User Activity Tracking in DS-CDMA Systems

Manuel A. Vázquez and Joaquín Míguez, *Member, IEEE*

***Abstract*—In modern multiuser communication systems, users are allowed to enter or leave the system at any given time. Thus, the number of active users is an unknown and time-varying parameter, and the performance of the system depends on how accurately this parameter is estimated over time. The so-called problem of user identification, which consists of determining the number and identities of users transmitting in a communication system, is usually solved prior to, and hence independently of, that posed by the detection of the transmitted data. Since both problems are tightly connected, a joint solution is desirable. In this paper, we focus on direct-sequence (DS) code-division multiple-access (CDMA) systems and derive, within a Bayesian framework, different receivers that cope with an unknown and time-varying number of users while performing joint channel estimation and data detection. The main feature of these receivers, compared with other recently proposed schemes for user activity detection, is that they are natural extensions of existing maximum *a posteriori* (MAP) equalizers for multiple-input–multiple-output communication channels. We assess the validity of the proposed receivers, including their reliability in detecting the number and identities of active users, by way of computer simulations.**

***Index Terms*—Activity detection, code-division multiple access (CDMA), joint channel and data estimation, per-survivor processing (PSP), sequential Monte Carlo (SMC) methods.**

I. Introduction

A. Background

MOST receivers for code-division multiple-access (CDMA) communication systems [1], [2] operate under the hypothesis that the number of users is a known and fixed parameter (possibly estimated in a previous step). This is not the case in many real-world systems in which users can enter or leave the system (i.e., they start or stop transmitting) at any time instant. Not accounting for the inactivity of some of the users leads to a performance penalty [3], [4] that aggravates as the difference between the number of active users and that of those considered by the receiver increases.

Several recent papers addressing the problem of user identification in both single-user and multiuser systems can be found in the literature [3]–[7]. This problem involves the determination of users that are currently active and transmitting over the communication link, and it is usually associated with the estimation of the corresponding channel response and the detection of transmitted data. In the sequel, we focus on multiuser schemes, which are of interest in many modern communication systems. In [4], a two-stage multiuser detector that separates the identification of active users from the detection of symbols they transmit is proposed. The authors do not consider any model for the time variation of the channel or the users' activity pattern, and hence, the method is tailored to static environments. A different approach based on hypotheses testing is presented in [3]. At every data frame, the receiver starts by deciding which of the users that were active at the previous frame are still active. Their contribution is then subtracted from the observations, and the result is, in turn, used to decide whether a new user has entered the system or not. Users are only allowed to enter or leave the system at the beginning of a frame, and moreover, only one user is allowed to enter the system at a given frame (although several users can leave). The communication channel is assumed to be static.

Angelosante *et al.* [6] advocate the use of random-set theory (RST) to tackle the problem of joint user identification and channel estimation in direct-sequence (DS) CDMA systems, and they derive different algorithms within this framework. In particular, they discuss the implementation of exact Bayesian estimation recursions for this problem (although they appear intractable due to their computational burden) and suboptimal algorithms [6, Sec. III]. The effectiveness of the proposed algorithms is shown through computer simulations for a transmission system with short frames (only ten symbols) and a model for the variation of the channel response based on first-order Markov fading. Within the same RST framework, an alternative receiver based on the use of tree search techniques to alleviate the complexity of the user and data detection processes is proposed in [7]. It is shown through computer simulations that this receiver performs well in a transmission scheme with short frames (ten symbols) and the same first-order Markov channel model as in [6]. However, the practical implementation of the algorithm involves important approximations, including a decision feedback scheme and the quantization of channel coefficients. The feedback of decisions makes the detector subject to error propagation phenomena when the signal-to-noise ratio (SNR) is not sufficiently high, whereas channel quantization introduces estimation errors and can compromise the complexity of the receiver (see the Appendix for a formal description of this method and the involved approximations). Another Bayesian approach to the problem of user identification is presented in [8]. However, only an additive white Gaussian noise (AWGN) channel is assumed, and thus, the channel estimation problem that arises in most practical systems is overlooked.

Manuscript received March 30, 2012; revised September 28, 2012 and January 11, 2013; accepted February 21, 2013. Date of publication March 7, 2013; date of current version September 11, 2013. This work was supported in part by the Ministry of Science and Innovation of Spain through the Consolider-Ingenio 2010 COMONSENS program under Grant CSD-2008-00010 and the DEIPRO project under Grant TEC-2009-14504-C02-01 and in part by a joint program of the *Comunidad de Madrid* and *Universidad Carlos III de Madrid* under Grant CCG10-UC3M/TIC-5225 ETORS. The review of this paper was coordinated by Dr. Y. Qian.

The authors are with the Departamento de Teoría de la Señal y Comunicaciones, Universidad Carlos III de Madrid, 28911 Madrid, Spain (e-mail: mvazquez@tsc.uc3m.es; jmiguez@tsc.uc3m.es).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TVT.2013.2251024

In [9], user identification in a DS-CDMA system is achieved by combining the use of a special set of signature sequences with a maximum-likelihood (ML) detector. However, the resulting receiver is single-user, specifically tailored for binary phase-shift keying (BPSK) modulation, and assumes an AWGN channel.

For static systems, Angelosante *et al.* [10] propose a sparse linear regression-based method for estimating several parameters of a DS-CDMA system, including the number and identities of the active users. It is meant to operate only during a training phase in which known transmitted symbols allow the estimation of relevant (static) parameters that would then be used by a detector. Thus, no data detection is performed, and neither the channel nor the user activity is allowed to change within a given frame.

A multiuser detection receiver for DS-CDMA systems that performs user identification is presented in [11] for systems in which the *a priori* probability of a user being active is low. The proposed method stems from applying a sparsity-aware maximum *a posteriori* (MAP) criterion and is, therefore, Bayesian. However, in addition to the low-activity hypothesis, the channel is assumed to be static and known from a previous training phase.

Another work on activity detection is introduced in [12]. Buzzi *et al.* propose a fully blind method based on the application of the generalized likelihood ratio test aimed at detecting whether or not a specific user (whose signature sequence is known) enters the system during a given frame.

B. Contributions

In this paper, we propose a novel approach to the problem of joint user identification, data detection, and channel estimation in DS-CDMA systems. Therefore, the aim and scope of this work is similar to that in [6] and [7]. However, while the contributions in the latter references rely on the theory of random sets, the models and algorithms herein introduced are built upon conventional (and elementary) probability concepts, which should be familiar and more intuitive for engineers working in the field. Furthermore, this is attained without compromising the flexibility and the scope of the transmission system model, which can effectively represent, at least, all the scenarios considered in [6] and [7].

We investigate systems where users can enter or leave the system at any time during transmission of a certain frame. For that purpose, a statistical dynamical model governing the users' activity is introduced, and the different algorithms that can be applied at the receiver exploit this dynamics to more accurately detect whether a user is transmitting or not.

Four different algorithms for user identification and data detection are derived. The first algorithm is the optimal MAP sequence detector, which is implemented by means of the Viterbi algorithm (VA) [13], assuming the channel is perfectly known. For those cases in which the latter hypothesis is not fulfilled, an analogous per-survivor processing (PSP) [14] method is suggested. The latter approach is appealing because it allows for a flexible tradeoff between performance and complexity. Additionally, two more algorithms for joint user identification, channel estimation, and data detection are proposed. They both build upon the sequential Monte Carlo (SMC) methodology, which is also known as particle filtering [15]. In particular, we derive the optimal SMC equalizer and, subsequently, a lower-complexity SMC receiver whose computational burden grows only linearly with the number of users. For the tracking of the time-varying channel, the last three algorithms rely on Kalman filtering (KF) [16], [17], and the channel coefficients need not be quantized.

Computer simulations considering a standard channel model and long data frames (that allow for significant variations in the channel coefficients) serve to assess the performance of the proposed algorithms in a realistic environment.

Although, in this paper, we are implicitly focusing on cellular networks, it should be remarked that DS-CDMA is also widely used in the context of wireless sensor networks (see, e.g., [18]–[21]). The communication nodes deployed in this kind of system can often switch on and off asynchronously during operation, and a fusion center acting as the receiver needs to be aware, at every time, of which nodes are transmitting. These nodes can be seen as the users of the system, and hence, the problem to solve is tantamount to the problem tackled in this paper.

C. Organization of the Paper

The remainder of this paper is organized as follows: In Section II, the discrete-time signal model of a DS-CDMA transmission system with a flat-fading and time-selective channel is described, and the proposed model describing the dynamics of the user activity is introduced. The optimal receiver for the case of a perfectly known channel is introduced in Section III, whereas an extension for the unknown channel scenario, which is based on the PSP approach, is obtained in Section IV. In Section V, a brief introduction to SMC methods is given, and two novel SMC algorithms that tackle the problem at hand are derived. Computer simulation results are shown in Section VI, and finally, Section VII is devoted to a brief summary and final remarks.

II. Signal Model

Notation: Throughout this paper, we use notation $p(\cdot)$ for probability density functions (pdfs) and probability mass functions (pmfs). The specific form of the pdf/pmf should always be clear from its argument. For example, $p(\mathbf{H}_{-1})$ is the *a priori* pdf for the channel matrix, whereas $p(u_t(i))$ is the prior pmf for the activity of the ith user at time t. The conditional pdf/pmf of random vector $\mathbf{a}$ given the observation of related vector $\mathbf{b}$ is written as $p(\mathbf{a}|\mathbf{b})$. Occasionally, we use a specific notation for Gaussian pdfs, by letting $\mathcal{N}(\mathbf{z}|\overline{\mathbf{z}}, \mathbf{K})$ be the Gaussian pdf of random vector $\mathbf{z}$ with mean $\overline{\mathbf{z}}$ and covariance matrix $\mathbf{K}$. Similarly, $\mathcal{U}(\mathcal{A})$ denotes the uniform probability distribution over set $\mathcal{A}$, whose cardinality is $|\mathcal{A}|$.

Additionally, for a time-dependent discrete signal, e.g., $\mathbf{a}$, we write $\mathbf{a}_{i:j}$ to denote the sequence of values of the signal between time instants i and j, i.e., $\mathbf{a}_{i:j} = \{\mathbf{a}_i, \mathbf{a}_{i+1}, \ldots, \mathbf{a}_j\}$.

A. Signal Model

We consider a synchronous DS-CDMA system with up to N users transmitting through a time-selective flat-fading channel. Each user is assigned a signature sequence of length L. The discrete-time baseband-equivalent model is [2]

$$\mathbf{y}_t = \mathbf{C}\mathbf{H}_t\mathbf{s}_t + \mathbf{z}_t \tag{1}$$

where $\mathbf{y}_t$ is the $L \times 1$ vector collecting the observations at discrete-time t, $\mathbf{C}$ is the $L \times N$ matrix of spreading codes whose columns are the signature sequences of all potential users of the system, $\mathbf{H}_t = \mathrm{diag}\{h_t(1), h_t(2), \ldots, h_t(N)\}$ is the time-varying $N \times N$ diagonal channel matrix, $\mathbf{s}_t = [s_t(1), \ldots, s_t(N)]$ is an $N \times 1$ vector with the symbols transmitted by all active users, and $\mathbf{z}_t = [z_t(1), \ldots, z_t(N)]$ is the AWGN with zero mean and covariance $\sigma_z^2\mathbf{I}_L$, with $\mathbf{I}_L$ being the identity matrix of order L. When user i is active at time t, symbol $s_t(i)$ is modeled as a discrete uniform random variable (r.v.) on the alphabet $\mathcal{A}$ (with cardinality $|\mathcal{A}|$), and hence, $s_t(i) \sim \mathcal{U}(\mathcal{A})$.

Although the number of active users in the system is time varying, a fixed value N (the maximum number of users that the system can hold) is assumed in every vector and matrix whose dimensions depend on it.[1] When a user is not active, we assume that its *transmitted* symbol is equal to 0 (no energy), which is intuitively true, on one hand, and greatly simplifies the model, on the other hand. Moreover, we are going to define the r.v., i.e.,

$$u_t(i) = \begin{cases} 1, & \text{if user } i \text{ is active at time } t \\ 0, & \text{if user } i \text{ is NOT active at time } t. \end{cases} \tag{2}$$

This indicator variable is related to the corresponding symbol transmitted by the ith user at time t according to

$$u_t(i) = 1 - \delta\left(s_t(i)\right) = \begin{cases} 1, & \text{if } s_t(i) \neq 0 \\ 0, & \text{if } s_t(i) = 0 \end{cases} \tag{3}$$

where $\delta(\cdot)$ denotes the Kronecker delta function. Notice that, from the definition given in (3), $s_t(i)$ completely determines $u_t(i)$. Specifically

$$p\left(u_t(i) = 0 \,|\, s_t(i)\right) = \delta\left(s_t(i)\right)$$
$$p\left(u_t(i) = 1 \,|\, s_t(i)\right) = 1 - \delta\left(s_t(i)\right). \tag{4}$$

Based on the r.v. $\{u_t(i), i = 1, \ldots, N, t \geq 0\}$, a dynamic model for the active users in the system at a given time instant can be easily defined through the following probabilities:

$$p\left(u_0(i) = 1\right) = \phi \tag{5}$$

$$p\left(u_t(i) = 1 \,|\, u_{t-1}(i) = 1\right) = \mu \tag{6}$$

$$p\left(u_t(i) = 1 \,|\, u_{t-1}(i) = 0\right) = \eta \tag{7}$$

[1]In practice, this value depends on the length of the signature sequences employed, i.e., L: a system with longer signature sequences can accommodate a larger number of users without the performance severely degrading due to multiple-access interference.

where ϕ is the *a priori* probability of a user being active at the beginning of the transmission, μ is the probability of a user being active at time t, conditional on the fact that it was already active at time $t-1$, and η is the probability of a user being active at time t given that it was not active at time $t-1$. From (5)–(7), we have, respectively

$$p\left(u_0(i) = 0\right) = 1 - \phi \tag{8}$$

$$p\left(u_t(i) = 0 \,|\, u_{t-1}(i) = 1\right) = 1 - \mu \tag{9}$$

$$p\left(u_t(i) = 0 \,|\, u_{t-1}(i) = 0\right) = 1 - \eta \tag{10}$$

and these six equations, i.e., (5)–(10), completely determine the dynamic model for the user activity.

The model described by (5)–(10) implies that sequence $u_t(i)$, for a fixed i, is a Markov chain. As for the symbols transmitted by the ith user, they are assumed to be independent and identically distributed (i.i.d., with uniform distribution $\mathcal{U}(\mathcal{A})$) conditional on the user being active, i.e., $u_t(i) = 1$. However, due to the sequence of activity indicators $u_t(i)$ being modeled as a Markov chain, the sequence of symbols $s_t(i)$ can be also shown to be a Markov chain itself. In particular

$$p\left(s_t(i) | s_{0:t-1}(i)\right) = p\left(s_t(i) | s_{0:t-1}(i), u_{t-1}(i)\right) \tag{11}$$

$$= p\left(s_t(i) | s_{t-1}(i), u_{t-1}(i)\right) \tag{12}$$

$$= p\left(s_t(i) | s_{t-1}(i)\right) \tag{13}$$

where it has been used, in (11) and (13), that $s_{t-1}(i)$ uniquely determines $u_{t-1}(i)$ [see (3)] and, in (12), that the symbol transmitted at time t by the ith user, i.e., $s_t(i)$, is independent of all the past transmitted symbols conditional on the activity of the user at the previous time instant, i.e., $u_{t-1}(i)$, being known. Furthermore, the different users are assumed to be independent, and hence, $p(\mathbf{s}_t | \mathbf{s}_{0:t-1}) = p(\mathbf{s}_t | \mathbf{s}_{t-1})$.

For the purpose of algorithm design, it is common to model the channel variation by means of an autoregressive (AR) process driven by white Gaussian noise [22]. We consider a second-order AR model, i.e.,

$$\mathbf{H}_t = \alpha_1\mathbf{H}_{t-1} + \alpha_2\mathbf{H}_{t-2} + \mathbf{V}_t \tag{14}$$

where α_1 and α_2 are the coefficients of the process, and $\mathbf{V}_t$ is an $N \times N$ diagonal matrix whose nonzero elements are i.i.d. Gaussian r.v. with zero mean and variance σ_v^2.

From (1) and (14) and assuming that the pdf of the channel at times $t = -1$ and $t = -2$ is Gaussian and known, it is apparent that the channel coefficients can be optimally estimated by KF [16] conditional on the transmitted symbols. For a convenient description of the corresponding algorithm, we rewrite (1) as

$$\mathbf{y}_t = \mathbf{B}_t\mathbf{h}_t + \mathbf{z}_t \tag{15}$$

where $\mathbf{h}_t = [h_t(1), h_t(2), \ldots, h_t(N)]^\top$ is a vector whose coefficients are the elements in the diagonal of $\mathbf{H}_t$, and

$$\mathbf{B}_t = \mathbf{C}\mathbf{S}_t \tag{16}$$

with $\mathbf{S}_t$ being a diagonal matrix whose elements are the symbols transmitted at time t, i.e.,

$$\mathbf{S}_t = \begin{pmatrix} s_t(1) & 0 & \cdots & 0 \\ 0 & s_t(2) & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & s_t(N) \end{pmatrix}. \quad (17)$$

III. Optimal Detection With Known Channel

We are interested in detecting, at every time instant t, which users are active along with their transmitted data. According to the signal model in Section II-B, all this information is contained in symbol vector $\mathbf{s}_t$. In particular, note that, for $i = 1, \ldots, N$, $s_t(i) = 0$ if and only if $u_t(i) = 0$ and $s_t(i) \in \mathcal{A}$ if and only if $u_t(i) = 1$. Thus, our ultimate goal is to find the most probable sequence of symbol vectors, i.e., $\mathbf{s}_{0:t}$, where $\mathbf{s}_t \in \{\mathcal{A} \cup \{0\}\}^N$ for every t, given the sequence of observations, i.e., $\mathbf{y}_{0:t}$. This can be seen as a MAP detection problem for the sequence of symbol vectors, i.e., we aim to compute

$$\mathbf{s}_{0:t}^{\text{MAP}} = \arg\max_{\mathbf{s}_{0:t}} p(\mathbf{s}_{0:t}|\mathbf{y}_{0:t}, \mathbf{H}_{0:t}). \quad (18)$$

For the sake of clarity, this section is divided into two different but closely related parts. In Section III-A, we derive an analytical expression for the posterior probability in (18) and obtain an explicit formula for its computation. Ultimately, this results in an equivalent optimization problem that is tackled in Section III-B by means of the VA.

A. Computation of Posterior Probabilities

The *a posteriori* probability of the sequence of symbol vectors can be decomposed by means of Bayes' theorem as

$$\begin{aligned} p(\mathbf{s}_{0:t}|\mathbf{y}_{0:t}, \mathbf{H}_{0:t}) &= \frac{p(\mathbf{y}_{0:t}|\mathbf{H}_t, \mathbf{s}_{0:t}) p(\mathbf{s}_{0:t}|\mathbf{H}_{0:t})}{p(\mathbf{y}_{0:t})} \\ &\propto p(\mathbf{y}_{0:t}|\mathbf{H}_{0:t}, \mathbf{s}_{0:t}) p(\mathbf{s}_{0:t}) \end{aligned} \quad (19)$$

where, on one hand, factor $p(\mathbf{y}_{0:t})$ does not depend on $\mathbf{s}_{0:t}$, and hence, it is just a normalization constant, and on the other hand, the channel and the symbols are *a priori* independent, which results in $p(\mathbf{s}_{0:t}|\mathbf{H}_{0:t}) = p(\mathbf{s}_{0:t})$.

Let us assume that the code matrix, i.e., $\mathbf{C}$, is known. Then, using the definition of conditional probability, the likelihood on the right-hand side of (19) can be decomposed as

$$\begin{aligned} &p(\mathbf{y}_{0:t}|\mathbf{H}_{0:t}, \mathbf{s}_{0:t}) \\ &\quad = p(\mathbf{y}_t|\mathbf{H}_{0:t}, \mathbf{s}_{0:t}, \mathbf{y}_{0:t-1}) p(\mathbf{y}_{0:t-1}|\mathbf{H}_{0:t}, \mathbf{s}_{0:t}) \\ &\quad = p(\mathbf{y}_t|\mathbf{H}_t, \mathbf{s}_t) p(\mathbf{y}_{0:t-1}|\mathbf{H}_{0:t-1}, \mathbf{s}_{0:t-1}) \\ &\quad = \prod_{k=0}^{t} p(\mathbf{y}_k|\mathbf{H}_k, \mathbf{s}_k) \\ &\quad = \left(2\pi\sigma_z^2\right)^{\frac{-L(t+1)}{2}} e^{-\sum_{k=0}^{t} \frac{|\mathbf{y}_k - \mathbf{C}\mathbf{H}_k \mathbf{s}_k|^2}{2\sigma_z^2}} \end{aligned} \quad (20)$$

where we have used the conditional independence of the observations $\mathbf{y}_t$ given channel $\mathbf{H}_t$ and symbols $\mathbf{s}_t$, i.e.,

$$\begin{aligned} p(\mathbf{y}_k|\mathbf{H}_{0:k}, \mathbf{s}_{0:k}, \mathbf{y}_{0:k-1}) &= p(\mathbf{y}_k|\mathbf{H}_k, \mathbf{s}_k) \\ &= \left(2\pi\sigma_z^2\right)^{-L/2} e^{-\frac{|\mathbf{y}_k - \mathbf{C}\mathbf{H}_k \mathbf{s}_k|^2}{2\sigma_z^2}}. \end{aligned} \quad (21)$$

As for the *a priori* probability of the symbols, i.e., $p(\mathbf{s}_{0:t})$ in (19), we note that this is not uniform. Recall from Section II [see (13)] that sequence $\mathbf{s}_0, \mathbf{s}_1, \ldots, \mathbf{s}_t$ forms a Markov chain because each vector $\mathbf{s}_t$ contains information not only about the data transmitted at time t but also about the activity of the users at that time instant.

By simply using the definition of conditional probability together with (13), the *a priori* probability of the sequence of symbol vectors can be written as

$$p(\mathbf{s}_{0:t}) = p(\mathbf{s}_0) \prod_{j=1}^{t} p(\mathbf{s}_j|\mathbf{s}_{j-1}). \quad (22)$$

Substituting (20) and (22) in (19) yields

$$p(\mathbf{s}_{0:t}|\mathbf{y}_{0:t}, \mathbf{H}_{0:t}) \propto p(\mathbf{s}_0) \prod_{j=1}^{t} p(\mathbf{s}_j|\mathbf{s}_{j-1}) e^{-\sum_{k=0}^{t} \frac{|\mathbf{y}_k - \mathbf{C}\mathbf{H}_k \mathbf{s}_k|^2}{2\sigma_z^2}} \quad (23)$$

which is the objective function that we have to maximize to solve the problem posed by (18).

Since maximizing a function is tantamount to minimizing the opposite of its logarithm, it follows that

$$\mathbf{s}_{0:t}^{\text{MAP}} = \arg\min_{\mathbf{s}_{0:t}} \{-\log p(\mathbf{s}_{0:t}|\mathbf{y}_{0:t}, \mathbf{H}_{0:t})\} \quad (24)$$

and substituting (23) into (24), we obtain

$$\begin{aligned} \mathbf{s}_{0:t}^{\text{MAP}} = \arg\min_{\mathbf{s}_{0:t}} \Bigg\{ &-\log p(\mathbf{s}_0) - \sum_{j=1}^{t} \log p(\mathbf{s}_j|\mathbf{s}_{j-1}) \\ &+ \sum_{k=0}^{t} \frac{|\mathbf{y}_k - \mathbf{C}\mathbf{H}_k \mathbf{s}_k|^2}{2\sigma_z^2} \Bigg\}. \end{aligned} \quad (25)$$

We can now arrange the terms in (25) in a more convenient way to arrive at

$$\begin{aligned} \mathbf{s}_{0:t}^{\text{MAP}} = \arg\min_{\mathbf{s}_{0:t}} \Bigg\{ &\frac{|\mathbf{y}_0 - \mathbf{C}\mathbf{H}_0 \mathbf{s}_0|^2}{2\sigma_z^2} - \log p(\mathbf{s}_0) \\ &+ \sum_{k=1}^{t} \left(\frac{|\mathbf{y}_k - \mathbf{C}\mathbf{H}_k \mathbf{s}_k|^2}{2\sigma_z^2} - \log p(\mathbf{s}_k|\mathbf{s}_{k-1}) \right) \Bigg\}. \end{aligned} \quad (26)$$

By taking into account that both the activity and the data transmitted by the different users are independent, the *a priori* probability of $\mathbf{s}_t$ is decomposed as

$$p(\mathbf{s}_t) = \prod_{i=1}^{N} p\left(s_t(i)\right) \quad (27)$$

where the *a priori* probability of the ith symbol in the vector can be computed as

$$\begin{aligned} p(s_t(i)) &= \sum_{u_t(i)\in\{0,1\}} p(s_t(i), u_t(i)) \\ &= \sum_{u_t(i)\in\{0,1\}} p(s_t(i)|u_t(i))\, p(u_t(i)) \\ &= p(s_t(i)|u_t(i)=0)\, p(u_t(i)=0) + \\ &= + p(s_t(i)|u_t(i)=1)\, p(u_t(i)=1). \end{aligned} \tag{28}$$

Since $s_t(i)=0$ whenever $u_t(i)=0$ (the user is not active), it follows that

$$p(s_t(i)|u_t(i)=0) = \delta(s_t(i)). \tag{29}$$

On the other hand, when a user is active ($u_t(i)=1$), we have assumed that all symbols from alphabet $\mathcal{A}$ are equally likely, and hence

$$p(s_t(i)|u_t(i)=1) = (1-\delta(s_t(i)))\frac{1}{|\mathcal{A}|}. \tag{30}$$

Using (5), (8), (29), and (30), (28) can be written as

$$p(s_t(i)) = \begin{cases} 1-\phi, & \text{if } s_t(i)=0 \\ \frac{1}{|\mathcal{A}|}\phi, & \text{if } s_t(i)\in\mathcal{A}. \end{cases} \tag{31}$$

The conditional terms $p(\mathbf{s}_j|\mathbf{s}_{j-1})$ in (25) can be analyzed using a similar argument. We exploit the independence of the users to write

$$p(\mathbf{s}_t|\mathbf{s}_{t-1}) = \prod_{i=1}^{N} p(s_t(i)|s_{t-1}(i)) \tag{32}$$

where

$$\begin{aligned} &p(s_t(i)|s_{t-1}(i)) \\ &\quad= \sum_{u_t(i)\in\{0,1\}} p(s_t(i)|u_t(i), s_{t-1}(i))\, p(u_t(i)|s_{t-1}(i)) \\ &\quad= p(s_t(i)|u_t(i)=0, s_{t-1}(i))\, p(u_t(i)=0|s_{t-1}(i)) \\ &\qquad+ p(s_t(i)|u_t(i)=1, s_{t-1}(i))\, p(u_t(i)=1|s_{t-1}(i)). \end{aligned} \tag{33}$$

Moreover, symbol $s_t(i)$ is conditionally independent of $s_{t-1}(i)$ given activity indicator $u_t(i)$ so that

$$p(s_t(i)|u_t(i)=0, s_{t-1}(i)) = p(s_t(i)|u_t(i)=0) \tag{34}$$

$$p(s_t(i)|u_t(i)=1, s_{t-1}(i)) = p(s_t(i)|u_t(i)=1) \tag{35}$$

and substituting (34) and (35) into (33) yields

$$\begin{aligned} p(s_t(i)|s_{t-1}(i)) &= p(s_t(i)|u_t(i)=0)\, p(u_t(i)=0|s_{t-1}(i)) \\ &\quad+ p(s_t(i)|u_t(i)=1)\, p(u_t(i)=1|s_{t-1}(i)). \end{aligned} \tag{36}$$

Using (29) and (30) again, (36) can be rewritten as

$$p(s_t(i)|s_{t-1}(i)) = \begin{cases} p(u_t(i)=0|s_{t-1}(i)), & \text{if } s_t(i)=0 \\ p(u_t(i)=1|s_{t-1}(i))\,\frac{1}{|\mathcal{A}|}, & \text{if } s_t(i)\in\mathcal{A} \end{cases} \tag{37}$$

where the two probabilities on the right-hand side of (37) can be easily derived from the model in Section II. In particular, if $s_{t-1}(i)=0$, then $u_{t-1}(i)=0$, whereas if $s_{t-1}(i)\in\mathcal{A}$, then $u_{t-1}(i)=1$; hence

$$p(u_t(i)|s_{t-1}(i)) = \begin{cases} p(u_t(i)|u_{t-1}(i)=0), & \text{if } s_{t-1}(i)=0 \\ p(u_t(i)|u_{t-1}(i)=1), & \text{if } s_{t-1}(i)\in\mathcal{A} \end{cases} \tag{38}$$

and substituting (38) into (37), together with a straightforward application of (6), (7), (9), and (10), yields

$$p(s_t(i)|s_{t-1}(i)) = \begin{cases} 1-\eta, & \text{if } s_t(i)=s_{t-1}(i)=0 \\ \eta\frac{1}{|\mathcal{A}|}, & \text{if } s_t(i)\in\mathcal{A}, s_{t-1}(i)=0 \\ 1-\mu, & \text{if } s_t(i)=0, s_{t-1}(i)\in\mathcal{A} \\ \mu\frac{1}{|\mathcal{A}|}, & \text{if } s_t(i), s_{t-1}(i)\in\mathcal{A}. \end{cases} \tag{39}$$

B. MAP Sequence Detection

The optimization problem given by (26) can be solved using the VA. Indeed, if we let

$$C_0^{(\mathbf{s}_0)} = \frac{|\mathbf{y}_0 - \mathbf{C}\mathbf{H}_0\mathbf{s}_0|^2}{2\sigma_z^2} - \log p(\mathbf{s}_0) \tag{40}$$

denote the metric associated with the first symbol vector, i.e., $\mathbf{s}_0$, and

$$C_t^{(\mathbf{s}_{k-1}\to\mathbf{s}_k)} = \frac{|\mathbf{y}_k - \mathbf{C}\mathbf{H}_k\mathbf{s}_k|^2}{2\sigma_z^2} - \log p(\mathbf{s}_k|\mathbf{s}_{k-1}), \quad k\geq 1 \tag{41}$$

denote the metric associated with the transition from $\mathbf{s}_{k-1}$ to $\mathbf{s}_k$, then (26) can be rewritten as

$$\mathbf{s}_{0:t}^{\mathrm{MAP}} = \arg\min_{\mathbf{s}_{0:t}} \left\{ C_0^{(\mathbf{s}_0)} + \sum_{k=1}^{t} C_k^{(\mathbf{s}_{k-1}\to\mathbf{s}_k)} \right\} \tag{42}$$

where both $C_0^{(\mathbf{s}_0)} \geq 0$ and $C_k^{(\mathbf{s}_{k-1}\to\mathbf{s}_k)} \geq 0$ for every k. Problem (42) can be interpreted as the sequential estimation of the state in a space-state system. To be precise, the state of the system at time t is determined by symbol vector $\mathbf{s}_t$, and the dynamics is governed by conditional distribution $p(\mathbf{s}_t|\mathbf{s}_{t-1})$. Since the state space is finite and the system is Markov, we can use the VA to solve (42) and compute $\mathbf{s}_{0:t}^{\mathrm{MAP}}$ exactly. A peculiarity of the algorithm in this particular case is that the initial state, i.e., $\mathbf{s}_0$, is unknown, and hence, the first step of the algorithm consists in computing the metric of every possible initial vector $\mathbf{s}_0$. Pseudocode 1 gives an overall view of the algorithm.

Pseudocode 1 Optimal receiver with known channel

1: **for** every possible initial state, $\mathbf{s}_0 \in (\mathcal{A} \cup \{0\})^N$ **do**
2: initialize its associated path as $\mathrm{p}_0^{(\mathbf{s}_0)} \leftarrow \mathbf{s}_0$
3: using (27) and (31), compute its initial cost as

$$C_0^{(\mathbf{s}_0)} = \frac{|\mathbf{y}_0 - \mathbf{C}\mathbf{H}_0\mathbf{s}_0|^2}{2\sigma_z^2} - \log p\left(\mathbf{s}_0\right)$$

4: **for** $t > 1$ **do**
5: **for** every state at time $t-1$, $\mathbf{s}_{t-1}$, **do**
6: **for** every possible symbol vector, $\mathbf{s}_t \in (\mathcal{A} \cup \{0\})^N$, **do**
7: using (32) and (39), compute the cost of transition $\mathbf{s}_{t-1} \rightarrow \mathbf{s}_t$ at time t as

$$C_t^{(\mathbf{s}_{t-1} \rightarrow \mathbf{s}_t)} = \frac{|\mathbf{y}_t - \mathbf{C}\mathbf{H}_t\mathbf{s}_t|^2}{2\sigma_z^2} - \log p\left(\mathbf{s}_t | \mathbf{s}_{t-1}\right)$$

8: **if** no path arrived yet at state $\mathbf{s}_t$ **then**
9: $C_t^{(\mathbf{s}_t)} \leftarrow C_{t-1}^{(\mathbf{s}_{t-1})} + C_t^{(\mathbf{s}_{t-1} \rightarrow \mathbf{s}_t)}$
10: set the path of state $\mathbf{s}_t$ at time t as $\mathrm{p}_t^{(\mathbf{s}_t)} \leftarrow \{\mathrm{p}_{t-1}^{(\mathbf{s}_{t-1})}, \mathbf{s}_t\}$
11: **else if** $C_t^{(\mathbf{s}_t)} > C_{t-1}^{(\mathbf{s}_{t-1})} + C_t^{(\mathbf{s}_{t-1} \rightarrow \mathbf{s}_t)}$ **then**
12: update the path and cost of state $\mathbf{s}_t$ using expressions in lines 9 and 10
13: choose the state with the lowest cost
14: **return** its associated path

Notice that, despite the channel being flat, the VA is needed to obtain the optimal solution to the posed detection problem due to the user activity model turning the sequence of transmitted symbol vectors, i.e., $\mathbf{s}_t$, into a Markovian process.

IV. Per-Survivor Processing: Sequence Detection With Unknown Channel

When the channel is unknown, we can extend the algorithm in Section III in the same way the classical PSP algorithm [14] extends the conventional VA. The key idea is to maintain a set of sequences of symbol vectors (termed *survivors*) for every possible state and every time instant t (instead of a single sequence as in the original VA), to cope with the uncertainty associated with the channel response that needs to be estimated.

As before, we aim at finding the sequence of symbol vectors that maximizes the *a posteriori* probability, i.e., the probability of the sequence of symbols given the sequence of observations. However, in this case, the channel is unknown, and we seek an approximate solution for the optimization problem, i.e.,

$$\mathbf{s}_{0:t}^{\mathrm{MAP}} = \arg\max_{\mathbf{s}_{0:t}} p(\mathbf{s}_{0:t} | \mathbf{y}_{0:t}). \tag{43}$$

To deal with the unknown channel, we replace the true (unknown) channel matrix at time k, i.e., $\mathbf{H}_k$, with an estimate $\hat{\mathbf{H}}_k$ so that (26) now becomes

$$\hat{\mathbf{s}}_{0:t}^{\mathrm{MAP}} = \arg\min_{\mathbf{s}_{0:t}} \left\{ \frac{\left|\mathbf{y}_0 - \mathbf{C}\hat{\mathbf{H}}_0, \mathbf{s}_0\right|^2}{2\sigma_z^2} - \log p(\mathbf{s}_0) + \sum_{k=1}^{t} \left(\frac{\left|\mathbf{y}_k - \mathbf{C}\hat{\mathbf{H}}_k\mathbf{s}_k\right|^2}{2\sigma_z^2} - \log p(\mathbf{s}_k | \mathbf{s}_{k-1}) \right) \right\}. \tag{44}$$

Similar to the classical PSP algorithm, we associate a channel estimator with every survivor to obtain the needed channel matrix estimates. Let P be the overall number of survivors and $\mathbf{s}_{0:t}^{(i)}$, $i \in \{1, \ldots, P\}$, the *path* (or symbol sequence) of the ith survivor up to time t. Using the sequence of symbol vectors $\mathbf{s}_{0:t}^{(i)}$ and observations $\mathbf{y}_{0:t}$, a conditional estimate of the channel matrix at time t, i.e., $\hat{\mathbf{H}}_t^{(i)}$, can be obtained (using the KF, for example). Since this estimate depends on the specific path of the survivor, i.e., $\mathbf{s}_{0:t}^{(i)}$, different survivors (even within the same state) can have different channel estimates, and hence, the superscript is needed.

From an estimate of the channel at time $t-1$, i.e., $\hat{\mathbf{H}}_{t-1}^{(i)}$, a prediction thereof at time t can be easily obtained by simply using the AR model of the channel given by (14). Once an estimate, i.e., $\hat{\mathbf{H}}_t^{(i)}$, of $\mathbf{H}_t$ for the ith survivor at time t is available, the cost of the transition from that survivor when $\mathbf{s}_t$ is transmitted is computed as

$$C_0^{(\mathbf{s}_{t-1}^{(i)} \rightarrow \mathbf{s}_t)} = \frac{\left|\mathbf{y}_t - \mathbf{C}\hat{\mathbf{H}}_t^{(i)}\mathbf{s}_t\right|^2}{2\sigma_z^2} - \log p\left(\mathbf{s}_t | \mathbf{s}_{t-1}^{(i)}\right). \tag{45}$$

The notation employed in (45) emphasizes that now transitions occur between a survivor, i.e., $\mathbf{s}_{t-1}^{(i)}$, at time $t-1$ and a state, i.e., $\mathbf{s}_t$, at time t (rather than between states). Similar to the previous algorithm, there is no initial state, and hence, in the first step of the algorithm, a metric must be computed for every survivor i, i.e.,

$$C_0^{\mathbf{s}_0^{(i)}} = \frac{\left|\mathbf{y}_0 - \mathbf{C}\hat{\mathbf{H}}_0^{(i)}\mathbf{s}_0^{(i)}\right|^2}{2\sigma_z^2} - \log p\left(\mathbf{s}_0^{(i)}\right), \quad i = 1, \ldots, P \tag{46}$$

using an initial estimate of the channel, i.e., $\hat{\mathbf{H}}_0^{(i)}$.

The PSP-based receiver is similar to the optimal receiver described in Section III. Pseudocodes 2 and 3 show, respectively, the initialization and the sequential step of the proposed receiver in more detail.

Pseudocodes 2 PSP `Initialization`

---

let $p \leftarrow P/|\mathcal{A} \cup \{0\}|^N$ be the number of survivors per state
initialize the overall survivor index $i \leftarrow 0$
for every possible state $\mathbf{s}_0 \in (\mathcal{A} \cup \{0\})^N$ **do**
for every survivor per state $j = 1, \ldots, p$ **do**

5: increment the overall survivor index $i \leftarrow i + 1$
6: initialize
- the path $\mathrm{p}_0^{(i)} \leftarrow \mathbf{s}_0$
- the channel matrix estimate $\hat{\mathbf{H}}_0^{(i)} = \mathbf{0}$
- the cost using (27) and (31) to compute

$$C_0^{(\mathbf{s}_0)^{(i)}} = \frac{\left|\mathbf{y}_0 - \mathbf{C}\hat{\mathbf{H}}_0^{(i)}\mathbf{s}_0^{(i)}\right|^2}{2\sigma_z^2} - \log p(\mathbf{s}_0^{(i)})$$

of the ith overall survivor (jth survivor within state $\mathbf{s}_0$)

7: update $\hat{\mathbf{H}}_0^{(i)}$ using $\mathbf{y}_0$ and $\mathbf{s}_0$

Pseudocodes 3 PSP `Sequential step`

1: let $p \leftarrow P/|\mathcal{A} \cup \{0\}|^N$ be the number of survivors per state
2: **for** $t > 1$ **do**
3: **for** every survivor, $i = 1, \ldots, P$ at time $t - 1$ **do**
4: compute $\hat{\mathbf{H}}_t^{(i)}$ from $\hat{\mathbf{H}}_{t-1}^{(i)}$ using (14)
5: let $\mathbf{s}_{t-1}$ be the state associated with the ith survivor at time $t - 1$
6: **for** every possible symbol vector at time t, $\mathbf{s}_t \in (\mathcal{A} \cup \{0\})^N$, **do**
7: using (32) and (39), compute the cost of the transition from the ith survivor at time $t - 1$ when $\mathbf{s}_t$ is transmitted

$$C_t^{(\mathbf{s}_{t-1}^{(i)} \to \mathbf{s}_t)} = \frac{\left|\mathbf{y}_t - \mathbf{C}\hat{\mathbf{H}}_t^{(i)}\mathbf{s}_t\right|^2}{2\sigma_z^2} - \log p\left(\mathbf{s}_t | \mathbf{s}_{t-1}\right)$$

8: **if** (number of survivors within state $\mathbf{s}_t$) $< p$ **then**
9: add a new survivor, identified as the kth overall survivor at time t, such that

$$\mathrm{p}_t^{(k)} \leftarrow \left\{\mathrm{p}_{t-1}^{(i)}, \mathbf{s}_t\right\}$$

$$C_t^{\mathbf{s}_0^{(k)}} \leftarrow C_{t-1}^{\mathbf{s}_0^{(i)}} + C_t^{(\mathbf{s}_{t-1}^{(i)} \to \mathbf{s}_t)}$$

10: update $\hat{\mathbf{H}}_t^{(k)}$ from $\hat{\mathbf{H}}_{t-1}^{(i)}$ using $\mathbf{y}_t$ and $\mathbf{s}_t$
11: **else**
12: find the survivor with the higher cost within state $\mathbf{s}_t$, identified as the kth overall survivor
13: **if** $C_t^{\mathbf{s}_0^{(k)}} > C_{t-1}^{\mathbf{s}_0^{(i)}} + C_t^{(\mathbf{s}_{t-1}^{(i)} \to \mathbf{s}_t)}$ **then**
14: update the kth survivor as in line 9
15: update $\hat{\mathbf{H}}_t^{(k)}$ from $\hat{\mathbf{H}}_{t-1}^{(i)}$ using $\mathbf{y}_t$ and $\mathbf{s}_t$
16: choose the overall survivor with the lowest cost
17: **return** its associated path

V. Sequential Monte Carlo Methods

A. Background

A different approach to the problem of detecting the active users in the considered DS-CDMA system, along with their transmitted data, is to use SMC methods, which are also known as particle filters (PFs) [23]–[25]. Here, we give a brief introduction to PFs and derive, afterward, the optimal SMC algorithm for the problem at hand.

Most particle filtering methods rely upon the principle of importance sampling (IS) [23] for building an empirical approximation of a desired pdf, e.g., $p(x)$, by drawing samples from a different distribution, which is known as *importance function* or *proposal pdf*, denoted by $q(x)$, with a domain that includes that of $p(x)$. These samples are then assigned appropriate normalized *importance* weights, i.e.,

$$x^{(i)} \sim q(x) \tag{47}$$

$$\tilde{w}^{(i)} \propto \frac{p(x^{(i)})}{q(x^{(i)})} \tag{48}$$

$$w^{(i)} = \frac{\tilde{w}^{(i)}}{\sum_{i=1}^{M} \tilde{w}^{(i)}} \tag{49}$$

where $x^{(i)}$ is the ith particle, M is the number of particles, $\tilde{w}^{(i)}$, $i = 1, \ldots, M$, are the unnormalized weights, and $w^{(i)}$ are the normalized weights, so that $\sum_{i=1}^{M} w^{(i)} = 1$. The particles drawn from $q(x)$ are said to be *properly weighted* with respect to $p(x)$ because

$$\begin{aligned}\frac{\mathbb{E}_q\left[f(x)\tilde{w}(x)\right]}{\mathbb{E}_q\left[\tilde{w}(x)\right]} &= \frac{\int f(x)\tilde{w}(x)q(x)dx}{\int \tilde{w}(x)q(x)dx} \\ &= \frac{\int f(x)k\frac{p(x)}{q(x)}q(x)dx}{\int k\frac{p(x)}{q(x)}q(x)dx} \\ &= \frac{\int f(x)p(x)dx}{\int p(x)dx} = \mathbb{E}_p\left[f(x)\right]\end{aligned} \tag{50}$$

where $f(x)$ is an arbitrary integrable function, and $\tilde{w}(x) = kp(x)/q(x)$, where k is an arbitrary constant. Intuitively, (50) means that the expectation of function $f(x)$ with respect to a pdf, i.e., $p(x)$, can be computed with respect to another pdf, i.e., $q(x)$, if the function is properly weighted using factor $\tilde{w}(x)$. Note that the normalization step in (49) naturally arises when the fraction on the left-hand side of (50) is approximated using samples from q, i.e.,

$$\begin{aligned}\frac{\mathbb{E}_q\left[f(x)\tilde{w}(x)\right]}{\mathbb{E}_q\left[\tilde{w}(x)\right]} &= \frac{\frac{1}{M}\sum_{i=1}^{M} f(x^{(i)})\tilde{w}(x^{(i)})}{\frac{1}{M}\sum_{i=1}^{M} \tilde{w}(x^{(i)})} \\ &= \sum_{i=1}^{M} f(x^{(i)})w(x^{(i)}).\end{aligned} \tag{51}$$

At the receiver, we are interested in detecting the sequence of symbol vectors $\mathbf{s}_{0:t}$ given the sequence of observations. Therefore, we need to approximate the *a posteriori* pmf[2] $p(\mathbf{s}_{0:t}|\mathbf{y}_{0:t})$, which contains all relevant statistical information for the optimal (Bayesian) estimation of $\mathbf{s}_{0:t}$ (and, hence, $\mathbf{u}_{0:t}$).

[2]Note that any pmf can be handled as a pdf by representing it using Dirac delta functions.

Accordingly, we choose an importance function with the form $q(\mathbf{s}_{0:t}|\mathbf{y}_{0:t})$.

One of the most appealing features of the particle filtering approach is its potential for online processing. Indeed, the IS principle can be sequentially applied by exploiting the recursive decomposition of the posterior distribution, i.e.,

$$p(\mathbf{s}_{0:t}|\mathbf{y}_{0:t}) \propto p(\mathbf{y}_t|\mathbf{s}_{0:t}, \mathbf{y}_{0:t-1})p(\mathbf{s}_t|\mathbf{s}_{t-1})p(\mathbf{s}_{0:t-1}|\mathbf{y}_{0:t-1}) \tag{52}$$

which is obtained by way of Bayes' theorem. The recursive algorithm called sequential IS (SIS) [24] combines the IS principle, decomposition (52), and an importance pmf that can be factored as

$$q(\mathbf{s}_{0:t}|\mathbf{y}_{0:t}) = q(\mathbf{s}_t|\mathbf{s}_{0:t-1}, \mathbf{y}_{0:t})q(\mathbf{s}_{0:t-1}|\mathbf{y}_{0:t-1}) \tag{53}$$

to build a discrete probability measure with random support that approximates the posterior pmf [24]. Let $\Omega_t = \{\mathbf{s}_{0:t}^{(i)}, w_t^{(i)}\}_{i=1}^M$ denote the set of weighted particles at time t. Then, the desired pmf is approximated as

$$\hat{p}(\mathbf{s}_{0:t}|\mathbf{y}_{0:t}) = \sum_{i=1}^{M} \delta\left(\mathbf{s}_{0:t} - \mathbf{s}_{0:t}^{(i)}\right) w_t^{(i)}. \tag{54}$$

When a new observation is collected at time $t+1$, the SIS algorithm proceeds as follows to recursively compute Ω_{t+1}:

1) IS: $\mathbf{s}_{t+1}^{(i)} \sim q(\mathbf{s}_{t+1}|\mathbf{s}_{0:t}^{(i)}, \mathbf{y}_{0:t+1})$;
2) Weight update: $\tilde{w}_{t+1}^{(i)} = w_t^{(i)} p(\mathbf{y}_{t+1}|\mathbf{s}_{0:t+1}^{(i)}, \mathbf{y}_{0:t})/q(\mathbf{s}_{t+1}^{(i)}| \mathbf{s}_{0:t}^{(i)}, \mathbf{y}_{0:t+1})$;
3) Weight normalization: $w_{t+1}^{(i)} = \tilde{w}_{t+1}^{(i)} / \sum_{k=1}^{M} \tilde{w}_{t+1}^{(k)}$.

The asymptotic convergence, for $M \to \infty$, of the SIS algorithm is proved in [26] under mild assumptions. It is straightforward to obtain data estimates from the approximate pmf $\hat{p}(\mathbf{s}_{0:t}|\mathbf{y}_{0:t})$. For example, a marginal MAP detector can be implemented as

$$\hat{\mathbf{s}}_t^{\mathrm{MAP}} = \arg\max_{\mathbf{s}_t} \left\{ \sum_{i=1}^{M} \delta(\mathbf{s}_t - \mathbf{s}_t^{(i)}) w_t^{(i)} \right\} \tag{55}$$

which amounts to selecting the particle with the highest accumulated weight (note that some particles can be replicated).

One major problem in the practical implementation of the SIS algorithm is that after a few time steps, most of the particles have importance weights with negligible values (very close to zero). This is known as the weight degeneracy phenomenon [24]. The common solution to this problem is to *resample* the particles. Resampling is an algorithmic step that stochastically discards particles with small weights while replicating those with significant weight [24], [27]. In its conceptually simplest form, resampling generates M new particles $\{\mathbf{s}_{0:t}^{(i)}, 1/M\}_{i=1}^M$ by drawing samples from the discrete pmf $p_w(\mathbf{s}_{0:t}^{(i)}) = w_t^{(i)}$.

B. Optimal SMC Receiver

It is well known that the performance of any particle filtering method depends to a great extent on the choice of the importance function. The optimal importance function is always the pdf itself that we want to approximate, and thus, in this case, we have[3]

$$q(\mathbf{s}_t|\mathbf{s}_{0:t-1}, \mathbf{y}_{0:t}) = p(\mathbf{s}_t|\mathbf{s}_{0:t-1}, \mathbf{y}_{0:t}). \tag{56}$$

By means of Bayes' theorem, the conditional probability on the right-hand side of (56) can be computed as

$$p(\mathbf{s}_t|\mathbf{s}_{0:t-1}, \mathbf{y}_{0:t}) = \frac{p(\mathbf{y}_t|\mathbf{s}_{0:t}, \mathbf{y}_{0:t-1})p(\mathbf{s}_t|\mathbf{s}_{0:t-1}, \mathbf{y}_{0:t-1})}{p(\mathbf{y}_t|\mathbf{s}_{0:t-1}, \mathbf{y}_{0:t-1})} \tag{57}$$

where the normalization constant $p(\mathbf{y}_t|\mathbf{s}_{0:t-1}, \mathbf{y}_{0:t-1})$ can be exactly evaluated; hence

$$p(\mathbf{s}_t|\mathbf{s}_{0:t-1}, \mathbf{y}_{0:t}) = \frac{p(\mathbf{y}_t|\mathbf{s}_{0:t}, \mathbf{y}_{0:t-1})p(\mathbf{s}_t|\mathbf{s}_{t-1})}{\sum_{\tilde{\mathbf{s}}_t} p(\mathbf{y}_t|\mathbf{s}_{0:t-1}, \tilde{\mathbf{s}}_t, \mathbf{y}_{0:t-1})p(\tilde{\mathbf{s}}_t|\mathbf{s}_{t-1})}. \tag{58}$$

Notice that in (58), we have also taken into account that when the corresponding observation, i.e., $\mathbf{y}_t$, is not available, the symbols transmitted at time t only depend on those transmitted at time $t-1$ (due to the user activity model), i.e., $p(\mathbf{s}_t|\mathbf{s}_{0:t-1}, \mathbf{y}_{0:t-1}) = p(\mathbf{s}_t|\mathbf{s}_{t-1})$.

The likelihood of the form $p(\mathbf{y}_t|\mathbf{s}_{0:t}, \mathbf{y}_{0:t-1})$ that appears in (58) can be analytically computed using a Kalman filter. Indeed, recalling that $\mathbf{h}_t$ is a vector built by taking all the elements in the diagonal of matrix $\mathbf{H}_t$, we can write

$$p(\mathbf{y}_t|\mathbf{s}_{0:t}, \mathbf{y}_{0:t-1}) = \int p(\mathbf{y}_t|\mathbf{h}_t, \mathbf{s}_{0:t}, \mathbf{y}_{0:t-1})p(\mathbf{h}_t|\mathbf{s}_{0:t}, \mathbf{y}_{0:t-1})d\mathbf{h}_t. \tag{59}$$

Both densities in the integrand are Gaussian, and therefore, the integral can be solved [28]. Specifically, it can be shown (see, e.g., [29, App.] for details) that

$$p(\mathbf{y}_t|\mathbf{s}_{0:t}, \mathbf{y}_{0:t-1}) = \mathcal{N}\left(\mathbf{y}_t|\mathbf{B}_t\left(\alpha_1\mathbf{h}_{t-2} + \alpha_2\mathbf{h}_{t-1}\right), \mathbf{B}_t\mathbf{K}_{t|t}\mathbf{B}_t^\top\right) \tag{60}$$

where $\mathbf{K}_{t|t}$ is the filtered covariance of $\mathbf{h}_t$ computed by the KF, and $\mathbf{B}_t = \mathbf{C}\mathbf{S}_t$.

The weight update equation for importance function (56) can be easily obtained combining (52), (53), (56), and (58) to yield

$$\tilde{w}_t^{(i)} = \frac{p\left(\mathbf{s}_{0:t}^{(i)}|\mathbf{y}_{0:t}\right)}{q\left(\mathbf{s}_{0:t}^{(i)}|\mathbf{y}_{0:t}\right)} = w_{t-1}^{(i)} \sum_{\tilde{\mathbf{s}}_t} p\left(\mathbf{y}_t|\mathbf{s}_{0:t-1}^{(i)}, \tilde{\mathbf{s}}_t, \mathbf{y}_{0:t-1}\right) p\left(\tilde{\mathbf{s}}_t|\mathbf{s}_{t-1}^{(i)}, \mathbf{y}_{0:t-1}\right). \tag{61}$$

[3]Notice that we are only defining the non-recursive part of the proposal function $q(\mathbf{s}_{0:t}|\mathbf{y}_{0:t})$ since the latter is *selected* to decompose according to (53).

Pseudocode 4 provides a complete description of the operation of the algorithm.

Pseudocode 4 Optimal SMC receiver

for every particle $i = 1, \ldots, M$ **do** // `initialization`
  initialize a KF to estimate the channel associated with particle i
  for every possible $\mathbf{s}_0 \in (\mathcal{A} \cup \{0\})^N$ **do**
    compute $p(\mathbf{s}_t|\mathbf{s}_{0:t-1}^{(i)}, \mathbf{y}_{0:t})|_{t=0} = p(\mathbf{s}_0|\mathbf{y}_0)$ using (58), noting that at time $t = 0$

$$p(\mathbf{s}_t|\mathbf{s}_{t-1}) = p(\mathbf{s}_0)$$
$$p(\mathbf{y}_t|\mathbf{s}_{0:t-1}, \tilde{\mathbf{s}_t}, \mathbf{y}_{0:t-1}) = p(\mathbf{y}_t|\tilde{\mathbf{s}_0})$$
$$p(\tilde{\mathbf{s}}_t|\mathbf{s}_{t-1}) = p(\tilde{\mathbf{s}_0})$$

  draw a sample, i.e., $\mathbf{s}_0^{(i)}$, from $q(\mathbf{s}_0|\mathbf{y}_0) = p(\mathbf{s}_0|\mathbf{y}_0)$
  compute the unnormalized weight of the particle $\tilde{w}_0^{(i)} = \sum_{\tilde{\mathbf{s}}0} p(\mathbf{y}_t|\tilde{\mathbf{s}}_0)p(\tilde{\mathbf{s}}_0)$
  update the KF associated with the ith particle using the drawn sample $\mathbf{s}_0^{(i)}$
compute the sum of the unnormalized weights $W \leftarrow \sum_i^M \tilde{w}_0^{(i)}$
normalize each weight $w_0^{(i)} \leftarrow \tilde{w}_0^{(i)}/W$
for $t > 1$ **do** // `sequential step`
  for every particle $i = 1, \ldots, M$ **do**
    for every possible $\mathbf{s}_t \in (\mathcal{A} \cup \{0\})^N$ **do**
      compute $p(\mathbf{s}_t|\mathbf{s}_{0:t-1}^{(i)}, \mathbf{y}_{0:t})$ using (58)
    draw a sample, i.e., $\mathbf{s}_t^{(i)}$, from $q(\mathbf{s}_t|\mathbf{s}_{0:t-1}^{(i)}, \mathbf{y}_{0:t}) = p(\mathbf{s}_t|\mathbf{s}_{0:t-1}^{(i)}, \mathbf{y}_{0:t})$
    compute the unnormalized weight of the ith particle, i.e., $\tilde{w}_t^{(i)}$, using (61)
    update the KF associated with the ith particle using sample $\mathbf{s}_t^{(i)}$
  if weight degeneracy phenomenon occurs **then**
    resample the particles
  else
    compute the sum of the unnormalized weights $W \leftarrow \sum_i^M \tilde{w}_t^{(i)}$
    normalize each weight $w_t^{(i)} \leftarrow \tilde{w}_t^{(i)}/W$
choose the particle with the highest weight
return its associated path

From (58), it is clear that the random sampling of the vector of symbols transmitted at time t requires the computation of $|\mathcal{A} \cup \{0\}|^N$ likelihood and $|\mathcal{A} \cup \{0\}|^N$ symbol vector probabilities. This entails an exponential complexity on the number of users, which is prohibitive in most practical cases. Therefore, alternative methods providing a tradeoff between performance and complexity are desirable. One of such methods, which is also based on particle filtering, is introduced in the following section.

C. Complexity-Constrained SMC Receiver

To get a complexity-constrained algorithm that solves the stated problem, we propose an SMC scheme based on the ideas of sampling in a higher dimension [30] and sequentially on the data space [31]. Specifically, we address the approximation of the joint pdf $p(\mathbf{s}_{0:t}, \mathbf{H}_t|\mathbf{y}_{0:t})$ using an importance function of the form $q(\mathbf{s}_{0:t}, \mathbf{H}_t|\mathbf{y}_{0:t})$, which will be defined as follows.

Using Bayes' theorem, the target pdf can be decomposed as

$$\begin{aligned} &p(\mathbf{s}_{0:t}, \mathbf{H}_t|\mathbf{y}_{0:t}) \\ &\quad= \frac{p(\mathbf{y}_t|\mathbf{H}_t, \mathbf{s}_{0:t}, \mathbf{y}_{0:t-1})p(\mathbf{s}_{0:t}, \mathbf{H}_t|\mathbf{y}_{0:t-1})}{p(\mathbf{y}_t|\mathbf{y}_{0:t-1})} \\ &\quad\propto p(\mathbf{y}_t|\mathbf{H}_t, \mathbf{s}_{0:t}, \mathbf{y}_{0:t-1})p(\mathbf{s}_{0:t}, \mathbf{H}_t|\mathbf{y}_{0:t-1}) \\ &\quad= p(\mathbf{y}_t|\mathbf{H}_t, \mathbf{s}_{0:t}, \mathbf{y}_{0:t-1})p(\mathbf{s}_t|\mathbf{s}_{0:t-1}, \mathbf{H}_t, \mathbf{y}_{0:t-1}) \\ &\qquad\times p(\mathbf{s}_{0:t-1}, \mathbf{H}_t|\mathbf{y}_{0:t-1}) \\ &\quad= p(\mathbf{y}_t|\mathbf{H}_t, \mathbf{s}_{0:t}, \mathbf{y}_{0:t-1})p(\mathbf{s}_t|\mathbf{s}_{t-1}) \\ &\qquad\times p(\mathbf{H}_t|\mathbf{s}_{0:t-1}, \mathbf{y}_{0:t-1})p(\mathbf{s}_{0:t-1}|\mathbf{y}_{0:t-1}) \end{aligned} \quad (62)$$

where the proportionality is due to $p(\mathbf{y}_t|\mathbf{y}_{0:t-1})$ being a normalization constant that does not depend on either the sequence of symbol vectors, i.e., $\mathbf{s}_{0:t}$, or on channel matrix $\mathbf{H}_t$. In addition, it has been taken into account that the channel and the symbols are *a priori* (when no observation connecting them is available) independent, and hence

$$p(\mathbf{s}_t|\mathbf{s}_{t-1}, \mathbf{H}_t, \mathbf{y}_{0:t-1}) = p(\mathbf{s}_t|\mathbf{s}_{t-1}). \quad (63)$$

At the sight of (1), the likelihood $p(\mathbf{y}_t|\mathbf{H}_t, \mathbf{s}_{0:t}, \mathbf{y}_{0:t-1})$ on the right-hand side of (62) is Gaussian with mean $\mathbf{C}\mathbf{H}_t\mathbf{s}_t$ and covariance $\sigma_z^2\mathbf{I}_L$ and, thus, can be easily computed. On the other side, the proposal function is, as usual, recursively defined. Specifically, we draw samples from an importance function of the form

$$\begin{aligned} &q(\mathbf{s}_{0:t}, \mathbf{H}_t|\mathbf{y}_{0:t}) \\ &\quad= q(\mathbf{s}_t|\mathbf{s}_{0:t-1}, \mathbf{H}_t, \mathbf{y}_{0:t})q(\mathbf{s}_{0:t-1}, \mathbf{H}_t|\mathbf{y}_{0:t}) \\ &\quad= q(\mathbf{s}_t|\mathbf{s}_{0:t-1}, \mathbf{H}_t, \mathbf{y}_{0:t})p(\mathbf{H}_t|\mathbf{s}_{0:t-1}, \mathbf{y}_{0:t-1}) \\ &\qquad\times q(\mathbf{s}_{0:t-1}|\mathbf{y}_{0:t-1}) \end{aligned} \quad (64)$$

and we only need to define the marginal proposal $q(\mathbf{s}_t|\mathbf{s}_{0:t-1}, \mathbf{H}_t, \mathbf{y}_{0:t})$. This last probability is conditional on channel matrix $\mathbf{H}_t$ and observation vectors $\mathbf{y}_{0:t}$, from which soft estimates of the symbols transmitted at time t can be obtained as

$$\hat{\mathbf{s}}_t = \mathbf{F}_t\mathbf{y}_t \quad (65)$$

where $\mathbf{F}_t$ is any linear filter, such as a decorrelator or a minimum mean square error (MMSE) filter [32]. Once a soft estimate, i.e., $\hat{s}_t$, of the jth symbol, i.e., s_t, is available, we draw a sample of the latter from the proposal

$$\begin{aligned} &q\left(s_t(j)|s_{t-1}(j), \hat{s}_t(j)\right) \\ &\quad= p\left(s_t(j)|s_{t-1}(j), \hat{s}_t(j)\right) \\ &\quad= \frac{p\left(\hat{s}_t(j)|s_t(j)\right) p\left(s_t(j)|s_{t-1}(j)\right)}{\sum_{\tilde{s}_t(j)} p\left(\hat{s}_t(j)|\tilde{s}_t(j)\right) p\left(\tilde{s}_t(j)|s_{t-1}(j)\right)}. \end{aligned} \quad (66)$$

The marginal proposal for the entire symbol vector is then defined as

$$q(\mathbf{s}_t|\mathbf{s}_{0:t-1}, \mathbf{H}_t, \mathbf{y}_{0:t}) = \prod_{j=1}^{N} q\left(s_t(j)|s_{t-1}(j), \hat{s}_t(j)\right) \tag{67}$$

and we have decoupled the sampling of the symbols within vector $\mathbf{s}_t$. Notice that probabilities $p(s_t(j)|s_{t-1}(j))$, $j = 1, \ldots, N$, in (66) can be easily calculated using (39). Regarding $p(\hat{s}_t(j)|s_t(j))$, we assume that the soft estimates follow a Gaussian distribution, and thus

$$q\left(\hat{s}_t(j)|s_t(j)\right) = \mathcal{N}\left(\hat{s}_t(j)|s_t(j), \sigma_q^2\right) \tag{68}$$

whose mean is the corresponding symbol that is being estimated and whose variance, i.e., σ_q^2, is matched to that of r.v. $\hat{s}_t(j) - s_t(j)$. Notice that this is just a model we assume for the sake of simplicity at the time of sampling. Soft estimate $\hat{s}_t(j)$ is not necessarily Gaussian distributed.

Equation (67) combined with (64) yields the (overall) proposal function, i.e.,

$$\begin{aligned} &q(\mathbf{s}_{0:t}, \mathbf{H}_t|\mathbf{y}_{0:t}) \\ &\quad = p(\mathbf{s}_t|\mathbf{s}_{t-1}) \prod_{j=1}^{N} \frac{p\left(\hat{s}_t(j)|s_t(j)\right)}{\sum_{\tilde{s}_t(j)} p\left(\hat{s}_t(j)|\tilde{s}_t(j)\right) p\left(\tilde{s}_t(j)|s_{t-1}(j)\right)} \\ &\qquad \times p(\mathbf{H}_t|\mathbf{s}_{0:t-1}, \mathbf{y}_{0:t-1}) q(\mathbf{s}_{0:t-1}|\mathbf{y}_{0:t-1}) \end{aligned} \tag{69}$$

and from (62) and (69), the resulting weight update equation is

$$\begin{aligned} \tilde{w}_t^{(i)} &= \frac{p\left(\mathbf{s}_{0:t}^{(i)}, \mathbf{H}_t^{(i)}|\mathbf{y}_{0:t}\right)}{q\left(\mathbf{s}_{0:t}^{(i)}, \mathbf{H}_t^{(i)}|\mathbf{y}_{0:t}\right)} \\ &= p\left(\mathbf{y}_t|\mathbf{H}_t^{(i)}, \mathbf{s}_{0:t}^{(i)}, \mathbf{y}_{0:t-1}\right) \\ &\quad \times \frac{\prod_{j=1}^{N} \sum_{\tilde{s}_t(j)} p\left(\hat{s}_t^{(i)}(j)|\tilde{s}_t(j)\right) p\left(\tilde{s}_t(j)|s_{t-1}^{(i)}(j)\right)}{\prod_{j=1}^{N} p\left(\hat{s}_t(j)|s_t(j)\right)} \\ &\quad \times w_{t-1}^{(i)}. \end{aligned} \tag{70}$$

Pseudocodes 5 and 6 summarize, respectively, the initialization and the sequential step of the algorithm.

Pseudocode 5 Complexity-constrained SMC receiver `Initialization`

for every particle $i = 1, \ldots, M$ **do**
  initialize a KF to estimate the channel associated with particle i
  draw a sample, i.e., $\hat{\mathbf{H}}_0^{(i)}$, using the filtered mean and covariance given by the KF
  from $\hat{\mathbf{H}}_0^{(i)}$ and $\mathbf{y}_0$, compute soft estimates, i.e., $\hat{\mathbf{s}}_0^{(i)}$, of the symbols transmitted
  for $j = 1, \ldots, N$ **do**
    for every possible $s_t(j)$ **do**
      compute $q(s_t(j)|s_{t-1}(j)^{(i)}, \hat{s}_t^{(i)}(j))|_{t=0} = q(s_0(j)|\hat{s}_0^{(i)}(j))$ using (66) noting that at time $t = 0$

$$p\left(s_t(j)|s_{t-1}(j)\right) = p\left(s_0(j)\right)$$
$$p\left(\tilde{s}_t(j)|s_{t-1}(j)\right) = p\left(\tilde{s}_0(j)\right)$$

    draw a sample, i.e., $s_0^{(i)}(j)$, from $q(s_0(j)|\hat{s}_0^{(i)}(j))$
  compute the unnormalized weight of the particle using (70)
  update the KF associated with the ith particle using the sample drawn $\mathbf{s}_0^{(i)} = [s_0^{(i)}(1), s_0^{(i)}(2), \ldots, s_0^{(i)}(N)]^\top$
compute the sum of the unnormalized weights $W \leftarrow \sum_i^M \tilde{w}_0^{(i)}$
normalize each weight $w_0^{(i)} \leftarrow \tilde{w}_0^{(i)}/W$

Pseudocode 6 Complexity-constrained SMC receiver `Sequential step`

for $t > 1$ **do**
  for every particle $i = 1, \ldots, M$ **do**
    draw a sample, i.e., $\hat{\mathbf{H}}_t^{(i)}$, using the predictive mean and covariance given by the KF
    from $\hat{\mathbf{H}}_t^{(i)}$ and $\mathbf{y}_t$, compute soft estimates, i.e., $\hat{\mathbf{s}}_t^{(i)}$, of the symbols transmitted
    for $j = 1, \ldots, N$ **do**
      for every possible $s_t(j)$ **do**
        compute $q(s_t(j)|s_{t-1}^{(i)}(j), \hat{s}_t^{(i)}(j))$ using (66)
      draw a sample, i.e., $s_t^{(i)}(j)$, from $q(s_t(j)|s_{t-1}^{(i)}(j), \hat{s}_t^{(i)}(j))$
    compute the unnormalized weight of the particle, i.e., $\tilde{w}_t^{(i)}$, using (70)
    update the KF associated with the ith particle using the sample drawn $\mathbf{s}_t^{(i)} = [s_t^{(i)}(1), s_t^{(i)}(2), \ldots, s_t^{(i)}(N)]^\top$
  if weight degeneracy phenomenon occurs **then**
    resample the particles
  else
    compute the sum of the unnormalized weights $W \leftarrow \sum_i^M \tilde{w}_t^{(i)}$
    normalize each weight $w_t^{(i)} \leftarrow \tilde{w}_t^{(i)}/W$
choose the particle with the highest weight
return its associated path

VI. Simulation Results

A. Setup

We have carried out a set of computer simulations to assess the performance of the proposed algorithms. For simplicity, the considered DS-CDMA system employs spreading codes of length $L = 8$ (randomly generated) and can allocate up to $N = 3$ users. The modulation format is BPSK, and transmission is carried out in frames of $K = 1000$ symbols per user.

The probability of a user becoming active (starting to transmit at a certain time instant given that it was not transmitting at the previous time instant) is $\eta = 0.01$, whereas the probability of a user staying active (transmitting at a certain time instant given that it was transmitting at the previous time instant) is $\mu = 0.99$.

The channel coefficients associated with different users are assumed to be independent. Their dynamics are modeled using the multidimensional AR process (14) for the purpose of algorithm design. However, to assess the performance of the proposed algorithms in a more realistic environment, we have carried out our computer simulations assuming channel impulse responses (CIRs) generated using the classical Clarke autocorrelation function [33]. Thus, the Doppler spectrum of every user coefficient is set to be the classical Jake's spectrum [34, Sec. 5.4], which yields Clarke's autocorrelation function, i.e.,

$$\rho(k) = \mathcal{J}_0\left(2\pi \frac{v_m}{c} F_c k T_s\right)$$

where k is the delay in symbol periods, $\mathcal{J}_0$ is the Bessel function of the first kind and zeroth order, $v_m = 50$ km/h is the relative speed of the transmitter with respect to the receiver, c is the speed of light, $F_c = 2$ GHz is the carrier frequency, and $T_s = 1/(500 \cdot 10^3)$ s is the symbol period. The latter entails a symbol rate of $f_s = 500 \cdot 10^3$ symbols (bits) per second.

Given a certain autocorrelation function (such as Clarke's autocorrelation function), the parameters of the AR process (coefficients and variance) that best approximate it can be obtained by solving the Yule–Walker equations [22]. However, plugging the values of the parameters they yield in this particular case ($\alpha_1 = 1.99999$, $\alpha_2 = -0.999998$, and $\sigma_v^2 = 3.84826 \cdot 10^{-11}$) into the KF results in poor performance. This is due to the small value assigned to variance σ_v^2, which does not account for the uncertainty in the estimation of the channel.[4] Through computer simulations, we have found that reasonable values for the parameters of the AR process are $\alpha_1 = 0.59999$, $\alpha_2 = 0.39999$, and $\sigma_v^2 = 0.0001$ (the value selected for σ_v^2 being the most critical one).

As for the signature sequences of the different users, the matrix of spreading codes, i.e., $\mathbf{C}$, is randomly generated at each frame. Each entry is drawn from a uniform distribution taking values on set $\{+1, -1\}$. The matrix is only dismissed (and a new matrix is generated following the same procedure) if it is rank deficient.

All the results presented herein are referred to a particular user of interest, which is chosen to be (without loss of generality) the first user. So that it is possible to compute the symbol error rate (SER) at every frame for this user, the *a priori* probability of this user being active is $\phi = 1$. However, after the first time instant, the reference user can leave or enter the system (according to the model) at any time as the rest of the users, and the receiver has to detect such events. As for the rest of the users, the (*a priori*) probability of each one being active at the beginning of a frame is $\phi = 0.5$.

A common issue in DS-CDMA systems is the so-called *near–far* problem [2]: the power received for the user of interest can be very small when compared with that received for interfering users. Since our goal is to compare the performance of the different algorithms rather than to tackle this well-known difficulty associated with CDMA and other wireless communication systems, all channel realizations that yield a signal-to-interference ratio below -30 dBs (the power of the interfering users is 1000 times bigger than that of the user of interest) are discarded in the simulations.

B. Algorithms and Figures of Merit

Within this simulation setup, we have compared

- the ML receiver when the channel is perfectly known, but the model governing the users' activity is ignored (hence, all the sequences of symbol vectors, i.e., $\mathbf{s}_{1:K}$, are assumed equally likely), labeled "Naive ML (known channel)";
- the optimal (MAP) receiver based on the VA that has full knowledge of the CIR, as developed in Section III, labeled "Optimal (known channel)";
- the PSP-based method explained in Section IV, labeled "PSP";
- the optimal SIS receiver derived in Section V-B, labeled "SIS-opt";
- the complexity-constrained SIS method based on the use of linear filters presented in Section V-C, here using MMSE filters, labeled "SIS-LF" (where LF stands for "linear filter");
- the finite random set (FRS)-based method developed in [7] considering a grid with $W = 20$ points equally spaced between $h_{\min} = -2.0$ and $h_{\max} = 2.0$ for the quantization of the channel coefficients, labeled "FRS ($W = 20$)."

The optimal (Viterbi-based) and PSP algorithms perform a search over a typical trellis diagram with $|\{\mathcal{A} \cup \{0\}\}|^N = 3^3 = 27$ states. The number of survivors employed by the PSP algorithm is $P = 54$ (which amounts to keeping two survivors per state), and this is also the number of particles employed by the two SMC algorithms (i.e., $M = P$).

We evaluate the performance of the receivers in terms of SER, activity detection error rate (ADER), and mean square error (MSE) of the channel estimates. When computing the SER, an error is computed at time t whenever the *symbol* $\hat{s}_t(1) \in \{\mathcal{A} \cup \{0\}\}$ estimated for the user of interest differs from the actual transmitted *symbol*, i.e., $s_t(1)$. This means that an error in activity detection also entails a symbol error.[5] Regarding the MSE, if we let $\hat{\mathbf{h}}_t$ denote the vector containing the coefficients in the diagonal of channel estimate $\hat{\mathbf{H}}_t$ and $\mathbf{h}_t$ denote the vector containing those in the diagonal of the

[4]Notice that uncertainty in the estimation of the channel is always present as long as *estimates* of the transmitted symbols (rather than the transmitted symbols themselves) are fed to the channel estimator.

[5]The opposite is not true, however, when calculating the ADER: if the user of interest is active, $\hat{s}_t(1) \neq s_t(1)$ entails no activity detection error as long as $\hat{s}_t(1) \neq 0$.

TABLE I
PARAMETERS OF THE SIMULATED SYSTEM

modulation format	BPSK
spreading codes length, L	8
max. number of users, N	3
frame length (symbol vectors), K	1000
a priori prob (user being active), ϕ	0.5
prob (idle user becoming active), η	0.01
prob (active user staying active), μ	0.99
AR process coeffs., α_1, α_2	0.59999, 0.39999
AR process variance, σ_v^2	0.0001
speed (km/h), v_m	50
carrier frequency (GHz), F_c	2
symbol rate (symbols/second), f_s	$500 \cdot 10^3$
number particles/survivors, $M = P$	54
grid size, W	20

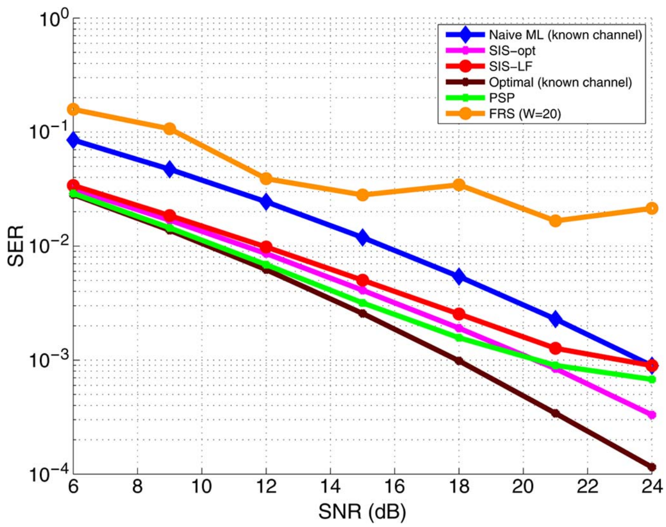

Fig. 1. SER for several values of the SNR (in decibels). The results are averaged over 45 000 independent data frames, where each one has its own associated channel realization and spreading-code matrix.

actual channel matrix $\mathbf{H}_t$, then the normalized MSE at time t is computed as

$$\mathrm{MSE}_t = \frac{\left(\mathbf{h}_t - \hat{\mathbf{h}}_t\right)^H \left(\mathbf{h}_t - \hat{\mathbf{h}}_t\right)}{\hat{\mathbf{h}}_t^H \hat{\mathbf{h}}_t}. \quad (71)$$

The overall normalized MSE is then obtained by averaging over all time steps, i.e.,

$$\mathrm{MSE} = \frac{1}{K} \sum_{t=0}^{K-1} \mathrm{MSE}_t. \quad (72)$$

All the algorithms that perform channel estimation rely on the KF to do so.

Table I summarizes the values of the model parameters that we have used in the computer simulations.

C. Numerical Results

Fig. 1 shows the SER achieved by all the receivers for different values of the SNR. For low SNRs (up to 10 dBs), the PSP receiver performs very close to the optimal receiver with known channel. The gap between the two algorithms widens as the SNR increases, but still, for SER $\approx 10^{-2.5}$, the curve of the PSP is only 1 dB away from that of the optimal receiver. The SMC algorithms, on the other hand, perform slightly worse for SNRs below 21 dBs. Specifically, the SIS-opt and SIS-LF receivers exhibit 3- and 4-dBs losses, respectively, when compared with the optimal receiver. It can be also observed that the slope of the PSP algorithm's curve becomes less steep for high SNRs (above 18 dBs), which translates into some performance penalty. The SIS-opt receiver does not suffer this performance degradation, and for SNRs higher than 21 dBs, it outperforms the PSP receiver. The FRS $(W = 20)$ algorithm displays poor performance in the whole range of SNRs, which is mainly caused by the channel estimation error introduced in quantization and, to a lesser extent, by the error propagation phenomenon. Further results for this receiver will be shown later.

One remarkable fact in Fig. 1 is the relatively poor performance of the Naive ML (known channel) algorithm, which points out the importance of exploiting the available statistical information concerning the dynamics of the users' activity.

Fig. 2 displays the channel estimation MSE attained by the PSP, optimal SMC equalizer, and complexity-constrained SMC equalizer, together with the corresponding error curve achieved by a KF fed with the true transmitted symbols (without errors), that serves as a lower bound and reference for comparison. The figure shows that the PSP algorithm is clearly outperformed by the two PFs for low SNRs, but only by the SIS-opt in high SNRs. However, for medium-high SNRs, the SIS-opt and PSP algorithms usually yield channel estimates with similar MSE, and the gap between the curves observed in Fig. 2 is mainly due to the MSE achieved by the PSP algorithm having a higher variance. The FRS $(W = 20)$ algorithm incurs, due to the quantization of channel space, on severe channel estimation errors (even for high SNRs), as illustrated by the MSE it attains, which is approximately two orders of magnitude above that of the remaining algorithms (recall that this is a *normalized* MSE).

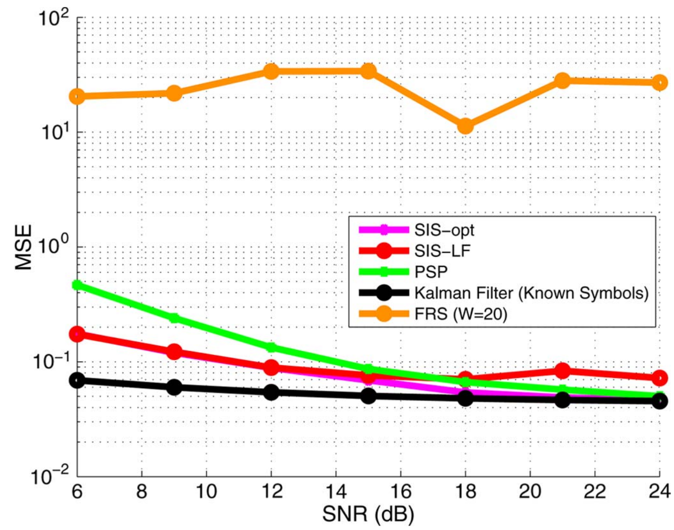

Fig. 2. MSE for several values of the SNR (in decibels). The results are averaged over 45 000 independent data frames, where each one has its own associated channel realization and spreading-code matrix.

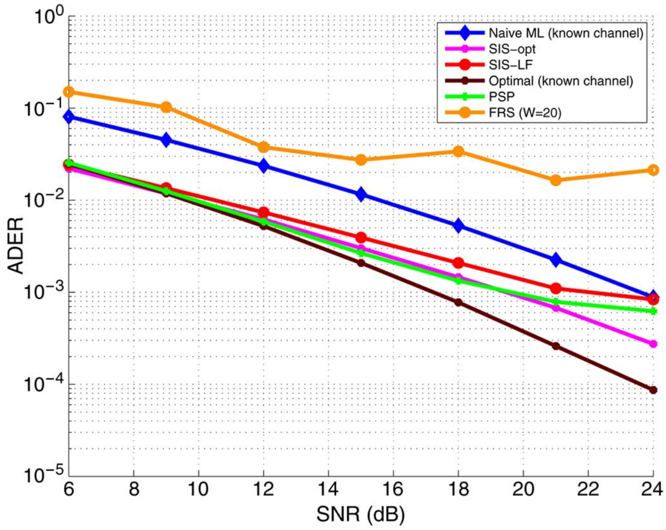

Fig. 3. ADER for several values of the SNR (in decibels). The results are averaged over 45 000 independent data frames, where one has its own associated channel realization and spreading-code matrix.

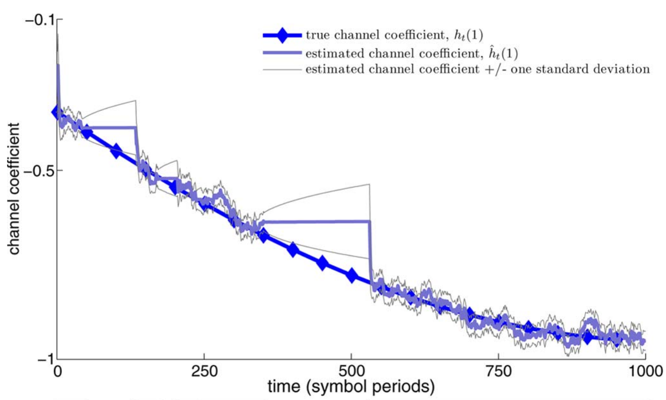

Fig. 4. Estimation of the user of interest's channel coefficient carried out by the SIS-opt algorithm for a specific frame when SNR = 15 dBs. The thick solid line below the plot indicates the activity of the user of interest: markers are only present when it is active; hence, the gaps correspond to periods of inactivity.

To evaluate the reliability of user activity detection, Fig. 3 displays the ADER against several values of the SNR. It can be seen that Fig. 3 is very similar to Fig. 1. Indeed, we have found that most *symbol* errors accounted when computing the SER were due to the algorithms not correctly detecting the activity of the user of interest.

Finally, we assess the effect of the users' inactivity in the channel estimation algorithms. Notice that the channel coefficient associated with a given user evolves regardless of the latter's activity, and hence, when the user is not transmitting, the channel can only be predicted using the AR model given by (14).

Fig. 4 shows the estimation of the user of interest's channel coefficient carried out by SIS-opt using the KF. The curve of the estimated channel coefficient is given by the filtered mean computed by the KF, whereas the two thin lines above and below it account for the mean plus and minus one standard deviation, respectively. As for the lower part of the figure, it shows the user activity (markers are present only when the user is transmitting an actual symbol.) The figure shows how

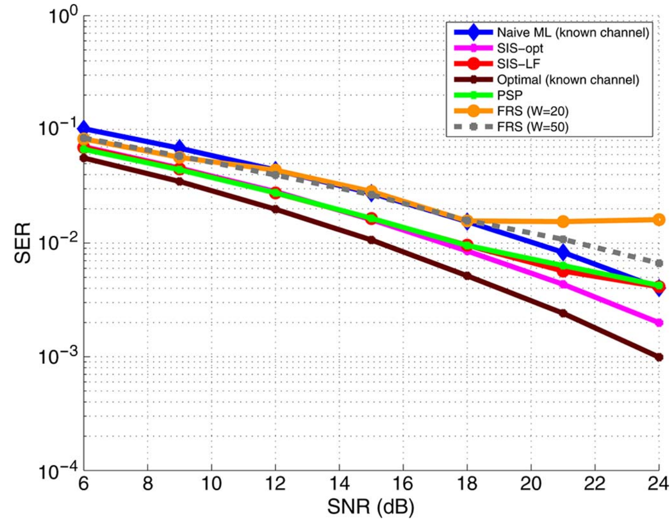

Fig. 5. SER for several values of the SNR (in decibels) when the frame length is $K = 10$. The results are averaged over 20 800 independent data frames, where each one has its own associated channel realization and spreading-code matrix.

the standard deviation of the channel estimate starts increasing when the user stops transmitting and, conversely, how the standard deviation curves tightly squeeze the channel estimate mean after the user has been transmitting for a few time instants. Additionally, when the user is not transmitting, it can be appreciated that the estimate of the corresponding channel coefficient stays approximately constant. This is due to the KF updating the estimate using only the *a priori* information given by the AR process that models the evolution of the channel, i.e., (14), which yields a slow variation.

Apparently, the performance of the FRS-based algorithm depends on both the length of the data frame and the number of points, i.e., W, of the grid that is used to quantize the channel coefficients. In the last experiment, we consider (as in [7]) short data frames of only $K = 10$ symbols per user to avoid the error propagation phenomena, and additionally, we simulate the FRS-based algorithm using a grid of $W = 50$ points (the corresponding curve to be labeled "FRS ($W = 50$)").

Fig. 5 shows the SER attained by the proposed receivers, the optimal sequence detector with a known channel, and the FRS-based receivers (with $W = 20$ and $W = 50$) when the frame length is $K = 10$. We first observe that the SER of the optimal receiver with a known channel increases (when compared with that in Fig. 1), as shorter sequences are harder to detect on average. As a consequence, the SER curves corresponding to all the other receivers described herein are also shifted upward, and the performance gaps between different algorithms are reduced. In particular, the SER of the FRS-based receiver appears closer to that of the PSP and SMC detectors. We also observe that using a finer grid for the discretization of channel coefficients ($W = 50$ instead of $W = 20$) leads to a clear performance improvement in the SER of the FRS-based approach, particularly noticeable for the higher SNRs. However, the computational burden of the FRS ($W = 50$) algorithm is prohibitive in practice since $(50 \times |\{-1, +1, 0\}|)^3 = 150^3 = 3\,375\,000$ different combinations of transmitted symbols and (discrete) channel coefficients need to be explored (for $N = 3$) in the worst-case scenario (if a systematic tree search procedure is applied, see the Appendix for more details), and furthermore, its performance is still inferior compared with that of the receivers introduced in this paper.

VII. Summary and Discussion

In this paper, we have tackled the problem of user identification in time-varying multiuser DS-CDMA systems. We have proposed a detailed model for the observed signals that explicitly takes into account the users' activity dynamics and the effect of a flat-fading and time-selective channel. At the receiver, the number and identity of the active users, their transmitted data, and their corresponding time-varying channel coefficients have to be jointly estimated at every time instant. For this task, we have proposed three different algorithms, which are based on the PSP and SMC methodologies. The optimal (MAP) sequence detector with a perfectly known channel has been also derived for comparison. A remarkable fact about the proposed receivers is that they all arise as natural extensions of simpler existing equalizers for multiple-input–multiple-output channels [35] in which the number of inputs (equivalent to the number of users in a DS-CDMA setup) is a fixed parameter.

The proposed receivers have been numerically assessed, by means of computer simulations, in terms of the SER (for data detection), the ADER (for user activity detection), and the MSE (for channel estimation). Both the PSP- and SMC-based equalizers attain a performance that is close to the genie-aided MAP sequence detector in terms of both the SER and the ADER, particularly in the low- and medium-SNR regions. Moreover, the computer simulations show that most symbol detection errors are due to the failure in the detection of the users' activity (both for the optimal and the proposed receivers). Therefore, keeping the ADER low is strictly necessary to attain a low SER.

We have also assessed the impact that ignoring the statistical model of the users' activity dynamics has on performance. To be specific, we have simulated an ML receiver that has perfect knowledge of the CIR (at every time step) but does not take into account the users' activity probabilities. All the proposed algorithms, including the PSP and SMC equalizers that have to carry out channel estimation, clearly outperform this *naive* ML receiver in the whole range of SNRs. Therefore, the appropriate modeling of the users' activity dynamics appears fundamental for the reliability of the receiver.

The aim and scope of this paper is similar to recent works [6], [7], where RST has been used for the representation of user activity patterns and the design of Bayesian detection algorithms. The system model proposed in this paper is built upon classical probability theory; hence, we expect that it appears more intuitive and easier to understand for many engineers working in the field. In addition, the proposed algorithms avoid certain approximations required by the RST-based receivers proposed in [7], including the feedback of hard decisions (with the subsequent risk of error propagation) and the discretization of channel coefficients. A detailed description of the methodology in [7] and a comparison with the techniques herein introduced are provided in Section VI and the Appendix.

Appendix
User Activity Detection Using Finite Random Sets

We briefly review the approach to multiuser activity detection based on RST. This methodology was introduced in [36] and later developed in [6], [7], and [37] by the same group of authors. We follow [7], which tackles the same problem as this paper and proposes practical algorithms for joint user activity detection, channel estimation, and multiuser data detection.

When comparing with the methodology of this paper, we will refer to the model in Section II as "elementary" because it is built using only notions and objects from the classical probability theory.

Model: We adapt the notation in [7] to be consistent with the rest of the paper. First, we define the (random) set of active users $\mathcal{S}_t = \cup_{i=1}^{N} \mathcal{S}_t^i$, where

$$\mathcal{S}_t^i = \begin{cases} \left\{ [i, h_t(i), s_t(i)]^\top \right\}, & \text{if user } i \text{ is active at time } t \\ \emptyset, & \text{otherwise.} \end{cases}$$

Assuming, for the sake of clarity, that the channel is real, each element in set $\mathcal{S}_t$ is a 3×1 vector in space $\mathbb{S} = \{1, \ldots, N\} \times \mathbb{R} \times \mathcal{A}$, containing the index, i.e., i, of a currently active user, its associated channel coefficient $h_t(i)$, and transmitted symbol $s_t(i)$. We denote the projection of $\mathcal{S}_t$ over the set of indexes $\{1, \ldots, N\}$ as $\pi(S_t)$ (hence, $i \in \pi(\mathcal{S}_t)$ if and only if user i is active).

Both in [7] and in the elementary model in Section II, the following assumptions are made.

- The observation noise is assumed to be a white Gaussian sequence.
- The symbols transmitted by the users are assumed mutually independent.
- The channel is assumed to be frequency flat and time varying.
- The maximum number of users that the system can hold, which is denoted by N, is fixed.

Additionally, in [7] and in this paper, we assume that $N \leq L$, but in both cases, the methodologies can be extended to overloaded scenarios.

The sequence $\{\mathcal{S}_t\}$ of FRSs can be shown to be a discrete-time Markov process with transition distribution [7]

$$\begin{aligned} p(\mathcal{S}_t|\mathcal{S}_{t-1}) = &|\mathcal{A}|^{-|\mathcal{S}_t|}\mu^{|\pi(\mathcal{S}_t)\cap\pi(\mathcal{S}_{t-1})|} \\ &\times (1-\mu)^{|\pi(\mathcal{S}_{t-1})|-\pi(\mathcal{S}_t)\cap\pi(\mathcal{S}_{t-1})} \\ &\times \eta^{|\pi(\mathcal{S}_t)\setminus\pi(\mathcal{S}_t)\cap\pi(\mathcal{S}_{t-1})|} \\ &\times (1-\eta)^{N-|\pi(\mathcal{S}_{t-1})|-|\pi(\mathcal{S}_t)\setminus\pi(\mathcal{S}_t)\cup\pi(\mathcal{S}_{t-1})|} \\ &\times \prod_{i\in\pi(\mathcal{S}_t)\cap\pi(\mathcal{S}_{t-1})} p(h_t(i)|h_{t-1}(i)) \\ &\times \prod_{i\in\pi(\mathcal{S}_t)\setminus\pi(\mathcal{S}_t)\cap\pi(\mathcal{S}_{t-1})} p\left(h_t(i)\right) \end{aligned} \qquad (73)$$

and *a priori* distribution (at time $t = 0$)

$$p(\mathcal{S}_0) = |\mathcal{A}|^{-|\mathcal{S}_0|}\phi^{|\mathcal{S}_0|}(1-\phi)^{N-|\mathcal{S}_0|} \prod_{i\in\pi(\mathcal{S}_0)} p\left(h_0(i)\right).$$

In (73), note the need for the marginal pdfs $p(h_t(i))$ to describe the dynamics of the FRS sequence. This is a difficulty (because most models are based on transition densities $p(h_t(i)|h_{t-1}(i))$, with an *a priori* density for the initial time instant) that does not exist in the proposed elementary model.

The assumptions on the RST-based model in [7] are exactly the same as those on the elementary model in Section II, and both models can represent the same kind of system.

Methodology: Using the given formulation, the MAP detector for the FRS $\mathcal{S}_t$ can be written as

$$\mathcal{S}_t^{\mathrm{MAP}} = \arg \max_{\mathcal{S}_t \in \{\mathbb{S} \cup \{\emptyset\}\}^N} p(\mathcal{S}_t | \mathbf{y}_{0:t}). \quad (74)$$

The (approximate) solution for this problem proposed in [7] is based on the Bayesian decomposition, i.e.,

$$p(\mathcal{S}_t|\mathbf{y}_{0:t}) \propto p(\mathcal{S}_t|\mathbf{y}_{0:t-1}) p(\mathbf{y}_t|\mathcal{S}_t)$$

where $p(\mathbf{y}_t|\mathcal{S}_t)$ is the likelihood of the candidate FRS $\mathcal{S}_t$, and

$$p(\mathcal{S}_t|\mathbf{y}_{0:t-1}) = \int p(\mathcal{S}_t|\mathcal{S}_{t-1}) p(\mathcal{S}_{t-1}|\mathbf{y}_{0:t-1}) \delta\mathcal{S}_{t-1} \quad (75)$$

where $\delta\mathcal{S}_{t-1}$ denotes the set integral operator in [38].

The methodology in [7] requires two types of approximations to handle the predictive distribution of (75). The first approximation consists in the discretization of the channel coefficients (see [7, Sec. 4]). To be specific, each coefficient $h_t(i) \in \mathbb{R}$, $i = 1, \ldots, N$, is assumed to take values on a grid with W points, which are denoted by $\mathcal{W} = \{a_1, \ldots, a_W\} \subset \mathbb{R}$. The discretized channel coefficients are denoted by $\check{h}_t(i) \in \mathcal{W}$, and the FRSs are redefined as $\mathcal{S}_t = \cup_{i=1}^N \mathcal{S}_{t-1}^i$, where

$$\check{\mathcal{S}}_t^i = \begin{cases} \left\{ \left[i, \check{h}_t(i), s_t(i)\right]^\top \right\}, & \text{if user } i \text{ is active at time } t \\ \emptyset, & \text{otherwise.} \end{cases}$$

Transition distribution $p(\check{\mathcal{S}}_t|\check{\mathcal{S}}_{t-1})$ is the same as in (73) but substituting $p(h_t(i)|h_{t-1}(i))$ and $p(h_t(i))$ by $p(\check{h}_t(i)|\check{h}_{t-1}(i))$ and $p(\check{h}_t(i))$, respectively. Note that both $p(\check{h}_t(i)|\check{h}_{t-1}(i))$ and $p(\check{h}_t(i))$ have to be computed from the original (continuous) channel model by integration, which is not necessarily a straightforward task.

After channel discretization, the predictive distribution becomes

$$p(\check{\mathcal{S}}_t|\mathbf{y}_{0:t-1}) = \sum_{\check{\mathcal{S}}_{t-1} \in (\emptyset \cup \check{\mathbb{S}})^N} p(\check{\mathcal{S}}_t|\check{\mathcal{S}}_{t-1}) p(\check{\mathcal{S}}_{t-1}|\mathbf{y}_{0:t-1}) \quad (76)$$

where $\check{\mathbb{S}} = \{1, \ldots, N\} \times \mathcal{W} \times \mathcal{A}$, which is still too complex because of the large number of terms in the summation. To alleviate this difficulty, Angelosante *et al.* [7] suggest a "zero-order approximation" (ZOA) that consists in substituting the sum on the right-hand side of (76) by a single term. Specifically, if we let $\tilde{\check{\mathcal{S}}}_{t-1}$ denote the FRS detected at time $t-1$, then the predictive distribution is approximated as

$$p(\check{\mathcal{S}}_t|\mathbf{y}_{0:t-1}) \approx p(\check{\mathcal{S}}_t|\tilde{\check{\mathcal{S}}}_{t-1}) p(\tilde{\check{\mathcal{S}}}_{t-1}|\mathbf{y}_{0:t-1})$$

and the original problem of MAP detection in (74) is replaced by (the approximation)

$$\tilde{\check{\mathcal{S}}}_t = \arg \max_{\check{\mathcal{S}}_t \in (\emptyset \cup \check{\mathbb{S}})^N} p(\check{\mathcal{S}}_t|\tilde{\check{\mathcal{S}}}_{t-1}) p(\tilde{\check{\mathcal{S}}}_{t-1}|\mathbf{y}_{0:t-1}) p(\mathbf{y}_t|\check{\mathcal{S}}_t). \quad (77)$$

Problem (77) can be solved using tree search procedures, as discussed in [7, App. A].

Discussion: The algorithms for detection of FRS proposed in [7] rely on the discretization of the channel coefficients and the so-called ZOA.

In addition to introducing a quantization error in the channel estimates, channel discretization can be understood as expanding the symbol alphabet, i.e., for each user, instead of detecting symbol $s_t(i)$ from set $\mathcal{A}$, it is necessary to detect a product symbol $h_t(i)s_t(i)$ from set $\mathcal{W} \times \mathcal{A}$. This is relevant because the size of the tree (over which solutions for problem (77) are searched) becomes $|\mathcal{W} \times \mathcal{A}|^N$. For example, with BPSK symbols and a known channel, the search tree has 2^N leaves. If, e.g., $W = |\mathcal{W}| = 10$, the search tree has 20^N leaves. While tree search methods (referred to as *sphere detection* in [7]) can be efficient, which means that they do not necessarily have to explore the complete tree to find the optimal path from the root to the leaves, their complexity is still given by the number of leaves. There is no optimal tree search method that can guarantee the optimal solution of (77) with nonexponential complexity for any observation vector $\mathbf{y}_t$. In particular, at low or even moderate SNR, it can be expected that a large proportion of the tree nodes have to be explored, while the search should be considerably faster at a high SNR. (See [39] for a review of tree search techniques applied to multiuser detection.)

The methods proposed in this paper, which are both based on the PSP principle and the SMC methodology, avoid discretization of the channel by using Kalman filters to estimate diagonal matrix $\mathbf{H}_t$ recursively. Therefore, they do not suffer from quantization error. Note, from the previous discussion, that in the FRS approach, reducing the quantization error (by increasing W) increases the complexity of the detector drastically.

The ZOA can be easily interpreted as a feedback scheme. It consists in "fixing" the set detected at time $t-1$ to ease detection at time t. As acknowledged in [7, Sec. 3.2], this kind of approximation can be expected to be adequate only for a sufficiently large SNR. At lower SNR ranges, it may lead to error propagation phenomena as any other decision feedback scheme. Within the framework of this paper, the ZOA in [7] is analogous to reducing the number of survivors in the PSP algorithm to one (for the whole set of possible states).

References

[1] S. Moshavi, "Multi-user detection for DS-CDMA communications," *IEEE Commun. Mag.*, vol. 34, no. 10, pp. 124–136, Oct. 1996.

[2] S. Verdú, *Multiuser Detection*. Cambridge, U.K.: Cambdridge Univ. Press, 1998.

[3] K. Halford and M. Brandt-Pearce, "New-user identification in a CDMA system," *IEEE Trans. Commun.*, vol. 46, no. 1, pp. 144–155, Jan. 1998.

[4] W.-C. Wu and K.-C. Chen, "Identification of active users in synchronous CDMA multiuser detection," *IEEE J. Sel. Areas Commun.*, vol. 16, no. 9, pp. 1723–1735, Dec 1998.

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

[6] D. Angelosante, E. Biglieri, and M. Lops, "Multiuser detection in a dynamic environment—Part II: Joint user identification and parameter estimation," *IEEE Trans. Inf. Theory*, vol. 55, no. 5, pp. 2365–2374, May 2009.

[7] D. Angelosante, E. Biglieri, and M. Lops, "Low-complexity receivers for multiuser detection with an unknown number of active users," *Signal Process.*, vol. 90, no. 5, pp. 1486–1495, May 2010.

[8] P. Pad, A. Mousavi, A. Goli, and F. Marvasti, "Simplified MAP-MUD for active user CDMA," *IEEE Commun. Lett.*, vol. 15, no. 6, pp. 599–601, Jun. 2011.

[9] A. Mousavi, P. Pad, P. Delgosha, and F. Marvasti, "A new decoding scheme for errorless codes for overloaded CDMA with active user detection," in *Proc. 18th IEEE ICT*, 2011, pp. 201–205.

[10] D. Angelosante, E. Grossi, G. Giannakis, and M. Lops, "Sparsity-aware estimation of CDMA system parameters," in *Proc. IEEE 10th Workshop IEEE SPAWC*, 2009, pp. 697–701.

[11] H. Zhu and G. Giannakis, "Exploiting sparse user activity in multiuser detection," *IEEE Trans. Commun.*, vol. 59, no. 2, pp. 454–465, Feb. 2011.

[12] S. Buzzi, L. Venturino, A. Zappone, and A. De Maio, "Blind user detection in doubly dispersive DS/CDMA fading channels," *IEEE Trans. Signal Process.*, vol. 58, no. 3, pp. 1446–1451, Mar. 2010.

[13] J. G. Proakis, *Digital Communications*, 4th ed. Singapore: McGraw-Hill, 2000.

[14] S. G. Glisic and P. A. Leppanen, *Wireless Communications: TDMA Versus CDMA*. Norwell, MA, USA: Kluwer, 1997.

[15] P. M. Djurić, J. H. Kotecha, J. Zhang, Y. Huang, T. Ghirmai, M. F. Bugallo, and J. Míguez, "Particle filtering," *IEEE Signal Process. Mag.*, vol. 20, no. 5, pp. 19–38, Sep. 2003.

[16] R. E. Kalman, "A new approach to linear filtering and prediction problems," *J. Basic Eng.*, vol. 82, pp. 35–45, 1960.

[17] B. D. O. Anderson and J. B. Moore, *Optimal Filtering*. Englewood Cliffs, NJ, USA: Prentice-Hall, 1979.

[18] K. Pandremmenou, L. Kondi, and K. Parsopoulos, "Optimal power allocation and joint source–channel coding for wireless DS-CDMA visual sensor networks using the Nash bargaining solution," in *Proc. IEEE ICASSP*, 2011, pp. 2340–2343.

[19] A. Katsenou, L. Kondi, and K. Parsopoulos, "Resource management for wireless visual sensor networks based on individual video characteristics," in *Proc. 18th IEEE ICIP*, 2011, pp. 149–152.

[20] E. Sim and L. Yang, "DS-CDMA with M-ary orthogonal modulation for wireless sensor networks simultaneously monitoring multiple events," in *Proc. IEEE 75th IEEE VTC Spring*, 2012, pp. 1–5.

[21] D. Torrieri and M. Valenti, "Guard zones and the near–far problem in DS-CDMA ad hoc networks," in *Proc. MILCOM*, 2012, pp. 1–6.

[22] C. Komnikakis, C. Fragouli, A. H. Sayeed, and R. D. Wesel, "Multi-input multi-output fading channel tracking and equalization using Kalman estimation," *IEEE Trans. Signal Process.*, vol. 50, no. 5, pp. 1065–1076, May 2002.

[23] J. S. Liu and R. Chen, "Sequential Monte Carlo methods for dynamic systems," *J. Amer. Stat. Assoc.*, vol. 93, no. 443, pp. 1032–1044, Sep. 1998.

[24] A. Doucet, S. Godsill, and C. Andrieu, "On sequential Monte Carlo sampling methods for Bayesian filtering," *Stat. Comput.*, vol. 10, no. 3, pp. 197–208, Jul. 2000.

[25] A. Doucet, N. de Freitas, and N. Gordon, *Sequential Monte Carlo Methods in Practice*. New York, NY, USA: Springer-Verlag, 2001.

[26] D. Crisan and A. Doucet, "A survey of convergence results on particle filtering," *IEEE Trans. Signal Process.*, vol. 50, no. 3, pp. 736–746, Mar. 2002.

[27] R. Douc, O. Cappe, and E. Moulines, "Comparison of resampling schemes for particle filtering," in *Proc. ISPA*, Sep. 2005, pp. 64–69.

[28] J. Zhang and P. M. Djurić, "Joint estimation and decoding of space–time trellis codes," *J. Appl. Signal Process.*, vol. 2002, no. 1, pp. 305–315, Jan. 2002.

[29] M. A. Vázquez and J. Míguez, "Maximum likelihood sequence detection in time and frequency selective MIMO channels with unknown order," *IEEE Trans. Veh. Technol.*, vol. 58, no. 1, pp. 499–504, Jan. 2009.

[30] M. K. Pitt and N. Shephard, "Auxiliary variable based particle filters," in *Sequential Monte Carlo Methods in Practice*, A. Doucet, N. de Freitas, and N. Gordon, Eds. New York, NY, USA: Springer-Verlag, 2001, ch. 13, pp. 273–293.

[31] J. Zhang, Y. Huang, and P. M. Djurić, "Multiuser detection with particle filtering," in *Proc. 11th EUSIPCO*, Sep. 2002, pp. 307–310.

[32] A. Sayed, *Fundamentals of Adaptive Filtering*. Hoboken, NJ, USA: Wiley-IEEE Press, 2003.

[33] T. Aulin, "A modified model for the fading signal at a mobile radio channel," *IEEE Trans. Veh. Technol.*, vol. VT-28, no. 3, pp. 182–203, Aug. 1979.

[34] J. D. Parsons, *The Mobile Radio Propagation Channel*. London, U.K.: Pentech, 1992.

[35] M. A. Vázquez, M. F. Bugallo, and J. Míguez, "Sequential Monte Carlo methods for complexity-constrained MAP equalization of dispersive MIMO channels," *Signal Process.*, vol. 88, no. 4, pp. 1017–1034, Apr. 2008.

[36] E. Biglieri and M. Lops, "Multiuser detection in a dynamic environment," *IEEE Trans. Inf. Theory*, vol. 53, no. 9, pp. 3158–3170, Sep. 2007.

[37] D. Angelosante and E. Biglieri, "Low-complexity receivers for multiuser detection with an unknown number of active users," in *Proc. IEEE ICASSP*, Mar. 31–Apr. 4, 2008, pp. 3481–3484.

[38] R. Mahler, *Statistical Multisource–Multitarget Information Fusion*. Boston, MA, USA: Artech House, 2007.

[39] G. Wei, L. Rasmussen, and R. Wyrwas, "Near optimum tree-search detection schemes for bit-synchronous multiuser CDMA Systems over Gaussian and two-path Rayleigh-fading channels," *IEEE Trans. Commun.*, vol. 45, no. 6, pp. 691–700, Jun. 1997.

Manuel A. Vázquez received the M.Sc. and the Ph.D. degrees in computer engineering from the University of A Coruña, A Coruña, Spain, in 2003 and 2008, respectively.

From April 2008 to March 2009, he was a Researcher with Tampere University of Technology, Tampere, Finland. From February to August 2011, he was a Visiting Researcher with Nanyang Technological University, Singapore. He is currently a Ph.D. Assistant Professor with the Departamento de Teoría de la Señal y Comunicaciones, Universidad Carlos III de Madrid, Madrid, Spain. His research interests include particle filtering, multiple-input–multiple-output systems, dynamic programming, positioning, and electroencephalogram signal processing.

Joaquín Míguez (S'98–A'00–M'03) received the M.Sc. and Ph.D. degrees in computer engineering from the Universidade da Coruña, Spain, in 1997 and 2000, respectively.

Late in 2000, he joined the Department of Electronics and Systems, Universidade da Coruña, where he became an Associate Professor in July 2003. Since October 2004, he has held the same position within the Department of Signal Theory and Communications of the Universidad Carlos III de Madrid, Madrid, Spain. He spent most of 2001 as a visiting researcher with the Department of Electrical Engineering, Stony Brook University, Stony Brook, NY, USA, and, at the time of this publication, he is an academic visitor with the Department of Mathematics, Imperial College London, London, U.K. He has also delivered invited lectures on his research in France, Germany, Spain, Finland, Brazil, and the UK. His research interests are in the fields of statistical signal processing, Bayesian analysis, dynamical systems, and the theory and applications of Monte Carlo methods. He has co-authored more than 35 international journal papers in the areas of electrical engineering, mathematical physics, and statistics.

Dr. Míguez co-received the IEEE Signal Processing Magazine Best Paper Award 2007.