PARTICLE FILTER ALGORITHMS FOR JOINT BLIND EQUALIZATION/DECODING OF CONVOLUTIONALLY CODED SIGNALS

Claudio J. Bordin Jr. and Luiz A. Baccalá

Escola Politécnica, Universidade de São Paulo, 05508-900, São Paulo - SP, Brazil.
E-mail: `bordin,baccala@lcs.poli.usp.br`

ABSTRACT

This work introduces the use of particle filters for joint blind equalization/decoding of convolutionally coded signals transmitted over frequency selective channels. As in the equalization-only case, we show how to evaluate the optimal importance function recursively via a bank of Kalman filters. Numerical simulation investigations using both stochastic and deterministic particle selection strategies show the outstanding superiority of the deterministic joint equalization/decoding method over approaches that perform blind equalization using particle filters prior to optimal decoding.

1. INTRODUCTION

Either because of slow convergence or excessive complexity, blind equalization still remains a practical challenge despite the many approaches developed in the last two decades (see [1] for a review). In this context, particle filters [2] - numerical techniques for the solution of Bayesian estimation problems - can play a major role as a compromise solution, that balances robustness and convergence speed with complexity of implementation.

Though not exactly new in an equalization-only setting [3], particle filters failed to attract much attention until recently, possibly due to the scarcity of theoretical results coupled with the need for sufficiently powerful computers that are only now becoming more widely available. The importance of particle filters is derived from their generality which allows them to handle elaborate signal models that are otherwise practically intractable (CDMA signal reception [4] for example).

As opposed to previous works [3] addressing equalization-only scenarios, we here propose the idea of using particle filters to jointly blind equalize and decode convolutionally coded signals transmitted over frequency selective channels, in an extension of Punskaya's work [5]. A further innovative aspect is our use of noncoherently noncatastrophic codes [6] to successfully deal with the phase ambiguities inherent to the blind equalization problem.

Claudio J Bordin Jr. was supported by the FAPESP grant 02/11457-7.

After stating the estimation problem (Sec. 2) and overviewing particle filters (Sec. 3), we show how to obtain the densities applicable to solving the joint blind equalization/decoding problem (Sec. 4). This is followed in Section 5 by numerical illustrations leaving our conclusions to Section 6.

2. SIGNAL MODEL AND PROBLEM STATEMENT

Consider a convolutionally coded digital communication system that transmits BPSK symbols over an additive noise frequency selective channel. Denoting the transmitted bit sequence by x_k and the code rate by $1/R$, the transmitted (binary) symbol sequence c_k is obtained as

$$c_{Rm+n} = \left(\sum_{l=0}^{K} x_{m-l} p_l^n \right) \bmod 2, \quad 0 \leq n < R, \qquad (1)$$

where K is the convolutional code constraint length and p_l^n the code generator coefficients. The standard transmitted BPSK signal is obtained by $s_k = 2\, c_k - 1$.

We assume a linear and time-invariant FIR transmission channel and perfect receiver synchronization, so that baud rate samples y_k of the received signal are expressed by the base-band equivalent model

$$y_k = \sum_{l=0}^{L-1} h_l s_{k-l} + v_k \ , \qquad (2)$$

where h_l is the channel impulse response, L its duration in symbol intervals and v_k the additive noise assumed to be zero-mean white circular gaussian of variance σ_v^2.

In this work, we aim at obtaining MAP estimates $\hat{x}_k$ of the transmitted bits given the observed data, i.e.,

$$\hat{x}_k = \arg\max_{x_k} p(x_k | \underline{y}_{0:k+d}), \qquad (3)$$

where $d \geq 0$ and $\underline{y}_{0:k} \triangleq (y_0, \ldots, y_{(k+1)R-1})$. In next sections we describe how particle filters can be employed to obtain estimates of the posterior densities $p(x_k | \underline{y}_{0:k+d})$ for the considered signal model.

3. PARTICLE FILTERS

Many estimation problems in discrete-time signal processing, including blind equalization, can be stated as that of tracking the state x_k of a dynamic system

$$\begin{cases} x_{k+1} = f_k(x_0, \ldots, x_k, u_k) \\ y_k = g_k(x_0, \ldots, x_k, v_k) \end{cases} \quad (4)$$

from the observation of its output $y_{0:k}$ sequence, where f_k and g_k, in full generality, may be nonlinear, time-varying functions and u_k and v_k stand respectively for system driving and observation noise.

The Bayesian solution to (4) involves finding the posterior density $p(x_k|\underline{y}_{0:k})$, which can be accomplished in closed-form only in a few restrictive cases. Particle filters provide iterative approximations to $p(x_{0:n}|\underline{y}_{0:n})$ (from which $p(x_n|\underline{y}_{0:n})$ can be obtained by marginalization) via a weighted sum of Dirac measures which approximates densities of arbitrary shape [2].

3.1. Principles

Most particle filtering algorithms rest on the importance sampling principle, whereby the probability density $p(x_{0:k}|\underline{y}_{0:k})$ can be consistently approximated [2] as the weighted sum of Dirac measures

$$\hat{p}(x_{0:k}|\underline{y}_{0:k}) = \frac{\sum_{i=1}^{M} w_k^{(i)} \delta_{x_{0:k}}(x_{0:k}^{(i)})}{\sum_{i=1}^{M} w_k^{(i)}}\ , \quad (5)$$

where $x_{0:k}^{(i)}$, $1 \le i \le M$, are independent samples (particles) from an arbitrary density $\pi(x_{0:k}^{(i)}|\underline{y}_{0:k}) > 0$ (importance function), with $\delta_{x_{0:k}}(x_{0:k}^{(i)})$ denoting Dirac measures in the variable $x_{0:k}$ centered in $x_{0:k}^{(i)}$, and where $w_k^{(i)} \triangleq \pi(x_{0:k}^{(i)}|\underline{y}_{0:k})/p(x_{0:k}^{(i)}|\underline{y}_{0:k})$.

The key to the renewed interest in particle filtering comes from the seminal work of Gordon *et al.* [7] showing that if the importance function $\pi(x_{0:k}|\underline{y}_{0:k})$ can be factored as

$$\pi(x_{0:k}|\underline{y}_{0:k}) = \pi(x_k|x_{0:k-1}, \underline{y}_{0:k})\pi(x_{0:k-1}|\underline{y}_{0:k-1}), \quad (6)$$

each element of the sequence $x_{0:k}^{(i)}$ can be sampled sequentially from the densities $\pi(x_m|x_{0:m-1}^{(i)}, \underline{y}_{0:m})$, $0 \le m \le k$. Likewise the weights $w_k^{(i)}$ can be sequentially updated, since from Bayes's law it follows that

$$p(x_{0:k}|\underline{y}_{0:k}) = p(x_{0:k-1}|\underline{y}_{0:k-1}) \frac{p(x_k, \underline{y}_k|x_{0:k-1}, \underline{y}_{0:k-1})}{p(\underline{y}_k|\underline{y}_{0:k-1})}, \quad (7)$$

which in combination with (6) produces

$$w_n^{(i)} \propto w_{k-1}^{(i)} \frac{p(x_k^{(i)}, \underline{y}_k|x_{0:k-1}^{(i)}, \underline{y}_{0:k-1})}{\pi(x_k^{(i)}|x_{0:k-1}^{(i)}, \underline{y}_{0:k})}\ . \quad (8)$$

The choice of the importance function, though essentially arbitrary, impacts the algorithm performance: one can show that $\pi(x_k|x_{0:k-1}, \underline{y}_{0:k}) = p(x_k|x_{0:k-1}, \underline{y}_{0:k})$ is optimal in minimizing the conditional variance of the (unnormalized) weights $w_k^{(i)}$ hence improving algorithm performance. After some iterations, due to a phenomenon known

Table 1. Stochastic Particle Filtering Algorithm

For $k = 0, \ldots, n$,
For $i = 1, \ldots, M$
-Draw $x_k^{(i)}$ from $\pi(x_k|x_{0:k-1}^{(i)}, y_{0:k})$.
-Update the weights as in (8).
-Normalize the weights.
-Estimate $p(x_{0:k}|y_{0:k})$ as in (5).

as *degeneracy*, particle filters can lead to particles $x_{0:k}^{(i)}$ that have negligible weights $w_k^{(i)}$ what severely compromise (5)'s estimation. To overcome this, the use of a *selection* scheme is mandatory when sample quality falls below a predefined threshold.

Most particle filter algorithms in the literature employ stochastic selection (namely, multinomial or residual resampling [2]). Recently, [5] proposed algorithms that employ deterministic selection, hence the name "deterministic particle filters". Despite the lack of rigorous convergence results, their chief interest is their excellent performance in certain situations. Motivation for them comes from observing that the optimal importance function can be expressed as

$$p(x_k|x_{0:k-1}, \underline{y}_{0:k}) = \frac{p(x_k, \underline{y}_k|x_{0:k-1}, \underline{y}_{0:k-1})}{\sum_{x_k} p(x_k, \underline{y}_k|x_{0:k-1}, \underline{y}_{0:k-1})}, \quad (9)$$

i.e., to determine $p(x_k|x_{0:k-1}, \underline{y}_{0:k})$, one must evaluate the term $p(x_k, \underline{y}_k|x_{0:k-1}, \underline{y}_{0:k-1})$ for each possible value of x_k. Hence, supposing that u_k is a discrete vector variable with D different possible values, at each iteration, the particle filter discards $M(D-1)$ "candidates" $x_k^{(i)}$ ($D-1$ for each particle). Thus the idea of performing particle selection at each iteration [8] as described in Table 2 comes to the fore.

As a final remark, it is worth mentioning the particle filtering algorithms described so far can be easily extended to provide fixed-lag smoothed estimates, since for $d > 0$ [9]:

$$p(x_{0:k}|\underline{y}_{0:k+d}) \approx \frac{\sum_{i=1}^{M} w_{k+d}^{(i)} \delta_{x_{0:k}}(x_{0:k}^{(i)})}{\sum_{i=1}^{M} w_{k+d}^{(i)}}. \quad (10)$$

4. BLIND EQUALIZATION AND DECODING

All densities needed to obtain estimates of $p(x_{0:k}|\underline{y}_{0:k})$ via both the stochastic particle filter (employing an optimal im-

Table 2. Deterministic Particle filter Algorithm

For $k = 0, \ldots, n$
  For $i = 1, \ldots, M$
    For $j = 1, \ldots, D$
      -Calculate $w^{(i,j)} = w_k^{(i)}\ p(x_k = x^{(j)}, y_k | x_{0:k-1}, y_{0:k-1})$.
  -Determine $(I, J)_t$, the sequence of M pairs (i, j) corresponding to the M largest $w^{(i,j)}$.
  For $t = 1, \ldots, M$
    -Make $x_{0:k-1}^{(t)} = x_{0:k-1}^{(I_t)}$ and $x_k^{(t)} = x^{(J_t)}$.
    -Make $w_k^{(t)} = w^{(I,J)_t}$.
  -Normalize the weights so that $\sum_{i=1}^{M} w_k^{(i)} = 1$.
  -Estimate $p(x_{0:k}|y_{0:k})$ as in (5).

portance function) and the deterministic particle filter described in Sec. 3.1 can be obtained from $p(x_k, \underline{y}_k | x_{0:k-1}, \underline{y}_{0:k-1})$ using (9). Assuming that x_k is an IID sequence one can show that

$$p(x_k, \underline{y}_k | x_{0:k-1}, \underline{y}_{0:k-1}) = p(\underline{y}_k | x_{0:k}, \underline{y}_{0:k-1})\ p(x_k)\ . \quad (11)$$

Moreover, as the bit sequence x_k uniquely defines the transmitted symbol sequence s_k, (11) implies (assuming that $x_k = 0$ for $k < 0$) that

$$p(\underline{y}_k | x_{0:k}, \underline{y}_{0:k-1}) = p(\underline{y}_k | S_{0:(k+1)R-1}, \underline{y}_{0:k-1})\ , \quad (12)$$

where $S_k \triangleq [\ s_k\ \ldots\ s_{k-L+1}\ ]^T$, and $(s_0, \ldots, s_{(k+1)R-1})$ is the symbol sequence corresponding to the bit sequence $x_{0:k}$. The density on the r.h.s. of (12) further decomposes as

$$p(\underline{y}_k | S_{0:(k+1)R-1}, \underline{y}_{0:k-1}) = \prod_{j=Rk}^{R(k+1)-1} p(y_j | S_{0:j}, y_{0:j-1})\ . \quad (13)$$

In (13) we exploited the fact that $p(y_j|S_{0:j}, y_{0:j-1}) = p(y_j| S_{0:k}, y_{0:j-1})$, $k > j$. In order to determine $p(y_j|S_{0:j}, y_{0:j-1})$ note that using definitions above, (2) can be rewritten as

$$\begin{cases} S_{j+1} &= A S_j + e_1\ s_{j+1} \\ y_j &= h^H S_j + v_j \end{cases} \quad (14)$$

where A is a $(L \times L)$ shift matrix (all entries zero, except the first subdiagonal, whose entries are ones), $e_1 = [\ 1\ 0\ \cdots\ 0\ ]^T$ and $h = [\ h_0\ \cdots\ h_{L-1}\ ]^T$.

From (14) one can see that y_j is conditionally gaussian given S_j and h. Assuming that the parameter h has gaussian prior distribution, and exploiting the fact it is conditionally gaussian given $S_{0:j}$ and $y_{0:j-1}$, it can be integrated out [4], resulting in

$$p(y_j|S_{0:j}, y_{0:j-1}) = \mathcal{N}_C\left(y_j \mid \hat{h}_{j-1}^H S_j\ ;\ S_j^H \Sigma_{j-1} S_j + \sigma_v^2\right)\ , \quad (15)$$

where $\hat{h}_j$ and Σ_j, respectively the conditional mean and variance of h are obtained by means of conventional Kalman filter iterations:

$$\hat{h}_j = \hat{h}_{j-1} + \frac{\underline{y}_j - S_j^H \hat{h}_{j-1}}{S_j^H \Sigma_{j-1} S_j + \sigma_v^2} \Sigma_{j-1} S_j\ . \quad (16)$$

$$\Sigma_j = \Sigma_{j-1} - \frac{\Sigma_{j-1} S_j S_j^H \Sigma_{j-1}}{S_j^H \Sigma_{j-1} S_j + \sigma_v^2}\ . \quad (17)$$

From (11)-(17), one can see that evaluating (11) for each particle requires the evaluation of $2R$ Kalman filter steps. Thus, if M particles are employed, the proposed algorithms require the evaluation of $2MR$ Kalman filter steps per bit, twice the complexity of an equalization-only algorithm (operating at the symbol rate).

5. SIMULATIONS

We implemented blind equalization and decoding algorithms employing particle filters for the signal model in Section 2, and evaluated their performance using Monte Carlo simulations that measured the bit error rates (BER) over 200 independent realizations. In our illustrative simulations, we used the channel $h = [\ 0.41\quad -0.82\quad 0.41\ ]^T$ and adopted two $R = 3$ noncoherently noncatastrophic convolutional codes [6] with constraint length $K = 3$ and $K = 4$ (with coefficients given in octal notation by (7,5,2) and (17,12,4), respectively) in order to avoid phase ambiguities.

To compute the mean BER the algorithms processed 150 message bits in each realization, and employed a fixed smoothing lag of 25 samples. The particles' initial states $S_{-1}^{(i)}$ were drawn from IID equiprobable ± 1 r.v., and we assumed that $\Sigma_{-1}^{(i)} = I$ and $h_{-1}^{(i)} \sim \mathcal{N}_C(h|0; I)$. The BER obtained by the proposed algorithms[1] as a function of the SNR (signal-to-noise ratio) using $M = 100$ particles and the (17,12,4) code is shown in Fig. 1. For comparison, we also show the performance obtained i) by the optimal MLSE equalizer (Viterbi) followed by a hard-decision decoder and ii) by a particle filter based blind equalization-only algorithm [4] (employing differential encoding) followed by an optimal soft decoder (based on the BCJR algorithm [10]).

As one can readily verify, the deterministic joint algorithm outperformed all the others. The stochastic algorithm, in turn, performed on average equivalently to the concatenated scheme, "failing" (i.e., obtaining BER of about 50%) in some realizations even at high SNR.

Figure 2 depicts the bit error performances of the proposed (deterministic) joint equalization and decoding algorithm employing $M = 50$ and $M = 100$ particles and the

[1]For the stochastic algorithm, we adopted that a multinomial resampling step that is carried out whenever the "effective sample size" [2] falls below 20%.

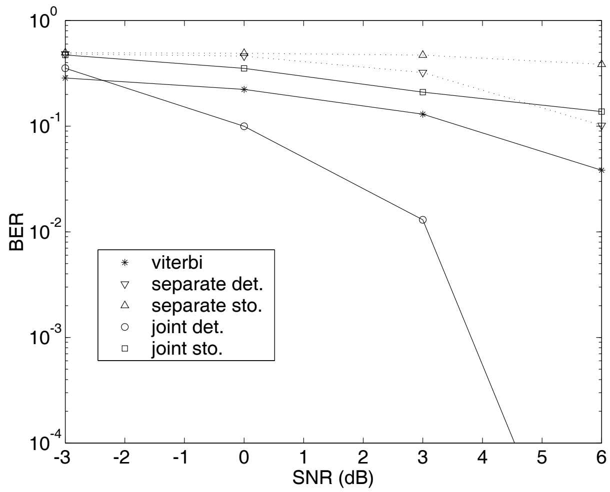

Fig. 1. Performance of the proposed joint equalization and decoding algorithms and of an alternative separate scheme using deterministic and stochastic particle filters as a function of the SNR.

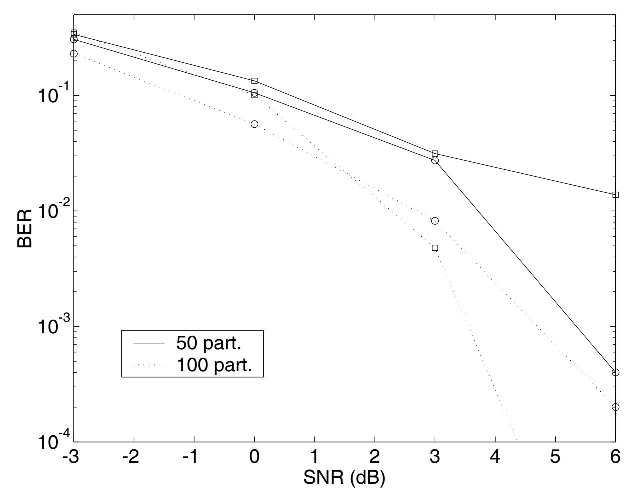

Fig. 2. Performance of the proposed (deterministic) joint equalization algorithm employing $M = 50$ and $M = 100$ particle for the codes (17,12,4) (square) and (7,5,2) (circle) as a function of the SNR.

(17,12,4) (square) and (7,5,2) (circle) codes. As one can verify, the performances obtained with $M = 100$ particles are very similar for both codes. With $M = 50$, however, the performance of the (17,12,4) code is worse, arguably because the number of possible states that need to be sampled is larger in this case.

6. CONCLUSIONS

After introducing the novel use of particle filtering algorithms for joint blind equalization and message decoding over frequency selective channels that employ convolutional codes, this work showed that deterministic particle filtering is the best alternative, greatly outperforming both joint stochastic particle filters and methods that equalize and decode messages separately.

7. REFERENCES

[1] Z. Ding and Y. Li, Eds., *Blind Equalization and Identification*, Marcel Dekker, New York, 2001.

[2] A. Doucet, J. de Freitas, and N. Gordon, Eds., *Sequential Monte-Carlo Methods in Practice*, Springer-Verlag, Berlin, Germany, 2000.

[3] J. S. Liu and R. Chen, "Blind deconvolution via sequential imputations," *Journal of the American Statistical Association*, vol. 90, no. 430, pp. 567–576, june 1995.

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

[5] E. Punskaya, *Sequential Monte Carlo Methods for Digital Communications*, Ph.D. thesis, University of Cambridge, 2003.

[6] Y. Kofman, E. Zehavi, and S. Shamai, "nd-convolutional codes - part ii: Structural analysis," *IEEE Transactions on Information Theory*, vol. 43, no. 2, pp. 576–589, march 1997.

[7] N. J. Gordon, D. J. Salmond, and A. F. M. Smith, "Novel-approach to nonlinear non-gaussian bayesian state estimation," *IEE Proceedings-F Image & Signal Processing*, vol. 140, no. 2, pp. 107–113, april 1993.

[8] B. Dong, X. Wang, and A. Doucet, "A new class of soft mimo demodulation algorithms," *IEEE Transactions on Signal Processing*, vol. 51, no. 11, pp. 2752–2763, november 2003.

[9] T. Clapp, *Statistical Methods for the Processing of Communication Data*, Ph.D. thesis, University of Cambridge, 2000.

[10] L. R. Bahl, J. Cocke, F. Jelinek, and J. Raviv, "Optimal decoding of linear codes for minimizing symbol error rate," *IEEE Transactions on Information Theory*, vol. 20, no. 2, pp. 284–287, march 1974.