A RAO-BLACKWELLIZED PARTICLE FILTER FOR BLIND EQUALIZATION OF FREQUENCY-SELECTIVE CHANNELS WITH UNKNOWN ORDER AND NOISE VARIANCE

Claudio J. Bordin Jr. and Marcelo G. S. Bruno

Instituto Tecnológico de Aeronáutica
São José dos Campos, 12228-900, Brazil. E-mail: {bordin,bruno}@ita.br

ABSTRACT

We propose in this paper a new particle filtering algorithm for blind equalization of FIR frequency-selective communication channels corrupted by additive Gaussian noise, assuming that both the channel order and noise variance are unknown. The proposed algorithm integrates out analytically the unknown parameters using a modified sequential importance sampling technique. We verify via numerical simulations that the proposed method leads to near optimal performance, greatly outperforming traditional methods under noise variance mismatch.

Index Terms— Adaptive equalizers, Sequential estimation, Monte Carlo methods, Bayes procedures.

1. INTRODUCTION

Particle filters have been extensively applied to solving blind equalization [1] [2] and related problems [3] on frequency-selective channels contaminated by Gaussian noise. However, differently from methods based on other techniques like MCMC [4], currently available particle filter methods require exact knowledge of the channel's additive noise variance, and can be shown to perform rather poorly when incorrect values are assumed for that quantity.

In this work, we fill this gap, developing an algorithm for blind equalization under unknown variance additive Gaussian noise. We start by developing an alternative method for optimal importance function and weight computation, and next evaluate the probability distributions needed for its application to the adopted signal model. While the proposed alternative method requires much simpler analytical evaluations, we verify that it does not incur in increased computation complexity or performance degradation compared to the existing methods.

The remainder of this article is organized as follows: in Sec. 2 we describe the adopted signal model and estimation objectives. After briefly reviewing elements of particle filter theory in Sec. 3, we describe the proposed sequential importance sampling method and present a recursive method for evaluating the needed probability densities in Sec. 4. Finally, in Sec. 5, we evaluate the performance of the proposed methods via numerical simulations, leaving our conclusions to Sec. 6.

The first author's work was funded by FAPESP, Brazil.

2. PROBLEM DESCRIPTION

Let b_n be an independent, identically distributed (i.i.d.) binary bit sequence and $s_n \in \{\pm 1\}$ the corresponding (i.i.d.) differentially encoded symbols [5]. Our aim in this work is to develop a recursive method for obtaining MAP smoothed estimates

$$\hat{b}_{n-d} = \arg\max_{b_{n-d}} p(b_{n-d}|y_{0:n}) \ , \qquad (1)$$

for $d \in \mathbb{N}^+$. The observations $y_{0:n} \triangleq \{y_0, ..., y_n\}$ are assumed to be the output of the additive noise frequency selective FIR channel

$$y_n = \sum_{i=0}^{L-1} h_i s_{n-i} + v_n \ , \qquad (2)$$

where L is the channel order, restricted to a known set $\mathcal{L}$, h_i are the (time-invariant) unknown channel impulse response terms, and v_n represents a Gaussian i.i.d zero-mean random noise process of unknown variance σ^2. For convenience, we rewrite (2) so that y_n is expressed as the output of the dynamic system

$$\begin{cases} S_{n+1} &= \mathcal{D}S_n + [s_{n+1}\ 0 \ldots 0]^T \\ y_n &= h^T S_n + v_n \end{cases} \qquad (3)$$

where $h \triangleq [h_0 \ldots h_{L-1}]^T$, $S_n \triangleq [s_n \ldots s_{n-L+1}]^T$, and $\mathcal{D} \in \mathbb{R}^{L\times L}$ is a displacement matrix[1]. The unknown parameters h, σ^2 and L are assumed to be, in turn, distributed a priori according to

$$p(\sigma^2) = \mathcal{IG}(\sigma^2|\alpha;\beta) \triangleq \frac{\beta^\alpha}{\Gamma(\alpha)}(\sigma^2)^{-(\alpha+1)} \exp\left(-\frac{\beta}{\sigma^2}\right) \quad (4)$$

$$p(L) = \mathcal{TP}(L|\lambda;\mathcal{L}) \triangleq \frac{\lambda^L}{c_{\mathcal{L},\lambda}\ L!}\mathcal{I}\left\{L \in \mathcal{L}\right\} \qquad (5)$$

$$p(h|L,\sigma^2) = \mathcal{N}_L(h|0; I\sigma^2/\epsilon^2) \qquad (6)$$

[1]An all-zero matrix except for its first lower diagonal, whose entries are unitary.

where $\mathcal{N}$ stands for a normal, $\mathcal{IG}$ an inverted-gamma, and $\mathcal{TP}$ a truncated Poisson distribution, $\mathcal{I}\{A\}$ for the indicator of event A and $c_{\mathcal{L},\lambda}$ is a normalizing term that does not depend on L.

3. PARTICLE FILTERS

Particle filters [6] are now a well-established numerical technique for solving stochastic filtering problems, i.e., estimating a sequence of hidden variables $S_{0:n}$ from the observations $y_{0:n}$, approximating the inferred variable posterior density via the weighted sum

$$p(S_{0:n}|y_{0:n}) \approx \frac{1}{\sum_{i=1}^{P} w_n^{(i)}} \sum_{i=1}^{P} w_n^{(i)} \delta\left(S_{0:n} - S_{0:n}^{(i)}\right) , \quad (7)$$

where $P \gg 1$, $\delta(.)$ denotes the Dirac delta function, $S_{0:n}^{(i)}$ are the particles, random samples of $\pi(S_{0:n}^{(i)}|y_{0:n})$ (importance function) and

$$w_n^{(i)} \triangleq p(S_{0:n}^{(i)}|y_{0:n})/\pi(S_{0:n}^{(i)}|y_{0:n}) \quad (8)$$

their respective weights. By defining the importance function as a product of marginals

$$\pi(S_{0:n}^{(i)}|y_{0:n}) \triangleq \prod_{j=0}^{n} \pi(S_j^{(i)}|S_{0:j-1}^{(i)}, y_{0:j}) , \quad (9)$$

recursive approximations to $p(S_{0:n}|y_{0:n})$ can be obtained as

$$\begin{aligned} S_n^{(i)} &\sim \pi(S_n^{(i)}|S_{0:n-1}^{(i)}, y_{0:n}) && (10) \\ w_n^{(i)} &\propto w_{n-1}^{(i)} \frac{p(S_n^{(i)}, y_n|S_{0:n-1}^{(i)}, y_{0:n-1})}{\pi(S_n^{(i)}|S_{0:n-1}^{(i)}, y_{0:n})} && (11) \end{aligned}$$

A problem with the traditional particle filter formulation of (10)-(11) is that $p(S_n^{(i)}, y_n|S_{0:n-1}^{(i)}, y_{0:n-1})$ may not be always simple to determine, which could impede the application of this inference technique to many specific inference problems.

4. PROPOSED METHOD

Supposing that $p(S_{0:n}^{(i)}, y_{0:n})$ can be determined for all $n > 0$, it can be verified that

$$\begin{aligned} w_n^{(i)} &= w_{n-1}^{(i)} \frac{p(S_{0:n}^{(i)}|y_{0:n})}{p(S_{0:n-1}^{(i)}|y_{0:n-1})\pi(S_n^{(i)}|S_{0:n-1}^{(i)}, y_{0:n})} && (12) \\ &\propto w_{n-1}^{(i)} \frac{p(S_{0:n}^{(i)}, y_{0:n})}{p(S_{0:n-1}^{(i)}, y_{0:n-1})\pi(S_n^{(i)}|S_{0:n-1}^{(i)}, y_{0:n})} && (13) \end{aligned}$$

While (12) is a consequence of (8) and (9), the proportionality sign in (13) accounts for the multiplying term $p(y_n|y_{0:n-1})$, common to all particles. The optimal importance function [6], in turn, can be obtained as

$$p(S_n^{(i)}|S_{0:n-1}^{(i)}, y_{0:n}) = \frac{p(S_{0:n}^{(i)}, y_{0:n})}{\sum_{\forall S_n^{(i)}} p(S_{0:n}^{(i)}, y_{0:n})}. \quad (14)$$

Obviously, this alternative formulation only makes sense when $p(S_{0:n}^{(i)}, y_{0:n})$ can be recursively determined. It is also worth stressing that although this method requires that $p(S_{0:n}^{(i)}, y_{0:n})$ can be directly evaluated, approximations such (7) are still needed as the number of possible sequences $S_{0:n}^{(i)}$ grows exponentially with n for the adopted signal model.

4.1. Determination of $p(S_{0:n}, y_{0:n})$

To determine[2] $p(S_{0:n}, y_{0:n})$, we first observe[3] that

$$p(S_{0:n}, y_{0:n}) = \sum_{L\in\mathcal{L}} \int_{\mathbb{R}^+} \int_{\mathbb{R}^L} p(S_{0:n}, y_{0:n}, L, h, \sigma^2)\, d\sigma^2\, dh. \quad (15)$$

Under the model assumptions described in Sec. 2, the integrand on the right-hand side (r.h.s) of (15) decomposes as

$$\begin{aligned} p(S_{0:n}, y_{0:n}, L, h, \sigma^2) &= p(y_{0:n}|S_{0:n}, L, h, \sigma^2) \\ &\quad \times p(S_{0:n})p(h|L, \sigma^2)p(L)p(\sigma^2) . \end{aligned} \quad (16)$$

The first term on the r.h.s of (16) is an $(n+1)$-variate Gaussian density with covariance matrix $I\sigma^2$. The distribution $p(S_{0:n})$, in turn, assumes the same value for all binary vector sequences that obey (3) and, therefore, is omitted in the sequel. Consequently, we can rewrite (16) as

$$\begin{aligned} p(S_{0:n}, y_{0:n}, L, h, \sigma^2) &\propto \left(\sigma^2\right)^{-\left(\frac{n+L+1}{2}+\alpha+1\right)} \\ &\quad \times (2\pi)^{-\left(\frac{n+L+1}{2}\right)} \frac{(\epsilon\lambda)^L}{L!} \exp\left\{-\frac{1}{2\sigma^2}\left(2\beta + Q_n\right)\right\} , \end{aligned} \quad (17)$$

where

$$\begin{aligned} Q_n &= \|\overline{y}_n - \overline{S}_n h\|^2 + \epsilon^2\|h\|^2 && (18) \\ \overline{S}_n &= [S_n \dots S_0]^T && (19) \\ \overline{y}_n &= [y_n \dots y_0]^T && (20) \end{aligned}$$

To integrate the channel parameter vector h, we must rewrite Q_n as a quadratic function of h

$$Q_n = \|h - \tilde{h}_n\|^2_{\Sigma_n^{-1}} + R_n \quad (21)$$

in which $\tilde{h}_n$, R_n and Σ_n do not depend on h. After some manipulations, these quantities can be determined as

$$\begin{aligned} \Sigma_n &= \left(\overline{S}_n^T\overline{S}_n + I_L\epsilon^2\right)^{-1} && (22) \\ \tilde{h}_n &= \Sigma_n\overline{S}_n^T\overline{y}_n && (23) \\ R_n &= \overline{y}_n^T\overline{y}_n - \overline{y}_n^T\overline{S}_n\Sigma_n\overline{S}_n^T\overline{y}_n && (24) \end{aligned}$$

[2] For clarity, we drop in this section the superscript index (i).

[3] Similar derivations can be found in [4] and [7]. Both of them, however, consider the model order L known and do not completely describe recursive evaluation methods.

Substituting (22)-(24) into (17), we get that

$$p(S_{0:n}, y_{0:n}, L, \sigma^2) \propto \left(\sigma^2\right)^{-\left(\frac{n+L+1}{2}+\alpha+1\right)} \times (2\pi)^{-\left(\frac{n+L+1}{2}\right)} \frac{(\epsilon\lambda)^L}{L!} \exp\left\{-\frac{1}{2\sigma^2}\left(2\beta + R_n\right)\right\} \times \int_{\mathbb{R}^L} \exp\left(-\frac{1}{2}\left\|h - \tilde{h}_n\right\|^2_{\sigma^{-2}\Sigma_n^{-1}}\right) dh \; . \tag{25}$$

The above integral evaluates to $\left(2\pi\sigma^2\right)^{L/2} |\Sigma_n|^{1/2}$, from which we obtain

$$p(S_{0:n}, y_{0:n}, L, \sigma^2) \propto$$

$$\propto C_n(L) \left(\sigma^2\right)^{-\left(\frac{n+1}{2}+\alpha+1\right)} \exp\left\{-\frac{1}{\sigma^2}\left(\beta + \frac{R_n}{2}\right)\right\} \tag{26}$$

$$\propto C_n(L)\, \Gamma(\alpha_n)\, \beta_n^{-\alpha_n}\, \mathcal{IG}(\sigma^2|\alpha_n; \beta_n) \; ,$$

where

$$\alpha_n = \alpha + (n+1)/2 \tag{27}$$

$$\beta_n = \beta + R_n/2 \tag{28}$$

$$C_n(L) = (2\pi)^{-\left(\frac{n+1}{2}\right)} |\Sigma_n|^{\frac{1}{2}} (\epsilon\lambda)^L / L! \tag{29}$$

Integrating out σ^2 in (26) amounts to discarding the inverted-gamma density, which finally leads to

$$p(S_{0:n}, y_{0:n}, L) \propto \frac{|\Sigma_{n,L}|^{\frac{1}{2}} (\epsilon\lambda)^L}{(\beta_{n,L})^{\alpha_n} L!} \; , \tag{30}$$

where the notation $\Sigma_{n,L}$, $\beta_{n,L}$ was introduced to stress the dependence of these variables on the order L. Finally, $p(S_{0:n}, y_{0:n})$ can be obtained by summing (30) for all $L \in \mathcal{L}$.

4.2. Recursive Evaluation

To recursively evaluate (30), one needs to obtain recursive expressions for $|\Sigma_{n,L}|$ and $R_{n,L}$ (needed to determine $\beta_{n,L}$). To this aim, observe that as a consequence of (19)-(20)

$$R_{n,L} = R_n^Y - (R_{n,L}^{SY})^T \Sigma_{n,L} R_{n,L}^{SY} \tag{31}$$

where

$$R_n^Y = R_{n-1}^Y + |y_n|^2 \tag{32}$$

$$R_{n,L}^{SY} = R_{n-1,L}^{SY} + S_{n,L} y_n \tag{33}$$

From (22), one can easily verify that

$$\Sigma_{n,L}^{-1} = \Sigma_{n-1,L}^{-1} + S_{n,L} S_{n,L}^T \; , \tag{34}$$

which after some manipulations [8] leads to

$$\Sigma_{n,L} = \Sigma_{n-1,L} - \frac{\Sigma_{n-1,L} S_{n,L} S_{n,L}^T \Sigma_{n-1,L}}{1 + S_{n,L}^T \Sigma_{n-1,L} S_{n,L}} \tag{35}$$

$$|\Sigma_{n,L}| = |\Sigma_{n-1,L}| \frac{1}{1 + S_{n,L}^T \Sigma_{n-1,L} S_{n,L}} \tag{36}$$

As a consequence of (32)-(36), (31) can be evaluated at a complexity $\mathcal{O}(P|\mathcal{L}|L^2)$, roughly the same of the method described in [1] if the channel order is assumed known. Finally, observe that for the recursions (32)-(36) to be consistent with previous assumptions, we must set $|\Sigma_{-1,L}| = \epsilon^L$, $\Sigma_{-1,L} = I_L \epsilon^2$, $R_{-1,L}^{SY} = 0$ and $R_{-1}^Y = 0$.

4.3. Data estimation

As each particle $S_{0:n}^{(i)}$ uniquely defines a corresponding bit sequence $b_{0:n}^{(i)}$ (by differentially decoding $s_{-1:n}^{(i)}$), one can verify that the latter sequence is a sample of $p(b_{0:n}|y_{0:n})$, which leads to an approximation of this distribution via (7). Smoothed estimates [9] of b_n can then be obtained as

$$\mathbb{P}(b_{n-d} = B|y_{0:n}) \approx \frac{\sum_{i=1}^P w_n^{(i)} \mathcal{I}\left\{b_{n-d,0:n}^{(i)} = B\right\}}{\sum_{j=1}^P w_n^{(j)}} \; . \tag{37}$$

where $b_{n-d,0:n}^{(i)}$ stands for the $(n-d)$-th element of the sequence $b_{0:n}^{(i)}$ (as available at time n) and $B \in \{0, 1\}$. The resulting estimation algorithm is summarized in Table 1.

5. SIMULATION RESULTS

To evaluate the performance of the proposed blind equalization method, we carried out numerical simulations evaluating the BER (bit error rate) as a function of the SNR (signal-to-noise ratio), defined as

$$SNR \triangleq 10 \log_{10} \|h\|^2/\sigma^2.$$

Simulations included 300 independent realizations, each consisting of a block of 250 i.i.d binary symbols. BER estimation was made after discarding the first 100 symbols to allow for algorithm convergence.

In the following simulations, we employed the $L = 3$ channel $h = [0.41 \;\; -0.82 \;\; 0.41]^T$ and assumed a non-informative prior density for the noise variance ($\alpha = \beta = 0$). The channel variance and order hyperparameters were empirically set to $\epsilon = 1$ and $\lambda = 0.001$. The filter employs $P = 300$ particles and a smoothing lag of $d = 5$ samples, and the resampling threshold was set to 90% of the effective sample size.

Figure 1 shows the results obtained for the algorithm in Table 1, with known (circle) and unknown channel order (square), with $\mathcal{L} = \{2, 3, 4\}$. In the same figure, we display the performance of the algorithm described in [1] (known channel order), operating with correct (triangles down) and underestimated variance (triangles up) $\sigma'^2 = 0.5\sigma^2$. For comparison, we also depict (dashed line) the performance of the optimal MAP detector (BCJR) operating with exact channel and noise variance parameters on blocks of 300 samples.

As one might verify, when the channel order is known, the proposed method perform similarly to that described in [1], much outperforming the latter when the employed noise variance value is incorrect. When the channel order is unknown, however, the proposed method suffers a performance penalty. This result qualitatively differs from those obtained by the method described in [2], which however adopts different priors for the channel parameters and assumes the noise variance known.

(Initialization)
For $L \in \mathcal{L}$
For $i = 1, ..., P$
Set $R^{Y(i)}_{-1,L} = 0$, $R^{SY(i)}_{-1,L} = 0, \Sigma^{(i)}_{-1,L} = I\epsilon^2$,
and $|\Sigma^{(i)}_{-1,L}| = \epsilon^{2L}$.
(Algorithm)
For $n \geq 0$
1. For $i = 0, ..., P-1$
a) For $L \in \mathcal{L}$
For $s^{(i)}_n = \{-1, 1\}$
a1) Update $R^{Y(i)}_{n,L}$, $R^{SY(i)}_{n,L}$, $\Sigma^{(i)}_{n,L}$ and $|\Sigma^{(i)}_{n,L}|$ via (32)-(36).
a2) Determine $\alpha^{(i)}_{n,L}$ and $\beta^{(i)}_{n,L}$ via (27)-(28).
b) Determine $p(S^{(i)}_{0:n-1}, s^{(i)}_n = \pm 1, y_{0:n})$ by summing (30) for all $L \in \mathcal{L}$
c) Determine the optimal importance function via (14), draw $S^{(i)}_n$ and select the corresponding values of $R^{Y(i)}_{n,L}$, $R^{SY(i)}_{n,L}$, $\Sigma^{(i)}_{n,L}$ and $|\Sigma^{(i)}_{n,L}|$.
d) Determine $b^{(i)}_n$ by differentially decoding $s^{(i)}_{n-1:n}$.
e) Update the weights $w^{(i)}_n$ via (13).
2. Normalize the weights.
3. Estimate b_{n-d} via (37).
4. Estimate the effective sample size
$N_{eff} = \left(P \sum_{i=1}^{P} |w^{(i)}_n|^2\right)^{-1}$
5. If N_{eff} is lesser than an arbitrary threshold, resample [6] the particles.

Table 1. Proposed Algorithm

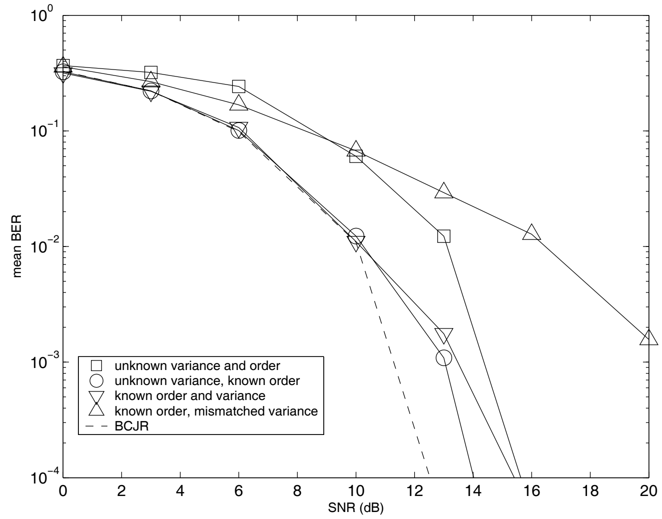

Fig. 1. Performance of blind equalization algorithms as a function of the signal-to-noise ratio (SNR) compared to the optimal (BCJR algorithm). The "unknown variance" curves refer to the proposed method, while the other results employ the method described in [1].

6. CONCLUSIONS

We proposed in this work a new particle filtering algorithm for blind equalization of unknown order communication channels subject to unknown variance additive noise. Via numerical simulations, we verified that the proposed method exhibits a near optimal performance for known order channels even in the absence of noise variance prior information, constituting therefore a more robust alternative to previous methods (e.g. [1]) that require exact knowledge of this quantity, without demanding increased computation effort.

7. REFERENCES

[1] J. S. Liu and R. Chen, "Blind Deconvolution via Sequential Imputations," *J. of the Am. Stat. Assoc.*, vol. 90, no. 430, pp. 567–576, June 1995.

[2] J. Míguez and P. M. Djurić, "Blind equalization of frequency-selective channels by sequential importance sampling," *IEEE Trans. on Sig. Proc.*, vol. 52, no. 10, October 2004.

[3] P. M. Djurić, J. H. Kotecha, J. Zhang, Y. Huang, T. Ghirmai, M. F. Bugallo, and J. Míguez, "Particle Filtering," *IEEE Sig. Proc. Mag.*, vol. 20, no. 5, pp. 19–38, September 2003.

[4] R. Chen, J. S. Liu, and X. Wang, "Convergence Analyses and Comparisons of Markov Chain Monte Carlo Algorithms in Digital Communications," *IEEE Trans. on Sig. Proc.*, vol. 50, no. 2, pp. 255–270, February 2002.

[5] J. G. Proakis and M. Salehi, *Communication systems engineering*, Prentice Hall, 1994.

[6] A. Doucet, S. Godsill, and C. Andrieu, "On sequential Monte Carlo sampling methods for Bayesian filtering," *Stat. Comput.*, vol. 10, pp. 197–208, July 2000.

[7] J. J. K. Ruanaidh and W. J. Fitzgerald, *Numerical Bayesian Methods Applied to Signal Processing*, Springer-Verlag, 1996.

[8] A. H. Sayed, *Fundamentals of Adaptive Filtering*, Wiley-IEEE Press, 2003.

[9] C. J. Bordin Jr. and L. A. Baccalá, "Particle Filters for Blind FIR Channel Equalization in Non-Gaussian Noise," in *Proc. of ICASSP 2006*, Toulouse - France, 2006, vol. IV, pp. 457–460.