14th European Signal Processing Conference (EUSIPCO 2006), Florence, Italy, September 4-8, 2006, copyright by EURASIP

SMC ALGORITHMS FOR APPROXIMATE-MAP EQUALIZATION OF MIMO CHANNELS WITH POLYNOMIAL COMPLEXITY

Manuel A. Vázquez†, *Joaquín Míguez*‡

†Departamento de Electrónica e Sistemas, Universidade da Coruña,
Campus de Elviña s/n, 15071 A Coruña, Spain. E-mail: mvazquez@udc.es
‡Departamento de Teoría de la Señal y Comunicaciones, Universidad Carlos III de Madrid.
Avenida de la Universidad 30, 28911 Leganés, Madrid, Spain. E-mail: jmiguez@ieee.org

ABSTRACT

Sequential Monte Carlo (SMC) schemes have been recently proposed in order to perform optimal equalization of multiple input multiple output (MIMO) wireless channels. The main feaures of SMC techniques that make them appealing for the equalization problem are (a) their potential to provide asymptotically optimal performance in terms of bit error rate and (b) their suitability for implementation using parallel hardware. Nevertheless, existing SMC equalizers still exhibit a very high computational complexity, relative to the dimensions of the MIMO channel, which makes them useless in practical situations. In this paper we introduce two new SMC equalizers whose computational load is only of polynomial order with respect to the channel dimensions, and avoid computationally heavy tasks such as matrix inversions. The performance of the proposed techniques is numerically illustrated by means of computer simulations.

1. INTRODUCTION

SMC methods, also known as particle filtering (PF) algorithms, are simulation based techniques that aim at approximating the *a posteriori* probability density function (pdf) of a time-varying signal of interest (SOI), given some related observations, using a discrete probability measure with a random support [1, 2, 3]. PF algorithms explore the space of the SOI by generating random samples (termed particles) from a proposal distribution. These particles are then assigned proper weights [1] and yield the discrete approximation of the *a posteriori* pdf.

Particle filtering is a very general methodology that has found numerous applications in digital communications (see [3] for an overview). The reason is that, although SMC techniques are computationally intensive in absolute terms (because usually many particles have to be generated in order to obtain some prescribed level of performance), they are inherently suitable for implementation using parallel hardware and, therefore, hold promise of very high processing speeds. Many of the most recently proposed digital transmission systems can benefit from the fast processing capabilities of PF methods. In particular, those that involve communication through multiple input multiple output (MIMO) channels, e.g., multi-antenna systems [4, 5], are well known to require fast and sophisticated signal processing techniques to carry out fundamental tasks such as channel equalization. Practical MIMO channel equalization poses several problems because the complexity of maximum *a posteriori* (MAP) data detection grows exponentially with the number of inputs (NoI) and the length of the channel impulse response (CIR). In [6, 7] quasi-MAP SMC equalizers with polynomial complexity (with respect to the NoI and the CIR length) have been described, but even these techniques can be prohibitive for certain classes of MIMO systems, since they involve running banks of Kalman filters (KFs) and perform successive matrix inversions.

In this work, we propose two new SMC equalizers for nearly-MAP equalization of MIMO systems with polynomial (quadratic) complexity. They are specially suitable for systems with a large number of transmitting and receiving antennas, or when there exist strict time requirements, such as in online detection. Compared with the methods in [6, 7], the new equalizers substitute the KF banks by less complex parallel adaptive channel estimation algorithms that avoid matrix inversions altogether.

The remaining of the paper is organized as follows. In the next section, the signal model for transmission over a frequency-selective MIMO channel is described. In section 3, the standard application of SMC methods to MIMO equalization is discussed. The fundamental ideas behind the proposed SMC equalizers are introduced in Section 4. Illustrative computer simulations are shown in Section 5 and, finally, concluding remarks are made in Section 6.

2. SIGNAL MODEL

The discrete-time equivalent model of a MIMO transmission system with frequency-selective and time-varying CIR can be written as [6]

$$\mathbf{x}_t = \sum_{i=0}^{m-1} \mathbf{H}_{i,t}\mathbf{b}_{t-i} + \mathbf{u}_t, \quad t \in \mathbb{N}, \tag{1}$$

where $\{\mathbf{H}_{i,t}\}_{i=0}^{m-1}$ is the $L \times N$-dimensional CIR, of length m, $\mathbf{b}_t$ is the $N \times 1$ vector containing the symbols transmitted at time t, $\mathbf{u}_t \sim N(\mathbf{u}_t|\mathbf{0}, \sigma_u^2\mathbf{I}_L)$ is an additive white Gaussian noise (AWGN) process with zero mean and covariance matrix $\sigma_u^2\mathbf{I}_L$ ($\mathbf{I}_L$ is the $L \times L$ identity matrix) and $\mathbf{x}_t$ is the $L \times 1$ vector of observations. The symbols are modeled as discrete uniform random variables in the alphabet $\mathscr{B}$, hence $\mathbf{b}_t \sim \mathscr{U}(\mathscr{B}^N)$. It is often convenient to use a more compact representation of (1), namely

$$\mathbf{x}_t = \mathbf{H}_t\overline{\mathbf{b}}_t + \mathbf{u}_t, \tag{2}$$

where $\mathbf{H}_t = [\mathbf{H}_{m-1,t} \cdots \mathbf{H}_{0,t}]$ is the $L \times Nm$ overall channel matrix at time t and $\overline{\mathbf{b}}_t = [\mathbf{b}_{t-m+1}^\top \cdots \mathbf{b}_t^\top]^\top$ is an $Nm \times 1$ vec-

This work was supported by *Ministerio de Educación y Ciencia* of Spain (project TEC2004-06451-C05-01).

tor that contains all the symbols involved in the the t-th observation.

The channel variation can be modeled with an autoregressive (AR) process [8], that we assume of first order for simplicity (higher orders are easily handled, except for the notational involvement), specifically

$$\mathbf{H}_t = \gamma \mathbf{H}_{t-1} + \mathbf{V}_t, \tag{3}$$

where $1-\varepsilon < \gamma < 1$ (for small $\varepsilon > 0$) and $\mathbf{V}_t$ is a matrix of independent and identically distributed (i.i.d.) Gaussian random variables with zero mean and variance σ_v^2.

Because of the channel frequency-selectivity, some type of smoothing is needed for reliable data detection. The design of smoothing detectors becomes simpler if we stack together several successive observation vectors, to yield the model

$$\mathbf{x}_{t,a} = \mathbf{H}_{t,a}\mathbf{b}_{t,a} + \mathbf{u}_{t,a}, \tag{4}$$

where $a \geq 1$ is the smoothing lag, $\mathbf{x}_{t,a} = [\mathbf{x}_t^\top \cdots \mathbf{x}_{t+a}^\top]^\top$ is the $L(a+1)\times 1$ vector of stacked observations, $\mathbf{b}_{t,a} = [\mathbf{b}_{t-m+1}^\top \cdots \mathbf{b}_{t+a}^\top]^\top$ has dimensions $N(m+a)\times 1$, $\mathbf{u}_{t,a} = [\mathbf{u}_t^\top \cdots \mathbf{u}_{t+a}^\top]^\top$ and

$$\mathbf{H}_{t,a} = \begin{bmatrix} \mathbf{H}_t(m-1) & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{H}_t(m-2) & \mathbf{H}_{t+1}(m-1) & \cdots & \mathbf{0} \\ \vdots & \mathbf{H}_{t+1}(m-2) & \ddots & \vdots \\ \mathbf{H}_t(0) & \vdots & \ddots & \mathbf{H}_{t+d}(m-1) \\ \vdots & \mathbf{H}_{t+1}(0) & \ddots & \mathbf{H}_{t+d}(m-2) \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{H}_{t+d}(0) \end{bmatrix}^\top \tag{5}$$

is the $L(a+1)\times N(m+a)$ stacked channel matrix.

3. MIMO CHANNEL EQUALIZATION

3.1 Sequential Importance Sampling

Most particle filtering methods rely upon the principle of Importance Sampling (IS) [1] for building an empirical approximation of a desired pdf[1], say $p(x)$, by drawing samples from a different distribution, known as *importance function* or *proposal pdf* and denoted $q(x)$. These samples are then assigned appropriate normalized *importance* weights, i.e.,

$$x^{(i)} \sim q(x) \quad \text{and} \quad w^{(i)} \propto \frac{p(x^{(i)})}{q(x^{(i)})},$$

where $\sum_{i=1}^M w^{(i)} = 1$, M being the number of particles. In order to detect the transmitted symbols, it is natural to aim at the approximation of the *a posteriori* marginal pdf of the data, $p(\mathbf{b}_{0:t}|\mathbf{x}_{0:t})$, which contains all relevant statistical information for the optimal (Bayesian) estimation of $\mathbf{b}_{0:t}$. In turn, an importance function of the form $q(\mathbf{b}_{0:t}|\mathbf{x}_{0:t})$ is needed.

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

$$p(\mathbf{b}_{0:t}|\mathbf{x}_{0:t}) \propto p(\mathbf{x}_t|\mathbf{b}_{0:t},\mathbf{x}_{0:t-1})p(\mathbf{b}_{0:t-1}|\mathbf{x}_{0:t-1}), \tag{6}$$

which is easily derived by taking into account the *a priori* uniform distribution of the symbols, and an adequate importance function that can be factored as

$$q(\mathbf{b}_{0:t}|\mathbf{x}_{0:t}) = q(\mathbf{b}_t|\mathbf{b}_{0:t-1},\mathbf{x}_{0:t})q(\mathbf{b}_{0:t-1}|\mathbf{x}_{0:t-1}). \tag{7}$$

The recursive algorithm that combines the IS principle and decompositions (6) and (7) to build a discrete random measure that approximates the posterior pdf is called sequential importance sampling (SIS) [2, 3]. Let $\Omega_t = \left\{\mathbf{b}_{0:t}^{(i)}, w_t^{(i)}\right\}_{i=1}^M$ denote the discrete measure at time t, where M is the number of particles. The desired pdf is approximated as

$$\hat{p}(\mathbf{b}_{0:t}|\mathbf{x}_{0:t}) = \sum_{i=1}^M \delta_i(\mathbf{b}_{0:t}) w_t^{(i)}, \tag{8}$$

where $\delta_i(\mathbf{b}_t) = \delta(\mathbf{b}_t - \mathbf{b}_t^{(i)})$ is the Dirac delta function. When a new observation is collected at time $t+1$, the SIS algorithm proceeds through the following steps to recursively compute Ω_{t+1},

1. Importance sampling: $\mathbf{b}_{t+1}^{(i)} \sim q(\mathbf{b}_{t+1}|\mathbf{b}_{0:t}^{(i)},\mathbf{x}_{0:t+1})$.
2. Weight update: $\tilde{w}_{t+1}^{(i)} = w_t^{(i)} \frac{p(\mathbf{x}_{t+1}|\mathbf{b}_{0:t+1}^{(i)},\mathbf{x}_{0:t})}{q(\mathbf{b}_{t+1}^{(i)}|\mathbf{b}_{0:t}^{(i)},\mathbf{x}_{0:t+1})}$
3. Weight normalization: $w_t^{(i)} = \frac{\tilde{w}_t^{(i)}}{\sum_{k=1}^N \tilde{w}_t^{(k)}}$

It is straightforward to obtain data estimates from the approximate pdf $\hat{p}(\mathbf{b}_{0:t}|\mathbf{x}_{0:t})$. In particular, the marginal MAP symbol detector is

$$\hat{\mathbf{b}}_t^{map} = \arg\max_{\mathbf{b}_t} \left\{ \sum_{i=1}^M \delta(\mathbf{b}_t - \mathbf{b}_t^{(i)}) w_t^{(i)} \right\}, \tag{9}$$

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 few time steps most of the particles have importance weights with negligible values (very close to zero). 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. In this paper, we consider only the conceptually simplest resampling scheme, that generates a set of M new and equally weighted particles, $\{\mathbf{b}_{0:t}^{(i)}, 1/M\}_{i=1}^M$, by drawing from the discrete probability distribution $p_{rsp}(\mathbf{b}_{0:t}^{(i)}) = w_t^{(i)}$.

3.2 Optimal importance function

The performance of the SIS algorithm considerably depends on the choice of importance function. The optimal proposal pdf for the described scheme is

$$q(\mathbf{b}_t|\mathbf{x}_{0:t},\mathbf{b}_{0:t-1}) = p(\mathbf{b}_t|\mathbf{x}_{0:t},\mathbf{b}_{0:t-1}) \propto p(\mathbf{x}_t|\mathbf{b}_{0:t},\mathbf{x}_{0:t-1}), \tag{10}$$

which contains all the information available at time t for the sampling of $\mathbf{b}_t$.

[1] We will always use the term *density*, even for discrete random variables, since any probability mass function can be expressed as a density using sums of Dirac delta functions.

In order to sample the importance function of (10), it is necessary to evaluate the likelihood $p(\mathbf{x}_t|\mathbf{b}_{0:t},\mathbf{x}_{0:t-1})$ for each possible value of vector $\mathbf{b}_t$, normalize and then draw from the resulting discrete distribution. Unfortunately, there are $|\mathscr{B}|^N$ possibilities for $\mathbf{b}_t$ and the evaluation of each likelihood involves running one step of a Kalman filter (KF) which, in turn, has a computational complexity $\mathscr{O}(L^3)$ because of matrix inversions [9]. Moreover, when the MIMO channel is dispersive ($m > 1$) the equalizer performance is very poor, because decisions made at time t are not reliable. Further details can be found in [6].

3.3 Delayed Sampling

Detection in dispersive channels usually requires some smoothing (i.e., to exploit posterior observations, $\mathbf{x}_{0:t+a}$, where $0 < a \leq m-1$ is a smoothing lag) in order to detect $\mathbf{b}_t$. In the context of particle filtering, smoothing is also referred to as *delayed sampling* [9] because particle $\mathbf{b}_t^{(i)}$ cannot be drawn until $\mathbf{x}_{t+a}$ is observed.

The optimal smoothing importance pdf is

$$q(\mathbf{b}_t|\mathbf{b}_{0:t-1},\mathbf{x}_{0:t+a}) = p(\mathbf{b}_t|\mathbf{b}_{0:t-1},\mathbf{x}_{0:t+a}) \\ \propto \sum_{\tilde{\mathbf{b}}_{t+1:t+a}\in\mathscr{B}^{Na}} \prod_{k=0}^{a} p\left(\mathbf{x}_{t+k}|\mathbf{b}_{0:t},\tilde{\mathbf{b}}_{t+1:t+k},\mathbf{x}_{0:t+k-1}\right), \quad (11)$$

where the evaluation of each factor $p\left(\mathbf{x}_{t+k}|\mathbf{b}_{0:t},\tilde{\mathbf{b}}_{t+1:t+k},\mathbf{x}_{0:t+k-1}\right)$ requires a KF step. The weight update equation becomes [6]

$$w_{t+a}^{(i)} = w_{t+a-1}^{(i)} \sum_{\tilde{\mathbf{b}}_{t:t+a}} \prod_{k=0}^{a} p\left(\mathbf{x}_{t+k}|\mathbf{b}_{0:t-1}^{(i)},\tilde{\mathbf{b}}_{t:t+k},\mathbf{x}_{0:t+k-1}\right). \quad (12)$$

where $\tilde{\mathbf{b}}_{t:t+d} \in \mathscr{B}^{N(a+1)}$. Therefore, sampling and updating a single particle with this method involves the computation of $|\mathscr{B}|^{N(a+1)}$ likelihoods. The complexity of the algorithm, therefore, grows exponentially with the number of antennas and the smoothing lag, i.e., it is $\mathscr{O}(|\mathscr{B}|^{N(a+1)})$.

4. A NEW SMC-MAP EQUALIZER

The SMC equalizer based on the optimal delayed importance function is limited by its practically intractable computational complexity. In [6, 7], new SMC smoothing schemes were proposed that avoid the exponential growth of complexity, but they still require to run banks of Kalman filters, as well as additional matrix inversions, that yield an $\mathscr{O}\left((Lm)^3\right)$ load.

In order to drastically reduce this stringent computational requirements, we propose a scheme to directly approximate, using samples, the joint smoothing distribution of the sequence of symbol vectors and channel matrices given the observations, i.e., $p(\mathbf{b}_{0:t},\mathbf{H}_{0:t}|\mathbf{x}_{1:t+a})$, instead of the marginal *a posteriori* pdf of (6). Let us start with the joint posterior *filtering* pdf, that admits the straightforward decomposition

$$\begin{aligned} p(\mathbf{b}_{0:t+a},\mathbf{H}_{0:t+a}|\mathbf{x}_{1:t+a}) \quad &\propto \quad p(\mathbf{x}_{t:t+a}|\mathbf{b}_{t-m+1:t+a},\mathbf{H}_{t:t+a}) \\ &\times \quad p(\mathbf{H}_{t:t+a}|\mathbf{H}_{t-1}) \\ &\times \quad p(\mathbf{b}_{0:t-1},\mathbf{H}_{0:t-1}|\mathbf{x}_{1:t-1}). \quad (13) \end{aligned}$$

Notice that $p(\mathbf{x}_{t:t+a}|\mathbf{b}_{t-m+1:t+a},\mathbf{H}_{t:t+a})$ and $p(\mathbf{H}_{t:t+a}|\mathbf{H}_{t-1})$ are Gaussian pdf's, straightforward to evaluate without the need of Kalman filtering. Assume also a suitable proposal pdf of the form

$$q_t(\mathbf{b}_{t:t+a},\mathbf{H}_{t:t+a}|\mathbf{b}_{0:t-1}^{(i)},\mathbf{H}_{0:t-1}^{(i)},\mathbf{x}_{1:t+a}) \propto \\ q_t(\mathbf{b}_{t:t+a}|\mathbf{H}_{t:t+a})q_t(\mathbf{H}_{t:t+a}), \quad (14)$$

with $q_t(\mathbf{b}_{t:t+a}|\mathbf{H}_{t:t+a})$ and $q_t(\mathbf{H}_{t:t+a})$ to be specified later, and the weight update rule

$$w_{t+a}^{(i)} \propto w_{t+a-1}\frac{p(\mathbf{x}_{t:t+a}|\mathbf{b}_{t-m+1:t+a}^{(i)},\mathbf{H}_{t:t+a}^{(i)})}{q_t(\mathbf{b}_{t:t+a}^{(i)}|\mathbf{H}_{t:t+a}^{(i)})} \times\frac{p(\mathbf{H}_{t:t+a}^{(i)}|\mathbf{H}_{t-1}^{(i)})}{q_t(\mathbf{H}_{t:t+a}^{(i)}|\mathbf{H}_{t-1}^{(i)})}. \quad (15)$$

The use of (14) and (15) yields a new set of weighted particles, $\tilde{\Omega}_{t+a} = \left\{\left(\mathbf{b}_{0:t+a}^{(i)},\mathbf{H}_{0:t+a}^{(i)}\right),w_{t+a}^{(i)}\right\}_{i=1}^{M}$, and the approximation

$$p(\mathbf{b}_{0:t+a},\mathbf{H}_{0:t+a}|\mathbf{x}_{1:t+a}) \approx \sum_{i=1}^{M} w_{t+a}^{(i)}\delta_i(\mathbf{b}_{0:t+a})\delta_i(\mathbf{H}_{0:t+a}). \quad (16)$$

Integrating (16) over $\mathbf{b}_{t+1:t+a}$ and $\mathbf{H}_{t+1:t+a}$, yields an estimate of the desired joint smooting pdf,

$$\begin{aligned} p(\mathbf{b}_{0:t},\mathbf{H}_{0:t}|\mathbf{x}_{1:t+a}) \quad &\approx \quad \int\int \sum_{i=1}^{M} w_{t+a}^{(i)}\delta_i(\mathbf{b}_{0:t+a})\delta_i(\mathbf{H}_{0:t+a}) \\ &= \quad \sum_{i=1}^{M} w_{t+a}^{(i)}\delta_i(\mathbf{b}_{0:t})\delta_i(\mathbf{H}_{0:t}). \quad (17) \end{aligned}$$

Therefore, it is only necessary to keep the weighted particles up to time t, $\Omega_{t+a} = \left\{\left(\mathbf{b}_{0:t}^{(i)},\mathbf{H}_{0:t}^{(i)}\right),w_{t+a}^{(i)}\right\}_{i=1}^{M}$, and apply (14), (15) and (17) sequentially, with resampling steps when needed. Approximate MAP, smooth symbol estimates are computed as

$$\hat{\mathbf{b}}_t^{map} = \arg\max_{\mathbf{b}_t}\left\{\sum_{i=1}^{M}\delta(\mathbf{b}_t-\mathbf{b}_t^{(i)})w_{t+a}^{(i)}\right\}. \quad (18)$$

4.1 Channel Sampling Scheme

Although the weight update equation (15) enables us to circumvent the use of the KF banks in [9, 6, 7], we still need to design a proposal pdf q_t that avoids matrix inversions and any other "heavy" computations. With that aim, we propose to use a bank of (simple) adaptive channel estimators that play the same role as the KF, but with less stringent requirements. A similar idea was applied, for single-user spread spectrum systems, in [10]. In this paper, we will consider both least mean squares (LMS) and recursive least squares (RLS) [11] channel estimation algorithms, to be briefly described in Sections 4.1.1 and 4.1.2, respectively.

Independently of the channel estimation method, at time t there is an available channel estimate, $\hat{\mathbf{H}}_{t-1}^{(i)}$, for each $i \in \{1,\ldots,M\}$. Taking into account the AR model of the channel variation, we propose to draw $\mathbf{H}_t^{(i)}$ from a Gaussian proposal pdf with mean $\gamma\hat{\mathbf{H}}_{t-1}^{(i)}$ and diagonal covariance matrix

$\sigma_H^2 \mathbf{I}$, where σ_H^2 is a design parameter. The remaining channel samples, $\mathbf{H}_{t+1:t+a}^{(i)}$, are then drawn using the AR model. In summary,

$$\mathbf{H}_{t:t+a}^{(i)} \sim q_t(\mathbf{H}_{t:t+a}) = N(\mathbf{H}_t|\gamma\hat{\mathbf{H}}_{t-1}^{(i)}, \sigma_H^2\mathbf{I}) \times \prod_{k=1}^{a} N(\mathbf{H}_{t+k}|\gamma\mathbf{H}_{t+k-1}, \sigma_v^2\mathbf{I}). \quad (19)$$

Given $\mathbf{H}_{t:t+a}^{(i)}$, we can draw the new symbol vector $\mathbf{b}_t^{(i)}$ (see Section 4.2 for details) and then update the bank of adaptive channel estimators, to yield $\hat{\mathbf{H}}_t^{(i)}$, $i = 1, \ldots, M$.

4.1.1 LMS Channel Estimation

Consider the minimum mean square error (MMSE) estimation of the channel given $\mathbf{b}_{0:t}^{(i)}$, $i \in \{1, \ldots, M\}$, and $\mathbf{x}_{1:t}$, i.e.,

$$\hat{\mathbf{H}}_t^{(i)} = \arg\min_{\mathbf{H}_t} \left\{ E\left[\left\| \mathbf{x}_t - \mathbf{H}_t \bar{\mathbf{b}}_t^{(i)} \right\|^2 \right] \right\}. \quad (20)$$

The simplest way to adaptively solve (20) is the LMS algorithm [11], which takes the form

$$\hat{\mathbf{H}}_t^{(i)} = \hat{\mathbf{H}}_{t-1}^{(i)} - \mu \left(\hat{\mathbf{H}}_{t-1}^{(i)} \bar{\mathbf{b}}_t^{(i)} - \mathbf{x}_t \right) \bar{\mathbf{b}}_t^{(i)H}, \quad (21)$$

where $\mu << 1$ is a step-size parameter.

4.1.2 RLS Channel Estimation

The LMS algorithm (21) has linear computational complexity, but it usually exhibits a slow convergence rate and poor tracking capabilities. To avoid these well-known drawbacks it is convenient to use the exponentially-weighted RLS algorithm [11]. The channel estimator is

$$\hat{\mathbf{H}}_t^{(i)} = \arg\min_{\mathbf{H}} \left\{ \sum_{k=0}^{t} \lambda^{t-k} \left\| \mathbf{x}_k - \mathbf{H}\bar{\mathbf{b}}_k \right\|^2 \right\}, \quad (22)$$

where $0 < \lambda < 1$ is a forgetting factor. The sequence of problems defined by (22) are recursively solved using the RLS algorithm, summarized in the following equations

$$\mathbf{R}_0^{(i)^{-1}} \propto \mathbf{I}_{Nm} \quad \text{(initialization)} \quad (23)$$

$$\mathbf{g}_t^{(i)} = \frac{\lambda^{-1}\mathbf{R}_{t-1}^{(i)^{-1}} \bar{\mathbf{b}}_t^{(i)}}{1 + \lambda^{-1} \bar{\mathbf{b}}_t^{(i)^H} \mathbf{R}_{t-1}^{(i)^{-1}} \bar{\mathbf{b}}_t^{(i)}} \quad (24)$$

$$\hat{\mathbf{H}}_t^{(i)^H} = \hat{\mathbf{H}}_{t-1}^{(i)^H} + \mathbf{g}_t^{(i)} \left(\mathbf{x}_t^H - \bar{\mathbf{b}}_t^{(i)^H} \hat{\mathbf{H}}_{t-1}^{(i)^H} \right) \quad (25)$$

$$\mathbf{R}_t^{(i)^{-1}} = \lambda^{-1} \left(\mathbf{I}_{Nm} - \mathbf{g}_t^{(i)} \bar{\mathbf{b}}_t^{(i)^H} \right) \mathbf{R}_{t-1}^{(i)^{-1}} \quad (26)$$

The complexity of the resulting equalizer is linear with respect to LNm. Since, normally, $L \approx N$, it becomes $\mathcal{O}(N^2)$.

4.2 Data Sampling Scheme

When the channel sample, $\mathbf{H}_{t:t+d}^{(i)}$, $i \in \{1, \ldots, M\}$, is available, we build a data proposal pdf based on linear MMSE detection, as suggested in [7], but avoiding the computation of inverse matrices. In particular, we exploit the matrix inversion lemma [11] to recursively approximate the inverse of the autocorrelation matrix

$$\mathbf{R}_{t,x}^{-1} = \left(\sum_{n=0}^{t} \alpha^{t-n} \mathbf{x}_{t,a}\mathbf{x}_{t,a}^H \right)^{-1} \quad (27)$$

as

$$\hat{\mathbf{R}}_{0,x}^{-1} \propto \mathbf{I}_{L(a+1)}, \quad \hat{\mathbf{R}}_{t,x}^{-1} = \alpha^{-1} \left(\mathbf{I}_{L(a+1)} - \mathbf{g}_{t,a}\mathbf{x}_{t,a}^H \right) \hat{\mathbf{R}}_{t-1,x}^{-1}, \quad (28)$$

where $0 < \alpha < 1$ is a forgetting factor and $\mathbf{g}_{t,a} = \frac{\alpha^{-1}\hat{\mathbf{R}}_{t-1}^{-1}\mathbf{x}_{t,a}}{1+\alpha^{-1}\mathbf{x}_{t,a}^H\hat{\mathbf{R}}_{t-1}^{-1}\mathbf{x}_{t,a}}$ is a gain vector. A bank of $N(a+1)$ MMSE linear filters is built,

$$\mathbf{W}_{t,a}^{(i)} = \sigma_b^2 \hat{\mathbf{R}}_{t,x}^{-1} \mathbf{H}_{t,a}^{(i)} \mathbf{E}, \quad (29)$$

for $i = 1, \ldots, M$, where $\mathbf{W}_{t,a}^{(i)}$ has dimensions $L(a+1) \times N(a+1)$, σ_b^2 is the symbol power and $\mathbf{E} = \begin{bmatrix} \mathbf{0}_{N(m-1)\times N(a+1)} \\ \mathbf{I}_{N(a+1)} \end{bmatrix}$, and $N(a+1)$ soft data estimates are computed,

$$\mathbf{y}_{t,a}^{(i)} = \mathbf{W}_{t,a}^{(i)^H} \mathbf{x}_{t,a}. \quad (30)$$

Let $y_{j,t,a}^{(i)}$ denote the j-th element in $\mathbf{y}_{t,a}^{(i)}$, let $b_{l,t}$ be the l-th symbol in $\mathbf{b}_t$ and let $j = Nk + q$ for integers $k, q \geq 0$. Then, $y_{j,t,a}^{(i)}$ is an estimate of $b_{q,t+k}$. If the symbols are binary, $b_{q,t} \in \{\pm 1\}$ (extension to higher order constellations is straightforward), we can assign probabilities $q_{+1,q,t}^{(i)} \propto \exp\{-\frac{1}{\sigma_y^2}|y_{j,q,t} - 1|^2$ (where σ_y^2 is a design parameter) and $q_{-1,q,t}^{(i)} = 1 - q_{+1,q,t}^{(i)}$, and draw a sample $b_{q,t}^{(i)}$ accordingly. Repeating this process for all symbols from time t to time $t+a$ we obtain the desired sample $\mathbf{b}_{t:t+a}^{(i)}$. The evaluation of $q_t(\mathbf{b}_{t:t+a}^{(i)}|\mathbf{H}_{t:t+a}^{(i)})$ is carried out by adequately multiplying the probabilities $q_{\pm 1,q,t}^{(i)}$ for $q = 1, \ldots, N$ and $t, \ldots, t+a$.

5. SIMULATIONS

In order to numerically illustrate the performance of the proposed SMC equalizers, we have considered a simple system with $N = 2$ transmitting antennas and $L = 3$ receiving antennas. The length of the CIR is $m = 2$, and the parameters of the channel AR process are $\gamma = 1 - 10^{-5}$ and $\sigma_v^2 = 10^{-4}$. We have also assumed a BPSK modulation format, thus $\mathscr{B} = \{\pm 1\}$, and each channel use consists of the transmission of a frame of 300 symbol vectors (i.e., 600 binary symbols overall), including a training sequence of 30 symbol vectors which are used to obtain an initial (rough) estimate of the CIR.

Within this simulation setup, we have compared the optimal smoothing SMC equalizer described in Section 3.3 (labeled "D-SIS opt") with the two low-complexity SMC smoothers proposed in this paper (labeled "LMS-D-SIS" and "RLS-D-SIS", depending on the type of adaptive channel estimator used). The smoothing lag is $a = 1$ for the three equalizers.

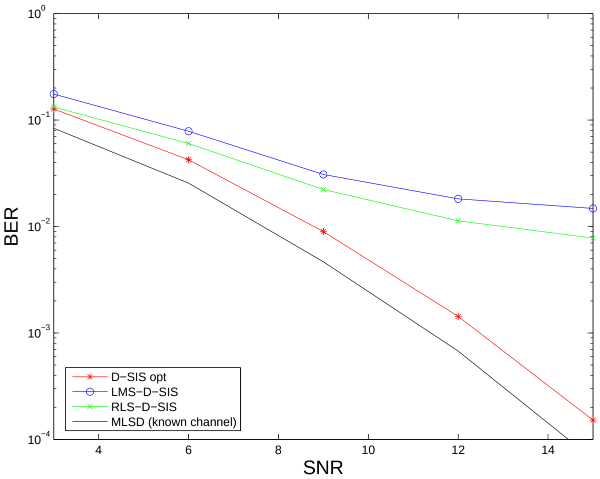

Figure 1: BER of the SMC equalizers and the MLSD for several values of the SNR (M=30 particles).

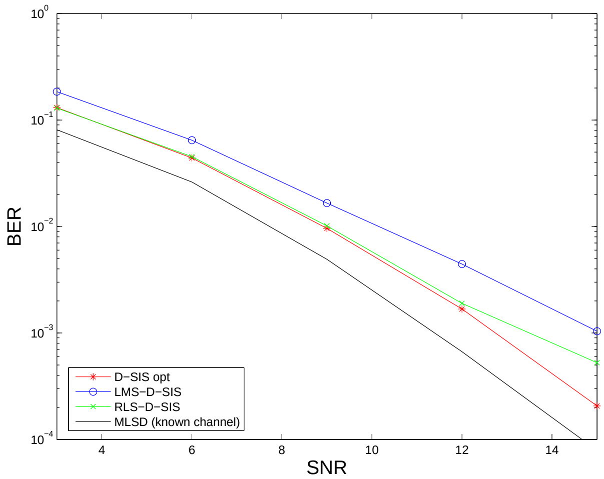

Figure 2: BER of the SMC equalizers and the MLSD for several values of the SNR (M=100 particles).

Figure 1 depicts the Bit Error Rate (BER) of the different algorithms for several SNR values, when the number of particles of the SMC equalizers is $M = 30$. The curve labeled "MLSD" shows the performance of the maximum likelihood sequence detector implemented via the Viterbi algorithm, with perfect knowledge of the time-varying CIR, and serves as a reference for performance. It can be seen that the optimal smoothing SMC algorithm clearly outperforms the proposed low-complexity equalizers. Indeed, the optimal delayed sampling scheme is much more efficient than the proposed techniques, although this is at the expense of an unaffordable computational load.

Nevertheless, the performance gap considerably reduces when we increase the number of particles. This is shown in Figure 2, where we can observe the BER curves obtained for $M = 100$, all other simulation parameters being the same. The RLS-D-SIS equalizer attains practically optimal performance up to SNR=12 dB, and suffers only a minor loss at SNR=15 dB. The LMS-D-SIS equalizer has an approximately constant loss of ≈ 2 dB for the whole range of SNR values, but has the advantage of its greater simplicity.

6. CONCLUSIONS

Existing particle filtering methods for MIMO channel equalization suffer from severe limitations because of their high computational requirements. In this paper we have introduced two low complexity sampling schemes that achieve an appealing trade-off between performance and complexity and are specially suitable for systems that operate under stringent time schedules. Moreover, our computer simulations show that for a sufficiently large number of particles, the proposed equalizers can achieve a nearly optimal performance.

REFERENCES

[1] J. S. Liu and R. Chen, "Sequential Monte Carlo methods for dynamic systems," *Journal of the American Statistical Association*, vol. 93, no. 443, pp. 1032–1044, September 1998.

[2] A. Doucet, N. de Freitas, and N. Gordon, Eds., *Sequential Monte Carlo Methods in Practice*, Springer, New York (USA), 2001.

[3] P. M. Djurić, J. H. Kotecha, J. Zhang, Y. Huang, T. Ghirmai, Mónica F. Bugallo, and J. Míguez, "Particle filtering," *IEEE Signal Processing Magazine*, vol. 20, no. 5, pp. 19–38, September 2003.

[4] G. J. Foschini, "Layered space-time arquitechture for wireless communications in a fading environment when using multi-element anntenas," *Bell Labs Technical Journal*, vol. 1, no. 2, pp. 41–59, Autumn 1996.

[5] G. G. Raleigh and J. M. Cioffi, "Spatio-temporal coding for wireless communication," *IEEE Transactions Communications*, vol. 46, no. 3, pp. 357–366, March 1998.

[6] Manuel A. Vázquez López and Joaquín Míguez, "A complexity-constrained particle filtering algorithm for MAP equalization of frequency-selective MIMO channels," in *Proceedings of the IEEE ICASSP*, May 2005.

[7] Manuel A. Vázquez López, Mónica F. Bugallo, and Joaquín Míguez, "Novel SMC techniques for blind equalization of flat-fading MIMO channels," in *Proceedings of the IEEE VTC*, June 2005.

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

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

[10] T. Bertozzi, D. Le Ruyet, C. Panazio, and H. Vu-Thien, "Channel tracking using particle filtering in unresolvable multipath environments," *EURASIP Journal of Applied Signal Processing*, vol. 2004, pp. 2328–2338, November 2004.

[11] S. Haykin, *Adaptive Filter Theory, 4th Edition*, Prentice Hall, Information and System Sciences Series, 2001.