A Sequential Monte Carlo Method for Adaptive Blind Timing Estimation and Data Detection

Tadesse Ghirmai, Mónica F. Bugallo, Joaquín Míguez, and Petar M. Djurić, *Senior Member, IEEE*

***Abstract*—Accurate estimation of synchronization parameters is critical for reliable data detection in digital transmission. Although several techniques have been proposed in the literature for estimation of the reference parameters, i.e., timing, carrier phase, and carrier frequency offsets, they are based on either heuristic arguments or approximations, since optimal estimation is analytically intractable in most practical setups. In this paper, we introduce a new alternative approach for blind synchronization and data detection derived within the Bayesian framework and implemented via the sequential Monte Carlo (SMC) methodology. By considering an extended dynamic system where the reference parameters and the transmitted symbols are system-state variables, the proposed SMC technique guarantees asymptotically minimal symbol error rate when it is combined with adequate receiver architectures, both in open-loop and closed-loop configurations. The performance of the proposed technique is studied analytically, by deriving the posterior Crámer–Rao bound for timing estimation and through computer simulations that illustrate the overall performance of the resulting receivers.**

***Index Terms*—Adaptive receivers, blind data detection, particle filtering, synchronization, timing recovery.**

I. Introduction

THERE are many scenarios where the wireless radio channel can be described by a set of well-defined physical parameters: the relative delay between the received signal and the local clock reference, the amplitude attenuation, and the carrier frequency and phase offsets. These parameters need to be estimated, and compensated for, prior to detection. The generalized synchronization problem [1], [2] consists of the recovery of the aforementioned reference parameters using the signal observed at the receiver front end, and it can be seen as a special case of channel equalization [2]. Synchronization algorithms obviously play a vital role in attaining reliable digital transmission.

Manuscript received October 29, 2003; revised September 3, 2004. This work was supported by the National Science Foundation (award CCR-0082607) and the Ministerio de Educación y Ciencia of Spain (project TEC2004-06451-C05-01). Part of the results here included were presented at the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2004). The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Xiaodong Wang.

T. Ghirmai is with the Department of Computer Engineering, Jackson State University, Jackson, MS 39217 USA (e-mail:tadesse.ghirmai@jsums.edu).

M. F. Bugallo and P. M. Djurić are with the Department of Electrical and Computer Engineering, State University of New York at Stony Brook, Stony Brook, NY 11794 USA (e-mail: monica@ece.sunysb.edu; djuric@ece.sunysb.edu).

J. Míguez is with the Departamento de Teoría de la Señal y Comunicaciones, Universidad Carlos III de Madrid, 28911 Leganés, Madrid, Spain (e-mail: jmiguez@ieee.org).

Digital Object Identifier 10.1109/TSP.2005.850372

Unfortunately, optimal estimators for synchronization parameters are impossible to obtain, in general [1], and, therefore, most of the existing techniques are based on approximate maximum-likelihood (ML) arguments and heuristic methods [1], [2]. Two classes of synchronization algorithms can be identified: decision directed (or data aided) methods, which require the transmission of pilot data, and nondecision directed (or nondata aided) techniques [1]. Unlike data-aided techniques, nondata-aided methods do not require pilot symbols and, instead, they exploit statistics of digital waveforms. Nondata-aided schemes are usually termed *blind* techniques, according to equalization terminology [3].

In recent years, much attention has been devoted to a group of techniques known as sequential Monte Carlo (SMC) algorithms (also referred to as *particle filtering* methods) [4]. All of these techniques are aimed at building a recursive Bayesian filter, which estimates the posterior probability density function (pdf) based on Monte Carlo simulations. Particle filters are an important alternative for predicting and estimating unknown parameters of interest in real-time applications, especially in systems with nonlinearities and non-Gaussianities where classical approaches based on the well-known Kalman filter [5], [6] provide solutions that may be far from optimal. Specifically, a main stream of research in the application of particle filtering to communications is currently under way [7].

In this paper, we propose the application of the SMC methodology to the generalized synchronization problem. In particular, we investigate SMC techniques for blind adaptive estimation of the synchronization parameters and the transmitted data sequence. We regard the application of particle filtering in this context very appealing, as it leads to improved numerical solutions for a problem where existing analytical approaches are suboptimal.

The new SMC blind receivers are derived by representing digital transmission through the wireless fading channel with an extended dynamic system. Thus, the symbol delay is modeled as a first-order autoregressive (AR) stochastic process [8], transmitted data are assumed independent and identically distributed (i.i.d.) random variables with a discrete uniform distribution, and the complex fading channel coefficients are samples from a second-order AR process driven by complex white Gaussian noise [9]. The latter is a broadly accepted way of modeling signal amplitude attenuation and carrier phase offset jointly.

The choice of a receiver architecture considerably constrains the attainable minimal symbol error rate (SER). Here, we suggest *open-* and *closed-loop* receiver structures which allow for complete removal of intersymbol interference (ISI) and achieve close-to-optimal SER. The performance of the resulting SMC-based receivers is assessed both analytically, by deriving the posterior Cramér–Rao bound (PCRB) for timing estimation, and through computer simulations.

The paper is organized as follows. Section II describes the signal model. The proposed SMC algorithm for joint synchronization and detection is presented in Section III. In Section IV, we describe the receiver architectures. Performance is studied analytically in Section V and numerically in Section VI. Finally, conclusions are presented in Section VII.

II. Signal Model

A. Received Signal

Let us consider a digital communication system where symbols from a discrete alphabet, $s_m \in \mathcal{S}$, are transmitted in frames of length M. The received signal at the front-end of the receiver, that consists of a matched filter, has the form

$$y(t) = h(t)e^{j\omega(t)t} \sum_{m=0}^{M-1} s_m g\left(t - mT + \tau(t)\right) + v(t) \tag{1}$$

where $h(t)$ is the complex multiplicative noise introduced by the frequency-flat Rayleigh fading channel, $\omega(t)$ is the carrier frequency error, s_m is the mth transmitted symbol, $g(t)$ is a raised-cosine pulse waveform, T is the symbol period, $0 \leq \tau(t) < T$ is the time-varying relative delay between the received signal and the local clock reference, and $v(t)$ is additive white Gaussian noise (AWGN) with power spectral density $N_0/2$. Sampling the matched filter output with rate $1/T$ and assuming that that the raised cosine waveform $g(t)$ has finite duration (which is always the case in practice), we can write the resulting discrete-time signal as

$$y_k = h_k e^{j\omega_k kT} \sum_{n=1-L}^{L} s_{k+n} g(-nT + \tau_k) + v_k \tag{2}$$

where $y_k = y(kT)$, $h_k = h(kT)$, $\omega_k = \omega(kT)$, $\tau_k = \tau(kT)$, $v_k = v(kT)$, $k = 0, 1, 2, \ldots$ denotes discrete-time, and L indicates the ISI span resulting from the limited time duration of $g(t)$. Note that, after symbol rate sampling, the noise term v_k remains white with variance σ_v^2. Using vector notation, we arrive at the convenient representation

$$y_k = h_k e^{j\omega_k kT} \mathbf{g}(\tau_k)^\top \mathbf{s}_k + v_k \tag{3}$$

by defining the $2L \times 1$ channel vector

$$\mathbf{g}(\tau_k) = [g\left((L-1)T + \tau_k\right), g\left((L-2)T + \tau_k\right), \ldots, g(-LT + \tau_k)]^\top$$

the symbol vector $\mathbf{s}_k = [s_{k-L+1}, s_{k-L+2}, \ldots, s_{k+L}]^\top$ and using the symbol $^\top$ to denote transposition.

The general objective is to jointly estimate the transmitted symbols $s_{0:M-1} = \{s_0, s_1, \ldots, s_{M-1}\}$, the signal timing $\tau_{0:M-1} = \{\tau_0, \tau_1, \ldots, \tau_{M-1}\}$, the complex fading coefficients $h_{0:M-1} = \{h_0, h_1, \ldots, h_{M-1}\}$, and the frequency offsets $\omega_{0:M-1} = \{\omega_0, \omega_1, \ldots, \omega_{M-1}\}$, using the set of received signals $y_{0:M-1} = \{y_0, \ldots, y_{M-1}\}$. For clarity of presentation, however, we will only consider here the detection of $s_{0:M-1}$ and the estimation of $h_{0:M-1}$ and $\tau_{0:M-1}$, while assuming $\omega_k = 0\ \forall k$. The proposed technique, however, can be easily extended to the general case where ω_k is also a time-varying magnitude to be estimated.[1]

B. State-Space Representation

Before applying SMC techniques, we model the signal of interest as a dynamic system in state space form. Following [8], we can model the symbol timing as a first order autoregressive (AR) process

$$\tau_k = \alpha\tau_{k-1} + u_k \tag{4}$$

where the perturbation variable u_k is assumed to be a zero-mean Gaussian with variance σ_u^2. It should be remarked that (4) is used to model time selectivity in the wireless link, and, therefore, it is not due to any timing corrections performed at the receiver. The process parameters α and σ_u^2 should be chosen to account for the physical channel rate of variation.

Similarly, the variation of the fading coefficient h_k can be modeled by a second order autoregression driven by a complex white Gaussian process [9]

$$h_k = \beta_1 h_{k-1} + \beta_2 h_{k-2} + e_k \tag{5}$$

where the values of the coefficients of the autoregressive process β_1 and β_2 and the variance of the zero-mean complex white Gaussian noise e_k are functions of the fading rate of the channel.

Taking into account the structure of $\mathbf{s}_k$, and combining (3) with $\omega_k = 0$, (4) and (5), we obtain the following dynamic state-space representation of the communication system:

$$\left.\begin{aligned} \tau_k &= \alpha\tau_{k-1} + u_k \\ \mathbf{h}_k &= \mathbf{A}\mathbf{h}_{k-1} + \mathbf{c}e_k \\ \mathbf{s}_k &= \mathbf{S}\mathbf{s}_{k-1} + \mathbf{d}_k \end{aligned}\right\} \quad \text{state equation} \tag{6}$$

$$y_k = \mathbf{c}^\top \mathbf{h}_k \mathbf{s}_k^\top \mathbf{g}(\tau_k) + v_k \quad \text{observation equation} \tag{7}$$

where

$$\mathbf{A} = \begin{bmatrix} \beta_1 & \beta_2 \\ 1 & 0 \end{bmatrix}$$

$\mathbf{h}_k = [h_k, h_{k-1}]^\top$, $\mathbf{c} = [1, 0]^\top$, and

$$\mathbf{S} = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ 0 & 0 & 0 & \cdots & 0 \end{bmatrix}$$

is a $2L \times 2L$ shifting matrix, $\mathbf{s}_k = [s_{k-L+1}, \ldots, s_{k+L}]^\top$ is an $2L \times 1$ vector, and $\mathbf{d}_k = [0, \ldots, 0, s_{k+L}]^\top$ is a $2L \times 1$ perturbation vector that contains the new symbol, s_{k+L}. Note that the system state at time k is given by ($\mathbf{s}_k$, $\mathbf{h}_k$, τ_k), while the model parameters, $\alpha, \sigma_u^2, \sigma_v^2, \beta_1, \beta_2, \sigma_e^2$, and L are assumed fixed and known.

[1] Specifically, the frequency offset can be handled in the same way as the delay τ_k, as described in the subsequent sections.

III. Sequential Monte Carlo Algorithm

A. Particle Filtering

The fading coefficients $\mathbf{h}_k$ in (6) and (7) can be considered nuisance parameters, i.e., they must be handled properly but we do not need to estimate them explicitly. Hence, we subsequently focus on the estimation of the sequence of *reduced* states up to time k, $\mathbf{x}_{0:k} = \{\mathbf{s}_{0:k}, \tau_{0:k}\}$, from the collected observations, $y_{0:k}$. From a Bayesian perspective, all necessary information is contained in the joint posterior probability distribution function,[2] $p(\mathbf{s}_{0:k}, \tau_{0:k}|y_{0:k})$. Since the latter distribution is, unfortunately, analytically intractable and prevents the derivation of closed-form Bayesian estimators, we use a particle filter to represent the posterior pdf by means of a discrete probability measure with random support. Specifically, we denote the measure at time k as

$$\Xi_k = \left\{(\mathbf{s}_{0:k}, \tau_{0:k})^{(n)}, w_k^{(n)}\right\}_{n=1}^{N} \tag{8}$$

where $\mathbf{s}_{0:k}^{(n)}$ and $\tau_{0:k}^{(n)}$ are sample trajectories, also referred as particles, and $w_k^{(n)}$ is the *importance weight*.

Using (8), the posterior pdf is approximated as

$$\hat{p}(\mathbf{s}_{0:k}, \tau_{0:k}|y_{0:k}) = \sum_{n=1}^{N} \delta^{(n)}(\mathbf{s}_{0:k}, \tau_{0:k}) w_k^{(n)} \tag{9}$$

where

$$\delta^{(n)}(\mathbf{s}_{0:k}, \tau_{0:k}) = \begin{cases} 1, & \text{if } \mathbf{s}_{0:k} = \mathbf{s}_{0:k}^{(n)}, \tau_{0:k} = \tau_{0:k}^{(n)} \\ 0, & \text{otherwise} \end{cases}$$

and Bayesian estimators of the symbols, the delay, or both, are straightforward to derive.

B. Sequential Importance Sampling

A major advantage of the particle filtering approach is the possibility to build the discrete measure Ξ_k sequentially, i.e., to compute Ξ_k recursively from Ξ_{k-1} when the kth observation y_k becomes available. One of the most popular particle filtering techniques is the so-called sequential importance sampling (SIS) algorithm [10].

According to the importance sampling (IS) principle [4], an empirical approximation of a desired pdf $p(x)$ can be obtained by drawing particles from an *importance function* or *proposal pdf*, $\pi(x)$, which is strictly positive $\pi(x) > 0$ and has the same domain as $p(x)$.

The resulting particles $x^{(n)} \sim \pi(x)$ are assigned (non-normalized) importance weights of the form

$$\tilde{w}^{(n)} = \frac{p\left(x^{(n)}\right)}{\pi\left(x^{(n)}\right)}. \tag{10}$$

In the synchronization application presented in this paper, the desired distribution is $p(\mathbf{s}_{0:k}, \tau_{0:k}|y_{0:k})$; hence, an importance function of the form $\pi(\mathbf{s}_{0:k}, \tau_{0:k}|y_{0:k})$ is needed.

[2]Since $\tau_k \in [0, T)$ is a continuous random process, the joint sequence $\{\mathbf{s}_{0:k}, \tau_{0:k}\}$ cannot be assigned a probability mass, despite the discrete alphabet of the symbols $\mathcal{S}$.

Assuming a proposal pdf that admits a factorization of the form

$$\pi(\mathbf{s}_{0:k}, \tau_{0:k}|y_{0:k}) = \pi(\mathbf{s}_k, \tau_k|\mathbf{s}_{0:k-1}, \tau_{0:k-1}, y_{0:k}) \times \pi(\mathbf{s}_{0:k-1}, \tau_{0:k-1}|y_{0:k-1}) \tag{11}$$

the sequential application of the IS principle has the following steps.

1) **Importance sampling**. We denote the probability measure at time $k-1$ as $\Xi_{k-1} = \{\mathbf{s}_{0:k-1}^{(n)}, \tau_{0:k-1}^{(n)}, w_{k-1}^{(n)}\}_{n=1}^{N}$. When y_k is observed, the state is propagated one time step according to

$$(\mathbf{s}_k, \tau_k)^{(n)} \sim \pi\left(\mathbf{s}_k, \tau_k|\mathbf{s}_{0:k-1}^{(n)}, \tau_{0:k-1}^{(n)}, y_{0:k}\right) = \pi_k(s_{k+L}, \tau_k). \tag{12}$$

Note that given $\mathbf{s}_{k-1}$ the only random variable in $\mathbf{s}_k$ is s_{k+L}.

2) **Weight update**. Once the new particles have been drawn, the importance weights are updated by

$$\tilde{w}_k^{(n)} = w_{k-1}^{(n)} \frac{p\left(y_k|\mathbf{s}_{0:k}^{(n)}, \tau_{0:k}^{(n)}, y_{0:k-1}\right) p\left(\tau_k^{(n)}|\tau_{k-1}^{(n)}\right)}{\pi_k\left(s_{k+L}^{(n)}, \tau_k^{(n)}\right)}. \tag{13}$$

Last, the importance weights are normalized.

C. Computation

The proposed SIS algorithm requires the numerical evaluation of the likelihood function in the weight update equation $p(y_k|\mathbf{s}_{0:k}, \tau_{0:k}, y_{0:k-1})$. Let $\mathcal{N}(\mathbf{x}; \boldsymbol{\mu}, \boldsymbol{\Sigma})$ denote the multivariate Gaussian distribution with mean $\boldsymbol{\mu}$ and covariance $\boldsymbol{\Sigma}$, and assume a Gaussian prior for the fading process, i.e., $\mathbf{h}_{-1} \sim \mathcal{N}(\mathbf{h}; \boldsymbol{\mu}_{-1}, \boldsymbol{\Sigma}_{-1})$. Then, it is straightforward to show that

$$p(y_k|\mathbf{s}_{0:k}, \tau_{0:k}, y_{0:k-1}) = \int_{\mathbf{h}_k} p(y_k|\mathbf{s}_k, \tau_k, \mathbf{h}_k) \times p(\mathbf{h}_k|\mathbf{s}_{0:k-1}, \tau_{0:k-1}, y_{0:k-1}) d\mathbf{h}_k \tag{14}$$

where

$$p(y_k|\mathbf{s}_k, \tau_k, \mathbf{h}_k) = \mathcal{N}\left(y_k; h_k \mathbf{g}^\top(\tau_k)\mathbf{s}_k, \sigma_v^2\right)$$

and

$$p(\mathbf{h}_k|\mathbf{s}_{0:k-1}, \tau_{0:k-1}, y_{0:k-1}) = \mathcal{N}(\mathbf{h}_k; \boldsymbol{\mu}_{k|k-1}, \boldsymbol{\Sigma}_{k|k-1})$$

is a complex Gaussian pdf. Notice that the predictive channel mean

$$\boldsymbol{\mu}_{k|k-1} = E_{p(\mathbf{h}_k|\mathbf{s}_{0:k-1}, \tau_{0:k-1}, y_{0:k-1})}[\mathbf{h}_k] \tag{15}$$

and the predictive covariance matrix

$$\boldsymbol{\Sigma}_{k|k-1} = E_{p(\mathbf{h}_k|\mathbf{s}_{0:k-1}, \tau_{0:k-1}, y_{0:k-1})} \times \left[(\mathbf{h}_k - \boldsymbol{\mu}_{k|k-1})(\mathbf{h}_k - \boldsymbol{\mu}_{k|k-1})^H\right] \tag{16}$$

can be obtained in closed-form using a Kalman filter [5], [6]. Therefore, the integral in (14) can be solved to yield

$$p(y_k|\mathbf{s}_{0:k}, \tau_{0:k}, y_{0:k-1}) = \mathcal{N}\left(y_k; \overline{y}_k, \sigma_{y_k}^2\right) \tag{17}$$

where

$$\overline{y}_k = \mathbf{g}(\tau_k)^\top \mathbf{s}_k \mathbf{c}^\top \boldsymbol{\mu}_{k|k-1}$$

and

$$\sigma_{y_k}^2 = \sigma_v^2 + \left(\mathbf{g}^\top(\tau_k)\mathbf{s}_k\mathbf{c}^\top\right) \boldsymbol{\Sigma}_{k|k-1} \left(\mathbf{g}^\top(\tau_k)\mathbf{s}_k\mathbf{c}^\top\right)^\top.$$

As for the importance function, a tradeoff between sampling efficiency and practical feasibility must be achieved. The optimal importance function, which minimizes the variance of the weights [10], is the posterior pdf

$$\pi\left(\mathbf{s}_k, \tau_k | \mathbf{s}_{0:k-1}^{(n)}, \tau_{0:k-1}^{(n)}, y_{0:k}\right) = p\left(\mathbf{s}_k, \tau_k | \mathbf{s}_{k-1}^{(n)}, \tau_{k-1}^{(n)}, y_k\right)$$

but it cannot be sampled directly due to the nonlinearity of the observations with respect to the delay, τ_k. On the other hand, the simplest proposal density is the prior pdf

$$\pi\left(\mathbf{s}_k, \tau_k | \mathbf{s}_{0:k-1}^{(n)}, \tau_{0:k-1}^{(n)}, y_{0:k}\right) = p(s_{k+L}) p\left(\tau_k | \tau_{k-1}^{(n)}\right)$$

which turns out to be very inefficient [10], meaning that a large number of particles is necessary in order to attain a good approximation of $p(\mathbf{s}_{0:k}, \tau_{0:k}|y_{0:k})$. Therefore, we propose a compromise importance function where the delay is sampled from the prior while the new datum s_{k+L} is drawn using likelihood information. Specifically, at time k, we will use the function

$$\begin{aligned}\left(s_{k+L}^{(n)}, \tau_k^{(n)}\right) &\sim \pi_k(s_{k+L}, \tau_k) \\ &= p\left(s_{k+L}|\mathbf{s}_{0:k-1}^{(n)}, \tau_k, \tau_{0:k-1}^{(n)}, y_{0:k}\right) p\left(\tau_k | \tau_{k-1}^{(n)}\right) \qquad (18)\end{aligned}$$

which can be sampled in two steps. First, we obtain a new delay particle

$$\tau_k^{(n)} \sim \mathcal{N}\left(\tau; \alpha\tau_{k-1}^{(n)}, \sigma_u^2\right) \qquad (19)$$

and then we draw a sample of the transmitted symbol from the first distribution on the right-hand side of (18). This is feasible because we can rewrite $p(s_{k+L}|\mathbf{s}_{0:k-1}^{(n)}, \tau_{0:k}^{(n)}, y_{0:k})$ as

$$\begin{aligned}p&\left(s_{k+L} = S|\mathbf{s}_{0:k-1}^{(n)}, \tau_{0:k}^{(n)}, y_{0:k}\right) \\ &\propto p\left(y_k | s_{k+L} = S, \mathbf{s}_{0:k-1}^{(n)}, \tau_{0:k}^{(n)}, y_{0:k-1}\right) \qquad (20)\end{aligned}$$

where $S \in \mathcal{S}$ is a symbol in the modulation alphabet. Notice that the likelihood on the right-hand side of (20) can be evaluated using (17). The resulting importance weights for the new particles are given by

$$w_k^{(n)} \propto w_{k-1}^{(n)} \sum_{S_j \in \mathcal{S}} \mathcal{N}\left(y_k; \overline{y}_k^{(n,j)}, \sigma_{y_k}^{2\ (n,j)}\right) \qquad (21)$$

where

$$\begin{aligned}\overline{y}_k^{(n,j)} &= \mathbf{g}^\top\left(\tau_k^{(n)}\right)\mathbf{s}_{k,j}^{(n)}\mathbf{c}^\top\boldsymbol{\mu}_{k|k-1}^{(n)} \qquad (22)\\ \sigma_{y_k}^{2\ (n,j)} &= \sigma_v^2 + \left(\mathbf{g}^\top\left(\tau_k^{(n)}\right)\mathbf{s}_{k,j}^{(n)}\mathbf{c}^\top\right) \\ &\quad \times \boldsymbol{\Sigma}_{k|k-1}^{(n)}\left(\mathbf{g}^\top\left(\tau_k^{(n)}\right)\mathbf{s}_{k,j}^{(n)}\mathbf{c}^\top\right)^\top \qquad (23)\end{aligned}$$

$\mathbf{s}_{k,j}^{(n)} = [s_{k-L+1}^{(n)}, \ldots, s_{k+L-1}^{(n)}, s_{k+L} = S_j]^\top$, and $\boldsymbol{\mu}_{k|k-1}^{(n)}$ and $\boldsymbol{\Sigma}_{k|k-1}^{(n)}$ are the predictive channel mean and covariance matrix, respectively, obtained by Kalman filtering from the observations and the nth state trajectory, $(\mathbf{s}_{0:k-1}, \tau_{0:k-1})^{(n)}$.

It is important to remark that the implementation of the proposed SIS algorithm requires a bank of Kalman filters (one for each particle) in order to compute the fading process statistics that are needed for the importance pdf and the weight update equation. The combination of the SIS algorithm and Kalman filtering has already been applied to other communication problems and is sometimes termed mixture Kalman filter (MKF) [11].

D. Smoothing

An important characteristic of the digital transmission system represented by (6) and (7) is that, for $\tau_k > 0$, each transmitted symbol contributes to $2L$ successive observations, where L is the ISI span parameter. Specifically, the symbol s_k, is a component of each data vector in the sequence

$$\mathbf{s}_{k-L:k+L-1} = \{\mathbf{s}_{k-L}, \mathbf{s}_{k-L+1}, \ldots, \mathbf{s}_{k+L-1}\} \qquad (24)$$

and, therefore, it is statistically dependent on the corresponding observations

$$y_{k-L:k+L-1} = \{y_{k-L}, y_{k-L+1}, \ldots, y_{k+L-1}\}. \qquad (25)$$

As a consequence, a reliable detection of the symbol sequence up to time k, $s_{0:k}$, requires adequate processing of (at least) the observations $y_{0:k+L-1}$.

A large ISI span has pernicious effects on the SIS algorithm, which requires a very large number of particles to attain a good performance. The usual strategy to account for this difficulty is to perform some type of *smoothing* [10], [12]–[14]. Fixed-interval smoothing algorithms [10], [15] do not fit the timing recovery problem requirements well because they involve reverse time processing.[3] Generic fixed-lag smoothing consists of estimating the state at time $k - D$ (where D is the smoothing lag) using the observations up to time k. The simplest approach to achieve this is to run the SIS algorithm, as described in the previous section, up to time k before detecting the symbol at time $k - D$. This method does not imply any extra computations but yields poor performance because it does not address the true problem, which is the accurate weighting of particles [10]. Alternative approaches, based on approximations, have been proposed [12], the most straightforward of which is to work with the posterior distribution $p(\mathbf{s}_{0:k-D}, \tau_{0:k}|y_{0:k})$ and its associated likelihood, $p(y_{k-D:k}|\mathbf{s}_{0:k-D}, \tau_{0:k}, y_{0:k-D-1})$, instead of $p(\mathbf{s}_{0:k}, \tau_{0:k}|y_{0:k})$ and $p(y_k|\mathbf{s}_{0:k}, \tau_{0:k}, y_{0:k-1})$, respectively [13], [14].

The proposed SIS algorithm can be used, with suitable modifications, to recursively compute a discrete smooth probability measure of the form

$$\Xi_k = \left\{(\mathbf{s}_{0:k-D}, \tau_{0:k})^{(n)}, w_k^{(n)}\right\}_{n=1}^N \qquad (26)$$

which yields an approximation of the desired pdf, $p(\mathbf{s}_{0:k-D}, \tau_{0:k}|y_{0:k})$. To obtain the new algorithm, we substitute the importance and likelihood functions as in the

[3]Let $x_{0:t_f}$ and $y_{0:t_f}$ denote the states and available observations. A fixed-interval smoothing algorithm approximates $p(x_{0:t_f}|y_{0:t_f})$ sequentially and then performs a second stage of processing where $p(x_{t_f-1}|x_{t_f}, y_{0:t_f}), \ldots, p(x_t|x_{t_f:-1:t+1}, y_{0:t_f}), \ldots, p(x_0|x_{t_f:-1:1}, y_{0:t_f})$ are approximated.

equation shown at the bottom of the page. As a result, the former importance sampling and weight update equations, (12) and (13), respectively, become

$$(\mathbf{s}_{k-D},\tau_k)^{(n)} \sim \pi\left(\mathbf{s}_{k-D},\tau_k|\mathbf{s}_{0:k-D-1}^{(n)},\tau_{0:k-1}^{(n)},y_{0:k}\right) \quad (27)$$

and

$$\tilde{w}_k^{(n)} = w_{k-1}^{(n)}\frac{p\left(\tau_k^{(n)}|\tau_{k-1}^{(n)}\right)L_k^{(n)}(D)}{\pi\left(\mathbf{s}_{k-D}^{(n)},\tau_k^{(n)}|\mathbf{s}_{0:k-D-1}^{(n)},\tau_{0:k-1}^{(n)},y_{0:k}\right)} \quad (28)$$

where $L_k^{(n)}(D)$ is the smooth likelihood given by

$$L_k^{(n)}(D) = \sum_{\mathcal{S}^D} p\left(y_{k-D:k}|\mathbf{s}_{0:k-D}^{(n)}, s_{k-D+L+1:k+L},\tau_{0:k}^{(n)},y_{0:k-D-1}\right). \quad (29)$$

The summation $\sum_{\mathcal{S}^D}(\cdot)$ in the above equation is carried out over all possible combinations of the sequence $s_{k-D+L+1:k+L} \in \mathcal{S}^D$. Given the nth particle $\mathbf{s}_{k-D}^{(n)}$, each symbol combination yields one sequence of subsequent data vectors $\mathbf{s}_{k-D+1:k}$.

A convenient form of the proposal distribution in (27) is

$$\pi\left(\mathbf{s}_{k-D},\tau_k|\mathbf{s}_{0:k-D-1}^{(n)},\tau_{0:k-1}^{(n)},y_{0:k}\right) = p\left(\tau_k^{(n)}|\tau_{k-1}^{(n-1)}\right)p\left(\mathbf{s}_{k-D}|\mathbf{s}_{0:k-D-1}^{(n)},\tau_{0:k}^{(n)},y_{0:k}\right) \quad (30)$$

where the delay is drawn from the corresponding prior $\tau_k^{(n)} \sim \mathcal{N}(\tau;\alpha\tau_{k-1}^{(n)},\sigma_u^2)$, while the data $\mathbf{s}_{k-D}^{(n)}$ are drawn from $p(\mathbf{s}_{k-D}|\mathbf{s}_{0:k-D-1}^{(n)},\tau_{0:k}^{(n)},y_{0:k})$. Recalling that the symbols are modeled as i.i.d. uniform random variables, the latter distribution can be decomposed as

$$p\left(\mathbf{s}_{k-D}|\mathbf{s}_{0:k-D-1}^{(n)},\tau_{0:k}^{(n)},y_{0:k}\right) \propto \sum_{\mathcal{S}^D} p\left(y_{k-D:k}|\mathbf{s}_{k-D}\,\mathbf{s}_{0:k-D-1}^{(n)},s_{k-D+L+1:k+L},\tau_{0:k}^{(n)},y_{0:k-D-1}\right) \quad (31)$$

and each likelihood in the summation can be numerically computed using

$$p(y_{k-D:k}|\mathbf{s}_{0:k},\tau_{0:k},y_{0:k-D-1}) \propto \prod_{j=0}^{D}\mathcal{N}\left(y_{k-j};\overline{y}_{k-j},\sigma_{y_{k-j}}^2\right) \quad (32)$$

where $\overline{y}_{k-j}$ and $\sigma_{y_{k-j}}^2$ are defined as in (22) and (23), respectively.

Some simplification can be achieved if we notice that in order to build $\mathbf{s}_{k-D}^{(n)}$ given $\mathbf{s}_{k-D-1}^{(n)}$ only one symbol, s_{k-D+L}, needs to be sampled. Hence, we can write

$$s_{k-D+L}^{(n)} \sim \rho^{(n)}(s_{k-D+L}) \quad (33)$$

and let

$$\mathbf{s}_{k-D}^{(n)} = \mathbf{S}\mathbf{s}_{k-D-1}^{(n)} + \left[0,\ldots,0,s_{k-D+L}^{(n)}\right]^\top \quad (34)$$

while the probability mass function (pmf) $\rho^{(n)}(s_{k-D+L})$ is defined according to (31) [see (35), shown at the bottom of the page], where the summation $\sum_{\mathcal{S}^{D+1}}(\cdot)$ is over all possible combinations of the sequence $s_{k-D+L:k+L} \in \mathcal{S}^{D+1}$.

The probability measure Ξ_k can be used to sequentially estimate the epoch and the received data, namely

$$\hat{\tau}_k^{\mathrm{MMSE}} = \sum_{n=1}^{N}\tau_k^{(n)}w_k^{(n)} \quad (36)$$

is the (approximate) minimum mean-square error (MMSE) estimate of the relative delay of the kth symbol, while

$$\hat{s}_{k-D}^{\mathrm{map}} = \arg\max_{s_{k-D}\in\mathcal{S}}\left\{\sum_{n=1}^{N}w_k^{(n)}\delta^{(n)}\left(s_{k-D}^{(n)}\right)\right\} \quad (37)$$

is the marginal MAP estimate of s_{k-D} given $y_{0:k}$.

E. Resampling

A major problem in the practical implementation of the SIS algorithm described so far is that the discrete measure Ξ_k degenerates quickly, i.e., after a few time steps, most of the importance weights have negligible values ($w_k^{(n)} \simeq 0$) and only a few particles with significant weights remain *useful*. The common solution to this problem is to *resample* the particles [10]. Resampling is an algorithmic step that stochastically eliminates particles with small weights, while those with larger weights are replicated. In its simplest form, resampling takes Ξ_k as an input and produces a new discrete measure $\tilde{\Xi}_k = \{(\tilde{\mathbf{s}}_{0:k-D},\tilde{\tau}_{0:k})^{(n)},1/N\}_{n=1}^N$, where $(\tilde{\mathbf{s}}_{0:k-D}^{(n)},\tilde{\tau}_{0:k}^{(n)}) = (\mathbf{s}_{0:k-D}^{(i)},\tau_{0:k}^{(i)})$ with probability $w_k^{(i)}$. The resampled trajectories are all equally weighted (i.e., all importance weights are reset to $1/N$).

Resampling at every time step is not needed in general. For the algorithms proposed in this paper, we have considered that

---

$$\begin{array}{rcl} \text{SIS} & & \text{SIS with smoothing} \\ \pi(\mathbf{s}_k,\tau_k|\mathbf{s}_{0:k-1},\tau_{0:k-1},y_{0:k}) & \longrightarrow & \pi(\mathbf{s}_{k-D},\tau_k|\mathbf{s}_{0:k-D-1},\tau_{0:k-1},y_{0:k}) \\ p(y_k|\mathbf{s}_{0:k},\tau_{0:k},y_{0:k-1}) & \longrightarrow & p(y_{k-D:k}|\mathbf{s}_{0:k-D},\tau_{0:k},y_{0:k-D-1}) \end{array}$$

$$\rho^{(n)}(s_{k-D+L}) = \frac{\sum_{\mathcal{S}^D} p\left(y_{k-D:k}|\mathbf{s}_{0:k-D-1}^{(n)},s_{k-D+L},s_{k-D+L+1:k+L},\tau_{0:k}^{(n)},y_{0:k-D-1}\right)}{\sum_{\mathcal{S}^{D+1}} p\left(y_{k-D:k}|\mathbf{s}_{0:k-D-1}^{(n)},s_{k-D+L:k+L},\tau_{0:k}^{(n)},y_{0:k-D-1}\right)} \quad (35)$$

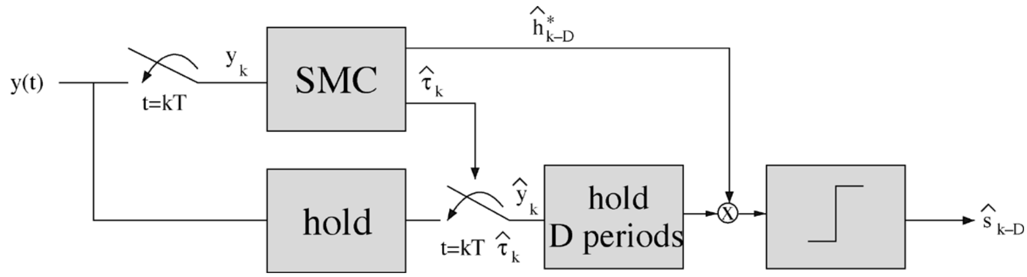

Fig. 1. Open-loop architecture.

resampling is carried out whenever the effective sample size of the particle filter [10], approximated as

$$\hat{N}_{\text{eff}} = \frac{1}{\sum_{i=1}^{N}\left(w_k^{(i)}\right)^2} \leq N \tag{38}$$

goes below a certain threshold (typically a fraction of N). Intuitively, the effective sample size represents the number of independent particles drawn from the true posterior distribution of interest that would be needed to compute signal estimates with the same quality as those obtained from Ξ_k.

F. Complexity

Comments regarding the complexity of the proposed method are in order. Equation (32) means that, at time k, a Kalman smoother has to be run for each particle and for each possible combination of symbols $s_{k-D+L:k+L}$, which involves D Kalman prediction and update steps for each symbol sequence. Therefore, the complexity of processing each particle grows exponentially with D, yielding $\mathcal{O}(|\mathcal{S}|^D)$, where $|\mathcal{S}|$ is the number of elements in the symbol alphabet.

On the other hand, it must be taken into account that the propagation of each particle in Ξ_k can be done independently, so the algorithm lends itself to efficient implementation using massively parallel hardware, except for the resampling step discussed in Section III-E (see [16] and [17] for a detailed discussion on the parallelization of resampling). Using state of the art techniques, resampling can be carried out with $\mathcal{O}(N)$ operations, where N is the number of particles; hence, the overall complexity of the proposed SMC smoothing algorithm grows with $\mathcal{O}(N|\mathcal{S}|^D)$.

IV. Receiver Architectures

Although the transmitted data with their timing can be estimated together using the SIS algorithm described above, it should be observed that the proposed method does not remove ISI. In other words, although the relative symbol delays τ_k are estimated, the sampling instants $t = kT$ are not corrected to attain a better timing and avoid ISI. As a consequence, the SER that can be attained by detecting according to (37) is lower bounded by the SER of the maximum likelihood sequence detector (MLSD) with perfect knowledge of the sequence of composite channel vectors $h_k\mathbf{g}(\tau_k)$. In general, with the assumption that the autocorrelation of the continuous-time delay process, $\tau(t)$ is wider than the symbol period T, a lower SER is achieved with a matched-filter receiver sampled at $t = kT - \tau_k$ because ISI-free observations $y_k = y(kT - \tau_k)$ can be obtained.

The SER of the matched-filter detector can be attained with the proposed SMC algorithm if an adequate receiver architecture is used. Below, we present two possible configurations that use the SMC approach for recovering the timing and the channel complex amplitude, thereby allowing the removal of the ISI before detection.

A. Open-Loop Receiver

We consider first the double branch structure depicted in Fig. 1. In the upper branch, the SMC block performs the proposed SIS algorithm, which is used to compute Monte Carlo MMSE estimates of the timing according to (36). The received signal is held in the lower branch, until the delay estimate is computed, and then sampled at $t = kT - \hat{\tau}_k^{\text{MMSE}}$ to obtain the observation

$$\hat{y}_k = y\left(kT - \hat{\tau}_k^{\text{MMSE}}\right) = h_k \mathbf{s}_k^\top \mathbf{g}\left(\tau_k - \hat{\tau}_k^{\text{MMSE}}\right) + v_k \tag{39}$$

where $h_k = h(kT + \tau_k - \hat{\tau}_k^{\text{MMSE}})$ and $\mathbf{s}_k = [s_{k-L+1}, \ldots, s_{k+L}]^\top$, if $\tau_k - \hat{\tau}_k^{\text{MMSE}} > 0$, while $\mathbf{s}_k = [s_{k-L}, \ldots, s_{k+L-1}]^\top$, if $\tau_k - \hat{\tau}_k^{\text{MMSE}} < 0$. Ideally, when $\hat{\tau}_k^{\text{MMSE}} \simeq \tau_k$, the ISI is removed and

$$\hat{y}_k \simeq h_k s_k + v_k. \tag{40}$$

Hence, minimal SER is achieved by multiplying by $\hat{h}_k^*$, followed by a simple threshold detector that makes a decision based on the minimal Euclidean distance between $\hat{y}_k$ and the elements in the symbol alphabet, $\mathcal{S}$. The aforementioned channel estimate is computed as

$$\hat{h}_k = \sum_{n=1}^{N} \mathbf{c}^\top \boldsymbol{\mu}_k^{(n)} w_{k+D}^{(n)} \tag{41}$$

where

$$\boldsymbol{\mu}_k^{(n)} = \mathrm{E}_{p\left(\mathbf{h}_k | \mathbf{s}_{0:k}^{(n)}, \tau_{0:k}^{(n)}, y_{0:k}\right)}[\mathbf{h}_k] \tag{42}$$

is the posterior channel estimate associated with the nth particle and calculated by Kalman filtering. Note, however, that this estimate cannot be obtained until the symbol sequences up to time k, $\{\mathbf{s}_{0:k}^{(n)}\}_{n=1}^N$ are available in the particle filter, which occurs at time $k + D$ due to the smoothing. Hence, there is need to hold $\hat{y}_k$ until it can be conveniently de-rotated for accurate detection, as depicted in Fig. 1.

TABLE I
SIS WITH RESAMPLING FOR THE OPEN-LOOP ARCHITECTURE

Initialization
 $\boldsymbol{\mu}_0 = 0$
 $\boldsymbol{\Sigma}_0 = \mathbf{I}$
For $k = 0$ to M *(total number of symbols)*
 For $n = 1$ to N *(total number of particles)*
  For each $s_{k-D+L+1:k+L} \in \mathcal{S}^D$
   Kalman prediction: $\left\{\boldsymbol{\mu}_{k-j|k-j-1}, \boldsymbol{\Sigma}_{k-j|k-j-1}\right\}_{j=0}^{D-1}$
  Draw $\tau_k^{(n)} \sim \mathcal{N}(\alpha\tau_{k-1}^{(n)}, \sigma_u^2)$
  Draw $s_{k-D+L}^{(n)} \sim \rho^{(n)}(s_{k-D+L})$ according to equation (35)
  Build $\mathbf{s}_{k-D}^{(n)}$
  Update weights $\tilde{w}_k^{(n)}$ according to equation (28)
  Normalize weights $w_k^{(n)} = \left(\sum_{i=1}^N \tilde{w}_k^{(i)}\right)^{-1} \tilde{w}_k^{(n)}$
  Kalman update:
$$\boldsymbol{\mu}_{k-D}^{(n)} = E_{p(\mathbf{h}_{k-D}|\mathbf{s}_{0:k-D},\tau_{0:k-D},y_{0:k-D})}\left[\mathbf{h}_{k-D}\right],$$
$$\boldsymbol{\Sigma}_{k-D}^{(n)} = E_{p(\mathbf{h}_{k-D}|\mathbf{s}_{0:k-D},\tau_{0:k-D},y_{0:k-D})}\left[(\mathbf{h}_{k-D} - \boldsymbol{\mu}_{k-D})(\mathbf{h}_{k-D} - \boldsymbol{\mu}_{k-D})^H\right],$$
 Resample if $\hat{N}_{eff} = \frac{1}{\Sigma_{n=1}^N \left(w_k^{(n)}\right)^2} < N/2$
 Estimate delay: $\hat{\tau}_k^{mmse} = \sum_{n=1}^N \tau_k^{(n)} w_k^{(n)}$
 Sample: $\hat{y}_k = y(t = kT - \hat{\tau}_k^{mmse})$
 Channel estimate: $\hat{h}_{k-D} = \sum_{n=1}^N \mathbf{c}^\top \boldsymbol{\mu}_{k-D}^{(n)} w_k^{(n)}$
 Detect symbol, $\hat{s}_{k-D} = \arg\min_{s\in\mathcal{S}} \parallel \hat{y}_{k-D}\hat{h}_{k-D}^* - s \parallel$

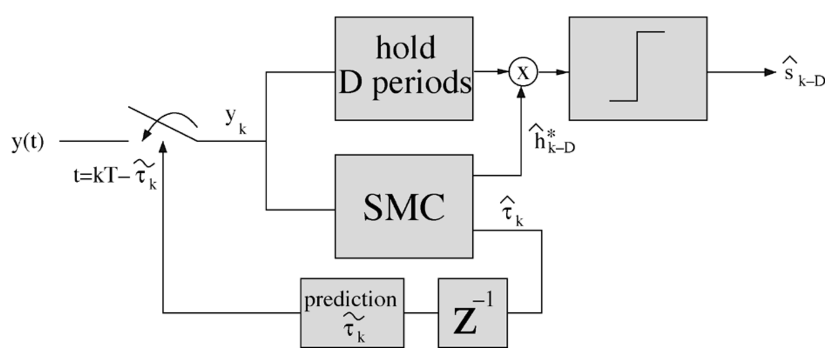

Fig. 2. Closed-loop architecture.

The recursive steps of the proposed SMC algorithm with the open-loop architecture are summarized in Table I.

B. Closed-Loop Receiver

The main drawback of the previous configuration is the necessity of sampling the received signal twice per symbol period. To avoid this drawback, the closed-loop receiver architecture shown in Fig. 2 can be used.

The SMC block represents the SIS algorithm with resampling described in Section III, which yields asymptotically optimal MMSE estimates of the relative symbol delay $\hat{\tau}_k^{\mathrm{MMSE}}$. This estimate is fed back and used to adjust the epoch of the next observation. Therefore, instead of sampling the received signal uniformly, to obtain $y_k = y(kT)$, the sampling time is adaptively selected according to the most recent estimate of the relative symbol delay, to yield $y_k = y(kT - \tilde{\tau}_k)$, where $\tilde{\tau}_k = \alpha\hat{\tau}_{k-1}^{\mathrm{MMSE}}$ is the MMSE prediction of τ_k.

The observations collected in this way have the form

$$y_k = h_k \mathbf{s}_k^\top \mathbf{g}(\tau_k - \tilde{\tau}_k) + v_k \qquad (43)$$

where $h_k = h(kT + \tau_k - \tilde{\tau}_k)$ and the symbol vector is $\mathbf{s}_k = [s_{k-L+1}, \ldots, s_{k+L}]^\top$, if $\tau_k - \tilde{\tau}_k > 0$, and $\mathbf{s}_k = [s_{k-L}, \ldots, s_{k+L-1}]^\top$ otherwise. Notice that if $\tilde{\tau}_k \simeq \tau_k$, the resulting observation $y_k \simeq h_k s_k + v_k$ is free of ISI and the corresponding symbol can be optimally detected multiplying y_k by $\hat{h}_k^*$ and using a simple threshold detector. As in the open-loop structure, the final detection step has to be delayed until $\hat{h}_k$ is available at time $k + D$.

The recursive steps of the proposed algorithm are summarized in Table II.

V. POSTERIOR CRAMÉR–RAO BOUND

Although available statistical results guarantee that the MMSE estimate of the delay $\hat{\tau}_k^{\mathrm{MMSE}}$ provided by the SIS algorithm converges asymptotically to the true MMSE delay estimate, it is apparent that, for $N < \infty$, we only obtain an approximation to the desired estimator and, as a consequence, a certain degradation in performance of the proposed adaptive receivers can be expected. In order to study the efficiency of the proposed estimation method, it is of a great interest to compute the

TABLE II
SIS WITH RESAMPLING FOR THE CLOSED-LOOP ARCHITECTURE

Initialization
 $\tilde{\tau}_0 = 0$
 $\boldsymbol{\mu}_0 = 0$
 $\boldsymbol{\Sigma}_0 = \mathbf{I}$
For $k = 0$ to M *(total number of symbols)*
 Sample $y_k = y(t = kT - \tilde{\tau}_k)$
 For $n = 1$ to N *(total number of particles)*
  For each $s_{k-D+L+1:k+L} \in \mathcal{S}^D$
   Kalman prediction: $\left\{\boldsymbol{\mu}_{k-j|k-j-1}, \boldsymbol{\Sigma}_{k-j|k-j-1}\right\}_{j=0}^{D-1}$
  Draw $\tau_k^{(n)} \sim \mathcal{N}(\alpha\tau_{k-1}^{(n)}, \sigma_u^2)$
  Draw $s_{k-D+L}^{(n)} \sim \rho^{(n)}(s_{k-D+L})$ according to equation (35)
  Build $\mathbf{s}_{k-D}^{(n)}$
  Update weights $\tilde{w}_k^{(n)}$ according to equation (28)
  Normalize weights $w_k^{(n)} = \left(\sum_{i=1}^N \tilde{w}_k^{(i)}\right)^{-1} \tilde{w}_k^{(n)}$
  Kalman update: $\boldsymbol{\mu}_{k-D}^{(n)}, \boldsymbol{\Sigma}_{k-D}^{(n)}$
 Resample if $\hat{N}_{eff} = \frac{1}{\sum_{n=1}^N \left(w_k^{(n)}\right)^2} < N/2$
 Timing recovery, channel estimation and symbol detection
  $\hat{\tau}_k^{mmse} = \sum_{n=1}^N \tau_k^{(n)} w_k^{(n)}$
  $\tilde{\tau}_{k+1} = \alpha\hat{\tau}_k^{mmse}$
  $\hat{h}_{k-D} = \sum_{n=1}^N \mathbf{c}^\top \boldsymbol{\mu}_{k-D}^{(n)} w_k^{(n)}$
  $\hat{s}_{k-D} = \arg\min_{s\in\mathcal{S}} \parallel y_{k-D}\hat{h}_{k-D}^* - s \parallel$
 Timing prediction: $\tilde{\tau}_{k+1} = \alpha\hat{\tau}_k$

variance bounds on the estimation errors and compare them with the lowest bounds corresponding to the optimal estimator.

When the parameter of interest is assumed fixed, the lower bound for the variance of any unbiased estimator is given by the well-known Cramér–Rao bound (CRB) [18] which, in turn, is obtained from the inverse of the Fisher information matrix (FIM). However, for Bayesian models where the parameter of interest is considered random, the lowest achievable variance is given by the PCRB [19], [20]. Therefore, we wish to derive the PCRB associated to the timing process $\tau_{0:k}$ in order to obtain a lower bound for the MSE of the delay estimates.

We define the $k+1$ dimensional vectors $\underline{\tau} = [\tau_0, \tau_1, \ldots, \tau_k]^\top$ and $\hat{\underline{\tau}} = [\hat{\tau}_0, \hat{\tau}_1, \ldots, \hat{\tau}_k]^\top$, where $\hat{\tau}_k$ is an arbitrary estimate of τ_k. For the signal model of interest in this paper, the PCRB can be stated as

$$\mathbf{P}_k = E_{p(\mathbf{s}_{0:k}, \tau_{0:k}, \mathbf{h}_{0,k}, y_{0:k})}\left[(\hat{\underline{\tau}} - \underline{\tau})(\hat{\underline{\tau}} - \underline{\tau})^\top\right] \geq \mathbf{J}(\tau_{0:k})^{-1}$$

where $\mathbf{J}(\tau_{0:k})$ is the $(k+1)\times(k+1)$ FIM, which is defined, element wise, as

$$[\mathbf{J}(\tau_{0:k})]_{ij} = E_{p(\mathbf{s}_{0:k}, \tau_{0:k}, \mathbf{h}_{0,k}, y_{0:k})} \times \left[\frac{-\partial^2 \log\left(p(\mathbf{s}_{0:k}, \tau_{0:k}, \mathbf{h}_{0,k}, y_{0:k})\right)}{\partial\tau_i\partial\tau_j}\right].$$

Notice that the ith element in the diagonal of $\mathbf{J}(\tau_{0:k})$, which we subsequently denote as $J_i = [\mathbf{J}(\tau_{0:k})]_{ii}$, corresponds to the inverse of the lowest achievable MSE in the estimation of τ_i.

It has been shown [20] that the direct computation of the FIM can be avoided by using a recursive method which sequentially evaluates the inverse MSE of τ_{k+1}, specifically

$$J_{k+1} = D_k^{22} - D_k^{21}\left(J_k + D_k^{11}\right)^{-1} D_k^{12}. \quad (44)$$

The terms in (44) are

$$\begin{aligned} D_k^{11} &= E_{p(\mathbf{s}_{0:k+1}, \tau_{0:k+1}, \mathbf{h}_{0,k+1}, y_{0:k+1})} \\ &\quad\times\left[-\triangle_{\tau_k}^{\tau_k} \log\left(p(\tau_{k+1}|\tau_k)\right)\right] & (45)\\ D_k^{12} &= E_{p(\mathbf{s}_{0:k+1}, \tau_{0:k+1}, \mathbf{h}_{0,k+1}, y_{0:k+1})} \\ &\quad\times\left[-\triangle_{\tau_k}^{\tau_{k+1}} \log\left(p(\tau_{k+1}|\tau_k)\right)\right] & (46)\\ D_k^{21} &= E_{p(\mathbf{s}_{0:k+1}, \tau_{0:k+1}, \mathbf{h}_{0,k+1}, y_{0:k+1})} \\ &\quad\times\left[-\triangle_{\tau_{k+1}}^{\tau_k} \log\left(p(\tau_{k+1}|\tau_k)\right)\right] & (47)\\ D_k^{22} &= E_{p(\mathbf{s}_{0:k+1}, \tau_{0:k+1}, \mathbf{h}_{0,k+1}, y_{0:k+1})} \\ &\quad\times\left[-\triangle_{\tau_{k+1}}^{\tau_{k+1}} \log\left(p(\tau_{k+1}|\tau_k)\right)\right] & (48)\\ &\quad + E_{p(\mathbf{s}_{0:k+1}, \tau_{0:k+1}, \mathbf{h}_{0,k+1}, y_{0:k+1})} \\ &\quad\times\left[-\triangle_{\tau_{k+1}}^{\tau_{k+1}} \log\left(p(y_{k+1}|\tau_{k+1}, \mathbf{s}_{k+1})\right)\right] & (49) \end{aligned}$$

where $\triangle$ denotes the second derivative operator, defined as $\triangle_{\tau_k}^{\tau_{k+1}} = \partial^2/\partial\tau_k\partial\tau_{k+1}$ and $\log(\cdot)$ is the natural logarithm. Recursion (44) is initialized at time $t = -1$, in the absence of observations, by considering $J_{-1} = 12/T^2$, which is the inverse of the variance of the uniform distribution in $[0, T)$. Notice that this is the only *a priori* information we use regarding the delay.

It is straightforward to numerically evaluate (45)–(48), which yield

$$D_k^{11} = \frac{\alpha^2}{\sigma_u^2} \quad (50)$$

$$D_k^{12} = D_k^{21} = -\frac{\alpha}{\sigma_u^2} \quad (51)$$

while, as for (49)

$$D_k^{22} = \left(\frac{1}{\sigma_u^2} + E_{p(\mathbf{s}_{0:k+1}, \tau_{0:k+1}, \mathbf{h}_{0,k+1}, y_{0:k+1})} \times\left[\frac{1}{|\mathcal{S}|^{2L}} \sum_{\mathbf{s}_k\in\mathcal{S}^{2L}} \frac{|h_k|^2\left(\mathbf{s}_k^\top \nabla_{\tau_{k+1}}\mathbf{g}(\tau_{k+1})\right)^2}{\sigma_v^2}\right]\right). \quad (52)$$

Unfortunately, it is not possible to obtain a closed-form expression for the expectation in the above equation. Instead, as suggested in [20], we can estimate it using Monte Carlo simulation. When Q i.i.d. state trajectories are generated, we approximate D_k^{22} as

$$\hat{D}_k^{22} = \left(\frac{1}{\sigma_u^2} + \frac{1}{|\mathcal{S}|^{2L} Q} \sum_{\mathbf{s}_k\in\mathcal{S}^{2L}} \sum_{j=1}^{Q} \frac{\left|h_k^{(j)}\right|^2\left(\mathbf{s}_k^\top \nabla_{\tau_{k+1}}\mathbf{g}\left(\tau_{k+1}^{(j)}\right)\right)^2}{\sigma_v^2}\right) \quad (53)$$

where $\nabla_{\tau_{k+1}} = \partial/\partial\tau_{k+1}$.

VI. COMPUTER SIMULATIONS

In order to obtain numerical results, we considered a differentially encoded binary modulation with symbol alphabet $\{\pm 1\}$ and symbol period $T = 10^{-4}$ s and a flat fading channel with fading rate 0.0022. The delay was modeled as a first order AR process with parameter $a = 0.999$ and noise variance $\sigma_u^2 = 3 \times 10^{-4}$ (unless otherwise specifically stated). The pulse waveform was a time-limited raised-cosine with a roll-off factor $a = 0.9$

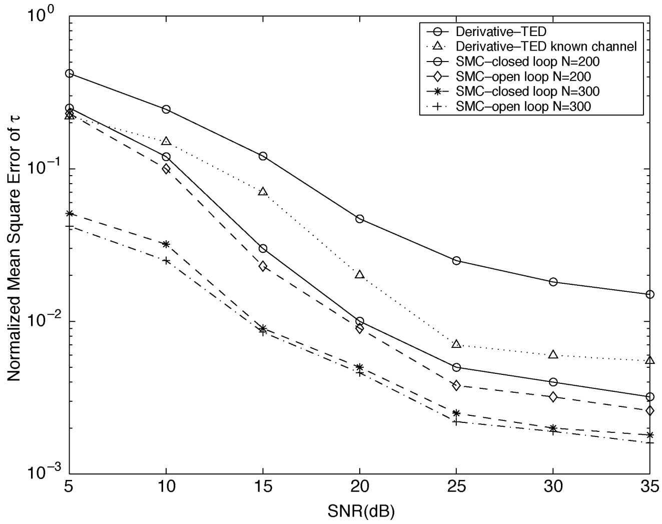

Fig. 3. NMSE for SMC-based timing estimators and AML TEDs.

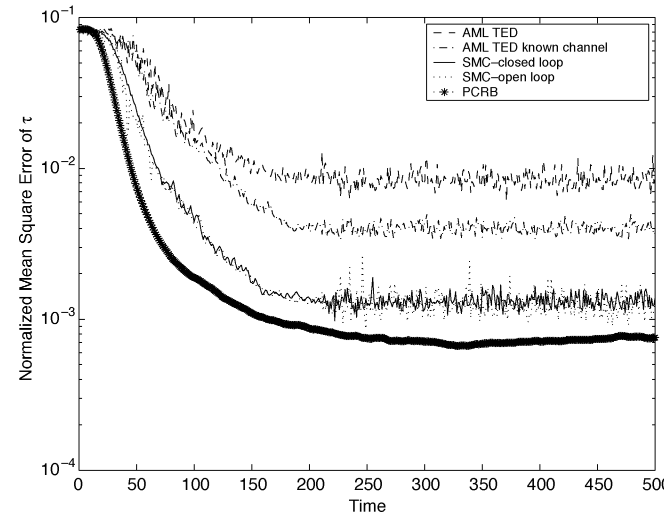

Fig. 4. Time evolution of the NMSE versus the PCRB.

that yields an ISI span of $2L = 4$ symbols. Transmission in frames of $M = 500$ bits was simulated.

Timing recovery was assessed in terms of the normalized mean-square error (NMSE) in the estimation of the symbol delays, τ_k. The NMSE was numerically estimated by running several independent simulations and averaging their results. Specifically, for Q independent trials, the MSE was approximated as $\text{NMSE} = (1/MQ)\sum_{i=1}^{Q}\sum_{k=0}^{M-1}((1/T)(\tau_k - \hat{\tau}_k^{(i)}))^2$ where $\hat{\tau}_k^{(i)}$ is the estimate of the delay for the kth symbol in the ith simulation. The NMSE of the proposed SMC techniques was compared with the approximate maximum likelihood (AML) timing error detector (TED) described in [1, Sec. 7.5]. The AML TED is based on a low signal-to-noise ratio (SNR) assumption and uses a receiver structure consisting of two matched filters (see [1] for details) to adaptively estimate the delays τ_k and suppress the ISI. For this classical technique, we considered two alternatives: one with known channel and another one where the channel coefficients were estimated using a Kalman filter.

Fig. 3 shows the NMSE attained by the proposed SMC blind receivers, both with open-loop and closed-loop configurations, and the AML TEDs for several values of SNR. The particle filtering algorithms were implemented with a smoothing lag $D = 2$ and $N = 200$, 300 particles. Each value of the NMSE was averaged over $Q = 100$ independent simulations. The SMC receivers with open- and closed-loop structures exhibited similar performances and attained consistently lower NMSEs than the AML TEDs.

Next, we considered the evolution of the NMSE within one frame $\text{NMSE}_k = (1/Q)\sum_{i=1}^{Q}((1/T)(\tau_k - \hat{\tau}_k^{(i)}))^2$, $k = 0, 1, \ldots, M-1$ and compared it with the lower bound given by the PCRB as described in Section V. Fig. 4 shows the results we obtained, for fixed SNR$=$ 25 dB, after $Q = 100$ independent simulation trials and using particle filters with $N = 300$ particles and lag $D = 2$. We observed that the proposed SMC algorithms perform close to the theoretical limit, while the NMSE of the AML TEDs fell considerably further from the PCRB.

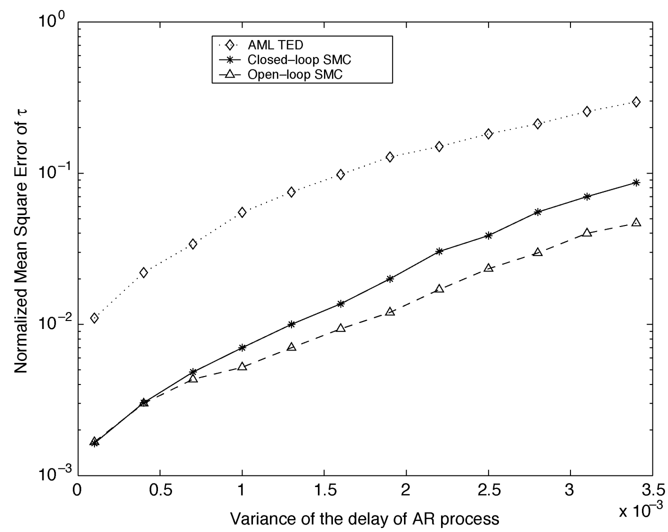

Fig. 5. MSE versus variance of the delay AR process.

In order to establish the relative merit of the open- and closed-loop architectures, we carried out a computer experiment where different values of the variance of the timing process σ_u^2 were considered. Fig. 5 shows the NMSE for the two proposed configurations and different values of σ_u^2. As before, we kept $N = 300$, $D = 2$ and show the average results of $Q = 100$ simulations. The two receivers performed almost identically for low values of σ_u^2 while, for higher values of the variance, the open-loop structure attained a clearly lower NMSE. This is consistent with the fact that the closed-loop receiver relies on a *prediction* of the delay in order to sample the observed signal, while in the open-loop scheme it is *estimated*.

Finally, we assessed the bit-error rate (BER) of the proposed receivers for increasing SNR. Let us remark that there is a delay ambiguity inherent to blind data detection. In particular, any shift of j time steps in the optimal (MAP) data sequence estimate is also a valid MAP estimate as long as the delay sequence is also time-shifted accordingly, i.e., $\hat{\tau} = jT + \tau$, with integer

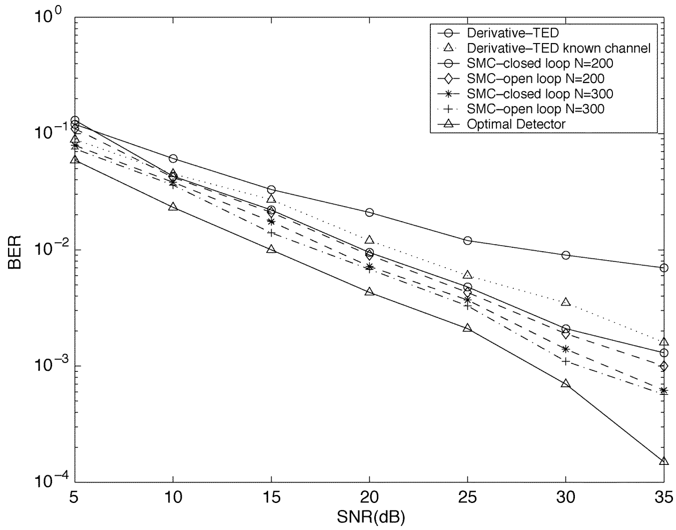

Fig. 6. BER versus SNR.

j. This ambiguity is easily removed if the data sequence length, M, is *a priori* known, as assumed in this paper.

In Fig. 6, the performance of the two SMC algorithms is compared with the two AML TEDs and also with the optimal matched-filter detector (with perfect knowledge of $\tau_{0:k}$ and $h_{0:k}$) that provides a lower bound for the BER. The parameters of the particle filters were N =200, 300 and $D = 2$. It is seen that the BER achieved by the SMC adaptive receivers is very close to the optimal BER, obtained using the genie-aided matched filter, and significantly lower than the BER of the receivers based on the AML TEDs.

VII. Conclusion

We have presented a novel adaptive algorithm for blind synchronization and data detection in frequency-flat fast-fading wireless channels based on a Bayesian estimation approach and the SMC methodology. Assuming that the relative delay between the received signal and the local clock reference and the fast fading complex channel vary according to autoregressive models, the proposed method obtains asymptotically optimal estimates and, when combined with adequate receiver architectures, minimal BER. In particular, we have proposed two candidate structures, with open-loop and closed-loop configurations. The performance of the resulting blind adaptive receivers is studied both analytically, through the derivation of the PCRB for timing estimation, and through computer simulations. The latter show a close-to-optimal performance of the proposed receivers both in terms of timing recovery and BER.

References

[1] U. Mengali and A. N. D'Andrea, *Synchronization Techniques for Digital Receivers*. New York: Plenum, 1997.

[2] G. Vázquez and J. Riba, "Non data-aided digital synchronization," in *Signal Processing Advances in Wireless Communications. Trends in Single and Multi-User Systems (Vol. II)*, G. B. Giannakis and Y. Hua, Eds. Upper Saddle River, NJ: Prentice-Hall, 2000, ch. 9.

[3] S. Haykin, Ed., *Blind Deconvolution*. Upper Saddle River, NJ: Prentice-Hall, 1994.

[4] A. Doucet, N. de Freitas, and N. Gordon, *Sequential Monte Carlo Methods in Practice*. New York: Springer, 2001.

[5] R. E. Kalman, "A new approach to linear filtering and prediction problems," *J. Basic Eng.*, vol. 82, pp. 35–45, 1983.

[6] S. Haykin, *Adaptive Filter Theory*, 3rd ed. Upper Saddle River, NJ: Prentice-Hall, 1996.

[7] P. M. Djurić, J. H. Kotecha, J. Zhang, Y. Huang, T. Ghirmai, M. F. Bugallo, and J. Míguel, "Applications of particle filtering to selected problems in communications," *IEEE Signal Process. Mag.*, vol. 20, no. 9, pp. 19–38, Sep. 2003.

[8] R. Iltis, "Joint estimation of PN code delay and multipath using the extended Kalman filter," *IEEE Trans. Commun.*, vol. 38, no. 10, pp. 1677–1685, Oct. 1990.

[9] M. J. Omidi, S. Pasupathy, and P. G. Gulak, "Joint data and Kalman estimation for Rayleigh fading channels," *Wireless Personal Commun.*, no. 10, pp. 319–339, 1999.

[10] 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, 2000.

[11] R. Chen, X. Wang, and J. S. 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.

[12] T. C. Clapp and S. J. Godsill, "Fixed-lag blind equalization and sequence estimation in digital communications systems using sequential importance sampling," in *Proc. ICASSP*, vol. 5, Mar. 1999, pp. 2495–2498.

[13] J. Zhang and P. M. Djurić, "Joint estimation and decoding of space-time trellis codes," *J. Appl. Signal Process.*, vol. 3, pp. 305–315, 2002.

[14] J. Míguez and P. M. Djurić, "Blind equalization of frequency selective channels by sequential importance sampling," *IEEE Trans. Signal Process.*, to be published.

[15] W. Fong, S. J. Godsill, A. Doucet, and M. West, "Monte Carlo smoothing with application to audio signal enhancement," *IEEE Trans. Signal Process.*, vol. 50, no. 2, pp. 438–448, Feb. 2002.

[16] M. Bolić, P. M. Djurić, and S. Hong, "New resampling algorithms for particle filters," presented at the IEEE ICASSP, Apr. 2003.

[17] ——, "Resampling algorithms for particle filters suitable for parallel VLSI implementation," presented at the CISS, Mar. 2003.

[18] C. W. Therrien, *Discrete Random Signals and Statistical Signal Processing*. Upper Saddle River, NJ: Prentice-Hall, 1992.

[19] N. Bergman, "Posterior Cramér-Rao bounds for sequential estimation," in *Sequential Monte Carlo Methods in Practice*, A. Doucet, N. de Freitas, and N. Gordon, Eds. New York: Springer, 2001, ch. 15, pp. 321–338.

[20] P. Tichavský, C. H. Muravchik, and A. Nehorai, "Posterior Cramér-Rao bounds for discrete-time nonlinear filtering," *IEEE Trans. Signal Process.*, vol. 46, no. 5, pp. 1386–1396, May 1998.

Tadesse Ghirmai received the B.Sc. degree from Addis Ababa University, Addis Ababa, Ethiopia, in 1988, the M.S.E.E. degree from University of South Florida, Tampa, in 1995, and Ph.D. degree from the State University of New York at Stony Brook in 2004.

He is currently an Assistant Professor at Jackson State University, Jackson, MS. He was with the University of Asmara, Asmara, Eritrea, from 1988 to 1993, where he was also Assistant Lecturer and as a Lecturer from 1995 to 1999, respectively. His research interests are in statistical signal processing with emphasis on the theory of Monte Carlo methods and their application to communications.

Mónica F. Bugallo received the Ph. D. degree in computer engineering from the Universidade da Coruña, Coruña, Spain, in 2001.

From 1998 to 2000, she was with the Departamento de Electrónica y Sistemas, Universidade da Coruña, where she worked in interference cancellation applied to multiuser communication systems. In 2001, she joined the Department of Electrical and Computer Engineering, State University of New York at Stony Brook, where she is currently an Assistant Professor and teaches courses in digital communications and information theory. Her research interests lie in the area of statistical signal processing and its applications to different disciplines, including communications and biology.

Joaquín Míguez was born in Ferrol, Galicia, Spain, in 1974. He received the Licenciado en Informatica (M.Sc.) and Doctor en Informatica (Ph.D.) degrees from Universidade da Coruña, Coruña, Spain, in 1997 and 2000, respectively.

In late 2000, he joined Departamento de Electrónica e Sistemas, Universidade da Coruña, where he became an Associate Professor in July 2003. His research interests are in the field of statistical signal processing, with emphasis in the topics of Bayesian analysis, sequential Monte Carlo methods, adaptive filtering, stochastic optimization, and their applications to multiuser communications, smart antenna systems, target tracking, and vehicle positioning and navigation.

Petar M. Djurić (SM'99) received the B.S. and M.S. degrees in electrical engineering from the University of Belgrade, Belgrade, Yugoslavia, in 1981 and 1986, respectively, and the Ph.D. degree in electrical engineering from the University of Rhode Island, Kingston, in 1990.

From 1981 to 1986, he was a Research Associate with the Institute of Nuclear Sciences, Vinca, Belgrade. Since 1990, he has been with the State University of New York at Stony Brook, where he is a Professor in the Department of Electrical and Computer Engineering. He works in the area of statistical signal processing, and his primary interests are in the theory of modeling, detection, estimation, and time-series analysis and its application to a wide variety of disciplines, including wireless communications and biomedicine.