BLIND EQUALIZATION BY SEQUENTIAL IMPORTANCE SAMPLING

Joaquín Míguez[*]

Dept. Electrónica e Sistemas,
Universidade da Coruña.
Facultade de Informática,
Campus de Elviña s/n,
15071 A Coruña, Spain.
e-mail: jmiguez@udc.es

Petar M. Djurić[†]

Dept. of Electrical and Computer Engineering,
Stony Brook University
Stony Brook, NY 11794, USA.
e-mail: djuric@ece.sunysb.edu

ABSTRACT

This paper introduces a novel blind equalization algorithm for frequency-selective channels based on a Bayesian formulation of the problem and the Sequential Importance Sampling (SIS) technique. SIS methods rely on building a Monte Carlo (MC) representation of the probability distribution of interest that consists of a set of samples and associated weights, computed recursively in time. We elaborate on this principle to derive a blind sequential algorithm that performs Maximum A Posteriori (MAP) symbol detection without explicit estimation of the channel parameters.

1. INTRODUCTION

The practicality of future wideband wireless communication systems greatly depends on the development of sophisticated coding and signal processing techniques that provide high spectral efficiencies and allow an approach to the theoretical capacity limits.

One fundamental problem in this context is the detection of a symbol sequence transmitted through a frequency-selective channel. This is a classical topic that has received considerable attention. When the channel parameters are known, Maximum Likelihood (ML) detection is optimal, and can be efficiently implemented by means of the Viterbi algorithm [1]. The straightforward way to acquire channel state information is to transmit training sequences which are known a priori by the transmitter and the receiver, but this approach results in an efficiency loss. Hence, a major stream of research has focused on *blind* methods where symbols are detected without knowledge of the channel coefficients and without using any training symbols. This includes both linear equalizers aimed at symbol detection without explicit channel estimation (see [2] and references therein), and joint channel estimation and symbol detection techniques [3].

This paper introduces a novel blind equalization algorithm based on a Bayesian approach to the problem and the Sequential Importance Sampling (SIS) technique [4, 5], a sequential Monte Carlo (MC) methodology recently revived in the field of statistics. SIS methods rely on building an MC representation of the desired probability distribution, consisting of samples (sometimes called *particles*) and associated weights, computed recursively in time as new observations are received. This MC representation exhibits uniform convergence to the true probability distribution as the number of particles grows [6] and allows to tightly approximate several types of estimators. We have elaborated on these principles to derive a sequential algorithm that attains Maximum A Posteriori (MAP) symbol detection without explicit estimation of the channel response. Although SIS methods have a weakness in that they are computationally very intensive, they lend themselves to implementation in massively parallel hardware, hence presenting themselves as potentially attractive for real-time implementation.

The remaining of this paper is organized as follows. Section 2 describes the signal model for the equalization problem. The Bayesian formulation underlying the proposed blind equalizer is developed in Section 3. In Section 4 we present a deterministic sequential solution to the problem of MAP symbol detection based on the Viterbi algorithm, and we identify its limitations. In Section 5, the SIS algorithm is briefly introduced and applied to the blind equalization problem. Illustrative computer simulations are presented in Section 6 and some concluding remarks are made in Section 7.

2. SIGNAL MODEL

Let us consider an uncoded digital communication system where binary symbols, $s_t \in \{\pm 1\}$, $t = 0, 1, 2, ...$, are transmitted in frames of length $T + 1$ through a frequency-selective fading channel. When the coherence time of the fading process is long enough compared to the frame size, it is common to assume that the channel impulse response is constant for the duration of the frame. At the receiving end, the observed signal is matched-filtered and sampled at the bit rate. The resulting discrete-time sequence constitutes a set of sufficient statistics for symbol detection and can be adequately represented by a dynamical state-space model of the form

$$\text{State equation:} \quad \mathbf{s}_t = \mathbf{T}\mathbf{s}_{t-1} + \mathbf{u}_t \qquad (1)$$

$$\text{Observation equation:} \quad y_t = \mathbf{s}_t^\top \mathbf{h} + v_t. \qquad (2)$$

In (1), the $m \times 1$ symbol vector

$$\mathbf{s}_t = [s_{t-m+1} \; s_{t-m+2} \; \dots \; s_t]^\top$$

[*]This work has been supported by the local government of Galicia, under grant no. PGIDT00PXI10504PR, and the Ministry of Science and Technology of Spain, under grant no. TIC2001-0751-C04-01.

[†]This work has been supported by the National Science Foundation, under awards no. CCR-9903120 and CCR-0082607.

is referred to as the *system state* at time t,

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

is the $m \times m$ *state-transition* matrix, and the $m \times 1$ vector $\mathbf{u}_t = [0\ 0\ \ldots\ s_t]^\top$ (where the new symbol s_t is a binary uniform random variable, i.e., $s_t \sim \mathcal{U}(\pm 1)$, and it is independent from previous and future symbols) is the *state perturbation*. In the observation equation, $\mathbf{h} = [h_{m-1}\ h_{m-2}\ \ldots\ h_0]^\top$ is an $m \times 1$ vector representing the discrete-time equivalent channel impulse response, $v_t \sim \mathcal{N}(0, \sigma^2)$ is a zero-mean Additive White Gaussian Noise (AWGN) component with variance σ^2, and y_t is the *observation* at time t. The channel order, m, will be often referred to as the channel memory because it determines the span of Inter-Symbol Interference (ISI).

The aim is to find a sequential algorithm to compute the MAP estimate of the transmitted symbols in a frame, $s_{0:T} = \{s_0, s_1, \ldots, s_T\}$, when the channel coefficients in $\mathbf{h}$ are unknown.

3. RECURSIVE COMPUTATION OF THE POSTERIOR PROBABILITY

Let $p[\mathbf{s}_{0:T}|y_{0:T}]$ represent the probability mass function (pmf) of the state sequence, $\mathbf{s}_{0:T}$, conditional on the corresponding series of observations. Hereafter, we will use the notation $p[\cdot|\cdot]$ to represent the probability density or mass function of the first argument, depending on whether it is continuous or discrete, respectively, conditional on the second argument. According to (1), the symbol and state sequences are equivalent as long as the bits preceding the current frame, $s_{-m+1:-1}$, are known. Under this assumption, we can write the MAP estimate of the transmitted symbols as

$$s_{0:T}^{(MAP)} = \arg\max_{s_{0:T}} \{p[\mathbf{s}_{0:T}|y_{0:T}]\}. \tag{3}$$

For reasons of computational complexity, it is desirable to solve problem (3) sequentially and recursively, i.e., obtaining $s_{0:t}^{(MAP)}$ from $s_{0:t-1}^{(MAP)}$ as y_t is observed. In order to achieve this goal, let us consider the following decomposition of the posterior pmf,

$$p[\mathbf{s}_{0:t}|y_{0:t}] \propto p[y_t|\mathbf{s}_{0:t}, y_{0:t-1}]p[\mathbf{s}_{0:t-1}|y_{0:t-1}]. \tag{4}$$

Equation (4) provides the basis for the sequential computation of $p[\mathbf{s}_{0:t}|y_{0:t}]$ as long as the likelihood function, $p[y_t|\mathbf{s}_{0:t}, y_{0:t-1}]$, can be analytically derived. This problem is addressed below.

Let us assume the channel vector is a priori distributed according to a Gaussian model with mean $\bar{\mathbf{h}}_{-1}$ and covariance $\mathbf{C}_{-1}$, i.e.,

$$\mathbf{h} \sim \mathcal{N}(\bar{\mathbf{h}}_{-1}, \mathbf{C}_{-1}). \tag{5}$$

The likelihood in (4) can be written as

$$p[y_t|\mathbf{s}_{0:t}, y_{0:t-1}] = \int_{\mathbb{R}^m} p[y_t|\mathbf{h}, \mathbf{s}_t]p[\mathbf{h}|\mathbf{s}_{0:t-1}, y_{0:t-1}]d\mathbf{h} \tag{6}$$

where $p[y_t|\mathbf{h}, \mathbf{s}_t] = \mathcal{N}(\mathbf{s}_t^\top \mathbf{h}, \sigma^2)$. It is important to realize that the posterior pdf of the channel at time t is also proportional to the above integrand,

$$p[\mathbf{h}|\mathbf{s}_{0:t} y_{0:t}] \propto p[y_t|\mathbf{h}, \mathbf{s}_t]\ p[\mathbf{h}|\mathbf{s}_{0:t-1}, y_{0:t-1}] \tag{7}$$

and, therefore, the posterior channel densities are also Gaussian, as a consequence of (5), i.e., $p[\mathbf{h}|\mathbf{s}_{0:t-1}, y_{0:t-1}] = \mathcal{N}(\bar{\mathbf{h}}_{t-1}, \mathbf{C}_{t-1})$ and $p[\mathbf{h}|\mathbf{s}_{0:t}, y_{0:t}] = \mathcal{N}(\bar{\mathbf{h}}_t, \mathbf{C}_t)$. Since both pdf's in the integrand of (6) are Gaussian, their product is also Gaussian and can be analytically found. We begin with the update of the posterior mean and covariance of $\mathbf{h}$ at time t. In this derivation, we exploit (8) (see top of next page). If we identify $\mathbf{U} = \frac{1}{\sigma^2}$, $\mathbf{H} = \mathbf{s}_t^\top$, $\mathbf{A} = \mathbf{C}_{t-1}$ and $\mathbf{z}_a = \bar{\mathbf{h}}_{t-1}$ it is straightforward to show that

$$\begin{aligned} \mathbf{C}_t &= \left(\frac{\mathbf{s}_t\mathbf{s}_t^\top}{\sigma^2} + \mathbf{C}_{t-1}^{-1}\right)^{-1} \\ \bar{\mathbf{h}}_t &= \mathbf{C}_t\left(\frac{\mathbf{s}_t y_t}{\sigma^2} + \mathbf{C}_{t-1}\bar{\mathbf{h}}_{t-1}\right) \end{aligned}$$

where $\bar{\mathbf{h}}_t$ is the optimal Bayesian estimate of the channel impulse response at time t given the observations and the symbols. Finally, since the application of the relationship (8) also yields the proportionality constant in (7), we can easily solve the integral in (6) to obtain the desired analytical expression for the likelihood, shown in (9).

4. EQUALIZATION VIA THE VITERBI ALGORITHM

In this section, we apply the Viterbi algorithm [1] to the sequential computation of $s_{0:T}^{(MAP)}$ using the recursive decomposition (4). This is a well-known tool for the solution of optimization problems that can be translated into a search of the *best* path in a special type of graph termed *trellis*. A trellis graph is arranged as a sequence of regularly connected stages, each one consisting of a fixed number of nodes, and a regular pattern of edges linking stage j to stage $j+1$. The Viterbi algorithm can be applied whenever the overall cost of a path across the trellis admits an additive decomposition, i.e., whenever we can assign a fixed cost to every edge in the graph and: (a) the cost of a particular path is given by the addition of the costs of the edges in the path, (b) either all edges have a non negative cost or all edges have a non positive cost. The algorithm is completely specified by a cost function and a set *branch metrics*, i.e., the costs of the edges in the graph.

In our case, there are 2^m nodes in the t-th stage of the trellis, one for each possible configuration of the state vector, $\mathbf{s}_t \in \{\pm 1\}^m$. The cost function is the logarithm of the posterior pmf, $\mathcal{L}(s_{0:T}) = \log\left(p[\mathbf{s}_{0:T}|y_{0:T}]\right)$, which, by virtue of (4), can be decomposed as an addition of $T+1$ terms

$$\mathcal{L}(s_{0:T}) = \log(p[y_0|\mathbf{s}_0]) + \sum_{t=1}^{T}\log(p[y_t|\mathbf{s}_{0:t}, y_{0:t-1}]).$$

As a consequence, the branch metrics are

- $\lambda_0 = \log(p[y_0|\mathbf{s}_0])$ for $t = 0$,
- $\lambda_t = \log(p[y_t|\mathbf{s}_{0:t}, y_{0:t-1}])$, $t = 1, 2, ..., T$.

Unfortunately, a closer look at the branch metrics reveals that they do not fulfill condition (b) above, i.e., they are not either always non negative or always non positive, since the likelihood function, $p[y_t|\mathbf{s}_{0:t}, y_{0:t-1}]$, can take values both smaller and greater than 1. Therefore, the Viterbi algorithm as specified by the cost function $\mathcal{L}(\cdot)$ and the branch metrics λ_t is not guaranteed to provide the MAP estimate of the data. Even if the algorithm can be applied anyway, we have verified through computer simulations (see Section 6) that it suffers from frequent misconvergence for medium and high Signal-to-Noise Ratio (SNR) values and attains a suboptimal Bit Error Rate (BER).

$$
\begin{aligned}
(\mathbf{y}-\mathbf{H}\mathbf{z})^\top \mathbf{U}(\mathbf{y}-\mathbf{H}\mathbf{z}) + (\mathbf{z}-\mathbf{z}_a)^\top \mathbf{A}(\mathbf{z}-\mathbf{z}_a) &= (\mathbf{z}-\mathbf{z}_p)^\top \mathbf{B}(\mathbf{z}-\mathbf{z}_p) + \mathbf{y}^\top \mathbf{U}\mathbf{y} + \mathbf{z}_a^\top \mathbf{A}\mathbf{z}_a \\
&\quad -(\mathbf{H}^\top \mathbf{U}\mathbf{y} + \mathbf{A}\mathbf{z}_a)^\top \mathbf{B}^{-1}(\mathbf{H}^\top \mathbf{U}\mathbf{y} + \mathbf{A}\mathbf{z}_a) \\
\text{where} \quad \mathbf{B} = \mathbf{H}^\top \mathbf{U}\mathbf{H} + \mathbf{A} \quad &\text{and} \quad \mathbf{z}_p = \mathbf{B}^{-1}(\mathbf{H}^\top \mathbf{U}\mathbf{y} + \mathbf{A}\mathbf{z}_a)
\end{aligned} \tag{8}
$$

$$
p[y_t|\mathbf{s}_{0:t}, y_{0:t-1}] = \frac{|\mathbf{C}_t|^{1/2}}{(2\pi\sigma^2|\mathbf{C}_{t-1}|)^{1/2}} e^{-\frac{1}{2}\left[\frac{y_t^2}{\sigma^2} + \bar{\mathbf{h}}_{t-1}^\top \mathbf{C}_{t-1}^{-1}\bar{\mathbf{h}}_{t-1} - \left(\frac{\mathbf{s}_t y_t}{\sigma^2} + \mathbf{C}_{t-1}^{-1}\bar{\mathbf{h}}_{t-1}\right)^\top \mathbf{C}_t \left(\frac{\mathbf{s}_t y_t}{\sigma^2} + \mathbf{C}_{t-1}^{-1}\bar{\mathbf{h}}_{t-1}\right)\right]}. \tag{9}
$$

---

5. EQUALIZATION VIA THE SIS ALGORITHM

The convergence of the Viterbi algorithm to the true MAP data estimate is not guaranteed, and therefore we consider the application of the SIS algorithm [4, 5] in order to solve problem (3).

The standard statement of the SIS algorithm [5, Section 2] is concerned with dynamical systems in state-space form where all fixed parameters are known and only the state sequence has to be estimated. This is not the case of model (1)-(2), so we present a slightly different derivation here. We begin with an Importance Sampling (IS) [7] scheme to draw M particles from the posterior pmf of the symbols,

$$
\begin{aligned}
\mathbf{s}_{0:T}^{(i)} &\sim q[\mathbf{s}_{0:T}|y_{0:T}] \\
\tilde{w}^{(i)} = \frac{p[\mathbf{s}_{0:T}^{(i)}|y_{0:T}]}{q[\mathbf{s}_{0:T}^{(i)}|y_{0:T}]} \quad &\text{and} \quad w^{(i)} = \frac{\tilde{w}^{(i)}}{\sum_{k=1}^M \tilde{w}^{(k)}}
\end{aligned}
$$

where $q[\mathbf{s}_{0:T}|y_{0:T}]$ is an *importance pmf* with the same support as $p[\mathbf{s}_{0:T}|y_{0:T}]$ but easier to sample from, and $\{w^{(i)}\}_{i=1,\dots,M}$ is a set of normalized *importance weights*. These particles are said to be *properly weighted*,[1] and they yield an MC estimate of the true posterior pmf [4],

$$
p[\mathbf{s}_{0:T}|y_{0:T}] \approx \hat{p}[\mathbf{s}_{0:T}|y_{0:T}] = \sum_{i=1}^M \delta_i w^{(i)} \tag{10}
$$

where $\delta_i = 1$ if $\mathbf{s}_{0:T} = \mathbf{s}_{0:T}^{(i)}$, and $\delta_i = 0$ otherwise.

The IS method can be modified so that it becomes possible to build the state trajectories, $\mathbf{s}_{0:T}^{(i)}$, and the importance weights, $w^{(i)}$, sequentially as new observations arrive. Let us consider an importance pmf that can be factorized as

$$
q[\mathbf{s}_{0:t}|y_{0:t}] = q[\mathbf{s}_t|\mathbf{s}_{0:t-1}, y_{0:t}]q[\mathbf{s}_{0:t-1}|y_{0:t-1}], \quad \forall t. \tag{11}
$$

Working with (4), (11) and the IS principle, it is simple to see that the importance weights can be evaluated recursively in time, leading to the SIS algorithm

$$
\begin{aligned}
\mathbf{s}_t^{(i)} &\sim q[\mathbf{s}_t|\mathbf{s}_{0:t-1}^{(i)}, y_{0:t}] \\
\tilde{w}_t^{(i)} = w_{t-1}\frac{p[y_t|\mathbf{s}_{0:t}^{(i)}, y_{0:t-1}]}{q[\mathbf{s}_t^{(i)}|\mathbf{s}_{0:t-1}^{(i)}, y_{0:t}]} \quad &\text{and} \quad w_t^{(i)} = \frac{\tilde{w}_t^{(i)}}{\sum_{k=1}^M \tilde{w}_t^{(k)}}
\end{aligned}
$$

[1] Meaning that $E_q[h(\mathbf{s}_{0:T}^{(i)})\tilde{w}^{(i)}] = E_p[h(\mathbf{s}_{0:T}^{(i)})]$, where E_p denotes statistical expectation with respect to the pmf in the subindex and $h(\cdot)$ is an arbitrary integrable function of the state sequence.

for $i = 1, ..., M$. The set of particles and normalized weights at time t yield an MC estimate of the posterior, $p[\mathbf{s}_{0:t}|y_{0:t}]$, analogous to (10). Hence, at time T the resulting approximate pmf can be used to perform inference on the posterior distribution. Namely, the MAP estimate of the symbols, $\mathbf{s}_{0:T}^{(MAP)}$, is approximated by the particle with the highest importance weight, i.e.,

$$
\begin{aligned}
\hat{\mathbf{s}}_{0:T}^{(MAP)} &:= \mathbf{s}_{0:T}^{(i_{max})} \\
\text{where} \quad i_{max} &= \arg\max_i \{w_T^{(i)}\}, \quad i = 1, ..., M.
\end{aligned}
$$

The steps involved in the SIS algorithm for channel equalization are described in detail in Table 1, including a resampling step [5] every time the *effective sample size* (M_{eff} in the table) goes below a certain threshold (ε). This operation is necessary to prevent the distribution of the weights from becoming too skewed, with all trajectories having negligible weights except for a few of them. Intuitively, the resampling operation consists of discarding those state sequences with very small importance weights, while those with a higher probability are replicated.

5.1. Optimal Importance PMF

The importance function $q[\cdot]$ is chosen by the designer, where the choice is based on a trade off between complexity and performance of the algorithm. In this section, we derive the optimal importance pmf that employs all the information available up to time t in order to propose new samples, i.e.,

$$
q[\mathbf{s}_t|\mathbf{s}_{0:t-1}^{(i)}, y_{0:t}] = p[\mathbf{s}_t|\mathbf{s}_{t-1}^{(i)}, y_{0:t}] \propto p[y_t|\mathbf{s}_t, \mathbf{s}_{t-1}^{(i)}, y_{0:t-1}].
$$

The likelihood in the above equation is shown in (12) (see top of next page), with $\tilde{\mathbf{C}}_t^{(i)^{-1}} = \frac{\mathbf{s}_t \mathbf{s}_t^\top}{\sigma^2} + \mathbf{C}_{t-1}^{(i)^{-1}}$. To be specific, we draw the new sample, $\mathbf{s}_t^{(i)}$, using the set of functions

$$
q[\mathbf{s}_t|\mathbf{s}_{t-1}^{(i)}, y_{0:t}] = \frac{p[y_t|\mathbf{s}_t, \mathbf{s}_{t-1}^{(i)}, y_{0:t-1}]}{p_{t,+} + p_{t,-}} \tag{13}
$$

where $p_{t,+} = p[y_t|s_t = +1, \mathbf{s}_{t-1}^{(i)}, y_{0:t-1}]$ and $p_{t,-} = p[y_t|s_t = -1, \mathbf{s}_{t-1}^{(i)}, y_{0:t-1}]$ are computed according to (12). Correspondingly, the weight update equation becomes

$$
\tilde{w}_t^{(i)} = w_{t-1}^{(i)}(p_{t,+} + p_{t,-})
$$

which therefore, does not depend on the new sample to be added to the trajectory and can be carried out in parallel with the sampling step.

$$p[y_t|\mathbf{s}_t, \mathbf{s}_{t-1}^{(i)}, y_{0:t-1}] = \frac{|\tilde{\mathbf{C}}_t^{(i)}|^{1/2}}{(2\pi\sigma^2|\mathbf{C}_{t-1}^{(i)}|)^{1/2}} e^{-\frac{1}{2}\left[\frac{y_t^2}{\sigma^2} + \overline{\mathbf{h}}_{t-1}^{(i)\top} \mathbf{C}_{t-1}^{(i)^{-1}} \overline{\mathbf{h}}_{t-1}^{(i)} - \left(\frac{\mathbf{s}_t y_t}{\sigma^2} + \mathbf{C}_{t-1}^{(i)^{-1}} \overline{\mathbf{h}}_{t-1}^{(i)}\right)^\top \tilde{\mathbf{C}}_t^{(i)} \left(\frac{\mathbf{s}_t y_t}{\sigma^2} + \mathbf{C}_{t-1}^{(i)^{-1}} \overline{\mathbf{h}}_{t-1}^{(i)}\right)\right]} \quad (12)$$

6. COMPUTER SIMULATIONS

Let us consider a system with $\mathbf{h} = [-0.7 \quad 1]^\top$, $T = 200$ and differentially encoded BPSK symbols. The reason for the differential modulation scheme is to avoid the well-known problem of phase ambiguity in blind equalization [2]. Fig. 1 plots the Bit Error Rate (BER) attained by:

- The linear Minimum Mean Square Error (MMSE) equalizer with known channel. It is a one-shot equalizer consisting of a Wiener matrix filter with dimensions $T \times (T + m - 1)$.
- The Maximum Likelihood Equalizer (MLE) implemented via the Viterbi algorithm with known channel [1].
- The blind MAP equalizer with unknown channel, implemented via the Viterbi algorithm in Section 4.
- The blind MAP equalizer with unknown channel that employs the SIS algorithm described in Section 5 with the optimal importance pmf, $p[\mathbf{s}_t|\mathbf{s}_{0:t-1}^{(i)}, y_{0:t}]$, $M = 100$ particles and resampling steps every time $M_{eff} < 0.2$.

We observe that the MAP equalizer implemented with the Viterbi algorithm has a poor performance in the higher SNR region because it does not necessarily converge to the true MAP estimate, as explained in Section 4. The blind MAP equalizer implemented with the SIS algorithm, instead, presents a much better BER curve that steadily tracks the performance limit given by the MLE. All nonlinear equalizers outperform the linear MMSE receiver.

7. CONCLUSIONS

We have introduced a novel blind equalization algorithm for frequency-selective channels based on a Bayesian formulation of the problem and the SIS methodology. A recursive algorithm that sequentially builds an MC representation of the symbol posterior pmf given the available observations has been derived. At any time, it is possible to draw from this representation, which consists of a set of data samples and corresponding weights, an arbitrarily tight approximation of the MAP estimate of the transmitted symbols. Hence, optimal blind equalization is achieved without explicit estimation of the channel response.

8. REFERENCES

[1] J. G. Proakis, *Digital Communications*, McGraw-Hill, Singapore, 1995.

[2] G. B. Giannakis and S. D. Halford, "Blind fractionally spaced equalization of noisy FIR channels: Direct and adaptive solutions," *IEEE Trans. on Signal Processing*, vol. 45, no. 9, pp. 2277–2292, September 1997.

[3] H. Zamiri-Jafarian and S. Pasupathy, "Adaptive MLSDE using the EM algorithm," *IEEE Trans. Communications*, vol. 47, no. 8, pp. 1181–1193, August 1999.

[4] 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.

[5] 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.

[6] D. Crisan and A. Doucet, "A survey of convergence results on particle filtering," *IEEE Trans. on Signal Processing*, vol. 50, no. 3, March 2002.

[7] J. Geweke, "Bayesian inference in econometric models using Monte Carlo integration," *Econometrica*, vol. 24, pp. 1317–1399, 1989.

T ≡frame size, $M \equiv$ number of trajectories, $i = 1, ..., M$.

Initialization:	$\mathbf{C}_{-1}^{(i)} := \mathbf{C}_{-1}$
	$\overline{\mathbf{h}}_{-1}^{(i)} := \overline{\mathbf{h}}_{-1}, \forall i.$
For $t = 0, 1, ..., T$:	
Importance sampling:	$\mathbf{s}_t^{(i)} \sim q[\mathbf{s}_t\|\mathbf{s}_{t-1}^{(i)}, y_{0:t}]$
Channel update:	$\mathbf{C}_t^{(i)} = \left(\frac{\mathbf{s}_t^{(i)}\mathbf{s}_t^{(i)\top}}{\sigma^2} + \mathbf{C}_{t-1}^{(i)^{-1}}\right)^{-1}$
	$\overline{\mathbf{h}}_t^{(i)} = \mathbf{C}_t^{(i)} \left(\frac{\mathbf{s}_t^{(i)} y_t}{\sigma^2} + \mathbf{C}_{t-1}^{(i)^{-1}}\right)$
Weight update:	$\tilde{w}_t^{(i)} = w_{t-1}^{(i)} \frac{p[y_t\|\mathbf{s}_{0:t}^{(i)}, y_{0:t-1}]}{q_t[\mathbf{s}_t^{(i)}\|\mathbf{s}_{t-1}^{(i)}, y_{0:t}]}$
Weight normalization:	$w_t^{(i)} = \frac{\tilde{w}_t^{(i)}}{\sum_{j=1}^M \tilde{w}_t^{(j)}}$
Resampling if:	$M_{eff} = \frac{1}{\sum_{j=1}^M w_t^{(j)^2}} < \varepsilon$,
	$0 < \varepsilon < 1$.

Table 1. SIS algorithm with resampling.

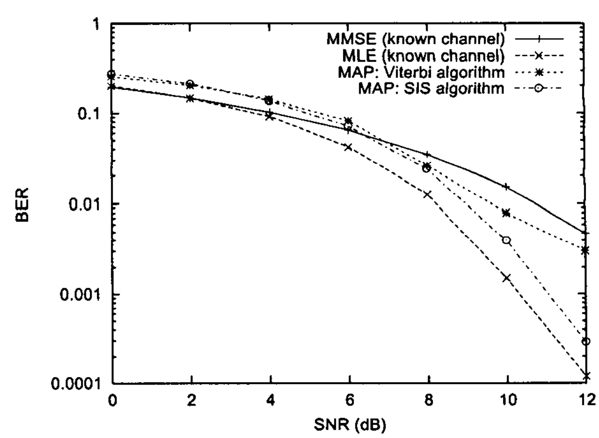

Fig. 1. BER vs. SNR for several equalizers.