A BLIND NETWORK OF EXTENDED KALMAN FILTERS FOR NONSTATIONARY CHANNEL EQUALIZATION

Rim Amara, Sylvie Marcos

Laboratoire des Signaux et Systèmes, CNRS-Supélec, plateau de Moulon,
91192, Gif-Sur-Yvette, FRANCE
amara@lss.supelec.fr , marcos@lss.supelec.fr

ABSTRACT

In this paper, a blind Network of Extended Kalman Filters (NEKF) is introduced for nonstationary linear channel equalization. The structure of NKF was recently suggested for optimal channel equalization. As the knowledge of the channel is the main constraint within the NKF equalizer, we here propose to extend the state to estimate, that was previously formed by the last M transmitted symbols, to the time-varying channel coefficients. The observation model becomes nonlinear suggesting thus extended Kalman filtering for state estimation. The proposed NEKF algorithm is completely blind towards any learning phase, with fast convergence properties. Compared to the blind Bayesian algorithm proposed by Iltis et al. in [2], the NEKF-based equalizer shows good performance with a really lower complexity.

1. INTRODUCTION

Signal processing literature provides us with a panoply of deconvolution algorithms adapted for the linear channel equalization. When channel equalization is viewed as an estimation procedure of the transmitted symbols $d(k)$, two main directions can be followed : hard and soft estimation of the symbols. In fact, one can think of the equalizer as a parameterized structure, such that the Linear Transverse Equalizer (LTE)[11], the Decision Feedback Equalizer (DFE)[11] or neural network one (MLP [10]), whose output, $\hat{d}(k-r)$, is a function f of the noisy observations $\{y(k), \dots, y(k-m+1)\}$ and of a set of structural parameters, $W(k)$, optimized according to a chosen criterion, via a learning phase. Generally, these are flexible structures with a relatively low complexity. Nevertheless, according to the chosen parameterization, the equalizer may require a long time of learning which is unacceptable in many communication systems. Therefore, a good generalization capacity of the equalizer, driven by the function f, out of the learning phase and especially in non stationary environments is required. The second approach is rather seen as a soft estimation of the transmitted symbols through the determination of their *a posteriori* probability density function (pdf) $p(d(k-r)/O)$[8][9][5][2], where O is a finite set of observations or all of them $y^k = \{y(0), \dots, y(k)\}$. In fact, deriving optimal Minimun Mean Square Error (MMSE) or Maximum *A Posteriori* (MAP) symbol-by-symbol estimators is based on the determination of the conditional symbol pdf. The structure consisting of a Network of Kalman Filters (NKF) emerges, with the same aim and according to a state formulation, when expanding the infinite horizon *a posteriori* pdf, $p(\mathbf{D}(k)/y^k)$ into a Weighted Gaussian Sum (WGS) [1] where $\mathbf{D}(k) = [d(k), \dots, d(k-M+1)]^T$ denotes the unknown symbol state vector. As most of the Bayesian estimators, the NKF-based equalizer [4][5] assumes a perfect knowledge of the communication system modelling and consequently of the channel coefficients. One solution is to identify the channel using the LMS algorithm [4]. If the latter is nonstationary, a Kalman or RLS identifying block can be used to track nonstationarities [6]. The solution, suggested here, consists in augmenting the state, previously considered in the last transmitted symbols $\mathbf{D}(k)$, to the channel coefficients. The observation model in the new state becomes a nonlinear one, and thus, a Network of Extended Kalman Filters (NEKF) can be used to give the MMSE state estimate. It appears that the approach by NEKF, in the context of channel identification, is similar to the blind Bayesian estimator developed by Iltis et al. in [2], which turns out to be structured into parallel Kalman filters too. Compared to this blind Bayesian equalizer, the NEKF achieves good performance with a lower complexity.

In Section 2, the NEKF algorithm is described with its associated state formulation. Section 3 comments the simulation results. An open discussion is presented in the fourth Section about the new blind equalization algorithm proposed in this paper. Finally, we give our conclusion.

2. BLIND EQUALIZATION BY NEKF

As stated in [4][5], the approach of linear channel equalization by parallel Kalman filtering, is based on a state formulation of the digital communication system as follows

$$\begin{aligned} \mathbf{D}(k+1) &= F\mathbf{D}(k) + \mathbf{G}d(k+1), \\ y(k) &= \mathbf{C}^T\mathbf{D}(k) + n(k), \end{aligned}$$

where

$F = \begin{bmatrix} 0\ 0\ 0 \dots 0 \\ 1\ 0\ 0 \dots 0 \\ 0\ 1\ 0 \dots 0 \\ \dots\dots\dots \\ 0 \dots 0\ 1\ 0 \end{bmatrix}$ is the one-step transition matrix,

$\mathbf{G} = [1, 0, \dots, 0]^T$, $\mathbf{D}(k) = [d(k), \dots, d(k-M+1)]^T$ denotes the symbol state to be estimated according to the MMSE criterion, $\mathbf{C}$ is the M-vector of the channel coefficients and $n(k)$ is an additive white Gaussian noise $\mathcal{N}(0, \sigma_n^2)$. The symbols $d(k)$ are in a finite alphabet $\gamma = \{d_i, i = 1, \dots, q\}$. When the channel is nonstationary, we propose in this paper to use an *a priori* known model for the varying channel, as for example the Markovian transition which is described by

$$\mathbf{C}(k+1) = \mathbf{C}(k) + \mathbf{w}(k) \tag{1}$$

where $\mathbf{C}(k) = [c_0(k), \ldots, c_{M-1}(k)]^T$ and $\mathbf{w}(k)$ is a Gaussian vector $\mathcal{N}(\mathbf{0}_M, R_w)$, $R_w = \sigma_c^2\ I_{M\times M}$, and to augment the state to be estimated, previously taken as $\mathbf{D}(k)$, to $\mathbf{X}(k) = [\mathbf{D}^T(k)\ \ \mathbf{C}^T(k)]^T$. The state formulation then becomes

$$\mathbf{X}(k+1) = F'\mathbf{X}(k) + \underbrace{\begin{bmatrix} \mathbf{G}d(k+1) \\ \mathbf{w}(k) \end{bmatrix}}_{\mathbf{z}(k)}, \tag{2}$$

$$y(k) = h(\mathbf{X}(k)) + n(k), \tag{3}$$

where

$$h(\mathbf{X}(k)) = \mathbf{X}(k)|_{1:M}^T \mathbf{X}(k)|_{M+1:2M} = \mathbf{D}^T(k)\mathbf{C}(k)$$

and $F' = \begin{bmatrix} F & 0_{M\times M} \\ 0_{M\times M} & I_{M\times M} \end{bmatrix}$.

The state noise pdf is generally not Gaussian due to $\mathbf{G}d(k+1)$ ([4][5]) and it is approximated by a WGS as below

$$p(\mathbf{z}(k)) = \sum_{i=1}^{q} \frac{1}{q}\mathcal{N}(\mathbf{z}(k) - \mathbf{G}'d_i, Q_i)$$

where $\mathbf{G}' = [\mathbf{G}^T\ \mathbf{0}_M^T]^T$ and Q_i is a small diagonal matrix.

Denoting the following partial derivatives evaluated in the predicted states $\mathbf{X}_i(k/k-1)$, as defined in (4),

$$\begin{aligned} \mathbf{H}_i(k) &= \frac{\partial h}{\partial \mathbf{X}(k)}(\mathbf{X}_i(k/k-1)) \\ &= [\mathbf{C}_i^T(k/k-1)\quad \mathbf{D}_i^T(k/k-1)]^T \end{aligned}$$

the linearized model (3) around the predicted states can be expressed by

$$\begin{aligned} y(k) &= \mathbf{H}_i^T(k)\mathbf{X}(k) \\ &\quad + h(\mathbf{X}_i(k/k-1)) - \mathbf{H}_i^T(k)\mathbf{X}_i(k/k-1) + n(k) \end{aligned}$$

Thus, the NKF applied to the linearized observation model becomes a Network of Extended Kalman Filters (NEKF) governed by the following equations.

Given the last estimated state $\hat{\mathbf{X}}(k-1)$ and its associated error covariance $\hat{P}(k-1)$ matrix, compute for $i = 1, \ldots, q$:

prediction step

$$\mathbf{X}_i(k/k-1) = F'\hat{\mathbf{X}}(k-1) + \mathbf{G}'d_i \tag{4}$$

$$P_i(k/k-1) = F'\hat{P}(k-1)F'^T + Q_i \tag{5}$$

$$\begin{aligned} e_i(k/k-1) &= y(k) - \hat{\mathbf{C}}(k-1)^T\mathbf{D}_i(k/k-1) \\ \sigma_i^2(k/k-1) &= \mathbf{H}_i(k)^T P_i(k/k-1)\mathbf{H}_i(k) + \sigma_n^2 \end{aligned}$$

filtering step

$$\begin{aligned} \mathbf{K}_i(k) &= P_i(k/k-1)\mathbf{H}_i(k)/\sigma_i^2(k/k-1) \\ P_i(k/k) &= (I_M - \mathbf{K}_i(k)\mathbf{H}_i(k)^T)P_i(k/k-1) \\ \mathbf{X}_i(k/k) &= \mathbf{X}_i(k/k-1) + \mathbf{K}_i(k)e_i(k/k-1) \end{aligned}$$

Then, the *MMSE state estimation*, $E\{\mathbf{X}(k)/y^k\}$, is derived by

$$\begin{aligned} \beta_i(k) &= \mathcal{N}(e_i(k/k-1), \sigma_i^2(k/k-1)) \\ \alpha_i(k) &= \frac{\beta_i(k)}{\sum_{j=1}^{q}\beta_j(k)} \\ \hat{\mathbf{X}}(k) &= \sum_{i=1}^{q}\alpha_i(k)\mathbf{X}_i(k/k) \\ \hat{P}(k) &= \sum_{i=1}^{q}\alpha_i(k)\{P_i(k/k) \\ &\quad + [\hat{\mathbf{X}}(k) - \mathbf{X}_i(k/k)][\hat{\mathbf{X}}(k) - \mathbf{X}_i(k/k)]^T\} \end{aligned}$$

In the above equations, the covariance Q_i, in the WGS expansion of the state noise pdf is taken the same for all i and equal to $\sigma_c^2\ I_{2M\times 2M}$. The estimated state $\hat{\mathbf{X}}(k)$ is the concatenation of the so-resulted estimations of both the symbol vector, $\hat{\mathbf{D}}(k)$, and the channel coefficients, $\hat{\mathbf{C}}(k)$ and it corresponds to a blind estimate of both of them. Due to the feedback in $\hat{P}(k-1)$ (5), the updated estimates of $\mathbf{D}(k)$ and $\mathbf{C}(k)$ are highly coupled through the Kalman gain. In fact, noting $\mathbf{K}_i(k) = [\mathbf{K}_{i,\mathbf{D}}^T(k)\ \ \mathbf{K}_{i,\mathbf{C}}^T(k)]^T$, both $\mathbf{K}_{i,\mathbf{D}}(k)$ and $\mathbf{K}_{i,\mathbf{C}}(k)$ are functions of the predicted symbol states $\mathbf{D}_i(k/k-1)$ and the last channel estimate $\hat{\mathbf{C}}(k-1)$.

In what follows, simulation results of the NEKF-based equalizer are exhibited for a binary transmission.

3. SIMULATIONS

The simulations are carried out for the time-varying channel depicted in Figure 1 with $\mathbf{C}(0) = [1; 0.2; 0.5]^T$ and $\sigma_c^2 = 5\ 10^{-5}$. For the NEKF algorithm, the state initialization of the symbol state to the true one is feasible (this is because we can set the transmitter and receiver shift registers to the same value at the beginning of transmission), the channel coefficients are initialized to zero. The NEKF performance are compared to those given by the blind Bayesian algorithm developed by Iltis et al. in [2]. This blind algorithm computes recursively all the possible conditional channel estimates $\mathbf{C}_i^{Iltis}(k/k) = E\{\mathbf{C}(k)/\mathbf{D}(k) = \mathbf{D}_i, y^k\}$, $\mathbf{D}_i$ is a possible M-binary vector, in order to determine the *a posteriori* symbol state vector probabilities $p_i(k) = p(\mathbf{D}(k) = \mathbf{D}_i/y^k)$ for $i = 1, \ldots, q^M$. Figures 2 and 3 show the Mean Square Error for both the NEKF and blind Bayesian equalizers. For the last, the MMSE channel estimate is determined according to

$$\hat{\mathbf{C}}_{Iltis}(k) = \sum_{i=1}^{q^M} p_i(k)\mathbf{C}_i^{Iltis}(k/k)$$

For both channel estimates, an empirical Mean Square Error (MSE) is computed as follows

$$MSE(k) = \frac{1}{card(\Omega)} \sum_{\omega \in \Omega} \{ \frac{1}{k} \sum_{i=1}^{k} ||\hat{\mathbf{C}}_\omega(i/i) - \mathbf{C}(i)||^2 \}$$

the index ω refers to a certain realization of the additive observation noise and Ω is the set of the considered realizations. The MSE, as well as Bit Error Rates (BERs), are obtained by averaging the results collected after 100 runs of a sequence of 10^4 bits. During the simulations, we have noticed some misconvergences of both algorithms to the opposite of the channel coefficients. The problem can be circumvented in fact by introducing an *a priori* information for example on the sign of a channel coefficient at the end of convergence. Even though, because the transitory length is not controllable, not all the misconvergences were avoided. Table 1 illustrates the rates of good convergence of the two algorithms and Table 2 shows the averaged BERs for the NEKF and blind Bayesian algorithm on only the good realizations. We can notice that the BERs achieved by the NEKF are comparable to those of Iltis et al. with a faster convergence.

4. DISCUSSION

The idea of augmenting the state to estimate to the unknown system parameters is certainly not new. Applied to parallel Kalman filtering, augmenting the state shows good promising results (Table 2) even though there is still a lot of research to do in dealing with the initialization and stability of such blind algorithm. However, some noteworthy remarks can be made :

• First, both the blind Bayesian algorithm of Iltis et al., which is the nearest optimal blind symbol detector to our knowledge in a nonstationary environment, and the NEKF algorithm determine, recursively, the *a posteriori* state pdf $p(\mathbf{D}(k)/y^k)$. More precisely, the NEKF determines, recursively, the means around which the state pdf is picked with their corresponding spreads.

• The estimated channel in each algorithm is given by

$$\begin{aligned} \hat{\mathbf{C}}(k) &= \sum_{i=1}^{q} \alpha_i(k)\mathbf{C}_i(k/k) \quad \text{for NEKF-equalizer} \\ \hat{\mathbf{C}}_{Iltis}(k) &= \sum_{i=1}^{q^M} p_i(k)\mathbf{C}_i^{Iltis}(k/k) \end{aligned}$$

with

$$\mathbf{C}_i(k/k) = \hat{\mathbf{C}}(k-1) + \mathbf{K}_{i,c}(k)(y(k) - \mathbf{D}_i^T(k/k-1)\hat{\mathbf{C}}(k-1))$$

$$\mathbf{C}_i^{Iltis}(k/k) = \mathbf{C}_i^{Iltis}(k/k-1) + \mathbf{K}_i^{Iltis}(k)(y(k) - \mathbf{D}_i^T\mathbf{C}_i^{Iltis}(k/k-1))$$

We note that both channel estimates are updated according to Kalman filters. The number of Kalman filters required by the blind Bayesian sequence estimator (q^M) is significantly higher compared to the one in the NEKF (just q). In fact, this is due to the prediction step led from only one admissible state, namely $\hat{\mathbf{C}}(k-1)$ for the channel coefficients and $\hat{\mathbf{D}}(k-1)$ for the symbol state ; and as the blind bayesian algorithm determines all the possible conditional channel estimates $E\{\mathbf{C}(k)/\mathbf{D}(k) = \mathbf{D}_i, y^k\}$, the NEKF computes a similar conditional mean, $E\{\mathbf{C}(k)/\mathbf{D}(k) = \mathbf{D}_i(k/k-1), y^k\}$, relatively to each predicted state. Hence, there is a real gain in complexity compared to the blind Bayesian algorithm. Also, for the NEKF-equalizer, assuming that symbols are i.i.d ($\forall\, i,\ Q_i$ is a constant), many equations in the NEKF-algorithm are the same. Besides, due to the special structure of $\mathbf{G}'$ and F', many matricial products can be seen as delays, thus achieving an additive gain in complexity.

• Extended Kalman filtering is a local approach to nonlinear filtering. The linearization of the observation model (3) is essentially dictated from a small deviation of the true state around nominal values of the state [3], namely in our equalization context, the predicted states $\mathbf{X}_i(k/k-1)$. Due to the discrete and rough character of part of the state space (combinations of $+1$ and -1 in a binary transmission), the linearization is accomplished between real scattered states $\mathbf{D}_i(k/k-1)$, thus risking the violation of the assumption of small deviations. This may, in fact, cause the instability noticed during simulations, especially for high SNR, when the real symbol predicted states, $\mathbf{D}_i(k/k-1)$, tend more towards M-binary vectors.

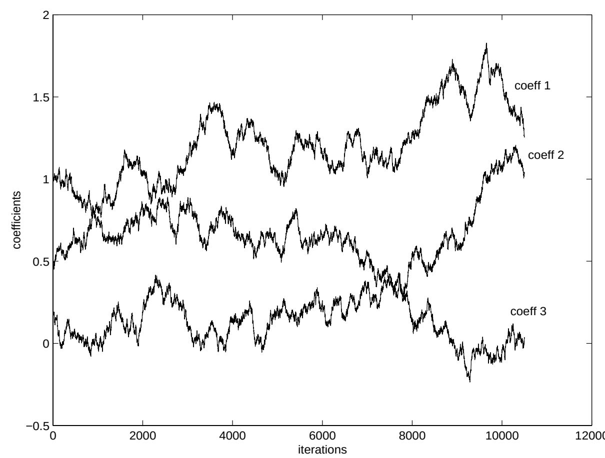

Fig. 1. *The Markovian channel used for simulations*

5. CONCLUSION

A Network of Extended Kalman Filters is proposed in this paper for nonstationary channel equalization. Coupling the MMSE estimation of the last transmitted symbols and the channel coefficients via a state formulation achieves good performance compared to the blind Bayesian algorithm of Iltis et al., with a lower complexity. Both the deduced channel estimates from the output of the NEKF and the conditional channel estimates of the blind Bayesian equalizer are updated according to similar Kalman filtering equations. The last derives all the possible conditional channel estimates whereas the NEKF structure considerably reduces the number of required Kalman filters thanks to the prediction step. As the prediction is the heart of the NEKF estimation as presented here, it will be worthwhile to better characterize admissible symbol state sets for the NEKF prediction stage.

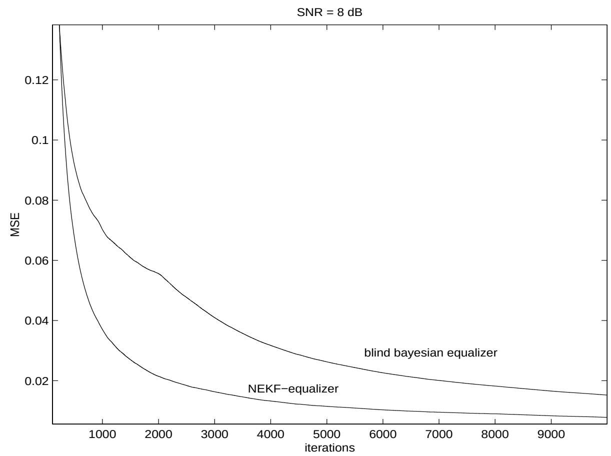

Fig. 2. *MSE for SNR = 8 dB, $M = 3$ and $r = 2$*

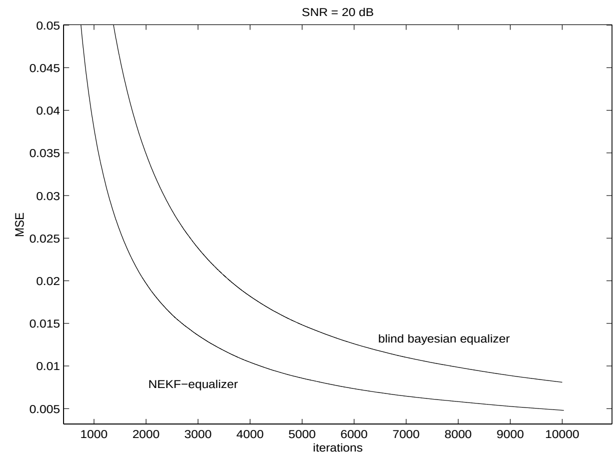

Fig. 3. *MSE for SNR = 20 dB, $M = 3$ and $r = 2$*

6. REFERENCES

[1] D.L Alspach, H.W Sorenson. *Nonlinear Bayesian Estimation Using Gaussian sum Approximations*. IEEE Transactions on Automatic control, Vol. 17, Aout 1972.

[2] R.A Iltis, J.J Shynk, K. Giridhar. *Bayesian algorithms for blind equalization*. IEEE trans. communications, vol. 42, no 2/3/4, pp. 1019-1032, Feb/Mar/Apr 1994.

[3] A.H Jaswinski. *Stochastic processes and filtering theory*. Academic press 1970.

[4] S.Marcos. *A network of adaptive Kalman filters for data channel equalization*. IEEE transactions on signal processing, September 2000.

[5] R.Amara , S.Marcos. *A network of Kalman filters for MAP symbol-by-symbol equalization*. ICASSP Istanbul, Turkey, June 2000.

[6] R.Amara , S.Marcos. *Parallel Kalman filtering for optimal equalization*. in preparation

[7] B. Mulgrew. *Nonlinear signal processing for adaptive equalization and multiuser detection*. EUSIPCO 98, Rhodes, Greece, pp. 537-544, set. 1998.

[8] S. Chen, B. Mulgrew and P.M Grant. *A clustering technique for digital communications channel equalization using Radial Basis Function Networks*. IEEE Trans. Neural Networks, vol. 4, no. 4, pp. 570-579, July 1993.

[9] J. Cid-Sueiro, A. Artes-Rodriguez, A.R Figueiras-Vidal. *Recurrent Radial Basis Function Networks for optimal symbol-by-symbol equalization*. Signal processing, vol. 40, pp. 53-63, 1994.

[10] M. Meyer, G. Pfeiffer. *Multilayer perceptron based decision feedback equalizers for channels with intersymbol interference*. IEE proceedings, vol.140, n6, pp. 420-424, Dec. 1993.

[11] J.G. Proakis. *Digital communication*. Mc Graw Hill,1989.

SNR (dB)	6	8	10	12	14	16	18	20	22	25
NEKF (%)	97	100	100	99	100	100	100	100	100	100
Iltis (%)	100	98	98	96	96	96	95	92	87	87

Table 1 : Rates of good convergence

SNR	0	2	4	6	8	10	12	14	16	18	20	25
NEKF	0.1747	0.1186	0.0848	0.0149	0.0041	$7.11e^{-4}$	$1.31e^{-4}$	$1.49e^{-4}$	$1.21e^{-4}$	$1.05e^{-4}$	$9.4e^{-5}$	$7.2e^{-5}$
Iltis	0.1229	0.0757	0.0416	0.0558	0.0681	0.0770	0.0339	0.0020	$4.17e^{-5}$	$7.82e^{-4}$	0.0014	$1.14e^{-4}$

Table 2 : Bit Error rates for $M = 3$ and $r = 2$