

WINDOW OPTIMIZATION ISSUES IN RECURSIVE LEAST-SQUARES ADAPTIVE FILTERING AND TRACKING

Tayeb Sadiki, Mahdi Triki, Dirk T.M. Slock*

Eurecom Institute
2229 route des Crêtes, B.P. 193, 06904 Sophia Antipolis Cedex, FRANCE
Email: {sadiki, triki, slock}@eurecom.fr

ABSTRACT

In this paper we consider tracking of an optimal filter modeled as a stationary vector process. We interpret the Recursive Least-Squares (RLS) adaptive filtering algorithm as a filtering operation on the optimal filter process and the instantaneous gradient noise (induced by the measurement noise). The filtering operation carried out by the RLS algorithm depends on the window used in the least-squares criterion. To arrive at a recursive LS algorithm requires that the window impulse response can be expressed recursively (output of an IIR filter). In practice, only two popular window choices exist (with each one tuning parameter): the exponential weighting (W-RLS) and the rectangular window (SWC-RLS). However, the rectangular window can be generalized at a small cost for the resulting RLS algorithm to a window with three parameters (GSW-RLS) instead of just one, encompassing both SWC- and W-RLS as special cases. Since the complexity of SWC-RLS essentially doubles with respect to W-RLS, it is generally believed that this increase in complexity allows for some improvement in tracking performance. We show that, with equal estimation noise, W-RLS generally outperforms SWC-RLS in causal tracking, with GSW-RLS still performing better, whereas for non-causal tracking SWC-RLS is by far the best (with GSW-RLS not being able to improve). When the window parameters are optimized for causal tracking MSE, GSW-RLS outperforms W-RLS which outperforms SWC-RLS. We also derive the optimal window shapes for causal and non-causal tracking of arbitrary variation spectra. It turns out that W-RLS is optimal for causal tracking of AR(1) parameter variations whereas SWC-RLS is optimal for non-causal tracking of integrated white jumping parameters, all optimal filter parameters having proportional variation spectra in both cases.

1. INTRODUCTION

The RLS algorithm is one of the basic tools for adaptive filtering. The convergence behavior of the RLS algorithm is now well understood. Typically, the RLS algorithm has a fast convergence rate, and is not sensitive to the eigenvalue spread of the correlation matrix of the input signal. However, when operating in a non-stationary environment, the adaptive filter has the additional task of tracking the variation in environmental conditions. In this context, it has been established that adaptive algorithms that exhibit good

convergence properties in stationary environments do not necessarily provide good tracking performance in a non-stationary environment; because the convergence behavior of an adaptive filter is a transient phenomenon, whereas the tracking behavior is a steady-state property [1, 2].

One fundamental non-stationary scenario involves a time-varying system in which the cross-correlation between the input signal and the desired response is time-varying. This case occurs in the system identification setup. To take into account system variation, two main variants of RLS algorithms exist. The first introduces a forgetting factor, and leads to the exponentially Weighted RLS (W-RLS) approach. The second uses a Sliding rectangular Window (SWC-RLS approach). In [3, 4], a generalized sliding window RLS (GSW-RLS) algorithm was introduced, that generalizes the W-RLS and SWC-RLS algorithms. The GSW-RLS uses a generalized window (see Fig. 1), which consists of an exponential window with a discontinuity at delay L . It can be seen that the exponential and rectangular windows are particular cases of the generalized window, for $\alpha = 0$ and $(\alpha, \lambda) = (1, 1)$ resp. In [3, 4],

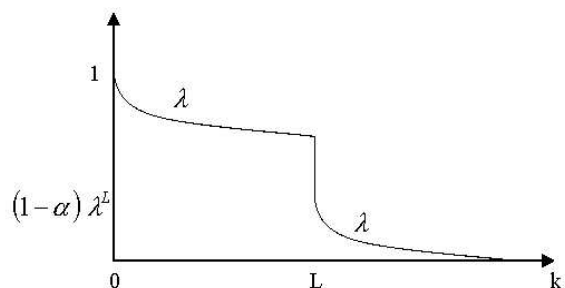


Fig. 1. The generalized sliding window

a tracking improvement for GSW-RLS was observed for different system variation models (AR(1), MA, and Random walk). On the one hand the initial portion of the window permits to emphasize the very recent past which allows very fast tracking. On the other hand, the GSW-RLS algorithm solves nevertheless an overdetermined system of equations and hence enjoys the fast convergence properties of RLS algorithms. Another effect of the exponential tail of the GSW is regularization. In fact, the rectangular window sample covariance matrix appearing in SWC-RLS can be particularly ill-conditioned compared to a sample covariance matrix based on an exponential window with compatible time constant. Finally, the GSW-RLS algorithm turns out to have the same structure and comparable computational complexity as the SWC-RLS algorithm.

*Eurecom's research is partially supported by its industrial partners: Bouygues Télécom, Fondation d'Entreprise Groupe Cegetel, Fondation Hasler, France Télécom, Hitachi, Sharp, ST Microelectronics, Swisscom, Texas Instruments, Thales.

This paper is organized as follows. In section 2, a tracking analysis in the frequency domain is presented. Uninformed and Informed Bayesian approaches are investigated respectively in sections 3, and 4. Finally a discussion and concluding remarks are provided in section 5.

2. TRACKING CHARACTERISTICS OF RLS ALGORITHMS

We consider the classic adaptive system identification problem (see Fig. 2). The adaptive system identification is designed for determining a (typically linear FIR) model of the transfer function for an unknown, time-varying digital or analog system.

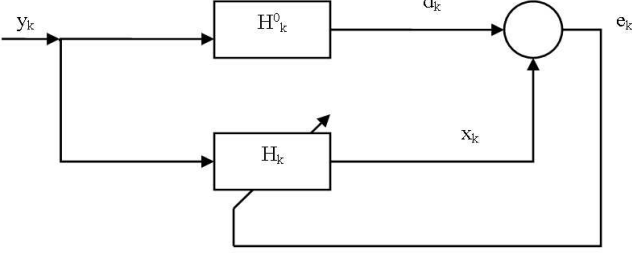


Fig. 2. System identification block diagram

The adaptive system identification problem can be described by:

$$\begin{cases} d_k = H_k^o Y_k + n_k \\ x_k = H_k^T Y_k \end{cases} \quad (1)$$

where

- n_k is an iid Gaussian noise sequence ($n_k \sim \mathcal{N}(0, \sigma_n^2)$) where σ_n^2 is the Minimum Mean Squared Error (MMSE)
- H_k^o denotes the optimal Wiener Filter
- H_k represents the adaptive Filter
- e_k is the a posteriori error given by:

$$e_k = d_k - x_k = \tilde{H}_k^T Y_k + n_k \quad (2)$$

where $\tilde{H}_k = H_k^o - H_k$ denotes the filter deviation.

In weighted RLS, the set of the N adaptive filter coefficients $H_k = [H_{1,k} \cdots H_{N,k}]^T$ gets adapted so as to minimize recursively the Weighted Least Squares criterion

$$J_k = F(q) e_k^2 = \sum_i f_i e_{k-i}^2 \quad (3)$$

where $F(z) = \sum_i f_i z^{-i}$ is the transfer function of the weighting window f_i characterizing the RLS algorithm, and $q^{-1} e_k^2 = e_{k-1}^2$. There are a number of references dealing with the performance of RLS algorithms in non-stationary environments [7, 6, 5, 4]. The basic idea is to focus on the model quality in terms of the output Excess MSE (EMSE). We consider stationary optimal filter variation models, hence the RLS algorithm will reach a stationary regime to which we limit attention. The EMSE is defined as:

$$EMSE = E \{ e_k^2 \} - \sigma_n^2 = E \left\{ Y_k^T \tilde{H}_k \tilde{H}_k^T Y_k \right\} \quad (4)$$

(in principle the a priori error signal should be considered for the EMSE, we shall stick to the a posteriori error signal to avoid the appearance of a delay in the notation). So, if we assume that the

system variation is a zero-mean, wide-sense stationary process H_k^o with a power spectral density matrix $S_{HH}(e^{j2\pi f})$, and if we invoke the independence assumption, in which Y_k and \tilde{H}_k are assumed to be independent (this works better for the a priori error), the EMSE can be expressed in the following form:

$$EMSE = \text{tr} \left\{ E \left[\tilde{H}_k^* \tilde{H}_k^T \right] R \right\} = \text{tr} \left\{ R \int_{-\frac{1}{2}}^{\frac{1}{2}} S_{\tilde{H}\tilde{H}}(e^{j2\pi f}) df \right\}.$$

By setting the gradient of J_k in (3) w.r.t. H_k to zero, we have

$$\begin{aligned} \left(F(q) Y_k Y_k^T \right) H_k &= F(q) Y_k d_k \\ &= F(q) Y_k Y_k^T H_k^o + F(q) Y_k n_k. \end{aligned}$$

Let's denote by $\tilde{F}(q) = \frac{F(q)}{F(1)}$ the (dc transfer) normalized weighting window. As this window is generally low-pass, $\tilde{F}(q)$ acts as an averaging operator, and we have

$$\tilde{F}(q) Y_k Y_k^T \approx R.$$

On the other hand, as the optimal system variation is independent of the input signal (in the system id setup), we approximate:

$$\begin{aligned} \tilde{F}(q) Y_k Y_k^T H_k^o &\approx \left(\tilde{F}(q) Y_k Y_k^T \right) \left(\tilde{F}(q) H_k^o \right) \\ &\approx R \tilde{F}(q) H_k^o. \end{aligned}$$

Hence the filter deviation can be expressed as:

$$\tilde{H}_k = H_k^o - H_k = \left(1 - \tilde{F}(q) \right) H_k^o - R^{-1} \tilde{F}(q) Y_k n_k$$

and the EMSE becomes:

$$\begin{aligned} EMSE &= N \sigma_n^2 \int_{-\frac{1}{2}}^{\frac{1}{2}} \left| \tilde{F}(e^{j2\pi f}) \right|^2 df \\ &\quad + \int_{-\frac{1}{2}}^{\frac{1}{2}} \left| 1 - \tilde{F}(e^{j2\pi f}) \right|^2 \text{tr} \left\{ R S_{HH}(e^{j2\pi f}) \right\} df \end{aligned} \quad (5)$$

Remark that the EMSE can be broken up into two terms:

- $E_{est} = N \sigma_n^2 \int_{-\frac{1}{2}}^{\frac{1}{2}} \left| \tilde{F}(e^{j2\pi f}) \right|^2 df$ corresponding to the estimation noise contribution; it can be interpreted as the estimation accuracy in time-invariant conditions,
- $E_{lag} = \int_{-\frac{1}{2}}^{\frac{1}{2}} \left| 1 - \tilde{F}(e^{j2\pi f}) \right|^2 \text{tr} \left\{ R S_{HH}(e^{j2\pi f}) \right\} df$ representing the estimation error resulting from low-pass filtering the system variations (lag noise, since in the causal window case this means lagging behind).

The estimation and lag noise terms can also be interpreted as the variance and the bias of the conditional estimation problem, for a given value of the optimal filter sequence. In fact,

$$\tilde{H}_k = \underbrace{\left(1 - \tilde{F}(q) \right)}_b H_k^o - R^{-1} \tilde{F}(q) Y_k n_k$$

where $b = E_{|H^o} \tilde{H}_k$ is the estimation bias. We get

$$R_{\tilde{H}\tilde{H}} = E_{|H^o} \left(\tilde{H}_k \tilde{H}_k^T \right) = \underbrace{b b^T}_{\text{bias}} + \underbrace{\sigma_n^2 \left(\sum_i \tilde{f}_i^2 \right)}_{\text{variance}} R^{-1}. \quad (6)$$

Then we see that:

- $E_{est} = N \sigma_n^2 \left(\sum_i \tilde{f}_i^2 \right)$ is the variance component,
- $E_{lag} = tr \{ R E_{H^o} [b b^T] \}$ is the bias component.

3. UNINFORMED APPROACH FOR RLS TRACKING ANALYSIS

In an uninformed approach we assume that little or no information about the system variations is available for the design of the RLS algorithm.

3.1. Uninformed Tracking Analysis of Causal RLS Algorithms

From (5), we can see that the following window characteristics characterize estimation and lag noises resp.:

- $l_\infty = \left(\sum_i \tilde{f}_i^2 \right) = \left\| \tilde{F} \right\|_2^2$ characterizing E_{est} ,
- $E_F(f) = \left| 1 - \tilde{F} \left(e^{j2\pi f} \right) \right|$, called *parameter tracking characteristic*, characterizing the lag noise.

To compare the tracking ability of RLS with different weighting windows, we shall choose the windows parameters such that the different algorithms behave identically under time-invariant conditions. In fact, comparing adaptive filters characterized by different values of l_∞ barely makes any sense (in the uninformed case) and it resembles "comparing runners that specialize in different distances" [8]. By normalizing the performance under time-invariant conditions, the tracking characteristic $E_F(w)$ depends only on the window shape.

Since the complexity of RLS with a rectangular window essentially doubles with respect to an exponential window, it is generally believed that this increase in complexity allows for some improvement in tracking performance. In contrast to this intuition, Fig. 3 shows that the plot of the normalized tracking characteristic of W-RLS lies below that corresponding to SWC-RLS. Thus, the tracking capability of W-RLS approach is better. This effect can be attributed to a higher degree of concentration of the exponential window around $i = 0$, which results in a smaller estimation delay, hence smaller bias error [8].

As we have mentioned in the Introduction, the SW-RLS approach can be generalized at a small cost for the resulting RLS algorithm to a window with three parameters (instead of just one). Compared to the Sliding and the Exponential windows, the Generalized Sliding Window introduces two extra degrees of freedom. The shape of the window depends on the choice of these degrees of freedom. Thus, they can be optimized to minimize the average parameter tracking characteristic. In other words, the window parameters are chosen so as to

$$\left\{ \begin{array}{l} \min_{\lambda, \alpha, L} \int_0^{f_0} \left| 1 - \tilde{F} \left(e^{j2\pi f} \right) \right|^2 df \\ \text{subject to } \sum_k \tilde{f}_k^2 = l_\infty \end{array} \right. \quad (7)$$

where f_0 is the assumed bandwidth of the system variations. In Fig. 3, we add the tracking characteristic of the GSW-RLS estimator, as a function of $f = f_0$. As expected, the optimized GSW-RLS outperforms the SWC-RLS and W-RLS approaches.

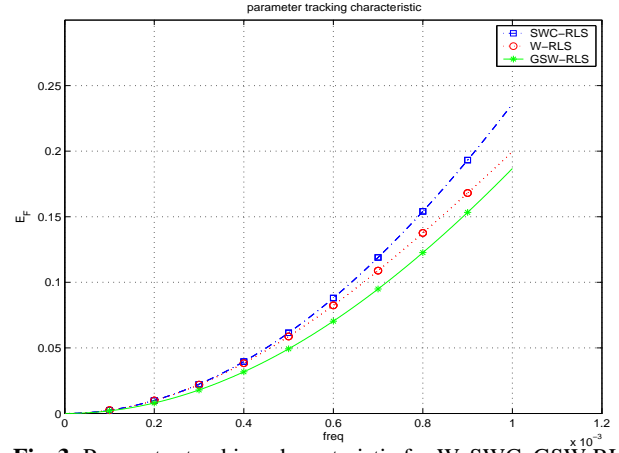


Fig. 3. Parameter tracking characteristic for W, SWC, GSW-RLS.

3.2. Uninformed Tracking Analysis for Non-Causal RLS Algorithms

Bias in the RLS algorithm is caused by two kinds of distortion: amplitude and phase distortions. Phase distortion can be considerably attenuated by introducing a suitable estimation delay (using non-causal filtering). With no information about the system characteristics, a suitable estimation delay can be determined as [8]:

$$\tau_e = \sum_k k \tilde{f}_k, \quad (8)$$

the mean of the \tilde{f}_k considered as a distribution. As before, under identical estimation noise, comparing the tracking capability of the non-causal RLS algorithms can be investigated by comparing what is called in [8] the *parameter matching characteristic* defined as:

$$\tilde{E}_F(w) = \left| e^{-j2\pi f \tau_e} - \tilde{F} \left(e^{j2\pi f} \right) \right| = \left| 1 - e^{j2\pi f \tau_e} \tilde{F} \left(e^{j2\pi f} \right) \right|.$$

Fig. 4 shows plots of normalized matching characteristics of SWC-RLS, W-RLS, and GSW-RLS (with optimized window parameters) algorithms. The curves show that in the non-causal adaptation case, the optimal shape for the generalized window becomes the rectangular one (and in particular, rectangular windowing outperforms exponential windowing). The better parameter matching properties of the SWC-RLS approach can be explained by the linearity of the associated phase characteristic (due to the window symmetry). The delay τ_e becomes the center of the window and after delay compensation, there is zero phase distortion left.

4. INFORMED APPROACH FOR RLS TRACKING OPTIMIZATION

Now we suppose the statistics of the system variation to be available. In this case, we can achieve an optimal tradeoff between the estimation and lag noises. The optimal tradeoff can be found by minimizing the EMSE:

$$\min_F \left(N \sigma_n^2 \int_{-\frac{1}{2}}^{\frac{1}{2}} df \left| F \left(e^{j2\pi f} \right) \right|^2 + \int_{-\frac{1}{2}}^{\frac{1}{2}} df \left| 1 - F \left(e^{j2\pi f} \right) \right|^2 tr \{ R S_{HH}(f) \} \right) \quad (9)$$

subject to $F(1) = 1$

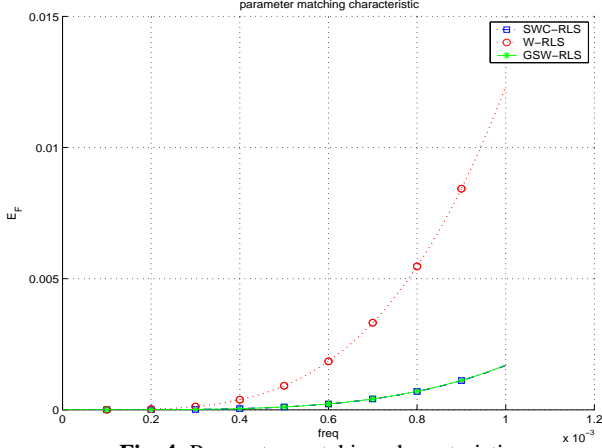


Fig. 4. Parameter matching characteristic.

4.1. Optimized Causal Parametric Windows, Separable Variation Spectrum Case

In a first instance, we propose to investigate a simple (but interesting) variation model. We assume that the impulse response coefficients have proportional Doppler spectrum; i.e.

$$S_{HH}(e^{j2\pi f}) = D S_{hh}(e^{j2\pi f}). \quad (10)$$

To simplify, we suppose also that the scalar spectrum S_{hh} is a flat low-pass spectrum; i.e.

$$S_{hh}(e^{j2\pi f}) = \begin{cases} 1 & |f| < f_0 \\ 0 & \text{elsewhere} \end{cases}$$

The matrix D is arbitrary but if it were diagonal (decorrelated filter coefficients) the diagonal would represent the power delay profile of the optimal filter (in wireless channel terminology). With the separable model, the Excess MSE (for the causal adaptation case) can be expressed as:

$$N\sigma_n^2 \int_{-\frac{1}{2}}^{\frac{1}{2}} df |F(e^{j2\pi f})|^2 + \text{tr}\{RD\} \int_{-\frac{1}{2}}^{\frac{1}{2}} df |1 - F(e^{j2\pi f})|^2.$$

The EMSE expressions for the different RLS variants become, as a function of the windows parameters:

$$\begin{aligned} EMSE^{SWCRLS} &= \frac{N\sigma_n^2}{L} \\ &+ 2 \text{tr}\{RD\} \left(\frac{L-1}{L} f_0 - \frac{1}{\pi L^2} \sum_{k=1}^{L-1} \sin(2\pi f_0 k) \right) \\ EMSE^{WRLS} &= N\sigma_n^2 \frac{1-\lambda}{1+\lambda} \\ &+ 2 \text{tr}\{RD\} \left(\lambda f - \frac{\lambda}{\pi} \frac{1-\lambda}{1+\lambda} \arctan\left(\frac{1+\lambda}{1-\lambda} \tan(\pi f)\right) \right) \\ EMSE^{GSWRLS} &= N\sigma_n^2 \frac{1-\lambda}{1+\lambda} \frac{1 + \alpha(\alpha-2)L^{2L}}{(1-\alpha\lambda^L)^2} \\ &+ 2 \text{tr}\{RD\} \left(f_0 \gamma_0 + \sum_{k=1}^{\infty} \frac{\gamma_k}{\pi k} \sin(2k\pi f_0) \right) \end{aligned}$$

where γ_k gets computed recursively as:

$$\begin{cases} \gamma_{k+1} = \lambda\gamma_k - \alpha\kappa^2 \lambda^{2L-1-k} & , k < L, \\ \gamma_{k+1} = \lambda\gamma_k & , k \geq L. \end{cases}$$

To investigate the tracking ability of the different RLS algorithms, we compare the minimum EMSE achieved by each variant (with optimized parameters). Fig. 5 plots the curves of minimized EMSE (as a function of the bandwidth f_0).

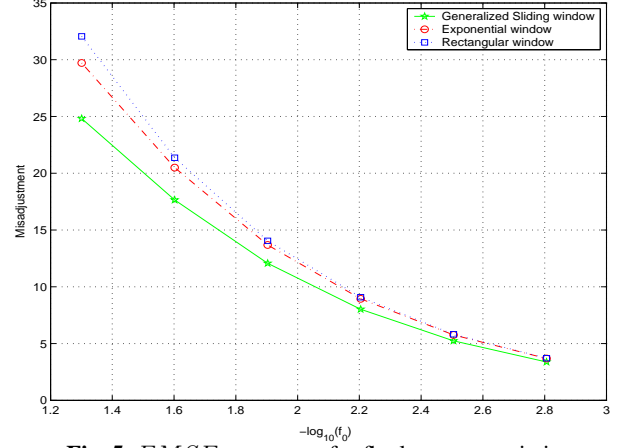


Fig. 5. $EMSE_{min}$ curves for flat low-pass variations.

This analysis shows that, for a flat low-pass spectrum, the exponential window performs better than the rectangular window, but also that the optimal generalized window performs even better.

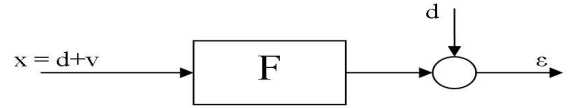


Fig. 6. Signal in noise problem.

4.2. Optimized Windows

Consider minimizing the Excess MSE with respect to the window coefficients. This problem can be interpreted in terms of Wiener filtering for a signal in noise problem, see Fig. 6, where the desired and noise signal spectra are respectively $S_{vv}(e^{j2\pi f}) = \sigma_v^2 = N\sigma_n^2$, and $S_{dd}(e^{j2\pi f}) = \text{tr}\{R S_{HH}(e^{j2\pi f})\}$. The causal Wiener solution for such problem is

$$F_{Wiener}(e^{j2\pi f}) = \frac{S_{dd}(e^{j2\pi f})}{S_{dd}(e^{j2\pi f}) + \sigma_v^2}. \quad (11)$$

The DC component of this Wiener solution is

$$F_{Wiener}(1) = \frac{S_{dd}(1)}{S_{dd}(1) + \sigma_v^2}. \quad (12)$$

Now, since S_{dd} is quite lowpass, we have $S_{dd}(1) \gg \sigma_v^2$. Thus, for an acceptable SNR, $F_{Wiener}(1) \approx 1$. Then,

$$F_{opt}(e^{j2\pi f}) \approx F_{Wiener}(e^{j2\pi f}) = \frac{S_{dd}(e^{j2\pi f})}{S_{dd}(e^{j2\pi f}) + \sigma_v^2}. \quad (13)$$

The associated Excess MSE is given by:

$$EMSE_{min} = \sigma_v^2 \int F_{opt} (e^{j2\pi f}) df = \sigma_v^2 f_0^{opt}. \quad (14)$$

If we impose a causality constraint, the Wiener solution becomes:

$$F_{Wiener}^c (e^{j2\pi f}) = 1 - \frac{\sigma_v^2}{\sigma^2 A (e^{j2\pi f})} \quad (15)$$

where $A(z)$ denotes the optimal prediction error filter for the signal $x = d + v$, and σ^2 the associate prediction error variance. Using the same arguments as before, one can show that, for an acceptable SNR, $F_{Wiener}^c(1) \approx 1$; and then $F_{opt}^c(e^{j2\pi f}) \approx F_{Wiener}^c(e^{j2\pi f})$. The associated Excess MSE is given by:

$$EMSE_{min}^c = \sigma_v^2 \int F_{opt}^c (e^{j2\pi f}) df = \sigma_v^2 f_0^{c,opt}. \quad (16)$$

4.3. Optimality Considerations for Classical Windows

The question we investigate in this section is "For which optimal filter variation model is the exponential or the rectangular window optimal?". To answer this question, we use the reverse engineering technique. We assume a separable optimal filter variation spectrum with uncorrelated optimal filter coefficients.

Equating the causal Wiener solution and the exponential window transfer function leads to a variation spectrum of the form:

$$S_{hh}(z) \propto \frac{1}{(1 - \alpha z^{-1})(1 - \alpha z)}. \quad (17)$$

Thus, for AR(1) "drifting" parameters, the exponential window optimizes the tracking performance (over the set of all causal windows).

Similarly, by equating the non-causal Wiener filter with the centered rectangular window, one can show that SWC-RLS is optimal for a variation spectrum of the form:

$$S_{hh}(z) \propto \frac{F_R(z)}{1 - F_R(z)} \quad (18)$$

where $F_R(z) = \frac{1}{L} \frac{z^{\frac{L}{2}} - z^{-\frac{L}{2}}}{z^{\frac{L}{2}} - z^{-\frac{L}{2}}}$ denotes the centered sliding window transfer function. $S_{hh}(z)$ in (18) can be interpreted to be the spectrum of a process g with spectrum $S_{gg}(z) = F_R(z)$, low-pass filtered by $H(z) = \left(\frac{1}{1 - F_R(z)}\right)^{1/2}$ (spectral factor). The input g can be interpreted as a white jumping process (with memory $L/2$), or as white noise viewed on the scale $L/2$: g stays constant for $L/2$ samples, then jumps to an uncorrelated value, stays constant again for $L/2$ samples and so on. On the other hand, around frequency zero

$$\frac{1}{1 - F_R(e^{j2\pi f})} \approx \frac{3}{\pi^2(L^2 - 1)} \frac{1}{f^2}. \quad (19)$$

Then, the low-pass filter $H(z)$ can be interpreted as an integrator.

5. CONCLUDING REMARKS

In this paper, we have considered tracking of a time-varying system modeled as a stationary vector process. In the Uninformed approach, we have investigated the tracking capability, by comparing the tracking and matching characteristics of the different RLS windows (under identical estimation noise). In the Informed approach, we have interpreted the RLS algorithm as a filtering operation on the optimal filter process and the instantaneous gradient noise. We have shown the optimality of the exponential window for AR(1) "drifting" parameters, and of the sliding window for integrated white "jumping" parameters. An open question remains: how to estimate and optimize simultaneously the adaptive filter and window parameters? Alternatively, one may opt for a two-step approach:

- Step 1: non-causal SWC-RLS (with short window), providing noisy but undistorted filter estimates.
 - Step 2: Wiener (Kalman) filtering to provide the optimal estimation noise/low-pass distortion compromise, e.g. as in [9].
- Such an approach allows for less constrained optimal filtering, that can be optimized, and tailored to individual (and possibly correlated) filter coefficients, whereas RLS has only one global window.

6. REFERENCES

- [1] O.M. Machhi and N.J. Bershad. "Adaptive Recovery of a Chirped Sinusoid in Noise: I. Performance of the RLS Algorithm," *IEEE Trans. Acoustics, Speech, and Signal Processing*, Vol.39, pp. 583–594, March 1991.
- [2] S. Haykin, A.H. Sayed, J.R. Zeidler, P. Yee, P.C. Wei. "Adaptive Tracking of Linear Time-Variant Systems by Extended RLS Algorithms," *IEEE Trans. Signal Processing*, Vol.45, pp. 1118–1128, May 1997.
- [3] K. Maouche and D.T.M. Slock. "Performance Analysis and FTF Version of the Generalized Sliding Window Recursive Least-Squares (GSWRLS) Algorithm," *In Proc. Asilomar Conference on Signals, Systems and Computers*, Vol.1, pp. 685–689, Pacific Grove, USA, Nov. 1995.
- [4] K. Maouche. "Algorithmes des Moindres Carrés Récursifs Doublement Rapides : Application à l'Identification de Réponses Impulsionnelles Longues," *PhD Thesis*, Ecole Nationale Supérieure des Télécommunications, Paris, March 1996.
- [5] M. Montazeri. "Une Famille d'Algorithmes Adaptatifs Comprenant les Algorithmes NLMS et RLS. Application à l'Annulation d'Echo Acoustique," *PhD Thesis*, Université de Paris-Sud, U.F.R Scientifique d'ORSAY, Sept. 1994.
- [6] E. Eleftheriou and D. Falconer. "Tracking Properties and Steady State Performance of RLS Adaptive Filter Algorithms," *IEEE Trans. Acoustics, Speech and Signal Processing*, Vol.34, No.5, pp. 1097–1110, Oct. 1986.
- [7] M. Niedzwiecki. "First-Order Tracking Properties of Weighted Least Squares Estimators," *IEEE Trans. Automatic Control*, Vol.33, No.1, pp. 94–96, Jan. 1988.
- [8] M. Niedzwiecki. "Identification of Time-Varying Processes," *Wiley & Sons*, 2000.
- [9] M. Lenardi and D. Slock. "Estimation of Time-Varying Wireless Channels and Application to the UMTS W-CDMA FDD Downlink," *In Proc. European Wireless (EW)*, Florence, Italy, Feb. 2002.