SEQUENTIAL MAP EQUALIZATION OF MIMO CHANNELS WITH UNKNOWN ORDER

Manuel A. Vázquez

Departamento de Electrónica e Sistemas
Universidade da Coruña
Campus de Elviña s/n, 15071
A Coruña (Spain).
mvazquez@udc.es

Joaquín Míguez

Depto. de Teoría de la Señal y Comunicaciones
Universidad Carlos III de Madrid
Avenida de la Universidad 30, 28911
Leganés, Madrid (Spain).
joaquin.miguez@uc3m.es

ABSTRACT

Practical equalization of multiple input multiple output (MIMO) channels poses several difficulties. Namely, it is well known that the complexity of maximum *a posteriori* (MAP) data detection grows exponentially with the number of inputs and the channel order, i.e., the length of the channel impulse response (CIR). Moreover, knowledge of the latter parameter is needed for reliable data detection, but its estimation is often a hard task and very few papers have tackled the problem. In this article, we propose the use of the sequential Monte Carlo (SMC) methodology to build quasi-MAP MIMO equalizers with polynomial complexity, that admit a parallel implementation and can handle the uncertainty in the channel order. In particular, we derive both optimal and complexity-constrained SMC algorithms for joint data detection, channel order and CIR estimation in frequency and time-selective MIMO channels. Computer simulation results are presented to illustrate the performance of the proposed techniques.

1. INTRODUCTION

The fact that the capacity of a wireless channel grows linearly with the minimum number of transmitting and receiving elements has attracted much attention on the multiple input multiple output (MIMO) channels that appear naturally in many common scenarios, such as multiuser and multiantenna systems. However, practical equalization of MIMO channels poses several difficulties. Namely, it is well known that the complexity of maximum *a posteriori* (MAP) data detection grows exponentially with the number of inputs and the channel order, i.e., the length of the channel impulse response (CIR). Moreover, knowledge of the latter parameter is needed for reliable data detection, but its estimation is often a hard task and relatively little work can be found in the literature. In [1], the conditional model order estimator (CME) criterion proposed by Kay [2] is used for estimating the MIMO channel order in a space-time coded system. CME is less restrictive, and has been shown to attain better performance, than the minimum description length (MDL) technique [2]. The mehod in [1] is based on the assumption that the CIR is fixed for the duration of a complete frame and the processing of the block of available observations is carried out off-line, in batch mode. Alternatively, a subspace-based technique for blind channel order estimation is derived in [3], but the channel in that work is also assumed static and the order estimator requires the batch processing of a large block of observations.

In this paper, we investigate a completely different approach to the problem, based on the sequential Monte Carlo (SMC) methodology, also known as particle filtering (PF) [4, 5]. SMC algorithms are simulation-based techniques for sequential and adaptive signal processing 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. These methods explore the space of the SOI by generating random samples (termed particles) from a proposal distribution. The particles are then assigned proper weights, which are recursively computed, and yield the discrete approximation of the *a posteriori* pdf.

The main advantage of the SMC methodology is its generality, which enables its application to numerous problems in the field of communications, including the adaptive equalization of MIMO channels [6]. In a previous work [7], we proposed the use of the SMC framework to build quasi-MAP MIMO adaptive equalizers with polynomial complexity and amenable to parallel implementation. Here, we extend these techniques to handle the uncertainty in the channel order. In particular, we derive both optimal and complexity-constrained SMC algorithms for the joint estimation of the transmitted data, the channel order and

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

This work was supported by *Comunidad Autónoma de Madrid* (project PRO-MULTIDIS-CM, ref. S0505/TIC/0223), *Ministerio de Educación y Ciencia* (projects DOIRAS, ref. TIC2003-02602, and MONIN, ref. TEC2006-13514-C02-01) and the EC (Network of Excellence 'CRUISE', ref. IST-4-027738).

the CIR estimation in a frequency and time-selective MIMO system.

The remaining of this paper is organized as follows. In Section 2, the discrete-time signal model of a MIMO transmission system with frequency and time-selective channel is described. The optimal adaptive MIMO equalizer that jointly estimates the transmitted symbols and the MIMO CIR, including its order, is introduced in Section 3. Since its computational load makes the latter technique impractical, complexity-constrained methods are proposed in Section 4. Computer simulation results are shown in Section 5 and, finally, Section 6 is devoted to the conclusions.

2. SIGNAL MODEL

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

$$\mathbf{x}_t = \sum_{i=0}^{m-1} \mathbf{H}_{i,t}\mathbf{s}_{t-i} + \mathbf{u}_t = \mathbf{H}_t\bar{\mathbf{s}}_t + \mathbf{u}_t, \tag{1}$$

where

- $\{\mathbf{H}_{i,t}\}_{i=0}^{m-1}$ is the $L \times N$-dimensional CIR, of length m;
- the $N \times 1$ input vector $\mathbf{s}_t = [s_{1,t}, \ldots, s_{N,t}]^\top$ contains the N symbols transmitted at time t, which are modeled as discrete uniform random variables with finite alphabet $\mathcal{S}$;
- $\mathbf{x}_t$ is the $L \times 1$ vector of observations;
- $\mathbf{H}_t = [\mathbf{H}_{m-1,t} \cdots \mathbf{H}_{0,t}]$ is an alternative $L \times Nm$ matrix representing the CIR;
- $\bar{\mathbf{s}}_t = [\mathbf{s}_{t-m+1}^\top \cdots \mathbf{s}_t^\top]^\top$ is an $Nm \times 1$ vector that includes all the symbols involved in the t-th observation;
- and $\mathbf{u}_t$ is an additive white Gaussian noise (AWGN) process with zero mean and covariance matrix $\sigma_u^2\mathbf{I}_L$ ($\mathbf{I}_L$ denotes the $L \times L$ identity matrix).

The CIR length, m, is modeled as a discrete random variable with uniform *a priori* probability distribution in a known finite set $\mathcal{M} = \{1, 2, \ldots, |\mathcal{M}|\}$. We assume a first-order auto-regressive (AR) model for the channel evolution [8], i.e.,

$$\mathbf{H}_t = \gamma\mathbf{H}_{t-1} + \mathbf{V}_t, \tag{2}$$

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

Because of the channel frequency-selectivity, some type of smoothing is needed for reliable data detection. In that case, it is useful to consider the stacked model

$$\mathbf{x}_{t,a} = \mathbf{H}_{t,a}\mathbf{s}_{t,a} + \mathbf{u}_{t,a}, \tag{3}$$

where $1 \leq a < m$ 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, $\mathbf{s}_{t,a} = [\mathbf{s}_{t-m+1}^\top \cdots \mathbf{s}_{t+a}^\top]^\top$ has dimensions $N(m+a) \times 1$, $\mathbf{u}_{t,a} = [\mathbf{u}_t^\top \cdots \mathbf{u}_{t+a}^\top]^\top$ and

$$\mathbf{H}_{t,a} = \begin{bmatrix} \mathbf{H}_t(m-1) & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{H}_t(m-2) & \mathbf{H}_{t+1}(m-1) & \cdots & \mathbf{0} \\ \vdots & \mathbf{H}_{t+1}(m-2) & \ddots & \vdots \\ \mathbf{H}_t(0) & \vdots & \ddots & \mathbf{H}_{t+a}(m-1) \\ \vdots & \mathbf{H}_{t+1}(0) & \ddots & \mathbf{H}_{t+a}(m-2) \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{H}_{t+a}(0) \end{bmatrix}^\top \tag{4}$$

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

3. SMC EQUALIZERS FOR UNKNOWN CHANNEL ORDER

3.1. Sequential Importance Sampling

Most SMC methods rely upon the principle of Importance Sampling (IS) 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 M is the number of samples, usually termed *particles*, $i = 1, \ldots, M$ and the weight normalization implies that $\sum_{i=1}^{M} w^{(i)} = 1$.

In order to detect the transmitted symbols, it is natural to aim at the approximation of the *a posteriori* marginal pdf of the data, $p(\mathbf{s}_{0:t}|\mathbf{x}_{0:t})$, which contains all relevant statistical information for the optimal (Bayesian) estimation of $\mathbf{s}_{0:t}$. Therefore, the application of the IS principle requires to choose an importance function of the form $q(\mathbf{s}_{0:t}|\mathbf{x}_{0:t})$, for which the dimension of the argument grows with time. Fortunately, one of the most appealing features of the SMC approach is its potential for online processing. Indeed, the

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

IS principle can be sequentially applied by exploiting the recursive decomposition of the posterior pdf

$$p(\mathbf{s}_{0:t}|\mathbf{x}_{0:t}) \propto p(\mathbf{x}_t|\mathbf{s}_{0:t}, \mathbf{x}_{0:t-1})p(\mathbf{s}_{0:t-1}|\mathbf{x}_{0:t-1}), \quad (5)$$

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

$$q(\mathbf{s}_{0:t}|\mathbf{x}_{0:t}) = q(\mathbf{s}_t|\mathbf{s}_{0:t-1}, \mathbf{x}_{0:t})q(\mathbf{s}_{0:t-1}|\mathbf{x}_{0:t-1}). \quad (6)$$

The recursive algorithm that combines the IS principle and decompositions (5) and (6) to build a discrete random measure that approximates the posterior pdf is called sequential importance sampling (SIS) [4, 5]. To be specific, let $\Omega_t = \left\{\mathbf{s}_{0:t}^{(i)}, w_t^{(i)}\right\}_{i=1}^{M}$ denote the set of weighted particles at time t. 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: $\mathbf{s}_{t+1}^{(i)} \sim q(\mathbf{s}_{t+1}|\mathbf{s}_{0:t}^{(i)}, \mathbf{x}_{0:t+1})$.
2. Weight update: $\tilde{w}_{t+1}^{(i)} = w_t^{(i)} \frac{p(\mathbf{x}_{t+1}|\mathbf{s}_{0:t+1}^{(i)}, \mathbf{x}_{0:t})}{q(\mathbf{s}_{t+1}^{(i)}|\mathbf{s}_{0:t}^{(i)}, \mathbf{x}_{0:t+1})}$
3. Weight normalization: $w_t^{(i)} = \frac{\tilde{w}_{t+1}^{(i)}}{\sum_{k=1}^{N} \tilde{w}_{t+1}^{(k)}}$

It has been shown in [7] that, assuming a known channel order $\tilde{m} \in \mathcal{M}$, the likelihood in the weight update step, $p(\mathbf{x}_{t+1}|\mathbf{s}_{0:t+1}^{(i)}, \mathbf{x}_{0:t})$, can be obtained analytically using the Kalman filter to integrate out the CIR.

Given Ω_t, it is straightforward to obtain a point-mass approximation of the *a posteriori* pdf, namely

$$p_M(\mathbf{s}_{0:t}|\mathbf{x}_{0:t}) = \sum_{i=1}^{M} \delta(\mathbf{s}_{0:t} - \mathbf{s}_{0:t}^{(i)})w_t^{(i)}, \quad (7)$$

where δ is the Dirac delta function, and estimators can be derived from $p_M(\mathbf{s}_{0:t}|\mathbf{x}_{0:t})$ easily. In particular, the marginal MAP symbol detector is

$$\hat{\mathbf{s}}_t^{map} = \arg\max_{\mathbf{s}_t} \left\{ \sum_{i=1}^{M} \delta(\mathbf{s}_t - \mathbf{s}_t^{(i)})w_t^{(i)} \right\}, \quad (8)$$

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

One well-known 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) [5]. 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, $\{\mathbf{s}_{0:t}^{(i)}, 1/M\}_{i=1}^{M}$, by drawing from the discrete probability distribution $p_{rsp}(\mathbf{s}_{0:t}^{(i)}) = w_t^{(i)}$. However, more efficient methods have been proposed in the literature and they can be directly applied to our problem, including those specifically designed to enable implementation with parallel-processing architectures [9].

3.2. Optimal Smoothing Equalizer with Unknown Order

We aim at approximating the posterior pdf of the data, $p(\mathbf{s}_{0:t}|\mathbf{x}_{0:t+a})$, where a is a smoothing lag. The channel order is a discrete random variable with support in the finite set $\mathcal{M}$, as described in Section 2. So as to account for all possible orders, we assume $a = |\mathcal{M}| - 1$, which guarantees that all observations vectors that involve the symbols in $\mathbf{s}_t$ are taken together for smoothing. It turns out that the desired pdf can be analytically calculated by "summing out" the random channel order m and the unknown symbols $\mathbf{s}_{t+1:t+a}$,

$$\begin{aligned} &p(\mathbf{s}_{0:t}|\mathbf{x}_{0:t+a}) \propto \\ &\sum_{m\in\mathcal{M}} \sum_{\mathbf{s}_{t+1:t+a}} p(\mathbf{x}_{t:t+a}|m, \mathbf{s}_{0:t+a}, \mathbf{x}_{0:t-1})\times \\ &\quad p(m|\mathbf{s}_{0:t-1}, \mathbf{x}_{0:t-1})p(\mathbf{s}_{0:t-1}|\mathbf{x}_{0:t-1}). \end{aligned} \quad (9)$$

The posterior pmf of the channel order at time $t-1$, $p(m|\mathbf{s}_{0:t-1}, \mathbf{x}_{0:t-1})$, can be computed recursively, up to a proportionality constant, using Bayes' theorem,

$$\begin{aligned} p(m|\mathbf{s}_{0:t-1}, \mathbf{x}_{0:t-1}) &\propto p(\mathbf{x}_{t-1}|m, \mathbf{s}_{0:t-1}, \mathbf{x}_{0:t-2}) \times \\ &\quad \times p(m|\mathbf{s}_{0:t-2}, \mathbf{x}_{0:t-2}) \\ &= p(m) \prod_{k=0}^{t-1} p(\mathbf{x}_k|m, \mathbf{s}_{0:k}, \mathbf{x}_{0:k-1}), \end{aligned} \quad (10)$$

where $p(m) = \frac{1}{|\mathcal{M}|}$ is the *a priori* pdf of m.

Decomposition in equation (9) enables the application of the SIS algorithm with the optimal importance function, which is summarized by the two steps given by (11) and (12) (see top of next page). The former defines the proposal function, from which symbol vector $\mathbf{s}_t$ is drawn, and the latter is the weight update equation.

The computation of the likelihoods in equations (9) and (10) is carried out using a bank of $|\mathcal{M}|$ Kalman filters (one per *a priori* possible channel order). Moreover, from (12) it is seen that the algorithm demands the computation of $|\mathcal{M}||\mathcal{S}|^{Na}$ likelihoods per particle, which yields a computational load that cannot be afforded in most practical cases.

4. A COMPLEXITY-CONSTRAINED ALTERNATIVE

In order to design SMC methods capable of handling an unknown channel order with a tractable complexity, we have

$$\mathbf{s}_t^{(i)} \sim q(\mathbf{s}_t|\mathbf{x}_{0:t+a}) = \frac{\sum_{m\in\mathcal{M}}\sum_{\tilde{\mathbf{s}}_{t+1:t+a}} p(\mathbf{x}_{t:t+a}|m, \mathbf{s}_{0:t-1}^{(i)}, \tilde{\mathbf{s}}_{t+1:t+a}, \mathbf{x}_{0:t-1}) p(m|\mathbf{s}_{0:t-1}^{(i)}, \mathbf{x}_{0:t-1})}{\sum_{\mathbf{s}_t}\sum_{m\in\mathcal{M}} p(m|\mathbf{s}_{0:t-1}^{(i)}, \mathbf{x}_{0:t-1}) \sum_{\tilde{\mathbf{s}}_{t+1:t+a}} p(\mathbf{x}_{t:t+a}|m, \mathbf{s}_{0:t-1}^{(i)}, \tilde{\mathbf{s}}_{t:t+a}, \mathbf{x}_{0:t-1})} \tag{11}$$

$$w_{t+a}^{(i)} = w_{t+a-1}^{(i)} \sum_{\mathbf{s}_t}\sum_{m\in\mathcal{M}} p(m|\mathbf{s}_{0:t-1}^{(i)}, \mathbf{x}_{0:t-1}) \sum_{\tilde{\mathbf{s}}_{t+1:t+a}} p(\mathbf{x}_{t:t+a}|m, \mathbf{s}_{0:t-1}^{(i)}, \tilde{\mathbf{s}}_{t:t+a}, \mathbf{x}_{0:t-1}) \tag{12}$$

extended the technique in [7] to incorporate and exploit the information of the *a posteriori* pdf of m. In particular, we aim at the approximation of the joint smoothing pdf of the sequence of symbol vectors and a set of sequences of channel matrices indexed by the elements of $\mathcal{M}$, i.e., $p(\mathbf{s}_{0:t+a}, \{\mathbf{H}_{m;0:t+a}\}_{m\in\mathcal{M}}|\mathbf{x}_{0:t+a})$ where $\mathbf{H}_{m;0:t+a}$, $m = 1,\ldots,|\mathcal{M}|$, is the sequence of $L\times Nm$ dimensional channel matrices obtained by assuming each possible channel order in $\mathcal{M}$. Using this notation, the desired pdf is decomposed as

$$\begin{gathered} p(\mathbf{s}_{0:t+a}, \{\mathbf{H}_{m;0:t+a}\}_{m\in\mathcal{M}}|\mathbf{x}_{0:t+a}) \propto \\ \sum_m p(\mathbf{x}_{t:t+a}|m, \mathbf{s}_{0:t+a}, \mathbf{H}_{m;0:t+a}, \mathbf{x}_{0:t-1})\times \\ p(m|\mathbf{s}_{0:t-1}, \mathbf{H}_{m;0:t-1}, \mathbf{x}_{0:t-1})\times \\ p(\{\mathbf{H}_{m;t:t+a}\}_{m\in\mathcal{M}}|\{\mathbf{H}_{m;t-1}\}_{m\in\mathcal{M}})\times \\ p\left(\mathbf{s}_{0:t-1}, \{\mathbf{H}_{m;0:t-1}\}_{m\in\mathcal{M}}|\mathbf{x}_{0:t-1}\right) \end{gathered} \tag{13}$$

where the term $p(m|\mathbf{s}_{0:t-1}, \mathbf{H}_{m;0:t-1}, \mathbf{x}_{0:t-1})$ can be computed recursively just like in (10), but whithout the need of Kalman filtering (due to the conditioning on the channel sequence $\mathbf{H}_{m;0:t-1}$), i.e.,

$$\begin{gathered} p(m|\mathbf{s}_{0:t-1}, \mathbf{H}_{m;0:t-1}, \mathbf{x}_{0:t-1}) \propto \\ p(m|\mathbf{s}_{0:t-2}, \mathbf{H}_{m;0:t-2}, \mathbf{x}_{0:t-2}) p(\mathbf{x}_{t-1}|m, \mathbf{s}_{t-1}, \mathbf{H}_{m;t-1}) \end{gathered} \tag{14}$$

where the likelihood function is Gaussian with mean $\mathbf{H}_{m;t-1}\mathbf{s}_{t-1}$ an covariance matrix $\sigma_u^2\mathbf{I}_L$.

Our goal is to take advantage of our ability to compute the channel order posterior pdf (14) so as to design an efficient importance function to draw both from the channel and the data processes. We propose a general scheme of the form

$$\begin{gathered} q(\mathbf{s}_{0:t+a}, \{\mathbf{H}_{m;0:t+a}\}_{m\in\mathcal{M}}|\mathbf{x}_{0:t+a}) = \\ q(\mathbf{s}_{t:t+a}|\mathbf{s}_{0:t-1}, \{\mathbf{H}_{m;0:t+a}\}_{m\in\mathcal{M}}, \mathbf{x}_{0:t+a}) \\ q\left(\{\mathbf{H}_{m;t:t+a}\}_{m\in\mathcal{M}}|\mathbf{s}_{0:t-1}, \{\mathbf{H}_{m;0:t-1}\}_{m\in\mathcal{M}}, \mathbf{x}_{0:t+a}\right) \\ q(\mathbf{s}_{0:t-1}, \{\mathbf{H}_{m;0:t-1}\}_{m\in\mathcal{M}}|\mathbf{x}_{0:t+a}) \end{gathered} \tag{15}$$

with $q(\mathbf{s}_{t:t+a}|\mathbf{s}_{0:t-1}, \{\mathbf{H}_{m;0:t+a}\}_{m\in\mathcal{M}}, \mathbf{x}_{0:t+a})$ (the data importance function), and $q(\{\mathbf{H}_{m;t:t+a}\}_{m\in\mathcal{M}}|\mathbf{s}_{0:t-1}, \{\mathbf{H}_{m;0:t-1}\}_{m\in\mathcal{M}}, \mathbf{x}_{0:t+a})$ (the channel importance function) to be specifically defined later. The general form of the weight update equation given (13) and (15) is

$$\begin{gathered} w_{t+a}^{(i)} \propto w_{t+a-1}^{(i)}\times \\ \frac{\sum_m p(\mathbf{x}_{t:t+a}|m, \mathbf{s}_{0:t+a}, \mathbf{H}_{m;0:t+a}, \mathbf{x}_{0:t-1})}{q(\mathbf{s}_{t:t+a}|\mathbf{s}_{0:t-1}, \{\mathbf{H}_{m;0:t+a}\}_{m\in\mathcal{M}}, \mathbf{x}_{0:t+a})}\times \\ \frac{p(m|\mathbf{s}_{0:t-1}, \mathbf{H}_{m;0:t-1}, \mathbf{x}_{0:t-1}) p\left(\{\mathbf{H}_{m;t:t+a}\}_{m\in\mathcal{M}}|\{\mathbf{H}_{m;t-1}\}_{m\in\mathcal{M}}\right)}{q\left(\{\mathbf{H}_{m;t:t+a}\}_{m\in\mathcal{M}}|\mathbf{s}_{0:t-1}, \{\mathbf{H}_{m;0:t-1}\}_{m\in\mathcal{M}}, \mathbf{x}_{0:t+a}\right)} \end{gathered} \tag{16}$$

Drawing from (15) and applying (16) yields a new set of weighted particles, $\tilde{\Omega}_{t+a} = \left\{\left(\mathbf{s}_{0:t+a}^{(i)}, \left\{\mathbf{H}_{m;0:t+a}^{(i)}\right\}_{m\in\mathcal{M}}\right), w_{t+a}^{(i)}\right\}_{i=1}^M$, and the approximation of the corresponding posterior pdf

$$\begin{gathered} p(\mathbf{s}_{0:t+a}, \mathbf{H}_{0:t+a}|\mathbf{x}_{1:t+a}) \approx \sum_{i=1}^M w_{t+a}^{(i)}\times \\ \delta(\mathbf{s}_{0:t+a} - \mathbf{s}_{0:t+a}^{(i)}) \sum_{\ell\in\mathcal{M}} \delta(\mathbf{H}_{0:t+a} - \mathbf{H}_{\ell;0:t+a}^{(i)}). \end{gathered} \tag{17}$$

Integrating (17) over $\mathbf{s}_{t+1:t+a}$ and $\{\mathbf{H}_{m;t+1:t+a}\}_{m\in\mathcal{M}}$, yields an estimate of the desired joint smooting pdf,

$$\begin{aligned} p(\mathbf{s}_{0:t}, \mathbf{H}_{0:t}|\mathbf{x}_{1:t+a}) &\approx \int\int \sum_{i=1}^M w_{t+a}^{(i)} \delta(\mathbf{s}_{0:t+a} - \mathbf{s}_{0:t+a}^{(i)}) \\ &\quad\times \sum_{\ell\in\mathcal{M}} \delta(\mathbf{H}_{0:t+a} - \mathbf{H}_{\ell;0:t+a}^{(i)}) \\ &= \sum_{i=1}^M w_{t+a}^{(i)} \delta(\mathbf{s}_{0:t} - \mathbf{s}_{0:t}^{(i)}) \sum_{\ell\in\mathcal{M}} \delta(\mathbf{H}_{0:t} - \mathbf{H}_{\ell;0:t}^{(i)}) \end{aligned} \tag{18}$$

Successively drawing from (15), updating the weights via (16) and approximating the smoothing pdf by integration as in (18) yields the new weighted particle set $\Omega_{t+a} = \left\{\left(\mathbf{s}_{0:t}^{(i)}, \left\{\mathbf{H}_{m;0:t}^{(i)}\right\}_{m\in\mathcal{M}}\right), w_{t+a}^{(i)}\right\}_{i=1}^M$. Approximate marginal MAP symbol estimates are computed as

$$\hat{\mathbf{s}}_t^{map} = \arg\max_{\mathbf{s}_t} \left\{\sum_{i=1}^M \delta(\mathbf{s}_t - \mathbf{s}_t^{(i)}) w_{t+a}^{(i)}\right\}. \tag{19}$$

Next, we elaborate the details of the importance functions from which the symbols and the channel CIRs are drawn.

4.1. Channel sampling scheme

The form of function

$$q\left(\{\mathbf{H}_{m;t:t+a}\}_{m\in\mathcal{M}}|\mathbf{s}_{0:t-1}, \{\mathbf{H}_{m;0:t-1}\}_{m\in\mathcal{M}}, \mathbf{x}_{0:t+d}\right) \tag{20}$$

in (15) implies that we need to draw samples of the channel matrices $\mathbf{H}^{(i)}_{m;t:t+a}$ for each one of the possible channel orders $m \in \mathcal{M}$.

For the sampling scheme to be sufficiently efficient (in order to avoid the need of a very large number of particles), we propose to run a bank of $|\mathcal{M}|$ adaptive channel estimators, one for each channel order. In particular, $\mathbf{H}^{(i)}_{m;t}$, $m = 1, \ldots, |\mathcal{M}|$, are drawn from Gaussian proposal pdf's

$$\mathbf{H}^{(i)}_{m;t} \sim N\left(\mathbf{H}_{m;t} | \gamma \hat{\mathbf{H}}^{(i)}_{m;t-1}, \sigma_H^2 \mathbf{I}\right), \tag{21}$$

where the mean, $\gamma \hat{\mathbf{H}}^{(i)}_{m;t-1}$ depends on the output of the channel estimator for particle i and order m, γ is the coefficient of the AR channel model and σ_H^2 is a design parameter. Then, for each $m \in \mathcal{M}$, subsequent samples, $\mathbf{H}^{(i)}_{m;t+1:t+a}$, are drawn using the AR model directly, i.e.,

$$\mathbf{H}^{(i)}_{m;t+k} \sim N(\mathbf{H}_{m;t+k} | \gamma \mathbf{H}^{(i)}_{m;t+k-1}, \sigma_v^2 \mathbf{I}), \quad k = 1, \ldots, a. \tag{22}$$

Moreover, since samples with different channel orders are statistically independent, we can write

$$\begin{aligned} q\left(\{\mathbf{H}_{m;t:t+a}\}_{m\in\mathcal{M}} | \mathbf{s}_{0:t-1}, \{\mathbf{H}_{m;0:t-1}\}_{m\in\mathcal{M}}, \mathbf{x}_{0:t+a}\right) = \\ \prod_{m\in\mathcal{M}} q\left(\mathbf{H}_{m;t:t+a} | \mathbf{s}_{0:t-1}, \mathbf{H}_{m;0:t-1}, \mathbf{x}_{0:t+a}\right) = \\ \prod_{m\in\mathcal{M}} N(\mathbf{H}_{m;t} | \gamma \hat{\mathbf{H}}^{(i)}_{m;t-1}, \sigma_H^2 \mathbf{I}) \times \\ \prod_{k=1}^{a} N(\mathbf{H}_{m;t+k} | \gamma \mathbf{H}^{(i)}_{m;t+k-1}, \sigma_v^2 \mathbf{I}). \end{aligned} \tag{23}$$

As an adaptive channel estimator, we propose to use the exponentially-weighted RLS algorithm [10]. Specifically, for each $m \in \mathcal{M}$ and each $i = 1, \ldots, M$, we need to compute

$$\hat{\mathbf{H}}^{(i)}_{m;t} = \arg\min_{\mathbf{H}} \left\{ \sum_{k=0}^{t} \lambda^{t-k} \|\mathbf{x}_k - \mathbf{H}\bar{\mathbf{s}}_{m;k}\|^2 \right\}, \tag{24}$$

where $0 < \lambda < 1$ is a forgetting factor, and the subindex m in $\bar{\mathbf{s}}_{m;k}$ stands for the number of stacked symbol vectors, namely, $\bar{\mathbf{s}}_{m;k} = [\mathbf{s}_{t-m+1}^\top \cdots \mathbf{s}_t^\top]^\top$ The sequence of problems defined by (24) are recursively solved using the following equations

$$\mathbf{R}^{(i)^{-1}}_{m;0} \;\propto\; \mathbf{I}_{Nm} \quad \text{(initialization)} \tag{25}$$

$$\mathbf{g}^{(i)}_{m;t} = \frac{\lambda^{-1} \mathbf{R}^{(i)^{-1}}_{m;t-1} \bar{\mathbf{s}}^{(i)}_{m;t}}{1 + \lambda^{-1} \bar{\mathbf{s}}^{(i)^H}_{m;t} \mathbf{R}^{(i)^{-1}}_{m;t-1} \bar{\mathbf{s}}^{(i)}_{m;t}} \tag{26}$$

$$\hat{\mathbf{H}}^{(i)^H}_{m;t} = \hat{\mathbf{H}}^{(i)^H}_{m;t-1} + \mathbf{g}^{(i)}_{m;t}\left(\mathbf{x}_t^H - \bar{\mathbf{s}}^{(i)^H}_{m;t} \hat{\mathbf{H}}^{(i)^H}_{m;t-1}\right) \tag{27}$$

$$\mathbf{R}^{(i)^{-1}}_{m;t} = \lambda^{-1}\left(\mathbf{I}_{Nm} - \mathbf{g}^{(i)}_{m;t} \bar{\mathbf{s}}^{(i)^H}_{m;t}\right) \mathbf{R}^{(i)^{-1}}_{m;t-1}. \tag{28}$$

4.2. Data sampling scheme

Following [7], we propose to use a bank of matrix linear filters to obtain soft symbol estimates that enable the efficient sampling of the symbol vectos $\mathbf{s}_{t:t+a}$. Specifically, $Nm \times 1$ dimensional vectors of symbol estimates are computed as

$$\mathbf{y}^{(i)}_{m;t} = \mathbf{W}^{(i)^H}_{m;t} \mathbf{x}_{t,m-1} \tag{29}$$

for $i = 1, \ldots, M$ and $m = 1, \ldots, |\mathcal{M}|$. Note that, for each $m \in \mathcal{M}$, the smoothing lag in the observations of (29), $\mathbf{x}_{t,m-1}$, is selected as $m-1$, which yields matrix filters, $\mathbf{W}^{(i)}_{m;t}$, with dimensions $Nm \times Lm$. When the filters are correctly built, the components of

$$\mathbf{y}^{(i)}_{m;t} = \left[y^{(i)}_{m;1,t}, \ldots, y^{(i)}_{m;N,t}, \ldots, y_{m;N,t+m-1}\right]^\top \tag{30}$$

are estimates of $s_{1,t}, \ldots, s_{N,t}, \ldots, s_{N,t+m-1}$.

Each matrix filter is designed as a *lightweight* version of the minimum mean square error (MMSE) detector, derived by applying the inversion lemma [10] to avoid the computation of inverse matrices. In particular, for each $m \in \mathcal{M}$, we recursively approximate the inverse of the autocorrelation matrix

$$\mathbf{R}^{-1}_{m;t,x} = \left(\sum_{n=0}^{t} \alpha^{t-n} \mathbf{x}_{t,m-1} \mathbf{x}^H_{t,m-1}\right)^{-1} \tag{31}$$

as

$$\hat{\mathbf{R}}^{-1}_{m;0,x} \;\propto\; \mathbf{I}_{Lm} \quad \text{(initialization)} \tag{32}$$

$$\hat{\mathbf{R}}^{-1}_{m;t,x} = \alpha^{-1}\left(\mathbf{I}_{Lm} - \mathbf{g}_{m;t}\mathbf{x}^H_{t,m-1}\right) \hat{\mathbf{R}}^{-1}_{m;t-1,x} \tag{33}$$

where $0 < \alpha < 1$ is a forgetting factor and

$$\mathbf{g}_{m;t} = \frac{\alpha^{-1} \hat{\mathbf{R}}^{-1}_{m;t-1} \mathbf{x}_{t,m-1}}{1 + \alpha^{-1} \mathbf{x}^H_{t,m-1} \hat{\mathbf{R}}^{-1}_{m;t-1} \mathbf{x}_{t,m-1}} \tag{34}$$

is a gain vector. The linear MMSE filters are then constructed as

$$\mathbf{W}^{(i)}_{m;t} = \sigma_s^2 \hat{\mathbf{R}}^{-1}_{m;t,x} \mathbf{H}^{(i)}_{m;t,m-1} \mathbf{E}, \tag{35}$$

for $i = 1, \ldots, M$ and $m = 1, \ldots, \mathcal{M}$, where σ_s^2 is the symbol power, $\mathbf{E} = \begin{bmatrix} \mathbf{0}_{N(m-1)\times Nm} \\ \mathbf{I}_{Nm} \end{bmatrix}$, and $\mathbf{H}^{(i)}_{m;t,m-1}$ is obtained by stacking matrices $\mathbf{H}^{(i)}_{m;t:t+m-1}$ in the same way as $\mathbf{H}_{t,d}$ results from stacking $\mathbf{H}_{t:t+d}$ in (4).

Since the maximum smoothing lag is $a = |\mathcal{M}| - 1$, corresponding to the lag needed for the maximum channel order, for each $m \in \mathcal{M}$ such that $m < |\mathcal{M}|$ it becomes necessary to apply a sequence of filters

$$\mathbf{W}^{(i)}_{m;t}, \ldots, \mathbf{W}^{(i)}_{m;t+(|\mathcal{M}|-m)}, \tag{36}$$

where

$$\mathbf{W}^{(i)}_{m;t+k} = \sigma_s^2 \hat{\mathbf{R}}^{-1}_{m;t+k,x} \mathbf{H}^{(i)}_{m;t+k,m-1} \mathbf{E}. \tag{37}$$

As a result, for each $i = 1, \ldots, M$ and each $m \in \mathcal{M}$ we obtain a sequence of estimates $y^{(i)}_{m;1,t}, \ldots, y^{(i)}_{m;N,t+a}$ for the symbols $s_{1,t}, \ldots, s_{N,t+a}$.

Let $y_{m;j,t+k}^{(i)}$ denote the linear estimate of $s_{j,t+k}$ assuming channel order m. If the symbols are binary, $s_{j,t+k} \in \{\pm 1\}$ (extension to higher order constellations is straightforward), we can assign probabilities $q_{+1|m,j,t+k}^{(i)} \propto \exp\{-\frac{1}{\sigma_y^2}|y_{m;j,t+k} - 1|^2$ (where σ_y^2 is a design parameter) and $q_{-1|m,j,t+k}^{(i)} = 1 - q_{+1,j,t+k}^{(i)}$. These probabilities are conditional (as indicated by the notation) on channel order m. We combine them by using the *a posteriori* pdf of the channel order, to yield

$$\begin{aligned} q_{+1,j,t+k}^{(i)} &= \sum_{m \in \mathcal{M}} q_{+1|m,j,t+k}^{(i)} \times \\ &\quad p(m|\mathbf{s}_{0:t-1}, \mathbf{H}_{m;0:t-1}, \mathbf{x}_{0:t-1}) \\ q_{-1,j,t+k}^{(i)} &= \sum_{m \in \mathcal{M}} q_{-1|m,j,t+k}^{(i)} \times \\ &\quad p(m|\mathbf{s}_{0:t-1}, \mathbf{H}_{m;0:t-1}, \mathbf{x}_{0:t-1}). \end{aligned} \quad (38)$$

These probabilities enable us to define a discrete binary probability distribution from which a sample $s_{j,t+k}^{(i)}$ can be drawn. Repeating this process for all symbols from time t to time $t+a$ we obtain the desired sample $\mathbf{s}_{t:t+a}^{(i)}$. The evaluation of $q(\mathbf{s}_{t:t+a}|\mathbf{s}_{0:t-1}, \{\mathbf{H}_{m;0:t+a}\}_{m\in\mathcal{M}}, \mathbf{x}_{0:t+a})$ is carried out by adequately multiplying the probabilities $q_{\pm 1,j,t}^{(i)}$ for $j = 1, ..., N$ and $t, ..., t+a$.

5. SIMULATION RESULTS

Consider a simple system with $N = 2$ transmitting antennas, $L = 3$ receiving antennas and CIR length $m = 3$. The parameters of the channel AR process are $\gamma = 1 - 10^{-5}$ and $\sigma_v^2 = 10^{-4}$. Also assume a BPSK modulation format and burst data transmission in blocks of $K = 300$ symbol vectors (i.e., 600 binary symbols overall), including $T = 30$ pilot symbols from each transmitter.

Within this simulation setup, we have compared:

- The two complexity-constrained SMC equalizers introduced in [7], that assume exact knowledge of the channel order, namely, the 'RLS-D-SIS' (RLS-Delayed sampling-SIS) and 'LMS-D-SIS' (Least Mean Squares - Delayed sampling - SIS). They only differ in the channel estimator they use (the RLS algorithm and the LMS algorithm, respectively).

- Two 'RLS-D-SIS' algorithms that work considering an incorrect channel order: one of them underestimating it ($m = 2$) and the other overestimating it ($m = 4$).

- A genie-aided maximum likelihood sequence detector (labeled 'MLSD') implemented using the Viterbi algorithm with perfect knowledge of the time varying CIR (including its order, m).

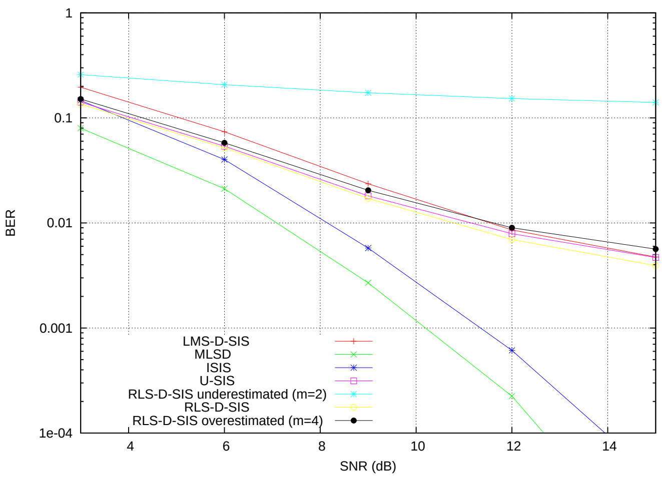

Fig. 1. BER of the SMC equalizers, with $M = 30$ particles, and the MLSD receiver for several values of the SNR (dB). The true channel order is $m = 3$ and the maximum channel order for the U-SIS is set as $|\mathcal{M}| = 4$.

- The two algorithms presented in this work for unknown channel order. The optimal procedure, described is Section 3.2, is labeled 'I-SIS' (Integrated-order SIS) while the complexity-constrained approach in Section 4 is labeled 'U-SIS' (Unknown channel order-SIS).

Figure 1 shows the bit error rate (BER) attained by the SMC equalizers, with $M = 30$ particles, and the genie-aided MLSD detector for several values of the signal-to-noise ratio (SNR). It is observed that the performance of the I-SIS and MLSD receivers is very similar (the BER curve presents practically the same slope). On the other hand, the U-SIS method appears suffers from some performance loss with respect to the optimal method (in exchange for the complexity reduction) and the RLS-D-SIS receiver with the true channel order ($m = 3$), but it outperforms the two RLS-D-SIS equalizers with mismatched order. Notice, in particular, the very poor behavior of the RLS-D-SIS technique with underestimated m.

For each t, the U-SIS algorithm computes the *a posteriori* pdf of the random channel order $m \in \mathcal{M}$. Figure 2 (left) depicts the evolution of this pdf for $t = 31 : 300$, i.e., the duration of a data frame (notice that no channel order estimation is carried out for $0 \leq t \leq 30$). It can be seen that the probability of the actual channel order, $m = 3$, quickly converges to a value very close to 1, while the other probabilities practically vanish. The results are the average of 50 independent simulations.

In order to check the robustness of the U-SIS technique, we have repeated the same experiment for a channel with actual order $m = 5$ and maximum order $|\mathcal{M}| = 8$. Figure 2 (right) shows the results. It can be seen that the proposed algorithm assigns a much higher probability (again, close to

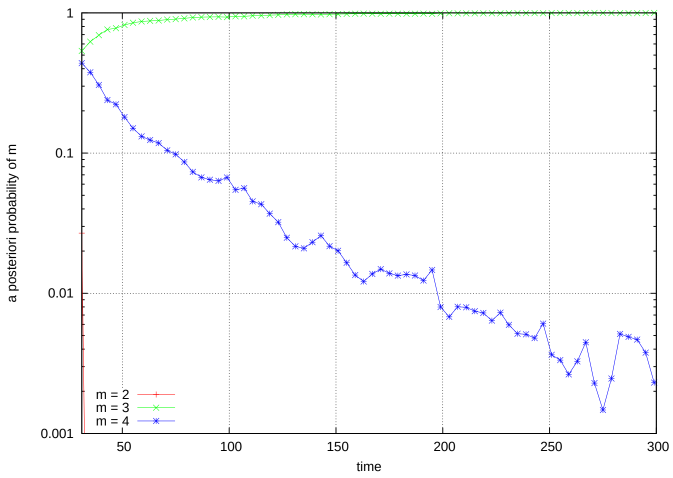

Fig. 2. **Left:** Time evolution of the *a posteriori* pdf of the different channel orders, as computed by the U-SIS method. The actual channel order is $m = 3$, the maximum possible order is $|\mathcal{M}| = 4$ and the number of particles is $M = 30$. **Right:** Time evolution of the *a posteriori* pdf of the different channel orders, as computed by the U-SIS method. The actual channel order is $m = 5$, the maximum possible order is $|\mathcal{M}| = 8$ and the number of particles is $M = 30$.

1) to the true channel order, while the other possible values receive very small probabilities (practically negligible for $m < 5$).

6. CONCLUSIONS

In this article, we have proposed the application of the SMC methodology to build quasi-MAP adaptive equalizers for MIMO systems that can handle the uncertainty in the channel order. In particular, we have derived both optimal (I-SIS) and complexity-constrained (U-SIS) SMC algorithms for joint data detection, channel order and CIR estimation in frequency and time-selective MIMO channels. The complexity of the optimal I-SIS equalizer grows exponentially with the channel dimensions, while U-SIS is constrained to a polynomial computational load. Besides the analytical derivation of the algorithms, we have shown computer simulation results that illustrate the performance of the proposed techniques.

Compared to existing methods for MIMO channel order estimation, the proposed SMC-based techniques:

- are designed for joint data detection and CIR estimation (including the order);
- perform sequential (adaptive) processing, rather than batch mode computations like the existing CME [1] and subspace-based [3] methods;
- produce soft order information (posterior probabilities) that can be exploited in different ways, instead of just choosing one particular order as in [1, 3]; and
- has been shown to work with relatively low and medium SNRs (below 10 dB), while the subspace technique in [3] is very sensitive to noise (according to the plots in that paper, accurate order detection is only achieved in general when SNR¿20 dB) and the CME-based algorithm of [1] attains only a relatively modest percentage of correct order detection (which is never higher than 85% for the range 10¡SNR¡20 dB and a simple, order 2, channel).

7. REFERENCES

[1] Daniel Ka Chun So and Roger S. Cheng, "Iterative EM receiver for space-time coded systems in MIMO frequency-selective fading channels with channel gain and order estimation," *IEEE Transactions on Wireless Communications*, vol. 3, no. 6, pp. 1928–1935, November 2004.

[2] S. Kay, "Conditional model order estimation," *IEEE Transactions on Signal Processing*, vol. 49, no. 9, pp. 1900–1917, September 2001.

[3] D. Kotoulas, P. Koukoulas, and N Kalouptsidis, "Subspace projection based blind channel order estimation of MIMO systems," *Automatica*, vol. 54, pp. 1351–1363, 2006.

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

[5] 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.

[6] D. Guo and X. Wang, "Blind detection in MIMO systems via sequential Monte Carlo," *IEEE Journal on Selected Areas in Communications*, vol. 21, no. 3, pp. 464–473, April 2003.

[7] Manuel A. Vázquez López and Joaquín Míguez, "SMC algorithms for approximate-MAP equalization of MIMO channels with polynomial complexity," in *Proceedings of the EURASIP EUSIPCO*, September 2006.

[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] M. Bolić, P. M. Djurić, and S. Hong, "Resampling algorithms and architectures for distributed particle filters," *IEEE Transactions Signal Processing*, vol. 53, no. 7, pp. 2442–2450, July 2005.

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