A COMPLEXITY-CONSTRAINED PARTICLE FILTERING ALGORITHM FOR MAP EQUALIZATION OF FREQUENCY-SELECTIVE MIMO CHANNELS

Manuel A. Vázquez[†], *Joaquín Míguez*[‡]

[†]Depto. Electrónica e Sistemas, Universidade da Coruña,
Campus de Elviña s/n, 15071 A Coruña, Spain. E-mail: mvazquez@udc.es
[‡]Depto. Teoría de la Señal y Comunicaciones, Universidad Carlos III de Madrid.
Butarque 15, 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. Unfortunately, for each simulated data sample, the complexity of existing algorithms grows exponentially with the number of input data streams. In this paper, we propose a novel SMC MIMO channel equalizer that avoids this limitation. An adequate design of the data sampling scheme leads to a reduction of the computational load per sample, which becomes linear in the number of channel inputs. Computer simulations that illustrate the nearly optimal bit error rate of the proposed SMC equalizer are presented.

1. INTRODUCTION

Very recently, the application of Sequential Monte Carlo (SMC) methods, also known as particle filtering (PF) [1], to Multiple Input Multiple Output (MIMO) channel equalization has been proposed in some papers, including [2, 3, 4]. In [2] it is shown that nearly optimal bit error rate (BER) can be achieved using particle filtering, but the complexity of the sampling scheme in [2] grows exponentially with the number of input data streams. In [3], complexity of the SMC algorithm is reduced by handling together identical data samples, or *particles*, which are represented as paths in a tree. In the same spirit, an stochastic M tree-search algorithm is proposed in [4].

Although the techniques in [3, 4] are successful in reducing the number of Monte Carlo samples to be processed, the complexity of generating and propagating each particle still grows exponentially with the number of data streams. In this paper, we propose a novel SMC MIMO equalizer that avoids this limitation. We adequately design the data sampling scheme in order to constrain the computational load per particle to be linear in the number of input data streams. Computer simulations that illustrate how the proposed SMC receiver attains nearly optimal bit error rate with a small number of particles are presented.

The remaining of the paper is organized as follows. In next section, the signal model for transmission over a MIMO dispersive channel is described. In section 3, The standard application of PF to MIMO equalization is discussed and the proposed SMC receiver is introduced. Illustrative computer simulations are shown in section 4 and, finally, concluding remarks are made in section 5.

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

2. SIGNAL MODEL

Transmission over a frequency-selective wireless channel with N input data streams and L output observation streams can be described by the discrete-time baseband-equivalent model (see, e.g., [2])

$$\mathbf{x}_t = \sum_{i=0}^{m-1} \mathbf{H}_t(i)\mathbf{s}_{t-i} + \mathbf{g}_t, \quad t = 0, 1, \ldots \qquad (1)$$

where $\mathbf{x}_t$ is the $L \times 1$ vector of observations collected at the receiving antennas, $\{\mathbf{H}_t(i)\}_{i=0}^{m-1}$ is the (time-varying) $L \times N$ matrix impulse response of the channel, $\mathbf{s}_t = [s_t(1), \ldots, s_t(N)]^\top$ is the $N \times 1$ vector of symbols transmitted at time t and $\mathbf{g}_t$ is an $L \times 1$ vector of independent Additive White Gaussian Noise (AWGN) components with variance $\sigma_{g,t}^2$.

It is common [1] to model the channel variation by means of a an AR (**A**uto**R**egressive) proccess driven by white Gaussian noise. Hence, we consider the first order AR model (see, e.g., [5])

$$\mathbf{H}_t(i) = \alpha\mathbf{H}_{t-1}(i) + \mathbf{V}_t(i) \quad \forall i \qquad (2)$$

where $\mathbf{V}_t(i)$ are $L \times N$ matrices of i.i.d. Gaussian elements with zero mean and variance σ_v^2. Model parameters α and σ_v^2 are selected to fit the field-measured autocorrelation function of the channel [1].

Using the above equations, MIMO transmission can be modeled as a dynamic system in state-space form. Specifically, the system state at time t consists of the channel impulse response, $\{\mathbf{H}_t(i)\}_{i=0}^{m-1}$, and the symbol vectors $\mathbf{s}_{t-m+1:t} = \{\mathbf{s}_{t-m+1}, \ldots, \mathbf{s}_t\}$. The channel state equation is (2), while the symbols are modeled as discrete uniform random variables in the alphabet $\mathcal{S}$, hence $\mathbf{s}_t \sim \mathcal{U}(\mathcal{S}^N)$. The observation equation is (1).

The dynamic system representation allows to use particle filtering in order to compute optimal joint estimates of the channel response and the transmitted data from the collected observations, as described below.

3. MIMO CHANNEL EQUALIZATION

3.1. Sequential Importance Sampling

Most particle filtering methods rely upon the principle of Importance Sampling (IS) [6] for building an empirical approximation of a desired PDF (say $p(x)$) by drawing samples from a different distribution, known as *importance function* or *proposal PDF* (denoted $q(x)$). These samples are then assigned appropriate normalized *importance* weights, i.e.,

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

where $\sum_{n=1}^{N} w^{(n)} = 1$. Since we are interested in detecting the transmitted symbols, we need to approximate the *a posteriori* PDF $p(\mathbf{s}_{0:t}|\mathbf{x}_{0:t})$ which contains all relevant statistical information for the optimal (Bayesian) estimation of $\mathbf{s}_{0:t}$, and the importance function has the form $q(\mathbf{s}_{0:t}|\mathbf{x}_{0:t})$.

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

$$p(\mathbf{s}_{0:t}|\mathbf{x}_{0:t}) \propto p(\mathbf{x}_t|\mathbf{s}_{0:t}, \mathbf{x}_{0:t-1})p(\mathbf{s}_{0:t-1}|\mathbf{x}_{0:t-1}), \qquad (3)$$

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

$$q(\mathbf{s}_{0:t}|\mathbf{x}_{0:t}) = q(\mathbf{s}_t|\mathbf{s}_{0:t-1}, \mathbf{x}_{0:t})q(\mathbf{s}_{0:t-1}|\mathbf{x}_{0:t-1}). \qquad (4)$$

The recursive algorithm that combines the IS principle and decompositions (3) and (4) to build a discrete random measure that approximates the posterior PDF is called Sequential Importance Sampling (SIS) [6]. Let $\Omega_t = \left\{\mathbf{s}_{0:t}^{(i)}, w_t^{(i)}\right\}_{i=1}^{M}$ denote the discrete measure at time t. The desired PDF is approximated as

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

where $\delta(\cdot)$ is Dirac's delta function. When a new observation is collected at time $t+1$, the SIS algorithm proceeds as follows to recursively compute Ω_{t+1}:

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

It can be shown that the particle filter computed with the SIS algorithm converges to the desired posterior pdf for a sufficiently large number of particles [6], i.e., $\hat{p}(\mathbf{s}_{0:t}|\mathbf{x}_{0:t}) \overset{N\to\infty}{\longrightarrow} p(\mathbf{s}_{0:t}|\mathbf{x}_{0:t})$. Besides, it is straightforward to obtain data estimates from the approximate PDF $\hat{p}(\mathbf{s}_{0:t}|\mathbf{x}_{0:t})$. E.g., a marginal MAP detector can be implemented as

$$\hat{\mathbf{s}}_t^{map} = \arg\max_{\mathbf{s}_t} \left\{ \sum_{i=1}^{M} \delta(\mathbf{s}_t - \mathbf{s}_t^{(i)})w_t^{(i)} \right\}, \qquad (5)$$

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 its simplest form, resampling generates N new particles $\{\mathbf{s}_{0:t}^{(i)}, 1/M\}_{i=1}^{M}$ by drawing samples from the discrete PDF $p_{resampling}(\mathbf{s}_{0:t}^{(i)}) = w_t^{(i)}$.

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 MIMO equalization problem is

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

which contains all the information available at time t for the sampling of $\mathbf{s}_t$. The likelihood on the right-hand side of (6) can be obtained in closed-form. Indeed, if we let $\mathbf{H}_t = [\mathbf{H}_t(m-1) \cdots \mathbf{H}_t(0)]$ be the $L \times Nm$ overall channel matrix and use $\mathbf{h}_t$ to denote the $LNm \times 1$ vector built by taking all elements in $\mathbf{H}_t$ row-wise, then we can write

$$p(\mathbf{x}_t|\mathbf{s}_{0:t}, \mathbf{x}_{0:t-1}) = \int_{\mathbf{h}_t} p(\mathbf{x}_t|\mathbf{s}_{t-m+1:t}, \mathbf{h}_t)p(\mathbf{h}_t|\mathbf{s}_{0:t-1}, \mathbf{x}_{0:t-1})d\mathbf{h}_t. \qquad (7)$$

Both densities in the integrand are Gaussian and, therefore, the integral can be solved (see, e.g., [5]). Specifically, $p(\mathbf{x}_t|\mathbf{s}_{t-m+1:t}, \mathbf{h}_t) = \mathcal{N}(\sum_{i=0}^{m-1} \mathbf{H}_t(i)\mathbf{s}_{t-i}, \sigma_{g,t}^2\mathbf{I})$ while $p(\mathbf{h}_t|\mathbf{s}_{0:t-1}, \mathbf{x}_{0:t-1})$ can be computed using the Kalman Filter (KF). The latter observation becomes apparent if we note that, given the symbols $\mathbf{s}_{0:t}$, the dynamic system (2)-(1) is linear in $\mathbf{h}_t$ and Gaussian. It is actually well-known that the KF can be integrated into the SIS algorithm for conditionally-linear Gaussian systems and the resulting method is termed Mixture Kalman Filter (MKF) [6, 7, 5].

The weight update equation for the importance function (6) can be easily derived, and the complete algorithm becomes

$$\mathbf{s}_t^{(i)} \sim q_t(\mathbf{s}_t) = \frac{p(\mathbf{x}_t|\mathbf{s}_t, \mathbf{s}_{0:t-1}^{(i)}, \mathbf{x}_{0:t-1})}{\sum_{\tilde{\mathbf{s}}_t \in \mathcal{S}^N} p(\mathbf{x}_t|\tilde{\mathbf{s}}_t, \mathbf{s}_{0:t-1}^{(i)}, \mathbf{x}_{0:t-1})} \qquad (8)$$

$$w_t^{(i)} \propto w_{t-1}^{(i)} \sum_{\tilde{\mathbf{s}}_t \in \mathcal{S}^N} p(\mathbf{x}_t|\tilde{\mathbf{s}}_t, \mathbf{s}_{0:t-1}^{(i)}, \mathbf{x}_{0:t-1}), \qquad (9)$$

with resampling when needed.

From (8), (9), it is seen that the MKF algorithm requires the computation of $|\mathcal{S}|^N$ different likelihoods (one for each possible value of $\mathbf{s}_t$) and each likelihood involves one Kalman filter step. As a consequence, the complexity of the method grows exponentially with the number of transmit antennas, which renders the algorithm impractical.

Moreover, algorithm (8), (9) yields a poor average performance when the MIMO channel is highly dispersive. Indeed, due to the channel convolutional effect, the energy of $\mathbf{s}_t$ is distributed over two or more symbol periods and decisions made at time t are necessarily unreliable.

3.3. Delayed Sampling

Whatever the approach, detection in dispersive channels usually requires smoothing, i.e., $\mathbf{s}_t$ is detected based on posterior observations $\mathbf{x}_{0:t+d}$, where $d \geq m-1$ is a smoothing lag. In the context of particle filtering, smoothing is also referred to as *delayed sampling* [6, 5] because particle $\mathbf{s}_t^{(i)}$ cannot be drawn until $\mathbf{x}_{t+d}$ is observed.

The optimal smoothing importance PDF is

$$q(\mathbf{s}_t|\mathbf{s}_{0:t-1}, \mathbf{x}_{0:t+d}) = p(\mathbf{s}_t|\mathbf{s}_{0:t-1}, \mathbf{x}_{0:t+d}) \propto \sum_{\tilde{\mathbf{s}}_{t+1:t+d} \in \mathcal{S}^{Nd}} \prod_{k=0}^{d} p\left(\mathbf{x}_{t+k}|\mathbf{s}_{0:t}, \tilde{\mathbf{s}}_{t+1:t+k}, \mathbf{x}_{0:t+k-1}\right), \qquad (10)$$

where factors $p\left(\mathbf{x}_{t+k}|\mathbf{s}_{0:t}, \tilde{\mathbf{s}}_{t+1:t+k}, \mathbf{x}_{0:t+k-1}\right)$ can be computed in the same way as (7), and the weight update equation becomes

$$w_{t+d}^{(i)} = w_{t+d-1}^{(i)} \sum_{\tilde{\mathbf{s}}_{t:t+d}} \prod_{k=0}^{d} p\left(\mathbf{x}_{t+k}|\mathbf{s}_{0:t-1}^{(i)}, \tilde{\mathbf{s}}_{t:t+k}, \mathbf{x}_{0:t+k-1}\right). \tag{11}$$

where $\tilde{\mathbf{s}}_{t:t+d} \in \mathcal{S}^{N(d+1)}$. Therefore, sampling and updating a single particle with this method involves the computation of $|\mathcal{S}|^{N(d+1)}$ likelihoods, one for each possible different sequence $\tilde{\mathbf{s}}_{t:t+d}$, and each likelihood requires $d+1$ Kalman filter steps. This means that the complexity of the algoritm grows exponentially with the number of antennas and the smoothing parameter, i.e., it is $\mathcal{O}(|\mathcal{S}|^{N(d+1)})$. Although the smoothed MKF procedure guarantees a practically optimal performance, its computational burden makes it intractable.

3.4. Constrained-complexity smoothing

In order to perform smoothing with a tractable complexity, we propose a novel SMC scheme based on the ideas of sampling in a higher dimension [8] and sequentially on the user (data) space [9]. Specifically, we address the approximation of the joint PDF $p(\mathbf{s}_{0:t+d}, \mathbf{h}_{t:t+d}|\mathbf{x}_{0:t+d})$ using an importance function $q_{t+d}(\mathbf{s}_{t:t+d}, \mathbf{h}_{t:t+d})$, to be defined below. In this case, the SIS algorithm is

$$\left(\mathbf{s}_{t:t+d}^{(i)}, \mathbf{h}_{t:t+d}^{(i)}\right) \sim q_{t+d}(\mathbf{s}_{t:t+d}, \mathbf{h}_{t:t+d})$$
$$w_{t+d}^{(i)} = w_{t+d-1}^{(i)} \frac{\prod_{k=0}^{d} p(\mathbf{x}_{t+k}|\mathbf{h}_{t+k}^{(i)}, \mathbf{s}_{t+k}^{(i)}) p(\mathbf{h}_{t+k}|\mathbf{s}_{0:t+k}^{(i)}, \mathbf{x}_{0:t+k-1})}{q_{t+d}(\mathbf{s}_{t:t+d}^{(i)}, \mathbf{h}_{t:t+d}^{(i)})} \tag{12}$$

for $i = 1, \ldots, M$, and resampling when needed. Note that the factors on the numerator of the weight update equation are all Gaussian and can be computed (using the KF for the right-most likelihood). The symbols $\mathbf{s}_{t+1:t+d}^{(i)}$ and the channel sequence $\mathbf{h}_{t:t+d}$ are just auxiliary variables [8] which are sampled for convenience and can be discarded in order to obtain a discrete random measure $\Upsilon_{t+d} = \{\mathbf{s}_{0:t}^{(i)}, w_{t+d}^{(i)}\}_{i=1}^{M}$ that approximates $p(\mathbf{s}_{0:t}|\mathbf{x}_{0:t+d})$.

In order to define the proposal PDF, $q_{t+d}(\cdot)$, we need to build a transformed observation model. We begin with the $L \times (d+1)$ stacked vector

$$\mathbf{x}_{t,d} = \mathbf{H}_{t,d}\mathbf{s}_{t,d} + \mathbf{g}_{t,d} \tag{13}$$

where $\mathbf{x}_{t,d} = \left[\mathbf{x}_t^\top, \ldots, \mathbf{x}_{t+d}^\top\right]^\top$, $\mathbf{s}_{t,d} = \left[\mathbf{s}_{t-m+1}^\top, \ldots, \mathbf{s}_{t+d}^\top\right]^\top$ is the $N(m+d) \times 1$ vector containing all contributing symbols, $\mathbf{g}_{t,d} = \left[\mathbf{g}_t^\top, \ldots, \mathbf{g}_{t+d}^\top\right]^T$ is the $L(d+1) \times 1$ AWGN vector, and

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

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

Let us initially factor the importance PDF as

$$q_{t+d}(\mathbf{h}_{t:t+d}^{(i)}, \mathbf{s}_{t:t+d}^{(i)}) \propto \prod_{k=1}^{d} p(\mathbf{h}_{t+k}^{(i)}|\mathbf{h}_{t+k-1}^{(i)}) \times$$
$$\times p(\mathbf{h}_t^{(i)}|\mathbf{s}_{0:t-1}^{(i)}, \mathbf{x}_{0:t-1}) q_{t+d}(\mathbf{s}_{t:t+d}^{(i)}) \tag{14}$$

so that the channel vector $\mathbf{h}_t^{(i)}$ is drawn from the Gaussian distribution $p(\mathbf{h}_t^{(i)}|\mathbf{s}_{0:t-1}^{(i)}, \mathbf{x}_{0:t-1})$, which is given by the KF operating on particle i, and $\mathbf{h}_{t+1:t+d}^{(i)}$ are predicted using the prior PDFs $p(\mathbf{h}_{t+k}^{(i)}|\mathbf{h}_{t+k-1}^{(i)})$.

Since $\mathbf{h}_{t:t+d}^{(i)}$ and $\mathbf{s}_{t-m+1:t-1}^{(i)}$ are already available, we can use them to suppress the causal inter-symbol interference in the observations and obtain

$$\bar{\mathbf{x}}_{t,d}^{(i)} = \mathbf{x}_{t,d} - \breve{\mathbf{H}}_{t,d}^{(i)} \breve{\mathbf{s}}_{t,d}^{(i)}$$

where $\breve{\mathbf{s}}_{t,d}^{(i)} = [\mathbf{s}_{t-m+1}^{(i)\top}, \cdots, \mathbf{s}_{t-1}^{(i)\top}]^\top$ is an $N(m-1) \times 1$ vector,

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

is an $L(d+1) \times N(m-1)$ matrix. Conditional on $\mathbf{h}_{t:t+d} = \mathbf{h}_{t:t+d}^{(i)}$ and $\mathbf{s}_{t-m+1:t-1} = \mathbf{s}_{t-m+1:t-1}^{(i)}$, the resulting model for the observations is

$$\bar{\mathbf{x}}_{t,d}^{(i)} = \bar{\mathbf{H}}_{t,d}^{(i)} \bar{\mathbf{s}}_{t,d} + \mathbf{g}_{t,d},$$

where $\bar{\mathbf{s}}_{t,d} = [\mathbf{s}_{t+d}^\top, \cdots, \mathbf{s}_t^\top]^\top$ and $\bar{\mathbf{H}}_{t,d}^{(i)}$ is an $L(d+1) \times N(d+1)$ matrix obtained by removing the first $N(m-1)$ columns of matrix $\mathbf{H}_{t,d}^{(i)}$, constructed according to model (13), and taking the remaining ones backwards.

Finally, the observations can be cast into the convenient form

$$\mathbf{z}_{t,d}^{(i)} = \mathbf{U}_{t,d}^{(i)^{-1}} \bar{\mathbf{H}}_{t,d}^{(i)^\top} \bar{\mathbf{x}}_{t,d}^{(i)} = \mathbf{U}_{t,d}^{(i)^\top} \bar{\mathbf{s}}_{t,d} + \acute{\mathbf{g}}_{t,d}^{(i)}, \tag{15}$$

where the upper-triangular $N(d+1) \times N(d+1)$ matrix $\mathbf{U}_{t,d}^{(i)}$ is the Cholesky factor of $\bar{\mathbf{R}}_{t,d}^{(i)} = \bar{\mathbf{H}}_{t,d}^{(i)^\top} \bar{\mathbf{H}}_{t,d}^{(i)} = \mathbf{U}_{t,d}^{(i)} \mathbf{U}_{t,d}^{(i)^\top}$, and $\acute{\mathbf{g}}_{t,d}^{(i)}$ is a Gaussian vector with zero mean and covariance matrix $\mathbf{\Sigma}_{t,d}^{(i)} = \mathbf{U}_{t,d}^{(i)^{-1}} \bar{\mathbf{H}}_{t,d}^{(i)^\top} \mathbf{D}_{g,t} \bar{\mathbf{H}}_{t,d}^{(i)} \mathbf{U}_{t,d}^{(i)^{-\top}}$, where $\mathbf{D}_{g,t}$ is an $L(d+1) \times L(d+1)$ diagonal matrix with $\sigma^2_{g,t+\lfloor k/L \rfloor}$ in the (k,k) position.

Symbols $\mathbf{s}_{t:t+d}$ can now be sequentially sampled (starting with $s_t(1)$) using model (15). Let $[\mathbf{Q}]_{i,j}$ denote the element in the i-th row and j-th column of $\mathbf{Q}$ (and similar notation, $[\mathbf{q}]_i$, for vectors). Then, we have (for $k = 1, ..., N(d+1)$)

$$[\mathbf{z}_{t,d}^{(i)}]_k = \sum_{l=1}^{k} u_{l,k}^{(i)} s_{t+\lfloor l/N \rfloor}(1 + (l \div N)) + [\acute{\mathbf{g}}_{t,d}^{(i)}]_k, \tag{16}$$

where $u_{l,k}^{(i)}$ is the element in the (l,k) position of matrix $\mathbf{U}_{t,d}^{(i)}$, $\lfloor \alpha \rfloor$ denotes the largest integer no greater than α and $(x \div y)$ is the remainder of x/y. Using (16) we can sample the symbols in $\mathbf{s}_{t:t+d}$ sequentially [9]. Assume, for simplicity, that $\mathcal{S} = \{\pm 1\}$ (the extension to a larger alphabet is straightforward), then symbol $s_{t+k}^{(i)}(n)$ is drawn conditional on $s_t^{(i)}(1), \ldots, s_{t+k}^{(i)}(n-1)$ according to the probabilities

$$\begin{aligned} Prob_{t+d}\{s_{t+k}(n)=1\} &= Prob\{\gamma_{t+k}^{(i)}(n)>0\} \\ Prob_{t+d}\{s_{t+k}(n)=-1\} &= Prob\{\gamma_{t+k}^{(i)}(n)<0\} \end{aligned}$$

where

$$\gamma_{t+k}(n) = \frac{[\mathbf{z}_{t,d}^{(i)}]_{kN+n} - \sum_{l=1}^{kN+n-1} u_{l,kN+n} s_{t+\lfloor l/N \rfloor}(1+(l \div N))}{u_{kN+n,kN+n}^{(i)}} \tag{17}$$

is a Gaussian random variable with mean $s_{t+k}(n)$ and variance $u_{kN+n,kN+n}^{(i)-2}[\mathbf{\Sigma}_{t,d}^{(i)}]_{kN+n,kN+n}$.

Finally, the proposal PDF for particle i can be recursively evaluated as

$$\begin{aligned} q_{t+d}(\mathbf{h}_{t:t+d}^{(i)}, \mathbf{s}_{t:t+d}^{(i)}) \propto \prod_{k=1}^{d} p(\mathbf{h}_{t+k}^{(i)}|\mathbf{h}_{t+k-1}^{(i)}) \times \\ \times p(\mathbf{h}_t^{(i)}|\mathbf{s}_{0:t-1}^{(i)}, \mathbf{x}_{0:t-1}) \prod_{k=0}^{d} \prod_{n=1}^{N} Prob_{t+d}\{s_{t+k}^{(i)}(n)\}. \end{aligned} \tag{18}$$

The propagation of a single particle in the proposed smoothing algorithm requires $d+1$ KF steps in order to evaluate the right-most likelihoods in the numerator of (12) and the evaluation of $|\mathcal{S}|N(d+1)$ symbol probabilities, instead of the $|\mathcal{S}|^{N(d+1)}$ symbol vector probabilities in the optimal smoothing algorithm (10), (11).

4. SIMULATION RESULTS

For our numerical experiments, we have considered a binary modulation format with symbol alphabet $\mathcal{S} = \{\pm 1\}$, and a frequency-selective MIMO channel with $N = 2$ inputs, $L = 3$ receiving antennas and a channel impulse response of length $m = 2$. The channel AR model parameters are $\alpha = 1 - 10^{-5}$ and $\sigma_v^2 = 10^{-4}$. We assume transmission is carried out in bursts of 300 bits and the first 15 bits in each frame are used as a trainig sequence to obtain an initial (rough) estimate of the channel response. Within this simulation setup, we have compared the optimal SMC filtering algorithm (8)-(9), the proposed complexity-constrained SMC smoother and the standard Maximum Likelihood Sequence Detector (MLSD) with known channel response, which is used as a performance reference. The smoothing lag for the smoothing equalizer is $d = 1$ and the number of particles for both SMC algorithms is $M = 30$, which perform marginal MAP detection according to (5).

Figure 1 shows the obtained results. It is seen that the proposed SMC smoother attains a nearly optimal bit error probability (less than 1 dB away from the reference curve at BER= 10^{-3}) and clearly outperforms the optimal SMC filter. This is particularly remarkable because the complexity per particle of both equalizers is similar. The optimal SMC smoother, which has not been included in the simulation, outperforms the proposed equalizer, but its complexity per particle is much higher.

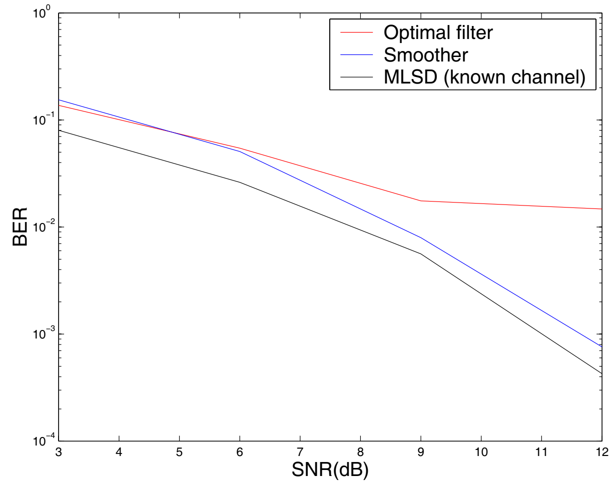

Fig. 1. BER for several values of the SNR (dB).

5. CONCLUSIONS

Existing particle filtering methods for MIMO channel equalization suffer from a stringent limitation because their computational complexity per particle grows exponentially with the number of input data streams. In this paper, we have introduced a novel sampling scheme that avoids this drawback by sampling data sequentially across data streams. Using this approach, we have designed a constrained-complexity SMC equalizer which attains nearly optimal BER, as illustrated by our computer simulations.

6. REFERENCES

[1] 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.

[2] 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.

[3] T. Bertozzi and D. Le Ruyet, "Iterative detection in MIMO channels using particle filtering," in *Proceedings of the IEEE ICC*, June 2004, vol. 5, pp. 2694–2698.

[4] Y. Huang, J. Zhang, and P. M. Djurić, "Bayesian detection for BLAST," *to appear in IEEE Transactions on Signal Processing*, 2004.

[5] 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.

[6] A. Doucet, S. Godsill, and C. Andrieu, "On sequential Monte Carlo Sampling methods for Bayesian filtering," *Statistics and Computing*, vol. 10, no. 3, pp. 197–208, 2000.

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

[8] M. K. Pitt and N. Shephard, "Auxiliary variable based particle filters," in *Sequential Monte Carlo Methods in Practice*, A. Doucet, N. de Freitas, and N. Gordon, Eds., chapter 13, pp. 273–293. Springer, 2001.

[9] J. Zhang, Y. Huang, and P. M. Djurić, "Multiuser detection with particle filtering," in *Proceedings of XI EUSIPCO*, September 2002.