RECURSIVE ESTIMATION OF SIGNALS IN AN AUTOREGRESSIVE NOISE

T. Gouraud†, F. Auger††, M. Guglielmi†

†Laboratoire d'Automatique de Nantes, URA C.N.R.S. 823, Ecole Centrale de Nantes/Université de Nantes, 1 rue de la Noë, 44072 Nantes cedex 03, France. E-mail : gouraud@lan.ec-nantes.fr

††LRTI-GE44, Bd de l'Université, B.P.406, 44602 Saint Nazaire cedex, France. E-mail : f.auger@ieee.org.

ABSTRACT

Most methods estimating noisy sinusoidal signals assume the noise to be white, and fail when they are used on real signals with colored noise. In this paper, we propose two new recursive algorithms, deduced from a recent work of Kay and Nagesha, for the estimation of sinusoidal signals embedded in an AR noise. The first one is a λ-RLS, whereas the second one uses Kalman Filtering. Their convergence speed, computational burden and statistical characteristics are compared and the advantages brought by these estimators for real signals are shown.

1. INTRODUCTION

The parameter estimation of sinusoidal signals from noisy data still receives a great attention. Many estimators have been proposed, among which the most studied and used are certainly the Prony, Pisarenko, MUSIC and minimum norm methods [1, 2]. They are asymptotically efficient and have excellent performances, but they assume that the additive noise is white. Unfortunatly, this hypothesis cannot be satisfied in some real applications, such as voltage and current measure of power supplies. For such cases, the method of the constrained maximum likelihood [3] or methods based on the fourth order cumulants [4, 5] solve the estimation problem of amplitudes and frequencies from data blurried with colored noise of unknown power spectral density. But since these estimators assume that the observed signals are wide sense stationary, they are unsuitable to track the slow variations of non-stationary signals. Furthermore, they need a large computational burden and they cannot be used for real time applications. On the other hand, the major work of Kay and Nagesha [6] yields an elegant maximum likelihood estimator (MLE) of the parameters of both the sinusoids and the AR noise. But this estimator requires computationally expensive matrix inversions and cannot track slowly-varying parameters. To avoid these drawbacks, we propose in this paper a recursive implementation of the MLE, and an algorithm based on the extended Kalman filter (EKF) approach.

After a brief introduction showing more precisely the motivation and the background of our work, section 2 presents the signal model and the new algorithms. In section 3, statistical characteristics, convergence speed and computational burden of these algorithms are compared. Finally, some conclusions are drawn which highlight the advantages of these algorithms for real signals.

2. SIGNAL MODELISATION AND ALGORITHMS

2.1. Signal modelisation

Many observed physical phenomenons lead to stationary signals which are composed of sinusoidal components embedded in an additive colored noise. When this noise is derived from a stable AR process of finite order, such signals write :

$$y[n] = \sum_{k=0}^{M} B_k e^{jn\theta_k} + w[n]; \quad n = 0, \cdots, N-1 \quad (1)$$

$$w[n] = -\sum_{l=1}^{p} a_l w[n-l] + u[n] \quad (2)$$

$$B_k = |B_k| e^{j\phi_k}; \quad k = 0, \cdots, M \quad (3)$$

where $|B_k|$, ϕ_k, θ_k and a_l are unknown and considered as deterministic, whereas the number of sinusoidal components and the AR model order p are assumed to be known. Thanks to a linear transformation of the parameters, (1) can also be written as :

$$y[n] = -\sum_{l=1}^{p} a_l y[n-l] - \sum_{k=0}^{M} \mu_k e^{jn\theta_k} + u[n] \quad (4)$$

$$\text{with} \quad \mu_k = -B_k \left(1 + \textstyle\sum_{l=1}^{p} a_l e^{-jl\theta_k}\right) \quad (5)$$

2.2. λ−RLS algorithm

This non-linear transformation (5) is the key point of the MLE proposed by Kay and Nagesha [6]. Then, when the frequencies are assumed to be known, the equation (1) is linear related to the unknown parameters and a λ-RLS can be used to estimate the amplitudes and the AR coefficients. Therefore, when n data are available, the usual quadratic criterion of the λ−RLS writes [7]:

$$J(\mu[n], a[n]) = \sum_{k=p}^{n} \lambda^{n-k} |v[k]|^2 \qquad 0 \le \lambda \le 1 \quad (6)$$

$$= V[n]^H \Lambda[n] V[n] \quad (7)$$

$$\begin{aligned}
V[n] &= Y[n] - H[n]\,a[n] - E[n]\,\mu[n] \\
\Lambda[n] &= diag(\lambda^{n-p}, \lambda^{n-p-1}, \cdots, 1) \\
[\mu[n]]_m &= \mu_{m-1}[n], \quad m = 1, \cdots, M+1 \\
[Y[n]]_k &= y[p+k-1], \quad k = 1, \cdots, n-p+1 \\
[V[n]]_k &= v[p+k-1] \\
[a[n]]_l &= a_l[n], \qquad l = 1, \cdots, p \\
[E[n]]_{km} &= e^{j(p+k-1)\theta_{m-1}} \\
[H[n]]_{kl} &= y[p+k-l-1]
\end{aligned}$$

and the resulting parameter estimator is given by :

$$\begin{bmatrix} \hat{a}[n] \\ \hat{\mu}[n] \end{bmatrix} = -P[n]\,[H[n] \mid E[n]]^H \Lambda[n] Y[n] \quad (8)$$

$$P[n] = \left([H[n] \mid E[n]]^H \Lambda[n] [H[n] \mid E[n]]\right)^{-1} \quad (9)$$

where $\hat{\mu}_k$ and $\hat{a}_l$ denote respectively the estimated values of μ_k and a_l.

The use of the forgetting factor λ is intended to ensure that the data in the far past are forgotten, to offer the tracking ability. With λ close to one, the algorithm has a large memory time constant. Therefore, the parameter estimation is accurate but the time variations of parameters are slowly followed. On the other hand, if λ is small, the λ-RLS is alert to follow the changes of parameters, but the estimates of a and μ are not very accurate. Thus, a trade-off between the convergence speed and the insensitivity to the noise should be found [7].

For computational efficiency and for real time applications, we are interested in estimating the unknown parameters recursively. Thus, thanks to the matrix inversion lemma [8], the λ−RLS method can be recursively implemented :

$$\begin{bmatrix} \hat{a}[n] \\ \hat{\mu}[n] \end{bmatrix} = \begin{bmatrix} \hat{a}[n-1] \\ \hat{\mu}[n-1] \end{bmatrix} - P[n] \begin{bmatrix} h[n] \\ e[n] \end{bmatrix} \epsilon[n] \quad (10)$$

$$\epsilon[n] = y[n] + h[n]^H \hat{a}[n-1] + e[n]^H \hat{\mu}[n-1] \quad (11)$$

$$h[n]^H = [y[n-1], \cdots, y[n-1-p]] \quad (12)$$

$$e[n]^H = e[n-1]^H\, T \quad (13)$$

$$T = diag\left(e^{j\theta_0}, \cdots, e^{j\theta_M}\right) \quad (14)$$

$$P[n] = \frac{P[n-1]}{\lambda} - \frac{P[n-1]\begin{bmatrix} h[n] \\ e[n] \end{bmatrix}\begin{bmatrix} h[n] \\ e[n] \end{bmatrix}^H P[n-1]}{\lambda\left(\lambda + \begin{bmatrix} h[n] \\ e[n] \end{bmatrix}^H P[n-1] \begin{bmatrix} h[n] \\ e[n] \end{bmatrix}\right)}$$

Theorically, $P[n]$, $\hat{\mu}[n]$ and $\hat{a}[n]$ should be initialized with (8) and (9) when $M+1+2p$ data are available. But, with such an initialization, we must wait the $M+1+2p$ data and inverse an almost sure invertible matrix. An alternative is to use :

$$\begin{aligned}
\hat{\mu}[p] &= 0 & e[p]^H &= [e^{jp\theta_0}, \cdots, e^{jp\theta_M}] \\
\hat{a}[p] &= 0 & h[p]^H &= [y[p-1], \cdots, y[0]] \\
P[p] &= I_{M+1+p}
\end{aligned}$$

which leads to a suboptimal algorithm. However, the initialization will be as rapidely forgotten as λ is small. Thus, λ has a great influence on both the transient behavior and the tracking ability. Both phases can be independently controlled and a faster convergence speed can be reached when the constant λ is substituted by one growing exponentially with time, such as [7] :

$$\begin{aligned}
\lambda[n] &= \alpha\,\lambda[n-1] + (1-\alpha)\,\lambda_\infty & 0 < \alpha \le 1 \\
&= \alpha^{n-p}\,\lambda[p] + (1-\alpha^{n-p})\,\lambda_\infty & 0 < \lambda_\infty \le 1
\end{aligned}$$

The parameter α mainly influences the convergence speed whereas λ_∞ controls the trade-off between alertness to follow the time-variations of $a[n]$ and $\mu[n]$ and the insensitivity to the noise. Moreover, the use of a time-varying forgetting factor does not increase the computational complexity : the constant factor λ is substituted by $\lambda[n]$ in the recursive equation (10).

2.3. EKF algorithm

This λ−RLS estimator assumes that the frequencies are known. Unfortunatly, this assumption is not satisfied by real signals : the frequencies may change slowly over time. For this, the transformation (5) can be used to derive a Kalman filter estimator [7], presented here for periodic signals whose frequencies are all multiple of a fundamental one θ.

Expression (1) can be interpreted as a non-linear state model :

$$Z[n] = A\,Z[n-1] \quad (15)$$

$$y[n] = C(Z[n]) + u[n] \quad (16)$$

$$C(Z[n]) = -\sum_{l=1}^{p} a_l y[n-l] - \sum_{k=0}^{M} \tilde{\mu}_k e^{jk\Phi[n]} \quad (17)$$

$$Z^T[n] = [\Phi[n],\ \theta[n],\ a[n],\ \tilde{\mu}[n]] \quad (18)$$

$$\Phi[n] = (n-p)\theta; \quad \tilde{\mu}_k = \mu_k e^{-jpk\theta} \quad (19)$$

$$[A]_{ij} = 1\ if\ (i=j)\ or\ (i=1; j=2) \quad (20)$$

The observed equation (16) is non-linear whereas the transition equation (15) is linear and corresponds to the time variations of the phase and to constant parameters. Furthermore, if the observed signal is non-stationary, then the time variations of parameters can be modelized by a random walk :

$$Z[n] = A\, Z[n\text{-}1] + W[n\text{-}1] \qquad (21)$$

where $W[n\text{-}1]$ is a random variable with covariance Q :

$$Q = diag(0, \sigma_\theta^2, \sigma_{a_1}^2, \cdots, \sigma_{a_p}^2, \sigma_{\mu_0}^2, \cdots, \sigma_{\mu_M}^2) \qquad (22)$$

Unlike the $\lambda-$RLS, the time variations of the parameters can be controlled independently. Unfortunatly, this approach increases the computational burden.

Since the observation equation (16) is not a linear function of the state $\Phi[n]$, we have to develop a first order Taylor expansion of (16), around the current estimate. Thus, according to the Kalman filtering theory [7], we get :

$$\begin{aligned}
Z[n/n\text{-}1] &= A\, Z[n\text{-}1/n\text{-}1] \qquad n \geq p & (23)\\
P[n/n\text{-}1] &= A\, P[n\text{-}1/n\text{-}1]\, A^T + Q & (24)\\
Z[n/n] &= Z[n/n\text{-}1] + K[n]\, \epsilon[n/n\text{-}1] & (25)\\
\epsilon[n/n\text{-}1] &= y[n] - C(Z[n/n\text{-}1]) & (26)\\
K[n] &= P[n/n\text{-}1]\, \delta C^H[n]\, \Sigma[n]^{-1} &\\
\Sigma[n] &= \delta C[n]\, P[n/n\text{-}1]\, \delta C^H[n] + \sigma^2 &\\
P[n/n] &= P[n/n\text{-}1] - K[n]\, \delta C[n]\, P[n/n\text{-}1] &\\
\delta C[n] &= \frac{\partial C}{\partial Z}(Z[n/n\text{-}1]) &
\end{aligned}$$

Because of the linearization and for an almost sure convergence, the initial value of $\theta[p/p\text{-}1]$ shall not be too far from the real one, otherwise the EKF may diverge. On the other hand, the other parameters can be initialized with any value. We choose to set $\mu[p/p\text{-}1] = 0$ and $a[p/p\text{-}1] = 0$.

3. COMPARISON OF THE ALGORITHMS

To compare both algorithms, a Monte-Carlo study is performed to investigate the statistical characteristics and the convergence speed of our algorithms. For this, we simulate a stationary periodic signal embedded in a first order AR noise whose coefficient is set to 0.95 and the resulting SNR is 15 dB. The deterministic component of this signal is made of a dc component, one fundamental of normalized frequency set to 1/50 and two harmonics. All amplitudes are constant and equal to 1.

To study the statistical characteristics of both algorithms, $\sigma^2[n] = \left|\hat{B}[n] - B\right|^2 / M + 1$ is computed at each instant and statistics over one hundred independent Monte-Carlo runs are estimated. Moreover, to compare both algorithms, these ones are initialized with the same value. Especially, in the EKF, the fundamental frequency parameter of the extended state is initialized to the real value.

Figure 1 shows that the asymptotic performances of parameter estimators based on a colored or a white noise assumption are identical, as mentionned in [6, 9]. The EKF (with $Q = 0$) and the recursive MLE ($\lambda-$RLS with $\lambda = 1$) are asymptotically efficient : they reach asymptotically the Cramer-Rao lower bound (CRB) [10] given for a white noise assumption. This latter is asymptotically equivalent to the Cramer-Rao lower bound based on colored noise. The algorithms assuming colored noise converge quicker than the other ones.

Finally, the $\lambda-$RLS algorithm is applied on the current distorted by a dimmer (figure 2) [10]. This real signal sampled at 3 kHz is a one-phase current whose fundamental frequency is 50 Hz and the study of the periodogram leads us to keep the 15 first harmonics to describe it. Figure 3 displays the square error between the real current and the estimated signal. We conclude that the $\lambda-$RLS algorithm based on colored noise converges faster than one ignoring the correlation of the noise. But the asymptotic behavior of both algorithms is identical. Thus, the conclusions drawn by simulations are confirmed.

From a computational point of view, the λ-RLS requires $(M + p + 1)^2$ multiplications and additions and no matrix inversion. Thereby, the comptational burden is proportional to $(M + p + 1)^2$ per data sample. We have to compare favourably this result to the computational burden of the MLE (8) (9) which is proportional to $\mathcal{O}[(p + M + 1)^3]$ per time sample.

4. CONCLUSION

In this paper, we propose two new recursive algorithms to estimate the parameters of sinuoidal signals embedded in an AR noise. In addition, white noise can also be considered by setting p to zero. The asymptotical performances of these estimators based on colored or white noise are identical. Moreover, since the recursive MLE ($\lambda-$RLS with $\lambda = 1$) and the EKF (with $Q = 0$) reach the Cramer-Rao lower bound (CRB), they are asymptotically efficient. But the convergence speed of both algorithms is quicker when the assumption of colored noise is made.

Unlike the $\lambda-$RLS algorithm, the MLE proposed by Kay and Nagesha [6] is a multistage estimator. Therefore, to reduce the computational complexity from $\mathcal{O}[(p+M+1)^2]$ to $\mathcal{O}[(M+1)^2] + \mathcal{O}[p^2]$, further works have already started to provide a recursive multistage estimator.

Finally, our recent works show that the λ−RLS is rather sensitive to errors on the fundamental frequency. Therefore, a further extension is to add to the λ−RLS algorithm a frequency estimator which can track slow variations of frequencies.

5. REFERENCES

[1] **S. Kay,** *"Modern spectral estimation - Theory and application,"* Prentice-Hall, Englewood Cliff, New Jersey, 1988.

[2] **R. Kumaresan,** *"On the zeros of the linear prediction-error filter for deterministic signals,"* IEEE Trans. on ASSP, Vol. 31, No 1, pp 576-579, Feb. 1983.

[3] **M.J. Turmon, M.I. Miller,** *"Maximum-likelihood estimation of complex sinusoids and Toeplitz covariance,"* IEEE Trans. on S.P., Vol. 42, No 5, pp 1074-1085, May 1994.

[4] **Z. Shi, F. Fairman,** *"Harmonic retrieval via state and forth order cumulants,"* IEEE Trans. on S.P., Vol. 42, No 5, pp 1109-1119, May 1994

[5] **A. Swami, J. Mendel,** *"Cumulant-based approach to the harmonic retrieval and the related problems,"* IEEE Trans. on S.P., Vol. 39, No 5, pp 1099-1109, May 1991.

[6] **S. Kay, V. Nagesha,** *"Maximum likelihood estimation of signals in autoregressive noise,"* IEEE Trans. on S.P. , Vol.42, No 1, pp 88-101, Jan. 1994.

[7] **L. Ljung, T. Södertröm,** *"Theory and practice of recursive identification,"* MIT Press, 1983.

[8] **G.H. Golub, C.F. Van Loan,** *"Matrix Computation,"* Johns Hopkins University Press, Baltimore, 1984.

[9] **P. Stoica, A. Nehorai,** *"Statistical analysis of two non-linear least-squares estimators of sine waves parameters in the colored noise case,"* proc. IEEE ICASSP, pp 2408-2411, 1988.

[10] **T. Gouraud, F. Auger, M. Guglielmi, M. Machmoum, S. Siala, M.F. Benkhoris,** *"A Maximum Likelihood Approach to Harmonics Measurements in Power Systems,"* IEEE PELS Workshop on computers in power electronics, Québec, Canada, Aug. 7-10, 1994.

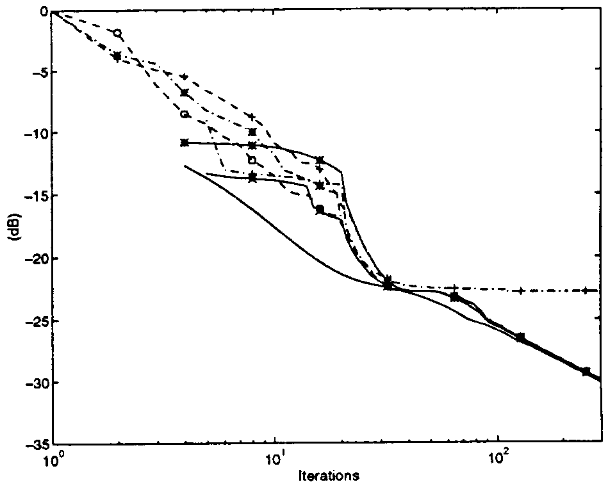

Fig.1 : $\sigma_B^2[n]$ versus iterations. Solid line and x (respectively $\star$) : MLE with a AR noise assumption (a white noise assumption); Dashdot line and x : Suboptimal MLE with a AR noise assumption. Dashed line and o (respectively +) : EKF with with a AR noise assumption (a white noise assumption); Dashdot line and + : $\lambda - RLS$ ($\lambda = 0.95$) with a AR noise assumption; Solid line : CRB

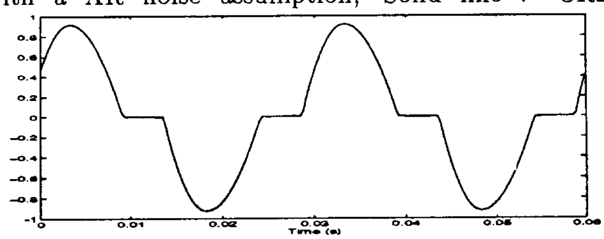

Fig.2 : a current distorted by a dimmer

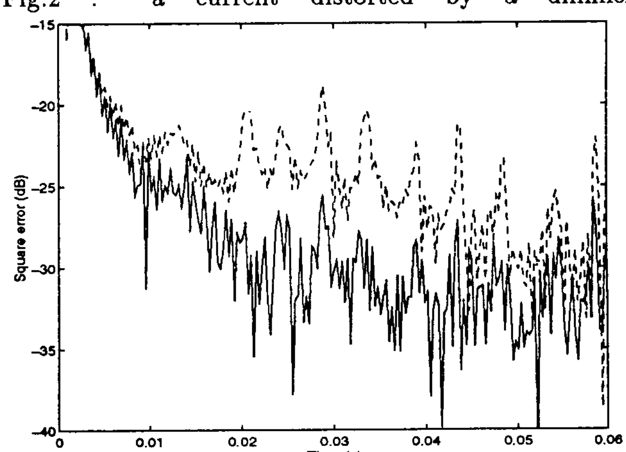

Fig.3 : Solid line : $\lambda - RLS$ ($\lambda = 0.95$) with a AR noise assumption (p = 2); Dashed line : $\lambda - RLS$ with a white noise assumption.