Round-Off Error Propagation in Four Generally Applicable, Recursive, Least-Squares-Estimation Schemes

M. H. Verhaegen

December 1987

Round-Off Error Propagation in Four Generally Applicable, Recursive, Least-Squares-Estimation Schemes

M. H. Verhaegen, Ames Research Center, Moffett Field, California

December 1987

NASA

Abstract

In this paper the numerical robustness of four generally applicable, recursive, least-squares-estimation schemes is analyzed by means of a theoretical round-off propagation study. This study highlights a number of practical, interesting insights of widely used recursive least-squares schemes. These insights have been confirmed in an experimental verification study as well.

Keywords

recursive least squares, divergence phenomenon, theoretical error analysis

1 Introduction

Recursive least squares (RLS) schemes are used in a broad class of practical applications, such as in the self-tuning regulator schemes developed by Åstrom et al. (1977). The simplicity of this class of schemes has been an important motive for their use. Some of the other reasons are 1) the versatility of implementations, allowing to exploit special (shift) structures in the regression model, first demonstrated by Levinson (1947), 2) the well-known statistical properties of the RLS estimates, and 3) robust ways to cope with time variations in the regression parameters, see T. Hägglund (1983) or R. Kulhavy (1985).

However, the numerical robustness of a number of RLS implementations is still not well understood. This can be concluded from the appearance of simulation studies reporting the *divergence* phenomenon in the implementation referred to in this paper as the conventional RLS implementation. Here the divergence phenomenon is a generic term used to indicate a whole class of problems where the quantities updated in the actual computer implementation lose their theoretical properties. A well-known example is the loss of symmetry of the parameter error covariance matrix.

Until now no explanation has been given for this phenomenon, only a number of "cures" have been proposed. These cures have two disadvantages: 1) they might unnecessarily complicate the numerical implementation and 2) they might give "satisfactory" results in simulation, but actually fail in real-time operation.

In numerical analysis, an error analysis is performed to understand the numerical robustness of algorithmic implementations. For RLS, such an analysis has been performed

1

in S. Ljung and L. Ljung (1985). Although, this analysis reveals some fundamental understanding of the robustness of RLS schemes, it does not answer the following important practical questions:

1. When does the loss of symmetry occurs in the conventional RLS scheme?

2. How does the loss of symmetry interferes with the loss of positive definiteness of the parameter error covariance matrix?

3. Does the conventional RLS scheme result in numerically less-accurate results compared to square root type of RLS?

4. When using a square root implementation do we choose a covariance-type or an information-type RLS?

To answer these questions, a new and detailed error analysis is performed in this paper. This analysis considers the following four generally applicable implementations of RLS:

1. Two implementations of the conventional RLS (indicated respectively as the CLS1 and CLS2).

2. The square-root-covariance RLS implementation (SCLS).

3. The square-root-information RLS implementation (SILS).

These implementations are precisely defined in section 2. Section 3 then presents the results of the round-off error propagation study for these four implementations. The new insights from this error analysis are evaluated experimentally in section 4. Finally, section 5 presents some concluding remarks.

2 Recursive Least Squares Estimation

2.1 The Linear Least Squares Problems

Let the linear scalar regression model be denoted as:

$$y_k = \vartheta' \varphi_k + e_k \tag{1}$$

where the regressor vector $\varphi_k \in R^n$, e_k is a zero mean discrete white noise sequence with variance σ_k^2 [2] and $(.)'$ denotes the transpose. When the observations of (y_k, φ_k) have been obtained for $k = 1, \cdots, N$ (with $N > n$), the least squares (LS) estimate of ϑ is defined as:

$$\hat{\vartheta}_N = \arg \min \sum_{i=1}^{N} \lambda^{N-i} (y_i - \vartheta' \varphi_i)^2 \tag{2}$$

Here $\{\lambda^{N-i}\}$ is a sequence of weighting coefficients that result if we discount old data by so-called "exponential forgetting ($\lambda < 1$)." For this case, simple calculations show that the LS estimate (equaion 2) is given by:

$$\hat{\vartheta}_N = R_N^{-1} f_N \tag{3}$$

$$R_N = \sum_{i=1}^{N} \varphi_i \varphi_i' \lambda^{N-i} \tag{4}$$

$$f_N = \sum_{i=1}^{N} \varphi_i y_i \lambda^{N-i} \tag{5}$$

For real-time operation, it is of primary importance to compute the estimates recursively. Therefore, the set of equations (3-5) is rearranged by simple manipulations to the following recursive form:

$$\hat{\vartheta}_k = \hat{\vartheta}_{k-1} + R_k^{-1} \varphi_k (y_k - \hat{\vartheta}_{k-1}' \varphi_k) \tag{6}$$

$$R_k = \lambda R_{k-1} + \varphi_k \varphi_k' \tag{7}$$

The set of equations (6-7) is updated from $k \geq 0$ on. This can be done when for $k = 0$ the initial conditions $\hat{\vartheta}_{-1}$ and R_{-1} are specified.

---

[2] In practice σ_k is generally not known a priori, but is estimated during the operation of the RLS. However, what actual value for σ_k is used in the RLS will turn out to be not important in the analysis of the numerical robustness

2.2 The Conventional RLS Method (CLS)

The recursive relationships (equations 6-7) are not well suited for computation, since the $n \times n$ information matrix R_k has to be inverted each time step. Therefore, it is more natural to introduce

$$P_k = R_k^{-1}$$

and by applying the matrix inversion lemma, Ljüng and Söderström (1983), (equations 6-7) become

$$\hat{\vartheta}_k = \hat{\vartheta}_{k-1} + K_k(y_k - \hat{\vartheta}'_{k-1}\varphi_k) \tag{8}$$

$$K_k = \frac{P_{k-1}\varphi_k}{\lambda\sigma_k^2 + \varphi'_k P_{k-1}\varphi_k} \tag{9}$$

$$P_k = \frac{1}{\lambda}\left(P_{k-1} - \frac{P_{k-1}\varphi_k\varphi'_k P_{k-1}}{\lambda\sigma_k^2 + \varphi'_k P_{k-1}\varphi_k}\right) \tag{10}$$

Remark 1: The linear regression model (equation 1) can be written in the state space form,

$$\vartheta_{k+1} = \vartheta_k \tag{11}$$

$$y_k = \varphi'_k\vartheta_k + \epsilon_k \quad \text{with} \quad E\left(\epsilon_k^2\right) = \sigma_k^2 \tag{12}$$

Together with the initial conditions $(\hat{\vartheta}_{-1}, P_{-1})$, this model is a special case of the general state space model used in the Kalman filter design (as described in Verhaegen and Van Dooren, 1986):

$$x_{k+1} = A_k x_k + B_k w_k \quad \text{with} \quad E(x_0) = \hat{\vartheta}_{-1} \ , \ E(x_0 x'_0) = P_{-1} \text{ and } E(w_k w'_k) = Q_k \tag{13}$$

$$y_k = C_k x_k + v_k \quad \text{with} \quad E(v_k v'_k) = R_k \tag{14}$$

for the following particular system matrices:

$$A_k = I \ , \ B_k = Q_k = 0 \ , \ C_k = \varphi'_k \text{ and } R_k = \sigma_k^2 \tag{15}$$

For these system matrices (equations 8-10) represent the so-called measurement update of the conventional Kalman filter as defined in Verhaegen and Van Dooren (1986). Therefore, equations (8-10) are conformally indicated in this paper as the conventional RLS (CLS). □

A commonly used implementation of the CLS is represented in Table 1 as the CLS1. Here we observe that for the calculation of the matrix K_k the symmetry of P_{k-1} is exploited.

The only reason for that is a slight reduction in computational complexity. Furthermore, the performed operation to calculate P_{k-1}^* is inherently symmetric. Therefore, this implementation "seemingly" forces the parameter error covariance matrix P_k to remain symmetric. However, the theoretical error analysis will prove that exactly the contrary is the case. Namely, that when implementing equations (8-10) exactly as they are, as done in the CLS2 implementation of Table 1, the sensitivity to the loss of symmetry of P_k vanishes.

quantity	mathematical expression	CLS1	CLS2
fiP	$\varphi_k' P_{k-1}$	$\varphi_k' \times P_{k-1}$	"
r_k^ϵ	$\lambda \sigma_k^2 + \varphi_k' P_{k-1} \varphi_k$	$\lambda.\sigma_k^2 + fiP \times \varphi_k$	"
K_k	$\dfrac{P_{k-1}\varphi_k}{\lambda\sigma_k^2 + \varphi_k' P_{k-1}\varphi_k}$	$\dfrac{fiP'}{r_k^\epsilon}$	$\dfrac{P_{k-1} \times \varphi_k}{r_k^\epsilon}$
P_{k-1}^*	$\dfrac{P_{k-1}\varphi_k\varphi_k' P_{k-1}}{\lambda\sigma_k^2 + \varphi_k' P_{k-1}\varphi_k}$	$K_k \times fiP$	"
P_k	$\frac{1}{\lambda}\left(P_{k-1} - \dfrac{P_{k-1}\varphi_k\varphi_k' P_{k-1}}{\lambda\sigma_k^2 + \varphi_k' P_{k-1}\varphi_k}\right)$	$\frac{1}{\lambda}\left(P_{k-1} - P_{k-1}^*\right)$	"

Table 1: The CLS1 and CLS2 implementation of the conventional RLS.

In this table the (") indicates that the same matrix-vector multiplication is performed for both implementations.

2.3 The Square Root Covariance RLS Method (SCLS)

This can be considered as a special case of the square-root-covariance Kalman-filter implementation, also defined in Verhaegen and Van Dooren (1986). For the sytem matrices given in equation (15), the recursive scheme characterizing this implementation becomes

$$\underbrace{\begin{pmatrix} \sqrt{\lambda}\sigma_k & \varphi_k' S_{k-1} \\ 0 & \frac{1}{\sqrt{\lambda}} S_{k-1} \end{pmatrix}}_{\text{pre-array}} U_1 = \underbrace{\begin{pmatrix} \sigma_k^\epsilon & 0 \\ G_k & S_k \end{pmatrix}}_{\text{post-array}} \tag{16}$$

with the parameter update equation

$$\hat{\vartheta}_k = \hat{\vartheta}_{k-1} + \frac{\sqrt{\lambda}}{\sigma_k^\epsilon} G_k(y_k - \hat{\vartheta}_{k-1}'\varphi_k) \tag{17}$$

and U_1 in equation (16) is an arbitrary orthogonal transformation that triangularizes the

prearray. Here S_k is the Choleski factor [3] of P_k, defined as follows:

$$P_k = S_k S_k'$$

(18)

Remark 2: Another type of square root covariance RLS is the so-called UD-factorization algorithm (Bierman, 1976). The operational complexity of this scheme is slightly less than the presented SCLS scheme. However, from the point of view of numerical reliability, preliminary simulation tests performed in Verhaegen (1985) demonstrated a similar round-off error pattern between these two schemes. Therefore, the latter implementation will not be discussed further in this paper. □

2.4 The Square Root Information RLS Method (SILS)

Similarly to the square-root-information Kalman-filter scheme, we can formulate for the LS problem recursions for the Choleski factor of the matrix R_k, defined in equation (4), and the parameter estimate $\hat{\vartheta}_k$ in a combined way. When the Choleski factor of R_k is defined as:

$$R_k = T_k' T_k$$

(19)

this implementation becomes

$$U_2 \begin{pmatrix} \sqrt{\lambda} T_{k-1} & \sqrt{\lambda}\hat{\xi}_{k-1} \\ \frac{\varphi_k'}{\sigma_k} & \frac{y_k}{\sigma_k} \end{pmatrix} = \begin{pmatrix} T_k & \hat{\xi}_k \\ 0 & \varepsilon_k \end{pmatrix}$$

(20)

with U_2 operating in a similar manner as U_1. The update of the parameter estimates is given by:

$$\hat{\vartheta}_k = T_k^{-1}(\hat{\xi}_k)$$

(21)

Remark 3: Other, so-called *fast LS* algorithms, such as the Levinson algorithm (Levinson, 1947) are not included in this paper. This is because their numerical behavior has already been analyzed by an error perturbation study. For example for Levinson related algorithms this has been done in Cybenko, 1980. Furthermore, these algorithms impose a special structure upon the regression model, and this restriction is not considered in this paper. □

---

[3] The "Choleski factor" is often called the "square root." In this paper the term "square root" is maintained as far as the names of the square-root RLS schemes are concerned because of the familiarity that it has acquired.

3 Theoretical Error Analysis

The *numerical robustness* with respect to round-off errors is analyzed for the four RLS schemes described in the previous section. Such an analysis can be split into three parts: (1) the round-off errors made in a single recursion, (2) the propagation of a single error in subsequent recursions and (3) the interaction and accumulation of the previous two error sources. In the paper of S. Ljung and L. Ljung (1985) only the second part has been treated. The following subsections present a much more refined analysis.

3.1 Propagation of a Single Error Under ∞ Computational Precision

In this subsection we consider the propagation of a single error at recursion instant $k-1$ to subsequent recursions, assuming that no additional round-off errors are made. The following theorem describes this propagation for the CLS1 and CLS2 implementation.

Theorem 1: Define the erroneous quantities, denoted by $\overline{(.)}$, at the $k-1$ recursion instant as,

$$\overline{P}_{k-1} = P_{k-1} + \delta P_{k-1} \quad \text{and} \quad \overline{\hat{\vartheta}}_{k-1} = \hat{\vartheta}_{k-1} + \delta\hat{\vartheta}_{k-1} \tag{22}$$

then these errors propagate to the next recursion instant k in the CLS1 implementation as:

$$\delta P_k = \frac{1}{\lambda}(I - K_k\varphi'_k)\delta P_{k-1}(I - K_k\varphi'_k)' + \frac{1}{\lambda}(\delta P_{k-1} - \delta P'_{k-1})\varphi_k K'_k + O(\delta^2) \tag{23}$$

$$\delta\hat{\vartheta}_k = (I - K_k\varphi'_k)\left[\delta\hat{\vartheta}_{k-1} + \frac{\delta P_{k-1}\varphi_k}{(\lambda\sigma_k^2 + \varphi'_k P_{k-1}\varphi_k)}(y_k - \varphi'_k\hat{\vartheta}_{k-1})\right] + O(\delta^2) \tag{24}$$

and in the CLS2 implementation as:

$$\delta P_k = \frac{1}{\lambda}(I - K_k\varphi'_k)\delta P_{k-1}(I - K_k\varphi'_k)' + O(\delta^2) \tag{25}$$

$$\delta\hat{\vartheta}_k = (I - K_k\varphi'_k)\left[\delta\hat{\vartheta}_{k-1} + \frac{\delta P_{k-1}\varphi_k}{(\lambda\sigma_k^2 + \varphi'_k P_{k-1}\varphi_k)}(y_k - \varphi'_k\hat{\vartheta}_{k-1})\right] + O(\delta^2) \tag{26}$$

where $O(\delta^2)$ indicates the order of magnitude of $\|\delta P_{k-1}\|^2$.

Proof: (The proof will only be given for the CLS1 implementation)

Consider the approximation,

$$\frac{1}{\lambda\sigma_k^2 + \varphi'_k P_{k-1}\varphi_k + \varphi'_k \delta P_{k-1}\varphi_k} = \frac{1}{\lambda\sigma_k^2 + \varphi'_k P_{k-1}\varphi_k}\left(1 - \frac{\varphi'_k \delta P_{k-1}\varphi_k}{\lambda\sigma_k^2 + \varphi'_k P_{k-1}\varphi_k}\right) + O(\delta^2)$$

Then the error on $\overline{K}_k$ updated in the CLS1 becomes

$$\delta K_k = \frac{\delta P'_{k-1}\varphi_k}{\lambda\sigma_k^2 + \varphi'_k P_{k-1}\varphi_k} + \frac{P'_{k-1}\varphi_k}{\lambda\sigma_k^2 + \varphi'_k P_{k-1}\varphi_k}\left(1 - \frac{\varphi'_k\delta P_{k-1}\varphi_k}{\lambda\sigma_k^2 + \varphi'_k P_{k-1}\varphi_k}\right) + O(\delta^2)$$

Similar matrix-error analysis then results in the error propagation model (equation 23) for δP_k.

Following the same line of derivation, the error model (equation 24) for $\delta\dot{\vartheta}_k$ results. □

Corollary 1: The error propagation model (equation 23) demonstrates that for an exponential forgetting factor $\lambda \leq 1$, the effect of the *loss of symmetry* causes a *blow up* (divergence) in the error on P_k. Therefore, when using the CLS1 implementation it is absolutely necessary to maintain symmetry of the matrix P_k to avoid divergence, no matter how accurately the computations are performed. This divergence phenomenon, which is due to loss of symmetry, is well known, but not understood, in literature. Theorem 1, now gives a theoretical explanation of this phenomenon. This theorem furthermore indicates that such a loss of symmetry does not occur when using the CLS2 implementation. □

Corollary 2: Since for the moment we assume that no round-off errors are performed in the recursions, one inherently computes the same equations for the SCLS scheme as for the CLS scheme. Therefore, starting with an error δS_{k-1}, which induces the following error on P_{k-1}:

$$\delta P_{k-1} = S_{k-1}\delta S'_{k-1} + \delta S_{k-1}S'_{k-1} + \delta S_{k-1}\delta S'_{k-1} \tag{27}$$

and the error $\delta\dot{\vartheta}_{k-1}$, the error propagation models (equations 25-26) hold for the SCLS implementation.

Remark 4: The same error models (equations 23-24) could be obtained by substituting the special system matrices (equation 15) in the error analysis results obtained for the conventional Kalman filter in Verhaegen and Van Dooren (1986). Repeating such an analysis particularly for the CLS, precisely indicates where the loss of symmetry comes from. □

For the SILS scheme, one *inherently* updates the information matrix R_k according to equation (7). Here an error T_{k-1}, similarly represented as in equation (22), induces an error δR_{k-1} given as:

$$\delta R_{k-1} = T'_{k-1}\delta T_{k-1} + \delta T'_{k-1}T_{k-1} + \delta T'_{k-1}\delta T_{k-1} \tag{28}$$

Again assuming no additional errors, this errors propagates according to:

$$\delta R_k = \lambda \delta R_{k-1} \tag{29}$$

Corollary 3: The error model (equation 29) immediately indicates that when using exponential forgetting and *infinite precision*, "information-matrix"-type estimation schemes are exponentially stable. For $\lambda = 1$ the error propagation does not increase.

For the CLS1, the CLS2 and the SCLS implementations, the matrix $(I - K_k \varphi'_k)$ plays a crucial role in the propagation of both the errors δP_{k-1} and $\delta \dot{\vartheta}_{k-1}$. Therefore, let us inspect this matrix more closely. Substituting the expression for K_k, obtained by comparing equations (6) and (8) and using equation (7) we find that:

$$(I - K_k \varphi'_k) = \lambda R_k^{-1} R_{k-1} \tag{30}$$

Now, consider the effect of an error induced at time instant $k_0 < k$ on the computed quantities $\overline{P}_k$ and $\overline{\vartheta}_k$. When $\overline{P}_k$ remains symmetric, this effect becomes according to equation (25):

$$\delta P_k = \frac{1}{\lambda^{k-k_0}} \phi(k, k_0) \delta P_{k_0} \phi(k, k_0)' + O(\delta^2) \tag{31}$$

and according to equation (24):

$$\delta \dot{\vartheta}_k = \phi(k, k_0) \delta \dot{\vartheta}_{k_0} + \sum_{j=0}^{k-k_0-1} \phi(k, k_0 + j).\mu_{k_0+1+j} + O(\delta^2) \tag{32}$$

with $\mu_i = \frac{\delta P_{i-1} \varphi_i}{(\lambda \sigma_i^2 + \varphi'_i P_{i-1} \varphi_i)}(y_i - \varphi'_i \dot{\vartheta}_{i-1})$ and $\phi(k, k_0 + j) = \prod_{i=k_0+j}^{k}(I - K_i \varphi'_i)$. Assume now that the conditions which guarantee that the residual $(y_i - \varphi'_i \dot{\vartheta}_{i-1})$ becomes a zero-mean random variable hold. Then the second term in the right-hand side of equation (32) vanishes when we consider only the mean of $\delta \dot{\vartheta}_k$, as denoted by $E(\delta \dot{\vartheta}_k)$. Under this assumption, we will only analyze the propagation of $E(\delta \dot{\vartheta}_k)$ in the sequel.

The errors δP_{k_0} and $E(\delta \dot{\vartheta}_{k_0})$ are attenuated if the transition matrix $\phi(k, k_0)$ is a *contraction*. In the next theorem it will be shown that this is indeed the case if the regressor vector φ_k is persistently exciting. Let us first define this condition of the regressor vector.

Definition 1: The regressor vector φ_i is persistently exciting over the observation interval $k_0 \leq i \leq k$ when using an exponential forgetting factor $\lambda \leq 1$, if the following condition is

fulfilled,

$$\alpha I \leq \sum_{i=k_0}^{k} \varphi_i \varphi_i' \lambda^{k-i} \leq \beta I \tag{33}$$

for some positive constants α and β.

Theorem 2: When the regressor vector φ_i satisfies condition equation (33) over the interval (k_0, k), the transition matrix $\phi(k, k_0)$ is a contraction.

Proof:

The matrix R_k is defined in equation (4) as,

$$R_k = \sum_{i=0}^{k} \varphi_i \varphi_i' \lambda^{k-i}$$

Since $k_0 < k$ we have that:

$$R_k = \lambda^{k-k_0} R_{k_0} + \sum_{i=k_0}^{k} \varphi_i \varphi_i' \lambda^{k-i}$$

Therefore, when equation (33) holds, the matrix difference $(R_k - \lambda^{k-k_0} R_{k_0})$ is *positive definite*, or alternatively denoted:

$$\lambda^{k-k_0} R_{k_0} < R_k$$

When we now express $\varphi(k, k_0)$ as $\lambda^{k-k_0} R_k^{-1} R_{k_0}$, using equation (30), $\phi(k, k_0)$ is clearly positive definite. This completes the proof.

Corollary 4: When the following conditions are met.

1. When the regressor vector is persistently exciting, and hence, R_k^{-1} remains bounded (which can easily be shown).

2. When $\lambda < 1$.

3. When no additional round-off errors are made in subsequent recursions.

then theorem 2 guarantees that a single error δP_{k_0} in equation (31) or $E(\delta \hat{\vartheta}_{k_0})$ in equation (32) decays exponentially when using the CLS2 and SCLS implementations. For the case $\lambda = 1$, the contraction of $R_k^{-1} R_{k_0}$ guarantees that the error propagation remains stable. For the CLS1 implementation this holds when additionally the symmetry on P_k is forced.

Another divergence phenomenon reported in literature is the *loss of positive definiteness* of P_k. The question we might ask is how this phenomenon is related or interferes with the loss of symmetry discussed so far. Some insight into this complicated matter is gained by the following theorem.

Theorem 3: When the errors on the parameter error covariance matrix P_{k-1} are symmetric and preserve its positive definiteness, the $\overline{P}_k$ updated by equation (10) remains positive definite.

Proof:

Since $\overline{P}_{k-1}$ is positive definite, it may be written as,

$$\overline{P}_{k-1} = XDX' \tag{34}$$

where $XX' = I_n$ and $D = diag(d_1, \cdots, d_n)$, with $d_i > 0$. Since X is of full rank, we might write the regressor vector φ_k as:

$$\varphi_k = X\mu \tag{35}$$

with $\mu = (\mu_1, \cdots, \mu_n)'$.

Substituting equations (34-35) in the update relationship for $\overline{P}_k$, we obtain:

$$\overline{P}_k = X(D - \frac{D\mu\mu'D}{1 + \mu'D\mu})X' \tag{36}$$

Evaluating the matrix between brackets results in:

$$(1 + \mu'D\mu)\overline{P}_k =$$

$$X \begin{pmatrix} \alpha_1 & -\mu_1\mu_2 d_1 d_2 & \cdots & -\mu_1\mu_n d_1 d_n \\ -\mu_1\mu_2 d_1 d_2 & \alpha_2 & \cdots & -\mu_2\mu_n d_2 d_n \\ & & \ddots & \\ -\mu_1\mu_n d_1 d_n & -\mu_2\mu_n d_2 d_n & \cdots & \alpha_n \end{pmatrix} X' \tag{37}$$

where $\alpha_i = \lambda\sigma_k^2 d_i + \sum_{j=1}^{n}{}^{j \neq i} \mu_j^2 d_i d_j$. The characteristic polynomial of the matrix between brackets in equation (37) can be evaluated for example by using the symbolic manipulation software package by MACSYMA (1983) for different values of n. For all values of n, this polynomial can be written as:

$$(-1)^n z^n + \sum_{i=0}^{n-1} (-1)^i a_i z^i = 0 \quad \text{with} \quad a_i > 0 \tag{38}$$

Obviously, this polynomial has only positive roots; therefore, $\overline{P}_k$ is positive definite.

3.2 Round-Off Errors Made in a Single Recursion

In the second part of the theoretical error analysis, we study the round-off errors made in a single recursion. The following theorem specifies bounds for the errors on the quantities updated each recursion in the four RLS schemes under investigation in this paper.

Theorem 4: Denoting the norms of absolute errors caused by round off during the construction of P_k, $\hat{\vartheta}_k$, S_k, T_k, and R_k by Δ_P, Δ_ϑ, Δ_S, Δ_T, and Δ_R, respectively, we obtain the following upperbounds (where norms are 2-norms),

1. CLS1 and CLS2

$$\Delta_P \leq \epsilon_1 \|P_k\|$$

$$\Delta_\vartheta \leq \epsilon_2 (\|\hat{\vartheta}_k\| + \|K_k\| \cdot \|y_k\|)$$

2. SCLS

$$\Delta_S \leq \epsilon_3 \|S_k\| / cos\phi_1$$

$$\Delta_P \leq \epsilon_4 \|P_k\| / cos\phi_1$$

$$\Delta_\vartheta \leq \epsilon_5 (\|\hat{\vartheta}_k\| + \|K_k\| \cdot \|y_k\|)$$

3. SILS

$$\Delta_T \leq \epsilon_6 \|T_k\| / cos\phi_2$$

$$\Delta_R \leq \epsilon_7 \|R_k\| / cos\phi_2$$

$$\Delta_P \leq \epsilon_8 \kappa(P_k) \|P_k\| / cos\phi_2$$

$$\Delta_\vartheta \leq \epsilon_9 [\kappa^2(T_k) |\varepsilon_k| + \kappa(T_k) \|\hat{\vartheta}_k\| + |\varepsilon_k| / cos\phi_3]$$

where $\kappa(.)$ denotes the condition number of the matrix $(.)$, ϵ_i are constants close to the machine precision ϵ, $cos\phi_i$ are defined as follows:

$$cos\phi_1 = \|S_k\| / \|[G_k \mid S_k]\|$$

$$cos\phi_2 = \|T_k\| / \left\| \begin{pmatrix} T_{k-1} \\ \varphi_k' / \sigma_k \end{pmatrix} \right\|$$

$$cos\phi_3 = |\varepsilon_k| / \left\| \begin{matrix} \dot{\xi}_k \\ \varepsilon_k \end{matrix} \right\|$$

which are usually close to one and ε_k is defined in equation (20).

Proof: The proof would be a repetition of the one given in Verhaegen and Van Dooren (1986) for the corresponding Kalman filter implementations, taking into account the special system matrices (equation 15). Therefore, the reader is referred to this paper.

3.3 Accumulation and Interaction of the Round-Off Errors

The third part of the analysis combines the local error upper bounds given in theorem 4 with the propagation of a single error, to yield bounds on the *total error* for the different RLS schemes. The total error is denoted by the prefix δ_{tot} instead of δ. Here we make the assumption that the total error at some instant k is the sum of the propagation of the previous errors (when introducing no additional round-off errors), plus the local errors made during these recursions.

For the CLS2, SCLS implementation the total error in the matrix P_k then satisfies:

$$\|\delta_{tot} P_k\| \leq \frac{1}{\lambda^{k-k_m}} \|\phi(k, k_m)\|^2 . \|\delta_{tot} P_{k_m}\| + \overline{\Delta_p^m} \tag{39}$$

In this equation k_m is the nearest time instant, smaller than k, for which $\phi(k, k_m)$ is a *contraction* and $\overline{\Delta_p^m}$ is given as:

$$\overline{\Delta_p^m} = \sum_{i=k_m}^{k} \epsilon_\ell \|P_i\| \tag{40}$$

This represents the accumulation of the local round-off errors made each recursion in the calculation of P_i in the interval (k_m, k). When P_i remains bounded, as imposed in condition 1 of corollary 4, $\overline{\Delta_p^m}$ also remains bounded.

In a similar way, the mean of the total error on the parameter estimates becomes:

$$\|\delta_{tot} \hat{\vartheta}_k\| \leq \|\phi(k, k_m)\| . \|\delta_{tot} \hat{\vartheta}_{k_m}\| + \overline{\Delta_\vartheta^m} \tag{41}$$

Both equations (39) and (41) model the propagation of the total round-off errors in the CLS2 and SCLS implementation by a linear model with bounded input signal.

Corollary 5: Error models (equations 39-41) indicate that when the regressor vector is persistently exciting, as defined in definition 1, the round-off errors made in the CLS2 and SCLS implementations remain bounded. Furthermore, parts 1 and 2 of theorem 4 show that the error models for both these implementations are identical. Therefore, only considering the numerical robustness no preference should be given to either of them. This latter condition generally does not hold for the CLS1 implementation since the "cures" to symmetrize P_k induce larger errors. This assertion has been confirmed in the experimental analysis made in Verhaegen and Van Dooren (1986) for Kalman filter implementations.

We now address the loss of positive definiteness under finite arithmetic precision. Using the assumption stated in this subsection to describe the total error on P_k, the following relationship results:

$$(P_k + \delta P_k) - \Delta P \le (P_k + \delta_{tot} P_k) \le (P_k + \delta P_k) + \Delta P \tag{42}$$

with ΔP according to theorem 4 bounded as $\|\Delta P\| \le \epsilon \|(P_k + \delta P_k)\|$.

According to theorem 3, the matrix $(P_k + \delta P_k)$ remains positive definite and we can arrange its eigenvalues as:

$$\lambda_1(P_k + \delta P_k) \ge \lambda_2(P_k + \delta P_k) \ge \cdots \ge \lambda_n(P_k + \delta P_k) > 0 \tag{43}$$

The possibility of negative eigenvalues of $(P_k + \delta_{tot} P_k)$ arises with the lower bound in equation (42). Insight in the effect of the perturbation ΔP on the eigenvalues of $(P_k + \delta P_k) - \Delta P$ is given is given by the following lemma, taken from Wilkinson, pp. 101-102 (1965).

Lemma 1: If A and $(A - E)$ are n by n symmetric matrices, then

$$\lambda_n(A - E) \ge \lambda_n(A) - \lambda_1(E)$$

Applying this lemma to the lower bound of equation (42), yields the following bound on the smallest eigenvalue of $(P_k + \delta_{tot} P_k)$:

$$\lambda_n(P_k + \delta_{tot} P_k) \ge \lambda_n(P_k + \delta P_k) - \epsilon \lambda_1(P_k + \delta P_k) \tag{44}$$

Therefore, this eigenvalue will be positive if

$$\frac{\lambda_1(P_k + \delta P_k)}{\lambda_n(P_k + \delta P_k)} \ge \frac{1}{\epsilon} \tag{45}$$

This condition indicates that $(P_k + \delta P_k)$ is numerically not singular, as defined in Wilkinson (1965).

Corollary 6: If the recursion of the parameter covariance matrix P_{k-1} in equation (10) under finite precision preserves the symmetry of this matrix and this matrix is numerically

not singular, as stated by equation (45), than the updating via equation (10) can not destroy the positive definiteness of the parameter covariance matrix.

From the computational scheme equation (20) of the SILS we clearly observe that it is not necessary to calculate the matrix P and the parameter estimate ϑ to continue recursions. Therefore, only the propagation of the total error on the information matrix R_k is of interest here. This model result from equation (29) and part 3 of theorem 4:

$$\|\delta_{tot}R_k\| \leq \lambda\|\delta_{tot}R_{k-1}\| + \epsilon_7\|R_{k-1}\| \tag{46}$$

When the observations are persistently exciting, R_k increases; therefore, the only thing of interest here are the relative errors. If these are denoted by δ_{tot}^r, equation (46) becomes

$$\|\delta_{tot}^r R_k\| \leq \gamma\lambda\|\delta_{tot}^r R_{k-1}\| + \epsilon_7 \tag{47}$$

where $\gamma = \|R_{k-1}\|/\|R_k\|$.

Corollary 7: When the observations are persistently exciting, the scalar γ remains smaller than 1. Therefore, the *linear relative* error model (equation 47) shows that the computations with "information" type of RLS implementations will remain exponentially stable for $\lambda \leq 1$. Each recursions, the quantities of interest P_{k-1} and $\hat{\vartheta}_{k-1}$ may be computed from T_{k-1} and $\hat{\xi}_{k-1}$. According to the bounds given in part 3 of theorem 4, these computations may be deteriorated by the condition number of P_{k-1} and T_{k-1}.

4 Experimental Evaluation

The purpose of this section is to experimentally validate the corollaries made in the theoretical error analysis of section 3.

To restrict such a verification study, we refer to a similar study made in Verhaegen and Van Dooren for Kalman filter implementations as an experimental verification of corollaries 2,3,5, and 7.

Of particular interest to the scientist dealing with RLS problems is the numerical robustness of the conventional RLS scheme, since this scheme is very attractive in a broad

class of applications. In the present analysis corollaries 1,4, and 5 focus on this topic. More precisely, these three corollaries provide insight into the following questions:

1. Whether a *"wrong"* way of implementing the conventional RLS has been the primary cause for its "high" sensitivity to the loss of symmetry (corollary 1)?

2. What is the influence of persistent excitation on the numerical robustness of the RLS schemes (corollary 4)?

3. What is the interference between the loss of symmetry and the loss of positive definiteness of P_k (corollary 6)?

To experimentally evaluate these three corollaries, a mixed-precision simulation study is performed. Here the single-precision quantities, as denoted by $\overline{(.)}$, represent the erroneous quantities, and the double-precision quantities are assumed to be error free. In this experimental analysis we will only focus on the errors on P_k and with the above convention the total error on this quantity, as denoted by δ_{tot} in equation (39), is approximated by:

$$\|\delta_{tot} P_k\| = \|\overline{P}_k - P_k\| \tag{48}$$

The regression model taken as a test vehicle in this analysis has the following form:

$$y_k = \sum_{i=1}^{4} \vartheta_i \varphi_k(i) \tag{49}$$

Two different time histories of $[y_k, \varphi_k(1), \cdots, \varphi_k(4)]$, for $k = 1, \cdots, 300$, are used in the following tests. Figure 1 displays these time histories, respectively, as measurement sequence 1 and 2.

The first test evaluates corollary 1. For this purpose use is made of measurement sequence 1, which guarantees the presistently excitation condition to be fulfilled during the whole test. Furthermore, an exponential forgetting factor $\lambda = 0.95$ was taken. The approximate error, equation (48), the loss of symmetry as computed by $\|\overline{P}_k - \overline{P}'_k\|$ and $\|\overline{P}_k\|$ are plotted for the CLS1 implementation in Fig. 2a and for the CLS2 implementation in Fig. 2b. These results clearly confirm corollary 1.

In the second test corollary 6 is evaluated. The same experimental condition as in the previous test are taken. Figures 3a and 3b, respectively, show the evolution of the

smallest eigenvalue and the third diagonal element of $\overline{P}_k$. Statistically these quantities have to be positive. In the error analysis, it has been reconfirmed that this remains true as long as $\overline{P}_k$ remains symmetric and numerically not singular. The first condition is easily violated by the CLS1 implementation; therefore, such a guarantee can not be stated for this implementation. This is confirmed in Fig. 3. Furthermore, such a loss of positive definiteness didn't occur with the CLS2 implementation for all the test performed in this simulation study. From Fig. 2b, we clearly see that the loss of symmetry is of different orders of magnitude smaller than the errors on $\overline{P}_k$. Therefore, the asymmetric part of ΔP in equation (42) can be neglected so that corollary 6 holds as well for the CLS2 as for the SCLS implementation.

The evaluation of corollary 4 in the third test makes use of the measurement sequence 2 in Fig. 1. From this figure we observe that the regressor vector φ_k is not persistently exciting during the time intervals $(1,40)$ and $(81,300)$. The impact of this on the round-off error propagation in the CLS2 implementation is pictured in Fig. 4. This figure clarifies that during the time intervals of lack of persistent excitation the error on $\overline{P}_k$ and the loss of symmetry increase "linearly." At the same time, $\|\overline{P}_k\|$ increases also linearly. Furthermore, the rate of increase is exactly the same in both cases, so that relatively no loss of precision occurs. Therefore, the practical impact of the lack of persistent excitation causing the so-called "burst" phenomenon, see Hughes and Jacobs (1974), is far more important than the numerical behavior of the considered RLS schemes which preserve the symmetry of P_k.

5 Concluding Remarks

In this paper, the numerical robustness of four different RLS schemes is analyzed by means of a theoretical error analysis. Apart from reconfirming the insights obtained in a similar analysis for Kalman filter implementations, given in Verhaegen and Van Dooren (1986). These are, respectively:

1. The conventional RLS (CLS) which preserves symmetry of the parameter error co-variance matrix P_k yields the same accuracy as the SCLS for all the corresponding quantities updated in both schemes.

2. The SILS is also numerically stable. The accuracy of computing P_k and the LS estimates is however penalized when the condition number of P_k is high. However, it is also not necessary to compute these quantities in order to continue the recursions with the SILS. Therefore, such large errors on P_k and the least squares estimates do not accumulate.

The new insights, of particular interest for applications with RLS, of this analysis are:

1. The CLS has been *"incorrectly"* implemented, making it very to the so-called loss of symmetry phenomenon. Correct implementation of this scheme overcomes this deficiency.

2. The *persistency of excitation* of the regressor vector in the regression model is required to guarantee the boundedness of the round-off errors. When this condition is violated, the errors on P_k will increase, however, with same rate of increase as $\|P_k\|$. Therefore, the increase of the errors is of minor importance.

3. The loss of positive definiteness is caused by the loss of symmetry

These conclusions have all been confirmed in an experimental verification study.

References

Åstrom, K.J., U. Borisson, L. Ljung, and B. Wittenmark (1977). Theory and applications of self-tuning regulators. *Automatica*, **13**, pp. 457-476.

Bierman, G.J. (1976). Factorization methods for discrete sequential estimation. *Academic Press*, New York.

Cybenko, G. (1980). The Numerical Stability of the Levison-Durbin algorithm for Toeplitz systems of equations. *SIAM J. Sci. Stat. Comput.*, **1**, 3, pp. 303-319.

Hägglund, T. (1983). New estimation techniques for adaptive control. *Ph.D. dissertation*, Lund Institute of Technology, Sweden.

Hughes, D.J. and O.L. Jacobs (1974). Turn-off, escape and probing in non-linear stochastic control. *IFAC Symp. on Stochastic control*, pp. 343-352.

Kulhavy, R. (1985). Restricted Exponential forgetting in Real-time identification. *Seventh IFAC Symposium on Syst. Identification and Parameter Estimation*, York, England.

Levinson, N. (1947). The Weiner RMS (root mean square) error criterion in filter design and prediction. *J. Math. Phys.*, **25**, pp. 261-278.

Ljung, L. and T. Söderström (1983). Theory and practice of recursive identification. *MIT Press*, Massachusetts.

Ljung, S. and L. Ljung (1985). Error propagation properties of recursive least squares adaptation algorithms. *Automatica*, **21**, 2, pp. 157-167.

MACSYMA Reference Manual (1983). The Mathlab Group Laboratory, Massachusetts Inst. of Tech., tenth ed.

Verhaegen, M.H. and P. Van Dooren (1986). Numerical aspects of different Kalman filter implementations. *IEEE Trans. Aut. Contr.*, AC-31, Vol. 10, pp. 907-917, Oct. 1986.

Verhaegen, M.H. (1985). A new class of algorithms in linear system theory, with application to real-time aerodynamic model identification. *Ph.D. dissertation*, Catholic University Leuven, Leuven, Belgium.

Wilkinson, J.H. (1965). The Algebraic Eigenvalue Problem. Clarendon Press, Oxford.

Figure Captions

Fig. 1: Two different measurement sequences for the variables (y_k, φ_u) in the regression model (1).

Fig. 2: Propagation of round-off errors in two implementations of the conventional RLS scheme ($\lambda = 0.95$, measurement sequence 1).

Fig. 3: Evolution of certain eigenvalues and diagonal elements of P_k in the CLS 1 implementation ($\lambda = 0.95$, measurement sequence 1).

Fig. 4: Influence off the persistency of excitation of the regressor vector $\varphi_(k)$ on the round-off propagation in CLS 2 implementation ($\lambda = 0.95$, measurement sequence 2).
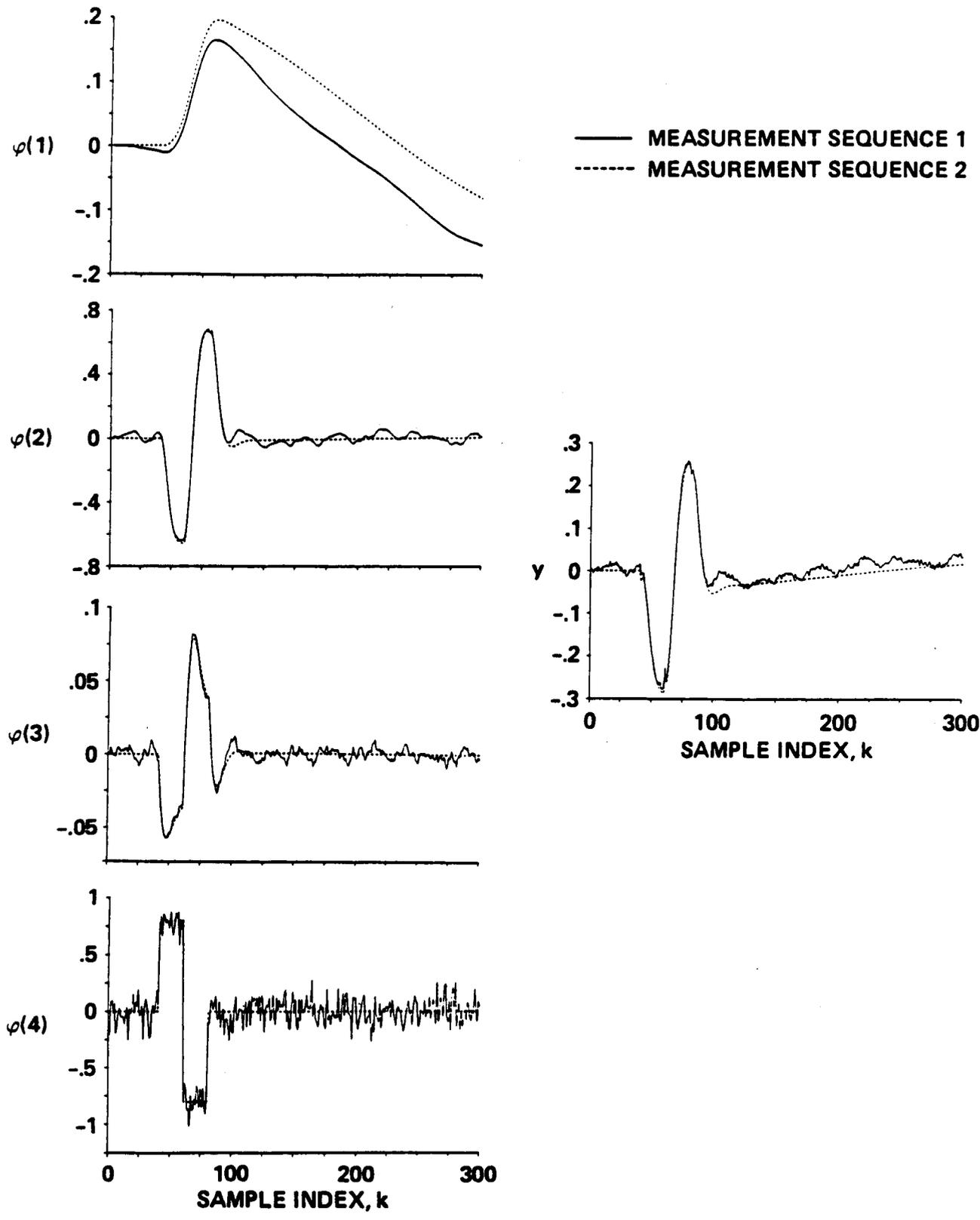
Fig. 1
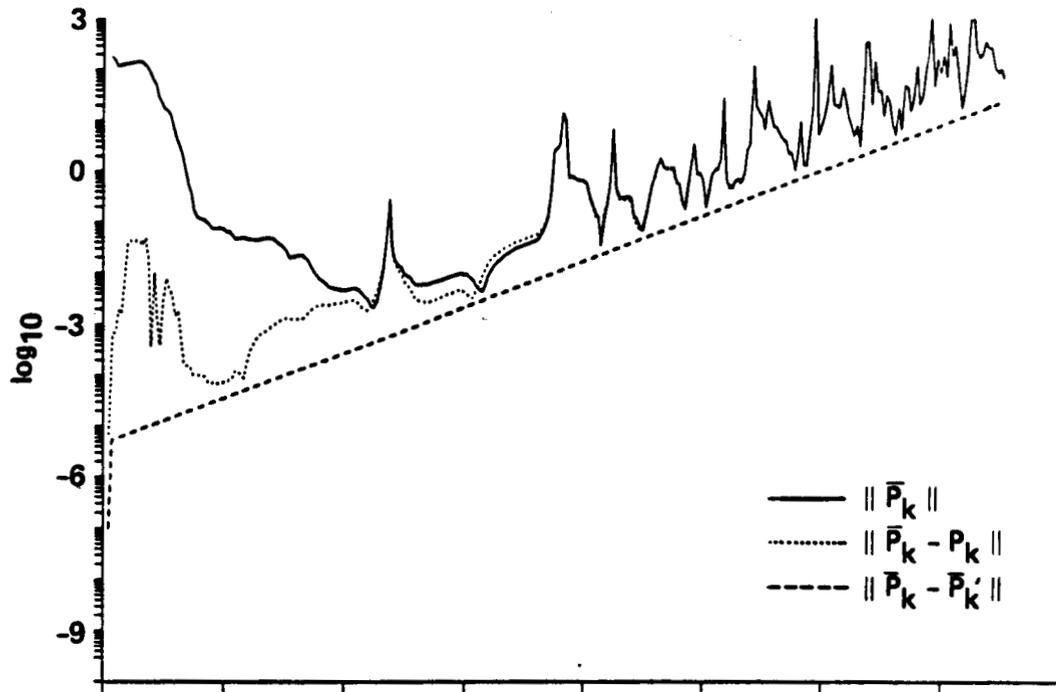
(a) THE CLS 1 IMPLEMENTATION
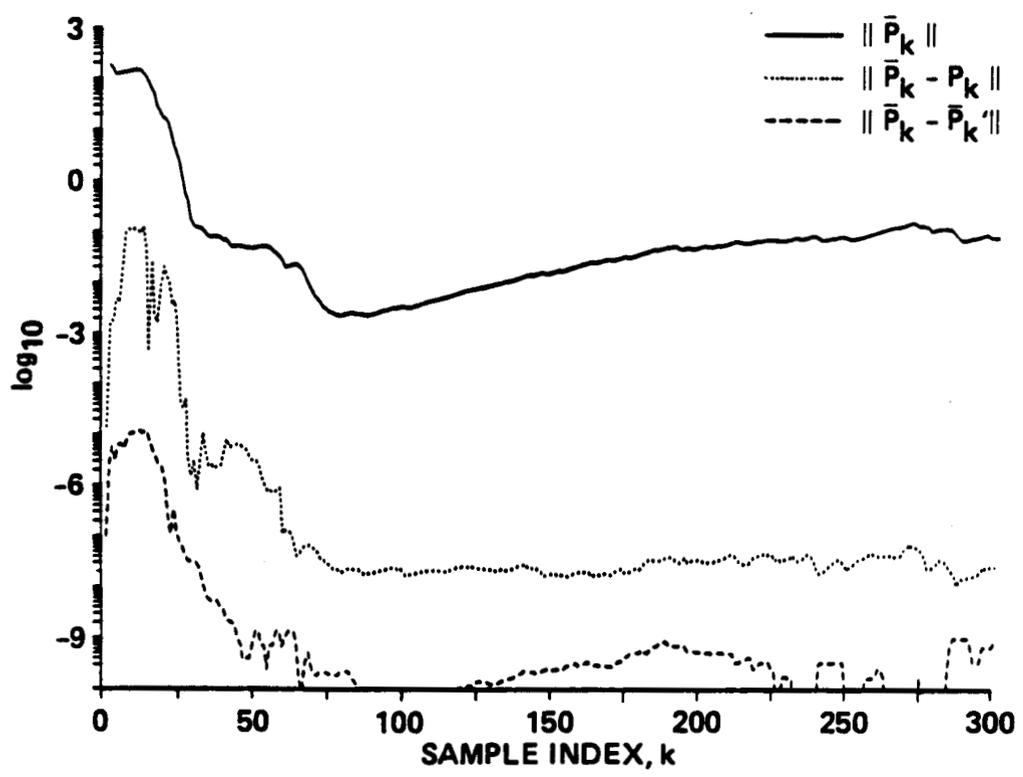
(b) THE CLS 2 IMPLEMENTATION

Fig. 2

22
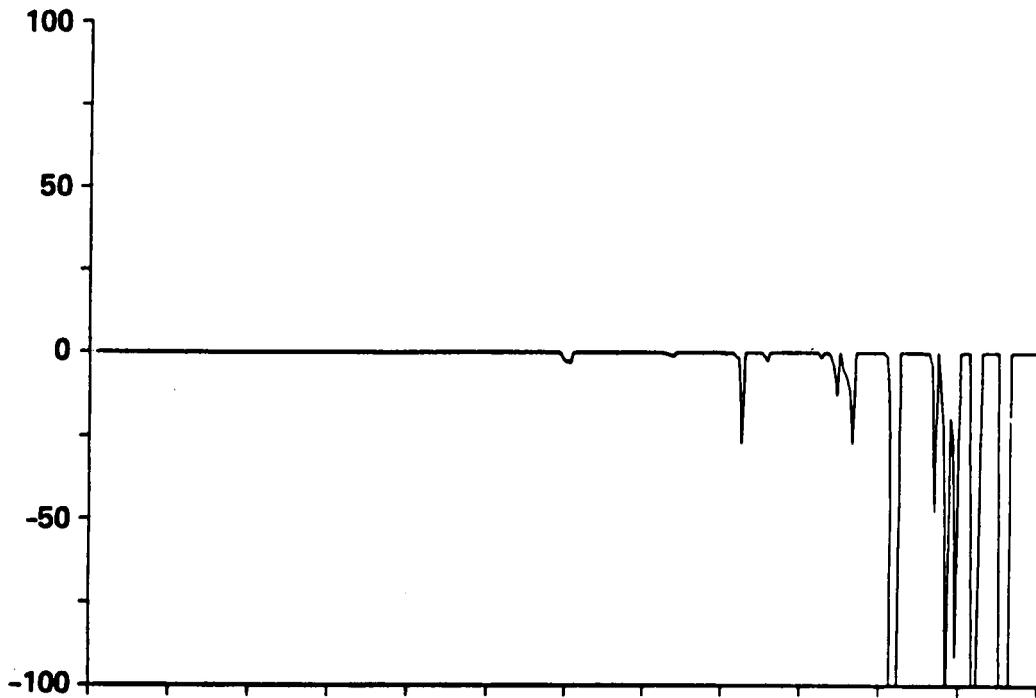
(a) EVOLUTION OF SMALLEST EIGEN VALUE OF $\bar{P}_k$
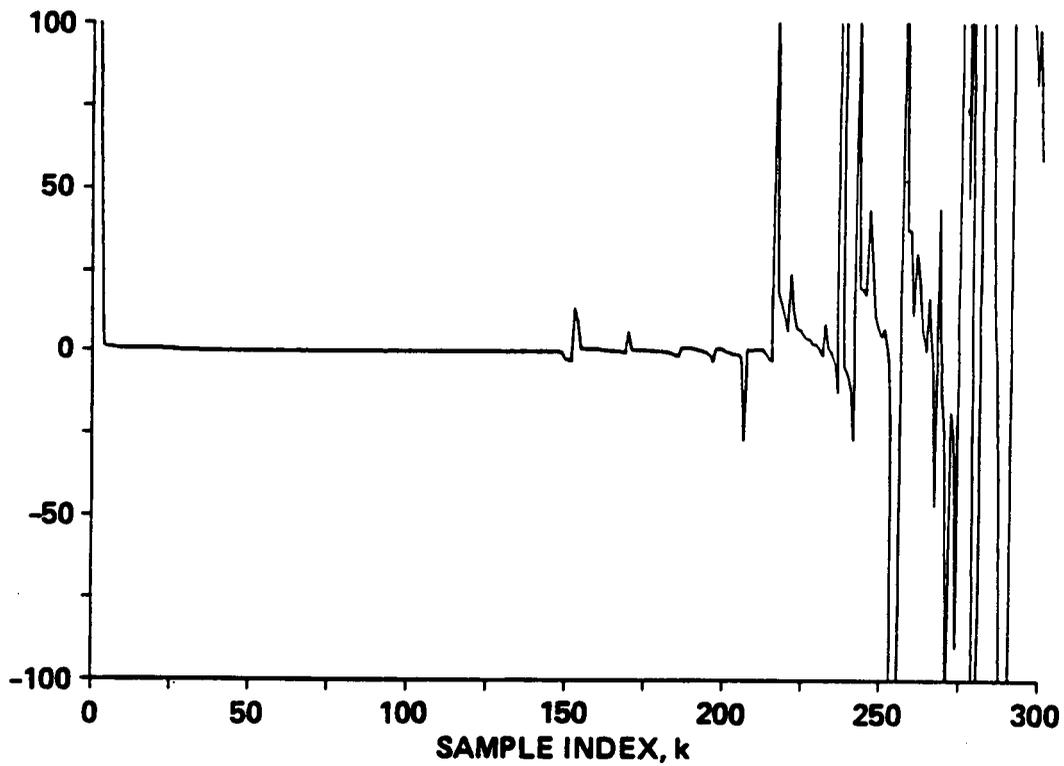
(b) EVOLUTION OF DIAGONAL ELEMENT 3 OF $\bar{P}_k$

Fig. 3
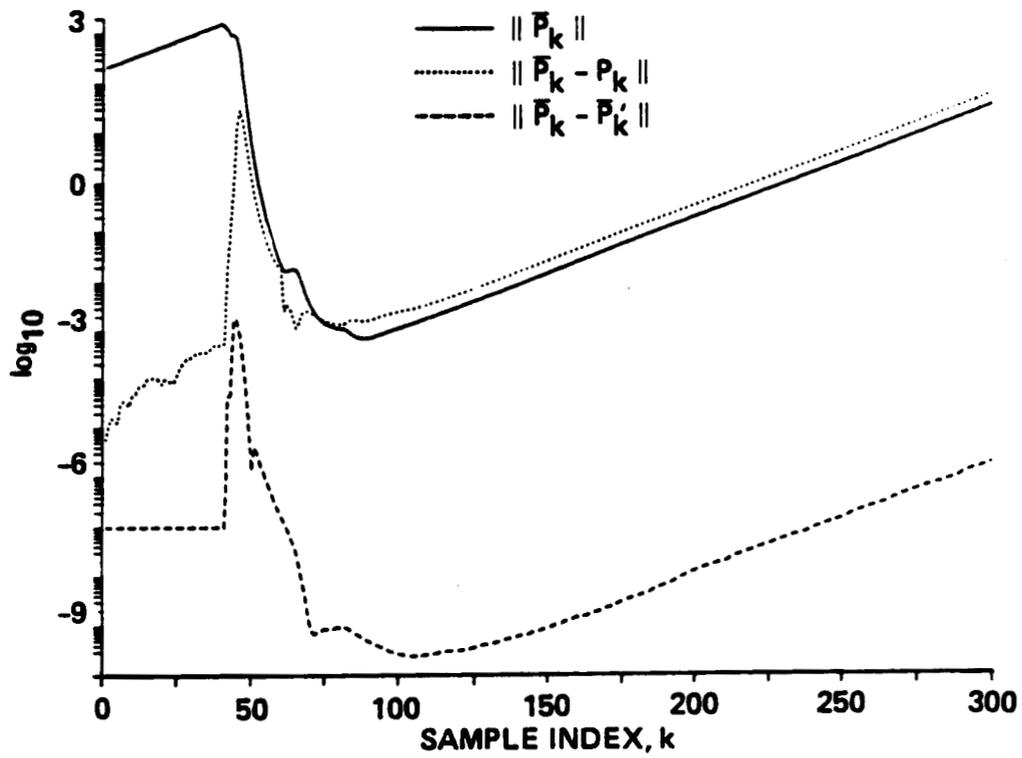
Fig. 4

Report Documentation Page

16. Abstract

In this paper the numerical robustness of four generally applicable, recursive, least-squares-estimation schemes is analyzed by means of a theoretical round-off propagation study. This study highlights a number of practical, interesting insights of widely used recursive least-squares schemes. These insights have been confirmed in an experimental verification study as well.