau}^{[i]},\boldsymbol{a}^{[i]}\right) & \\ \quad\times P\left(d_{k'}[\ell']=\mu'|\boldsymbol{r},\boldsymbol{\tau}^{[i]},\boldsymbol{a}^{[i]}\right); & \tau_k^{[i]} \leq \tau_{k'}^{[i]} \\ P\left(d_{k'}[\ell']=\mu' \mid d_k[\ell]=\mu,\boldsymbol{r},\boldsymbol{\tau}^{[i]},\boldsymbol{a}^{[i]}\right) & \\ \quad\times P\left(d_k[\ell]=\mu|\boldsymbol{r},\boldsymbol{\tau}^{[i]},\boldsymbol{a}^{[i]}\right); & \tau_k^{[i]} > \tau_{k'}^{[i]}. \end{cases} \quad (11)$$

Then the Gibbs sampling theorem is applied to approximate the conditional APP in (11) given the user constraints only, plus the single-user APP given the code constraints. The latter we already have solved in (8). We will see that above method supports very high system loads.

For $\tau_k^{[i]} \leq \tau_{k'}^{[i]}$,

$$\begin{aligned} P&\left(d_k[\ell]=\mu \mid d_{k'}[\ell']=\mu',\boldsymbol{r},\boldsymbol{\tau}^{[i]},\boldsymbol{a}^{[i]}\right) \\ &\approx \sum_{\boldsymbol{d}_{\overline{k,k'}}\in\{-1,+1\}} P\left(d_k[\ell]=\mu \mid d_{k'}[\ell']=\mu',\boldsymbol{d}_{\overline{k,k'}},\boldsymbol{r},\boldsymbol{\tau}^{[i]},\boldsymbol{a}^{[i]}\right) \\ &\quad P\left(\boldsymbol{d}_{\overline{k,k'}} \mid \boldsymbol{r},\boldsymbol{\tau}^{[i]},\boldsymbol{a}^{[i]}\right) \\ &\approx \frac{1}{M}\sum_{m=1}^{M} \delta\left(d_k[\ell]=\mu \mid d_{k'}[\ell']=\mu',\boldsymbol{d}_{\overline{k,k'}}^{[i,m]},\boldsymbol{r},\boldsymbol{\tau}^{[i]},\boldsymbol{a}^{[i]}\right), \end{aligned} \quad (12)$$

$\ell=1,\ldots,L$, $\ell' \in \{\ell-1,\ell,\ell+1\}$, and $\boldsymbol{d}_{\overline{k,k'}} \triangleq \boldsymbol{d}\backslash\{\boldsymbol{d}_k,\boldsymbol{d}_{k'}\}$. Incorporating the user constraints only, the conditional symbol posterior $P(d_k[\ell]=\mu \mid d_{k'}[\ell']=\mu',\boldsymbol{d}_{\overline{k,k'}},\boldsymbol{r},\boldsymbol{\tau}^{[i]},\boldsymbol{a}^{[i]})$ in (12) can be factored as

$$P\left(d_k[\ell]=\mu \mid d_{k'}[\ell']=\mu',\boldsymbol{d}_{\overline{k,k'}}^{[i,m]},\boldsymbol{r},\boldsymbol{\tau}^{[i]},\boldsymbol{a}^{[i]}\right) = \check{C}\sum_{\mathcal{S}_{\ell,\ell'}^{\mu\mu'}} \check{\alpha}_{s'}[\ell-1]\check{\gamma}_{s',s}[\ell]\check{\beta}_s[\ell]. \quad (13)$$

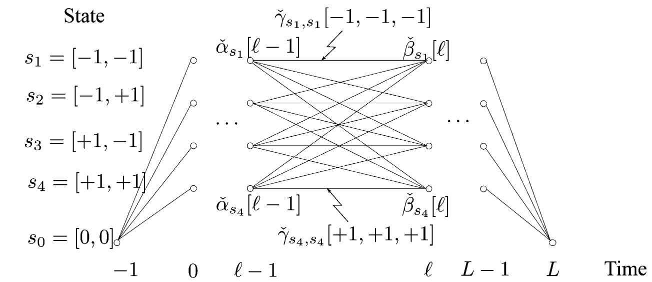

Fig. 1. Joint trellis for two uncoded users k and k' in the Monte Carlo SAGE scheme with $\tau_k \leq \tau_{k'} < T_d$.

The normalization constant $\check{C}$ results from the boundary condition $\sum_{d_k[\ell]\in\{-1,+1\}} P\left(d_k[\ell] \mid d_{k'}[\ell'], \boldsymbol{d}^{[i,m]}_{\overline{k,k'}}, \boldsymbol{r}, \boldsymbol{\tau}^{[i]}, \boldsymbol{a}^{[i]}\right) = 1$. The definition of the probabilities $\check{\alpha}$, $\check{\beta}$, and $\check{\gamma}$ is analogous to that of α, β, and γ in Section III-D-2). From (4) and the block-diagonal structure of $\boldsymbol{S}(\boldsymbol{\tau})^\dagger\boldsymbol{S}(\boldsymbol{\tau})$, the logarithm of the transition probability $\check{\gamma}_{s',s}[\ell]$ in (13) yields

$$\ln \check{\gamma}_{s',s}[\ell] \propto \frac{2}{N_0} d_k[\ell]\Re \times \left\{ (\boldsymbol{g}_k[\ell]^{[i]})^\dagger \boldsymbol{r} - \sum_{j\neq k,k'} \mathcal{J}^{[m]}_{k,j}[\ell] - \mathcal{J}_{k,k'}[\ell] \right\}.$$

Note that the interference term $\mathcal{J}^{[m]}_{k,j}[\ell]$, defined in (10), depends on the sample index m, while $\mathcal{J}_{k,k'}[\ell]$ does not i.e.,

$$\mathcal{J}_{k,k'}[\ell] \triangleq (\boldsymbol{g}_k[n]^{[i]})^\dagger \sum_{\Delta n=-1}^{0} \boldsymbol{g}_{k'}[n+\Delta n]^{[i]} d_{k'}[n+\Delta n]. \quad (14)$$

For the problem at hand, the trellis has five states $s_n[d_k[\ell], d_{k'}[\ell]]$, $n = 0,\ldots,4$, and 16 state transitions denoted by

$$\check{\gamma}_{s',s}[d_k[\ell], d_{k'}[\ell], d_{k'}[\ell-1] \triangleq \check{\gamma}_{s',s}[\ell] \quad (15)$$

as shown in Fig. 1. Notice that the system is causal.

For $\tau^{[i]}_k > \tau^{[i]}_{k'}$, the statements (12)–(15) are the same but with interchanged indexes k and k'.

From the *sequence* of joint probabilities

$$P\left(d_k[\ell] = \pm\mu, d_{k'}[\ell'] = \pm\mu' \mid \boldsymbol{d}^{[i,m]}_{\overline{k,k'}}, \boldsymbol{r}, \boldsymbol{\tau}^{[i]}, \boldsymbol{a}^{[i]}\right),$$

$\ell = 1,\ldots,L$, $\ell' = \{\ell-1, \ell, \ell+1\}$ in (11), the Gibbs sampler draws the matrix $(\boldsymbol{D})^{[i,m]}_{k,k'} \in \{-1,+1\}^{L\times 3}$ with $[(d_k[\ell]\ d_{k'}[\ell-1])^{[i,m]}, (d_k[\ell]\ d_{k'}[\ell])^{[i,m]}, (d_k[\ell]\ d_{k'}[\ell+1])^{[i,m]}]$ as its ℓth row, and computes the corresponding indicator functions in (12) at epoch t.

When no symbol interleavers are present, the MCMC implementation of the joint APPs in (11) yields

$$P\left(d_k[\ell] = \mu, d_{k'}[\ell'] = \mu' \mid \boldsymbol{r}, \boldsymbol{\tau}^{[i]}, \boldsymbol{a}^{[i]}\right) \approx \frac{1}{M}\sum_{m=1}^{M} \delta\left(d_k[\ell] = \mu, d_{k'}[\ell'] = \mu' \mid \boldsymbol{d}^{[i,m]}_{\overline{k,k'}}, \boldsymbol{r}, \boldsymbol{\tau}^{[i]}, \boldsymbol{a}^{[i]}\right), \quad (16)$$

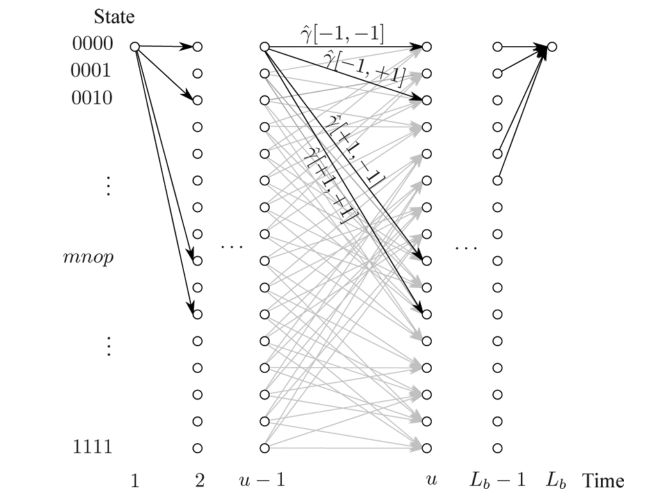

Fig. 2. Joint trellis for two coded users k and k' with $\tau_k, \tau_{k'} < T_d$ and $G = [5_8 7_8]$.

$\ell = 1,\ldots,L$ and $\boldsymbol{d}_{\overline{k,k'}} \triangleq \boldsymbol{d}\backslash\{\boldsymbol{d}_k, \boldsymbol{d}_{k'}\}$. The computation of (16) requires the evaluation of the joint probability $P\left(d_k[\ell] = \mu, d_{k'}[\ell'] = \mu' \mid \boldsymbol{d}^{[i,m]}_{\overline{k,k'}}, \boldsymbol{r}, \boldsymbol{\tau}^{[i]}, \boldsymbol{a}^{[i]}\right)$ which again can be factored as

$$P\left(d_k[\ell] = \mu, d_{k'}[\ell'] = \mu' \mid \boldsymbol{d}^{[i,m]}_{\overline{k,k'}}, \boldsymbol{r}, \boldsymbol{\tau}^{[i]}, \boldsymbol{a}^{[i]}\right) = \hat{C} \sum_{\mathcal{S}^{\mu\mu'}_{\ell,\ell'}} \hat{\alpha}_{s'}[u-1]\hat{\gamma}_{s',s}[u]\hat{\beta}_s[u]$$

where $\mathcal{S}^{\mu\mu'}_{\ell,\ell'}$ is the set of state pairs (s', s) in the joint trellis so that $d_k[\ell] = \mu, d_{k'}[\ell'] = \mu'$. The joint trellis T for the user k and user k' codes is the Shannon-product of the individual trellis [31]. The normalization constant $\hat{C}$ results from the boundary condition $\sum_{x,y\in\mathcal{D}} P\left(d_k[\ell] = x, d_{k'}[\ell'] = y \mid \boldsymbol{d}^{[i,m]}_{\overline{k,k'}}, \boldsymbol{r}, \boldsymbol{\tau}^{[i]}, \boldsymbol{a}^{[i]}\right) = 1$. The definition of the probabilities $\hat{\alpha}$, $\hat{\beta}$, and $\hat{\gamma}$ is similar to that of α, β, and γ in Section III-D-2). Exploiting the block-diagonal structure of $\boldsymbol{S}(\boldsymbol{\tau})^\dagger\boldsymbol{S}(\boldsymbol{\tau})$ again, it follows for the logarithm of the transition probability $\gamma_{s',s}[u]$ that

$$\ln \hat{\gamma}_{s',s}[u] \propto \frac{2}{N_0} \sum_{\ell=R^{-1}u}^{R^{-1}(u+1)-1} \sum_{j\neq k,k'} \Re\{(\boldsymbol{g}_k[\ell]^{[i]})^\dagger\boldsymbol{r} - d_k[\ell]\mathcal{J}^{[m]}_{k,j}[\ell] + (\boldsymbol{g}_{k'}[\ell]^{[i]})^\dagger\boldsymbol{r} - d_{k'}[\ell]\mathcal{J}^{[m]}_{k',j}[\ell] - \begin{cases} d_k[\ell]\mathcal{J}_{k,k'}[\ell]; & \tau^{[i]}_k < \tau^{[i]}_{k'} \\ d_{k'}[\ell]\mathcal{J}_{k',k}[\ell]; & \tau^{[i]}_k \geq \tau^{[i]}_{k'}. \end{cases}$$

The interference terms $\mathcal{J}^{[m]}_{k,k'}[\ell]$ and $\mathcal{J}_{k,k'}[\ell]$ are defined in (10) and (14), respectively. The joint trellis is illustrated in Fig. 2 for two users k and k' with generator matrix $G = [5_8 7_8]$ in octal notation. The trellis has 16 states $s_n[\mathcal{D}_k[u], \mathcal{D}_{k'}[u]]$ with $\mathcal{D}_k[u]$ denoting the state of user k at time interval u in binary notation, $n = 0,\ldots,15$, and 64 state transitions denoted by $\hat{\gamma}[b_k[u], b_{k'}[u]] \triangleq \hat{\gamma}_{s',s}[u]$. It can be clearly seen that $\hat{\gamma}_{s',s}[u]$ incorporates user *and* coding constraints.

4) The M-Step: The M-step of the SAGE algorithm is realized by first maximizing (5) with respect to the transmission delays τ_k:

$$\tau_k^{[i+1]} = \arg\max_{\tau_k} \left| \sum_{\ell=1}^{L} \lambda(\ell, \tau_k) \right|. \quad (17)$$

Then by inserting (17) into (5), taking derivatives with respect to the a_k's, setting the results equal to zero, and solving yields

$$a_k^{[i+1]} = \frac{1}{L + \frac{N_0}{\sigma_k^2}} \sum_{\ell=1}^{L} \lambda(\ell, \tau_k^{[i+1]}). \quad (18)$$

5) Uncoded Transmission: To obtain a low-complex Monte Carlo implementation of the SAGE scheme, when channel coding is not present i.e., $\boldsymbol{d} = \boldsymbol{b}$, samples from the *symbol* posterior $P\left(d_k[\ell] = \mu \mid \bar{\boldsymbol{d}}^{[i,m]}, \boldsymbol{r}, \boldsymbol{\tau}^{[i]}, \boldsymbol{a}^{[i]}\right)$ with $\bar{\boldsymbol{d}} = \boldsymbol{d} \backslash d_k[\ell]$, can be used to approach the APP in (6), i.e.,

$$P\left(d_k[l] = \mu \mid \boldsymbol{r}, \boldsymbol{\tau}^{[i]}, \boldsymbol{a}^{[i]}\right) \approx \frac{1}{M} \sum_{m=1}^{M} \delta\left(d_k[\ell] = \mu \mid \bar{\boldsymbol{d}}^{[i,m]}, \boldsymbol{r}, \boldsymbol{\tau}^{[i]}, \boldsymbol{a}^{[i]}\right). \quad (19)$$

After some algebra, the symbol posterior can be expressed as

$$P\left(d_k[\ell] = \mu \mid \bar{\boldsymbol{d}}_k^{[i,m]}, \boldsymbol{r}, \boldsymbol{\tau}^{[i]}, \boldsymbol{a}^{[i]}\right) = \frac{1}{1 + \exp\left\{ -\frac{4}{N_0} \mu \Re\{ (\boldsymbol{g}_k[\ell]^{[i]})^\dagger \boldsymbol{r} - \sum_{j \neq k} \mathcal{J}_{k,j}^{[m]}[\ell] \} \right\}} \quad (20)$$

with the interference term $\mathcal{J}_{k,k'}^{[m]}[\ell]$ defined in (10). Similarly, the joint symbol posterior [4]

$$P\left(d_k[\ell] = \mu, d_{k'}[\ell'] = \mu' \mid \bar{\bar{\boldsymbol{d}}}^{[i,m]}, \boldsymbol{r}, \boldsymbol{\tau}^{[i]}, \boldsymbol{a}^{[i]}\right)$$

with $\bar{\boldsymbol{d}} = \boldsymbol{d} \backslash \{d_k[\ell], d_{k'}[\ell']\}$ can be used to approximate the APP in (7) according to

$$\begin{aligned} P\left(d_k[\ell] = m, d_{k'}[\ell'] = n \mid \boldsymbol{r}, \boldsymbol{\tau}^{[i]}, \boldsymbol{a}^{[i]}\right) &\approx \frac{1}{M} \sum_{m=1}^{M} \delta\left(d_k[\ell] = \mu \mid d_{k'}[\ell'] = \mu', \bar{\bar{\boldsymbol{d}}}^{[i,m]}, \boldsymbol{r}, \boldsymbol{\tau}^{[i]}, \boldsymbol{a}^{[i]}\right) \\ &\quad \times P\left(d_{k'}[\ell'] = \mu' \mid \boldsymbol{r}, \boldsymbol{\tau}^{[i]}, \boldsymbol{a}^{[i]}\right). \end{aligned} \quad (21)$$

To evaluate (21), we need to compute the conditional probability density

$$P\left(d_k[\ell] = \mu \mid d_{k'}[\ell'] = \mu', \bar{\boldsymbol{d}}^{[i,m]}, \boldsymbol{r}, \boldsymbol{\tau}^{[i]}, \boldsymbol{a}^{[i]}\right) = \frac{1}{1 + \exp\{ -\frac{4}{N_0} (\mu \Re\{ \mathcal{M}_{k,k'}^{[m]}[\ell, \ell'] \} - \mu\mu' (\boldsymbol{g}_k[\ell]^{[i]})^\dagger \boldsymbol{g}_{k'}[\ell']^{[i]}) \}}$$

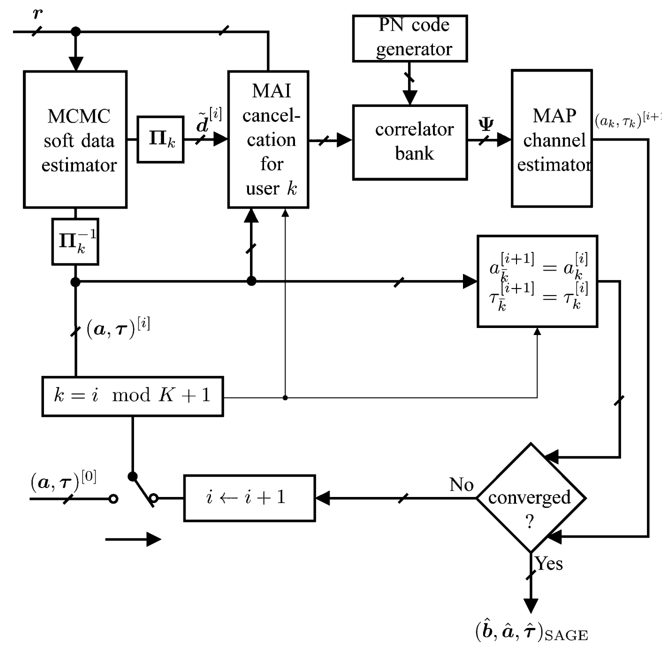

Fig. 3. Block diagram of the MCMC-SAGE scheme.

with the abbreviated notation

$$\mathcal{M}_{k,k'}^{[m]}[\ell, \ell'] \triangleq \left(\boldsymbol{g}_k[\ell]^{[i]}\right)^\dagger \boldsymbol{r} - \sum_{j \neq k,k'} \mathcal{J}_{k,j}^{[m]}[\ell].$$

The APPs $P\left(d_{k'}[\ell'] \mid \boldsymbol{r}, \boldsymbol{\tau}^{[i]}, \boldsymbol{a}^{[i]}\right)$ in (21) are given in (19).

According to (17) and (18), this so called MCMC-SAGE scheme updates user $k = i \mathrm{mod} K + 1$ at iteration i. First, the MCMC-SAGE scheme iteratively approaches the vector estimate of the users' data symbols $\boldsymbol{d}^{[i]}$ and its measurement error covariance matrix by using the Monte Carlo steps in (8) and (12), respectively. Based on these new estimates, a guess of the k^{th} user's interference is computed and stripped away from the received signal in (5). Then, the cleaned signal is fed into a correlator bank that matches the input signal with $N_c Q$ time-shifted versions of the k^{th} user's signature waveform $\boldsymbol{S}_k$. The largest output signal is selected by subsequent channel estimator providing $a_k^{[i+1]}$ and $\tau_k^{[i+1]}$. A block diagram of the MCMC-SAGE receiver is shown in Fig. 3.

E. Initialization

A user specific preamble of L_p symbols, embedded in every user's data block, is used to initialize the MCMC-SAGE scheme. Starting from $\boldsymbol{\tau}^{[-1]} = \boldsymbol{0}$ and $\boldsymbol{a}^{[-1]} = \boldsymbol{0}$, the timing and channel coefficients are updated alternatively according to (17) and (18) in successive order until either convergence is achieved or the number of stages is two. In one stage, every user is updated once i.e., one stage corresponds to K iterations. The final parameter estimates are assigned to $\boldsymbol{\tau}^{[0]}$ and $\boldsymbol{a}^{[0]}$ respectively. The initial estimate of the data vector $\boldsymbol{d}$ is set randomly.

IV. Performance Analysis

A. Convergence Rate

The convergence rate affects both the length of the burn-in period and the efficiency of the posterior estimation based on Monte Carlo samples. Low convergence rate results in inaccurate estimates in the empirical *a posteriori* probabilities and joint *a posteriori* probabilities in (11) and (8), respectively, computed by means of the Monte Carlo samples. In the following subsections, the convergence properties of the MCMC algorithms in general settings are established and a technique to compute the convergence rate of the MCMC-SAGE scheme exactly is given. Showing that the computations of the exact convergence rate is mathematically intractable for practical cases, an approximate technique is then presented to compute the convergence rate based on Gaussian approximation which reduces the computational complexity substantially.

1) Exact Convergence Rate of the Gibbs Sampler: We now construct a Markov transition rule in the MCMC algorithm presented in Section III-C so that its limiting distribution is the desired posterior distribution. For simplicity of notation we drop the indexes i and k used for the SAGE iterations and for the user k, respectively. That is, $\boldsymbol{d}_k^{[i,m]}$ is denoted by $\boldsymbol{d}^{[m]}$. For ease of exposition we also assume that the channel coding is not present. However, the results can be easily extended to the coded case.

Let the ergodic pseudo-random samples $\boldsymbol{d}^{[1]}, \boldsymbol{d}^{[2]}, \ldots, \boldsymbol{d}^{[m]}, \ldots, \boldsymbol{d}^{[M]}$ be generated by the kth($k = 1, 2, \ldots, K$) user, as explained in Section III-D, from a discrete-time Markov source, where $\boldsymbol{d}^{[m]} = \left[d^{[m]}[1], d^{[m]}[2], \ldots, d^{[m]}[L]\right]^T$. The 2^L distinct states of the Markov chain are denoted by the L-tuple $\boldsymbol{\delta}_\zeta = [\delta_\zeta(1), \delta_\zeta(2), \ldots, \delta_\zeta(L)]^T$, $\delta_\zeta(.) \in \{-1, +1\}$, $\zeta = 0, 1, \ldots, 2^L - 1$ and state transitions of the chain are governed by the transition probabilities $P\left(\boldsymbol{d}^{[m]} = \boldsymbol{\delta}_\zeta \mid \boldsymbol{d}^{[m-1]} = \boldsymbol{\delta}_{\zeta'}\right) = p_{\zeta\zeta'}$, $\zeta, \zeta' = 0, 1, \ldots, 2^L - 1$, from which an $2^L \times 2^L$ Markov probability transition matrix $\mathbf{P} = [p_{\zeta\zeta'}]$ is formed. We also define the m-step transition probability matrix as $\mathbf{P}^{[m]} = [p_{\zeta\zeta'}^{[m]}]$ where $p_{\zeta\zeta'}^{[t]} = P\left(\boldsymbol{d}^{[m]} = \boldsymbol{\delta}_\zeta \mid \boldsymbol{d}^{[0]} = \boldsymbol{\delta}_{\zeta'}\right)$.

A Markov chain is said to be *irreducible* if its state space is a single communicating class; in other words, if it is possible to reach any state from other state. Formally, state ζ' is accessible from state ζ if there exists an integer $m \geq 0$ such that $p_{\zeta\zeta'}^{[m]} > 0$. Similarly, a Markov state is said to be *aperiodic* if the greatest common divisor of the set $\{m : p_{\zeta\zeta'}^{[m]}\}$ is 1. That is the returns to state ζ can occur at irregular times. Clearly, if a state ζ is aperiodic and the chain is irreducible, then every state of the Markov chain is aperiodic [32].

We now state the following convergence result without proof. A proof can be found in [33]. Suppose a discrete-time Markov chain is irreducible and aperiodic and $\mathbf{p}_\zeta^{[m]}, \zeta = 0, 1, \ldots, 2^L - 1$, denotes the ζth column vector of $\mathbf{P}^{[m]}$. Then for every ζ, $\mathbf{p}_\zeta^{[m]}$ converges to the stationary probability distribution $\boldsymbol{\pi}$ ($\boldsymbol{\pi}\mathbf{P} = \boldsymbol{\pi}$) geometrically. That is, there exist $0 < \rho < 1$ and $c > 0$ such that for $\zeta = 0, 1, \ldots, 2^L - 1, \| \mathbf{p}_\zeta^{[m]} - \boldsymbol{\pi} \| \leq c\rho^t$, implying that $\lim_{t \to \infty} \mathbf{P}^{[m]} = \mathbf{P}$. On the other hand ρ can be determined as follows. The Jordan decomposition [34] of the Markov transition matrix $\mathbf{P} \in \mathcal{R}^{2^L \times 2^L}$ is

$$\mathbf{X}\mathbf{P}\mathbf{X}^{-1} = \mathrm{diag}(\mathbf{J}_1, \mathbf{J}_2, \ldots, \mathbf{J}_\kappa) \tag{22}$$

where, $\mathbf{X} \in \mathcal{R}^{2^L \times 2^L}$ is a nonsingular matrix and $\mathbf{J}_l \in \mathcal{C}^{m_l \times m_l}$'s with $m_1 + m_2 + \cdots + m_\kappa = 2^L$ are referred to Jordan blocks corresponding to $|\lambda_1| \geq |\lambda_2| \geq \cdots \geq |\lambda_\kappa|$, the eigenvalues of $\mathbf{P}$, respectively. Since the transition probability matrix $\mathbf{P}$ is nonnegative, it is shown by the Perron–Frobenius theorem [35] that the largest eigenvalue of $\mathbf{P}$ is unity in absolute value, that is $\mid \lambda_1 \mid = 1$, and has the multiplicity of one if $\mathbf{P}$ is irreducible and, furthermore, $\mid \lambda_2 \mid < 1$ if $\mathbf{P}$ is aperiodic. It is clear from (1) that $\mathbf{P}^{[m]} = \mathbf{X}^{-1}\, \mathrm{diag}\,(1, \mathbf{J}_2^m, \ldots, \mathbf{J}_\kappa^m)\, \mathbf{X}$ and consequently as $m \to \infty$, the limit of $\mathbf{P}^{[m]}$ approaches its stationary distribution $\boldsymbol{\pi}$ with convergence rate $\rho = \mid \lambda_2 \mid$.

For the uncoded case, the transition probability matrix $\mathbf{P}_k$ for the Markov chain generated by the user k, $k = 1, 2, \ldots, K$, can be obtained from the Gibbs sampler as follows:

$$\begin{aligned}\mathbf{P}_k &= P\left(\boldsymbol{d}_k^{[m]} \mid \boldsymbol{d}_k^{[m-1]}\right) \triangleq P\left(\boldsymbol{d}_k^{[m]} \mid \boldsymbol{d}_k^{[m-1]}, \boldsymbol{d}_{\bar{k}}^{[m-1]}, \boldsymbol{r}, \boldsymbol{\tau}, \boldsymbol{a}\right)\\ &= \prod_{l=1}^{L} P\left(d_k^{[m]}[l] \mid d_k^{[m]}[1], \ldots, d_k^{[m]}[l-1], d_k^{[m-1]}[l+1]\right.\\ &\qquad \left. \ldots, d_k^{[m-1]}[L], \boldsymbol{d}_{\bar{k}}^{[m-1]}, \boldsymbol{r}, \boldsymbol{\tau}, \boldsymbol{a}\right).\end{aligned} \tag{23}$$

The conditional probabilities on the right-hand side of (23) are given by (20). Notice that the computation of the transition matrix $\mathbf{P}$ in (22) is bounded by $\mathcal{O}(2^L)$ operations which makes a practical implementation impossible. Hence, we rely on sub-optimal methods.

2) Convergence Rate of the Gibbs Sampler by Gaussian Approximation: Suppose no channel coding is employed. If we assume that $\boldsymbol{d}$ in (1) is approximately Gaussian with mean zero and unit covariance matrix then $p(\boldsymbol{d} \mid \boldsymbol{r})$ is also multivariate Gaussian with mean $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$. It can be easily shown that $\boldsymbol{\mu} = \boldsymbol{\Sigma}\boldsymbol{A}^\dagger\boldsymbol{S}^\dagger\boldsymbol{r}$ and $\boldsymbol{\Sigma} = (\boldsymbol{A}^\dagger\boldsymbol{S}^\dagger\boldsymbol{S}\boldsymbol{A} + N_0\boldsymbol{I})^{-1}$.

Under the Gaussian assumption the rate of convergence can be obtained as follows [36]. Let $\boldsymbol{\Sigma}^{-1} = \mathbf{L} - \mathbf{U}$ where $\mathbf{L}$ is the lower triangular part of matrix $\boldsymbol{\Sigma}^{-1}$, and $\mathbf{U}$ is upper triangular with null diagonal elements. Further, let $\mathbf{c} = (\boldsymbol{I} - \mathbf{C})\boldsymbol{\mu}$ where $\mathbf{C} = \mathbf{L}^{-1}\mathbf{U}$. Then, it can be shown that the Markov chain induced by the Gibbs sampling has a normal transition density with mean $E(\boldsymbol{d}^{[m+1]} \mid \boldsymbol{d}^{[m]}) = \mathbf{C}\boldsymbol{d}^{[m]} + \mathbf{c}$ and covariance $\boldsymbol{\Sigma} - \mathbf{C}\boldsymbol{\Sigma}\mathbf{C}^\dagger$. Thus, $\{\boldsymbol{d}^{[m]}\}$ is a multivariate AR(1) process. It follows that the rate of the convergence of the Gibbs sampling algorithm is $\rho(\mathbf{C})$, the spectral radius of the matrix $\mathbf{C}$. An upper bound to the spectral radius is given as follows. For each $k \in \aleph$

$$\rho(\mathbf{C}) \leq \|\mathbf{C}^k\|_F^{1/k}$$

where the Frobenius norm $\|\mathbf{C}\|_F$ of a complex-valued matrix $\mathbf{C}$ is defined as $\|\mathbf{C}\|_F = \sqrt{\mathrm{trace}(\mathbf{C}^\dagger\mathbf{C})}$. Consequently, $\rho(\mathbf{C}) = \lim_{k \to \infty} \|\mathbf{C}^k\|_F^{1/k}$.

B. Modified Cramér–Rao Bound for the Estimated Parameters

The task is to derive the Cramér–Rao bound (CRB) on the variances of any unbiased estimates $\hat{\boldsymbol{\theta}}$ of the parameter vector $\boldsymbol{\theta}$

based on the observation $\boldsymbol{r}$. It is shown in [37] that for $\theta_p \in \boldsymbol{\theta}$, $\text{var}(\hat{\theta}_p - \theta_p) \geq [\boldsymbol{I}^{-1}(\boldsymbol{\theta})]_{pp}$, where $\boldsymbol{I}(\boldsymbol{\theta})$ is the $3K \times 3K$ Fisher information matrix whose (p,q)th component is defined by

$$[\boldsymbol{I}(\boldsymbol{\theta})]_{pq} \triangleq -E_{\boldsymbol{r}}\left\{\frac{\partial^2 \ln p(\boldsymbol{r} \mid \boldsymbol{a}, \boldsymbol{\tau})}{\partial\theta_p \partial\theta_q}\right\}, \text{ for } p,q = 1,2,\ldots,3K.$$

As $\boldsymbol{r}$ depends on the nuisance parameter vector $\boldsymbol{d}$, as well, $p(\boldsymbol{r} \mid \boldsymbol{a}, \boldsymbol{\tau}, \boldsymbol{d})$ should first be averaged over the random data vector $\boldsymbol{d}$ i.e., $p(\boldsymbol{r} \mid \boldsymbol{a}, \boldsymbol{\tau}) = E_{\boldsymbol{d}}\{p(\boldsymbol{r} \mid \boldsymbol{a}, \boldsymbol{\tau}, \boldsymbol{d})\}$. This task, however, is cumbersome to solve. To circumvent this drawback, we resort to the so-called modified CRB (MCRB) [38] that relies on the definition of the modified Fisher information matrix

$$[\tilde{\boldsymbol{I}}(\boldsymbol{\theta})]_{pq} \triangleq -E_{\boldsymbol{d}}\left\{E_{\boldsymbol{r}|\boldsymbol{d}}\left\{\frac{\partial^2 \ln p(\boldsymbol{r} \mid \boldsymbol{a}, \boldsymbol{\tau}, \boldsymbol{d})}{\partial\theta_p\, \partial\theta_q}\right\}\right\}.$$

For the joint likelihood function in (4), one gets after some manipulation

$$[\tilde{\boldsymbol{I}}(\boldsymbol{\theta})]_{pq} = \frac{2}{N_0} E_{\boldsymbol{d}}\left\{\Re\left[\frac{\partial\boldsymbol{\mu}^\dagger(\boldsymbol{\theta}, \boldsymbol{d})}{\partial\theta_p}\frac{\partial\boldsymbol{\mu}(\boldsymbol{\theta}, \boldsymbol{d})}{\partial\theta_q}\right]\right\}, \tag{24}$$

$p, q = 1, 2, \ldots, 3K$. On taking the expectations with respect to the data $\boldsymbol{d}$ after taking the partial derivatives in (24) with respect to θ_p and θ_q, for different regions of p and q values, under the assumption that the data sequences are independent and equally likely and the fact that $\boldsymbol{S}^\dagger(\tau_p, \ell)\boldsymbol{S}(\tau_p, \ell) = 1$, for $p = 1, 2 \ldots K$; $\ell = 0, 1, \ldots, L-1$, the modified Fisher information matrix becomes diagonal whose (p,p)th component can be evaluated as

$$[\tilde{\boldsymbol{I}}(\boldsymbol{\theta})]_{pp} = \frac{2}{N_0}\begin{cases} L; & p = 1, \ldots, K \\ L; & p = K+1, \ldots, 2K \\ \sigma_p^2 \sum\limits_{\ell=0}^{L-1} \mid \boldsymbol{S}'(\ell) \mid^2; & p = 2K+1, \ldots, 3K \end{cases} \tag{25}$$

with the abbreviated notation $\boldsymbol{S}'[\ell] \triangleq \partial\boldsymbol{S}_p(\tau_p, \ell)/\partial\tau_p \quad |_{t=\ell T_b + \hat{\tau}_p}$. The final result for the MCRBs on the estimates of the channel coefficients and the transmission delays is obtained by inverting the diagonal matrix $\tilde{\boldsymbol{I}}(\boldsymbol{\theta})$ in (25) as follows:

$$\text{var}(\hat{a}_k) \geq \frac{N_0}{L}, \tag{26}$$

$$\text{var}(\hat{\tau}_k) \geq \frac{N_0}{(8\pi^2 L B_{s_k}^2)}, \tag{27}$$

$k = 1, 2, \ldots, K$. The Gabor bandwidth of the kth user's spreading code waveform, $s_k(t)$, is given by

$$B_{s_k} \triangleq \left(\int_{-\infty}^{+\infty} f^2 \mid S_k(f) \mid^2 df\right)^{1/2},$$

and $S_k(f)$ is the Fourier transform of $s_k(t),\ t \in [0, T_b]$. Note that the Gabor bandwidth B_{s_k} tends to infinity for rectangular-shaped (continuous-time) chip waveforms.

C. Numerical Examples

To illustrate the performance of the MCMC-SAGE scheme, we consider a CDMA system with random spreading. All users employ $R = 1/2$ terminated convolutional codes with generator matrix $G = [5_8, 7_8]$. Data blocks of length $L = 160$ code bits comprising $L_p = 6$ pilot bits are fed into a symbol interleaver of size $L_\Pi = 3200$ bits and sent over asynchronous quasi-static flat-Rayleigh fading channels. The receiver processes $Q = 4$ samples per chip.[1] The Gibbs sampler draws $M = 5$ samples per SAGE iteration. The number of burn-in samples is set to 0.

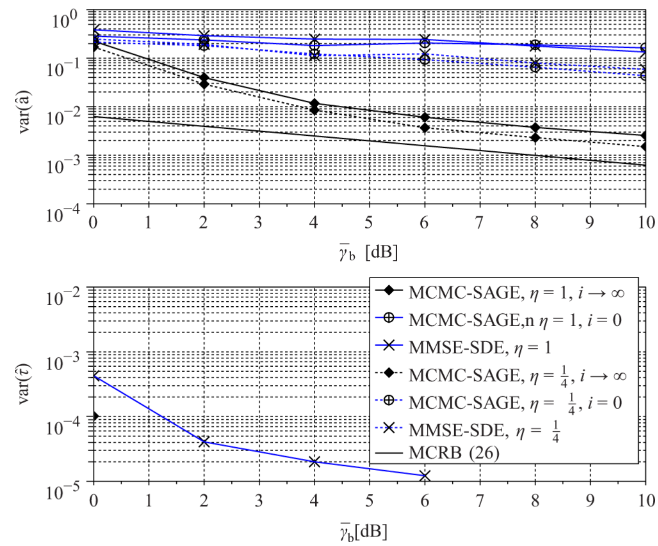

Fig. 4. MSE performance versus SNR in quasi-static Rayleigh block-fading for various effective system loads η: $N_c = 8$, $\tau/T_d \in]0, 0.5[$, $L_p = 6$, $L = 160$, $L_\Pi = 3200$, CC(5,7).

For comparison purpose, the performance of a so-called minimum mean-square error (MMSE) separate decoding and channel estimation (MMSE-SDE) scheme has been added to the plots as well. The MMSE-SDE scheme comprises an MMSE channel estimator, a separate linear MMSE equalizer [40], neglecting code constraints, and a separate max-logMAP data decoder, neglecting user constraints. As the MMSE channel estimate is a function of the (unknown) code timing, we proceed as follows. First, a sequence of K-dimensional MMSE channel estimates is computed based on the preamble symbols, and the received signal $\boldsymbol{r}$ for different delay values of user 1 while the delays of the other users are kept zero. Based on the strongest delay for user 1, the procedure is repeated for user 2 and so on and so forth until the last user is reached.

Fig. 4 shows the average mean-square-error (MSE) of the channel estimates $\hat{\boldsymbol{a}}$ and transmission delay estimates $\hat{\boldsymbol{\tau}}$ as a function of the *effective* SNR $\bar{\gamma}_b = L - L_p/L\mathcal{E}/N_0$ with the *effective* system load $\eta \triangleq KR/N_c$ as parameter. The normalized transmission delays are uniformly distributed in $\tau/T_d \in]0,\ 0.5[$. Starting from a gap of larger than 20 dB with respect to the MCRB (26) at $i = 0$, the MSE of $\boldsymbol{a}$ approaches the MCRB up to 3.0 dB (5.0 dB) for $\eta = 0.25$ ($\eta = 1$) at $\bar{\gamma}_b = 10$ dB. For the MMSE-SDE scheme, the gap is larger than 20 dB with respect to the MCRB over the entire range of SNR. For the MSE of $\boldsymbol{\tau}$, the MCMC-SAGE is capable of finding the correct transmission delays already at $i = 0$ over the entire range of SNR, indicating that the MCMC-SAGE scheme is robust against delay estimation errors. The MMSE-SDE scheme does the same for small

[1]The value $Q = 4$ is a widely used compromise among time resolution, delay estimate error variance, and computational complexity [39].

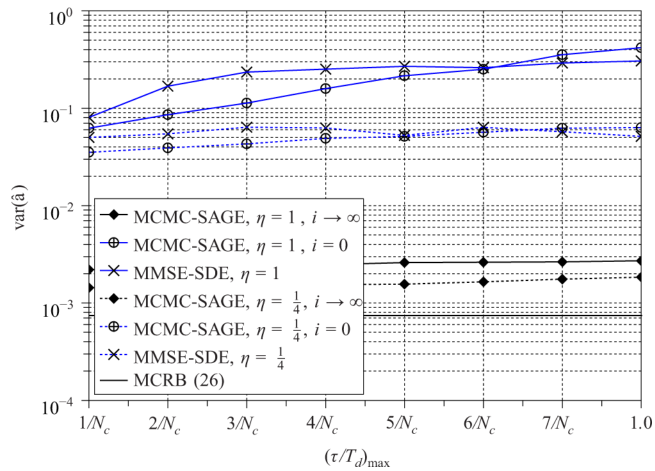

Fig. 5. MSE performance versus transmission delays in quasi-static Rayleigh fading for various effective system loads η. The normalized transmission delays are uniformly distributed on $\tau/T_d \in]0, (\tau/T_d)_{\max}[$. Parameters: $N_c = 8$, $L_p = 6$, $L = 160$, $L_\Pi = 3200$, CC(5,7), $\bar{\gamma}_b = 10$ dB.

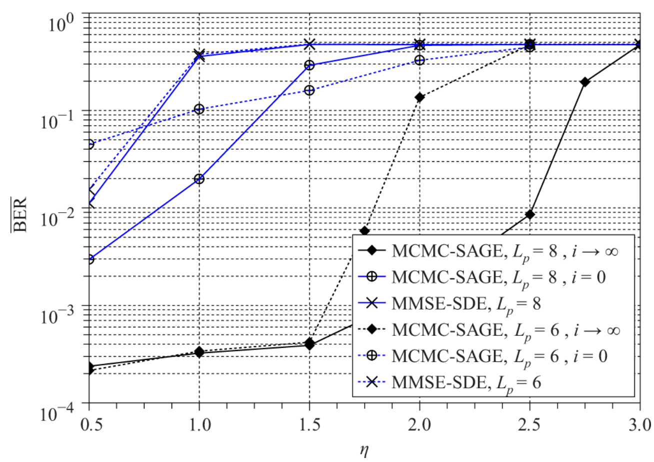

Fig. 6. Supported effective system load η in quasi-static flat Rayleigh block-fading: $N_c = 8$, $\tau/T_d \in]0, 0.5[$, $L = 160$, $L_\Pi = 3200$, CC(5,7), $\bar{\gamma}_b = 10$ dB.

loads ($\eta = 0.25$) but has difficulties with handling high loads ($\eta = 1$).

In Fig. 5, the MSE of the channel estimates is plotted versus the maximum transmission delay $(\tau/T_d)_{\max}$, $1/N_c \le (\tau/T_d)_{\max} \le 1$ and $\tau/T_d \in]0, (\tau/T_d)_{\max}[$ with the system load as parameter. All users are received with average SNR $\bar{\gamma}_b = 10$ dB. The receiver is not synchronized. It can be seen that the performance of the MCMC-SAGE scheme is roughly independent of $(\tau/T_d)_{\max}$ as i tends to ∞.

The average bit error rate $\overline{\mathrm{BER}}$ is plotted in Fig. 6 versus the effective system load of η for $\bar{\gamma}_b = 10$ dB when no synchronization information is available at the receiver. Notice that averaging is performed over all the realizations of the channel coefficients, transmission delays, and the users. When the load is increased, the MCMC-SAGE receiver performs nearly optimally until a load threshold $\eta_{\max}$. Beyond $\eta_{\max}$, the increasing cross-correlations among different users overwhelm the receiver and the average bit error rate $\overline{\mathrm{BER}}$ tends to values near 0.5. A theoretical justification of this effect can be found in [41].

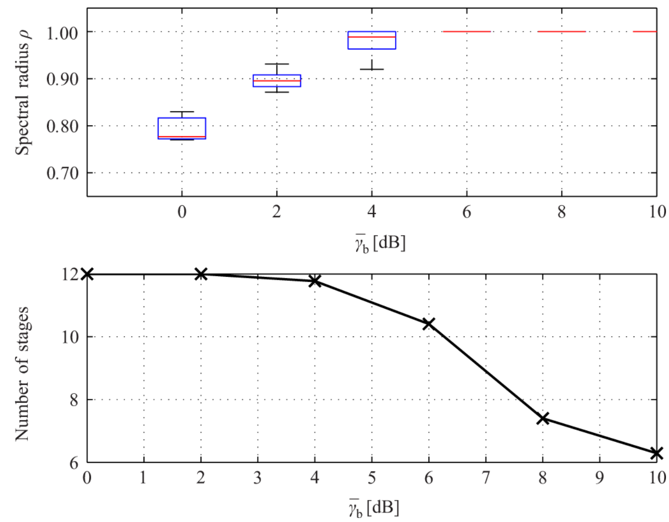

Fig. 7. a) Convergence rate of the Gibbs sampler by Gaussian approximation; b) Convergence behavior of the MCMC-SAGE scheme: flat Rayleigh block-fading, $\eta = 1$, $N_c = 8$, $\tau/T_d \in]0, 0.5[$, $L = 160$, $L_\Pi = 3200$, CC(5,7).

For the cut-off error rate at the traditional erroneous rejection level of at most 1%, the MCMC-SAGE scheme supports $\eta = 1.75$ ($K = 28$), while the MMSE-SDE scheme manages $\eta < 0.5$ ($K < 8$). When the number of pilot symbols is increased to $L_p = 8$ i.e., 5% overhead, the proposed MCMC-SAGE supports $\eta = 2.5$ ($K = 40$). It can be seen that the proposed scheme is capable of handling extreme system loads in contrast to previously published schemes for joint multiuser estimation and synchronization [21].

Fig. 7(a) shows a box-plot of the Gibbs sampler's convergence rate ρ by Gaussian approximation versus the effective SNR for $\eta = 1$. It can be seen that the median value of the spectral radius is $\rho = 0.77$ for $\bar{\gamma}_b = 0$ dB (low SNR) and $\rho = 1$ for $\bar{\gamma}_b \le 6$ (high SNR). The poor performance at high SNR mainly arises from the ambiguity of the system. Note that $p(\boldsymbol{d}, \boldsymbol{a}, \boldsymbol{\tau}|\boldsymbol{r}) = p(-\boldsymbol{d}, -\boldsymbol{a}, \boldsymbol{\tau}|\boldsymbol{r})$ for all $\boldsymbol{d}$, $\boldsymbol{a}$, and $\boldsymbol{\tau}$. With increasing SNR the gap between the two modes becomes larger, making it more difficult for the MCMC to move from one mode to the other [42]. The convergence behavior of the overall MCMC-SAGE scheme is depicted in Fig. 7(b) for the same settings. With increasing SNR the iterative MCMC-SAGE scheme converges more often to a global maximum of the likelihood function with fewer iterations [6]. For $\bar{\gamma}_b = 6$ dB and $\bar{\gamma}_b = 10$ dB, respectively, the MCMC-SAGE scheme requires 10.4 and 6.3 stages to converge.

V. Summary

A computationally efficient algorithm has been proposed for joint time acquisition, multi-channel estimation, and multiuser soft-data decoding based on the SAGE algorithm. At each iteration the joint *a posteriori* probabilities of all users' data symbols are forwarded to one particular user's joint channel coefficient and timing estimator and vice versa. A Gibbs sampling technique from Markov Chain Monte Carlo statistical signal processing is used to compute the joint *a posteriori* probability. Exact analytical expressions have been obtained for the estimates of transmission delays and the channel coefficients. Convergence in likelihood is guaranteed for the proposed algorithm.

Monte Carlo simulations for asynchronous coded DS/CDMA over flat Rayleigh fading channels show that the proposed Monte Carlo SAGE scheme supports the remarkable system load of $\eta = 2.5$ for a small, say 5%, pilot overhead, taking into account the rate of the channel codes.

REFERENCES

[1] T. G. Giallorenzi and S. G. Wilson, "Multiuser ML sequence estimator for convolutional coded asynchronous DS-CDMA systems," *IEEE Trans. Commun.*, vol. 44, pp. 997–1008, Aug. 1996.

[2] P. D. Alexander, A. J. Grant, and M. C. Reed, "Iterative detection in code-division multiple-access with error control coding," *Eur. Trans. Telecommun.*, vol. 9, pp. 419–425, Sep.-Oct. 1998.

[3] C. Berrou and A. Glavieux, "Near optimum error correcting coding and decoding: Turbo-codes," *IEEE Trans. Commun.*, vol. 44, pp. 1261–1271, Oct. 1996.

[4] X. Wang and H. V. Poor, "Iterative (turbo) soft interference cancellation and decoding for coded CDMA," *IEEE Trans. Commun.*, vol. 47, pp. 1046–1061, Jul. 1999.

[5] J. Boutros and G. Caire, "Iterative multiuser joint decoding: Unified framework and asymptotic analysis," *IEEE Trans. Inf. Theory*, vol. 48, pp. 1772–1793, Jul. 2002.

[6] A. Kocian, I. Land, and B. H. Fleury, "Joint channel estimation, partial successive interference cancellation, and data decoding for DS-CDMA based on the SAGE algorithm," *IEEE Trans. Commun.*, vol. 55, pp. 1231–1241, Jun. 2007.

[7] J. J. Spilker and D. T. Magill, "The delay-lock discriminator-An optimum tracking device," *Proc. IRE*, vol. 49, pp. 1403–1416, Sep. 1961.

[8] A. Radovic and B. Aazhang, "Iterative algorithms for joint data detection and delay estimation for code division multiple access communication systems," in *Proc. 31st Annu. Conf. Communication, Control, Computing*, Monticello, IL, Sep. 1993, pp. 1–10.

[9] E. G. Ström, S. Parkvall, S. L. Miller, and B. E. Ottersten, "Propagation delay estimation in asynchronous direct-sequence code-division multiple access systems," *IEEE Trans. Commun.*, vol. 44, pp. 84–93, Jan. 1996.

[10] R. A. Iltis, "A DS-CDMA tracking mode receiver with joint channel/delay estimation and MMSE detection," *IEEE Trans. Commun.*, vol. 49, no. 10, pp. 1770–1779, Oct. 2001.

[11] A. Dempster, N. Laird, and D. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *J. Roy. Stat. Soc., Ser. B*, vol. 39, pp. 1–38, Jan. 1977.

[12] J. A. Fessler and A. O. Hero, "Complete-spaces and generalized EM algorithm," in *Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing (ICASSP)*, Minneapolis, MN, Apr. 1993, pp. IV/1–IV/4.

[13] H. V. Poor, "On parameter estimation in DS/SSMA formats," in *Proc. Int. Conf. Advances Communications Control Systems*, Baton Rouge, LA, Oct. 1988, pp. 98–109.

[14] L. B. Nelson and H. V. Poor, "Iterative multiuser receivers for CDMA channels: An EM-based approach," *IEEE Trans. Commun.*, vol. 44, pp. 1700–1710, Dec. 1996.

[15] A. Kocian and B. H. Fleury, "EM-based joint data detection and channel estimation of DS-CDMA signals," *IEEE Trans. Commun.*, vol. 51, pp. 1709–1720, Oct. 2003.

[16] C. N. Georghiades, "Optimum delay and sequence estimation from incomplete data," *IEEE Trans. Inf. Theory*, vol. 36, pp. 202–208, Jan. 1990.

[17] A. J. Viterbi, "Error bounds for convolutional codes and an asymptotically optimum decoding algorithm," *IEEE Trans. Inf. Theory*, vol. 13, pp. 260–269, Apr. 1967.

[18] I. Sharfer and A. O. Hero, "A maximum likelihood digital receiver using coordinate ascent and the discrete wavelet transform," *IEEE Trans. Signal Process.*, vol. 47, pp. 813–825, Mar. 1999.

[19] N. Noels *et al.*, "A theoretical framework for soft-information-based synchronization in iterative (turbo) receivers," *EURASIP J. Wireless Commun. Netw.*, no. 2, pp. 117–129, 2005.

[20] L. R. Bahl, J. Cocke, F. Jelinek, and J. Raviv, "Optimal decoding of linear codes for minimizing symbol error rate," *IEEE Trans. Inf. Theory*, vol. 20, pp. 284–287, Mar. 1974.

[21] R. A. Iltis, S. Kim, and A. Thomas, "EM-based acquisition of DS-CDMA waveforms for radiolocation," in *Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing (ICASSP)*, Salt Lake City, UT, May 2001, pp. 2229–2232.

[22] A. Gelfand and A. Smith, "Sampling-based approaches to calculating marginal densities," *J. Amer. Stat. Assoc.*, vol. 85, pp. 398–409, 1990.

[23] A. Doucet and X. Wang, "Monte Carlo methods for signal processing," *IEEE Signal Process. Mag.*, vol. 153, pp. 152–170, Nov. 2005.

[24] G. Wei and M. Tanner, "A Monte Carlo implementation of the EM algorithm and the poor man's data augmentation algorithms," *J. Amer. Stat. Assoc.*, vol. 85, no. 411, pp. 699–704, Sep. 1990.

[25] J. A. Fessler and A. O. Hero, "Space-alternating generalized expectation-maximization algorithm," *IEEE Trans. Signal Process.*, vol. 42, pp. 2664–2677, Oct. 1994.

[26] G. McLachlan and T. Krishnan, *The EM Algorithm and Extensions*. New York: Wiley, 1997.

[27] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, "Equation of state calculations by fast computing machines," *J. Chem. Phys.*, vol. 21, no. 6, pp. 1087–1092, Jun. 1953.

[28] W. K. Hastings, "Monte Carlo sampling methods using Markov chains and their applications," *Biometrica*, vol. 57, no. 1, pp. 97–109, Apr. 1970.

[29] S. Geman and D. Geman, "Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 6, no. 6, pp. 721–741, Nov. 1984.

[30] L. Tierney, "Markov chains for exploring posterior distributions," *Ann. Stat.*, vol. 22, no. 4, pp. 1701–1728, Dec. 1994.

[31] G. Markarian, B. Honorary, and P. Benachour, "Trellis decoding technique for binary-adder channel with M users and its application in LANs," *Proc. Inst. Electr. Eng.—Commun.*, vol. 144, no. 2, pp. 65–69, Apr. 1997.

[32] R. Chen, J. S. Liu, and X. Wang, "Convergence analysis and comparisons of Markov chain Monte Carlo algorithms in digital communications," *IEEE Trans. Signal Process.*, vol. 50, pp. 255–270, Feb. 2002.

[33] R. N. Bhattacharya and E. C. Waymire, *Stochastic Processes with Applications*. New York: Wiley, 1990.

[34] G. H. Golub and C. F. van Loan, *Matrix Computations*, 3rd ed. Baltimore, MD: The John Hopkins Univ. Press, 1996.

[35] A. Berman and R. J. Plemmons, *Nonnegative Matrices in the Mathematical Sciences*. Philadelphia, PA: SIAM, 1994.

[36] S. K. Sahu and G. O. Roberts, "On convergence of the EM algorithm and the Gibbs sampler," *Kluwer Stat. Comput.*, vol. 9, pp. 55–64, 1999.

[37] H. L. Van Trees, *Detection, Estimation and Modulation Theory*. New York: Wiley, 1968.

[38] A. N. D'Andrea, U. Mengali, and R. Reggiannini, "The modified Cramer-Rao bound and its application to synchronisation problems," *IEEE Trans. Commun.*, vol. 42, no. 2/3/4, pp. 1391–1399, Feb./Mar./Apr. 1994.

[39] M. Kiessling and S. A. Mujtaba, "Performance enhancements to the UMTS (W-CDMA) initial cell search algorithm," in *Proc. IEEE Int. Conf. Communications (ICC)*, New York City, May 2002, vol. 1, pp. 590–594.

[40] 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, pp. 276–287, May 1996.

[41] M. Moher, "An iterative multiuser decoder for near-capacity communications," *IEEE Trans. Commun.*, vol. 46, pp. 870–880, Jul. 1998.

[42] A. Frigessi, P. D. Stefano, C.-R. Hwang, and S.-J. Sheu, "Rates of the Gibbs sampler, the Metropolis algorithm and other single-site updating dynamics," *J. Roy. Stat. Soc., Ser. B*, vol. 55, no. 1, pp. 205–219, 1993.

Alexander Kocian (S'98–M'04) received the Dipl. Ing. degree (with distinction) from the Vienna University of Technology, Austria, in 1997 and the Ph.D. degree from Aalborg University (AAU), Denmark, in 2003, both in electrical engineering.

In 2001, AAU sponsored him for a six-month Visiting Research Scholarship at the Wireless Systems Laboratory, Georgia Institute of Technology, Atlanta. After serving the Digital Communications Laboratory at AAU as Assistant Professor for another two years, he joined the Department of Electronics Engineering of the Muscat branch in the Sultanate of Oman of Birla Institute of Technology (BIT), India, as a Reader/Associate Professor. Since 2008, he has been a Research Fellow with the Department of Electronics Engineering, University of Rome "Tor Vergata," Italy. He has consulted for Elektrobit, RTX Telecom, IMT, the Italian Ministry of University and Research (MIUR), and the European Space Agency (ESA). His research interests include iterative information processing in multiple-access communication systems, analog signal processing for satellite communication payloads, characterization of multiple-input multiple-output (MIMO) channels, and networking.

Erdal Panayırci (S'73–M'80–SM'91–F'03) received the Diploma Engineering degree in electrical engineering from the Istanbul Technical University, Istanbul, Turkey, and the Ph.D. degree in electrical engineering and system science from Michigan State University, East Lansing.

Until 1998, he has been with the Faculty of Electrical and Electronics Engineering at the Istanbul Technical University, where he was a Professor and Head of the Telecommunications Chair. Currently, he is Professor of Electrical Engineering and Head of the Electronics Engineering Department at Kadir Has University, Istanbul, Turkey. His recent research interests include communication theory, synchronization, advanced signal processing techniques and their applications to wireless communications, coded modulation, and interference cancellation with array processing. He published extensively in leading scientific journals and international conferences. He has coauthored the book *Principles of Integrated Maritime Surveillance Systems* (Boston, MA: Kluwer Academic, 2000). He spent two years in 1980 and 1981 with the Department of Computer Science, Michigan State University, as a Fulbright-Hays Fellow and a NATO Senior Scientist. Between 1990 and 1991, he was with the Center for Communications and Signal Processing, New Jersey Institute of Technology, Newark, as Visiting Professor, and from 1998 to 2000, he was Visiting Professor at the Department of Electrical Engineering, Texas A&M University, College Station. During 2008 and 2009, he was a Research Scholar at the Department of Electrical Engineering, Princeton University, Princeton, NJ. He has been the principal coordinator of a 6th and 7th Frame European project called NEWCOM (Network of Excellent on Wireless Communications) and WIMAGIC Strep project representing Kadir Has University.

Dr. Panayirci was an Editor for the IEEE TRANSACTIONS ON COMMUNICATIONS in the areas of Synchronizations and Equalizations from 1995 to 1999. He served as a Member of the IEEE Fellow Committee from 2005 to 2008. He was the Technical Program Chair of the IEEE International Conference on Communications (ICC) held in Istanbul, Turkey, in 2006. He is Technical Program Chair of the upcoming IEEE PIMRC to be held in Istanbul, Turkey, in 2010. Presently, he is head of the Turkish Scientific Commission on Signals and Systems of URSI (International Union of Radio Science).

H. Vincent Poor (S'72–M'77–SM'82–F'87) received the Ph.D. degree in electrical engineering and computer science from Princeton University, Princeton, NJ, in 1977.

From 1977 until 1990, he was on the faculty of the University of Illinois at Urbana-Champaign. Since 1990, he has been on the faculty at Princeton, where he is the Michael Henry Strater University Professor of Electrical Engineering and Dean of the School of Engineering and Applied Science. His research interests are in the areas of stochastic analysis, statistical signal processing, and their applications in wireless networks and related fields. Among his publications in these areas are the recent books *MIMO Wireless Communications* (Cambridge, U.K.: Cambridge Univ. Press, 2007) and *Quickest Detection* (Cambridge, U.K.: Cambridge Univ. Press, 2009).

Dr. Poor is a member of the National Academy of Engineering, a Fellow of the American Academy of Arts and Sciences, and an International Fellow of the Royal Academy of Engineering (U.K.). He is also a Fellow of the Institute of Mathematical Statistics, the Optical Society of America, and other organizations. In 1990, he served as President of the IEEE Information Theory Society, and from 2004 to 2007, he served as the Editor-in-Chief of the IEEE TRANSACTIONS ON INFORMATION THEORY. He was the recipient of the 2005 IEEE Education Medal. Recent recognition of his work includes the 2007 Technical Achievement Award of the IEEE Signal Processing Society, the 2009 Edwin Howard Armstrong Award of the IEEE Communications Society, the 2010 IET Ambrose Fleming Medal, and the 2011 IEEE Eric E. Sumner Award.

Marina Ruggieri (S'84–M'85–SM'94) was formerly with FACE-ITT and GTC-ITT, Roanoke, VA, in the High Frequency Division from 1985 to 1986 and was a Research and Teaching Assistant at the University of Roma Tor Vergata, Italy, from 1986 to 1991; and an Associate Professor in telecommunications at the University of L'Aquila, Italy, from 1991 to 1994 and at Tor Vergata, Italy, from 1994 to 2000. Since November 2000, she has been a Full Professor in telecommunications at Tor Vergata. Since 2003, she has directed a Master's in Advanced Satellite Communications and Navigation Systems at Tor Vergata. Her research focuses on space communications and navigation systems, integrated systems, mobile and multimedia networks, and ICT for biotechnology and energy. She is author of about 250 papers in international journals/transactions and proceedings of international conferences and eight book chapters and books.

Dr. Ruggieri was in the Technical-Scientific Committee of the Italian Space Agency (ASI) from 2004 to 2006. From 2007 to 2008, she was Vice-President of the ASI Technical-Scientific Committee. Since December 2007, she has been an Expert with the Italian Superior Council of Telecommunications. Since March 2010, she has been a member of the Committee of Experts for the Research Policy (CEPR) of the Ministry of University and Research (MIUR). Since January 2010, she has been a President of the IEEE Aerospace and Electronic Systems (AES) Society. Since December 2006, she has been a Vice-President of the AFCEA Rome Chapter. She is Director of CTIF-Italy, the Italian branch of the Center for Teleinfrastruktur (CTIF) in Aalborg, Denmark, opened at the University of Roma Tor Vergata. She is Editor of the IEEE TRANSACTIONS ON AEROSPACE AND ELECTRONIC SYSTEMS for Space Systems, Assistant Editor of the *IEEE Aerospace and Electronics Systems Magazine*. She was awarded the 1990 Piero Fanti International Prize, and she was nominated for the Harry M. Mimmo Award in 1996 and the Cristoforo Colombo Award in 2002. She received the 2009 Pisa Donna Award. She is a member of the AFCEA and IIN.