STEADY-STATE PERFORMANCE ANALYSIS OF THE LMS ADAPTIVE TIME-VARYING SECOND ORDER VOLTERRA FILTER

Mounir Sayadi*, Farhat Fnaiech* , Sébastien Guillon** and Mohamed Najim**

*E.S.S.T.T, 5 Av. Taha Hussein 1008, Tunis, Tunisia, Tel (216) 1 392 559, Fax (216)1 391 166
**Equipe signal / image, ENSERB 351 cours de la libération, 33405 Talence cedex, France.
Tel (33) 56 84 66 74, Fax (33) 56 84 84 06, email : najim@goelette.tsi.u-bordeaux.fr

ABSTRACT:

In this paper, the steady-state performance of the Least Mean Square (LMS) adaptive second order Volterra filter, with constant step-size μ, in a time-varying setting, is analysed. The quantitative evaluation of the steady-state Excess Mean Square Error (EMSE), where the contribution of the gradient misadjustment and the tracking error are well characterized, is established . The optimum step-size for time-varying second order Volterra filter is then given. Thus, we can study the correlation between the Excess MSE and the optimum step-size in one hand and the parameters of the time-varying nonlinear system, in the other hand. Furthermore, the steady-state behavior predicted by the analysis is in good agreement with the experimental results. The adaptive filter was used in a second order Volterra system identification in a non stationary environment.

1. INTRODUCTION

The convergence analysis of the LMS adaptive algorithm in the mean and the mean square context has been studied for stationary linear filter in many works [1][2]. In the non-stationary linear filtering, the tracking properties and the problem of finding an optimal value of the step size of the adaptive LMS algorithm has been investigated widely in the literature [3][4]. It is shown that the tracking Mean Square Error (MSE) results from the tradeoff between the gradient part which is μ-increasing and the lag contribution which is μ-decreasing. The decoupled character of the gradient and the lag errors in the case of linear filter is proved in [5].

However, in the non linear case and more precisely the second order Volterra filter [6][7], few number of papers dealing with the convergence analysis of the LMS adaptive algorithms were presented and limited only to the stationary environment [8]. The aim of this paper is the study of the convergence of the LMS second order Volterra filter. In a first step, we present a quantitative evaluation of the Excess Mean Squar Error and in a second step, the optimum value of the step-size is derived.

2. THE LMS ADAPTIVE NON LINEAR FILTER

The general nonlinear filtering problem may be represented by a second order Volterra series [8] of gaussian input $x(n)$ and zero-mean output $y(n)$:

$$y(n) = A^T X(n) + tr\{B[X(n)X^T(n) - R_x]\} \qquad (1)$$

Where $X(n) = [x(n),\ldots\ldots,x(n-N+1)]^T$ (2)

$$A = [a(0),\ldots\ldots,a(n-1)]^T \qquad (3)$$

and $B = \begin{bmatrix} b(0,0) & ... & b(0,N-1) \\ .. & .. & .. \\ b(N-1,0) & ... & b(N-1,N-1) \end{bmatrix}$ (4)

R_x denotes the N by N autocorrelation matrix of $x(n)$ i.e. $[R_x]_{i,j} = r_x(j-i)$. A and B are respectively the linear and quadratic filter operators. According [8], and assuming the hypothesis of a Gaussian input, its shown that the linear and quadratic filter weights which minimize the mean square error between filter output $y(n)$ and desired output $s(n)$ is given by

$$A_0 = R_x^{-1} R_{sx} \text{ and } B_0 = (1/2) R_x^{-1} T_{sx} R_x^{-1} \qquad (5)$$

Where $R_{sx} = \left[r_{sx}(0),....,r_{sx}(N-1) \right]^T$ (6)

$$T_{sx} = \begin{bmatrix} t_{sx}(0,0) & ... & t_{sx}(0,N-1) \\ .. & .. & .. \\ t_{sx}(N-1,0) & ... & t_{sx}(N-1,N-1) \end{bmatrix}$$ (7)

$r_{sx}(i) = E\left[s(n)x(n-i) \right]$ and

$t_{sx}(i,j) = E\left[s(n)x(n-i)x(n-j) \right]$ (8)

Then; the LMS algorithm for the adaptive Second order Volterra filter is given by [8]:

$$\begin{cases} A(n+1) = A(n) - 2\mu_A e(n) X(n) \\ B(n+1) = B(n) - \mu_B e(n) X(n) X^T(n) \end{cases}$$ (9)

Where $e(n) = y(n) - s(n)$ denotes the residual error of the filter. μ_A and μ_B are the step size of the algorithm, they determine the rate of adaptation of the linear and quadratic filter part. The LMS algorithm is considered as a stochastic variant of the steepest descent method [1] by assuming the independence between $A(n)$, $B(n)$ and $X(n)$. $\mu_A < 1/\lambda_{max}$ and $\mu_B < 1/\lambda^2_{max}$ are the algorithm stability conditions in a stationary environment [1][5][8]. λ_{max} is the largest eigenvalue of R_x.

3. EVALUATION OF THE EXCESS MEAN SQUARE ERROR IN THE TIME-VARYING CONTEXT

Let's now examine the ability of the LMS algorithm to operate in a non stationary environment. Assume that the linear and quadratic optimum filter weights vary following:

$$\begin{cases} A_0(n+1) = A_0(n) + \Delta A_0(n) \\ B_0(n+1) = B_0(n) + \Delta B_0(n) \end{cases}$$ (10)

$\sigma^2_{NSA} = \text{cov}(\Delta A_0(n))$ and $\sigma^2_{NSB} = \text{cov}(\Delta B_0(n))$ are defined as the scalar covariance of the filter increment standard diviation. Then the LMS algorithm has the task not only seeking the minimum point of the error performance surface but also tracking the continually changing position of this minimum point. Since gradient and lag errors are decoupled [5] and according to the procedure used in [8], the total MSE is:

$$J(n) = J_{min} + j_A(n) + j_B(n)$$ (11)

where:

$J_{min} = \text{cov}(e(n))$ if we omit the additive noise.

$$J_A(n) = E\left[tr\left\{ \Delta A^T(n) R_x \Delta A(n) \right\} \right]$$ (12)

$$J_B(n) = E\left[tr\left\{ \Delta B^T(n) R_x \Delta B(n) R_x \right\} \right]$$ (13)

$$\Delta A(n) = A(n) - A_0(n)$$ (14)

$$\Delta B(n) = B(n) - B_0(n)$$ (15)

The evaluation of the steady-state MSE for the linear part, i.e. $j_A = \lim_{n\to\infty}(j_A(n))$, is given by [2][5]:

$$J_A = J_{min}\left[\mu_A N\sigma_x^2 \right] + \frac{N\sigma^2_{NSA}}{4\mu_A}$$ (16)

In a similar fashion, we shall evaluate now the steady-state MSE for the quadratic part, i.e. $j_B = \lim_{n\to\infty}(j_B(n))$. Hence by using (9)(10) and (15), yields:

$$\Delta B(n+1) = B(n+1) - B_0(n+1) =$$

$$\Delta B(n) - \mu_B e(n) X(n) X^T(n) - \Delta B_0(n)$$ (17)

which can be reformulated as follows:

$$\Delta B(n+1) = \Delta B(n) - \mu_B\left[2R_x \Delta B(n) R_x + W(n) \right]$$
$$- \Delta B_0(n)$$ (18)

where

$$W(n) = e(n) X(n) X^T(n) - 2R_x \Delta B(n) R_x$$ (19)

Assuming that the measurement noise is uncorrelated with $\Delta B(n)$ and $\Delta B_0(n)$ is uncorrelated with the inputs $X(n)$ and with $\Delta B(n)$, this gives:

$$E\left[\left\| \Delta B(n+1) \right\|^2 \right] = E\left[\left\| \Delta B(n) \right\|^2 \right]$$
$$- 4\mu_B E\left[tr\left\{ \Delta B(n) R_x \Delta B(n) R_x \right\} \right]$$
$$+ \mu_B^2 E\left[\left\| W(n) \right\|^2 \right] + E\left[\left\| \Delta B_0(n) \right\|^2 \right]$$
$$+ 4\mu_B^2 E\left[\left\| R_x \Delta B(n) R_x \right\|^2 \right]$$ (20)

Where the matrix norm of a given matrix A is denoted by $\|A\|$ and defined by the quantity: $\|A\| = \left[tr(AA^T) \right]^2$.

In the steady state of the adaptation, and assuming slow time variation of the optimal filter weights, we can assume that :

$$E\left[\left\| \Delta B(n+1) \right\|^2 \right] = E\left[\left\| \Delta B(n) \right\|^2 \right]$$ (21)

In addition, for $\mu_B << \lambda_{max}^{-2}$, it can be proved [8] that:

$$\mu_B E\left[\left\| R_x \Delta B(n) R_x \right\|^2 \right] << E\left[tr\left\{ \Delta B(n) R_x \Delta B(n) R_x \right\} \right]$$

More over, from (20), we write:

$$4\mu_B E\left[tr\left\{\Delta B(n)R_x \Delta B(n)R_x\right\}\right] \approx \mu_B^2 E\left[\left\|W(n)\right\|^2\right]$$
$$+ E\left[\left\|\Delta B_0(n)\right\|^2\right] \quad (22)$$

Furthermore, by using the asymptotic assumption $B(n) = B_0(n)$, $W(n) = e(n)X(n)X^T(n)$, the independence between $e(n)$ and $x(n)$, and the hypothesis of a Gaussian input, it results:

$$E\left[\left\|W(n)\right\|^2\right] = J_{\min}\left[tr\left(R_x^2\right) + tr\left(R_x^2\right) + tr^2(R_x)\right] \quad (23)$$
$$= J_{\min}\left[2tr(R_x^2) + N^2(\sigma_x^2)^2\right]$$

The second term of the right hand side of (22) may be evaluated by:

$$E\left[\left\|\Delta B_0(n)\right\|^2\right] = E\left[tr\left(\Delta B_0(n)\Delta B_0^T(n)\right)\right] = N\sigma_{NSB}^2 \quad (24)$$

Then:

$$J_B = \lim_{n\to\infty} E\left[tr\left\{\Delta B(n)R_x \Delta B(n)R_x\right\}\right] =$$
$$J_{\min}\left[\frac{\mu_B}{4}\left[2tr(R_x^2) + N^2(\sigma_x^2)^2\right]\right] + \frac{N\sigma_{NSB}^2}{4\mu_B} \quad (25)$$

Note that the second term of the right hand side of (25) is the nonlinear *Excess MSE* Lag caused by variation ΔB_0 of $B_0(n)$. The first term corresponds to the stationary gradient misadjustment [8]. The resulting expression of the total steady-state *Excess MSE* is given by:

$$J = \lim_{n\to\infty}\left(j(n) - j_{\min}\right) = J_A + J_B =$$
$$J_{\min}\left[\mu_A N\sigma_x^2 + \frac{\mu_B}{4}\left[2tr(R_x^2) + N^2(\sigma_x^2)^2\right]\right] \quad (26)$$
$$+ \frac{N\sigma_{NSA}^2}{4\mu_A} + \frac{N\sigma_{NSB}^2}{4\mu_B}$$

In this simplified expression for the steady-state *Excess MSE,* the contribution of the gradient misadjustment and the tracking error are well characterized. Clearly, there is an optimum of the adaptation step size when $\mu_A = \mu_B = \mu$ which is:

$$\mu_{opt} = \sqrt{\frac{N\left(\sigma_{NSA}^2 + \sigma_{NSB}^2\right)}{J_{\min}\left[4N\sigma_x^2 + 2tr\left(R_x^2\right) + N^2(\sigma_x^2)^2\right]}} \quad (27)$$

Note that this optimum is exact only in the case of slow variation [9]. The dependence between the Excess MSE and the optimum step-size in one hand and the parameters of the time-varying nonlinear system, in the other hand, can be studied.

4. COMPUTER SIMULATIONS

In order to illustrate the performance of the algorithm, we have performed extensive computer simulations. The adaptive LMS filter was used to estimate a second order Volterra system with gaussian input x_k and output y_k in a non stationary and noisy environment:

$$y_k = 0.8x_k - 0.5x_{k-1} + 0.7x_k^2 + 0.1x_{k-1}^2 - 0.4x_k x_{k-1}$$
$$+ n_k$$

For all the experiments, the signal-to-noise ratio is defined as: SNR dB = $10\log(\sigma_y^2/\sigma_{nk}^2)$. σ_{nk}^2 is the variance of the additive noise n_k. $\mu_A = \mu_B = \mu$.

The norm of the coefficient-error vector defined as : Er(dB): $10\log(\left\|\Theta - \hat{\Theta}\right\|^2 /\left\|\Theta\right\|^2)$ is plotted in Fig 3 versus time iterations for different 'variance of the non stationarity' σ_{NS} (θ is a vector containing all the time varying weights).

Fig.2 shows that the theoretical and experimental Excess MSE versus step-size are in quite good agreement. The theoretical total Excess MSE is plotted in Fig.1 versus step-size μ and in Fig.4 for different input power σ_x^2. The optimal step-size is plotted in Fig.5 versus minimum MSE $J0 = J_{\min}$.

5. CONCLUSION

In this paper, we have studied the steady-state performance of the adaptive second order Volterra filter controlled by the LMS algorithm, with constant step-size μ, in a time-varying context. We have provided a method to determine a quantitative evaluation of the steady-state *Excess MSE*. The optimum step-size for time-varying second order Volterra filter is then given. Extensive computer simulation have been performed and show the correlation between the Excess MSE and the optimum step-size in one hand and the parameters of the time-varying nonlinear system, such as input power, variance of non stationarity and step size, in the other hand. These results are exact in the case of a slow non stationarity. Further works concerning the analysis of the nonlinear LMS algorithm in the non stationary context need to be done. Note that the independence between the spectral characteristics of the input signal and the convergence behavior of the LMS algorithm imposes a bound on the step-size of the algorithm depending on the eigenvalue spread of the autocorrelation matrix of the input signal.

REFERENCES:

[1] S.Haykin " *Adaptive Filter Theory* " second edition Englewood Clifsfs, NJ: Prentice-Hall, 1991.

[2] M.G.Bellanger "*Adaptive Digital Filter and Signal Analysis*" Marcel Dekker Inc, USA, 1987.

[3] B.Widrow and J.M.McCool " Stationary and non Stationary learning characteristics of the LMS adaptive filter " *Proc. IEEE, ,* Vol. 64, pp.1151-1162, *August 1976*.

[4] O.Macchi and E.Eweda "Second-order convergence analysis of stochastic adaptive linear filter" *IEEE Trans. Automatic Control,* Vol-AC 28, no.1, pp. 76-85, *January 1983*.

[5] O.Macchi " Optimization of Adaptive Identification for Time-Varying Filters " *IEEE Trans. on Automatic Control*, vol. AC-31, pp. 283-287, 1986.

[6] G.L.Sicuranza "Quadratic filters for signal processing" *Proc. IEEE,* vol.80, no.8, pp. 1263-1285, August 1992.

[7] V.J.Mathews "Adaptive polynomial filters" *IEEE Signal Processing Magasine*, pp.10-26, July 1991.

[8] T.Koh and J.E.Powers " Second order Volterra filtering and its application to nonlinear system identification " *IEEE Trans. on Acoustic, Speech and Signal Processing*, Vol. ASSP-33, pp. 1445-1455, December 1985.

[9] E.Eweda and O.Macchi, " Tracking error bounds of adaptive nonstationary filtering," Automatica, vol.21, no.3, pp.293-302, 1985.
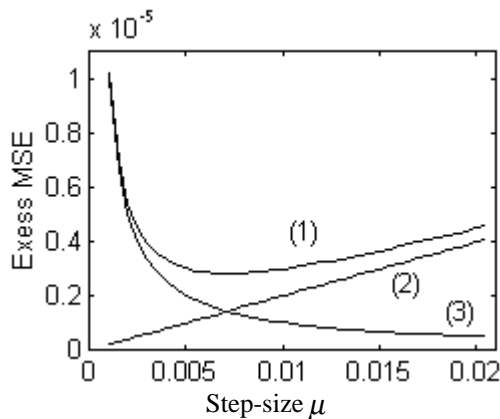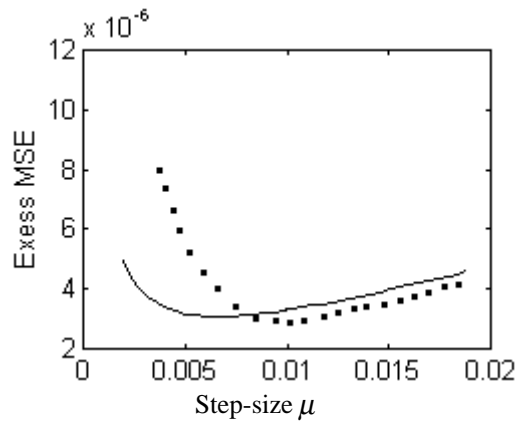
Fig.2: Theoretical (—) and experimental (■■) Excess MSE versus step-size. (SNR=25dB, $\sigma_x^2 = 0.17$, $\mu_{opt} = 0.0071$, $\sigma_{NSA} = \sigma_{NSB} = 10^{-4}$)
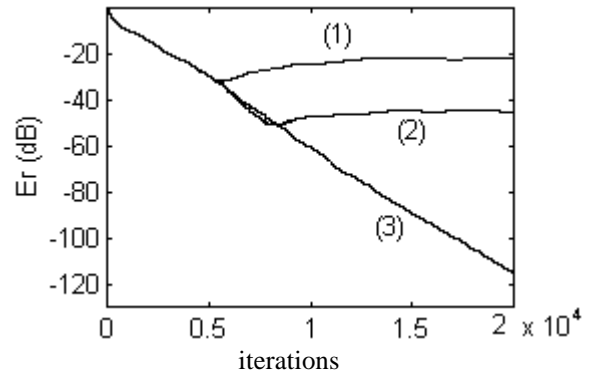

iterations

Fig.3:Norm of coef-error vector versus time for different σ_{NS} : (1)$\sigma_{NS}.=10^{-4}$, (2)$\sigma_{NS}.=10^{-5}$, (3)$\sigma_{NS}.=0$ (SNR=∞DB,σ_x^2=0.17,$\mu = \mu_{opt}$=0.0071,$\sigma_{NS}=\sigma_{NSA}=\sigma_{NSB}$)

Fig.1: Theoretical total Excess MSE (1) = stationary EMSE (2) + Lag MSE versus step-size (3) (SNR= 25dB, σ_x^2=0.17,μ_{opt}= 0.0071, $\sigma_{NSA} = \sigma_{NSB} = 10^{-4}$)
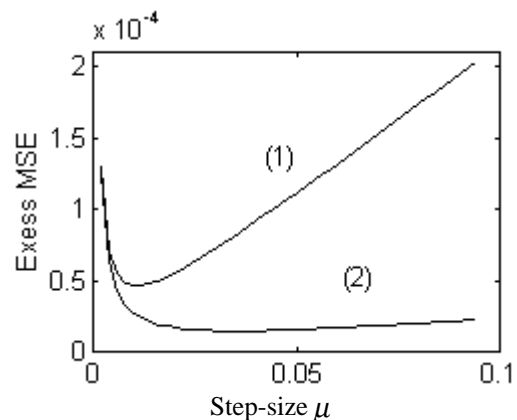

Fig.4: Theoretical total Excess MSE versus step-size for different σ_x^2 (SNR=10dB, $\sigma_{NSA}=\sigma_{NSB}$=5.$10^{-4}$) (1)$\sigma_x^2$=0.011 (2)$\sigma_x^2$=0.0017
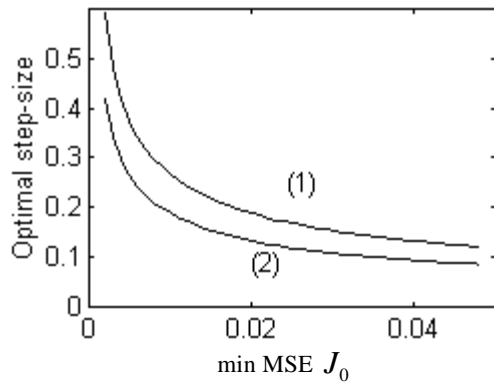
Fig.5: Optimal step-size versus min MSE J_0 for different σ^2_{NSA} and σ^2_{NSB} ($\sigma^2_x=0.0017$)
 (1)$\sigma^2_{NSA}= \sigma^2_{NSB}=2.5\ 10^{-6}$ (2)$\sigma^2_{NSA}=0$, $\sigma^2_{NSB}=2.5\ 10^{-6}$