Performance Analysis and FTF Version of the Generalized Sliding Window Recursive Least-Squares (GSWRLS) Algorithm

Karim Maouche and Dirk T. M. Slock
Institut EURECOM, 2229 route des Crêtes,
B.P. 193, 06904, Sophia Antipolis Cedex, FRANCE

Abstract

In this paper, we derive a new RLS algorithm: the Generalized Sliding Window RLS (GSW RLS) algorithm and its fast numerically stabilized version: the GSW SFTF algorithm. The generalized window used consists of an exponential decay with base λ for the first L lags, a decrease by a factor $1-\alpha$ at lag L, and a further exponential decay with base λ beyond lag L. The exponential and rectangular windows are special cases of the generalized window. We analyze the steady-state Excess Mean Squared Error components due to estimation noise and lag noise with different models for the time-varying optimal filter coefficients. This analysis shows that the exponential window performs better than the rectangular window, but also that the optimal generalized window performs even better.

1 Introduction

Tracking ability is a desired feature in adaptive filtering, especially when the system to be identified can vary quickly (relative to the duration of the FIR filter impulse response) as in the case of acoustic echo channels. The current state of the art in Recursive Least-Squares (RLS) adaptive filtering algorithms gives a choice of two possible windows: the exponential window (WRLS) and the rectangular window (SWC RLS). Though SWC RLS is twice as complex as WRLS, simulation experience leads one to believe that the rectangular window allows for better tracking of sudden changes. One element that helps explain this is that the SWC RLS algorithm has a strictly finite memory that forgets the past completely after a finite time. The SWC RLS algorithm solves recursively an overdetermined system of linear equations. More recently, the Affine Projection (AP) algorithm has been introduced. The AP algorithm can be seen to be related to the RLS family in that it provides recursively the minimum modification solution to an underdetermined set of linear equations. The criterion that the AP algorithm minimizes is related to the criterion of SWC RLS in the sense that the window length would be taken shorter than the FIR filter length. The AP algorithm has been found in simulations to show excellent tracking behavior, better than any of the existing RLS algorithms. This can be intuitively explained on the basis of the short window length: due to the short window length, there is only an averaging operation over a short time period and hence the lag is small. However, besides this window issue, there is also a convergence issue. Indeed, on the basis of the abover reasoning, the Normalized LMS (NLMS) algorithm, which corresponds to the AP algorithm with a window length equal to one, should have the fastest tracking capability. But we know that this is normally not true because the convergence of the NLMS algorithm is hampered by its dependence on the eigenvalue spread of the input covariance matrix. So the best tracking is obtained by the AP algorithm with some intermediate window length that depends on a number of parameters. In this paper, we propose the GSW RLS algorithm, a new RLS algorithm that generalizes the WRLS and SWC RLS algorithms. This algorithm uses a generalized window (see Fig.1) which consists of the superposition of an exponential window for the L most recent data and the same but attenuated exponential window for the rest of the data. The tracking improvement w.r.t. SWC RLS comes from the fact that the length L of the first part of the window can be smaller than the filter length. So it is possible to emphasize the L most recent data and hence to approach the situation of the AP algorithm. On the other hand, the GSW RLS algorithm solves an overdetermined system of equations and hence enjoys the fast convergence properties of RLS algorithms. In particular, in the noise-free perfect modeling case, perfect convergence will occur in a finite amount of time. The AP algorithm, though faster than the NLMS algorithm, will have to go through an infinite number of adaptations for perfect convergence to occur.

1058-6393/96 $5.00 © 1996 IEEE
Proceedings of ASILOMAR-29

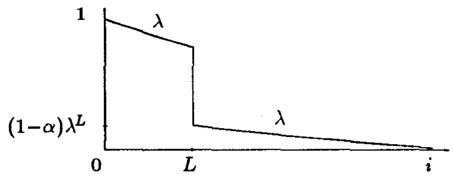

Figure 1: The generalized window.

Another effect of the exponential tail of the GSW is regularization. The finite window sample covariance matrices appearing in the SWC RLS and AP algorithms can be particularly ill-conditioned compared to a sample covariance matrix based on an exponential window with compatible time constant. So the exponential tail in the GSW RLS algorithm should reduce the noise enhancement phenomenon occurring in the AP algorithm. The GSW RLS algorithm turns out to have the same structure and computational complexity as the SWC RLS algorithm. The shift-invariance property inherent in adaptive filtering allows for the derivation of a fast version, the GSW FTF algorithm, that follows as a straightforward generalization of the SWC SFTF algorithm. We analyze the performance of the GSW algorithm for tracking time-varying optimal filter coefficients that vary according to different models that are the random walk, an AR(1) process and an MA(M) process. We compute the estimation noise and lag noise components of the steady-state excess MSE and prove for these parameter variations that the generalized window performs best. We also prove that the exponential window behaves better than the rectangular window which is an unexpected and presumably a new result.

2 The GSW RLS algorithm

An adaptive (FIR) filter $W_{N,k}$ combines linearly N consecutive input samples $\{x(i-n), n=0,\ldots,N-1\}$ to approximate (the negative of) the desired-response signal $d(i)$. The resulting error signal is given by

$$\epsilon_N(i|k) = d(i) + \sum_{n=0}^{N-1} W_{N,k}^{n+1} x(i-n) \ , \qquad (1)$$

where $X_N(i) = \left[x^H(i)\, x^H(i-1)\cdots x^H(i-N+1)\right]^H$ is the input data vector and superscript ^{H} denotes Hermitian (complex conjugate) transpose. In the WRLS algorithm, the set of N transversal filter coefficients $W_{N,k} = \left[W_{N,k}^1 \cdots W_{N,k}^N\right]$ are adapted so as to minimize recursively the LS criterion

$$\xi_N(k) \;=\; \sum_{i=1}^{k} \lambda^{k-i} \|\epsilon_N(i|k)\|^2 \ , \qquad (2)$$

where $\lambda \in (0,1]$ is the exponential weighting factor, $\|v\|_\Lambda^2 = v\Lambda v^H$, $\|.\| = \|.\|_I$. On the other hand, the SWC RLS algorithm minimizes recursively the following criterion:

$$\xi_{N,L}(k) = \sum_{i=k-L+1}^{k} \|\epsilon_N(i|k)\|^2 \ , \qquad (3)$$

where L, the length of the sliding window, must be greater than the filter length: $L > N$, in which case, the associated covariance matrix is invertible. Now, consider the criterion associated with the generalized window:

$$\xi_{N,L}(k) = \sum_{i=1}^{k} w_{k-i} \|\epsilon_N(i|k)\|^2 \ , \ w_i = \begin{cases} \lambda^i & 0 \le i \le L \\ (1-\alpha)\lambda^i & i > L \end{cases} . \qquad (4)$$

The new criterion generalizes the WRLS and SWC RLS criteria since the WRLS criterion (2) is obtained from (4) by setting $\alpha = 0$ and the SWC RLS criterion (4) is the one given by the generalized criterion when $\alpha = 1$ and $\lambda = 1$. Let $W_{N,L,k}$ be the adaptive RLS filter provided by such a window, the minimization of (4) leads to the following

$$W_{N,L,k} = -P_{N,L,k}^H \, R_{N,L,k}^{-1} \ , \qquad (5)$$

where $R_{N,L,k} = \sum_{i=1}^{k} w_i X_N(i) X_N^H(i)$ and $P_{N,L,k} = \sum_{i=1}^{k} w_i X_N(i) d^H(i)$. The use of the same forgetting factor for the two windows allows the following recursions for the sample second order moments

$$\begin{aligned} R_{N,L+1,k} &= \lambda R_{N,L,k-1} + X_N(k)X_N^H(k) \\ &= R_{N,L,k} - \alpha\lambda^L X_N(k-L)X_N^H(k-L) \end{aligned} \qquad (6)$$

$$\begin{aligned} P_{N,L+1,k} &= \lambda P_{N,L,k-1} + X_N(k)d^H(k) \\ &= P_{N,L,k} - \alpha\lambda^L X_N(k-L)d^H(k-L) \ . \end{aligned} \qquad (7)$$

Hence, we can derive the new algorithm by applying the strategy for the usual WRLS algorithm twice. The first step will be devoted to the time and order update $(k-1, L) \rightarrow (k, L+1)$, which is analogous to the update of the usual WRLS algorithm while the second step will be the order downdate $(k, L+1) \rightarrow (k, L)$. The downdate scheme is obtained as follows: By using (7), one has

$$-P_{N,L,k}^H = W_{N,L+1,k}R_{N,L+1,k} - \alpha\lambda^L d(k-L)X_N^H(k-L). \qquad (8)$$

Using (6) for $R_{N,L+1,k}$ in term of $R_{N,L,k}$, we get

$$W_{N,L,k} = W_{N,L+1,k} + \alpha\lambda^L \nu_{N,L+1}(k)\widetilde{D}_{N,L+1,k} \ , \qquad (9)$$

where $\nu_{N,L+1}(k) = d(k-L) + W_{N,L+1,k}X_N(k-L)$ and $\widetilde{D}_{N,L+1,k} = -X_N^H(k-L)R_{N,L,k}^{-1}$ are the a posteriori error signal and the a priori Kalman gain of the downdate part. Applying the MIL to (6) gives

$$R_{N,L,k}^{-1} = R_{N,L+1,k}^{-1} - D_{N,L+1,k}^H \delta_{N,L+1}^{-1}(k) D_{N,L+1,k} \ , \tag{10}$$

$D_{N,L+1,k} = -X_N^H(k-L)R_{N,L+1,k}^{-1}$ and $\delta_{N,L+1}(k) = \alpha^{-1}\lambda^{-L} - D_{N,L+1,k}X_N(k-L)$ are respectively the a posteriori Kalman gain and the likelihood variable associated with the downdate part. Now, it is straigthtforward to find that $\widetilde{D}_{N,L+1,k} = \alpha^{-1}\lambda^{-L}\delta_{N,L+1}^{-1}(k)D_{N,L+1,k}$ and that the a priori error is $\nu_{N,L+1}^p(k) = d(k-L) + W_{N,L,k}X_N(k-L) = \alpha^{-1}\lambda^{-L}\delta_{N,L+1}^{-1}(k)\nu_{N,L+1}(k)$. By associating the update part to the downdate part, we find the GSW RLS algorithm

$$\begin{aligned}
\widetilde{C}_{N,L,k} &= -\lambda^{-1}X_N^H(k)R_{N,L,k-1}^{-1} \\
\gamma_{N,L}^{-1}(k) &= 1 - \widetilde{C}_{N,L,k}X_N(k) \\
\epsilon_{N,L}^p(k) &= d(k) + W_{N,L,k-1}X_N(k) \\
\epsilon_{N,L}(k) &= \epsilon_{N,L}^p(k)\gamma_{N,L}(k) \\
W_{N,L+1,k} &= W_{N,L,k-1} + \widetilde{C}_{N,L,k}\epsilon_{N,L}(k) \\
R_{N,L+1,k}^{-1} &= \lambda^{-1}R_{N,L,k-1}^{-1} - \widetilde{C}_{N,L,k}^H\gamma_{N,L}(k)\widetilde{C}_{N,L,k} \\
D_{N,L+1,k} &= -X_N^H(k-L)R_{N,L+1,k}^{-1} \\
\delta_{N,L+1}(k) &= \alpha^{-1}\lambda^{-L} - D_{N,L+1,k}X_N(k-L) \\
\nu_{N,L+1}(k) &= d(k-L) + W_{N,L+1,k}X_N(k-L) \\
\nu_{N,L+1}^p(k) &= \alpha^{-1}\lambda^{-L}\delta_{N,L+1}^{-1}(k)\nu_{N,L+1}(k) \\
W_{N,L,k} &= W_{N,L+1,k} + \alpha\lambda^L D_{N,L+1,k}\nu_{N,L+1}^p(k) \\
R_{N,L,k}^{-1} &= R_{N,L+1,k}^{-1} - D_{N,L+1,k}^H\delta_{N,L+1}^{-1}(k)D_{N,L+1,k} \ .
\end{aligned} \tag{11}$$

The algorithm is initialized with $R_{N,L,0} = \mu I$ where μ is a small scalar quantity. The GSW RLS algorithm has the same structure as the SWC RLS algorithm. It shows a computational complexity of $O(N^2)$ operations.

3 The GSW SFTF algorithm

In [1], a fast numerically stabilized version of the SWC RLS algorithm is given. Using the same technique, it is straightforward to derive the fast numerically stabilized version of the GSW RLS algorithm that is the GSW SFTF algorithm

$$\begin{aligned}
&\left(A_{N,L,k}, \alpha_{N,L}^{-1}(k), \widetilde{C}_{N+1,L,k}, \gamma_{N+1,L}^{-s}(k)\right) = \\
&f_U\left(A_{N,L',k'}, \lambda^{-1}\alpha_{N,L'}^{-1}(k'), \left[0\,\widetilde{C}_{N,L,k'}\right], \gamma_{N,L}^{-s}(k), X_{N+1}(k)\right) \\
&\left(B_{N,L,k}, \beta_{N,L}(k), \left[\widetilde{C}_{N,L,k}\ 0\right], \gamma_{N,L}^{-s}(k)\right) = \\
&f_D\left(B_{N,L',k'}, \lambda\beta_{N,L'}(k'), \widetilde{C}_{N+1,L,k}, \gamma_{N+1,L}^{-s}(k), X_{N+1}(k)\right) \\
&\left(A_{N,L',k}, \alpha_{N,L'}^{-1}(k), \widehat{D}_{N+1,L,k}, --\widehat{\delta}_{N+1,L}^{s}(k)\right) = \\
&f_U\left(A_{N,L,k}, \alpha_{N,L}^{-1}(k), \left[0\,\widehat{D}_{N+1,L,k'}\right], -\widehat{\delta}_{N,L}^{s}(k), X_{N+1}(k_L)\right) \\
&\left(B_{N,L',k}, \beta_{N,L'}(k), \left[\widehat{D}_{N,L,k}\ 0\right], -\widehat{\delta}_{N,L}^{s}(k)\right) = \\
&f_D\left(B_{N,L,k}, \beta_{N,L}(k), \widehat{D}_{N+1,L,k}, -\widehat{\delta}_{N+1,L}^{s}(k), X_{N+1}(k_L)\right) \\
&k \text{ even:}\ \ \gamma_{N,L}(k) = \alpha\lambda^{N+L-1}\widehat{\delta}_{N,L}^{s}(k)\beta_{N,L'}(k)\alpha_{N,L'}^{-1}(k) \\
&\qquad\qquad \widehat{\delta}_{N,L}(k) = \widehat{\delta}_{N,L}^{s}(k) \\
&k \text{ odd:}\ \ \widehat{\delta}_{N,L}^{-1}(k) = \alpha\lambda^{N+L-1}\gamma_{N,L}^{-s}(k)\beta_{N,L'}(k)\alpha_{N,L'}^{-1}(k) \\
&\qquad\qquad \gamma_{N,L}^{-1}(k) = \gamma_{N,L}^{-s}(k) \\
&(W_{N,L,k}, \epsilon_{N,L}(k)) = f_J\left(W_{N,L',k'}, \widetilde{C}_{N,L,k}, \gamma_{N,L}(k), d(k), X_N(k)\right) \\
&\left(W_{N,L',k}, -\alpha\lambda^{L'}\nu_{N,L'}(k)\right) \\
&= f_J\left(W_{N,L,k}, \widehat{D}_{N,L,k}, --\widehat{\delta}_{N,L}^{-1}(k), d(k_L), X_N(k_L)\right) \ ,
\end{aligned} \tag{12}$$

where $k' = k-1$, $L' = L-1$, $k_L = k-L+1$, $\widehat{D}_{N,L,k} = \lambda^{1-L}D_{N,L,k}$ et $\widehat{\delta}_{N,L}(k) = \lambda^{-2(L-1)}\delta_{N,L}(k)$. $A_{N,L,k}$ and $B_{N,L,k}$ are the forward and backward prediction filters. $\alpha_{N,L}(k)$ and $\beta_{N,L}(k)$ are the corresponding prediction error energies. f_U, f_D and f_J are linear transformations and are given in [2]. The numerical complexity of the GSW SFTF algorithm is $16N$ operations, which is the complexity of the SWC SFTF algorithm.

4 Performance analysis

For the purpose of analysis, we consider the following classical identification model for the desired signal

$$d(k) = W_{N,k-1}^o X_N(k) + n(k) \ , \tag{13}$$

where $n(k)$ is a centered Gaussian i.i.d. sequence with variance σ_n^2 $\left(n(k) \sim \mathcal{N}(0, \sigma_n^2)\right)$ and $W_{N,k}^o$ is the unknown filter that is time-varying according to a certain model of variation. Considering a general window whose coefficients are denoted by w_i, one can show from (5) that the deviation filter $\widetilde{W}_{N,k} = W_{N,k} + W_{N,k}^o$ is given by

$$\widetilde{W}_{N,k} = \left(\sum_{i=0}^{\infty} w_i \left(\Delta_{k,i}X_{k-i} + n_{k-i}\right) X_{k-i}\right) R_k^{-1} \ , \tag{14}$$

where $\Delta_{k,i} = W_{N,k}^o - W_{N,k-i-1}^o$. Let COV_k be the covariance matrix of the deviation filter: $COV_k =$

$\mathrm{E}\widetilde{W}_{N,k}^H\widetilde{W}_{N,k}$. Now, assume that k is big enough so that: $R_{N,k} \approx \mathrm{E}R_{N,k} = (\sum_{i=0}^{\infty} w_i)R = \tau^{-1}R$. By using the fact that $\mathrm{E}\widetilde{W}_k^H\widetilde{W}_k = \mathrm{E}_X\mathrm{E}_{n,Z|X}\widetilde{W}_k^H\widetilde{W}_k$, one find easily that

$$COV_k = \tau R^{-1}(\sigma_n^2\sum_{i=0}^{\infty} w_i^2 R + \sum_{i,j=0}^{\infty} w_i w_j C_{i,j})\tau R^{-1} \ , \tag{15}$$

with $C_{i,j} = \mathrm{E}X_{k-i}X_{k-i}^H\Delta(i,j)X_{k-j}X_{k-j}^H$ and $\Delta(i,j) = \mathrm{E}\Delta_{k,i}^H\Delta_{k,j}$. In the case where X_k is a real Gaussian vector, one has

$$C_{i,j} = R\Delta(i,j)R + T_{i-j}\Delta(i,j)T_{i-j} + T_{i-j}\mathrm{tr}\,(T_{j-i}\Delta(i,j))\ , \tag{16}$$

where $T_{i-j} = \mathrm{E}X_{k-i}X_{k-j}^T$. Assuming statistical independence between X_{k-1} and $W_{N,k}$, one can express the variance of the a priori error signal as

$$\begin{aligned} \mathrm{var}(\epsilon_N^p(k)) &= \sigma_n^2 + \mathrm{tr}\,(R\,COV_{k-1}) \qquad (17)\\ &= \sigma_n^2 + N\sigma_n^2\sum_{j=0}^{\infty}\widetilde{w}_j^2 + \sum_{i,j=0}^{\infty}\widetilde{w}_i\widetilde{w}_j\mathrm{tr}\,(C_{i,j}R^{-1}) \ . \end{aligned}$$

with $\widetilde{w}_i = \tau w_i$. It clearly appears that there are two contributions in the MSE: the first one $\xi_n = (1 + N\sum_{i=0}^{\infty}\widetilde{w}_i)\sigma_n^2$ is the contribution due to estimation noise while the second one $\xi_W = \sum_{i,j=0}^{\infty}\widetilde{w}_i\widetilde{w}_j\mathrm{tr}\,(C_{i,j}R^{-1})$ is due to lag noise and originates from the variation of the echo path response. We can compute the EMSE due to estimation noise for the WRLS, SWC RLS and GSW RLS algorithms:

$$\begin{aligned} GSWRLS &: \ \sigma_n^2\left(1 + N\tfrac{1-\lambda}{1+\lambda}\tfrac{1+\alpha(\alpha-2)\lambda^{2L}}{(1-\alpha\lambda^L)^2}\right)\\ WRLS &: \ \sigma_n^2\left(1 + N\tfrac{1-\lambda}{1+\lambda}\right)\\ SWCRLS &: \ \sigma_n^2\left(1 + \tfrac{N}{L}\right) \ . \end{aligned} \tag{18}$$

The EMSE of WRLS can be recovered from the EMSE of the GSW RLS by letting $\alpha = 0$. The same thing holds for the SWC RLS algorithm for $\alpha = 1$ and $\lambda \to 1$. Now, if one considers the case where the input signal is a white noise with variance σ_x^2 and $\Delta(i,j)$ is a diagonal matrix, then

$$\xi_W = \sigma_x^2\sum_{i,j=0}^{\infty}\widetilde{w}_i\widetilde{w}_j(1 + (N+1)\delta_{i,j})\mathrm{tr}\,(\Delta(i,j)) \ . \tag{19}$$

The value of $\Delta(i,j)$ depends on the optimal filter variation. For instance, we begin the analysis by considering an AR(1) variation.

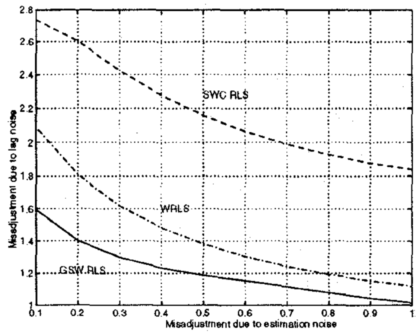

Figure 2: Misadjustment curves for an AR(1) variation ($N = 50$, $\sigma_x^2 = \sigma_z^2 = 0.01$, $\sigma_n^2 = 0.1$ and $\rho = 0.99$).

4.1 AR(1) variation of the optimal filter

When the optimal filter varies acording to an AR(1) process, one has

$$W_{N,k}^o = \rho W_{N,k-1}^o + Z_k \ , \ Z_k \sim \mathcal{N}(0, (1-\rho^2)\,\sigma_z^2 I) \ , \tag{20}$$

and $\Delta(i,j) = \frac{\sigma_z^2}{1-\rho^2}\left(1 + \rho^{|j-i|} - \rho^{i+1} - \rho^{j+1}\right)I$. It follows from (19) that the EMSE du to lag noise for the WRLS and SWCRLS algorithms are

$$\begin{aligned} WRLS &: \ \tfrac{N\sigma_x^2\sigma_z^2}{1-\rho^2}\left(1 + \tfrac{1-\lambda}{1+\lambda}\left(\tfrac{1-\rho(2+\lambda)}{1-\rho\lambda} + \tfrac{2(N+1)(1-\rho)}{1-\rho\lambda^2}\right)\right)\\ SWCRLS &: \ \tfrac{N\sigma_x^2\sigma_z^2}{L^2(1-\rho^2)}\Big(L(2N + L + 2) -\\ &\qquad 2\rho\tfrac{(1-\rho^L)(N+L+1)}{1-\rho} + \tfrac{L(1-\rho^2)-2\rho(1-\rho^L)}{(1-\rho)^2}\Big) \ . \end{aligned} \tag{21}$$

The EMSE due to lag noise for the generalized window is given when $\rho \neq \lambda$ by

$$\begin{aligned} \xi_W &= m\Big((2N+3)\tfrac{1+\alpha(\alpha-2)\lambda^{2L}}{1-\lambda^2} + \tfrac{1-\alpha\lambda^L}{1-\lambda}\\ &\left(\tfrac{1-\alpha\lambda^L}{1-\lambda} - 2\rho\tfrac{1-\alpha(\rho\lambda)^L}{1-\rho\lambda}\right) - 2(N+1)\rho\tfrac{\alpha(\alpha-2)(\rho\lambda^2)^L}{1-\rho\lambda^2} + 2\rho\lambda\\ &\left((1-\alpha)\lambda^{2(L-1)} + \tfrac{1-\lambda^{2(L-1)}}{1-\lambda^2} - \alpha(\rho\lambda)^{L-1}\tfrac{1-(\lambda/\rho)^{L-1}}{1-(\lambda/\rho)}\right)\Big) \ , \end{aligned} \tag{22}$$

with $m = \frac{N\sigma_x^2\sigma_z^2(1-\lambda)^2}{(1-\alpha\lambda^L)^2(1-\rho^2)}$. Fig.(2) shows curves that give lag noise misadjustment vs. estimation noise misadjustment for $R = \sigma_x^2 I, Q = \sigma_z^2 I, \sigma_x^2 = \sigma_z^2 = .01, N = 50$, and $\sigma_n^2 = .1$. The curve related to the GSW RLS algorithm is obtained by minimizing the EMSE due to lag noise w.r.t λ and L. The value of α being the solution of a second order equation which is obtained for a given value of the EMSE due estimation noise. These curves show that for the same value of the estimation noise misadjustment, the GSW RLS algorithm has a lower lag noise misadjustment. It shows also that the

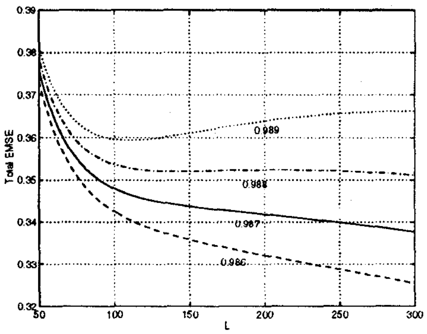

Figure 3: Total MSE of the SWC RLS algorithm for different values of ρ ($N = 50$, $\sigma_x^2 = \sigma_z^2 = 0.01$ and $\sigma_n^2 = 0.1$).

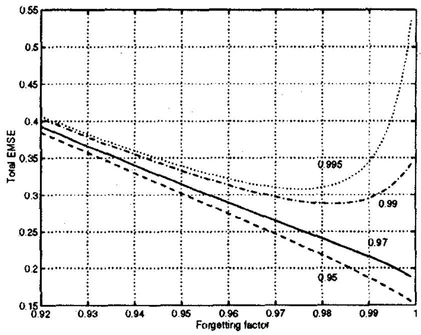

Figure 4: Total MSE of the WRLS algorithm for different valuse of ρ ($N = 50$, $\sigma_x^2 = \sigma_z^2 = 0.01$ and $\sigma_n^2 = 0.1$).

exponential window behaves better than the rectangular window. On Fig.(4) and (3), we give for different values of ρ, the total EMSE for the exponential and rectangular windows as a function of respectively λ and L. These curves show that for a relatively fast variation ($\rho < 0.98$), the EMSE is strictly decreasing with λ (resp. L). In this situation, the limit of the EMSE when $\lambda \rightarrow 1$ (resp. $L \rightarrow \infty$) is the variance of the desired signal : $\text{var}(d_k) = \sigma_n^2 + \frac{N\sigma_x^2\sigma_z^2}{1-\rho^2}$. For slower variations($\rho > 0.98$), one finds the classical behavior of adaptive filtering algorithms.

4.2 Random walk variation

In the case of a random walk variation, the optimal filter is given by

$$W_{N,k}^o = W_{N,k-1}^o + Z(k) \quad , \quad Z(k) \text{ i.i.d.} \sim \mathcal{N}(0, \sigma_z^2 I) \ . \tag{23}$$

and $\Delta(i,j) = \min(i+1, j+1)\sigma_z^2 I$. This gives for the EMSE due to lag noise (19)

$$\begin{aligned} GSWRLS \ &: \ m\Big((N+2)\tfrac{\alpha(2-\alpha)\lambda^{2L}\theta(\lambda)+1}{(1+\lambda)^2} - \tfrac{2\alpha\lambda^L P(\lambda)}{1-\lambda} \\ &+2(1-\alpha)L\lambda^{2L-1}(1-\lambda) + 2\tfrac{\lambda P(\lambda^2)-(1-\alpha)^2\lambda^{2L+1}\theta(\lambda)}{(1+\lambda)^2(1-\lambda)}\Big) \\ WRLS \ &: \ N\sigma_x^2\sigma_z^2\tfrac{2+N(1-\lambda)}{(1+\lambda)^2(1-\lambda)} \\ SWCRLS \ &: \ N\sigma_x^2\sigma_z^2\tfrac{(L+1)(3N+2L+4)}{L} \ , \end{aligned} \tag{24}$$

with $\theta(\lambda) = L\lambda^2 - L - 1, P(\lambda) = 1 - L\lambda^{L-1} + (L-1)\lambda^L$ and $m = \frac{N\sigma_x^2\sigma_z^2}{(1-\alpha\lambda^L)^2}$. The optimization of the generalized window shows that the GSW RLS algorithm has a better tracking ability than the WRLS and SWC RLS algorithms. We finf also in this case that the exponential window performs better than the rectangular window.

4.3 MA(M) variation

Here the optimal filter varies according to an MA model of order M: $W_k^o = \sum_{i=k-M+1}^{k} Z_i \quad , \quad Z_k \sim \mathcal{N}(0, \sigma_z^2 I)$, thus $\Delta(i,j) = \sigma_z^2(M - q_{j+1} - q_{i+1} + q_{|i-j|})I$, where $q_l = (M - |l|)$ for $|l| \geq M-1$ and 0 elsewhere. The EMSE due to lag noise for the WRLS and the SWC RLS are

$$\begin{aligned} WRLS: \quad & 2N\sigma_x^2\sigma_z^2\tfrac{1-\lambda^M}{1+\lambda}\left(\tfrac{1}{1-\lambda} + (N+1)\tfrac{1+\lambda^M}{1+\lambda}\right) \\ SWCRLS(L < M): \ & N\sigma_x^2\sigma_z^2\tfrac{1}{L}\left(\tfrac{2}{3}L^2 + (N+2)L + (N+4/3)\right) \\ (L \geq M): \ & N\sigma_x^2\sigma_z^2\tfrac{M}{L^2}\left(L^2 + (2N+3)L - \tfrac{M-1}{3}(3N+M+4)\right) \ , \end{aligned} \tag{25}$$

In [3], one can find the expression of the lag noise for the GSW RLS algorithm. Here again, it turns out that the generalized window tracks better and that the exponential window is better than the rectangular window. In [3], we give also the second order analysis in the case where the optimal filter varies according to an AR(1) process and the input signal is another AR(1) process. This analysis leads us to the same conclusions as before.

References

[1] D.T.M. Slock and T. Kailath "A modular prewindowing framework for covariance FTF RLS algorithms". *Signal Processing*, 28(1):47–61, 1992.

[2] D.T.M. Slock and T. Kailath "A modular multichannel multiexperiment fast transversal filter RLS algorithm". *Signal Processing*, 28(1):25–45, 1992.

[3] K. Maouche and Dirk T.M. Slock. "Analysis of Recursive Least-Squares FIR Algorithms in a non-stationary context". Technical Report RR N° 95-041, Institut Eurécom, Sophia Antipolis, France, October, 1995.