EURASIP Journal on Wireless Communications and Networking 2005:2, 130–140

Adaptive Blind Multiuser Detection over Flat Fast Fading Channels Using Particle Filtering

Yufei Huang

Department of Electrical Engineering, The University of Texas at San Antonio, San Antonio, TX 78249-06615, USA
Email: yhuang@utsa.edu

Jianqiu (Michelle) Zhang

Department of Electrical and Computer Engineering, University of New Hampshire, Durham, NH 03824, USA
Email: jianqiu.zhang@unh.edu

Isabel Tienda Luna

Departamento de Física Aplicada, Universidad de Granada, Granada 18071, Spain
Email: isabelt@ugr.es

Petar M. Djurić

Department of Electrical and Computer Engineering, Stony Brook University, Stony Brook, NY 11794-2350, USA
Email: djuric@ece.sunysb.edu

Diego Pablo Ruiz Padillo

Departamento de Física Aplicada, Universidad de Granada, Granada 18071, Spain
Email: druiz@ugr.es

Received 30 April 2004; Revised 16 September 2004

We propose a method for blind multiuser detection (MUD) in synchronous systems over flat and fast Rayleigh fading channels. We adopt an autoregressive-moving-average (ARMA) process to model the temporal correlation of the channels. Based on the ARMA process, we propose a novel time-observation state-space model (TOSSM) that describes the dynamics of the addressed multiuser system. The TOSSM allows an MUD with natural blending of low-complexity particle filtering (PF) and mixture Kalman filtering (for channel estimation). We further propose to use a more efficient PF algorithm known as the stochastic M-algorithm (SMA), which, although having lower complexity than the generic PF implementation, maintains comparable performance.

Keywords and phrases: multiuser detection, time-observation state-space model, fading channel estimation, particle filtering, mixture Kalman filter.

1. INTRODUCTION

When multiuser detection (MUD) was introduced in the eighties, it has received a great deal of attention due to its ability to reduce multiple access interference (MAI) and potential for increasing the capacity of CDMA systems. Since then, numerous detectors have been proposed in the literature for both synchronous and asynchronous transmission and some popular ones include the decorrelating detector, the minimum mean square error (MMSE) detector, the multistage detector, and the decision feedback detector [1].

In practice, distortion in signal strength due to time-varying fading channels must be attended while performing MUD. Even though noncoherent detection methods as proposed in [2] are often appealing owing to their simplicity since no inference on fading channels is needed, coherent detection has been proved to deliver better performance [3]. With coherent detection, estimation of channels can be obtained with or without pilot signals. Between them, significant amount of research has been devoted to schemes without using pilot signals, or blind MUD methods. Blind MUD methods are bandwidth more efficient and the approaches proposed, to name a few, include the recursive least square (RLS) [4, 5], subspace-based [6], expectation-maximization [7], genetic algorithm [8] and Kalman filtering [9, 10, 11, 12, 13, 14]. However, most of the approaches cited above assume slow or quasi-static fading channels.

In this paper, we focus on blind MUD for fast flat Rayleigh fading channels and in synchronous systems. In particular, we assume to know *a priori* the second-order statistics of the underlying channel, based on which a mathematical tractable approximation using autoregressive-moving-average (ARMA) model is adopted. The approximation enables a dynamic state-space modeling (DSSM) of the problem, which lends itself naturally to a Kalman-filtering-related detection solution. The use of Kalman filtering for blind MUD on similar modeling has been seen in [11, 12, 14], where the decision-directed approach was used to estimate the channel variable necessary for the Kalman filtering. One inherent drawback with the decision-directed approach is the error propagation, which greatly limits the performance of such implementation.

Recently, the combined (mixture) Kalman filtering and sequential importance sampling (particle filtering) algorithms have been applied to blind detection of convolutional codes [15], space-time trellis codes [16], and blind MUD [17] over fading channels. The mixture Kalman filtering (MKF) approach is shown to greatly reduce the error propagation of the decision-directed implementations and thus exhibits considerable performance improvement. However, in the proposed use of the MKF to blind MUD in [17], particle filtering (PF) was mainly intended for channel tracking and the embedded MUD at a symbol interval was achieved by an optimum detector, which has exponential complexity with the number of users. Consequently, the proposed MKF algorithm becomes prohibitively complex even for systems with moderate number of users.

In this paper, unlike all existing Kalman filtering detectors, a completely different viewpoint to multiuser systems is taken and we propose a novel time-observation state-space model (TOSSM). Even though the TOSSM is equivalent to the common DSSM, it allows the PF-based multiuser detection to be naturally blended with the mixture Kalman filtering for channel estimation. The new mixture Kalman filtering algorithm samples one user at a time and therefore permits efficient implementation. We further propose to use a more efficient PF algorithm known as the stochastic M-algorithm (SMA), which has shown to attain additional complexity reduction over the generic PF implementation and yet maintain comparable performance.

The rest of the paper is organized as follows. In Section 2, the problem of blind MUD is formulated. In Section 3, a novel TOSSM is described and in Section 4, the optimum solution is discussed. Particle filtering and SMA solutions are proposed in Sections 6 and 7, respectively. The simulation results are presented in Section 8. Section 9 contains some concluding remarks.

2. PROBLEM FORMULATION

Consider a synchronous CDMA system with a processing gain C and K users. Let T denote the symbol duration and $s_k(t)$ the normalized deterministic signature waveform assigned to the kth user. Then, at the nth symbol interval, the received signal $y(t)$ can be expressed as a summation of K antipodally modulated synchronous signature waveforms plus noise, that is,

$$y(t) = \sum_{k=1}^{K} a_{n,k} b_{n,k} s_k(t) + u(t), \quad t \in \big[(n-1)T, nT\big], \tag{1}$$

where $b_{n,k} \in \{-1, +1\}$ is the BPSK modulated bit transmitted by the kth user, $a_{k,n}$ the CSI (fading coefficient) of the kth user, and $u(t)$ the received zero mean additive complex white Gaussian noise with variance σ^2. The cross-correlation between the signature waveforms of the users is given by the cross-correlation matrix $\mathbf{R}$, where element r_{k_1,k_2} represents the cross-correlation between the signature waveform of the k_1th and the k_2th user and is defined as

$$r_{k_1 k_2} = \langle s_{k_1}, s_{k_2} \rangle = \int_{(n-1)T}^{nT} s_{k_1}(t) s_{k_2}(t)\,dt. \tag{2}$$

The channel for each user is considered as Rayleigh flat fading channel and ARMA processes can be adopted to model its time correlation with satisfaction [11, 15, 18]. Given an ARMA(r_1, r_2) process, the CSI of the kth user at the nth interval $a_{k,n}$ can be represented as

$$\begin{aligned} a_{n,k} + \phi_{k,1} a_{n-1,k} \cdots \phi_{k,r_1} a_{n-r_1,k} \\ = \rho_{k,0} v_{n,k} + \cdots + \rho_{k,r_2} v_{n-r_2,k}, \end{aligned} \tag{3}$$

where $v_{n,k}$ is an i.i.d. random complex Gaussian process that drives the ARMA process, $\{\phi_{k,1}, \ldots, \phi_{k,r_1}\}$ and $\{\rho_{k,1}, \ldots, \rho_{k,r_2}\}$ are the AR and MA coefficients of the model. We assume that we know a priori the second-order statistics of the underlying fading channel, and therefore the coefficients of the ARMA model can be precomputed so that the power spectral density of the ARMA process matches that of the fading channel. For convenience, we assume that $r_1 = r_2 = r$; otherwise zeros can be padded to the coefficients to make the orders equal.

An equivalent form of (1) consists of a set of sufficient statistics represented by the matched filter output,

$$y_{n,k} = \langle y(t), s_k(t) \rangle = \int_{(n-1)T}^{nT} y_n(t) s_k(t)\,dt. \tag{4}$$

The set of matched filter outputs $\mathbf{y}_n = [y_{n,1}, \ldots, y_{K,n}]^{\mathrm{T}}$, where $(\cdot)^{\mathrm{T}}$ stands for matrix transpose, can be represented in vector-matrix form as

$$\mathbf{y}_n = \mathbf{R}\mathbf{A}_n \mathbf{b}_n + \mathbf{u}_n, \tag{5}$$

where $\mathbf{A}_n = \operatorname{diag}\{a_{n,1}, \ldots, a_{n,K}\}$ is the diagonal matrix of the channel state information, $\mathbf{b}_n = [b_{n,1}, \ldots, b_{n,K}]^{\mathrm{T}}$ is the user date vector, and $\mathbf{u}_n$ is the complex Gaussian noise vector with independent real and imaginary components and with covariance matrix equal to $\sigma^2\mathbf{R}$. Our objective is to perform sequential symbol detection without knowing the CSI $a_{n,k}$, that is, blind multiuser detection.

3. TIME-OBSERVATION STATE-SPACE SYSTEM MODELING

A succinct mathematical representation of a time-varying system is the dynamic state-space model (DSSM). The state-space representation of CDMA systems in flat fading channels can be found in the existing literatures [11] and it can be expressed as

$$\begin{aligned} \mathbf{h}_{k,n} &= \mathbf{Q}_k \mathbf{h}_{k,n-1} + \mathbf{g} v_{k,n} \quad \forall k, \\ a_{k,n} &= \boldsymbol{\rho}_k^{\mathrm{T}} \mathbf{h}_{k,n} \quad \forall k, \\ \mathbf{y}_n &= \mathbf{R}\mathbf{A}_n \mathbf{b}_n + \mathbf{u}_n, \end{aligned} \tag{6}$$

where $\mathbf{h}_{n,k}^{\mathrm{T}} = [h_{n,k} \ \cdots \ h_{n-r,k}]$ is an $(r+1) \times 1$ channel state vector, $\boldsymbol{\rho}_k^{\mathrm{T}} = [\rho_{k,0} \ \cdots \ \rho_{k,r}]$,

$$\mathbf{Q}_k = \begin{bmatrix} -\phi_{k,1} & \cdots & -\phi_{k,r} & 0 \\ 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & 1 & 0 \end{bmatrix}, \qquad \mathbf{g} = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}. \tag{7}$$

In (6), $\mathbf{h}_{k,n}$ for all k and $\mathbf{b}_n$ are the unknowns to be estimated. Note that the observation $\mathbf{y}_n$ is not linear in $\mathbf{h}_{k,n}$ for all k and $\mathbf{b}_n$, and therefore the Kalman filter cannot provide the optimum solution. In fact, the optimum solution can be obtained by a so-called splitting Kalman filter, where, at time n, 2^n Kalman filters are required. The complexity of the splitting Kalman filter is exponential with both time and users and thus computational prohibited. Instead, particle filtering can be used to obtain good approximations of the optimum solution with reduced complexity. PF algorithms on (6) incorporated with Kalman filtering were proposed in [17]. However, as mentioned in the introduction, due to the structure of (6), particles of $\mathbf{b}_n$ must be sampled jointly, and the complexity becomes exponential with the number of users. The prohibitive complexity on large user systems implies that this PF algorithm is infeasible for practical applications. To circumvent this difficulty, in the following we introduce a time-observation state-space model (TOSSM) for the system:

$$\begin{aligned} p(\mathbf{b}_{1:N}, \bar{y}_{1:NK}) &= p(\bar{y}_{NK} | \mathbf{b}_{1:N}, \bar{y}_{1:NK-1}) p(\mathbf{b}_{1:N}, \bar{y}_{1:NK-1}) \\ &= p(\bar{y}_{NK} | \mathbf{b}_{1:N}, \bar{y}_{1:NK-1}) \\ &\quad \times p(b_{N,K} | \mathbf{b}_{N,1:K-1}, \mathbf{b}_{1:N-1}, \bar{y}_{1:NK-1}) \\ &\quad \times p(\mathbf{b}_{N,1:K-1}, \mathbf{b}_{1:N-1}, \bar{y}_{1:NK-1}) \\ &= p(\bar{y}_{NK} | \mathbf{b}_{1:N}, \bar{y}_{1:NK-1}) p(b_{N,K}) \\ &\quad \times p(\mathbf{b}_{N,1:K-1}, \mathbf{b}_{1:N-1}, \bar{y}_{1:NK-1}) \\ &= p(\bar{y}_{NK} | \mathbf{b}_{1:N}, \bar{y}_{1:NK-1}) \\ &\quad \times p(\mathbf{b}_{N,1:K-1}, \mathbf{b}_{1:N-1}, \bar{y}_{1:NK-1}). \end{aligned} \tag{8}$$

In developing the TOSSM, we start with the Cholesky factorization of the cross-correlation matrix $\mathbf{R}$ as

$$\mathbf{R} = \mathbf{F}^{\mathrm{T}} \mathbf{F}, \tag{9}$$

where $\mathbf{F}$ is a uniquely defined $K \times K$ lower triangular matrix. Now, right multiplying $(\mathbf{F}^{\mathrm{T}})^{-1}$ with the matched filter output, we obtain

$$\bar{\mathbf{y}}_n = (\mathbf{F}^{\mathrm{T}})^{-1} \mathbf{y}_n = \mathbf{F}\mathbf{A}_n \mathbf{b}_n + \bar{\mathbf{u}}_n \tag{10}$$

or, equivalently,

$$\bar{\mathbf{y}}_n = \mathbf{F}\mathbf{B}_n \mathbf{a}_n + \bar{\mathbf{u}}_n, \tag{11}$$

where $\mathbf{B}_n = \operatorname{diag}\{b_{n,1}, \ldots, b_{n,K}\}$ is the diagonal user data matrix, and $\mathbf{a}_n = [a_{n,1}, \ldots, a_{n,K}]$ is the $K \times 1$ vector of CSI. Since the covariance matrix of $\bar{\mathbf{u}}_n$ becomes $E[\bar{\mathbf{u}}_n \bar{\mathbf{u}}_n^{\mathrm{T}}] = \sigma^2 \mathbf{F}^{-\mathrm{T}} \mathbf{R} \mathbf{F}^{-1} = \sigma^2 \mathbf{I}$, where $\mathbf{I}$ is an identity matrix, $\bar{\mathbf{y}}_n$ is called the whitened matched filter (WMF) output. Next, define a tall channel vector of $K(r+1) \times 1$ as $\mathbf{h}_n = [\mathbf{h}_{1,n}^{\mathrm{T}} \cdots \mathbf{h}_{K,n}^{\mathrm{T}}]^{\mathrm{T}}$ and the channel transition becomes

$$\mathbf{h}_n = \mathbf{Q}\mathbf{h}_{n-1} + \mathbf{G}\mathbf{v}_n, \tag{12}$$

where $\mathbf{v}_n = [v_{1,n}, \ldots, v_{K,n}]^{\mathrm{T}}$, $\mathbf{Q} = \operatorname{diag}(\mathbf{Q}_1, \ldots, \mathbf{Q}_K)$, and $\mathbf{G} = \operatorname{diag}(\underbrace{\mathbf{g}, \ldots, \mathbf{g}}_{K})$ are $K(r+1) \times K(r+1)$ and $K(r+1) \times K$ matrices. We can thus express $\mathbf{a}_n$ by $\mathbf{h}_n$ in a compact form by

$$\mathbf{a}_n = \mathbf{P}\mathbf{h}_n, \tag{13}$$

where $\mathbf{P} = \operatorname{diag}(\boldsymbol{\rho}_1^{\mathrm{T}}, \ldots, \boldsymbol{\rho}_K^{\mathrm{T}})$ is of dimension $K \times K(r+1)$. Now by replacing $\mathbf{a}_n$ in (11) by (13), we have

$$\bar{\mathbf{y}}_n = \mathbf{F}\mathbf{B}_n \mathbf{P}\mathbf{h}_n + \bar{\mathbf{u}}_n. \tag{14}$$

If we denote the kth row of $\mathbf{F}$ by $\mathbf{f}_k^{\mathrm{T}}$, the kth WMF output $\bar{\mathbf{y}}_n$ can be written as

$$\bar{y}_{n,k} = \mathbf{f}_k^{\mathrm{T}} \mathbf{B}_n \mathbf{P}\mathbf{h}_n + \bar{u}_{n,k}, \tag{15}$$

where $\bar{u}_{n,k}$ is the kth element of $\bar{\mathbf{u}}_n$. Now, instead of considering the system evolving only along time, we imagine a system progressing alternately along the path of time and the WMF observations $\bar{y}_{n,k}$. The concept is further illustrated in Figure 1. To describe this new system, we must collapse the time index n and the observation index k into one time-observation index l, where $l = (n-1)K + k$. This conversion is reversible or, in other words, we can also calculate k and n from l by $k = \operatorname{mod}(l, K)$ and $n = (l-k)/K + 1$, where $\operatorname{mod}(k, K)$ is the k modulo K operation. In the following description of the TOSSM indexed by l, all k and n are assumed to be obtained from the corresponding l. Now, we introduce a $K \times K$ auxiliary matrix $\widetilde{\mathbf{B}}_l = \operatorname{diag}\{b_{n,1}, \ldots, b_{n,k}, 0 \ldots, 0\}$. The state-space representation for the new time-observation system indexed by l can be then constructed as

$$\begin{aligned} \mathbf{h}_l &= \begin{cases} \mathbf{Q}\mathbf{h}_{l-1} + \mathbf{G}\mathbf{v}_l & \text{if } k = 1, \\ \mathbf{h}_{l-1} & \text{if } k \neq 1, \end{cases} \\ \bar{y}_l &= \mathbf{f}_k^{\mathrm{T}} \widetilde{\mathbf{B}}_l \mathbf{P}\mathbf{h}_l + \bar{u}_l \end{aligned} \tag{16}$$

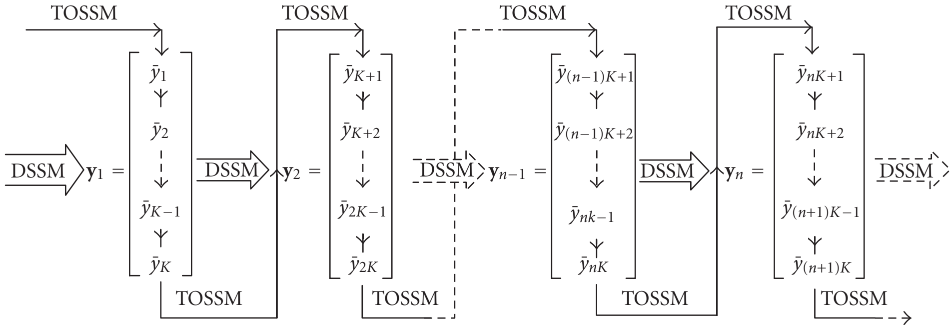


Figure 1: Illustrative plot of the TOSSM.

and we call (16) the TOSSM. Note that (16) and (6) describe the same system. There are, however, key differences between the two models. Unlike (6), the state transitions of $\mathbf{h}_l$ in the TOSSM are time (or index) varying, that is, at different l, different transition is applied. Specifically, when $k = 1$ or, equivalently, n increases by 1, $\mathbf{h}_l$ updates according to the ARMA channel model, and otherwise when $k \neq 1$ and n remains unchanged from $l-1$, $\mathbf{h}_l$ is assumed to be static. Additionally, in the TOSSM, the number of the unknown user bits changes with l and especially, only one new unknown signal $b_{n,k}$ is included each time when l is incremented by one. Therefore, if we assume perfect detection at $l-1$, that is, $b_{n,1}, \ldots,$ and $b_{n,k-1}$ are known exactly, then there is only one unknown user bit to be detected. Note that in the conventional DSSM (6), K unknown users bits need to be detected altogether as the system evolves to time n. This is the key of the model that leads to efficient particle filtering solutions. We, however, want to stress that the decision on $b_{n,k}$ (except $k = K$) is not finalized at l. Since the observations from y_{l+1} up to y_{l+r} with $r = K - k$ all contain information about $b_{n,k}$, the final decision is reached only at $l + r$, or in general, when $k = K$.

4. OPTIMUM BAYESIAN BLIND DETECTION

In a Bayesian framework, the optimum decision on $\mathbf{b}_N$ can be obtained by the marginalized posterior mode (MPM) criterion, which is expressed as

$$(\hat{b}_{N,k})_{\text{MPM}} = \operatorname{sgn}\left(\sum_{\mathbf{b}_N \in \{-1,1\}^K} b_{n,k}\, p(\mathbf{b}_N \mid \bar{y}_{1:NK})\right), \qquad (17)$$

where $p(\mathbf{b}_N \mid \bar{y}_{1:NK})$ is the posterior distribution that is essential for computing (17) and the subscript $1:NK$ denotes a collection of the variable indexed from 1 to NK, e.g., $\bar{y}_{1:NK} = \{\bar{y}_1, \ldots, \bar{y}_{NK}\}$. Notice that the posterior distribution $p(\mathbf{b}_N \mid \bar{y}_{1:NK})$ is independent of $\mathbf{b}_{1:(N-1)}$, that is, the bits transmitted prior to time n. Further, the marginalization in (17) suggests that $(\hat{b}_{N,k})_{\text{MPM}}$ is also independent of other users' bits transmitted at n. Therefore, the MPM solution is immune to decision errors on $\mathbf{b}_{1:(N-1)}$ and other users' bits transmitted at n.

Now, to derive $p(\mathbf{b}_n \mid \bar{y}_{1:NK})$, marginalization on $p(\mathbf{b}_{1:N} \mid \bar{y}_{1:NK})$ over $\mathbf{b}_{1:(N-1)}$ is needed, that is,

$$\begin{aligned} p(\mathbf{b}_N \mid \bar{y}_{1:NK}) &= \sum_{\mathbf{b}_{1:N-1}} p(\mathbf{b}_{1:N} \mid \bar{y}_{1:NK}) \\ &= \frac{\sum_{\mathbf{b}_{1:N-1}} p(\mathbf{b}_{1:N}, \bar{y}_{1:NK})}{\sum_{\mathbf{b}_{1:N}} p(\mathbf{b}_{1:N}, \bar{y}_{1:NK})}. \end{aligned} \qquad (18)$$

Considering the TOSSM (16), we found the joint distribution in (8), where the last equation was obtained by assuming the noninformative priors for $\mathbf{b}_{N,K}$, that is, $p(b_{N,K} = 1) = 0.5$. Equation (8) indicates a recursive calculation of $p(\mathbf{b}_{1:N}, \bar{y}_{1:NK})$ from $l = 1$ to NK through multiplying the marginal likelihood $p(\bar{y}_l \mid \mathbf{b}_{n,1:k}, \mathbf{b}_{1:n-1}, \bar{y}_{1:l})$ at each recursion. These likelihoods $p(\bar{y}_l \mid \mathbf{b}_{n,1:k}, \mathbf{b}_{1:n-1}, \bar{y}_{1:l})$ for $l = 1, \ldots, NK$ are obtained by marginalizing the channel state vector $\mathbf{h}_l$ from $p(\bar{y}_l, \mathbf{h}_l \mid \mathbf{b}_{n,1:k}, \mathbf{b}_{1:n-1}, \bar{y}_{1:l})$, and we show in the appendix that

$$p(\bar{y}_l \mid \mathbf{b}_{n,1:k}, \mathbf{b}_{1:n-1}, \bar{y}_{1:l}) = \mathcal{N}(m_l, c_l) \qquad (19)$$

and the mean m_l and variance c_l can be calculated sequentially through the Kalman filter. This is equivalent to say that $p(\bar{y}_l \mid \mathbf{b}_{n,1:k}, \mathbf{b}_{1:n-1})$ can be calculated from a run of the Kalman filter. Now, revisiting (18), we see that, to calculate $p(\mathbf{b}_N \mid \bar{y}_{1:NK})$, $p(\bar{y}_l \mid \mathbf{b}_{n,1:k}, \mathbf{b}_{1:n-1})$ must be evaluated for 2^{NK} combinations of $\mathbf{b}_{1:N}$, or 2^{NK} Kalman filters are needed, each of which corresponding to one possible combination. As a result, totally 2^{NK} Kalman filters are required for the MPM solution. The expansion of the numbers of the Kalman filters with l presents a tree structure illustrated in Figure 2. The MPM solution has thus a complexity exponentially increasing with both time n and the number of users K. This is apparently a formidable task not possible for real applications. We, therefore, must resort to suboptimum solutions with manageable complexity. One choice is particle filtering.

5. A DECISION-DIRECTED APPROACH TO BLIND MUD

A decision-directed approach to blind MUD was proposed in [11] based on DSSM (6). We describe in the following a corresponding decision-directed approach on the TOSSM (16).

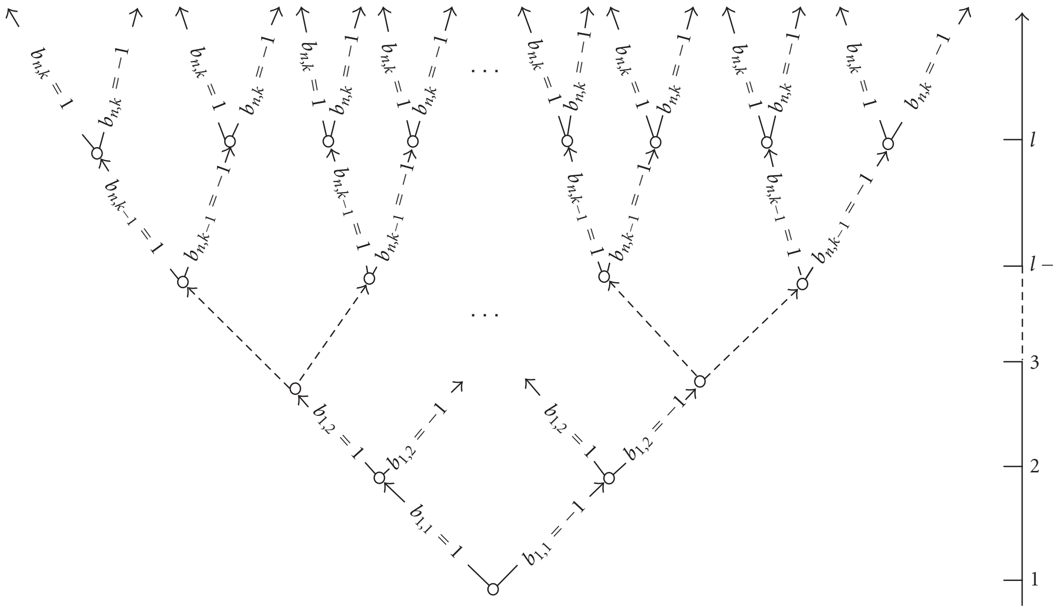


Figure 2: The tree structure of the optimum solution. Each path in the tree represents a run of the Kalman filter.

Predictive step:

$$\widetilde{\mathbf{h}}_l = \begin{cases} \mathbf{Q}\widehat{\mathbf{h}}_{l-1} + \mathbf{G}\mathbf{u}_l & \text{if } k = 1, \\ \widehat{\mathbf{h}}_{l-1} & \text{if } k \neq 1, \end{cases}$$

$$\widehat{\boldsymbol{\Sigma}}_l = \begin{cases} \mathbf{Q}\widehat{\boldsymbol{\Xi}}_{l-1}\mathbf{Q}^{\mathrm{T}} + \sigma^2\mathbf{G}\mathbf{G}^{\top} & \text{if } k = 1, \\ \widehat{\boldsymbol{\Xi}}_{l-1} & \text{if } k \neq 1. \end{cases}$$

Detection step:

$$\widehat{b}_{n,k} = \operatorname{sgn}(z_{n,k});$$
$$z_{n,k} = \Big(\bar{y}_l - \sum_{j=1}^{k-1} f_{k,j} a_{i,j} b_{i,j}\Big) a_{i,k}^*;$$
$$a_{i,k} = \boldsymbol{\rho}_k \widetilde{\mathbf{h}}_l.$$

Update step:

$$\widehat{\mathbf{K}}_l = \widehat{\boldsymbol{\Sigma}}_l \widehat{\mathbf{C}}_l^{\mathrm{H}} / \widehat{c}_l \text{ with } \widehat{c}_l = \widehat{\mathbf{C}}_l \widehat{\boldsymbol{\Sigma}}_l \widehat{\mathbf{C}}_l^{\mathrm{H}} + \sigma^2,$$
$$\widehat{\mathbf{h}}_l = \widetilde{\mathbf{h}}_l + \widehat{\mathbf{K}}_l(\bar{y}_l - \widehat{\mathbf{C}}_l \mathbf{h}_l),$$
$$\widehat{\boldsymbol{\Xi}}_l = (\mathbf{I} - \widehat{\mathbf{K}}_l \widehat{\mathbf{C}}_l)\widehat{\boldsymbol{\Sigma}}_l,$$

where $\widehat{\mathbf{C}}_l = \mathbf{f}_k^{\mathrm{T}} \widehat{\mathbf{B}}_l \mathbf{P}$ and $\widehat{\mathbf{B}}_l = \operatorname{diag}\{[\widehat{b}_{n,1}, \dots, \widehat{b}_{n,k}, 0 \dots, 0]\}$.

Algorithm 1: Decision-directed detector (DD).

One distinct feature of the decision-directed approach on the TOSSM is that the decision on only one user's bit is made at each l. Specifically, let $\widehat{b}_{n,k-1}$ and $\widehat{\mathbf{h}}_{l-1}$ represent the decisions on $b_{n,k-1}$ and $\mathbf{h}_{l-1}$ at $l-1$, then the decision-directed approach at l can be summarized in Algorithm 1. Clearly, the above decision-directed algorithm is equivalent to one run of the Kalman filter, and therefore it is a lot simpler than the optimum MPM solution. Nevertheless, the user bit is determined based on the prediction of the channel states and the decisions on previous users' bits, and thus it is not optimum. Compared with the algorithm based on DSSM (6), at time k with k from 1 to K, the above algorithm makes a decision on one user at a time and updates the channel state vector $\mathbf{h}_l$ whenever a decision is reached. The updated $\mathbf{h}_l$ will then influence the decision on $b_{n,k+1}$. Therefore, in both a good and a bad way, decisions at early stages (smaller k) would have more impact on decisions at later stages (larger k) than those made by the algorithm on DSSM. If detection error exists in early stages, they will be propagated into later stages. It is therefore beneficial to rank the users according to the estimated SNR. The performance of the decision-directed algorithm is, however, ultimately limited by error propagation.

6. PARTICLE FILTERING DETECTOR FOR BLIND MUD

Particle filtering belongs to the family of Monte Carlo sampling which aims at using samples to approximate posterior distribution. However, particle filtering distinguishes itself by employing a sequential importance sampling scheme, and in particular, it is designed for nonlinear and non-Gaussian systems described through state-space modeling such as the problem concerned.

In the context of the proposed problem, when $\mathbf{y}_N$, or equivalently $\bar{\mathbf{y}}_N$, is observed at time N, the objective of particle filtering is to draw, say, J weighted random samples $\{\mathbf{b}_{1:N}^{(j)}, w_{NK}^{(j)}\}_{j=1}^{J}$ from $p(\mathbf{b}_{1:N} \mid \bar{y}_{1:nK})$, where $w_{NK}^{(j)}$ is the weight of the jth sample $\mathbf{b}_{1:N}^{(j)}$. With the samples, $p(\mathbf{b}_N \mid \bar{y}_{1:NK})$ can be approximated by

$$p(\mathbf{b}_{1:N} \mid \bar{y}_{1:nK}) \approx \sum_{j=1}^{J} w_{NK}^{(j)} \prod_{l=1}^{NK} \delta\big(b_{n,k} - b_{n,k}^{(j)}\big), \tag{20}$$

where $\delta(\cdot)$ is the Dirac delta function, and hence the MPM solution of $\mathbf{b}$ by a simple weighted summation is

$$(\hat{b}_{N,k})_{\text{MPM}} \approx \text{sgn}\left(\sum_{j=1}^{J} w_{NK}^{(j)} b_{N,k}^{(j)}\right) \tag{21}$$

for $k = 1, \ldots, K$. By the law of large numbers, the approximation will converge to the true MPM solution with the increase of the number of samples J. If these samples are taken directly from the posterior distribution, then all the samples have equal weights. However, direct sampling from $p(\mathbf{b}_{1:N} | \bar{y}_{1:NK})$ is prohibited since all possible combinations of $\mathbf{b}_{1:N}$ must be evaluated on $p(\mathbf{b}_{1:N} | \bar{y}_{1:NK})$, which again requires 2^{NK} Kalman filters. To circumvent the difficulty, importance sampling is performed where samples are taken from a proposal importance function $\pi(\mathbf{b}_{1:KN} | \bar{y}_{1:KN})$ and weighted according to

$$w_{KN}^{(j)} = \frac{p(\mathbf{b}_{1:KN}^{(j)} | \bar{y}_{1:KN})}{\pi(\mathbf{b}_{1:KN}^{(j)} | \bar{y}_{1:KN})} \quad \forall j. \tag{22}$$

Notice that $\pi(\mathbf{b}_{1:KN} | \bar{y}_{1:KN})$ is a very high-dimensional distribution and it is burdensome to sample the variables and calculate the weights altogether. Fortunately, the TOSSM allows a Markovian factorization on the posterior distribution as

$$\begin{aligned} p(\mathbf{b}_{1:N}, \bar{y}_{1:NK}) &\propto p(\bar{y}_{NK} | \mathbf{b}_{1:N}, \bar{y}_{1:NK-1}) p(b_{N,K}) \\ &\quad \times p(\mathbf{b}_{N,1:K-1}, \mathbf{b}_{1:N-1} | \bar{y}_{1:NK-1}) \\ &= p(\bar{y}_{NK} | \mathbf{b}_{1:N}, \bar{y}_{1:NK-1}) \\ &\quad \times p(\mathbf{b}_{N,1:K-1}, \mathbf{b}_{1:N-1} | \bar{y}_{1:NK-1}). \end{aligned} \tag{23}$$

Then, if we choose the importance distribution as

$$\begin{aligned} \pi(\mathbf{b}_{1:N} | \bar{\mathbf{y}}_{1:NK}) &= p(b_{N,k} | \mathbf{b}_{N,1:K-1}, \mathbf{b}_{1:N-1}, \bar{y}_{1:NK}) \\ &\quad \times \pi(\mathbf{b}_{N,1:K-1}, \mathbf{b}_{1:N-1} | \bar{y}_{1:NK-1}), \end{aligned} \tag{24}$$

the weight can be calculated by

$$\begin{aligned} w_{KN}^{(j)} &= \frac{p(\bar{y}_{NK} | \mathbf{b}_{1:N}^{(j)}, \bar{y}_{1:NK-1}) p(b_{N,K}^{(j)})}{p(b_{N,K}^{(j)} | \mathbf{b}_{N,1:K-1}^{(j)}, \mathbf{b}_{1:N-1}^{(j)}, \bar{y}_{1:NK})} \\ &\quad \times \frac{p(\mathbf{b}_{N,1:K-1}^{(j)}, \mathbf{b}_{1:N-1}^{(j)} | \bar{y}_{1:NK-1})}{\pi(\mathbf{b}_{N,1:K-1}^{(j)}, \mathbf{b}_{1:N-1}^{(j)} | \bar{y}_{1:NK-1})} \\ &= \frac{p(\bar{y}_{NK} | \mathbf{b}_{1:N}^{(j)}, \bar{y}_{1:NK-1}) p(b_{N,K}^{(j)})}{p(b_{N,K}^{(j)} | \mathbf{b}_{N,1:K-1}^{(j)}, \mathbf{b}_{1:N-1}, \bar{y}_{1:NK})} w_{KN-1}^{(j)} \\ &\propto p(\bar{y}_{NK} | \mathbf{b}_{N,1:K-1}^{(j)}, \mathbf{b}_{1:N-1}^{(j)}, \bar{y}_{1:NK-1}) w_{KN-1}^{(j)} \\ &= \sum_{b_{N,K}} p(\bar{y}_{NK} | \mathbf{b}_{1:N}^{(j)}, \bar{y}_{1:NK-1}) w_{KN-1}^{(j)} \\ &= \mu_{KN-1}^{(j)} w_{KN-1}^{(j)}, \end{aligned} \tag{25}$$

where $\mu_{KN-1}^{(j)}$ is the weight update factor. Examining (24) and (25), we find that given $w_{KN-1}^{(j)}$ and $p(\mathbf{b}_{N,1:K-1}, \mathbf{b}_{1:N-1} | \bar{y}_{1:NK-1})$, the importance function (24) and the weights (25) are known exactly as long as $p(\bar{y}_{NK} | \mathbf{b}_{1:N}^{(j)}, \bar{y}_{1:NK-1})$ can be derived. In fact, we have indicated in Section 4 that $p(\bar{y}_{NK} | \mathbf{b}_{1:N}, \bar{y}_{1:NK-1})$ can be calculated through the Kalman filter as

$$\begin{aligned} \lambda_{NK}(i) &= p(\bar{y}_{NK} | b_{N,K} = 2 * i - 3, \mathbf{b}_{N,1:K}^{(j)}, \mathbf{b}_{1:N}^{(j)}, \bar{y}_{1:NK-1}) \\ &= \mathcal{N}_c(m_{NK}^{(j)}(i), c_{NK}^{(j)}(i)) \end{aligned} \tag{26}$$

for $i = 1, 2$ where $m_l^{(j)}(i)$ and $c_l^{(j)}(i)$ are calculated the same way as shown in the appendix but for a set of $\mathbf{b}_{1:NK}$ given in (26). We can therefore obtain samples and weights using a recursive algorithm. To put the idea in concrete procedure, we assume that at $l - 1$, we have obtained from a previous recursion the trajectories (samples) $\{\mathbf{b}_{0:l-1}^{(j)}\}_{j=1}^{J}$ appropriately weighted with the weights $\{w_{l-1}^{(j)}\}_{j=1}^{J}$. Using the recent observations $\bar{y}_l$, we update the trajectories and weights as in Algorithm 2. This process of recursively obtaining particles is called particle filtering. After each recursion, the mean $\boldsymbol{\eta}_l^{(j)}$ and covariance vectors $\Xi_l^{(j)}$ are passed on to the next recursion. From (21), we also see that to calculate all the elements of $\{\mathbf{b}_N\}_{\text{MPM}}$, $w_{NK}^{(j)}$ is required. Therefore the decision on all the elements can only be made after recursion $l = KN$ and the particles for $b_{N,k}$ for $k = 1, 2, \ldots, K - 1$ must be stored.

In the above derivation of particle filtering, the adopted importance function is known as optimum in the sense that minimizes the variance of the weights. The above particle filtering procedure suffers from particle impoverishment, that is, after several recursions, some weights of the samples become negligible and stop contributing to the overall evaluation. To prevent it, we insert a residue resampling step [15] after every fixed recursion. Particularly, during the resampling at recursion l, the particles for $\mathbf{b}_{n,1:k}^{(j)}$, the mean vectors, and covariance matrices must be treated as a set in the resampling process.

7. STOCHASTIC *M*-DETECTOR FOR BLIND MUD

Recently, a every efficient particle filtering algorithm called stochastic M-algorithm (SMA) was proposed in [19] for problems with discrete unknowns. SMA can provide similar performance as generic particle filtering but with much reduced complexity. SMA can be considered as a particle filtering algorithm with the discrete delta functions as importance functions. In addition, each trajectory produces two samples (-1 and 1) for the binary case rather than one sample as in the generic PF. A key feature with SMA is that no two trajectories are identical, which is however rarely true with the generic PF. As a result, the SMA can provide more sample diversities with less trajectories than the generic PF. Nonetheless, notice that the number of trajectories doubles after each sampling and therefore a selection step is required

For $j = 1$ to J, do as follows.

(1) Predictive step:

Calculate

$$\boldsymbol{\mu}_l^{(j)} = \begin{cases} \mathbf{Q}\boldsymbol{\eta}_{l-1}^{(j)} & \text{if } k = 1, \\ \boldsymbol{\eta}_{l-1}^{(j)} & \text{if } k \neq 1, \end{cases}$$

and

$$\boldsymbol{\Sigma}_l^{(j)} = \begin{cases} \mathbf{Q}\boldsymbol{\Xi}_{l-1}^{j}\mathbf{Q}^\top + \sigma^2\mathbf{G}\mathbf{G}^\top & \text{if } k = 1, \\ \boldsymbol{\Xi}_{l-1}^{(j)} & \text{if } k \neq 1. \end{cases}$$

(2) Sampling step.

(a) For $i = 1$ and -1, calculate

(i) $m_l^{(j)}(i) = \mathbf{c}_l^{(j)}(i)\boldsymbol{\mu}_l^{(j)}$ and

$c_l^{(j)}(i) = \mathbf{c}_l(i)\boldsymbol{\Sigma}_l^{j}\mathbf{c}_l^{(j)\mathrm{H}}(i) + \sigma^2$,

where $\mathbf{c}_l^{(j)} = \mathbf{f}_k^{\mathrm{T}}\mathbf{B}_l^{(j)}(i)\mathbf{P}$, $\mathbf{B}_l^{(j)}(i) = \operatorname{diag}\{b_{n,1}^{(j)}, \ldots, b_{n,k-1}^{(j)}, i, 0, \ldots, 0\}$;

(ii) $\lambda_l^{(j)}(i) = \mathcal{N}_c(m_l^{(j)}(i), c_l^{(j)}(i))$.

(b) Sample $m \in \{-1, 1\}$ with probability proportional to $\lambda_l^{(j)}(i)\ \forall i$.

(c) Set $b_l^{(j)} = m$.

(d) Calculate $\mu_l^{(j)} = \sum_{i\in\{-1,1\}} \lambda_l^{(j)}(i)$ and the unnormalized weight $\bar{w}_l^{(j)} = \mu_l^{(j)} w_{l-1}^{(j)}$.

(3) Updating step. Calculate

(i) $\mathbf{K}_l^{(j)} = \boldsymbol{\Sigma}_l^{(j)}\mathbf{c}^{(j)}(m)_l^{\mathrm{H}}/c_l^{j}(m)$;

(ii) $\boldsymbol{\eta}_l^{(j)} = \boldsymbol{\mu}_l^{(j)} + \mathbf{K}_l(\bar{y}_l - \mathbf{c}_l^{(j)}(m)\boldsymbol{\mu}_l^{(j)})$;

(iii) $\boldsymbol{\Xi}_l^{(j)} = (\mathbf{I} - \mathbf{K}_l^{(j)}\mathbf{c}_l^{(j)}(m))\boldsymbol{\Sigma}_l^{(j)}$.

Form the new trajectories $\mathbf{b}_{0:l}^{(j)} = \{b_l^{(j)}, \mathbf{b}_{0:l-1}^{(j)}\}\ \forall j$.

Normalize the weight as $w_l^{(j)} = \bar{w}_l^{(j)}/\sum_{j=1}^{J}\bar{w}_l^{(j)}$.

Algorithm 2: Particle filtering detector (PFD).

Trajectory expansion

(1) For $j = 1$ to J,

(i) perform the predictive step in the PFD Algorithm;

(ii) perform (2)(a) in Algorithm PFD;

(iii) set $b_l^{(2j-1)} = 1$ and calculate the weight by $\bar{w}_l^{(2j-1)} = \lambda_l^{(j)}(1)w_{l-1}^{(j)}$;

(iv) set $b_l^{(2j)} = -1$ and calculate the weight by $\bar{w}_l^{(2j-1)} = \lambda_l^{(j)}(-1)w_{l-1}^{(j)}$;

(v) form $2J$ new trajectories by setting $\mathbf{b}_l^{(2j-1)} = \{b_l^{(2j-1)}, \mathbf{b}_{0:l}^{(j)}\}$ and $\mathbf{b}_l^{(2j)} = \{b_l^{(2j)}, \mathbf{b}_{0:l}^{(j)}\}$.

(2) Normalize the weights $\bar{w}_k^{(j)}$ to obtain $w_k^{(j)}$.

(3) Trajectory selection: select J trajectories from $2M$ trajectories using the optimal resampling algorithm.

(4) Updating step: for $j = 1$ to J; perform the updating step in the PFD Algorithm.

Algorithm 3: Stochastic M detector (SMD).

to avoid the exponential increase of trajectories. Here, we use the optimal resampling algorithm [20] since it is a sampling-without-replacement algorithm and does not produce replicates of the same trajectories, the feature that is required by SMA. The SMA for the problem concerned at the lth recursion is outlined as in Algorithm 3.

The structure of the SMA resembles the popular M-algorithm. However, since the SMA is still a PF algorithm, it can provide probability information about the unknowns and thus can be applied to iterative MUD of a coded system.

7.1. Discussion on the MPM, decision-directed, and particle filtering solutions

Comparing the PFD and the SMD with the decision-directed algorithm, we see that the processes along each trajectory is almost as identical as a decision-directed algorithm except that a sampling step is used in the place of the detection step, and they all resemble one run of Kalman filter which corresponds to a path in the tree of Figure 2. There are two paths going out at every note in the tree, and in selecting a path, the decision-directed algorithm uses a deterministic approach, while PFD and SMD adopt a soft measure which is based on probability. What is more, each trajectory is also associated with a weight which indicates the significance of the trajectory in final decision. Although trajectories with small weight do not seem to contribute much to current decision making at the present stage, they, however, might flourish in later recursions and carry significant weights in decision. The soft measure can apparently prevent current decision errors from greatly influencing the future decision, a key advantage over the decision-directed approach.

Comparing the PFD and the SMD with the optimum MPM solution, PFD, especially the SMD, has clear edge in complexity since it only maintains J trajectories or equivalently J Kalman filter at all times, but the required Kalman filter for the MPM grows exponentially with time. Further, the PFD and the SMD achieve every effective and efficient approximation to the true posterior distribution and therefore provide decision performance closer to optimum. Since the two detectors produce soft (probabilistic) results, they are readily applied in turbo MUD.

8. SIMULATION RESULTS

In this section, the bit error rate (BER) performance of the proposed PFDs and SMDs are studied through experiments. In all the experiments, the transmitted signal was differential BPSK modulated. The number of users was 15. For the PFDs, 151 trajectories were maintained, whereas 4 and 32 trajectories were tested for SMDs. Further, an AR model was adopted for the fading process, which was normalized to have a unit power, and thus the signal-to-noise ratio (SNR) was obtained by $10\log(1/\sigma^2)$.

In Figure 3, we provide the BER versus SNR for the different algorithms on a scenario of $\Omega_d = 0.03$. The genie-aided detector is included as a lower bound. We notice that the PFDs and SMDs with 32 trajectories are of the same

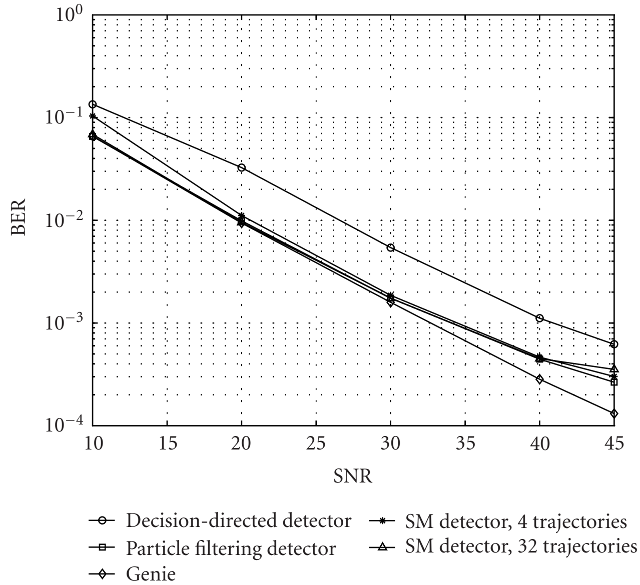


FIGURE 3: BERs versus SNR performance for various detectors. $\Omega = 0.03$.

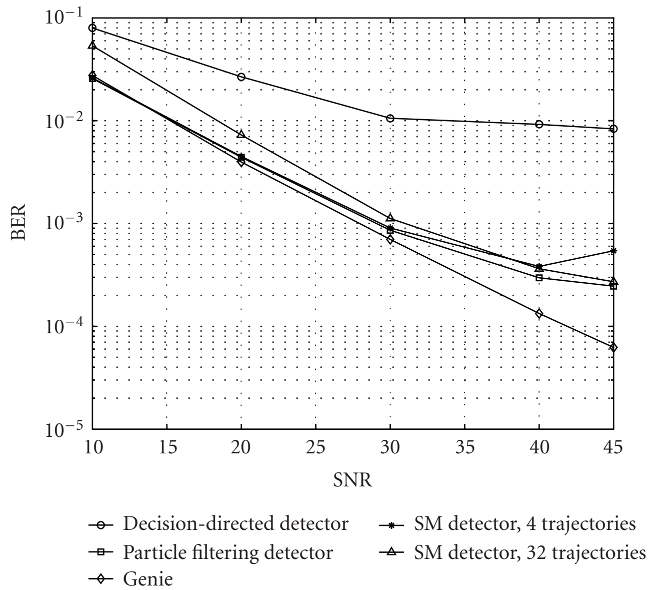


FIGURE 5: BERs versus SNR performance for various detectors for users with different power. $\Omega = 0.03$.

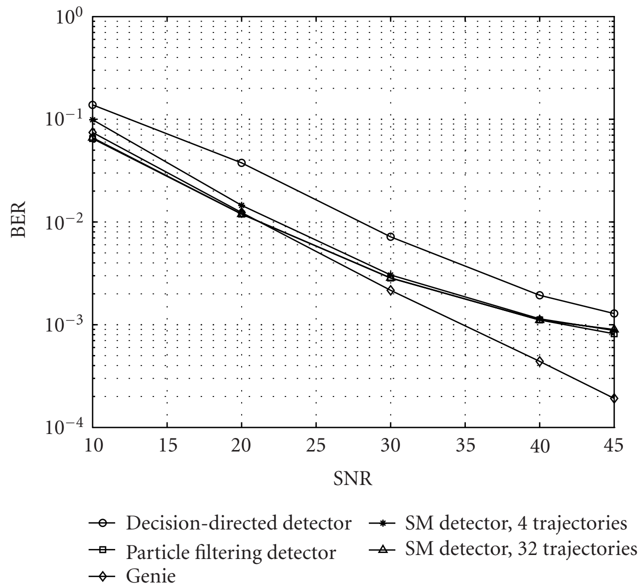


FIGURE 4: BERs versus SNR performance for various detectors. $\Omega = 0.05$.

order of magnitude as that of the genie-aided detector at low SNR (less than 30 dB). On the other hand, the results obtained by the SMDs with 4 and 32 trajectories are very close, especially after 20 dB, and comparable to that of the PFD. The SMD with 4 trajectories is obviously more favorable since it requires only about 1/35 of complexity of the PFD. As a final note, the PFD and SMDs achieve about 7 dB gain over the decision-directed detectors at 10^{-3} BER. In Figure 4, we provide the BER versus SNR performance for a higher Doppler frequency of $\Omega_d = 0.05$. Similar observations can be drawn as for the previous case even though the overall performance of the detectors is worse, which is reasonable considering that the channels are fading faster.

In Figure 5, we provide the BER versus SNR of the first user for the different algorithms on a scenario of $\Omega_d = 0.03$. In addition, the users have different power. The difference between the power of the first user and that of the last user is 10 dB and the other users' powers are equally spaced in between. The genie-aided detector is also included as a lower bound. In this case, the PFDs and SMDs with 32 trajectories are approximately of the same order of magnitude as that of the genie-aided detector at SNRs of the first user less than 30 dB. As in the case of equal power, the results obtained by the SMDs with 4 and 32 trajectories are very close, especially after 30 dB, and comparable to that of the PFD. Again, the SMD with 4 trajectories is obviously more favorable since it requires only about 1/35 of complexity of the PFD. In this experiment, the performance of the decision-directed detector is much worse compared to the performance of the PDF and SMDs. For example, the latter achieves about 11 dB gain over the former at 10^{-2} BER. In Figure 6, we provide the BER versus SNR performance for a Doppler frequency of $\Omega_d = 0.05$. Since the channels considered are fading faster, the performance of the detectors is worse. However, in general, similar observations to the tested detectors can be drawn. It is important to outline that the performance of the decision-directed detector gets worse in this case, for example, the PFD and SMDs achieve about 20 dB gain over the decision-directed detectors at 10^{-2} BER.

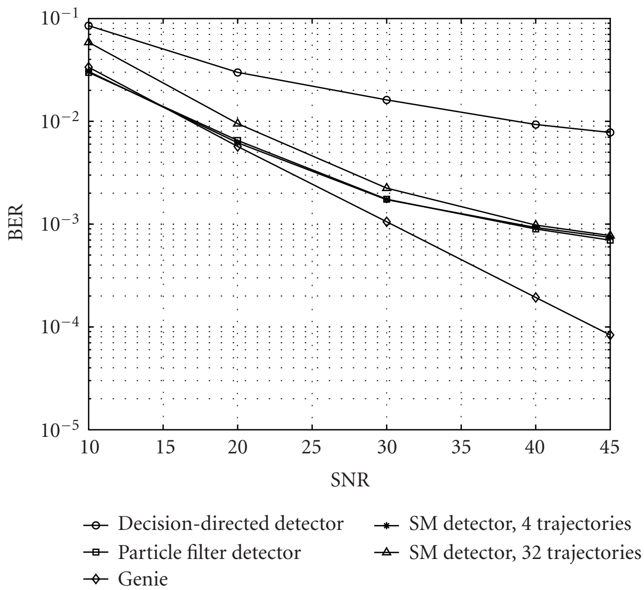


FIGURE 6: BERs versus SNR performance for various detectors for users with different power. $\Omega = 0.05$.

9. CONCLUSION

In this paper, we proposed to solve blind MUD over flat fast fading channels. We constructed a novel time-observation state-space model, based on which efficient particle filtering and stochastic M detectors were proposed. Particularly, the detectors based on the SMA demonstrated greater potential than those using generic PF. The former can provide comparable performance as the latter but with much smaller complexity.

APPENDIX

DERIVATION OF THE LIKELIHOOD $p(\bar{y}_l|\mathbf{b}_{n,1:k}, \mathbf{b}_{1:n-1}, \bar{y}_{1:l})$

The likelihood $p(\bar{y}_l|\mathbf{b}_{n,1:k}, \mathbf{b}_{1:n-1}, \bar{y}_{1:l})$ can be obtained as

$$\begin{aligned} p(\bar{y}_l|\mathbf{b}_{n,1:k}, \mathbf{b}_{1:n-1}, \bar{y}_{1:l-1}) &= \int p(\bar{y}_l, \mathbf{h}_l|\mathbf{b}_{n,1:k}, \mathbf{b}_{1:n-1}, \bar{y}_{1:l})\, d\mathbf{h}_t \\ &= \int p(\bar{y}_l|\mathbf{h}_l, \mathbf{b}_{n,1:k})\, p(\mathbf{h}_l|\mathbf{b}_{n,1:k-1}, \mathbf{b}_{1:n-1}, \bar{y}_{1:l-1})\, d\mathbf{h}_l, \end{aligned} \tag{A.1}$$

where the last equality is arrived by the fact that, given $\mathbf{h}_l$, and $\mathbf{b}_{n,1:k}$, $\bar{y}_l$ is independent of other variables, and $\mathbf{h}_l$ is independent of $b_{n,k}$. In (A.1), two distributions are involved in the integral. The first distribution is the likelihood defined by the observation equation which is

$$p(\bar{y}_l|\mathbf{h}_l, \mathbf{b}_{n,1:k}) = \mathcal{N}(\mathbf{C}_l\mathbf{h}_l, \sigma^2), \tag{A.2}$$

where $\mathbf{C}_l = \mathbf{f}_k^{\top}\widetilde{\mathbf{B}}_l\mathbf{P}$. The second distribution $p(\mathbf{h}_l|\mathbf{b}_{n,1:k-1}, \mathbf{b}_{1:n-1}, \bar{y}_{1:l-1})$ is the predictive density which can be obtained from the predictive step of the Kalman filter [21, 22], that is,

$$p(\mathbf{h}_l|\mathbf{b}_{n,1:k-1}, \mathbf{b}_{1:n-1}, \bar{y}_{1:l-1}) = \mathcal{N}(\boldsymbol{\mu}_l, \boldsymbol{\Sigma}_l), \tag{A.3}$$

where

$$\boldsymbol{\mu}_l = \begin{cases} \mathbf{Q}\boldsymbol{\eta}_{l-1} & \text{if } k = 1, \\ \boldsymbol{\eta}_{l-1} & \text{if } k \neq 1, \end{cases} \tag{A.4}$$

and

$$\boldsymbol{\Sigma}_l = \begin{cases} \mathbf{Q}\Xi_{l-1}\mathbf{Q}^{\top} + \sigma^2\mathbf{G}\mathbf{G}^{\top} & \text{if } k = 1, \\ \Xi_{l-1} & \text{if } k \neq 1. \end{cases} \tag{A.5}$$

In (A.4) and (A.5), $\boldsymbol{\eta}_{l-1}$ and Ξ_{l-1} are computed from the update steps of the Kalman filter expressed in terms of l as

$$\boldsymbol{\eta}_l = \boldsymbol{\mu}_l + \mathbf{K}_l(\bar{y}_l - m_l), \tag{A.6}$$

and

$$\Xi_l = (\mathbf{I} - \mathbf{K}_l\mathbf{C}_l)\boldsymbol{\Sigma}_l, \tag{A.7}$$

where $m_l = \mathbf{C}_l\boldsymbol{\mu}_l$ and $\mathbf{K}_l = \boldsymbol{\Sigma}_l\mathbf{C}_l^{\mathrm{H}}/c_l$ with $c_l = \mathbf{C}_l\boldsymbol{\Sigma}_l\mathbf{C}_l^{\mathrm{H}} + \sigma^2$. Now the integration in (A.1) is readily derived as

$$p(\bar{y}_l|\mathbf{b}_{n,1:k}, \mathbf{b}_{1:n-1}, \bar{y}_{1:l}) = \mathcal{N}(m_l, c_l). \tag{A.8}$$

ACKNOWLEDGMENTS

This work was supported by the National Science Foundation under Awards no. CCR-9903120 and no. CCR-0082607 and partially supported by the "Ministerio de Ciencia y Tecnología," Spain, under Project TIC 2001-2902.

REFERENCES

[1] S. Verdú, *Multiuser Detection*, Cambridge University Press, New York, NY, USA, 1998.

[2] A. Russ and M. K. Varanasi, "Noncoherent multiuser detection for nonlinear modulation over the Rayleigh-fading channel," *IEEE Trans. Inform. Theory*, vol. 47, no. 1, pp. 295–307, 2001.

[3] H.-Y. Wu and A. Duel-Hallen, "On the performance of coherent and noncoherent multiuser detectors for mobile radio CDMA channels," in *Proc. 5th IEEE International Conference on Universal Personal Communications*, vol. 1, pp. 76–80, Cambridge, Mass, USA, September–October 1996.

[4] M. Honig, U. Madhow, and S. Verdu, "Blind adaptive multiuser detection," *IEEE Trans. Inform. Theory*, vol. 41, no. 4, pp. 944–960, 1995.

[5] R. A. Iltis, "Exact and approximate maximum-likelihood parameter estimation for quasi-synchronous CDMA signals," *IEEE Trans. Commun.*, vol. 48, no. 7, pp. 1208–1216, 2000.

[6] X. Wang and H. V. Poor, "Blind multiuser detection: a subspace approach," *IEEE Trans. Inform. Theory*, vol. 44, no. 2, pp. 677–690, 1998.

[7] U. Fawer and B. Aazhang, "A multiuser receiver for code division multiple access communications over multipath channels," *IEEE Trans. Commun.*, vol. 43, no. 234, pp. 1556–1565, 1995.

[8] K. Yen and L. Hanzo, "Genetic algorithm assisted joint multiuser symbol detection and fading channel estimation for synchronous CDMA systems," *IEEE J. Select. Areas Commun.*, vol. 19, no. 6, pp. 985–998, 2001.

[9] T. J. Lim, L. K. Rasmussen, and H. Sugimoto, "An asynchronous multiuser CDMA detector based on the Kalman filter," *IEEE J. Select. Areas Commun.*, vol. 16, no. 9, pp. 1711–1722, 1998.

[10] T. J. Lim and Y. Ma, "The Kalman filter as the optimal linear minimum mean-squared error multiuser CDMA detector," *IEEE Trans. Inform. Theory*, vol. 46, no. 7, pp. 2561–2566, 2000.

[11] P. H.-Y. Wu and A. Duel-Hallen, "Multiuser detectors with disjoint Kalman channel estimators for synchronous CDMA mobile radio channels," *IEEE Trans. Commun.*, vol. 48, no. 5, pp. 752–756, 2000.

[12] S. Vasudevan and M. K. Varanasi, "Achieving near-optimum asymptotic efficiency and fading resistance over the time-varying Rayleigh-faded CDMA channel," *IEEE Trans. Commun.*, vol. 44, no. 9, pp. 1130–1143, 1996.

[13] K. J. Kim and R. A. Iltis, "Joint detection and channel estimation algorithms for QS-CDMA signals over time-varying channels," *IEEE Trans. Commun.*, vol. 50, no. 5, pp. 845–855, 2002.

[14] B. Flanagan, C. Suprin, S. Kumaresan, and J. Dunyak, "Performance of a joint Kalman demodulator for multiuser detection," in *Proc. 56th IEEE Vehicular Technology Conference (VTC '02)*, vol. 3, pp. 1525–1529, Vancouver, Canada, 2002.

[15] R. Chen, X. Wang, and J. S. Liu, "Adaptive joint detection and decoding in flat-fading channels via mixture Kalman filtering," *IEEE Trans. Inform. Theory*, vol. 46, no. 6, pp. 2079–2094, 2000.

[16] J. Zhang and P. M. Djurić, "Joint estimation and decoding of space-time trellis codes," *EURASIP J. Appl. Signal Process.*, vol. 2002, no. 3, pp. 305–315, 2002.

[17] E. Punskaya, C. Andrieu, A. Doucet, and W. J. Fitzgerald, "Particle filtering for multiuser detection in fading CDMA channels," in *Proc. 11th IEEE Signal Processing Workshop on Statistical Signal Processing*, pp. 38–41, Singapore, August 2001.

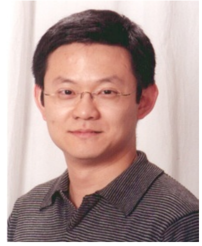
[18] Y. Huang and P. M. Djurić, "A blind particle filtering detector of signals transmitted over flat fading channels ," *IEEE Trans. Signal Processing*, vol. 52, no. 7, pp. 1891–1900, 2004.

[19] Y. Huang, J. Zhang, and P. M. Djurić, "Bayesian detection for BLAST," *IEEE Transactions on Signal Processing*, vol. 53, no. 3, pp. 1086–1096, March 2005.

[20] P. Fearnhead, *Sequential Monte Carlo method in filter theory*, Ph.D. thesis, Oxford University, Oxford, UK, 1998, available from http://www.maths.lancs.ac.uk/~fearnhea/.

[21] M. H. Hayes, *Statistical Digital Signal Processing and Modeling*, John Wiley & Sons, New York, 1996.

[22] S. Haykin, *Adaptive Filter Theory*, Prentice Hall, Upper Saddle River, NJ, USA, 4th edition, 2002.



Yufei Huang received the B.S. degree in applied electronics from Northwestern Polytechnical University, Xi'an, China, in 1995, and the M.S. and Ph.D. degrees in electrical engineering from the State University of New York at Stony Brook (SUNYSB), Stony Brook, NY, in 1997 and 2001, respectively. He is now an Assistant Professor in the Department of Electrical Engineering, The University of Texas at San Antonio. From 2001 to 2002, he worked as a Postdoctoral Researcher in the Department of Electrical and Computer Engineering, SUNYSB. His current research interests are in Bayesian inference, Monte Carlo methods, signal processing for communications, and bioinformatics.

Jianqiu (Michelle) Zhang received her B.S. and M.S. degrees in electrical engineering in 1992 and 1995, respectively, from Zhejiang University, Hangzhou University, and Zhongshan University, Guangzhou, China. From 1995 to 1997, she worked as a Software Engineer in R&D, Guangdong Nortel, China. She received her Ph.D. degree in electrical engineering from the State University of New York at Stony Brook in 2002. Currently, she is an Assistant Professor in the Department of Electrical and Computer Engineering at the University of New Hampshire. Her general research interest lies in the fields of wireless communications, information theory, and statistical methods in signal processing. She had been working on topics including CDMA multiuser detection, turbo-coding, particle filtering, space-time coding, and channel capacity.

Isabel Tienda Luna was born in Dona Mencía, Córdoba, Spain, on October 5th 1978. She received her B.S. and M.S. degrees from the University of Granada, Spain, in 1999 and 2001, respectively. Now she is a predoctoral researcher within the Systems, Signals and Waves Research Group, Department of Applied Physics, the University of Granada. Her predoctoral research is founded by the "Ministerio de Educaciṍn y Ciencia" of Spain. Her research interests are in the areas of digital communications and signal processing including particle filtering applications, channel estimation, and multiuser detection.

Petar M. Djurić received his B.S. and M.S. degrees in electrical engineering from the University of Belgrade, in 1981 and 1986, respectively, and his Ph.D. degree in electrical engineering from the University of Rhode Island in 1990. From 1981 to 1986 he was a Research Associate with the Institute of Nuclear Sciences, Vinca, Belgrade. Since 1990 he has been with Stony Brook University, where he is a Professor in the Department of Electrical and Computer Engineering. He works in the area of statistical signal processing, and his primary interests are in the theory of modeling, detection, estimation, and time series analysis and its application to a wide variety of disciplines including wireless communications and biomedicine. Professor Djurić has served on numerous technical committees for the IEEE and SPIE and has been invited to lecture at universities in the United States and overseas. He is the Area Editor of special issues of the Signal Processing Magazine and Associate Editor of the IEEE Transactions on Signal Processing. He is also Chair of the IEEE Signal Processing Society Committee on Signal Processing—Theory and Methods and is on the Editorial Board of Digital Signal Processing, the EURASIP Journal on Applied Signal Processing, and the EURASIP Journal on Wireless Communications and Networking. Professor Djurić is a Member of the American Statistical Association and the International Society for Bayesian Analysis.

Diego Pablo Ruiz Padillo was born in Cabra, Córdoba, Spain, on July 17th 1968. He received his B.S. and M.S. degrees in electronic physics from the University of Granada, in 1991 and 1993, respectively, and his Ph.D. degree with honors in 1995. He was a granted national researcher from the "Ministerio de Ciencia y Tecnología" of Spain, and an Assistant Professor in the Universities of Malaga and Granada from 1991 to 1998. Now Dr. Diego P. Ruiz is an Associate Professor within the Systems, Signals and Waves Research Group of the Department of Applied Physics in the University of Granada, Spain. His research interests include statistical signal processing and its applications to wireless communications, blind channel identification and equalization, adaptive algorithms, higher-order statistics, and radar signal processing applied to transient electromagnetic problems.