# SAGE Algorithms for Multipath Detection and Parameter Estimation in Asynchronous CDMA Systems

A. Logothetis* and C. Carlemalm†

**Abstract**

Algorithms to estimate blindly the parameters and input sequences for a multiuser asynchronous CDMA system are proposed. The communication system is modeled as a frequency selective fading channel with multipath propagation where the number of active propagation channels is unknown and highly time-varying. Estimates of input sequences and unknown parameters are jointly computed via the SAGE algorithm. The SAGE algorithm provides a unified approach in combining jointly multiuser detection and parameter estimation in an optimal way. Depending on the realizations of the SAGE algorithm well-known state and parameter estimation schemes of various complexities are optimally combined.

**Key-words:** asynchronous CDMA, SAGE, multiuser detection, parameter estimation, HMM, Viterbi, iterative decoding.

**EDICS: 3-CEQU (Channel Modeling, Estimation, and Equalization)**

## 1 Introduction

There are today several operational CDMA communication systems for both military and commercial applications including satellite networking, cellular mobile radio systems, and indoor wireless communications [1]. In a CDMA system, all users transmit simultaneously and at the same frequency and the transmitted signals occupy the entire system bandwidth. Code sequences are used to separate one user from another and it is assumed that the receiver has knowledge of these codes of some or all users. Demodulation, based on the reception of the transmitted signals of all users in the presence of additive noise, must occur to ascertain the information transmitted by each user. In most CDMA systems the users transmit independently, thereby causing the transmitted signals to

*Automatic Control, Department of Signals, Sensors and Systems, Royal Institute of Technology, SE-100 44 Stockholm, SWEDEN.

†Division of Signals and Systems, Uppsala Univeristy, SE-751 20 Uppsala, SWEDEN. Telephone: +46 18 4713068. Fax: +46 18 555096. Email: catharina.carlemalm@signal.uu.se. *Corresponding author.*

arrive asynchronously at the receiver. Many radio channels also exhibit severe multipath propagation which causes frequency selective fading [2, 3].

The interest in the problem of multiuser detection was motivated by the work of Verdú [4] where he devised both the minimum error probability detector and the maximum likelihood (ML) sequence detector for CDMA systems over Gaussian noise channels. Many suboptimal algorithms, computationally less intensive than their optimal counterparts, followed from the development of these optimal detectors, see *e.g.* [5, 6, 7, 8, 9]. In [10], ML sequence estimation with unknown random phase or fading parameters treated as missing data is derived. Similar to [10], the expectation maximization (EM) algorithm [11] is used in [12] for detection in a multiuser synchronous channel, rather than estimation of unknown parameters.

In [13], multiuser receivers iterate between the EM algorithm for amplitude estimation and multistage data detection. [14] uses the EM algorithm for amplitude estimation of direct sequence CDMA systems. The approach in [14] differs from the one in [13] in that the users' data are considered probabilistically as missing data. In [15], the EM algorithm is applied to the problem of ML sequence estimation when symbol timing information is not present. Most of the work done on timing acquisition for CDMA systems focuses on jointly estimating the necessary parameters for all users [13, 16]. These techniques involve solving a multidimensional optimization problem for a large number of parameters, which can be computationally intensive. In [17], a subspace-based scheme is derived for estimation of channel parameters in CDMA communication systems operating over channels with either single or multiple propagation paths. Subspace-based schemes for parameter estimation and data detection of asynchronous CDMA systems are also proposed in [18].

In this paper, we consider an asynchronous multipath and multiuser CDMA system. Here, the channel is modeled as an additive white Gaussian noise (AWGN) frequency selective multipath channel with a fixed number of users, but with a time-varying number of propagation channels. We assume that the input data belong to a finite symbol alphabet and that phase-shift keying (PSK) is used for modulation.

The approach in the current paper is a generalized extension of the CDMA joint data and channel estimation scheme presented in [13]. In [13], the number of multipath channels is fixed and known, an assumption not true in practice. Here, we use a SAGE based algorithm to jointly estimate the channel parameters, demodulate the transmitted messages, and estimate the number of active propagation paths. This paper is an extension of [19]. Similar work can be found in [20] and [21], where data and channel estimation in frequency non-selective synchronous channels is studied.

*Contributions:* In this paper, we propose three iterative, optimal in a maximum a posteriori (MAP) sense, off-line schemes that yield the joint data and channel estimates of an asynchronous multipath and multiuser CDMA system. The estimates are computed via the SAGE algorithm

[22], and depending on the choices of the hidden data spaces and the parameter index sets, three different algorithms are proposed. Algorithm I iteratively computes the MAP parameter estimates and at the same time computes estimates of the transmited data and detects the active propagation paths using a HMM smoother. Algorithms II and III compute the MAP channel-data estimates jointly. In particular, Algorithm II iterates between a channel parameter estimator and the Viterbi algorithm, which computes MAP sequence estimates of the data and the active propagation paths, while Algorithm III computes each of the unknown data symbols and detects the active propagation paths iteratively one at a time.

In computer simulation studies, we evaluate the performance of our detection and estimation methods. We compare the errors of the channel estimates to the derived CRLB. The BER in detecting the input sequence and the presence of multipaths is compared to results achieved by using a Viterbi algorithm when all other parameters are assumed to be known.

The outline of this paper is as follows. In Section 2, the system model is described and the problem and objectives are stated. Our suggested estimation and detection schemes are formulated in Sections 3 and 4. Section 5 then provides the performance analysis of the proposed algorithms. Finally, Section 6 contains some concluding remarks and potential directions for future research.

# 2 Problem Description

We begin with a description of the communication system and we outline our estimation and detection objectives and the proposed schemes.

## 2.1 CDMA System Model

In this paper, we assume that the multiuser and multipath CDMA system operates over AWGN frequency selective multipath channels [17, 18]. Let $K$ denote the known number of asynchronous users. Each user transmits $N$ data symbols belonging to the finite symbol alphabet $\{0, 1, \ldots \mathcal{M}-1\}$, where $\mathcal{M}$ is known. The input data are modulated using PSK. The $k$th user's modulated signal (for $k = 1, \ldots, K$) is given by

$$s_k(t) = \sqrt{2P_k}\exp\{j\phi_k\}\sum_{n=1}^{N}\left\{\exp\left\{j\frac{2\pi}{\mathcal{M}}m_{n,k}\right\}\sum_{l=0}^{L-1}\mathrm{rect}_{T_c}(t-lT_c-nT)d_k(l)\right\},\ -\infty < t < \infty, \tag{1}$$

where $m_{n,k}$ is the $n$th symbol from the $k$th user, $T$ is the symbol duration, $T_c = T/L$ is the chip duration, and $L$ is the number of chips per symbol. The symbols $m_{n,k} \in \{0, 1, \ldots \mathcal{M}-1\}$ are modeled as a finite-state independent and identically distributed (iid) equiprobable random variables. Here, $d_k(l), l = 0, \ldots L-1$, denotes the known code sequence of the $k$th user. $P_k$ is the transmitted power, $\phi_k$ is the carrier phase relative to the local oscillator, and $\mathrm{rect}_{T_c}(\cdot)$ is a rectangular pulse of duration

$T_c$, given by

$$\mathrm{rect}_{T_c}(t) = \begin{cases} 1, & \text{if } 0 \leq t \leq T_c \\ 0, & \text{otherwise} \end{cases} . \tag{2}$$

The modulated signals are subject to attenuation and propagation delays. The propagation delay, carrier phase, received power and channel response of the signal are assumed to be unknown parameters requiring estimation. The received signal is given by

$$r(t) = \sum_{k=1}^{K} r_k(t) + w(t), \ -\infty < t < \infty, \tag{3}$$

where $r_k(t)$ is the received, attenuated and delayed signal of the $k$th user and $w(t)$ is assumed to be white complex Gaussian noise with zero mean and a two-sided power spectral density $\sigma^2$. The signal $r_k(t)$, as illustrated in Fig. 2, is given by

$$r_k(t) = \sum_{p=1}^{S} \alpha_{k,p} \ \delta_{k,p}(t) \ s_k(t - \tau_{k,p}), \tag{4}$$

where $S$ is the maximum number of possible propagation paths, $\alpha_{k,p}$ is the attenuation factor, $\tau_{k,p}$ is the propagation delay of the $k$th user via the $p$th propagation path, and $\delta_{k,p}(t) \in \{0, 1\}$ denotes whether or not a path is active. We assume $\tau_{k,p}$ to be uniformly distributed within the interval $[0, T)$, which implies that the length of the channel is always lower than the processing gain. The number of transmission paths at any time instant is assumed to be unknown and time-varying. We use $S$ Markov chains as indicator functions to model the presence or non-existence of the possible paths.

The received continuous-time signal is converted into discrete-time by sampling of the output matched to the chip waveform, which is a rectangular pulse for PSK [17]. Because of time-delays in the system, each observation will contain at least the end of the previous symbol and the beginning of the current symbol for each user and path. Furthermore, we assume that $\delta_{k,p}(t)$ is constant during one sampling period. Extending the results in [17], it can be shown that the discrete-time vector, $y_n$, as depicted in Fig. 3, is given by

$$y_n = \sum_{k=1}^{K} \sum_{p=1}^{S} \xi_{k,p} \ \delta_{n,k,p} \left( \exp\left\{ j \frac{2\pi}{\mathcal{M}} m_{n-1,k} \right\} \rho_{k,p} + \exp\left\{ j \frac{2\pi}{\mathcal{M}} m_{n,k} \right\} \varrho_{k,p} \right) + w_n, \tag{5}$$

where

$$\rho_{k,p} = (1 - \eta_{k,p}) a_k^r(\nu_{k,p}) + \eta_{k,p} a_k^r(\nu_{k,p} + 1), \tag{6}$$

$$\varrho_{k,p} = (1 - \eta_{k,p}) a_k^l(\nu_{k,p}) + \eta_{k,p} a_k^l(\nu_{k,p} + 1), \tag{7}$$

and $y_n \in C^L$ is the $n$th vector output from the integrate-and-dump chip, $w_n \in C^L$ where each element is a zero mean white complex Gaussian noise sequence with variance $\sigma^2/T_c$ and

$$\xi_{k,p} = \alpha_{k,p} \sqrt{2P_k} \exp\{j\phi_k\}. \tag{8}$$

$\delta_{n,k,p} \in \{0,1\}$ is modeled as a two-state (on/off) Markov chain, which indicates whether the $p$th propagation path of the $k$th user is active at time instant $n$. The transition probability matrix of $\delta_{n,k,p}$ is denoted as $\Pi^{(k,p)} = \left[t_{ij}^{(k,p)}\right]$, where

$$t_{ij}^{(k,p)} = \Pr(\delta_{n+1,k,p} = j | \delta_{n,k,p} = i), \qquad \forall i,j \in \{0,1\}, \tag{9}$$

$$0 \leq t_{ij}^{(k,p)} \leq 1, \qquad \forall i,j \in \{0,1\}, \tag{10}$$

and

$$t_{i0}^{(k,p)} + t_{i1}^{(k,p)} = 1, \qquad \forall i \in \{0,1\}. \tag{11}$$

The time delay $\tau_{k,p}$ is separated into an integer part $\nu_{k,p} \in \{0,1,\ldots,L-1\}$ and a fractional part $\eta_{k,p} \in [0,1)$, given by

$$\frac{\tau_{k,p}}{T_c}\mathrm{mod}(L) = \nu_{k,p} + \eta_{k,p}, \tag{12}$$

where mod denotes the modulus operator. Finally,

$$a_k^r(\nu) \triangleq \begin{cases} [d_k(L-\nu)\cdots d_k(L-1)\ 0_{1\times(L-\nu)}]', & \text{for } \nu \in \{1,\ldots,L\} \\ 0_{L\times 1}, & \text{for } \nu = 0 \end{cases}, \tag{13}$$

$$a_k^l(\nu) \triangleq \begin{cases} [0_{1\times\nu}\ d_k(0)\cdots d_k(L-\nu-1)]', & \text{for } \nu \in \{0,\ldots,L-1\} \\ 0_{L\times 1}, & \text{for } \nu = L \end{cases}. \tag{14}$$

*Prior Distributions:* Let $\theta = (\eta_{k,p}, \nu_{k,p}, \xi_{k,p}, t_{ii}^{(k,p)}; k = 1,\ldots,K,\ p = 1,\ldots,S,\ i = 0,1)$ denote the unknown parameter vector and $f(\theta)$ denote the prior distribution of $\theta$. Note that if $t_{00}^{(k,p)}$ and $t_{11}^{(k,p)}$ are estimated then $t_{01}^{(k,p)}$ and $t_{10}^{(k,p)}$ are estimated using Eq. (11). We assume the parameters are mutually independent, thus

$$f(\theta) = \prod_{k=1}^{K}\prod_{p=1}^{S} f(\eta_{k,p})f(\nu_{k,p})f(\xi_{k,p})f(t_{00}^{(k,p)})f(t_{11}^{(k,p)}). \tag{15}$$

We assume uniform priors on the transition probabilities and on time delays $\tau_{k,p}$, and consequently on $\eta_{k,p}$ and $\nu_{k,p}$. We assume $\xi_{k,p}$ to be Gaussian distributed with known mean $\mu_{k,p}$ and variance $\sigma_{k,p}^2$. When $\mu_{k,p}$ is zero this implies that the fading on path $p$ for the $k$th user is Rayleigh distributed. Alternatively, when $\mu_{k,p}$ is nonzero, then we consider Rician fading channels.

*Notation:* Using $N$ samples, let the sequence of measurements $(y_1,\ldots,y_N)$ be denoted as $Y$. Let $Y_n \triangleq (y_1,\ldots,y_n)$ and $Y_n^N \triangleq (y_n,\ldots,y_N)$. Let $M_k$ and $\Delta_{k,p}$ denote the symbol sequence $(m_{1,k},\ldots,m_{N,k})$ and the indicator functions $(\delta_{1,k,p},\ldots,\delta_{N,k,p})$ of the $k$th user and $p$th propagation path, respectively. Finally, let $m_n \triangleq \{m_{n,k}; k = 1,\ldots,K\}$, $\delta_n \triangleq \{\delta_{n,k,p}; k = 1,\ldots,K,\ p = 1,\ldots,S\}$, $M \triangleq \{M_k; k = 1,\ldots,K\}$ and $\Delta \triangleq \{\Delta_{k,p}; k = 1,\ldots,K,\ p = 1,\ldots,S\}$.

## 2.2 Objectives and Proposed Schemes

Given the observed data $Y$, it is desired to perform the following.

**1.** *Parameter Estimation:*

- Estimate the complex constants $\xi_{k,p}$, defined in (8), for each user $k$ and each path $p$.
- Estimate the time-delays $\tau_{k,p}$ for each user $k$ and each path $p$.
- Estimate the transition probabilities $\Pi^{(k,p)}$ of the $S$ different propagation paths and $K$ users.

**2.** *Estimation of Input Sequence:* Estimate the transmitted input sequence $m_{n,k}$, for $n = 1, \ldots, N$ and $k = 1, \ldots, K$.

**3.** *Detection of Active Propagation Paths:* Estimate the active propagation paths $\delta_{n,k,p}$ for $n = 1, 2, \ldots, N$, $k = 1, \ldots, K$ and $p = 1, 2, \ldots, S$.

In this paper, we seek to compute the MAP data and channel estimates as follows

$$\theta^{\text{ MAP}} = \arg\max_{\theta} f(\theta|Y), \tag{16}$$

$$(\theta, M, \Delta)^{\text{ MAP}} = \arg\max_{(\theta,M,\Delta)} f(\theta, M, \Delta|Y). \tag{17}$$

We propose to use the SAGE algorithm [22] to yield the desired data and channel estimates as given in Eqs. (16)–(17). The SAGE algorithm is suited to problems where subsets of the parameter vector can be updated sequentially. The SAGE algorithm is to be preferred to the EM algorithm, since convergence of the likelihood function is significantly faster, and the maximization step is often simplified.

Associated with each small group of parameters is a *hidden data* space, which together with the observations form the *complete data* space. The simplicity of the SAGE algorithm relies on the maximization of an objective function of the complete data, which is computationally more tractable than maximizing the likelihood function directly.

We now outline the SAGE algorithm. Let $\phi_\Upsilon$ denote the parameter components indexed by the set $\Upsilon \subset \Omega = \{1, 2, \ldots, |\phi|\}$, where $|\phi|$ denotes the cardinality of parameter vector $\phi$, and let $\bar{\Upsilon}$ denote the compliment of $\Upsilon$, such that $\Upsilon \cup \bar{\Upsilon} = \Omega$ and $\Upsilon \cap \bar{\Upsilon} = \emptyset$. Starting from an initial parameter estimate $\phi^{(0)}$, the SAGE algorithm yields on the $l$th iteration (for $l = 0, 1, \ldots$) the parameter estimate $\phi^{(l)}$ by performing the steps outlined in Fig. 1.

In [22], it is proven that the sequence of estimates, $\left\{\phi^{(l)}\right\}$, generated by the SAGE algorithm, monotonically increases the penalized likelihood objective function. Asymptotic convergence rates are derived, and sufficient conditions for monotonic convergence in norm are determined. See also [12].

In this paper, we propose three different detection schemes according to the choice of the index sets $\Upsilon$ and the desired channel and data estimates given in Eqs. (16)–(17). In Section 3, we propose a scheme (called *Algorithm I*) for computing Eq. (16). In Section 4, we consider the problem formulation stated in (17) and two schemes (called *Algorithm II* and *Algorithm III*) are proposed based on two different choices of index sets, $\Upsilon$.

The SAGE algorithm provides a unified approach in combining jointly multiuser detection and parameter estimation in an optimal way. Depending on the realizations of the SAGE algorithm well-known state and parameter estimation schemes, such as the HMM smoother, the Viterbi algorithm, the Yule-Walker and Baum-Welch re-estimation formulae, are optimally combined.

# 3 Parameter Estimation via SAGE: Algorithm I

In this section, we consider the problem formulation stated in Eq. (16), that is, we seek to compute the MAP parameter estimates as follows

$$\theta^{\mathrm{MAP}} = \arg\max_{\theta} f(\theta|Y). \tag{18}$$

As a by-product of the SAGE algorithm we compute MAP estimates of $M, \Delta$ as follows

$$\begin{aligned}
\hat{\delta}_{n,k,p}^{\mathrm{MAP}} &= \arg\max_{\delta_{n,k,p}\in\{0,1\}} \Pr\left(\delta_{n,k,p} \mid \theta^{\mathrm{MAP}}, Y\right), && (19)\\
\hat{m}_{n,k}^{\mathrm{MAP}} &= \arg\max_{m_{n,k}\in\{0,\ldots,\mathcal{M}-1\}} \Pr\left(m_{n,k} \mid \theta^{\mathrm{MAP}}, Y\right), && (20)
\end{aligned}$$

for $n = 1, \ldots N,\ p = 1, \ldots S$ and $k = 1 \ldots K$.

Using the notation in Subsection 2.1, the parameter vector is chosen as $\phi = \theta$ and the hidden data space as $X^{\Upsilon} = (M, \Delta)$. The index sets are chosen from a partition $\{\Upsilon_1, \ldots, \Upsilon_4\}$ such that

$$\begin{aligned}
\phi_{\Upsilon_1} &= \left(t_{ii}^{(k,p)};\ k = 1, \ldots, K,\ p = 1, \ldots, S,\ i = 0, 1\right), && (21)\\
\phi_{\Upsilon_2} &= (\eta_{k,p};\ k = 1, 2, \ldots, K,\ p = 1, 2, \ldots, S), && (22)\\
\phi_{\Upsilon_3} &= (\nu_{k,p};\ k = 1, 2, \ldots, K,\ p = 1, 2, \ldots, S), && (23)\\
\phi_{\Upsilon_4} &= (\xi_{k,p};\ k = 1, 2, \ldots, K,\ p = 1, 2, \ldots, S). && (24)
\end{aligned}$$

We now present the details of the SAGE algorithm for computing MAP parameter estimates for our asynchronous CDMA system defined in Subsection 2.1.

**Expectation Step:**

Evaluate

$$\mathcal{Q}\left(\theta; \theta^{(l)}\right) \triangleq \mathrm{E}\left\{\ln f(Y, M, \Delta, \theta) | Y, \theta^{(l)}\right\}, \tag{25}$$

where

$$\ln f(Y, M, \Delta, \theta) = -N \ln\left(\frac{\pi\sigma^2}{T_c}\right) \tag{26}$$
$$+ \sum_{k=1}^{K}\sum_{p=1}^{S}\left(\ln f(\eta_{k,p}) + \ln f(\nu_{k,p}) + \ln f(t_{00}^{(k,p)}) + \ln f(t_{11}^{(k,p)}) - \ln(\pi\sigma_{k,p}^2) - \frac{|\xi_{k,p}-\mu_{k,p}|^2}{\sigma_{k,p}^2}\right)$$
$$+ \sum_{n=2}^{N}\sum_{k=1}^{K}\sum_{p=1}^{S}\ln f(\delta_{n,k,p}|\delta_{n-1,k,p}) + \sum_{k=1}^{K}\sum_{p=1}^{S}\ln f(\delta_{1,k,p}) + \sum_{n=1}^{N}\sum_{k=1}^{K}\ln f(m_{n,k})$$
$$- \sum_{n=1}^{N}\frac{T_c}{\sigma^2}\left|y_n - \sum_{k=1}^{K}\sum_{p=1}^{S}\xi_{k,p}\delta_{n,k,p}\exp\left\{j\frac{2\pi}{\mathcal{M}}m_{n-1,k}\right\}\rho_{k,p} - \sum_{k=1}^{K}\sum_{p=1}^{S}\xi_{k,p}\delta_{n,k,p}\exp\left\{j\frac{2\pi}{\mathcal{M}}m_{n,k}\right\}\varrho_{k,p}\right|^2.$$

This step requires the evaluation of the following probabilities

$$\gamma_n^{(l)}(i_m, j_m, i_\delta, j_\delta) \triangleq \Pr\left(m_{n-1} = i_m, m_n = j_m, \delta_{n-1} = i_\delta, \delta_n = j_\delta|Y, \theta^{(l)}\right), \tag{27}$$

where $i_m \triangleq (i_{m,1}, i_{m,2}, \ldots, i_{m,K}), j_m \triangleq (j_{m,1}, j_{m,2}, \ldots, j_{m,K}) \in \{0, 1, \ldots, \mathcal{M}-1\}^K$ and $i_\delta \triangleq (i_{\delta,1,1}, \ldots, i_{\delta,K,S}), j_\delta \triangleq (j_{\delta,1,1}, \ldots, j_{\delta,K,S}) \in \{0,1\}^{KS}$. Thus,

$$\mathcal{Q}\left(\theta; \theta^{(l)}\right) \triangleq \sum_{i_m, j_m}\sum_{i_\delta, j_\delta}\sum_{n=1}^{N}\gamma_n^{(l)}(i_m, j_m, i_\delta, j_\delta)\left\{-\ln\left(\frac{\pi\sigma^2}{T_c}\right)\right. \tag{28}$$
$$+ \sum_{k=1}^{K}\sum_{p=1}^{S}\frac{1}{N}\left(\ln f(\eta_{k,p}) + \ln f(\nu_{k,p}) + \ln f(t_{00}^{(k,p)}) + \ln f(t_{11}^{(k,p)}) - \ln(\pi\sigma_{k,p}^2) - \frac{|\xi_{k,p}-\mu_{k,p}|^2}{\sigma_{k,p}^2}\right)$$
$$+ \sum_{k=1}^{K}\sum_{p=1}^{S}(1 - \mathrm{I}\{n-1\})\ln t_{i_{\delta,k,p}j_{\delta,k,p}}^{(k,p)} + \sum_{k=1}^{K}\sum_{p=1}^{S}\mathrm{I}\{n-1\}\ln f(\delta_{1,k,p} = j_{\delta,k,p}) + \sum_{k=1}^{K}\ln f(m_{n,k} = j_{m,k})$$
$$\left.-\frac{T_c}{\sigma^2}\left|y_n - \sum_{k=1}^{K}\sum_{p=1}^{S}\xi_{k,p}j_{\delta,k,p}\exp\left\{j\frac{2\pi}{\mathcal{M}}i_{m,k}\right\}\rho_{k,p} - \sum_{k=1}^{K}\sum_{p=1}^{S}\xi_{k,p}j_{\delta,k,p}\exp\left\{j\frac{2\pi}{\mathcal{M}}j_{m,k}\right\}\varrho_{k,p}\right|^2\right\}.$$

Here, $\mathrm{I}\{\cdot\}$ is the indicator function, such that

$$\mathrm{I}\{x\} = \begin{cases} 1, & \text{if } x = 0, \\ 0, & \text{otherwise} \end{cases}. \tag{29}$$

The probabilities in (27) are efficiently computed via the forward backward recursions of an HMM smoother. Following [23], we first define the forward $\alpha$ and the backward $\beta$ unnormalized densities, as follows

$$\begin{aligned}\alpha_n^{(l)}(i_m, j_m, i_\delta, j_\delta) &= f\left(m_{n-1} = i_m, m_n = j_m, \delta_{n-1} = i_\delta, \delta_n = j_\delta, Y_n|\theta^{(l)}\right), \\ \beta_n^{(l)}(i_m, j_m, i_\delta, j_\delta) &= f\left(Y_{n+1}^N|m_{n-1} = i_m, m_n = j_m, \delta_{n-1} = i_\delta, \delta_n = j_\delta, Y_n, \theta^{(l)}\right).\end{aligned} \tag{30}$$

Let

$$i_m \triangleq \left[\begin{array}{cccc} i_{m,1} & i_{m,2} & \cdots & i_{m,K}\end{array}\right]', \tag{31}$$

and let similar notations hold for $j_m$ and $l_m$. Furthermore, let

$$i_\delta \triangleq \left[\begin{array}{cccccccc} i_{\delta,1,1} & i_{\delta,1,2} & \cdots & i_{\delta,1,S} & i_{\delta,2,1} & \cdots & i_{\delta,K,S} \end{array}\right]'. \tag{32}$$

Similar notation holds for $j_\delta$ and $l_\delta$. Define the following quantity

$$b_n^{(l)}(i_m, j_m, j_\delta) \triangleq f(y_n|m_{n-1} = i_m, m_n = j_m, \delta_n = j_\delta) = \left(\frac{\pi\sigma^2}{T_c}\right)^{-1} \times \tag{33}$$
$$\exp\left\{-\frac{T_c}{\sigma^2}\left|y_n - \sum_{k=1}^{K}\sum_{p=1}^{S}\xi_{k,p}j_{\delta,k,p}\exp\left\{j\frac{2\pi}{\mathcal{M}}i_{m,k}\right\}\rho_{k,p} - \sum_{k=1}^{K}\sum_{p=1}^{S}\xi_{k,p}j_{\delta,k,p}\exp\left\{j\frac{2\pi}{\mathcal{M}}j_{m,k}\right\}\varrho_{k,p}\right|^2\right\}.$$

Using similar arguments as in [23], we show that the following recursions hold for the updating of the unnormalized densities

$$\begin{aligned}
\alpha_{n+1}^{(l)}(j_m, l_m, j_\delta, l_\delta) &= \sum_{i_m}\sum_{i_\delta}\alpha_n^{(l)}(i_m, j_m, i_\delta, j_\delta)\Pr(m_{n+1} = l_m)\Pr(\delta_{n+1} = l_\delta|\delta_n = j_\delta)b_{n+1}^{(l)}(j_m, l_m, l_\delta), \\
&\quad \text{for } n = 1, \ldots, N-1, && (34) \\
\beta_n^{(l)}(i_m, j_m, i_\delta, j_\delta) &= \sum_{l_m}\sum_{l_\delta}\beta_{n+1}^{(l)}(j_m, l_m, j_\delta, l_\delta)\Pr(m_{n+1} = l_m)\Pr(\delta_{n+1} = l_\delta|\delta_n = j_\delta)b_{n+1}^{(l)}(j_m, l_m, l_\delta), \\
&\quad \text{for } n = N-1, \ldots, 1 && (35)
\end{aligned}$$

for $i_m, j_m, l_m \in \{0, 1, \ldots, \mathcal{M}-1\}^K$, $i_\delta, j_\delta, l_\delta \in \{0, 1\}^{KS}$, where

$$\begin{aligned}
\Pr(m_{n+1} = l_m) &= \mathcal{M}^{-K}, && (36) \\
\Pr(\delta_{n+1} = l_\delta|\delta_n = j_\delta) &= \prod_{k=1}^{K}\prod_{p=1}^{S} t_{j_{\delta,k,p}l_{\delta,k,p}}^{(k,p)}. && (37)
\end{aligned}$$

The forward and backward densities are initialized as follows

$$\begin{aligned}
\alpha_1^{(l)}(i_m, j_m, i_\delta, j_\delta) &= b_1^{(l)}(i_m, j_m, j_\delta), \text{ for } i_m, j_m \in \{0, 1, \ldots, \mathcal{M}-1\}^K, i_\delta, j_\delta \in \{0, 1\}^{KS}, \\
\beta_N^{(l)}(i_m, j_m, i_\delta, j_\delta) &= 1, \text{ for } i_m, j_m \in \{0, 1, \ldots, \mathcal{M}-1\}^K, i_\delta, j_\delta \in \{0, 1\}^{KS}.
\end{aligned}$$

Finally, the state probabilities $\gamma_n^{(l)}(\cdot)$ are given by

$$\gamma_n^{(l)}(i_m, j_m, i_\delta, j_\delta) = \frac{\alpha_n(i_m, j_m, i_\delta, j_\delta)\beta_n(i_m, j_m, i_\delta, j_\delta)}{\sum_{i_m, j_m}\sum_{i_\delta, j_\delta}\alpha_n(i_m, j_m, i_\delta, j_\delta)\beta_n(i_m, j_m, i_\delta, j_\delta)}, \tag{38}$$

for $n = 1, \ldots, N$ and $i_m, j_m \in \{0, 1, \ldots, \mathcal{M}-1\}^K$, $i_\delta, j_\delta \in \{0, 1\}^{KS}$.

This completes the Expectation step of the algorithm given in Fig. 1.

**Maximization Steps:**

Instead of maximizing (25) directly as a function of $\theta$, the SAGE algorithm updates subsets of parameters sequentially, keeping the other parameters fixed at their previous values. Depending on the choice of the parameter index set, $\Upsilon$, the parameters are updated as follows.

*Transition Probability Matrix Update:* Here, the index set $\Upsilon_1$ is studied. Given the constraints $0 \leq t_{ij}^{(k,p)} \leq 1$, $\forall i,j \in \{0,1\}$ and $\sum_{j=0}^{1} t_{ij}^{(k,p)} = 1$, for each $i \in \{0,1\}$, compute

$$\left(t_{ii}^{(k,p)}; i=0,1\right)^{(l+1)} = \arg \max_{\left(t_{ii}^{(k,p)}; i=0,1\right)} \mathcal{Q}^{\Upsilon_1}\left(\theta_{\Upsilon_1}; \theta^{(l)}\right). \tag{39}$$

Using standard HMM theory [23], it turns out that the Markov chain transition probability estimates on the $(l+1)$th iteration are given by

$$\left(t_{ii}^{(k,p)}\right)^{(l+1)} = \frac{\sum_{n=2}^{N} \Pr\left(\delta_{n,k,p} = i, \delta_{n-1,k,p} = i | Y, \theta^{(l)}\right)}{\sum_{n=1}^{N-1} \Pr\left(\delta_{n,k,p} = i | Y, \theta^{(l)}\right)}, \text{ for } i \in \{0,1\} \tag{40}$$

for $k = 1, 2, \ldots, K$ and $p = 1, 2, \ldots, S$. Where

$$\Pr\left(\delta_{n,k,p} = i, \delta_{n-1,k,p} = i | Y, \theta^{(l)}\right) = \sum_{i_m, j_m} \sum_{i_\delta \neq i, j_\delta \neq i} \gamma_n^{(l)}(i_m, j_m, i_\delta, j_\delta). \tag{41}$$

*Update of the Fractional Part of the Time Delay:* Here, the index set $\Upsilon_2$ is studied. Compute $\eta_{k,p}^{(l+1)}$ for $p \in \{1, 2, \ldots S\}$, $k \in \{1, \ldots K\}$, as follows

$$(\eta_{k,p}; k = 1, \ldots, K,\ p = 1, \ldots, S)^{(l+1)} = \arg \max_{(\eta_{k,p}; k=1,\ldots,K,\ p=1,\ldots,S)} \mathcal{Q}^{\Upsilon_2}\left(\phi_{\Upsilon_2}; \phi^{(l)}\right). \tag{42}$$

This is solved by differentiating $\mathcal{Q}^{\Upsilon_2}\left(\phi_{\Upsilon_2}; \phi^{(l)}\right)$ with respect to $\eta_{k,p}$ $\forall k \in \{1, \ldots, K\}, p \in \{1, \ldots, S\}$ and setting each equation equal to zero. The resulting $KS$ simultaneous equations are solved according to the following:

Let $e^{(l+1)}$ be a $K \times S$ matrix such that the $(k', p')$th element is $e_{k',p'}^{(l+1)}$ and

$$e_{k',p'}^{(l+1)} = \mathrm{E}\left\{\varepsilon_{k',p'}^{(l)} \middle| Y, \theta^{(l)}\right\}, \tag{43}$$

where

$$\begin{aligned} \varepsilon_{k',p'}^{(l)} = \sum_{n=1}^{N} \mathrm{Re}\Bigg\{ (c_{n,k',p'}^{(l)})^H \Bigg( y_n &- \sum_{k=1}^{K} \sum_{p=1}^{S} \xi_{k,p}^{(l)} \delta_{n,k,p} \exp\left\{j \frac{2\pi}{\mathcal{M}} m_{n-1,k}\right\} a^r(\nu_{k,p}^{(l)}) \\ &- \sum_{k=1}^{K} \sum_{p=1}^{S} \xi_{k,p}^{(l)} \delta_{n,k,p} \exp\left\{j \frac{2\pi}{\mathcal{M}} m_{n,k}\right\} a^l(\nu_{k,p}^{(l)}) \Bigg) \Bigg\}, \end{aligned} \tag{44}$$

where $(\cdot)^H$ denotes the conjugate transpose operator, and

$$\begin{aligned} c_{n,k,p}^{(l)} = \delta_{n,k,p} \xi_{k,p}^{(l)} \exp\left\{j \frac{2\pi}{\mathcal{M}} m_{n-1,k}\right\} \left[-a^r(\nu_{k,p}^{(l)}) + a^r(\nu_{k,p}^{(l)} + 1)\right] \\ + \delta_{n,k,p} \xi_{k,p}^{(l)} \exp\left\{j \frac{2\pi}{\mathcal{M}} m_{n,k}\right\} \left[-a^l(\nu_{k,p}^{(l)}) + a^l(\nu_{k,p}^{(l)} + 1)\right]. \end{aligned} \tag{45}$$

Let $W^{(l+1)}$ be a $KS \times KS$ matrix, given by

$$W^{(l+1)} = \mathrm{E}\left\{\sum_{n=1}^{N} \left|W_n^{(l+1)}\right|^2 \middle| Y, \theta^{(l)}\right\}, \tag{46}$$

where the $(K(p-1)+k)$th column of the $L \times KS$ matrix $W_n^{(l+1)}$ is given by $c_{n,k,p}^{(l)}$. The fractional part of the time-delays are given by the following modified Yule-Walker equation

$$\mathrm{vec}\left(\eta^{(l+1)}\right) = \left(W^{(l+1)}\right)^{-1} \mathrm{vec}\left(e^{(l+1)}\right), \tag{47}$$

where $\eta^{(l+1)}$ is a $K \times S$ matrix, and the $(k,p)$th element is equal to $\eta_{k,p}^{(l+1)}$.

*Update of the Integer Part of the Time Delay:* Here, the index set $\Upsilon_3$ is studied. Compute $(\nu_{k,p}; k = 1, \dots K, p = 1, \dots S)^{(l+1)}$ as follows

$$(\nu_{k,p}; k = 1, \dots K,\ p = 1, \dots S)^{(l+1)} = \arg \max_{(\nu_{k,p}; k=1,\dots K,\ p=1,\dots S)} \mathcal{Q}^{\Upsilon_3}\left(\phi_{\Upsilon_3}; \phi^{(l)}\right). \tag{48}$$

The joint maximization of $\mathcal{Q}^{\Upsilon_3}\left(\phi_{\Upsilon_3}; \phi^{(l)}\right)$ with respect to $(\nu_{k,p}; k = 1, \dots K, p = 1, \dots S)$ is a discrete combinatorial optimization problem with computational cost $O(L^{KS})$.

*Update of the Complex Constant $\xi$ (which is a known function of the unknown parameters carrier phase, transmitted power, channel response):* Here, the index set $\Upsilon_4$ is studied. Compute $(\xi_{k,p}; k = 1 \dots K, p = 1, \dots S)^{(l+1)}$ as follows

$$(\xi_{k,p}; k = 1, \dots K,\ p = 1, \dots S)^{(l+1)} = \arg \max_{(\xi_{k,p}; k=1,\dots K,\ p=1,\dots S)} \mathcal{Q}^{\Upsilon_4}\left(\phi_{\Upsilon_4}; \phi^{(l)}\right). \tag{49}$$

The joint maximization of $\xi_{k,p}$, $\forall k \in 1, \dots, K$ and $\forall p \in \{1, \dots, S\}$ is computed by solving a $K \times S$ linear system of equations derived from taking the derivative of $\mathcal{Q}^{\Upsilon_4}\left(\phi_{\Upsilon_4}; \phi^{(l)}\right)$ with respect to each $\xi_{k,p}$ and setting the results equal to zero.

Let $\xi^{(l+1)}$ denote the $K \times S$ matrix, such that the $(k,p)$th element of $\xi^{(l+1)}$ is equal to $\xi_{k,p}^{(l+1)}$. $\xi^{(l+1)}$ is given by the following modified Yule-Walker equation

$$\mathrm{vec}\left(\xi^{(l+1)}\right) = \left(R^{(l+1)}\right)^{-1} \mathrm{vec}\left(h^{(l+1)}\right), \tag{50}$$

where $h^{(l+1)}$ is a $K \times S$ matrix, and its $(k,p)$th element is given by

$$\begin{aligned} h_{k,p}^{(l+1)} &= \frac{\mu_{k,p}}{\sigma_{k,p}^2} + \frac{T_c}{\sigma^2} \times \\ &\sum_{n=1}^{N} \mathrm{E}\left\{ \left(\delta_{n,k,p}\rho_{k,p}^{(l)}\right)^H y_n \exp\left\{-j\frac{2\pi}{\mathcal{M}} m_{n-1,k}\right\} + \left(\delta_{n,k,p}\varrho_{k,p}^{(l)}\right)^H y_n \exp\left\{-j\frac{2\pi}{\mathcal{M}} m_{n,k}\right\} \middle| Y, \theta^{(l)} \right\}, \end{aligned} \tag{51}$$

and $R^{(l+1)}$ is an $KS \times KS$ matrix, given by

$$R^{(l+1)} = \frac{T_c}{\sigma^2} E\left\{ \sum_{n=1}^{N} \left| R_n^{(l)} \right|^2 \middle| Y, \theta^{(l)} \right\} + \Sigma, \tag{52}$$

where $\Sigma$ is a $KS \times KS$ diagonal matrix given by

$$\Sigma = \mathrm{diag}(1/\sigma_{1,1}^2, 1/\sigma_{2,1}^2, \dots, 1/\sigma_{K,1}^2, 1/\sigma_{1,2}^2, \dots, 1/\sigma_{K,S}^2). \tag{53}$$

$R_n^{(l)}$ has dimensions $L \times KS$, where the $(K(p-1)+k)$th column is given by

$$\delta_{n,k,p}\rho_{k,p}^{(l)} \exp\left\{j\frac{2\pi}{\mathcal{M}} m_{n-1,k}\right\} + \delta_{n,k,p}\varrho_{k,p}^{(l)} \exp\left\{j\frac{2\pi}{\mathcal{M}} m_{n,k}\right\}. \tag{54}$$

# 4 Joint Data and Channel Estimation via SAGE: Algorithms II and III

In this section, we consider the problem formulation stated in Eq. (17), that is, we seek to compute the joint MAP data and channel estimates as follows

$$(\theta, M, \Delta)^{\text{MAP}} = \arg \max_{(\theta, M, \Delta)} f(\theta, M, \Delta | Y). \tag{55}$$

Using the notation in Section 2.1, the parameter vector is chosen as $\phi = (M, \Delta, \theta)$ and the hidden data space as $X^{\Upsilon} = \phi_{\Upsilon}$. Consequently, the proposed SAGE algorithms reduce to coordinate descent methods. Depending on the choice of the partition set we propose two joint data and channel parameter estimation schemes. In the first scheme called Algorithm II, presented in Section 4.1, the data estimation and the active path detection is performed jointly via a Viterbi algorithm. The second data estimation problem, presented in Section 4.2, is performed iteratively following a similar approach to [13]. We call this Algorithm III.

## 4.1 Algorithm II

In this subsection, the index sets are chosen from a partition $\{\Upsilon_1, \ldots, \Upsilon_5\}$ such that

$$\phi_{\Upsilon_1} = \left(t_{ii}^{(k,p)};\ k = 1, 2, \ldots, K,\ p = 1, 2, \ldots, S,\ i = 0, 1\right) \tag{56}$$

$$\phi_{\Upsilon_2} = (\eta_{k,p};\ k = 1, 2, \ldots, K,\ p = 1, 2, \ldots, S) \tag{57}$$

$$\phi_{\Upsilon_3} = (\nu_{k,p};\ k = 1, 2, \ldots, K,\ p = 1, 2, \ldots, S) \tag{58}$$

$$\phi_{\Upsilon_4} = (\xi_{k,p};\ k = 1, 2, \ldots, K,\ p = 1, 2, \ldots, S) \tag{59}$$

$$\phi_{\Upsilon_5} = (M, \Delta) \tag{60}$$

The key idea is to iterate between channel (transition probabilities, time delay, channel amplitude) estimation and data/active channel detection assuming all other parameters known and fixed to their previous best estimates. The details of the SAGE algorithm for computing the joint MAP data and channel estimates for our asynchronous CDMA system is given below.

**Expectation Step:**

Evaluate for the $i$th, $i = 1, \ldots, 5$, index set:

$$\mathcal{Q}^{\Upsilon_i}\left(\phi_{\Upsilon_i}; \phi^{(l)}\right) \triangleq \ln f(Y, \phi_{\Upsilon_i}, \phi_{\bar{\Upsilon}_i}). \tag{61}$$

**Maximization Steps:**

*Transition Probability Matrix Update:* Given the constraints $0 \le t_{ij}^{(k,p)} \le 1$ and $\sum_{j=0}^{1} t_{ij}^{(k,p)} = 1$ for each $i, j \in \{0, 1\}$, compute

$$\left(t_{ii}^{(k,p)}\right)^{(l+1)} = \arg \max_{t_{ii}^{(k,p)}} \mathcal{Q}^{\Upsilon_1}\left(\phi; \phi^{(l)}\right). \tag{62}$$

The Markov chain transition probabilities are then given by

$$\left(t_{ii}^{(k,p)}\right)^{(l+1)} = \frac{\sum_{n=2}^{N} \mathrm{I}\left\{\delta_{n,k,p}^{(l)} - i\right\} \mathrm{I}\left\{\delta_{n-1,k,p}^{(l)} - i\right\}}{\sum_{n=1}^{N-1} I\left\{\delta_{n,k,p}^{(l)} - i\right\}}, \quad \text{for } i \in \{0,1\}, \tag{63}$$

for $k = 1, 2, \ldots, K$ and $p = 1, 2, \ldots, S$.

*Update of the Fractional Part of the Time Delay:* Compute $\eta_{k,p}^{(l+1)}$ for $p \in \{1, 2, \ldots S\},\ k \in \{1, \ldots K\}$, as follows

$$(\eta_{k,p}; k = 1, \ldots, K,\ p = 1, \ldots, S)^{(l+1)} = \arg \max_{(\eta_{k,p}; k=1,\ldots,K,\ p=1,\ldots,S)} \mathcal{Q}^{\Upsilon_2}\left(\phi_{\Upsilon_2}; \phi^{(l)}\right). \tag{64}$$

Let $e^{(l+1)}$ be a $K \times S$ matrix such that the $(k', p')$th element is $e_{k',p'}^{(l+1)} = \varepsilon_{k',p'}^{(l+1)}$, where $\varepsilon_{k',p'}^{(l+1)}$ is given in Eq. (44). Let $W^{(l+1)}$ be a $KS \times KS$ matrix, given by

$$W^{(l+1)} = \sum_{n=1}^{N} \left|W_n^{(l+1)}\right|^2, \tag{65}$$

where the $(K(p-1)+k)$th column of the $L \times KS$ matrix $W_n^{(l+1)}$ is given by $c_{n,k,p}^{(l)}$. The fractional parts of the time-delays are given by the following modified Yule-Walker equation

$$\text{vec}\left(\eta^{(l+1)}\right) = \left(W^{(l+1)}\right)^{-1} \text{vec}\left(e^{(l+1)}\right). \tag{66}$$

*Update of the Integer Part of the Time Delay:* Compute $(\nu_{k,p}; k = 1, \ldots K, p = 1, \ldots S)^{(l+1)}$ as follows

$$(\nu_{k,p}; k = 1, \ldots K,\ p = 1, \ldots S)^{(l+1)} = \arg \max_{(\nu_{k,p}; k=1,\ldots K,\ p=1,\ldots S)} \mathcal{Q}^{\Upsilon_3}\left(\phi_{\Upsilon_3}; \phi^{(l)}\right), \tag{67}$$

where $\mathcal{Q}^{\Upsilon_3}\left(\phi_{\Upsilon_3}; \phi^{(l)}\right)$ is given in (61). The joint maximization of $\mathcal{Q}^{\Upsilon_3}\left(\phi_{\Upsilon_3}; \phi^{(l)}\right)$ with respect to $(\nu_{k,p}; k = 1, \ldots K, p = 1, \ldots S)$ is a discrete combinatorial optimization problem with computational cost $O(L^{KS})$. In the subsection 4.2 below, we propose an alternative SAGE scheme, that reduces the cost to $O(LKS)$ per iteration.

*Update of the Complex Constant $\xi$ (which is a known function of the unknown parameters carrier phase, transmitted power, channel response):* The $K \times S$ matrix $(\xi)^{(l+1)}$ is given by the following modified Yule-Walker equation

$$\text{vec}((\xi)^{(l+1)}) = (R^{(l+1)})^{-1}\text{vec}(h^{(l+1)}), \tag{68}$$

where $h^{(l+1)}$ is a $K \times S$ matrix, and its $(k, p)$th element is given by

$$h_{k,p}^{(l+1)} = \frac{\mu_{k,p}}{\sigma_{k,p}^2} + \frac{T_c}{\sigma^2} \sum_{n=1}^{N} \left\{ \left(\delta_{n,k,p}^{(l)} \rho_{k,p}^{(l)}\right)^H y_n \exp\left\{-j\frac{2\pi}{\mathcal{M}} m_{n-1,k}^{(l)}\right\} + \left(\delta_{n,k,p}^{(l)} \varrho_{k,p}^{(l)}\right)^H y_n \exp\left\{-j\frac{2\pi}{\mathcal{M}} m_{n,k}^{(l)}\right\} \right\} \tag{69}$$

and $R^{(l+1)}$ is an $KS \times KS$ matrix, given by

$$R^{(l+1)} = \sum_{n=1}^{N} \left|R_n^{(l)}\right|^2 + \Sigma. \tag{70}$$

$\Sigma$ is given by Eq. (53) and $(R_n^r)^{(l)}$ has dimensions $L \times KS$, where the $(K(p-1)+k)$th column is given by

$$\delta_{n,k,p}^{(l)} \rho_{k,p}^{(l)} \exp\left\{ j \frac{2\pi}{\mathcal{M}} m_{n-1,k}^{(l)} \right\} + \delta_{n,k,p}^{(l)} \varrho_{k,p}^{(l)} \exp\left\{ j \frac{2\pi}{\mathcal{M}} m_{n,k}^{(l)} \right\}. \tag{71}$$

*Joint Estimation of Input Sequences and Detection of Active Propagation Paths:* The joint data estimation and active path detection is a discrete optimization problem, efficiently solved via dynamic programming (also known as the Viterbi Algorithm). The dynamic programming algorithm that yields the Markov chain sequence estimate $(M, \Delta)^{(l+1)}$ on the $(l+1)$th iteration, based on the measurements $Y$ and the parameter estimate $\theta^{(l)}$, is given in Fig. 4.

## 4.2 Algorithm III

The index sets are chosen from a partition $\{\Upsilon_1, \ldots, \Upsilon_{3+KS+NK+NKS}\}$ such that

$$\begin{aligned}
\phi_{\Upsilon_1} &= (t_{ii}^{(k,p)};\ k = 1, 2, \ldots, K,\ p = 1, 2, \ldots, S,\ i = 0, 1) && (72)\\
\phi_{\Upsilon_2} &= (\eta_{k,p};\ k = 1, 2, \ldots, K,\ p = 1, 2, \ldots, S) && (73)\\
\phi_{\Upsilon_3} &= \nu_{1,1} && (74)\\
&\ \vdots \\
\phi_{\Upsilon_{2+KS}} &= \nu_{K,S} && (75)\\
\phi_{\Upsilon_{3+KS}} &= (\xi_{k,p};\ k = 1, 2, \ldots, K,\ p = 1, 2, \ldots, S) && (76)\\
\phi_{\Upsilon_{4+KS}} &= m_{1,1} && (77)\\
&\ \vdots \\
\phi_{\Upsilon_{3+KS+NK}} &= m_{N,K} && (78)\\
\phi_{\Upsilon_{4+KS+NK}} &= \delta_{1,1,1} && (79)\\
&\ \vdots \\
\phi_{\Upsilon_{3+KS+NK+NKS}} &= \delta_{N,K,S} && (80)
\end{aligned}$$

The details of the SAGE algorithm for computing the joint MAP data and channel estimates for our asynchronous CDMA system are given below.

**Expectation Step:**

Evaluate for the $i$th index set ($i = 1, \ldots, 3+KS+NK+NKS$):

$$\mathcal{Q}^{\Upsilon_i}(\phi_{\Upsilon_i}; \phi^{(l)}) \ \triangleq\ \ln f(Y, \phi_{\Upsilon_i}, \phi_{\bar{\Upsilon}_i}). \tag{81}$$

**Maximization Steps:**

Instead of maximizing (81) directly as a function of $\theta$, the SAGE algorithm updates some parameters sequentially by keeping the others fixed to their previous values. In particular, depending on the choice of the parameter index set, $\Upsilon$, the parameters are updated as follows.

*Transition Probability Matrix Update:* The update formulae for $(t_{ii}^{k,p})^{(l+1)}$ are given by (63).

*Update of the Fractional Part of the Time Delay:* The update formula for $\eta_{k,p}^{(l+1)}$ is given by (66).

*Update of the Integer Part of the Time Delay:* Compute $\nu_{k,p}^{(l+1)}$ for $k = 1, \dots K,\ p = 1, \dots S$ as follows

$$\nu_{k,p}^{(l+1)} = \arg \max_{\nu_{k,p} \in \{1,\dots,L\}} \mathcal{Q}^{\Upsilon_{2+(k-1)S+p}}(\phi_{\Upsilon_{2+(k-1)S+p}}; \phi^{(l)}), \tag{82}$$

where $\mathcal{Q}^{\Upsilon_i}$ is given in (81). This is a discrete optimization problem with cost $O(L)$ per iteration for each $k$ and $p$.

*Update of the Complex Constant $\xi$ (which is a known function of the unknown parameters carrier phase, transmitted power, channel response):* The update formula of $\xi^{(l+1)}$ is given by (68).

*Estimation of Input Sequences:* Compute $m_{n,k}^{(l+1)}$ for $n = 1, \dots, N,\ k = 1, \dots K$ as follows

$$m_{n,k}^{(l+1)} = \arg \max_{m_{n,k} \in \{0,1,\dots,\mathcal{M}-1\}} \mathcal{Q}^{\Upsilon_{3+KS+(n-1)K+k}}(\phi_{\Upsilon_{3+KS+(n-1)K+k}}; \phi^{(l)}), \tag{83}$$

where $\mathcal{Q}^{\Upsilon_i}$ is given in (81). This is a discrete optimization problem with cost $O(\mathcal{M})$ per iteration for each $k$ and $p$.

*Detection of Active Propagation Paths:* Compute $\delta_{n,k,p}^{(l+1)}$ for $n = 1, \dots, N,\ k = 1, \dots K,\ p = 1, 2, \dots, S$ as follows

$$\delta_{n,k,p}^{(l+1)} = \arg \max_{\delta_{n,k,p} \in \{0,1\}} \mathcal{Q}^{\Upsilon_{3+NK+nKS+(k-1)S+p}}(\phi_{\Upsilon_{3+NK+nKS+(k-1)S+p}}; \phi^{(l)}), \tag{84}$$

where $\mathcal{Q}^{\Upsilon_i}$ is given in (81). This is a discrete optimization problem with cost $O(2)$ per iteration for each $(n, k, p)$.

# 5 Performance Analysis

## 5.1 Cramer Rao Lower Bounds

It is well-known that the CRLB specifies the lowest estimation error, in the mean square error sense, for any unbiased estimator of an unknown random parameter [24]. The mean square error of any unbiased estimate $\hat{\theta}$ satisfies the inequality [24]

$$E\{(\theta - \hat{\theta})(\theta - \hat{\theta})^H\} \geq C^{-1}, \tag{85}$$

where

$$C = E\left\{-\frac{\partial^2 \ln f(Y, \theta)}{\partial\theta\partial\theta'}\right\} \tag{86}$$

Computing $C$ in Eq. (86) is not straightforward for our given model assumptions. Instead of Eq. (86), it is a standard practice in the CDMA literature to compute the Fisher Information Matrix (FIM) assuming that the transmitted data are exactly known [18, 13], *i.e.*,

$$\text{FIM}(M,\theta) = E\left\{ -\frac{\partial^2 \ln f(Y|M,\theta)}{\partial\theta\partial\theta'} \middle| M,\theta \right\}. \tag{87}$$

Assuming the data sequence $M$ and the active propagation paths $\Delta$ are known, then the FIM is given by

$$\text{FIM}(M,\Delta,\theta) = E\left\{ -\frac{\partial^2 \ln f(Y|M,\Delta,\theta)}{\partial\theta\partial\theta'} \middle| M,\Delta,\theta \right\}. \tag{88}$$

In practice $M$ and $\Delta$ are not known. In this paper, we compute the complete data information matrix defined as follows

$$C_{\text{com}} = E\left\{ -\frac{\partial^2 \ln f(Y,M,\Delta,\theta)}{\partial\theta\partial\theta'} \right\}. \tag{89}$$

The complete data information matrix is related to Eq. (86) as follows (known as the "missing information principle") [11]

$$E\left\{ -\frac{\partial^2 \ln f(Y,\theta)}{\partial\theta\partial\theta'} \right\} = E\left\{ -\frac{\partial^2 \ln f(Y,M,\Delta,\theta)}{\partial\theta\partial\theta'} \right\} - E\left\{ -\frac{\partial^2 \ln f(M,\Delta|Y,\theta)}{\partial\theta\partial\theta'} \right\}. \tag{90}$$

The elements of $C_{\text{com}}$ are given in the appendix.

## 5.2 Numerical Results

One thousand Monte Carlo simulations have been carried out to evaluate the performance of our proposed detection and estimation schemes. The errors of the channel estimates were compared to the computed estimation bounds derived in the previous section. The BER in detecting the input sequence and the presence of multipaths were compared to results achieved using the Viterbi algorithm assuming all other parameters known.

The number of users was chosen to be $K = 2$ and the maximum number of paths for each user was $S = 2$. Furthermore, $\mathcal{M} = 2$ and the transmitted data are assumed to be iid, equiprobable random variables. The signal to noise ratio (SNR) of the first user was fixed at 8 dB, and the ratio of the second user to the first user was 8 dB. This scenario was investigated in [13]. In [13], the propagation paths are always present. In our formulation, we consider time varying number of paths. The transition probability of $\delta_{n,k,p}$ was chosen as

$$\Pi^{(k,p)} = \begin{bmatrix} 0.9 & 0.1 \\ 0.1 & 0.9 \end{bmatrix} \quad \text{for } k = 1,2,\ p = 1,2.$$

The number of chips per symbol was $L = 31$. The spreading sequences $d_k(l)$ and the channel parameters were randomly chosen from one simulation to the next. We assumed uniform prior on all the parameters and assumed the channel attenuations to be zero mean. Furthermore, $E\{\delta_{0,k,p}\} = 0.5, \forall k \in \{1, \dots, K\}, p = \{1, \dots, S\}$. The algorithms were iterated until convergence (*i.e.* $\phi^{(l+1)} = \phi^{(l)}$) or up to a maximum number of 20 iterations. All channel parameters were randomly initialized.

In Fig. 5, the root mean square (RMS) estimation errors of the channel coefficients, $\xi$, as a function of the number of bits for each of the three algorithms are depicted. Comparisons with the lower bounds computed in the previous section were made. Clearly, the performance of Algorithm I is better than those of Algorithm II and III. Furthermore, the performance of Algorithm II and III are very similar. Next, in Fig. 6, the RMS of the time delays as a function of the number of bits for each of the three algorithms are depicted. It is clearly shown, that all our proposed detection and estimation algorithms yield satisfactory channel estimates. In Fig. 7 and Fig. 8, the bit error rate (BER) of the input sequences and the BER in detecting the presence of different multipaths are shown, respectively. We compare the results with the BER obtained using a Viterbi algorithm to compute the transmitted data and active propagation paths, assuming all channel parameters known. Here, once again Algorithm I outperforms Algorithms II and III. Finally, in Fig. 9, the RMS of the fractional time delay of Algorithm I is depicted as a function of the SNR of the first user. Here, the ratio of the second user to the first user was 8 dB and $N$ was fixed to 200. It was observed that the RMS of $\xi$ and the BER of the input sequences and the BER of the detection of multipaths showed similar results. Algorithms II and III yield similar results.

To conclude, the convergence of Algorithm I is definitely better than the convergence of Algorithms II and III. Though, the computational complexity of Algorithms I and II is much higher than that of Algorithm III. The BER of the detection of multipath sequence is much less for Algorithm I than it is for Algorithm III and II. If only channel estimates are of interest and it is crucial having low computational complexity, Algorithm III should preferably be used, otherwise Algorithm I should be considered.

# 6 Conclusions

In this paper, we studied the problem of multipath detection and parameter estimation in an asynchronous CDMA system with unknown and time-varying number of multipaths. Three algorithms, called Algorithm I, II and III in the paper, were proposed. They were all based on a generalization of the EM algorithm, known as the SAGE algorithm. The SAGE algorithm combines jointly multiuser detection and parameter estimation. Depending on the realizations of the SAGE algorithm well-known state and parameter estimation, such as the HMM smoother, Viterbi algorithm, Yule-Walker and Baum-Welch re-estimation formulae, were optimally combined. These three algorithms differed in the choices of hidden data spaces. Algorithm I computes the MAP parameter estimates as

$$\theta^{\mathrm{MAP}} = \arg\max_{\theta} f(\theta|Y) \tag{91}$$

and by using a hidden Markov model smoother, the MAP estimates of input sequences, $\mathcal{M}$, and the detection of the active propagation paths, $\Delta$, were computed as follows

$$\hat{\delta}_{n,k,p}^{\mathrm{MAP}} = \arg\max_{\delta_{n,k,p}} \Pr\left(\delta_{n,k,p} \mid \theta^{\mathrm{MAP}}, Y\right), \tag{92}$$

$$\hat{m}_{n,k}^{\mathrm{MAP}} = \arg\max_{m_{n,k}} \Pr\left(m_{n,k} \mid \theta^{\mathrm{MAP}}, Y\right). \tag{93}$$

Algorithms II and III, on the other hand, compute the joint MAP data and channel estimates as follows

$$(\theta, M, \Delta)^{\mathrm{MAP}} = \arg\max_{(\theta,M,\Delta)} f(\theta, M, \Delta|Y). \tag{94}$$

Algorithms II and III differed in the choice of hidden data spaces. Algorithm II uses the Viterbi algorithm to achieve estimates of $\mathcal{M}$ and $\Delta$ while Algorithm III does it in a much less computationally expensive way. Algorithm III has a computational cost that grows linearly with the number of users and propagation paths. Simulation studies, where RMS and BER were compared to theoretical bounds, showed as expected that Algorithm I outperformed Algorithms II and III in terms of BER. The RMS errors of the channel parameter $\xi$ and the time delay $\tau$ were close to the CRLB for all the proposed schemes.

As future work we intend to extend our schemes to handle fading stochastically time-varying multipath channels. Furthermore, future work will also address the case of severe intersymbol interference (ISI) when the channel impulse response spans more than one symbol period.

# Appendix

## A Cramer Rao Lower Bounds

Let $\xi^R_{k,p}$ and $\xi^I_{k,p}$, denote the real and imaginary part of $\xi_{k,p}$, respectively. The elements of $C_{\text{com}}$ ($\forall k,k' \in \{1,\ldots,K\}, p,p' \in \{1,\ldots,S\}$) defined in Eq. (89) are given by

$$E\left\{-\frac{\partial^2 \ln f(Y,M,\Delta,\theta)}{\partial\xi^R_{k,p}\partial\xi^R_{k',p'}}\right\} = E\left\{-\frac{\partial^2 \ln f(Y,M,\Delta,\theta)}{\partial\xi^I_{k',p'}\partial\xi^I_{k,p}}\right\} = \frac{2}{\sigma^2_{k,p}}\mathrm{I}\{k-k'\}\mathrm{I}\{p-p'\}$$
$$+ \quad 2\frac{T_c}{\sigma^2}E\left\{\sum_{n=1}^{N}\delta_{n,k',p'}\delta_{n,k,p}\left\{\rho^H_{k',p'}\rho_{k,p}\cos\left(\frac{2\pi}{\mathcal{M}}(m_{n-1,k}-m_{n-1,k'})\right) + \varrho^H_{k',p'}\rho_{k,p}\cos\left(\frac{2\pi}{\mathcal{M}}(m_{n-1,k}-m_{n,k'})\right)\right.\right.$$
$$\left.\left. +\rho^H_{k',p'}\varrho_{k,p}\cos\left(\frac{2\pi}{\mathcal{M}}(m_{n,k}-m_{n-1,k'})\right) + \varrho^H_{k',p'}\varrho_{k,p}\cos\left(\frac{2\pi}{\mathcal{M}}(m_{n,k}-m_{n,k'})\right)\right\}\right\}, \tag{95}$$

$$E\left\{-\frac{\partial^2 \ln f(Y,M,\Delta,\theta)}{\partial\eta_{k,p}\partial\eta_{k',p'}}\right\} = 2\frac{T_c}{\sigma^2}\mathrm{Re}\left\{\sum_{n=1}^{N}E\left\{\delta_{n,k,p}\delta_{n,k',p'}\xi_{n,k,p}\xi^H_{n,k',p'}\right.\right. \times$$
$$\left((-a^r(\nu_{k',p'}) + a^r(\nu_{k',p'}+1))\exp\left\{j\frac{2\pi}{\mathcal{M}}m_{n-1,k'}\right\} + (-a^l(\nu_{k',p'}) + a^l(\nu_{k',p'}+1))\exp\left\{j\frac{2\pi}{\mathcal{M}}m_{n,k'}\right\}\right)^H \times$$
$$\left.\left.\left((-a^r(\nu_{k,p}) + a^r(\nu_{k,p}+1))\exp\left\{j\frac{2\pi}{\mathcal{M}}m_{n-1,k}\right\} + (-a^l(\nu_{k,p}) + a^l(\nu_{k,p}+1))\exp\left\{j\frac{2\pi}{\mathcal{M}}m_{n,k}\right\}\right)\right\}\right\}, \tag{96}$$

$$E\left\{-\frac{\partial^2 \ln f(Y,M,\Delta,\theta)}{\partial\eta_{k,p}\partial\xi_{k',p'}}\right\} = 2\frac{T_c}{\sigma^2}\mathrm{Re}\left\{\sum_{n=1}^{N}E\left\{\delta_{n,k,p}\delta_{n,k',p'}\xi_{n,k,p}\times\right.\right.$$
$$\left((-a^r(\nu_{k,p}) + a^r(\nu_{k,p}+1))\exp\left\{j\frac{2\pi}{\mathcal{M}}m_{n-1,k}\right\} + (-a^l(\nu_{k,p}) + a^l(\nu_{k,p}+1))\exp\left\{j\frac{2\pi}{\mathcal{M}}m_{n,k}\right\}\right)^H \times$$
$$\left\{\left(\rho_{k',p'}\exp\left\{j\frac{2\pi}{\mathcal{M}}m_{n-1,k'}\right\} + \varrho_{k',p'}\exp\left\{j\frac{2\pi}{\mathcal{M}}m_{n,k'}\right\}\right) + \mathrm{I}\{k-k'\}\mathrm{I}\{p-p'\}\times\right.$$
$$\left.\left.\left.\left(y_n - \sum_{k=1}^{K}\sum_{p=1}^{S}\delta_{n,k,p}\xi_{k,p}\rho_{k,p}\exp\left\{j\frac{2\pi}{\mathcal{M}}m_{n-1,k}\right\} - \sum_{k=1}^{K}\sum_{p=1}^{S}\delta_{n,k,p}\xi_{k,p}\varrho_{k,p}\exp\left\{j\frac{2\pi}{\mathcal{M}}m_{n,k}\right\}\right)\right\}\right\}\right\}, \tag{97}$$

$$E\left\{-\frac{\partial^2 \ln f(Y,M,\Delta,\theta)}{\partial t^{(k,p)}_{ij}\partial t^{(k,p)}_{ij}}\right\} = E\left\{\frac{\sum_{n=1}^{N-1}\Pr\{\delta_{n,k,p}=i\}}{t^{(k,p)}_{ij}}\right\}, \tag{98}$$

for $i,j \in \{0,1,\ldots,\mathcal{M}-1\}$, where $\Pr\{\delta_{n,k,p}=j\}$ is computed as follows

$$\begin{pmatrix}\Pr\{\delta_{n,k,p}=0\}\\ \Pr\{\delta_{n,k,p}=1\}\end{pmatrix} = \left(\left(\Pi^{(k,p)}\right)^H\right)^n\begin{pmatrix}\Pr\{\delta_{0,k,p}=0\}\\ \Pr\{\delta_{0,k,p}=1\}\end{pmatrix} \tag{99}$$

and $\Pr\{\delta_{0,k,p}=i\}$ is assumed known.

The following can be shown

$$\left(\Pi^{(k,p)}\right)^n = \begin{pmatrix}(t^{(k,p)}_{11}-1)/(t^{(k,p)}_{00}-1) & -1\\ 1 & 1\end{pmatrix}\begin{pmatrix}1 & 0\\ 0 & \left(t^{(k,p)}_{11}+t^{(k,p)}_{00}-1\right)^n\end{pmatrix}\begin{pmatrix}(t^{(k,p)}_{11}-1)/(t^{(k,p)}_{00}-1) & -1\\ 1 & 1\end{pmatrix}^{-1} \tag{100}$$

Thus,

$$\begin{aligned} E\{\delta_{n,k,p}\} &= E\left\{E\{\delta_{n,k,p}|t_{00}^{(k,p)}, t_{11}^{(k,p)}\}\right\} = E\left\{\frac{\left(t_{00}^{(k,p)}-1\right)\left(1-\left(t_{00}^{(k,p)}+t_{11}^{(k,p)}-1\right)^n\right)}{t_{00}^{(k,p)}+t_{11}^{(k,p)}-2}\right\}\Pr\{\delta_{0,k,p}=0\} \\ &+ E\left\{\frac{\left(t_{00}^{(k,p)}-1\right)+\left(t_{11}^{(k,p)}-1\right)\left(t_{00}^{(k,p)}+t_{11}^{(k,p)}-1\right)^n}{t_{00}^{(k,p)}+t_{11}^{(k,p)}-2}\right\}\Pr\{\delta_{0,k,p}=1\}. \end{aligned} \tag{101}$$

Assuming uniform prior on $t_{00}^{(k,p)}$ and $t_{11}^{(k,p)}$, it can be shown

$$E\{\delta_{n,k,p}\} = \frac{1}{2}\left(1-\frac{1+(-1)^n}{(n+1)(n+2)}\right)\Pr\{\delta_{0,k,p}=0\} + \frac{1}{2}\left(1+\frac{1+(-1)^n}{(n+1)(n+2)}\right)\Pr\{\delta_{0,k,p}=1\}. \tag{102}$$

Other important results are as follows

$$\begin{array}{rr} |a_k^r(\nu_{k,p})|^2 = \nu_{k,p} & \left|a_k^l(\nu_{k,p})\right|^2 = L-\nu_{k,p}, \\ (a_k^r(\nu_{k,p}))^H a_k^l(\nu_{k,p}) = 0, & (a_k^r(\nu_{k,p}))^H a_k^l(\nu_{k,p}+1) = 0, \\ (a_k^r(\nu_{k,p}+1))^H a_k^l(\nu_{k,p}) = d_k(0)d_k(L-1), & (a_k^r(\nu_{k,p}+1))^H a_k^l(\nu_{k,p}+1) = 0, \\ (a_k^r(\nu_{k,p}))^H a_k^r(\nu_{k,p}+1) = \sum_{n=1}^{\nu_{k,p}} d_k(L-n)d_k(L-n-1), & \\ \left(a_k^l(\nu_{k,p})\right)^H a_k^l(\nu_{k,p}+1) = \sum_{n=1}^{L-\nu_{k,p}-1} d_k(n)d_k(n-1) & \end{array} \tag{103}$$

Finally, it can be shown that

$$E\{\varrho_{k',p'}^H \rho_{k,p}\} = 0 \tag{104}$$

$$E\{\rho_{k',p'}^H \rho_{k,p}\} = \begin{cases} L/3, & \text{if } k=k', p=p' \\ 0, & \text{otherwise} \end{cases} \tag{105}$$

$$E\{\varrho_{k',p'}^H \varrho_{k,p}\} = \begin{cases} L/3, & \text{if } k=k', p=p' \\ 0, & \text{otherwise} \end{cases} \tag{106}$$

# References


[1] R.L. Pickholtz, L.B. Milstein, D.L. Shilling, "Spread spectrum for mobile communications", *IEEE Transactions on Vehicular Technology*, Vol. 40, pp. 313–322, 1991.

[2] T.S. Rappaport, S.Y. Seidel, K. Takamizawa, "Statistical channel impulse response models for factory and open plan building radio communication design", *IEEE Transactions on Communications*, Vol. 39, pp. 794–807, 1991.

[3] A. Radovic, E. Zollinger, "Measured time-variant characteristics of radio channels in the indoor environment", *Proceedings of the Mobile Radio Conference (MRC)*, pp. 267–274, 1991.

[4] S. Verdù, "Minimum probability of error for asynchronous Gaussian multiple access channels", *IEEE Transactions on Information Theory*, Vol. 32, pp. 85–96, 1986.

[5] H.V. Poor, S. Verdù, "Single-user detectors for multiuser channels", *IEEE Transactions on Communicationss*, Vol. 36, pp. 50–60, 1988.

[6] R. Lupas and S. Verdú, "Linear multiuser detectors for synchronous code-division multiple-access channels", *IEEE Transactions on Information Theory*, Vol. 35, no. 1, pp. 123–136, 1989.

[7] M.K. Varanasi, B. Aazhang, "Multistage detection in asynchronous code-division multiple access communications", *IEEE Transactions on Communications*, Vol. 38, pp. 509–519, 1990.

[8] R. Lupas and S. Verdú, "Near-far resistance of multiuser detectors in asynchronous channels", *IEEE Transactions on Communications*, Vol. 38, no. 4, pp. 496–508, 1990.

[9] Z. Xie, C. Rushforth and R. Short, "Multiuser signal detection using sequential decoding", *IEEE Transactions on Communications*, Vol. 38, no. 5, pp. 578–583, 1990.

[10] C.N. Georghiades, J. Choong Han, "Sequence estimation in the presence of random parameters via the EM algorithm", *IEEE Transactions on Communications*, Vol. 45, pp. 300–308, 1997.

[11] A.P. Dempster, N.M. Laird and D.B. Rubin, "Maximum likelihood from incomplete data via the EM algorithm", *Journal of the Royal Statistical Society*, Ser. B, Vol. 39, No. 1, pp. 1–38, 1977.

[12] L.B. Nelson, H.V. Poor, "Iterative multiuser receivers for CDMA channels: An EM based approach", *IEEE Transactions on Communications*, Vol. 44, pp. 1700–1710, 1996.

[13] U. Fawer and B. Aazhang, "A Multiuser Receiver for Code Division Multiple Access Communications over Multipath Channels", *IEEE Transactions on Communications*, Vol. 43, No. 2/3/4, pp. 1556–1565, 1995.

[14] H.V. Poor, "On parameter estimation in DS/SSMA formats", in *Advances in Communications and Signal Processing*, Springer Verlag, New York, pp. 59–70, 1988.

[15] C. Georghiades and D. Snyder, "The expectation-maximization algorithm for symbol unsynchronized sequence detection", *IEEE Transactions on Communications*, Vol. 39, no. 1, pp. 54–61, 1991.

[16] M.K. Varanasi, S. Vasudevan, "Multiuser detectors for synchronous CDMA communication over nonselective Rician fading channels", *IEEE Transactions on Communications*, Vol. 42, pp. 711–722, 1994.

[17] S. Bensley and B. Aazhang, “Subspace-based channel estimation for code division multiple access communication systems”, *IEEE Transactions on Communications*, Vol. 44, no. 8, pp. 1009–1020, 1996.

[18] S. Parkvall, *Near-Far Resistant DS-CDMA Systems: Parameter Estimation and Data Detection*, PhD thesis, Stockholm, 1996.

[19] C. Carlemalm and A. Logothetis, ”Blind Signal Detection and Parameter Estimation for an Asynchronous CDMA System with Time-Varying Number of Transmission Paths”, *Proceedings 9th IEEE Signal Processing Workshop on Statistical Signal and Array Processing*, Portland, USA, September 1998.

[20] A. Chkeif and G. Kawas Kaleh, ”Iterative Multiuser Detector with Antenna Array: An EM-Based Approach”, *Proceedings International Symposium on Information Theory, ISIT'98*, Cambridge, USA, 1998.

[21] R. Wang and S.D. Blostein, ”Spatial-Temporal CDMA Receiver Structures for Rayleigh Fading Channels”, *Proceedings International Conference on Communications, ICC'99*, Vancouver, Canada, 1999.

[22] J.A. Fessler and A.O. Hero, “Space-alternating generalized expectation-maximization algorithm”, *IEEE Transactions on Signal Processing*, Vol. 42, No. 10, pp. 2664–2677, 1994.

[23] L.R. Rabiner, “A tutorial on hidden Markov models and selected applications in speech recognition”, *Proceedings of the IEEE*, Vol. 77, pp. 257–286, 1989.

[24] H.L. van Trees, *Detection, Estimation and Modulation Theory*, New York, NY: Wiley, 1968.

**1.** Choose an index set $\Upsilon \subset \Omega$.

**2.** Choose an admissible hidden-data space $X^\Upsilon$ for $\phi_\Upsilon$.

**3. Expectation step:** Compute

$$\mathcal{Q}^\Upsilon \left(\phi_\Upsilon; \phi^{(l)}\right) \triangleq E\left\{\log f\left(Y, X^\Upsilon, \phi_\Upsilon, \phi_{\tilde{\Upsilon}}^{(l)}\right) |Y, \phi^{(l)}\right\}. \tag{107}$$

**4. Maximization step:**

$$\phi_\Upsilon^{(l+1)} = \arg\max_{\phi_\Upsilon} \mathcal{Q}^\Upsilon \left(\phi_\Upsilon; \phi^{(l)}\right), \tag{108}$$

$$\phi_{\tilde{\Upsilon}}^{(l+1)} = \phi_{\tilde{\Upsilon}}^{(l)}. \tag{109}$$

**5.** Optional: Repeat steps 2–4 until convergence.

Figure 1: The SAGE Algorithm.

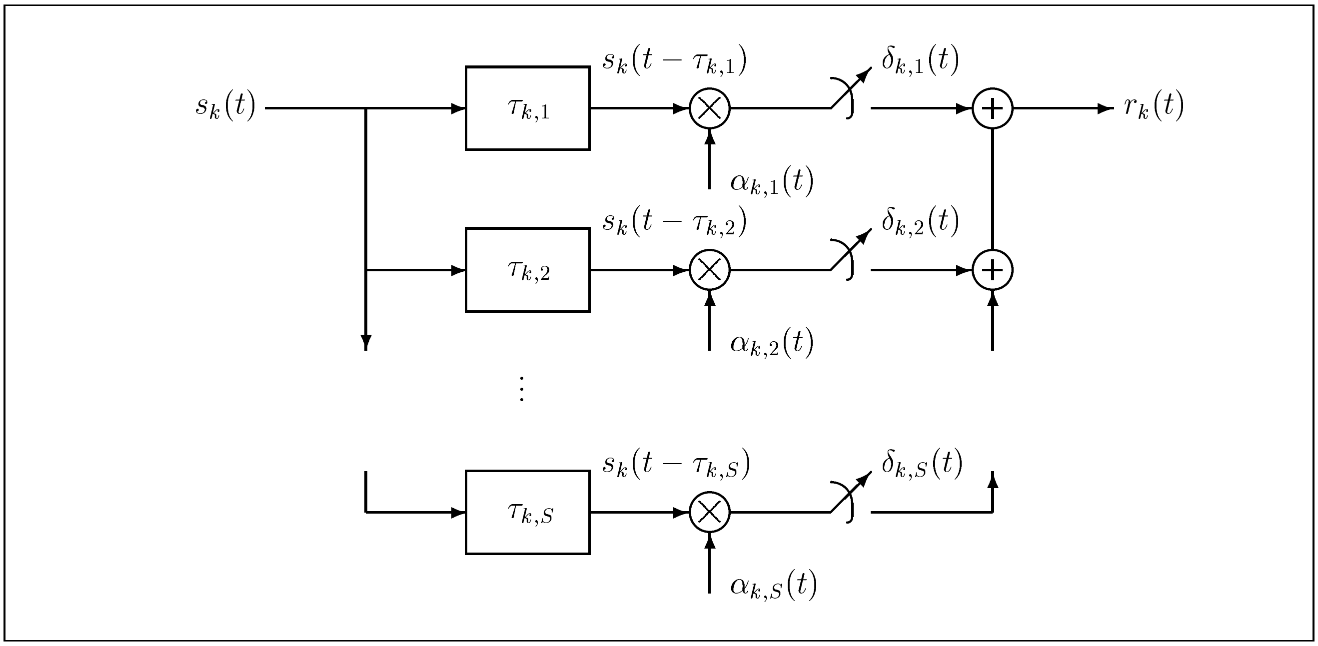


Figure 2: The continuous-time channel model for the $k$th user.

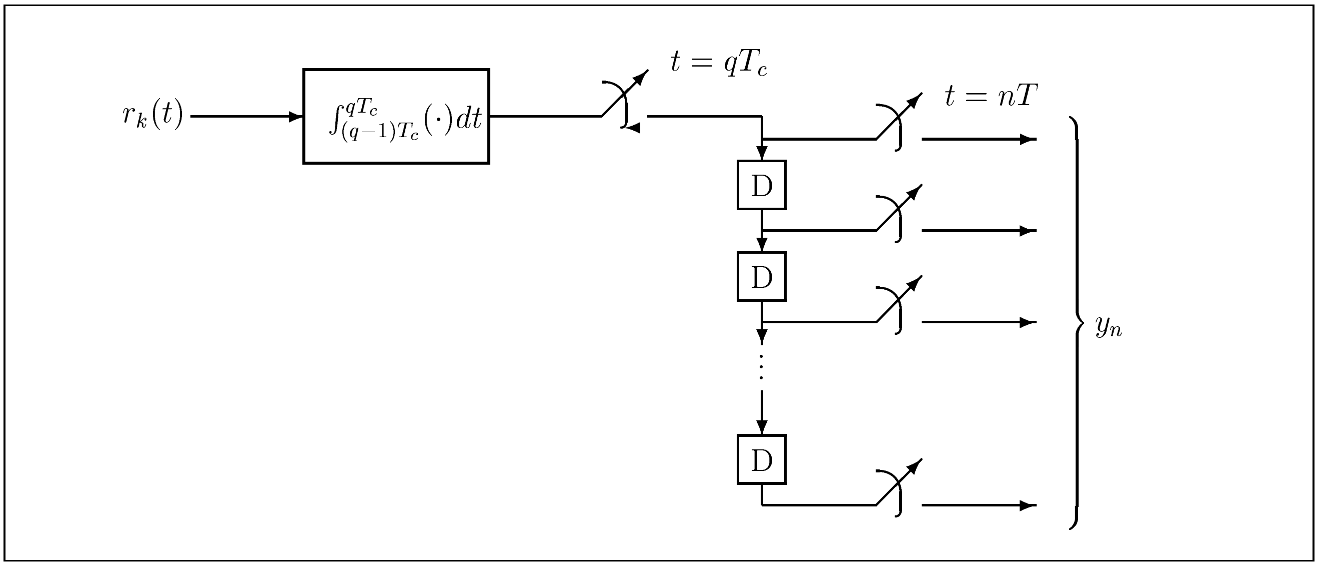


Figure 3: Sampling at the receiver. The discrete signals at the output of the integrate-and-dump filter are generated at chip rate. The discrete output vector process $y_n$ is generated at bit rate.

**Viterbi algorithm:**

**1. Initialization:**
For $i_m, j_m \in \{0, 1, \ldots, \mathcal{M}-1\}^K$ and $i_\delta \in \{0,1\}^{KS}$, set

$$\zeta_1(i_m, j_m, i_\delta) \quad = \quad \ln b_1^{(l)}(i_m, j_m, i_\delta) \tag{110}$$

**2. Recursion:**
For $2 \leq n \leq N$, $j_m, l_m \in \{0, 1, \ldots, \mathcal{M}-1\}^K$ and $j_\delta \in \{0,1\}^{KS}$

$$\begin{aligned}\zeta_n(j_m, l_m, j_\delta) &= \ln b_n^{(l)}(j_m, l_m, j_\delta) \\ &+ \max_{\substack{i_m \in \{0,1,\ldots,\mathcal{M}-1\}^K \\ i_\delta \in \{0,1\}^{KS}}} \left[\zeta_{n-1}(i_m, j_m, i_\delta) + \ln \Pr(\delta_n = j_\delta | \delta_{n-1} = i_\delta)\right]\end{aligned} \tag{111}$$

$$\begin{aligned}&\psi_n(j_m, l_m, j_\delta) = \\ &\quad \arg \max_{\substack{i_m \in \{0,1,\ldots,\mathcal{M}-1\}^K \\ i_\delta \in \{0,1\}^{KS}}} \left[\zeta_{n-1}(i_m, j_m, i_\delta) + \ln \Pr(\delta_n = j_\delta | \delta_{n-1} = i_\delta)\right]\end{aligned} \tag{112}$$

**3. Termination:**

$$(m_{N-1}, m_N, \delta_N)^{(l+1)} = \arg \max_{\substack{i_m, j_m \in \{0,\ldots,\mathcal{M}-1\}^K \\ i_\delta \in \{0,1\}^{KS}}} \zeta_N(i_m, j_m, i_\delta) \tag{113}$$

**4. Backtracking:**
For $n = N-1, N-2, \ldots, 2$

$$(m_{n-1}, m_n, \delta_n)^{(l+1)} = \psi_{n+1}\left(m_n^{(l+1)}, m_{n+1}^{(l+1)}, \delta_{n+1}^{(l+1)}\right) \tag{114}$$

Figure 4: Joint estimation of input sequence and detection of active propagation paths for Algorithm II

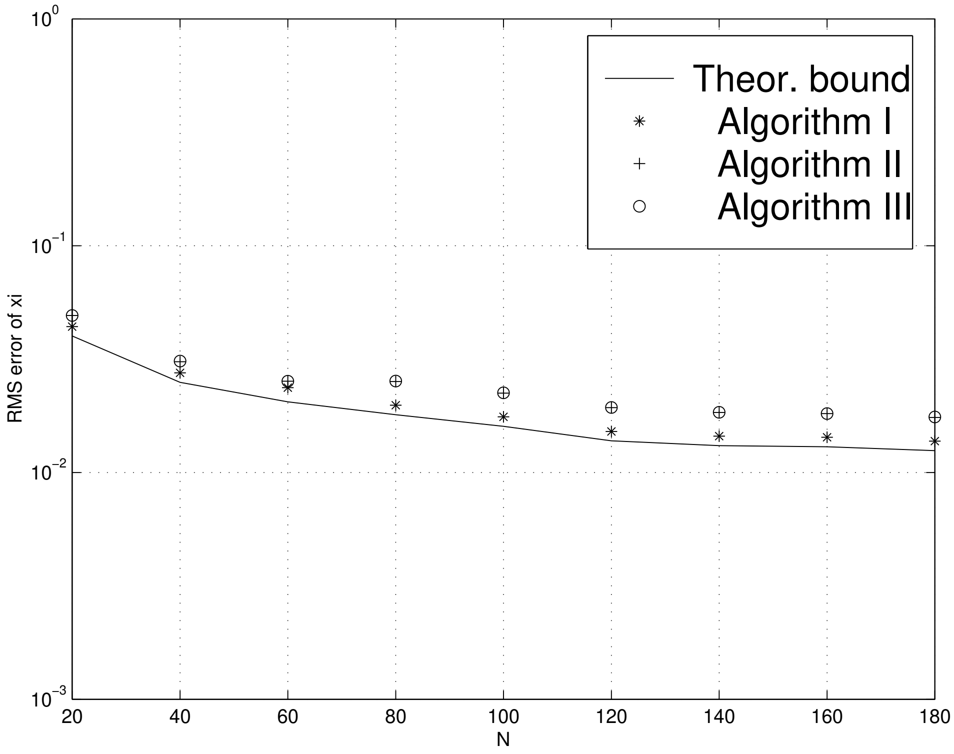


Figure 5: Simulated RMS error of the estimated coefficient $\xi$ versus the number of samples $N$.

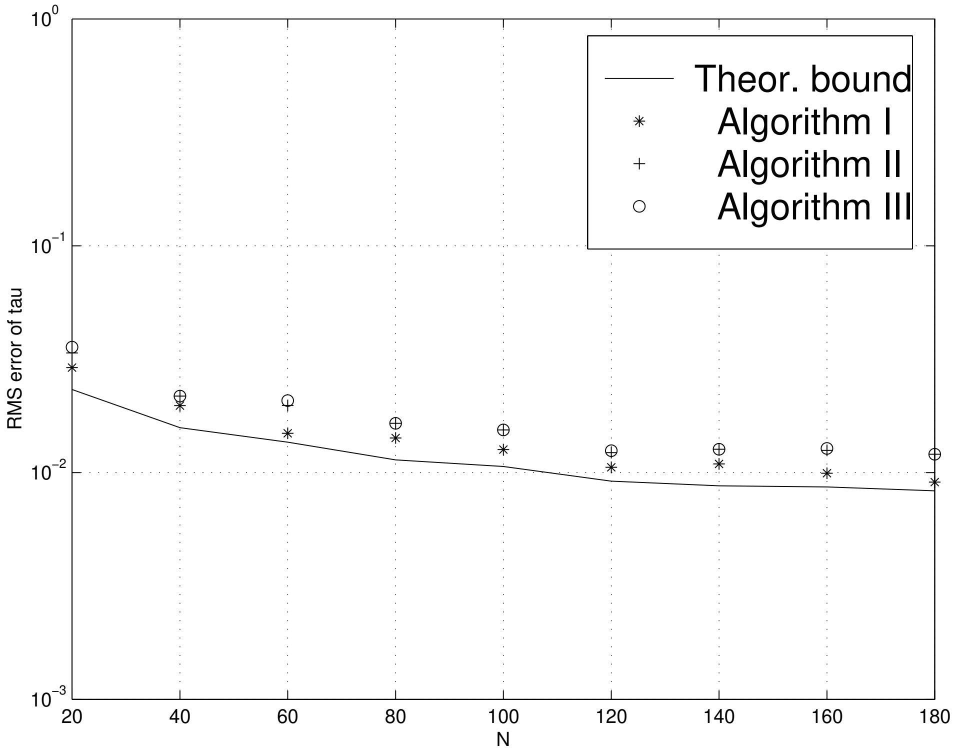


Figure 6: Simulated RMS error of the estimated delays $\tau$ versus the number of samples $N$.

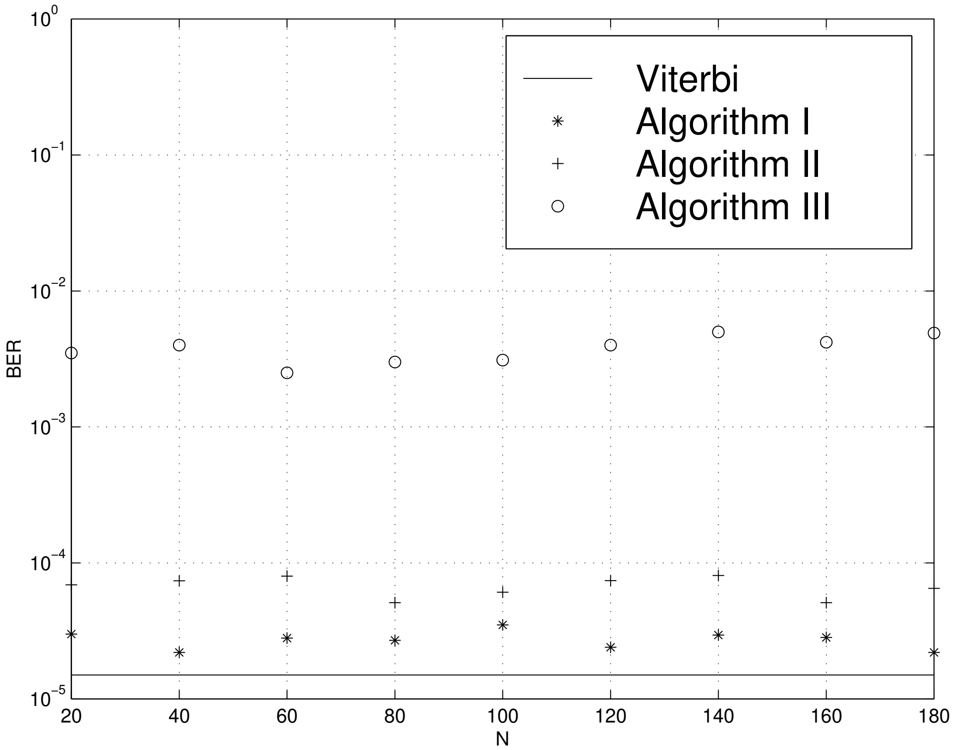


Figure 7: Simulated BER of the estimated data versus the number of samples $N$. The solid line illustrates the BER computed using a Viterbi algorithm assuming the channel parameters are known.

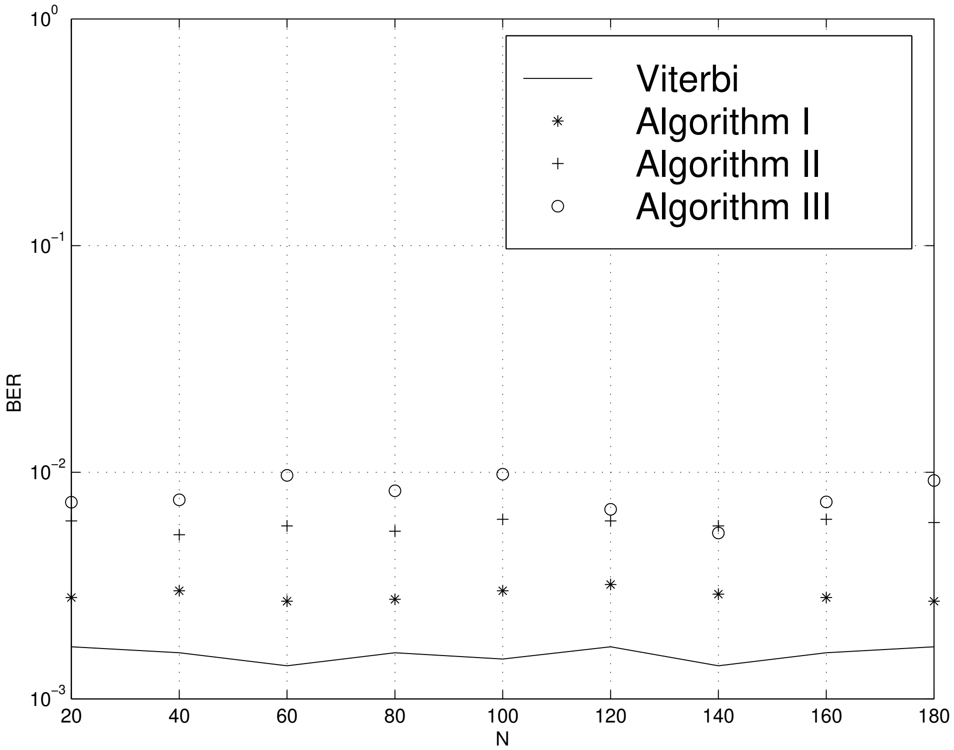


Figure 8: Simulated BER in detecting the presence of active propagation paths versus the number of samples $N$. The solid line illustrates the BER computted using a Viterbi algorithm assuming the channel parameters are known.

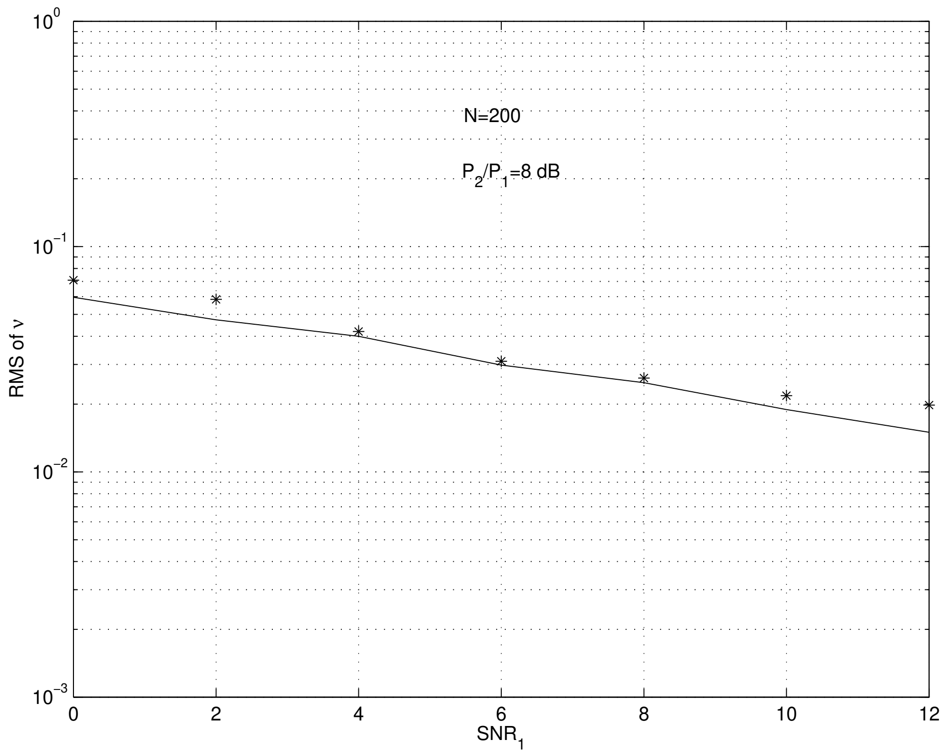


Figure 9: **Algorithm I:** Simulated RMS error of the estimated delays $\nu$ versus the SNR of the first user.