Analysis of fast recursive least squares algorithms for adaptive filtering

M. Arezki, A. Benallal, P. Meyueis, A. Guessoum and D. Berkani

***Abstract*—**In this paper, we present new version of numerically stable fast recursive least squares (NS-FRLS) algorithm. This new version is obtained by using some redundant formulae of the fast recursive least squares (FRLS) algorithms. Numerical stabilization is achieved by using a propagation model of first order of the numerical errors. A theoretical justification for this version is presented by formulating new conditions on the forgetting factor. An advanced comparative method is used to study the efficiency of this new version relative to RLS algorithm by calculating their normalized square norm gain error. It will be followed by an analytical analysis of the convergence of this version and we show, both theoretically and experimentally, their robustness. The simulation over a very long duration for a stationary signal did not reveal any tendency to divergence.

***Keywords*—**Fast RLS, Estimation, Adaptive Filtering, Propagation of Errors, Numerical Stability.

I. Introduction

In general the problem of system identification involves constructing an estimate of an unknown system given only two signals, the input signal and a reference signal. Typically the unknown system is modelled linearly with a finite impulse response (FIR), and adaptive filtering algorithms are employed to iteratively converge upon an estimate of the response. If the system is time-varying, then the problem expands to include tracking the unknown system as it changes over time. The system identification problem has numerous applications in control systems, digital communications, and signal processing, and a recent survey of adaptive filtering algorithms highlights the rich diversity of techniques available in the literature [1]. Adaptive filtering has been, and still is, an area of active research, playing important roles in an ever increasing number of applications [1], [2]. Numerous algorithms for the solution of the adaptive filtering problem have been proposed over the years. The recursive least squares (RLS) algorithms are used in a broad class of applications. The RLS algorithm solves this problem, but at the expense of increased computational complexity. A large number of fast RLS (FRLS) algorithms have been developed over the years, but, unfortunately, it seems that the better a FRLS algorithm is in terms of computational efficiency, the more severe is its problems related to numerical stability [3]. Several numerical solutions of stabilization, with stationary signals, are proposed in the literature [5]–[10]. In the following section, we propose a new version of numerically stable FRLS (NS-FRLS) algorithm. This new version is obtained by using some redundant formulae of the fast recursive least squares FRLS algorithms. Numerical stabilization is achieved by using a propagation model of first order of the numerical errors [5], [8]. We provide a theoretical justification for this version by formulating new conditions on forgetting factor. It will be followed by an analytical analysis of the convergence of this version and we show, both theoretically and experimentally, their robustness.

Manuscript received October 9, 2006: Revised version received March 4, 2007.
This work was supported by the Laboratoire de Traitement de Signal et Image LATSI of the Department of Electronics, University of Blida Algeria.

M. Arezki, A. Benallal and A. Guessoum are with the Department of Electronics, University of Blida Algeria. (e-mail: md_arezki@hotmail.com; a_benallal@hotmail.com; guessouma@hotmail.com).
P. Meyrueis is with the Laboratoire des Systemes Photoniques, University of Louis Pasteur, ENSPS, Bd. Sébastien Brant-BP 10413 ILLKIRCH 67412 France (e-mail: meyruies@sphot.u-strasbg.fr).
D. Berkani, is with the the Department of Electrical and Computer Engineering of ENP Algiers, Algeria (e-mail: dberkani@hotmail.com).

II. FRLS Algorithms

The main identification block diagram of a linear system with finite impulse response (FIR), by adaptive filtering using an adaptation algorithm, is represented in Fig.1.

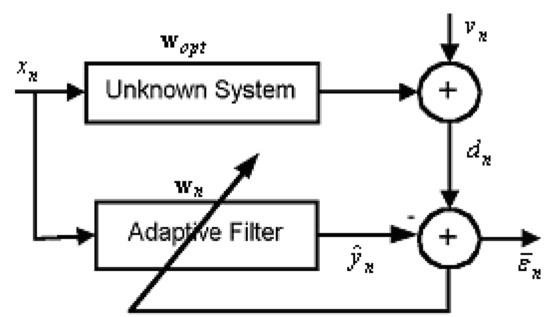

Fig. 1: System identification block diagram

The output a priori error $\bar{\varepsilon}_n$ of this system at time n is:

$$\bar{\varepsilon}_n = d_n - \hat{y}_n \quad (1)$$

where

$$\hat{y}_n = \mathbf{w}_{n-1}^{\mathrm{T}} \mathbf{x}_n \quad (2)$$

is the model filter output, $\mathbf{x}_n = [x_{n-1}, x_{n-2}, ..., x_{n-L+1}]^{\mathrm{T}}$ is a vector containing the last L samples of the input signal x_n,

$\mathbf{w}_{n-1} = [w_{1,n-1}, w_{2,n-1}, \ ..., w_{L,n-1}]^{\mathrm{T}}$ is the coefficient vector of the adaptive filter and L is the filter length. We assume that the desired signal from the model is:

$$d_n = v_n + \mathbf{w}_{opt}^{\mathrm{T}} \mathbf{x}_n \tag{3}$$

where $\mathbf{w}_{opt} = [w_{opt,1}, w_{opt,2}, \ ..., w_{opt,L}]^{\mathrm{T}}$ is the unknown system impulse response vector and v_n is a stationary, zero-mean, and independent noise sequence that is uncorrelated with any other signal. The superscript $^{\mathrm{T}}$ describes transposition. The filter $\mathbf{w}_n$ is calculated by minimizing the weighted least squares criterion according to [1]:

$$J_n(\mathbf{w}) = \sum_{i=1}^{n} \lambda^{n-i} \left(d_i - \mathbf{w}_n^T \mathbf{x}_i \right)^2 \tag{4}$$

where λ denotes the exponential forgetting factor ($0<\lambda\leq 1$). The recursive solution is written as follows:

$$\mathbf{w}_n = \mathbf{w}_{n-1} + \mathbf{g}_n \, \bar{\varepsilon}_n \tag{5}$$

where $\mathbf{g}_n$ represents the adaptation gain, given by:

$$\mathbf{g}_n = \underbrace{\mathbf{R}_n^{-1}\mathbf{x}_n}_{\text{RLS}} = \underbrace{\gamma_n \tilde{\mathbf{k}}_n}_{\text{FRLS}} \tag{6}$$

with

$$\mathbf{R}_n = \sum_{i=1}^{n} \lambda^{n-i} \mathbf{x}_i \mathbf{x}_i^T = \lambda \mathbf{R}_{n-1} + \mathbf{x}_n \mathbf{x}_n^{\mathrm{T}} \tag{7}$$

The quantity $\mathbf{R}_n$ is the L-by-L sample covariance matrix of the input signal x_n. The variables γ_n and $\tilde{\mathbf{k}}_n$ respectively indicate the likelihood variable and normalized Kalman gain vector. This latter is calculated, independently of the filtering part $\mathbf{w}_n$, by a FRLS algorithm using forward/backward linear prediction analysis over the signal x_n [1]. The calculation complexity of a FRLS algorithm is of order L. This reduction of complexity, compared to that of RLS algorithms, which have a complexity of order L^2, have made all FRLS algorithms numerically unstable.

III. Numerically Stable FRLS Algorithm

In this section, we present a new version of numerically stable fast recursive least squares (NS-FRLS) algorithms. This new version is obtained by using some redundant formulae of the fast recursive least squares FRLS algorithms. Numerical stabilization is achieved by using a propagation model of first order of the numerical errors [5], [8].

A. Proposed Approach

Any adaptive algorithm can be seen as a nonlinear dynamic system in discrete time, which is theoretically written in the form of state:

$$\boldsymbol{\varphi}_n = \mathrm{f}[\boldsymbol{\varphi}_{n-1}; \mathbf{x}_n] \tag{8}$$

where the function $\mathrm{f}[.]$ depends on the structure of the algorithm. The vector $\boldsymbol{\varphi}_n$ containing all quantities (scalar and vectorial) of the recursive variables of FRLS algorithms corresponding to the analysis, by forward and backward linear prediction of order L, of a signal x_n.

In finite precision, system (8) is written as:

$$\hat{\boldsymbol{\varphi}}_n = \mathrm{f}[\hat{\boldsymbol{\varphi}}_{n-1}; \mathbf{x}_n] + \mathbf{e}(n) \tag{9}$$

where the vector $\hat{\boldsymbol{\varphi}}_n$ is calculated by the algorithm in the presence of the instantaneous numerical errors represented by the vector $\mathbf{e}(n)$.

The introduced error will be noted $\Delta\boldsymbol{\varphi}_n$:

$$\Delta\boldsymbol{\varphi}_n = \hat{\boldsymbol{\varphi}}_n - \boldsymbol{\varphi}_n \tag{10}$$

this is a vector containing the errors in the states of the system. Using the following assumptions:

- Analysis of the errors to the 1st order:

$$\Delta\boldsymbol{\varphi}_n = \mathbf{F}(n)\,\Delta\boldsymbol{\varphi}_{n-1} + \mathbf{e}(n) \tag{11}$$

where $\mathbf{F}(n)$ represents a transition matrix for the error propagation, it is expressed by:

$$\mathbf{F}(n) = \frac{\partial\, \mathrm{f}[\boldsymbol{\varphi}, \mathbf{x}_n]}{\partial \boldsymbol{\varphi}} \Big|_{\boldsymbol{\varphi}=\boldsymbol{\varphi}_{n-1}} \tag{12}$$

This assumption makes it possible to obtain a linear system of propagation depending on time.

- Calculation in infinite precision: $\mathbf{e}(n) = 0$

$$\Delta\boldsymbol{\varphi}_n = \mathbf{F}(n)\,\Delta\boldsymbol{\varphi}_{n-1} \tag{13}$$

- The solution existence of the least squares

We can write the state vector of the errors at the time n as follows:

$$\Delta\boldsymbol{\varphi}_n = \left[\Delta\boldsymbol{\varphi}_n^{a\,\mathrm{T}} \ \ \Delta\boldsymbol{\varphi}_n^{k\,\mathrm{T}} \ \ \Delta\boldsymbol{\varphi}_n^{b\,\mathrm{T}} \right]^{\mathrm{T}} \tag{14}$$

where

$$\Delta\boldsymbol{\varphi}_n^a = \begin{bmatrix} \Delta\mathbf{a}_n \\ \Delta\alpha_n \end{bmatrix}, \ \Delta\boldsymbol{\varphi}_n^k = \begin{bmatrix} \Delta\tilde{\mathbf{k}}_n \\ \Delta\gamma_n \end{bmatrix}, \ \Delta\boldsymbol{\varphi}_n^b = \begin{bmatrix} \Delta\mathbf{b}_n \\ \Delta\beta_n \end{bmatrix} \tag{15}$$

represent respectively the errors cumulated up until the time n in the forward, Kalman and backward recursive variables. The $(3L+3)\mathrm{x}(3L+3)$ dimensional matrix $\mathbf{F}(n)$ given by:

$$\mathbf{F}(n) = \begin{bmatrix} \mathbf{F}_{11}(n)\,\mathbf{F}_{12}(n)\,\mathbf{F}_{13}(n) \\ \mathbf{F}_{21}(n)\,\mathbf{F}_{22}(n)\,\mathbf{F}_{23}(n) \\ \mathbf{F}_{31}(n)\,\mathbf{F}_{32}(n)\,\mathbf{F}_{33}(n) \end{bmatrix} \tag{16}$$

represents the propagation matrix of the errors.

The stability of system (13) depends on the study of the matrix properties. Then, if all of the eigenvalues of the matrix $\mathbf{F}(n)$ are less than one in magnitude, the algorithm is numerically stable locally about its optimum solution.

B. Numerical Stability

Using some redundant formulae of FRLS algorithms, we can calculate differently the backward a priori prediction errors in tree ways: $\bar{r}_n$ is given by definition using the backward prediction $\mathbf{b}_n$ and the input signal; $\bar{r}_n^{\mathrm{f}_0}$ is calculated by using in its formula the backward prediction error variance β_n and $\tilde{k}_{L+1,n}^{+}$ the $(L+1)^{\mathrm{th}}$ coefficient of the normalized Kalman gain vector $\tilde{\mathbf{k}}_{L+1,n}^{+}$ of order $L+1$; $\bar{r}_n^{\mathrm{f}_1}$ is calculated by using in its formula the likelihood variable γ_n, the forward predictor error variance α_n and $\tilde{k}^{+}_{L+1,n}$.

By making the difference between these backward a priori prediction errors, we have defined a variable called divergence indicator ξ_n [8]:

$$\xi_n = \bar{r}_n - \bar{r}^{\mathrm{f}}_n \quad \begin{cases} = 0 & \text{theory} \\ \neq 0 & \text{practical} \end{cases} \tag{17}$$

where

$$\bar{r}^{\mathrm{f}}_n = [(1-\mu_s)\bar{r}^{\mathrm{f}_0}_n + \mu_s \bar{r}^{\mathrm{f}_1}_n] \tag{18}$$

with

$$0 \leq \mu_s \leq 1 \tag{19}$$

In practice, the variable ξ_n that depends of the recursive variable is never null, due to the precision of machines used. This variable, theoretically null, does not modify the structure of the algorithm. Also, its introduction in an unspecified point of the algorithm modifies its numerical properties. We define three backward a priori prediction errors ($\bar{r}^{\gamma}_n$, $\bar{r}^{\beta}_n$ and $\bar{r}^{b}_n$), theoretically equivalents, which will be used to calculate the variables γ_n, β_n and $\mathbf{b}_n$ respectively. By introducing our variable into the algorithm and using suitably the scalar parameters $(\mu^{\gamma}, \mu^{\beta}, \mu^{b})$ and μ_s, the propagation matrix $\mathbf{F}(n)$ is modified to obtain the numerical stability.

A version of numerically stable (NS-FRLS) algorithms [9] is given in Table I. The resulting stabilized FRLS algorithms have a complexity of 8*L*. Let us note that for $\mu_s = 0$ and $\mu^{\gamma} = \mu^{\beta} = \mu^{b} = -1$, the algorithm corresponds to the FTF (Fast Transversal Filter) [4], numerically unstable.

C. Analysis Prediction Part

The study of the stability of matrix for the numerical errors propagation $\mathbf{F}(n)$ in all its general case is a very difficult task because of its complexity. However, for version that we developed, we can deduce that the matrix $\mathbf{F}_{11}(n)$ has all its eigenvalues lower than one and that the matrix $\mathbf{F}_{13}(n)$ is null. We can thus say that the matrices $\mathbf{F}_{12}(n)$ and $\mathbf{F}_{23}(n)$ can be made negligible by choosing of a forgetting factor λ very close to one. In this case, the matrix $\mathbf{F}(n)$ is very similar to a lower triangular matrix per blocks; we could therefore write $\mathbf{F}(n)$ as follows:

$$\mathbf{F}(n) = \begin{bmatrix} \mathbf{F}_{11}(n) & \mathbf{0}_{L+1} & \mathbf{0}_{L+1} \\ \mathbf{F}_{21}(n) & \mathbf{F}_{22}(n) & \mathbf{0}_{L+1} \\ \mathbf{F}_{31}(n) & \mathbf{F}_{32}(n) & \mathbf{F}_{33}(n) \end{bmatrix} \tag{20}$$

Thus we can approach the errors cumulated until the time *n* in Kalman $\Delta\boldsymbol{\varphi}^{k}_n$ and backward $\Delta\boldsymbol{\varphi}^{b}_n$ recursive variables by:

$$\Delta\boldsymbol{\varphi}^{k}_n = \mathbf{F}_{22}(n)\Delta\boldsymbol{\varphi}^{k}_{n-1} + \mathrm{T}_k\left[\Delta\boldsymbol{\varphi}^{a}_{n-1}\right] \tag{21a}$$

$$\Delta\boldsymbol{\varphi}^{b}_n = \mathbf{F}_{33}(n)\Delta\boldsymbol{\varphi}^{b}_{n-1} + \mathrm{T}_b\left[\Delta\boldsymbol{\varphi}^{a}_{n-1}; \Delta\boldsymbol{\varphi}^{k}_{n-1}\right] \tag{21b}$$

where $\mathrm{T}_k[.]$ and $\mathrm{T}_b[.]$, non explicit functions, which do not depend on $\Delta\boldsymbol{\varphi}^{k}_{n-1}$ and $\Delta\boldsymbol{\varphi}^{b}_{n-1}$ respectively. The matrices $\mathbf{F}_{22}(n)$ and $\mathbf{F}_{33}(n)$ can be written as follows:

$$\mathbf{F}_{22}(n) = \begin{bmatrix} \mathbf{M}_k(n) & \mathbf{0}_L \\ * & c_{\gamma}(n) \end{bmatrix} \tag{22a}$$

$$\mathbf{F}_{33}(n) = \begin{bmatrix} \mathbf{M}_b(n) & \mathbf{u}(n) \\ * & c_{\beta}(n) \end{bmatrix} \tag{22b}$$

with

$$c_{\gamma}(n) = (\rho_n \theta_n^{-1})^2 + (1 + 2\mu^{\gamma}\mu_s)\rho_n \theta_n^{-1}(1-\theta_n^{-1}) \tag{23}$$

$$\rho_n = \frac{\lambda \alpha_{n-1}}{\alpha_n} \quad (24); \qquad \theta_n = \frac{\lambda \beta_{n-1}}{\beta_n} \tag{25}$$

$$\mathbf{M}_b(n) = \left[\mathbf{I}_L \theta_n^{-1} - c_b(n)\mathbf{R}_n^{-1}\mathbf{x}_n\mathbf{x}_n^{\mathrm{T}}\right] \tag{26a}$$

$$c_b(n) = (1+\mu^{b}) - 2(1-\theta_n^{-1})(1+\mu^{\gamma}) \tag{26b}$$

$$c_{\beta}(n) = \lambda\left\{1 + 2(1-\mu_s)(1-\theta_n^{-1})\left(\mu^{\beta} - \mu^{\gamma}(1-\theta_n^{-1})\right)\right\} \tag{27}$$

$$\mathbf{u}(n) = \lambda(1-\mu_s)\left\{2(1-\theta_n^{-1})\mu^{\gamma} - \mu^{b}\right\}\gamma_n \tilde{k}^{+}_{L+1,n}\tilde{\mathbf{k}} \tag{28}$$

where $\mathbf{M}_k(n)$ called companion matrix with eigenvalues equal to the poles of the backward predictor $\mathbf{b}_{n-1}$, the quantities (*) do not influence on the study and $\mathbf{I}_L$ represents the identity matrix.

By taking the following expression:

$$\tilde{k}^{+}_{L+1,n} = \mathbf{q}^{\mathrm{T}}_{L+1}\tilde{\mathbf{k}}^{+}_{L+1,n} = \lambda^{-1}\mathbf{q}^{\mathrm{T}}_{L+1}\mathbf{R}^{-1}_{L+1,n-1}\mathbf{x}_{L+1,n} \tag{29}$$

where the vector $\mathbf{q}_{L+1} = [0,0, \ ...,0,1]^{\mathrm{T}}$ makes it possible to extract the (*L*+1)[th] component from a vector of (*L*+1) order and by using the following approximation:

$$\tilde{k}^{+}_{L+1,n} \cong -(1-\lambda^{-1})\mathbf{q}^{\mathrm{T}}_{L+1}\mathbf{R}^{-1}_{L+1,X}\mathbf{x}_{L+1,n} \tag{30}$$

with $\mathbf{R}_{L+1,X} = \mathrm{E}\left\{\mathbf{x}_{L+1,n}\mathbf{x}^{\mathrm{T}}_{L+1,n}\right\}$, where $\mathrm{E}\{.\}$ denotes expected value. A forgetting factor λ close to one make it possible to weaken the influence of a vector $\mathbf{u}(n)$, and to approach the numerical errors in the calculation $\Delta\boldsymbol{\varphi}^{k}_n$ and $\Delta\boldsymbol{\varphi}^{b}_n$ by first order models deduced respectively from $\mathbf{F}_{22}(n)$ and $\mathbf{F}_{33}(n)$:

$$\Delta\gamma_n = c_{\gamma}(n)\Delta\gamma_{n-1} + p_{\gamma}(n) \tag{31a}$$

$$\Delta\mathbf{b}_n = \mathbf{M}_b(n)\Delta\mathbf{b}_{n-1} + \mathbf{p}_b(n) \tag{31b}$$

$$\Delta\beta_n = c_{\beta}(n)\Delta\beta_{n-1} + p_{\beta}(n) \tag{31c}$$

By assuming that the perturbation terms ($p_{\gamma}(n)$, $\mathbf{p}_b(n)$ and $p_{\beta}(n)$) remain limited. The choice of the control parameters, so that the system is stable, amounts to studying the scalars gain ($c_{\gamma}(n)$, $c_{\beta}(n)$) and matrix gain $\mathbf{M}_b(n)$. More precisely, it is necessary that these gains are lower than one. There are an infinity number of solutions for the choice of the control parameters from these equations and for each solution there is a stability condition on λ.

The analysis using the mean behavior of system (21) does not make it possible to find the true stability condition, because the numerical errors can be of null mean but of unlimited variance. Let us calculate the variance of $\Delta\mathbf{b}_n$, for that we use the statistical approach, $\mathrm{E}\{\Delta\mathbf{b}_n\Delta\mathbf{b}_n^{\mathrm{T}}\}=\sigma^2_{\Delta b,n}\mathbf{I}_L$, we can write:

$$\begin{aligned}\mathrm{E}\{\Delta\mathbf{b}_n\Delta\mathbf{b}_n^{\mathrm{T}}\} &= \mathrm{E}\{(\theta_n^{-1})^2\}\mathrm{E}\{\Delta\mathbf{b}_{n-1}\Delta\mathbf{b}_{n-1}^{\mathrm{T}}\}\\ &-\mathrm{E}\{\theta_n^{-1}\,c_b(n)\}\mathrm{E}\{\mathbf{R}_{N,t}^{-1}\mathbf{x}_n\mathbf{x}_n^T\}\mathrm{E}\{\Delta\mathbf{b}_{n-1}\Delta\mathbf{b}_{n-1}^{\mathrm{T}}\}\\ &-\mathrm{E}\{\theta_n^{-1}\,c_b(n)\}\mathrm{E}\{\Delta\mathbf{b}_{n-1}\Delta\mathbf{b}_{n-1}^{\mathrm{T}}\}\mathrm{E}\{\mathbf{R}_{N,t}^{-1}\mathbf{x}_n\mathbf{x}_n^T\}^{\mathrm{T}}\\ &+\mathrm{E}\{c_b^2(n)\}\mathrm{E}\left\{\mathbf{R}_n^{-1}\mathbf{x}_n\mathbf{x}_n^T\Delta\mathbf{b}_{n-1}\Delta\mathbf{b}_{n-1}^{\mathrm{T}}\left(\mathbf{R}_n^{-1}\mathbf{x}_n\mathbf{x}_n^T\right)^{\mathrm{T}}\right\}\\ &+\mathrm{E}\{\mathbf{p}_b(n)\,\mathbf{p}_b^T(n)\}\end{aligned} \tag{32}$$

We assume that the elements of disturbing vector $\mathbf{p}_b(n)$ are sequences of white noise, mean-zero and of known variance $\sigma^2_{p_b}$. We notice that the variables θ_n and $\mathbf{R}_n$ are slow variable quantities compared to the input signal x_n. And we assume that the components of vector $\Delta\mathbf{b}_n$ are independent between them and independent of the various theoretical variables given in the algorithm. Moreover, the input signal is a sequence of uncorrelated Gaussian variables, $\mathrm{E}\{\mathbf{x}_n\mathbf{x}_n^{\mathrm{T}}\}=\mathbf{R}_{xx}=\sigma_x^2\mathbf{I}_L$, and the variable θ_n has an asymptotic value λ and $\mathrm{E}\{\mathbf{R}_n^{-1}\mathbf{x}_n\mathbf{x}_n^{\mathrm{T}}\}=(1-\lambda)\mathbf{I}_L$ [12]. We can write:

$$\begin{aligned}\sigma^2_{\Delta b,n}\mathbf{I}_L &= \lambda^{-2}\sigma^2_{\Delta b,n-1}\mathbf{I}_L \;-2\lambda^{-1}(1-\lambda)\,c_b\,\sigma^2_{\Delta b,n-1}\mathbf{I}_L\\ &+c_b^2(1-\lambda)^2\sigma_x^{-4}\mathbf{H}(n)+\sigma^2_{p_b}\mathbf{I}_L\end{aligned} \tag{33}$$

with

$$c_b=\left(1+\mu^b\right)-2\left(1-\lambda^{-1}\right)\left(1+\mu^\gamma\right) \tag{34}$$

where

$$\mathbf{H}(n)=\mathrm{E}\{\mathbf{x}_n\mathbf{x}_n^T\Delta\mathbf{b}_{n-1}\Delta\mathbf{b}_{n-1}^{\mathrm{T}}\mathbf{x}_n\mathbf{x}_n^T\} \tag{35}$$

is a $(L\times L)$ square matrix whose elements are given by :

$$h_{ij}=\sum_{k=1}^{L}\sum_{m=1}^{L}\mathrm{E}\{\Delta b_{k,n-1}\Delta b_{m,n-1}x_{n-k+1}x_{n-m+1}x_{n-i+1}x_{n-j+1}\} \tag{36}$$

By developing the expression (36), we find:

$$\begin{aligned}h_{ij}=\sum_{k=1}^{L}\Big[&\sigma^2_{\Delta b,n-1}\,\mathrm{E}\{x^2_{n-k+1}\}\mathrm{E}\{x_{n-i+1}x_{n-j+1}\}\\ &+\sigma^2_{\Delta b,n-1}\,\mathrm{E}\{x_{n-k+1}x_{n-i+1}\}\mathrm{E}\{x_{n-i+1}x_{n-j+1}\}\\ &+\sigma^2_{\Delta b,n-1}\,\mathrm{E}\{x_{n-k+1}x_{n-j+1}\}\mathrm{E}\{x_{n-i+1}x_{n-j+1}\}\Big]\end{aligned} \tag{37}$$

If $i\neq j \;\Rightarrow\; h_{ij}=0$; and if $i=j\;\Rightarrow$

$$h_{ii}=\sigma^2_{\Delta b,n-1}\sigma_x^2\sum_{k=1}^{L}\left[\mathrm{E}\{x^2_{n-k+1}\}+2\mathrm{E}\{x_{n-k+1}x_{n-i+1}\}\right] \tag{38}$$

We get:

$$h_{ij}=\begin{cases}\sigma^2_{\Delta b,n-1}\,\sigma_x^4\;(L+2) & i=j\\ 0 & i\neq j\end{cases} \tag{39}$$

Finally, we obtain the following expression:

$$\sigma^2_{\Delta b,n}=G\;\sigma^2_{\Delta b,n-1}+\sigma^2_{p_b} \tag{40}$$

where

$$G=\lambda^{-2}-2\,c_b\lambda^{-1}(1-\lambda)+c_b^2(1-\lambda)^2(L+2) \tag{41}$$

By assuming that $\sigma^2_{p_b}$ is limited, the stability condition of equation (40) is given by the solution of the following inequality:

$$|\,G\,|<1 \tag{42}$$

By studying the stability of system (40) for a suitable choice of the control parameters, then expression (41) will be a function of λ and L only. For appropriate choices we selected the following control parameters:

$$\mu^\gamma=0,\;\;\mu^\beta=\mu^b=1\;\text{ and }\;0\leq\mu_s\leq 1 \tag{43}$$

By applying condition (42), we get:

$$\lambda>\frac{4L+5}{4L+7}=1-\frac{1}{2L+3.5} \tag{44}$$

These conditions can be written in another simpler form:

$$\lambda=1-\frac{1}{pL} \tag{45}$$

where the parameter p is a real number strictly greater than 2 to ensure numerical stability.

D. Analysis Filtering Part

The analysis uses the common independence assumption that the current input signal vector is statistically independent of the current coefficient vector of the adaptive filter [1]. We define the weight-error vector at time n as:

$$\Delta\mathbf{w}_n=\mathbf{w}_{opt}-\mathbf{w}_n \tag{46}$$

The output a priori error $\bar{\varepsilon}_n$ can be written as:

$$\bar{\varepsilon}_n=v_n+\mathbf{x}_n^T\Delta\mathbf{w}_{n-1} \tag{47}$$

The recursion in (5) on the coefficient error vector is:

$$\Delta\mathbf{w}_n=[\mathbf{I}_L-\mathbf{g}_n\mathbf{x}_n^{\mathrm{T}}]\Delta\mathbf{w}_{n-1}-\mathbf{g}_n v_n \tag{48}$$

The mean behavior of the RLS coefficient error vector can now be determined by taking the expected value of both sides of (48) and using the independence assumption to yield [1]

$$\mathrm{E}\{\Delta\mathbf{w}_n\}=\mathrm{E}\{\left[\mathbf{I}_L-\mathbf{R}_n^{-1}\mathbf{x}_n\mathbf{x}_n^T\right]\Delta\mathbf{w}_{n-1}\}-\mathrm{E}\{\mathbf{R}_n^{-1}\mathbf{x}_n v_n\} \tag{49a}$$

We obtain:

$$\mathrm{E}\{\Delta\mathbf{w}_n\}=\lambda\,\mathrm{E}\{\Delta\mathbf{w}_{n-1}\} \tag{49b}$$

The steady-state solution of (49b) is:

If $\lambda<1 \;\Rightarrow\; \mathrm{E}\{\Delta\mathbf{w}(\infty)\}=0_L$; from which we obtain the steady-state mean coefficient vector of the RLS adaptive filter as:

$$\mathrm{E}\{\mathbf{w}(\infty)\}=\mathbf{w}_{\mathrm{opt}} \tag{50}$$

The mean square error $MSE(n)=\mathrm{E}\{\bar{\varepsilon}_n^2\}$ can be written as:

$$MSE(n)=\sigma_v^2+\mathrm{tr}\left[\mathbf{R}_{xx}\,\mathrm{E}\{\Delta\mathbf{w}_{n-1}\Delta\mathbf{w}_{n-1}^{\mathrm{T}}\}\right] \tag{51}$$

We obtain:

$$MSE(n)=\sigma_v^2+\sigma_x^2\,\mathrm{E}\left\{\|\Delta\mathbf{w}_{n-1}\|^2\right\} \tag{52}$$

where $\mathrm{E}\{v_n^2\}=\sigma_v^2$, tr[.] represents the trace operator and $\|.\|$

denotes the 2-norm vector.

Table I: New version of numerically stable fast recursive least squares (NS-FRLS) algorithm

Initialization:

$\mathbf{w}_0 = \mathbf{a}_0 = \mathbf{b}_0 = \tilde{\mathbf{k}}_0 = 0_L$; $\gamma_0 = 1$; $\alpha_0 = \lambda^L E_0$; $\beta_0 = E_0$;

$E_0 \geq \sigma_x^2 \frac{L}{100}$

Variables available at the discrete-time index n**:**

$\mathbf{a}_{n-1}; \mathbf{b}_{n-1}; \tilde{\mathbf{k}}_{n-1}; \gamma_{n-1}; \alpha_{n-1}; \beta_{n-1}; \mathbf{w}_{n-1}$.

New information: x_n and d_n.

- **Prediction Part:**

Modeling of x_n and x_{n-L}

$\bar{e}_n = x_n - \mathbf{a}_{n-1}^{\mathrm{T}} \mathbf{x}_{n-1}$;

$\tilde{\mathbf{k}}_{L+1,n}^{+} = \begin{bmatrix} \tilde{\mathbf{k}}_n^{+} \\ \tilde{k}_{L+1,n}^{+} \end{bmatrix} = \begin{bmatrix} 0 \\ \tilde{\mathbf{k}}_{n-1} \end{bmatrix} + \frac{\bar{e}_n}{\lambda \alpha_{n-1}} \begin{bmatrix} 1 \\ -\mathbf{a}_{n-1} \end{bmatrix}$;

$\mathbf{a}_n = \mathbf{a}_{n-1} + \bar{e}_n \gamma_{n-1} \tilde{\mathbf{k}}_{n-1}$;

$\alpha_n = \lambda \alpha_{n-1} + \gamma_{n-1} \bar{e}_n^2$

$\bar{r}_n = x_{n-L} - \mathbf{b}_{n-1}^{\mathrm{T}} \mathbf{x}_n$;

$\bar{r}_n^{\mathrm{f}_0} = \lambda \beta_{n-1} \tilde{k}_{L+1,n}^{+}$; $\bar{r}_n^{\mathrm{f}_1} = \lambda^{-L+1} \gamma_{n-1} \alpha_{n-1} \tilde{k}_{L+1,n}^{+}$;

Numerical Stability

$\xi_n = \bar{r}_n - [(1-\mu_s)\bar{r}_n^{\mathrm{f}_0} + \mu_s \bar{r}_n^{\mathrm{f}_1}]$;

$\bar{r}_n^{\gamma} = \bar{r}_n + \mu^{\gamma} \xi_n$; $\bar{r}_n^{\beta} = \bar{r}_n + \mu^{\beta} \xi_n$; $\bar{r}_n^{b} = \bar{r}_n + \mu^{b} \xi_n$;

$\tilde{\mathbf{k}}_n = \tilde{\mathbf{k}}_n^{+} + \tilde{k}_{L+1,n}^{+} \mathbf{b}_{n-1}$;

$\gamma_n = \frac{\lambda \alpha_{n-1}}{\alpha_n - \lambda^L \bar{r}_n^{\gamma} \bar{r}_n^{\mathrm{f}_1}} \gamma_{n-1}$;

$\mathbf{b}_n = \mathbf{b}_{n-1} + \bar{r}_n^{b} \gamma_n \tilde{\mathbf{k}}_n$;

$\beta_n = \lambda \beta_{n-1} + \gamma_n (\bar{r}_n^{\beta})^2$;

- **Filtering Part:**

$\bar{\varepsilon}_n = d_n - \mathbf{w}_{n-1}^{\mathrm{T}} \mathbf{x}_n$;

$\mathbf{w}_n = \mathbf{w}_{n-1} + \bar{\varepsilon}_n \gamma_n \tilde{\mathbf{k}}_n$.

We define the normalized misalignment in dB as follows:

$$NM(n) = 10 \log_{10} \mathrm{E} \left\{ \frac{\left\| \Delta \mathbf{w}_n \right\|^2}{\left\| \mathbf{w}_{opt} \right\|^2} \right\} \tag{53}$$

which measures the mismatch between the true impulse response and the modeling filter. For that, we need to determine the next expressions:

$$\begin{aligned} \mathrm{E}\{\Delta \mathbf{w}_n \Delta \mathbf{w}_n^{\mathrm{T}}\} = \ & \mathrm{E}\{\Delta \mathbf{w}_{n-1} \Delta \mathbf{w}_{n-1}^{\mathrm{T}}\} - \mathrm{E}\{\Delta \mathbf{w}_{n-1} \Delta \mathbf{w}_{n-1}^{\mathrm{T}}\} \mathrm{E}\{\mathbf{g}_n \mathbf{x}_n^{\mathrm{T}}\}^{\mathrm{T}} \\ & - \mathrm{E}\{\mathbf{g}_n \mathbf{x}_n^{\mathrm{T}}\} \mathrm{E}\{\Delta \mathbf{w}_{n-1} \Delta \mathbf{w}_{n-1}^{\mathrm{T}}\} \\ & + \mathrm{E}\{\mathbf{g}_n \mathbf{x}_n^{\mathrm{T}} \Delta \mathbf{w}_{n-1} \Delta \mathbf{w}_{n-1}^{\mathrm{T}} \mathbf{g}_n \mathbf{x}_n^{\mathrm{T}}\} + \mathrm{E}\{\mathbf{g}_n \mathbf{g}_n^{\mathrm{T}}\} \mathrm{E}\{\nu_n^2\} \end{aligned} \tag{54}$$

where

$$\mathrm{E}\{\mathbf{g}_n \mathbf{g}_n^{\mathrm{T}}\} = \mathrm{E}\{\mathbf{R}_n^{-1} \mathbf{x}_n \mathbf{x}_n^{\mathrm{T}} \mathbf{R}_n^{-1}\} = (1-\lambda)^2 \mathbf{R}_{xx}^{-1} \tag{55}$$

So,

$$\mathrm{E}\{\Delta \mathbf{w}_n \Delta \mathbf{w}_n^{\mathrm{T}}\} = \lambda^2 \, \mathrm{E}\{\Delta \mathbf{w}_{n-1} \Delta \mathbf{w}_{n-1}^{\mathrm{T}}\} + (1-\lambda)^2 \frac{\sigma_\nu^2}{\sigma_x^2} \mathbf{I}_L \tag{56}$$

By taking the trace of both sides of (56), we find:

$$\mathrm{tr}\left[\mathrm{E}\{\Delta \mathbf{w}_n \Delta \mathbf{w}_n^{\mathrm{T}}\}\right] = \lambda^2 \, \mathrm{tr}\left[\mathrm{E}\{\Delta \mathbf{w}_{n-1} \Delta \mathbf{w}_{n-1}^{\mathrm{T}}\}\right] + (1-\lambda)^2 \frac{\sigma_\nu^2}{\sigma_x^2} L \tag{57}$$

Finally we can write the mean square norm coefficient error vector as:

$$\mathrm{E}\left\{\left\| \Delta \mathbf{w}_n \right\|^2\right\} = \lambda^2 \, \mathrm{E}\left\{\left\| \Delta \mathbf{w}_{n-1} \right\|^2\right\} + (1-\lambda)^2 \frac{\sigma_\nu^2}{\sigma_x^2} L \tag{58}$$

The stability of the recursion (58) is guaranteed if $\lambda < 1$.

IV. Simulations

To confirm the validity of our analysis and demonstrate the improved numerical performance, some simulations are carried out. In order to evaluate the numerical stability of the different algorithms, all simulations were performed in 32-bit single-precision floating-point representation. The input signal x_n used in our simulation is a white Gaussian noise, with mean zero and variance one.

A. Prediction Part

We are only interested in right process of the prediction part algorithm because divergences only concerns the prediction part; the filtering part is robust compared to the numerical implementation, it only requires a forgetting factor lower than one [2] and the right functioning of the adaptation gain (6) provided by the prediction part. It is stable if the latter one is. For that, we evaluate for the good behavior of the divergence indicator ($\xi_n \to 0$) and the likelihood variable ($0 < \gamma_n \leq 1$). We define thereafter the normalized square norm gain-error in dB by:

$$NGE(n) = 10 \log_{10} \left(\frac{\mathrm{E}\{\| \Delta \mathbf{g}_n \|^2\}}{\mathrm{E}\{\| \mathbf{g}_n \|^2\}} \right) \tag{59}$$

where $\Delta \mathbf{g}_n = \left(\mathbf{R}_n^{-1} \mathbf{x}_n - \gamma_n \tilde{\mathbf{k}}_n\right)$ is gain-error vector. Variable $NGE(n)$ makes it possible to measure the mismatch between the gains $\mathbf{R}_n^{-1} \mathbf{x}_n$ and $\gamma_n \tilde{\mathbf{k}}_n$ calculated by RLS and NS-FRLS algorithms respectively. For a suitable choice of the parameters, we checked the validity of the numerical stability condition on λ given by equation (45). In addition, if this condition is not satisfied then the proposed algorithm diverges.

For a filter of order L=32 and μ_s =0.5, Fig.2 and Fig.3, show the evolution of variables γ_n, ξ_n and $NGE(n)$ correspond to the case numerically unstable (p<2) and numerically stable (p>2) respectively.

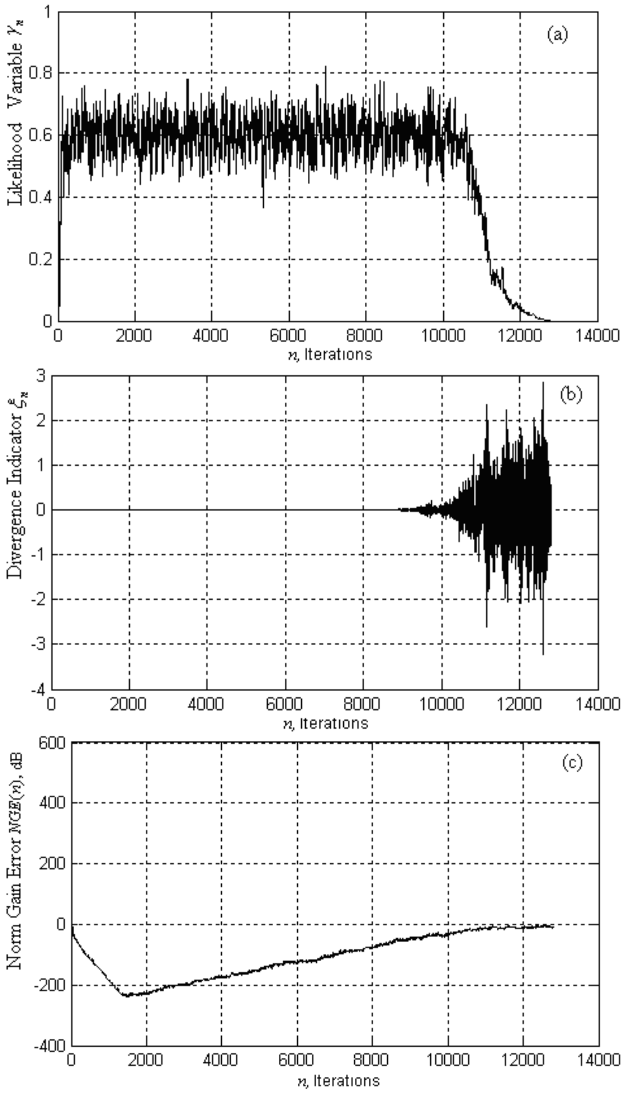

Fig. 2: Evolution of the variables γ_n , ξ_n and $NGE(n)$ for L = 32, p = 1. 95, μ_s = 0. 5

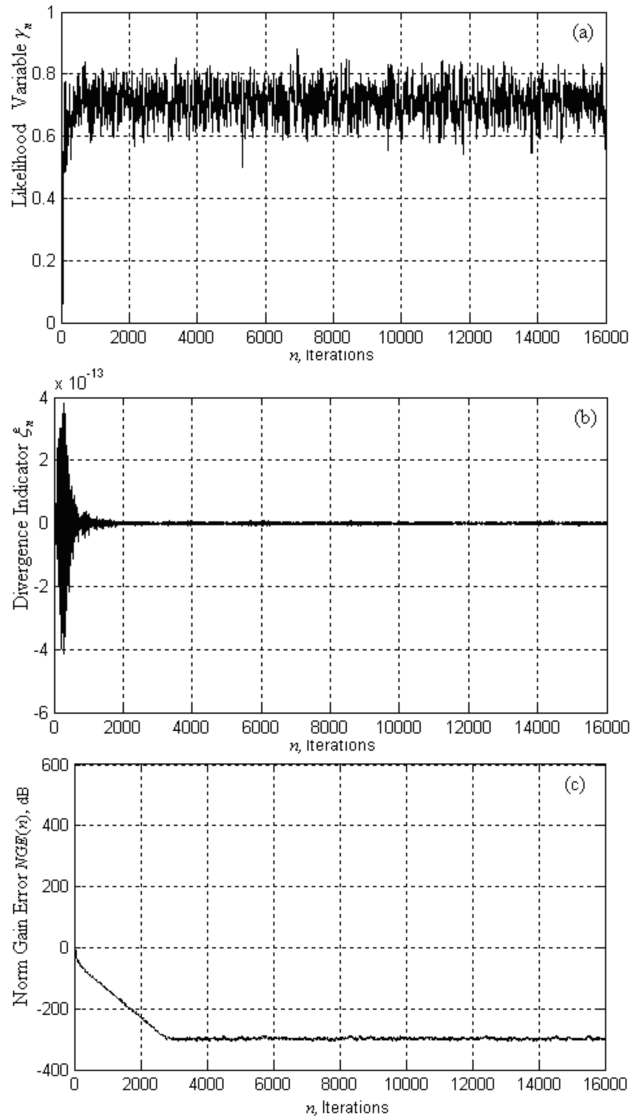

Fig. 3 Evolution of the variables γ_n , ξ_n and $NGE(n)$ for L = 32, p = 3, μ_s = 0. 5

During the first iterations, the likelihood variable γ_n (Fig.2a) is almost constant (asymptotic value), then it starts to oscillate, to finish with a divergence. The divergence indicator ξ_n (Fig.2b) increases indefinitely until total divergence of the algorithm. For the normalized squared norm gain-error $NGE(n)$ (Fig.2c), we notice that it diverges well before the other variables; it increases indefinitely until the total divergence of the algorithm. Fig.3 illustrates the stability of this version, where the divergence indicator ξ_n remains very weak, and the likelihood variable γ_n fluctuates around its optimal value. We notice that variable $NGE(n)$ converges and remains stable (simulations were run for more than 10^7 samples). The algorithm was tested successfully by simulations at very long term and for different orders of filter to improve the stability of the algorithm.

B. Filtering Part

We try to estimate an impulse response $\mathbf{w}_{\text{opt}}$ of length L=32 the same length is used for the adaptive filter $\mathbf{w}_n$. Performance of the estimation is measured by the mean square error $MSE(n)$ and normalized misalignment $NM(n)$. The reference signal d_n is obtained by convolving $\mathbf{w}_{\text{opt}}$ with x_n and adding a white Gaussian noise signal v_n with the signal-to-noise ratio (*SNR*) is equal to 50 dB. We run the NS-FRLS algorithm with a forgetting factor λ (45) where p=3. Fig.4 illustrates the behaviour of the mean square error $MSE(n)$ obtained from simulations and determined from the theoretical expression in (52). From this plot, we observe that simulation and theoretical curves agree very well. Fig.6 plots the normalized misalignment $NM(n)$ as obtained from the theoretical analysis (53) and from simulation results. It can be seen that, for adaptive filter, there is a good agreement between the actual behavior of the algorithm and that predicted by the theoretical expression. Note that the noise effect does not severely degrade the performance of adaptive estimator. This indicates that the proposed algorithm is robust.

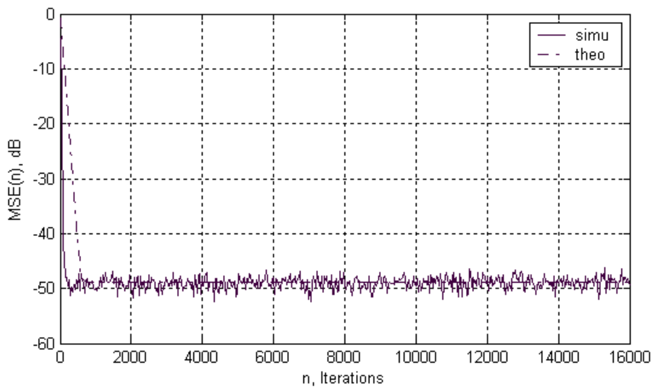

Fig. 4: Comparison of theoretical and simulation curves of the $MSE(n)$ for L=32, p = 3, μ_s = 0. 5

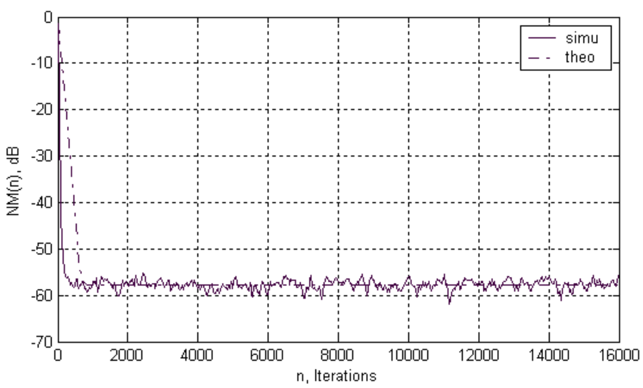

Fig. 5: Comparison of theoretical and simulation curves of the $NM(n)$ for L=32, p = 3, μ_s = 0. 5

V. Conclusion

We have analyzed the numerical properties of the stabilized FRLS algorithm by using a propagation model of first order of the numerical errors. We have presented new version of a numerically stable fast recursive least squares (NS-FRLS) algorithm. The condition of stabilization was shown to be capable of maintaining a good convergence performance by way of computer simulations. This algorithm thus modified is stable numerically for a suitable choice of the control parameters and the forgetting factor. The introduced variable $NGE(n)$, allowed us to compare and to show the stability evolution of the algorithms. Simulation experiments related to the mean square error and the normalized misalignment proved well the validity of the analysis.Numerical stability was checked by simulation over very long duration of stationary signal and for different orders of filter.

References

[1] S. Haykin, *Adaptive Filter Theory,* 4th ed. Englewood Cliffs, NJ: Prentice-Hall, 2002.

[2] A. H. Sayed, *Fundamentals of Adaptive Filtering*, JohnWiley &Sons, New York, NJ, USA, 2003.

[3] J.R. Treichler, C.R. Johnson and M.G. Larimore, *Theory and Design of Adaptive Filter,* Prentice Hall, 2001.

[4] J. Cioffi and T. Kailath, "Fast RLS Transversal Filters for adaptive filtering," *IEEE press. On ASSP*, vol. 32, no., pp. 304–337, 1984.

[5] A. Benallal and A. Gilloire,, "A New method to stabilize fast RLS algorithms based on a first-order model of the propagation of numerical errors," *Proc. ICASS*,, New York, 1988, pp.1365-1368.

[6] D.T.M. Slock and T. Kailath, "Numerically stable fast transversal filters for recursive least squares adaptive filtering," *IEEE transactions on signal processing*, vol. 39, no. 1, pp. 92–114, 1991.

[7] A.P.Liavas and P.A.Regalia, "Numerical stability issues of the conventional recursive least squares algorithm," in *Proc. ICASSP*, Seattle, WA, 1998

[8] M. Arezki, A. Benallal, A. Guessoum and D. Berkani, "Three New Versions of Numerically Stable Fast Least Squares Algorithms (NS-FRLS) for Adaptive Filtering," in *Proc. 4th international symposium on Communication Systems Networks and Digital Signal Processing,* Newcastle, U.K, 2004, pp.528-532.

[9] M. Arezki, F. Ykhlef, P. Meyrueis and D. Berkani, "Numerically Stable Fast Least Squares Algorithms in Adaptive Filtering," in *Proc. 10th World Multi-Conference on Systemics Cybernetics and Informatics*, vol.5, Orlando, Florida, USA, 2006, pp.221-224.

[10] S. Binde, "A numerically stable fast transversal filter with leakage correction," *IEEESignal Processing Letters,* vol. 2, no. 6, pp. 114-116, 1995.

[11] G.Carayannis, D.Manolakis and N.Kalouptsidis, "A Unified View of Parametric Processing Algorithms for Prewindowed Signals," *Signal Processing,* vol.10, no. 4, 1986.

[12] O. Macchi and M. Bellanger, "Le point sur le filtrage adaptatif transverse," *Proc. 11^{e}Colloque GRETSI*, Nice, 1987, 1G-4G.

M. Arezki has received the degree of Engineer in electronics from the Ecole Nationale Polytechique ENP, Algiers, Algeria, in 1986. He obtained a Master in Electrical Engineering of university of Blida, Algeria in 1997.
He has been with the University of Blida, Algeria, as a teacher Assistant, in the department of Electronics since 1989 and a Research Assistant in the Digital Signal and Image Processing Research Laboratory since 2000. His current research interests include adaptive filtering, acoustic echo cancellation, identification, fast algorithms.