On the Sensitivity of Transversal RLS Algorithms to Random Perturbations in the Filter Coefficients

Sasan Ardalan

Center for Communications and Signal Processing
Dept. of Electrical and Computer Engineering
North Carolina State University

CCSP-TR-88/5

January 1988

Abstract

Transversal Recursive Least Squares (RLS) algorithms estimate filter coefficients which minimize the accumulated sum of the square of the error residuals termed the error power. In this paper the sensitivity of this error power to random perturbations about the optimum filter coefficients is investigated. Expressions are derived for the mean and variance of the deviation from the optimum error power. It is shown that for the prewindowed growing memory RLS algorithm ($\lambda = 1$) the mean value of the deviation increases linearly with the number of iterations. The variance of the deviation increases in proportion to the square of the number of iterations. Expressions are also derived for the variance for correlated signals. These expressions show that the variance of the deviation for correlated signals increases compared to uncorrelated white signals by a term related to the sum of the square of the off-diagonal elements of the sample autocorrelation matrix. Expressions are also derived for the mean and variance of the deviation for the exponentially windowed RLS algorithm ($\lambda < 1$). In this case the deviations are bounded and inversly proportional to $1-\lambda$.

I. Introduction

The transversal Recursive Least Squares algorithm [1,2,3] has received considerable attention as a high performance alternative to the Least Mean Square (LMS) based adaptive filter [4]. However, a major impedement to the implementation and realization of systems based on this algorithm has been the unstable behaviour of the efficient fast transversal RLS algorithms [1,2,3]. Recent efforts have been directed at predicting the effects of finite precision roundoff errors in RLS algorithms. In [5] the error propagation of the RLS algorithm including the fast Kalman algorithm [6] was studied. It was shown that by using a forgetting factor $\lambda < 1$ the RLS algorithm is stablized. However, the fast Kalman algorithm (and its variants) was shown to be inherently unstable for $\lambda < 1$. This is also demonstrated in [12].

A number of efforts have been made towards stablizing the fast RLS algorithms, for example, [7,8]. Nevertheless, analysis is always needed to determine the deviation from the infinite precision performance. In particular, this deviation must be related to the register sizes used to implement the algorithm. In the course of such analysis some insight can be gained into the divergence mechanism in these algorithms.

In [9] and [10] expressions were derived for the mean square roundoff error. However, the variance of the random fluctuations about the mean square roundoff error was not determined. Furthermore, the analysis were carried out for white input signals only.

In this paper the sensitivity of the RLS algorithm is studied by deriving expressions for both the mean deviation from the optimum and the variance about the mean due to random perturbations in the optimum filter coefficients of RLS algorithms. Results are derived for both $\lambda=1$ and $\lambda<1$ and for white and correlated signal.

The method is based on expanding the expression for the accumulated sum of the squared error, which is minimized by the RLS algorithm, and which is a function of the transversal filter coefficients, about the optimum filter in a Taylor series using second order terms. The deviation due to random perturbations in the filter coefficients is found from this expansion and is itself a random quantity. The method was inspired by the work of [11].

The results show that the mean deviation increases linearly with time for the prewindowed growing memory ($\lambda = 1$) RLS algorithm. The variance of this deviation increases with the square of the number of iterations. In other words, the sensitivity of the RLS algorithm to random perturbations in the filter coefficients increases with time. Hence, the algorithm may diverge in a finite precision implementation. On the other hand, the mean deviation for the exponentially windowed RLS algorithm ($\lambda<1$) is bounded and decreases as λ is decreased: it is inversely proportional to $1-\lambda$. It is shown that the sensitivity of the RLS algorithm increases for correlated signals such that the variance of the deviation increases by a term related to the sum of the square of the off-diagonal elements of the sample autocorrelation matrix. This term is of course zero for white signals.

The RLS Problem

Consider the adaptive filtering problem of estimating the desired response $d(n)$ by forming a linear combination of the current and previous input samples $x(n)$. Define the vector $\mathbf{d}_n$ which collects all samples of $d(n)$ from time 0 to n in a single $M \times 1$ column vector ($M \gg n$)

$$\mathbf{d}_n = [\mathrm{d}(\mathrm{n}), \mathrm{d}(\mathrm{n}-1), \ldots, \mathrm{d}(0), \; 0, \ldots, 0]^{\mathrm{T}} \tag{1.1}$$

Also, define the input vector,

$$\mathbf{x}_n = [x(n), x(n-1), \ldots, x(0), 0, \ldots, 0]^T \tag{1.2}$$

Now, define the $M \times N$ matrix, $X_{N,n}$

$$X_{N,n} = [\mathbf{x}_n \; \mathbf{x}_{n-1} \cdots \mathbf{x}_{n-N+1}] \tag{1.3}$$

where

$$\mathbf{x}_{n-1} = q^{-1} \mathbf{x}_n = [x(n-1) \; x(n-2), \ldots, x(0), 0 \ldots 0]^T \tag{1.4}$$

and q^{-1} is unit delay operator.

Now consider the weight vector, $\mathbf{w}_{\mathrm{N,n}}$, which forms a linear combination of the current and past input samples to produce an estimation $\hat{d}(n)$ of $d(n)$. Then the error vector is

$$\epsilon_n = \mathbf{d}_n - X_{\mathrm{N,n}} \; \mathbf{w}_{\mathrm{N,n}} \tag{1.5}$$

The Euclidean length of the vector ϵ_n is define by

$$\xi_N(n) = \epsilon_n^T \; \epsilon_n . \tag{1.6}$$

which is also termed the error power. It is well known that the weight vector which minimizes the Euclidean length of the vector ϵ_n is given by [1],

$$\mathbf{w}_{N,n} = \mathbf{d}_n^T X_{N,n} (X_{N,n}^T X_{N,n})^{-1} \tag{1.7}$$

The above expression can also be derived by taking the derivative of (1.6) with respect to $\mathbf{w}_{N,n}$ and setting it to zero. Thus, if we consider $\xi_N(n)$ as a performance measure, then for RLS algorithms, its first derivative with respect to the weight vector is zero.

Substituting for $\mathbf{w}_{N,n}$ based on (1.7) into (1.5) we obtain,

$$\epsilon_n = [I - X_{N,n} (X_{N,n}^T X_{N,n})^{-1} X_{N,n}^T]\mathbf{d}_n \tag{1.8}$$

Introducing the projection operators,

$$P_{N,n} = X_{N,n}(X_{N,n}^T X_{N,n})^{-1} X_{N,n}^T \tag{1.9}$$

and,

$$P_{N,n}^{\perp} = I - P_{N,n} \tag{1.10}$$

we can write,

$$\epsilon_n = P_{N,n}^{\perp} \mathbf{d}_n \tag{1.11}$$

In a vector space interpretation, ϵ_n is the orthogonal projection of the vector $\mathbf{d}_n$ onto the sub-space spanned by the columns of $X_{N,n}$. The projection operator $P_{N,n}$ projects $\mathbf{d}_n$ onto the subspace to form $\hat{\mathbf{d}}_n$ the least squares estimation of $\hat{\mathbf{d}}_n$.

$$\hat{\mathbf{d}}_n = P_{N,n} \mathbf{d}_n \tag{1.12}$$

II. Fast RLS Algorithms

In the derivation of Fast Recursive Least Squares algorithms, reference [2] introduces transversal filters which operate on the input samples to produce filtered errors. The forward prediction filter $\mathbf{f}_{N,n}$ acts on the previous samples $x(n)$ to produce the forward prediction of x_n. Using projection operators, the forward prediction error vector is,

$$\mathbf{e}_n = P^{\perp}_{N,n-1} \quad \mathbf{x}_n \tag{2.1}$$

The Euclidean length of $\mathbf{e}_n$ is referred to as the forward residual power. Thus,

$$\alpha_N(n) = \mathbf{e}_n^T \mathbf{e}_n = \mathbf{x}_n^T \; P^{\perp}_{N,n-1} \quad \mathbf{x}_n \tag{2.2}$$

The filter $\mathbf{f}_{N,n}$ is given by

$$\mathbf{f}_{N,n} = \mathbf{x}_n^T \; X_{N,n-1} \; (X^T_{N,n-1} \;\; X_{N,n-1})^{-1} \tag{2.3}$$

Define the filter operator [2],

$$K_{N,n} = X_{N,n} \;\; (X^T_{N,n} \;\; X_{N,n})^{-1} \tag{2.4}$$

Then

$$\mathbf{f}_{N,n} = \mathbf{x}_n^T \; K_{N,n-1} \tag{2.5}$$

The backward prediction filter $\mathbf{b}_{\mathrm{N,n}}$ is defined by

$$\mathbf{b}_{\mathrm{N,n}} = \mathbf{x}^T_{n-N} \;\; K_{N,n} \tag{2.6}$$

and the backward prediction error is

$$\mathbf{r}_n = P^{\perp}_{N,n} \quad \mathbf{x}_{n-N} \tag{2.7}$$

The backward residual power is,

$$\beta_N(n) = \mathbf{r}_n^T \mathbf{r}_n = \mathbf{x}^T_{n-N} \;\; P^{\perp}_{N,n} \quad \mathbf{x}_{n-N} \tag{2.8}$$

Define the pinning vector, $(M \times 1)$,

$$\boldsymbol{\sigma}_n = [1, \;\; 0, \;\; 0, \;\; 0 \ldots 0]^T \tag{2.9}$$

Then we can project the pinning vector onto the subspace spanned by the columns of $X_{N,n}$. The associated error vector is,

$$\boldsymbol{\gamma}_n = P_{N,n} \quad \boldsymbol{\sigma}_n \tag{2.10}$$

and its Euclidean length is

$$\gamma_N(n) = \sigma_n^T \; P_{N,n} \sigma_n \tag{2.11}$$

The filter associated with the least squares prediction of σ_n is,

$$c_{N,n} = \sigma_n^T \; K_{N,n} \tag{2.12}$$

In Table I, the predictions errors, their associated filters and definitions are summarized.

Table I

Prediction Error Vector	Definition	Residual Power	Transversal Filter	Transversal Filter Definition
ϵ_n	$P^{\perp}_{N,n} \; \mathbf{d}_n$	$\xi_N(n) = \mathbf{d}_n^T P^{\perp}_{N,n} \; \mathbf{d}_n$	$\mathbf{w}_{N,n}$	$\mathbf{d}_n^T K_{N,n}$
$\mathbf{e}_n$	$P^{\perp}_{N,n-1} \; \mathbf{x}_n$	$\alpha_N(n) = \mathbf{x}_n^T P^{\perp}_{N,n-1} \; \mathbf{x}_n$	$\mathbf{f}_{N,n}$	$\mathbf{x}_n^T K_{N,n-1}$
$\mathbf{r}_n$	$P^{\perp}_{N,n} \; \mathbf{x}_{n-N}$	$\beta_N(n) = \mathbf{x}_{n-N}^T P^{\perp}_{N,n} \mathbf{x}_{n-N}$	$\mathbf{b}_{N,n}$	$\mathbf{x}_{n-N}^T K_{N,n}$
γ_n	$P^{\perp}_{N,n} \; \sigma_n$	$\gamma_N(n) = \sigma_n^T P^{\perp}_{N,n} \sigma_n$	$\mathbf{c}_{N,n}$	$\sigma_n^T K_{N,n}$

III. Random Perturbations in Transversal Filter Coefficients

We now pose the following question. If we consider the residual powers as performance measures, then how much deviation results if the optimum transversal filter coefficients are perturbed by a small random variable?

Consider the case where the weight vector, $\mathbf{w}_{N,n}$ is perturbed by the random vector $\delta_N(n)$. We are interested in the deviation, $\Delta_\xi(n)$, from the optimal residual power $\xi_N(n)$ due to these perturbations. The perturbations are zero mean, independent random variables. Thus,

$$\delta_N(n) = [\delta_0(n) \quad \cdots \quad \delta_{N-1}(n)]^T \tag{3.1}$$

$$E\{\delta_i(n)\} = 0 \tag{3.2}$$

$$E\{\delta_i(n) \ \delta_j(n)\} = \begin{cases} \sigma_\delta^2 & i = j \\ 0 & i \neq j \end{cases} \tag{3.3}$$

The ideal infinite precision residual power is a function of the weight coefficients

$$\xi_N(n) = f\,[w_0(n), \ w_1(n), \ \cdots, \ w_{N-1}(n)] \tag{3.4}$$

Therefore, consider the general case of a performance measure f_∞ which is a function of the coefficients, c_i, $i=0,\ N-1$. Thus,

$$f_\infty = f\,(c_0, \ c_1, \ldots, \ c_{N-1}) \tag{3.5}$$

Now, if each coefficient c_i, is perturbed by dc_i, then the degradation in the performance measure, f_∞, is

$$df\,(c_0, c_1, \cdots c_{N-1}) \approx \sum_{i=0}^{N-1} \left(\frac{\partial f}{\partial c_i} \Big|_\infty dc_i \right), \tag{3.6}$$

The above expression is used to estimate the spectral deviation of fixed digital filters due to the quantization of filter coefficients. In the case of RLS algorithms, the first derivative of the performance measures $\xi_N(n)$ with respect to the weight filter coefficient, $w_i(n)$, is zero. Hence, the above approach cannot be used. Instead, consider the performance measure $J(c_0, c_1, \cdots c_{N-1})$ where $\dfrac{\partial J}{\partial c_i} = 0 \ \ i=0, \cdots N-1$.

Then using second order terms,

$$dJ \approx \frac{1}{2} \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} \left(\frac{\partial^2 J}{\partial c_i \partial c_j} \Big|_\infty dc_i dc_j \right) \tag{3.7}$$

Now, the perturbations, dc_i are zero mean, independent random variables. Hence dJ is also random and its mean value is [11],

$$\overline{dJ} = E\{dJ\} = \frac{1}{2}\sum_{i=0}^{N-1}\left(\frac{\partial^2 J}{\partial c_i^2}\Big|_{b_o}\right) E\left\{(dc_i)^2\right\} \neq 0 \tag{3.8}$$

As in [11], before we derive an expression for dJ, let us define the random variable ϵ_i to be the square of dc_i.

$$\epsilon_i = (dc_i)^2 \tag{3.9}$$

Define its mean and variance as,

$$E\{\epsilon_i\} = \bar{\epsilon}$$

Thus

$$\overline{dJ} = \frac{1}{2}\sum_{i=0}^{N-1}\left(\frac{\partial^2 J}{\partial c_i^2}\Big|_{b_o}\right)\bar{\epsilon} \tag{3.10}$$

Define,

$$E\{\epsilon_i^2\} = \overline{\epsilon^2} \tag{3.11}$$

Thus,

$$E\{(dJ)^2\} = \frac{\overline{\epsilon^2}}{4}\sum_{i=0}^{N-1}\left(\frac{\partial^2 J}{\partial c_i^2}\Big|_{b_o}\right)^2 + \tag{3.12}$$

$$\frac{(\bar{\epsilon})^2}{4}\sum_{\substack{i=0\\ i\neq j}}^{N-1}\sum_{j=0}^{N-1}\left(\frac{\partial^2 J}{\partial c_i^2}\Big|_{b_o}\right)\left(\frac{\partial^2 J}{\partial c_j^2}\Big|_{b_o}\right) + \frac{(\bar{\epsilon})^2}{2}\sum_{\substack{i=0\\ i\neq j}}^{N-1}\sum_{j=0}^{N-1}\left(\frac{\partial^2 J}{\partial c_i \partial c_j}\Big|_{b_o}\right)^2$$

Hence the variance of dJ is [11],

$$\sigma_{dJ}^2 = \overline{\epsilon^2}\frac{-}{4}\sum_{i=0}^{N-1}\left(\frac{\partial^2 J}{\partial c_i^2}\Big|_{b_o}\right)^2 + (\bar{\epsilon})^2\sum_{\substack{i=0\\ i\neq j}}^{N-1}\sum_{j=0}^{N-1}\left(\frac{\partial^2 J}{\partial c_i \partial c_j}\Big|_{b_o}\right)^2 \tag{3.13}$$

IV. Application to Transversal RLS Algorithms

In the RLS algorithm, $\xi_N(n)$ is the performance measure which is minimized. Define the deviation from $\xi_N(n)$ as $\Delta_\xi(n)$, so that

$$\xi_N'(n) = \xi_N(n) + \Delta_\xi(n) \tag{4.1}$$

where the prime denotes the value due to perturbations. Then since the perturbations are random variables $\Delta_\xi(n)$ is also random. We are interested in deriving the mean value and variance of this deviation from the performance measure. Based on (3.10) and (3.13), the deviation $\Delta_\xi(n)$ from the optimum depends on the terms,

$$\frac{\partial^2 \xi_N(n)}{\partial w_i^2(n)} \quad i=0,\ldots,N-1 \tag{4.2}$$

and

$$\frac{\partial^2 \xi_N(n)}{\partial w_i(n) \partial w_j(n)} \quad i=0,\ldots,N-1 \ ; \ j=0,\ldots,N-1 \tag{4.3}$$

These terms form the diagonal and off-diagonal elements of the following matrix.

$$\frac{\partial^2 \xi_N(\mathrm{n})}{\partial \mathbf{w}_{\mathrm{N,n}}^2} \tag{4.4}$$

Now, from (1.5) and (1.6),

$$\xi_N(n) = \boldsymbol{\epsilon}_n^T \boldsymbol{\epsilon}_n = \mathbf{d}_\mathrm{n}^T \mathbf{d}_\mathrm{n} - \mathbf{d}_\mathrm{n}^T \mathrm{X}_{\mathrm{N,n}} \mathbf{w}_{\mathrm{N,n}} - \tag{4.5}$$

$$\mathbf{w}_{\mathrm{N,n}}^T \mathrm{X}_{\mathrm{N,n}}^T \mathbf{d}_\mathrm{n} - \mathbf{w}_{\mathrm{N,n}}^T \mathrm{X}_{\mathrm{N,n}}^T \mathrm{X}_{\mathrm{N,n}} \mathbf{w}_{\mathrm{N,n}}$$

$$\frac{\partial \xi_N(\mathrm{n})}{\partial \mathbf{w}_{\mathrm{N,n}}} = -2\mathbf{d}_\mathrm{n}^T \mathrm{X}_{\mathrm{N,n}} + 2\mathrm{X}_{\mathrm{N,n}}^T \mathrm{X}_{\mathrm{N,n}} \mathbf{w}_{\mathrm{N,n}} \tag{4.7}$$

Thus,

$$\frac{\partial^2 \xi_N(\mathrm{n})}{\partial \mathbf{w}_{\mathrm{N,n}}^2} = 2\mathrm{X}_{\mathrm{N,n}}^T \mathrm{X}_{\mathrm{N,n}} \tag{4.8}$$

Define the matrix $Q(n)$,

$$Q(n) = X_{N,n}^T X_{N,n} \tag{4.7}$$

Therefore, the properties of $Q(n)$ determine the sensitivity of the algorithm to perturbations in the weight coefficients. In Table II the matrix for each filter of the Fast RLS algorithm is presented.

Table II

Transversal Filter	Residual Power	Matrix $Q(n)$
$\mathbf{w}_{N,n}$	$\xi_N(n)$	$X_{N,n}^T X_{N,n}$
$\mathbf{f}_{N,n}$	$\alpha_N(n)$	$X_{N,n-1}^T X_{N,n-1}$
$\mathbf{b}_{N,n}$	$\beta_N(n)$	$X_{N,n}^T X_{N,n}$
$\mathbf{c}_{N,n}$	$\gamma_N(n)$	$X_{N,n}^T X_{N,n}$

V. Derivation of Results

Prewindowed Growing Memory Algorithm

The matrix $Q(n)$ determines the sensitivity of the algorithm to perturbations in the filter coefficients. In this section we will attempt to analyze this matrix. Define

$$\bar{R}(n) = \frac{1}{n} \sum_{i=0}^{n} \mathbf{x}(i)\mathbf{x}^T(i) \tag{5.1}$$

where

$$\mathbf{x}(n) = [x(n) \;\; x(n-1) \;\; \cdots \;\; x(n-N+1)]^T \tag{5.2}$$

Then,

$$Q(n) = X_{N,n}^T \ X_{N,n} = \sum_{i=0}^{n} \mathbf{x}(i)\mathbf{x}^T(i) = n\bar{R}(n) \tag{5.3}$$

Now, as n becomes large,

$$\lim_{n \to \infty} \bar{R}(n) \approx E\{\mathbf{x}(i)\mathbf{x}^T(i)\} = R_x \tag{5.4}$$

where R_x is the sample autocorrelation matrix. Then,

$$Q(n) \approx n \ R_x \tag{5.5}$$

If $x(n)$ is a white random process,

$$R_x = \sigma_x^2 I \tag{5.6}$$

Therefore, from (3.10) the mean deviation from the optimum due to random perturbations $\boldsymbol{\delta}_N(n)$ in the weight vector is,

$$\overline{\Delta_\xi} = n\bar{\epsilon} \ N\sigma_x^2 \tag{5.7}$$

In the above expression we have defined,

$$\epsilon_i(n) = \delta_i^2(n)$$

from which,

$$\bar{\epsilon} = E\{\delta_i^2(n)\}$$

and

$$\overline{\epsilon^2} = E\{\epsilon_i^2(n)\}$$

The variance of the deviation is,

$$\sigma_{\Delta_\xi}^2 = n^2\overline{\epsilon^2} \ N\sigma_x^4 \tag{5.8}$$

If the perturbations are assumed to be zero mean uncorrelated random processes, then

$$\overline{\Delta_\xi} = \frac{n}{2} N \sigma_\delta^2 \sigma_x^2 \tag{5.9}$$

Exponentially Windowed Algorithm

For the exponentially windowed RLS algorithm,

$$Q(n) = X_{N,n}^T X_{N,n} = \sum_{i=0}^{n} \lambda^{n-i} \mathbf{x}(i)\mathbf{x}^T(i) \tag{5.10}$$

Hence,

$$Q(n) \approx \sum_{i=0}^{n} \lambda^{n-i} \sigma_x^2 I = \frac{1-\lambda^{n+1}}{1-\lambda} \sigma_x^2 I \tag{5.11}$$

Now for n large, since $\lambda<1$,

$$Q(n) \approx \frac{1}{1-\lambda} \sigma_x^2 I \tag{5.12}$$

This leads to

$$\overline{\Delta_\xi} = \frac{N}{1-\lambda} \bar{\epsilon} \sigma_x^2 \tag{5.13}$$

The variance is,

$$\sigma_{\Delta_\xi}^2 = N \sigma_x^4 \epsilon^2 \frac{\overline{\quad}}{(1-\lambda)^2} \tag{5.14}$$

VI. Deviation for Correlated Signals

For correlated signals, we have for $\lambda=1$ from (5.5),

$$Q(n) \approx nR_x \tag{6.1}$$

where

$$R_x = \begin{bmatrix} r_x(0) & r_x(1) & \cdots & r_x(N-1) \\ r_x(1) & r_x(0) & \cdots & r_x(N-2) \\ \vdots & & & \vdots \\ r_x(N-1) & \cdots & & r_x(0) \end{bmatrix} \quad (6.2)$$

and

$$r_x(i) = E\{x(n)x(n-i)\} \quad (6.3)$$

In particular

$$r_x(0) = \sigma_x^2 \quad (6.4)$$

Define the scalar function of any square matrix A, $g(A)$ such that

$$g(A) = \sum_{i \neq j}\sum a_{ij}^2 \quad (6.5)$$

Then from (3.10), (6.1), and (6.4)

$$\overline{\Delta}_\xi(n) = nN\overline{\epsilon}\sigma_x^2 \quad (6.6)$$

From (3.13)

$$\sigma_{\Delta_\xi}^2(n) = n^2\overline{\epsilon^2}N\sigma_x^4 + 4n^2(\overline{\epsilon})^2 g(R_x) \quad (6.7)$$

Comparing (6.6) and (6.7) with the results for white random inputs, (5.7) and (5.8), we observe that the mean of the deviation does not change for equal power correlated and white signals. However, the variance of the deviation increases for correlated signals. This increase is a function of the square of the off-diagonal element of the autocorrelation matrix. Thus, as the correlation increases, the performance of the RLS algorithm degrades compared to white signals.

Since R_x is symmetric we can write (6.7) as

$$\sigma^2_{\Delta_\xi}(n) = n^2\overline{\epsilon^2}N\sigma_x^4 + 4n^2(\bar{\epsilon})^2N\sum_{i=1}^{N-1} r_x^2(i) \tag{6.8}$$

VII. Interpretation of Results

In order to interprete the results derived so far and to come to some useful conclusions we reexamine the expression for the error residual power:

$$\xi_N(n) = \epsilon_n^T\epsilon_n = [\mathbf{d}_n - X_{N,n}\ \mathbf{w}_{N,n}]^T[\mathbf{d}_n - X_{N,n}\ \mathbf{w}_{N,n}] \tag{7.1}$$

From this expression it is seen that the RLS algorithm estimates the filter $\mathbf{w}_{N,n}$ at each iteration such that the error power using this filter acting on all samples up to time n is minimized. Now if in (7.1) $\mathbf{w}_{N,n}$ is perturbed then as it acts on the input samples from time 0 to n an error will occur which based on the expression will accumulate with time. Thus, the error power will grow as n is increased which is precisely what the expression for the deviation (5.9) predicts.

In an infinite precision implementation of the fast RLS algorithm, the algorithm will compute the residual powers in Table I. The joint process error power $\xi_N(n)$ may or may not be computed in an RLS application but assume that it was. The point is that in the fast algorithms the error powers are computed. Now, in a finite precision implementation of the algorithm at each time step the filter coefficients will deviate due to roundoff arithmetic from the optimum. Since in essence the algorithm is evaluating the error powers based on Table I and expressions like (7.1), then as predicted by (5.9) roundoff errors will accumulate and the deviation from the optimum error power will increase with the number of iterations.

For the exponentially windowed RLS algorithm, the effects of perturbations decay due to the exponential weighting and the deviation becomes bounded as predicted by (5.13). This is of course the result predicted in [5].

At this point it is instructive to connect the results of this paper to the results in [9,10] where the roundoff error due to fixed point and floating point implementations of the RLS algorithms were derived. In both cases it was shown that the mean square roundoff error due to rounding in the updating of the weight vector increased linearly with the number of iterations for $\lambda=1$ and was bounded and inversely proportional to $1-\lambda$ for $\lambda<1$. These results, which were confirmed by simulations, are exactly what is predicted by this paper through a straight forward approach. It must be emphasized that the results of this paper only indicate how the deviations depend on the number of iterations and λ. They do not predict the actual mean square roundoff error as in [9,10]. Significantly, however, in [9] and [10] the variance of the fluctuations about the mean square roundoff error was not computed. The results of this paper point out that this variance increases with the square of the number of iterations for $\lambda=1$ as shown in (5.8). In [9] simulations are presented in which the variance of the roundoff error increases as the number of iterations is increased. This was not commented on in [9] however.

Another important result of this paper is that it shows that the deviation increases when the input signal is correlated. The case of correlated signals was not examined in [9] and [10].

In conclusion, a straight forward approach inspired by [11] was introduced in this paper to investigate the sensitivity of the transversal RLS algorithm to perturbations in the optimum filter coefficients. Although the results to not predict the deviation from the optimum due to finite precision effects as in [9] and [10] they do support those results and further extend the general conclusions about the behaviour of the RLS algorithm to correlated signals.

VIII. References

[1] L. Ljung and T. Soderstrom, *Theory and Practice of Recursive Identification*, MIT Press, 1983.

[2] John M. Cioffi and A Thomas Kailath, "Fast, recursive-least-squares transversal filters for adaptive filtering" *IEEE Transactions on Acoustics, Speech, and Signal Processing*, Vol ASSP-32, No. 2, April 1984, pp. 304-337.

[3] G. Carayannis, D.G. Manolakis, N. Kalouptsids, "A fast sequential algorithm for least-squares filtering and prediction," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, Vol ASSP-32, No. 2, April 1984, pp. 304-337.

[4] T. A. C. M. Claasen and W. F. G. Mecklenbrauker, "Adaptive techniques for signal processing in communications," *IEEE Communications Magazine*, Vol. 23, No. 11, November 1985.

[5] S. Ljung and L. Ljung, "Error propagation properties of recursive least-squares adaptation algorithms," *Automatica*, Vol. 21, No. 2, March 1985.

[6] D.D. Falconer and L. Ljung, "Application of Fast-Kalman estimation to adaptive equalization" *IEEE Transactions on Communication*, Vol. COMM-26, No. 110, Oct. 1978.

[7] M. Bellanger and C. C. Evci, "Coefficient wordlength limitation in FLS adaptive filters," *Proceedings of ICASSP 86*, April 7, 1986, Tokyo, Japan, pp. 3011-3014.

[8] J. Botto, "Stabilization of fast Recursive Least Squares transversal filters for adaptive filtering," *Proc. IEEE ICASSP 87*, Dallas, Texas. pp. 403-407.

[9] Sasan Ardalan, S.T. Alexander, "Fixed point roundoff error analysis of the exponentially windowed RLS algorithm for time-varying systems," *IEEE Trans. on Acoustics Speech and Sig. Proc.*, Vol. ASSP-35, No. 6, June 1987.

[10] Sasan H. Ardalan,"Floating point roundoff error analysis of the RLS and LMS adaptive algorithms," *IEEE Trans. on Circuits and Systems*, Vol. CAS-33, No. 12, December 1986, pp. 1192-1208.

[11] P. Moroney, *Issues in the Implementation of Digital Feedback Compensators*, MIT Press, Cambridge, MA, 1983.

[12] John M. Cioffi, "Limited precision effects in adaptive filtering", *IEEE Trans. on Circuits and Systems*, Vol. CAS-34, No. 7, July 1987.