Proceedings of the 17th World Congress
The International Federation of Automatic Control
Seoul, Korea, July 6-11, 2008

Recursive parameter estimation by means of the SG-algorithm ⋆

Magnus Evestedt * **Alexander Medvedev** *

* *Department of Information Technology, Uppsala University, P. O. Box 337, SE-751 05, SWEDEN*
magnus.evestedt@it.uu.se; alexander.medvedev@it.uu.se

Abstract: Recursive parameter estimation in linear regression models by means of the Stenlund-Gustafsson algorithm is considered. The manifold of stationary solutions to the parameter update equation is parameterized in terms of excitation properties. It is shown that the parameter estimation error vector does not diverge under lack of excitation, therefore achieving the purpose of anti-windup. Furthermore, an elementwise form of the parameter vector estimate is suggested revealing the effect of individual matrix entries in the Riccati equation on the parameter estimation updates. Simulations are performed to illustrate the loss of convergence rate in the estimates versus the decrease of computational power needed for two specific approximations of the Riccati equation in the elementwise form.

Keywords: Recursive identification, Parameter estimation, Kalman filter

1. INTRODUCTION

The task of a recursive parameter estimation algorithm is to track dynamical properties of signals and systems. A number of methods have been suggested in the past and the most common ones are the recursive least squares with forgetting factor (RLS) and normalized least mean squares (N-LMS). These are shown in Ljung and Gunnarsson [1990] to be *ad hoc* variants of the optimal Kalman filter algorithm.

A significant complication in the practical use of the Kalman filter in parameter estimation is its sensitivity to the system input excitation. If the input does not provide sufficient information for parameter identification, a phenomenon called covariance windup occurs. *Ad hoc* solutions to this issue have been proposed *e.g.* in Bittanti et al. [1990], Hägglund [1983]. An interesting specialization of the Riccati equation with anti-windup properties, in the sequel referred to as the Stenlund-Gustafsson (SG) algorithm, had been suggested in Stenlund and Gustafsson [2002]. Later, a variation of this idea was presented in Cao and Schwartz [2004].

The special form of the Riccati equation in the SG-algorithm has made it possible to show its non-divergence under lack of excitation in Medvedev [2004], completely parameterize its stationary behaviour in Evestedt and Medvedev [2006a] and also enabled elementwise decoupling and convergence analysis in Medvedev and Evestedt [2008]. However, the properties of the parameter estimation equation of the SG-algorithm have not been addressed so far. Now, the fact that both the Riccati equation and the parameter estimate of the SG algorithm are governed by a certain matrix operator (elementary matrix transformation), facilitates the analysis since some of the results from the Riccati equation analysis carry over to the parameter estimation equation.

The advantageous anti-windup properties of the SG-algorithm make it a good candidate for engineering applications such as active vibration control Olsson [2005], acoustic echo cancellation Evestedt et al. [2005] and change detection Evestedt and Medvedev [2006b].

In this paper, the close relationship between the recursive parameter estimation and the Riccati equation of the SG-algorithm is exploited to formulate new results concerning stationary solutions, convergence properties and elementwise decoupling of the parameter estimate. The elementwise decoupling property described in Medvedev and Evestedt [2008] is utilized to lower the computational power demands of the SG-algorithm for the case of periodic regressor and is illustrated by simulation.

2. PRELIMINARIES

The focus of this paper is on linear regression models of the following type

$$y(t) = \varphi^T(t)\theta + e(t) \tag{1}$$

where $y(t)$ is the scalar output measured at discrete time instances $t = [0, \infty)$, $\varphi \in R^n$ is the regressor vector, $\theta \in R^n$ is the parameter vector to be estimated and the scalar e is the disturbance.

If $e(t)$ is white and the parameter vector is subject to the random walk model driven by a zero-mean white sequence $w(t)$

$$\theta(t) = \theta(t-1) + w(t)$$

then the optimal, in the sense of minimum of the *a posteriori* parameter error covariance matrix, estimate is yielded by

$$\begin{aligned}\hat{\theta}(t) &= \hat{\theta}(t-1) + K(t)\left(y(t) - \varphi^T(t)\hat{\theta}(t-1)\right) \\ &= \left(I - K(t)\varphi^T(t)\right)\hat{\theta}(t-1) + K(t)y(t) \end{aligned} \tag{2}$$

with the Kalman gain

$$K(t) = \frac{P(t-1)\varphi(t)}{r(t) + \varphi^T(t)P(t-1)\varphi(t)} \tag{3}$$

and $P(t)$, $t = [1, \infty)$, the solution to the Riccati equation

⋆ This work has been carried out with financial support by Swedish Research Council.

$$P(t) = P(t-1) - \frac{P(t-1)\varphi(t)\varphi^T(t)P(t-1)}{r(t) + \varphi^T(t)P(t-1)\varphi(t)} + Q(t) \quad (4)$$

for some $P(0) = P^T(0)$, $P(0) \geq 0$ describing the covariance of the initial guess of $\hat{\theta}(t)$, $t = 0$. Optimality of the estimate is guaranteed only when Q is the covariance matrix of w and $r(t) = \text{var}\ e(t)$, Ljung and Gunnarsson [1990].

Introducing the matrix-valued matrix function

$$A_t(X) = I + r^{-1}(t)X\varphi(t)\varphi^T(t)$$

(2) can be rewritten as

$$\hat{\theta}(t) = A_t^{-1}(P(t-1))\hat{\theta}(t-1) + K(t)y(t) \quad (5)$$

Apparently, the geometrical properties of the function $A_t(\cdot)$ are of fundamental importance for the dynamics of the recursive identification algorithm.

The regressor vector sequence is called persistently exciting Söderström and Stoica [1989], if there exists a $c \in R^+$ and integer $m > 0$ such that for all t

$$cI \leq \sum_{k=t}^{t+m} \varphi(k)\varphi^T(k) \quad (6)$$

This condition is important for the dynamic behavior of (4) as when it is not satisfied, some eigenvalues of $P(\cdot)$ increase linearly. This is usually referred to as covariance windup.

2.1 The SG-algorithm

In the approach taken in Stenlund and Gustafsson [2002], a special choice of $Q(t)$ is suggested to deal with the windup problem and to control the convergence point of the solution to (4)

$$Q(t) = \frac{P_d\varphi(t)\varphi^T(t)P_d}{r(t) + \varphi^T(t)P_d\varphi(t)}$$

where $P_d \in R^{n\times n}$, $P_d > 0$.

The difference $E(t) = P(t) - P_d$ is shown, in Stenlund and Gustafsson [2002] for non-singular $P(t)$ and in Evestedt and Medvedev [2006a] for a general case, to obey the recursion

$$E(t+1) = A_t^{-1}(P(t))E(t)A_t^{-1}(P_d) \quad (7)$$

or, in a vectorized form

$$\text{vec}\ E(t+1) = M(P_d, P(t))\text{vec}\ E(t) \quad (8)$$

$$M(P_d, P(t)) = A_t^{-1}(P(t-1)) \otimes A_t^{-1}(P_d) \quad (9)$$

where $\otimes$ denotes Kronecker product. Thus the matrix equation in the SG-algorithm can be rewritten as a linear discrete time-varying system which form facilitates its analysis. In Medvedev [2004], this structure is utilized to show non-divergence of the algorithm and in Evestedt and Medvedev [2006a] to examine its stationary properties under lack of excitation.

The fact that the matrix function $A_t(X)$ appears both in the Riccati equation of the SG-algorithm and in the parameter estimate is striking. It was a key to the analysis of the Riccati equation and will in the sequel also be proved to facilitate the analysis of the parameter update equation.

3. STATIONARY POINTS

Using regression model (1), estimator (2) can be written as

$$\hat{\theta}(t) = \hat{\theta}(t-1) + K(t)(e(t) + \varphi^T(t)(\theta - \hat{\theta}(t-1))) \quad (10)$$

Subtracting θ from both sides of the equation and defining

$$\tilde{\theta}(t) \triangleq \hat{\theta}(t) - \theta$$

yields

$$\begin{aligned}\tilde{\theta}(t) &= \tilde{\theta}(t-1) + K(t)(e(t) - \varphi^T(t)\tilde{\theta}(t-1)) \\ &= A_t^{-1}(P(t-1))\tilde{\theta}(t-1)) + K(t)e(t) \quad (11)\end{aligned}$$

In order to separate the direction of excitation at each particular time instant from the excitation intensity, introduce a re-parametrization of the matrix function $A_t(X)$

$$A_t(X) = I + \rho XU(t)$$

where $\rho(t) = r^{-1}(t)\varphi(t)^T\varphi(t)$ and

$$U(t) = \frac{\varphi(t)\varphi^T(t)}{\varphi^T(t)\varphi(t)}$$

The matrix $U(t)$ is a Hermitian projection with rank $U(t) = 1$. Define the normalized eigenvectors of $U(t)$ as $\xi_i(t)$, $i = 1, \ldots, n$, where $\xi_1(t)$ corresponds to the unit eigenvalue of $U(t)$ and $\xi_2(t), \ldots, \xi_n(t)$ correspond to the zero eigenvalues of $U(t)$. Then $\rho(t)$ describes the energy in the regressor vector at time t and $\xi_1(t)$ characterizes the direction.

Excitation is called sufficient at time t when the following rank condition is satisfied

$$\text{rank}\ [\xi_1(t+n-1)\ \ldots\ \xi_1(t)] = n \quad (12)$$

which is a stricter condition than persistent excitation since it demands that each sequence of n consequent regressor vectors is linearly independent.

Now, assuming $e(t) = 0$, consider a stationary point of (10) *i.e.* $\hat{\theta}(t) = \hat{\theta}(t-1) = \hat{\theta}^* = const$. The proposition below characterizes the space of all possible stationary solutions.

Proposition 1. For $e(t) = 0$, any stationary solution of (10) can be decomposed as

$$\hat{\theta}^* = \theta + \sum_{k=2}^{n} m_k\xi_k \quad (13)$$

for some scalars $m_2, \ldots, m_n$.

Proof: Omitted.

Notice that the result above is valid even when condition (6) for persistent excitation is not satisfied. The actual stationary solution is dependent on the current excitation properties of the regressor vector. For the case of persistent excitation it follows that $\hat{\theta}^* = \theta$.

4. DYNAMICS OF THE PARAMETER ESTIMATE

The similarities between (5) and (7) suggest the use of the Lyapunov transformation, earlier utilized in Medvedev [2004] for analysis of the Riccati matrix equation, in order to bring the parameter estimation equation to a structure revealing form.

4.1 Lyapunov Transformation

Suppose that at each $\tau = 1, n+1, 2n+1, \ldots$, the sequence $\{U(t)\}$ is known n steps in advance. If the sequence is sufficiently exciting on each interval of n consecutive steps, a matrix $T(\tau)$ can be defined as follows

$$\det T(\tau) \neq 0\ T(\tau) = [\xi_1(\tau)\ \ldots\ \xi_1(\tau+n-1)]$$

If, however, the sequence is not sufficiently exciting such that the set $\{\xi_1(t), t = \tau \ldots, \tau+n-1\}$ includes only $k < n$ linearly independent vectors, the matrix T must be constructed differently. Let the matrix $T_k(\tau)$ consist of the k linearly independent $\xi_1(t)$ on $t \in [\tau, \dots, \tau + n - 1]$ and μ^i, $i = k+1, \dots, n$ form an orthonormal basis of the left nullspace of T_k. The transformation then has the following form

$$T(\tau) = [T_k(\tau)\ \mu^{k+1} \dots \mu^n]$$

The matrix $T(\tau)$ is a Lyapunov transformation and preserves the stability properties of the original dynamic system. The columns of the transformation matrix are denoted as

$$T(t) = [\xi_1^1(t)\ \dots\ \xi_1^n(t)]$$

with time variable dropped when appropriate to save space.

Introduce the state matrix

$$Z(t) = T^T(t)E(t)T(t) \tag{14}$$

and denote the elements of $Z(t)$ by

$$Z(t) = \{z_{kl}, k = 1, \dots, n; l = 1, \dots, n\}$$

Since $E(t)$ is a symmetric matrix, the transformed matrix $Z(t)$ is also symmetric.

Let the column position in $T(t)$ of the current direction of excitation, *i.e.* $\xi_1^i(t)$ be denoted by i. Then, for some $X \in R^{n\times n}$, consider $X \geq 0$ the vectors

$$d_i^T(X) = \left[{\xi_1^1}^T X \xi_1^i \dots {\xi_1^n}^T X \xi_1^i\right]$$
$$D_i^T(X) = [D_i^1(X)\ \dots\ D_i^n(X)]$$

which are related to each other as

$$D_i(X) = \frac{\rho_i d_i(X)}{1 + \rho_i {\xi_1^i}^T X \xi_1^i}$$

It is shown in Medvedev and Evestedt [2008] that if $P(t)$ is a solution to (7), then

$$D_i^k(P(t)) = \frac{\rho_i \left({\xi_1^k}^T P_d \xi_1^i + z_{ik}(t)\right)}{1 + \rho_i \left({\xi_1^k}^T P_d \xi_1^i + z_{ii}(t)\right)} \tag{15}$$

After the transformation $\bar{\theta}(t) = T^T\hat{\theta}(t)$ of linear time-varying discrete system (2), the transformed system matrix becomes $\bar{A}_t^{-1}(\cdot) = T^T A_t^{-1}(\cdot) T^{-T}$. The system matrix is calculated in Medvedev [2004] to be

$$\bar{A}_i^{-1}(X) = I - [0_{n\times(i-1)}\ D_i(X)\ 0_{n\times(n-i)}] \tag{16}$$

Thus, at each step $t = 1, \dots, n$, $\bar{A}_t^{-1}(X)$ is the sum of a unit matrix and a matrix whose nonzero elements are all in one column. The position i of the column vector $D_i(X)$ in (16) is defined by the current excitation direction ξ_1^i.

In the sequel, the notion of vector element in the direction of excitation comes in handy. Consider the vector $x(t) = [x_1(t)\ x_2(t)\ \dots\ x_n(t)]^T$, recursively updated according to

$$x(t) = \bar{A}_t^{-1}(\cdot)x(t-1) = x(t-1) - x_i(t-1)D_i(\cdot) \tag{17}$$

This means that at each step, the vector $x(t)$ is updated by the vector $D_i(\cdot)$ weighted by the vector element $x_i(t-1)$. The particular element i is defined by the current direction of excitation and thus, the scalar $x_i(t-1)$ is defined as the vector element in the direction of excitation.

4.2 Non-divergence of Parameter Estimate

For a parameter estimation algorithm to perform well when excitation lacks in some directions, it is desirable that the estimation error does not diverge. The following propositions show, utilizing the transformation matrix $T(t)$, that this is the case for the SG-algorithm as well as for any algorithm of the form (2), (3).

Proposition 2. In the ideal case of perfect measurements, *i.e.* $e(t) = 0$, the parameter estimation error, $\tilde{\theta}$ is non-diverging.

Proof: Omitted.

The next proposition shows that the increase in the parameter error vector elements outside the current excitation direction is bounded at each step and the value of the bound is defined by the element of the parameter estimation error in the current direction of excitation.

Proposition 3. Define the transformed parameter error vector $v(t) = T^T(t)\tilde{\theta}(t)$. If $e(t) = 0$, for each element v_k of $v(t)$, $k \neq i$, the following inequality applies

$$|v_k(t+1) - v_k(t)| \leq |v_i(t)| \tag{18}$$

Proof: Omitted.

5. ELEMENTWISE FORM

In (11), it is not clear how the individual elements of the matrix P in the Riccati equation of the SG-algorithm affect the parameter updates. This is clarified by the following proposition.

Introduce a new parameter state vector as $\bar{\theta}(t) = T^T\hat{\theta}(t)$.

Proposition 4. The elementwise parameter update equation can be written as

$$\bar{\theta}_k(t) = \bar{\theta}_k(t-1) - \frac{\rho_i \left({\xi_1^k}^T P_d \xi_1^i + z_{ik}(t-1)\right)}{1 + \rho_i \left({\xi_1^k}^T P_d \xi_1^i + z_{ii}(t-1)\right)} \tag{19}$$
$$\times \left(\bar{\theta}_i(t-1) + \frac{1}{||\varphi(t)||_2} y(t)\right)$$

or, for the special case of $k = i$

$$\bar{\theta}_i(t) = \frac{\bar{\theta}_i(t-1)}{1 + \rho_i \left({\xi_1^i}^T P_d \xi_1^i + z_{ii}(t-1)\right)}$$
$$+ \frac{1}{||\varphi||_2} \frac{\rho_i \left({\xi_1^i}^T P_d \xi_1^i + z_{ii}(t-1)\right)}{1 + \rho_i \left({\xi_1^i}^T P_d \xi_1^i + z_{ii}(t-1)\right)} y(t)$$

Proof: See Appendix A

As can be seen above, for each time instant, the dynamics of each individual parameter update is only dependent upon two elements of the transformed Riccati equation, z_{ii} and z_{ik}, in the row corresponding to the current direction of excitation.

5.1 Comparison to the Normalized Least Mean Squares (N-LMS)

Consider the case when $P(0) = P_d$ in the SG-algorithm. Then the gain matrix in (2) becomes

$$K(t) = \frac{P_d\varphi(t)}{r(t) + \varphi^T(t) P_d \varphi(t)}$$

In Ljung and Gunnarsson [1990], it is shown that the N-LMS is a special case of the Kalman filter with the choices $P_d = \mu I$, $\mu \in R^+$ and $r(t) = 1$. A general choice of P_d yields a matrix step-size N-LMS algorithm.

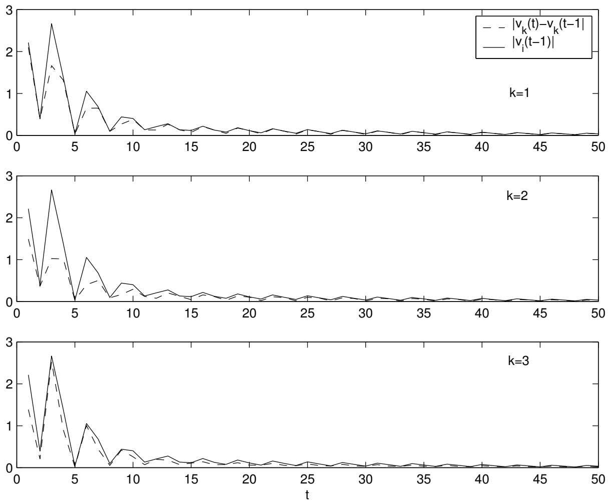

Fig. 1. Bound (18) on the change of the estimated parameter vector elements.

Robustness conditions of such an algorithm are given in Rupp and Cezanne [2000]. Ways of choosing the gain matrix are suggested in Mokino [1993], Gay [1998]. In (19), it is clear that P_d works as a weighting matrix of the regressor directions affecting the updates of the individual parameter estimate vector elements.

6. SIMULATION

In this section, simulations are performed to illustrate the possibilities of decreasing the computational complexity of recursive parameter estimation by employing the elementwise calculations. Two types of approximations are made and analyzed in terms of performance loss compared to the full SG-algorithm.

6.1 Periodic Excitation

In order to facilitate the implementation of the elementwise SG-algorithm, the special case of periodic regressor vectors is considered. This assumption is often made in the literature, Ramos et al. [2007], Akçay and At [2006] and is reasonable in some engineering applications. Here the number of estimated parameters equals the input period as in Akçay and At [2006]. For the SG-algorithm, it means that the transformation T is a constant matrix.

6.2 Bound on Parameter Error Increase

The bound stated in Proposition 3 is illustrated in Fig. 1 for a simulation of a system with $n = 3$. The subplots correspond to different elements in the parameter error vector. As can be seen, both the bound and the estimates converge at what seems to be exponential rate.

6.3 Band Matrix Approximation

According to Medvedev and Evestedt [2008], transformed Riccati equation (14) can be updated elementwise as

$$z_{kl}(t+1) = z_{kl}(t) + \rho_i \left(\frac{{\xi_l^i}^T P_d \xi_1^i {\xi_1^k}^T P_d \xi_1^i}{1 + \rho_i {\xi_1^i}^T P_d \xi_1^i} - \frac{({\xi_l^i}^T P_d \xi_1^i + z_{il}(t))({\xi_1^k}^T P_d \xi_1^i + z_{ik}(t))}{1 + \rho_i {\xi_1^i}^T P_d \xi_1^i + z_{ii}(t)} \right) \quad (20)$$

or, for the special case of $l = i$

$$z_{ki}(t+1) = \frac{(1 + \rho_i {\xi_1^i}^T P_d \xi_1^i) z_{ki}(t) - \rho_i {\xi_1^k}^T P_d \xi_1^i z_{ii}(t)}{\left(1 + \rho_i {\xi_1^i}^T P_d \xi_1^i\right)\left(1 + \rho_i ({\xi_1^i}^T P_d \xi_1^i + z_{ii}(t))\right)} \quad (21)$$

which reads for $k = l = i$ as

$$z_{ii}(t+1) = \frac{z_{ii}(t)}{\left(1 + \rho_i {\xi_1^i}^T P_d \xi_1^i\right)\left(1 + \rho_i ({\xi_1^i}^T P_d \xi_1^i + z_{ii}(t))\right)} \quad (22)$$

The equations above allow to choose which elements of the Riccati equation to employ in the parameter estimation algorithm and it is tempting to exploit this possibility in order to decrease its computational complexity.

One obvious approach is to only use a limited number of super- and subdiagonals in the matrix Z for updating the parameter estimates. These elements are guaranteed to be non-diverging since they obey the same equations as the full SG-algorithm. All other elements in Z are considered to be zero, *i.e.* the corresponding elements of P are assumed to already have converged to the elements of P_d at these positions, making Z a band matrix.

$$Z_q = \begin{bmatrix} z_{11} & z_{12} & \cdots & z_{1q} & 0 & & \cdots & 0 \\ z_{21} & z_{22} & z_{23} & \cdots & z_{2(q+1)} & 0 & \cdots & 0 \\ & z_{32} & z_{33} & & & \ddots & \ddots & \vdots \\ \vdots & & & \ddots & & & & 0 \\ z_{q1} & \vdots & & & & & & z_{(n-q+1)n} \\ 0 & z_{(q+1)2} & & & & & & z_{(n-q+2)n} \\ & 0 & \ddots & & & & & \vdots \\ \vdots & \vdots & \ddots & & & & \ddots & \\ 0 & 0 & \cdots & 0 & z_{n(n-q+1)} & z_{n(n-q+2)} & \cdots & z_{nn} \end{bmatrix} \quad (23)$$

The selected z_{kl} can either be computed online or, if the regressor sequence is known, in advance. The procedure for updating the parameter vector between $t = \tau + 1 \ldots \tau + n$ is the following.

Algorithm 1

(1) Select q as the number of required diagonals.
(2) Let $t = \tau + 1$.
(3) Transform the parameter vector by $\bar{\theta}(t) = T^T \hat{\theta}(t)$.
(4) Update the parameter estimates according to (19).
(5) Update the Riccati equation elements according to (20) and (23).
(6) Increase $t = t + 1$.
(7) if $t \leq \tau + n$ goto (3).
(8) Transform the parameter vector by $\hat{\theta}(\tau + n) = T^{-T} \bar{\theta}(\tau + n)$.

This procedure is similar to that of AKFA (Average Kalman Filter Algorithm) in Wigren [1998], where the adaptation gains in the Kalman filter are the solution of an averaged diagonal Riccati equation, breaking the latter down to a small number of scalar equations. Here it is however possible to add super- and subdiagonals to improve the estimation performance. An extensive comparison of the SG-algorithm to other parameter estimation methods is given in Evestedt et al. [2005].

A 3-dimensional system was simulated with 3-periodic input signal $u(t)$. The output signal was constructed

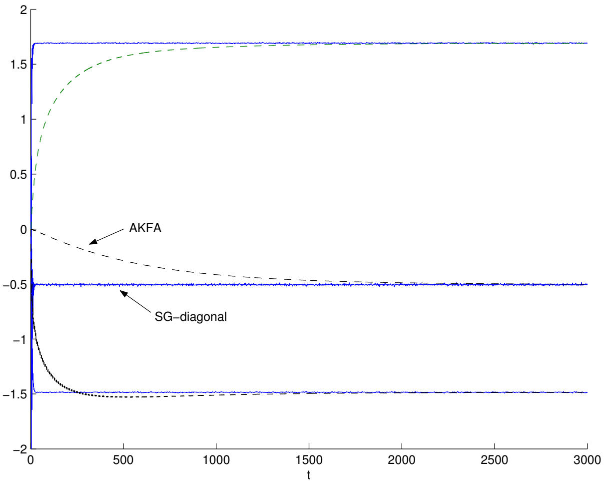

Fig. 2. Comparison between AKFA and the diagonal SG-algorithm with $q = 1$.

with (1), $\varphi(t) = [u(t-1)\ u(t-2)\ u(t-3)]$ and white measurement noise (SNR= $15dB$). The matrices P_d and $P(0)$ in the SG-algorithm were selected randomly and no tuning other than that described in Wigren [1998] was performed for the AKFA. In Fig. 2, it can be seen that both estimation algorithms reach the true parameter vector values but the diagonal element ($q = 1$) SG-algorithm converges faster at the cost of larger estimate variance. In general, AKFA seems to be more sensitive to the actual excitation directions used in the construction of the regressor vectors. The performance of the SG-algorithm is on the other hand highly dependent on the choice of P_d in the same way as the Kalman filter algorithm is dependent on r and Q.

The main difference between employing the full SG-algorithm Riccati equation and the super- subdiagonal ones lies in the transient period. When P has converged to P_d, the transformed matrix Z becomes zero which gives the same parameter estimation updates in both cases. In Fig. 3 illustrates this for the main diagonal SG-algorithm. By increasing the number of diagonals used, the deviation from the full SG-algorithm is decreased. For a higher dimensional system, a small difference persists over a longer period of time.

6.4 Small Matrix Elements Approximation

Yet another approximation of the SG-algorithm is provided in Medvedev and Evestedt [2008]. There it is argued that (22) implies fast convergence of the diagonal elements of Z in the direction of excitation. Assuming small z_{ii} is thus justified and an approximation of (22) is given as

$$z_{ii}(t+1) = \frac{z_{ii}(t)}{\left(1 + \rho_i {\xi_1^i}^T P_d \xi_1^i\right)^2}$$

and for the off-diagonal elements of Z in the direction of excitation

$$z_{ki}(t+1) = \frac{z_{ki}(t)}{1 + \rho_i {\xi_1^i}^T P_d \xi_1^i}$$

This means that also z_{ki}, $k = 1, \ldots, n$, $k \neq i$ are small. Then (21) becomes

$$z_{kl}(t+1) = z_{kl}(t), \ k \neq i, \ l \neq i$$

The convergence of Z to zero is guaranteed by the nature of the algorithm.

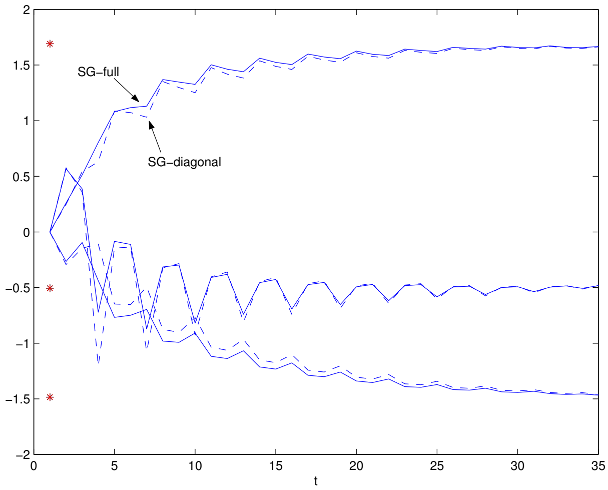

Fig. 3. Comparison between the parameter estimates using the full SG-algorithm and only the main diagonal, $q = 1$. Asterisks give the true values of the parameters.

Combining the above approximations into matrix form yields

$$Z(t+1) = \tag{24}$$

$$\begin{bmatrix} z_{11}(t) & z_{12}(t) & \cdots & \dfrac{z_{1i}(t)}{1+\rho_i {\xi_1^i}^T P_d \xi_1^i} & \cdots & z_{1n}(t) \\ z_{21}(t) & z_{22}(t) & z_{23}(t) & & & \vdots \\ & z_{32}(t) & z_{33}(t) & & & \\ \vdots & \vdots & & \ddots & & \\ & & & \dfrac{z_{ii}(t)}{\left(1+\rho_i {\xi_1^i}^T P_d \xi_1^i\right)^2} & & \\ & & & \vdots & & \\ z_{n1}(t) & \cdots & & \dfrac{z_{ni}(t)}{1+\rho_i {\xi_1^i}^T P_d \xi_1^i} & \cdots & z_{nn}(t) \end{bmatrix}$$

The adaptation gains can be calculated beforehand if ρ_i and P_d are known. The algorithm becomes

Algorithm 2

(1) Let $t = \tau + 1$.
(2) Transform the parameter vector by $\bar{\theta}(t) = T^T \hat{\theta}(t)$.
(3) Update the parameter estimates according to (19).
(4) Update the Riccati equation elements according to (20) and (24).
(5) Increase $t = t + 1$.
(6) if $t \leq \tau + n$ goto (2).
(7) Transform the parameter vector by $\hat{\theta}(\tau + n) = T^{-T} \bar{\theta}(\tau + n)$.

The transient employing this procedure is shown in Fig. 4 for the 3-dimensional system described above.

7. CONCLUSIONS

The parameter estimation by means of the SG-algorithm is studied for sufficiently and insufficiently exciting regressor vector sequences. In absence of measurement disturbance, the stationary solution to the parameter update equation is shown to belong to a manifold defined by the properties of the regressor vector sequence. The parameter error vector is proved to be non-divergent under lack of excitation.

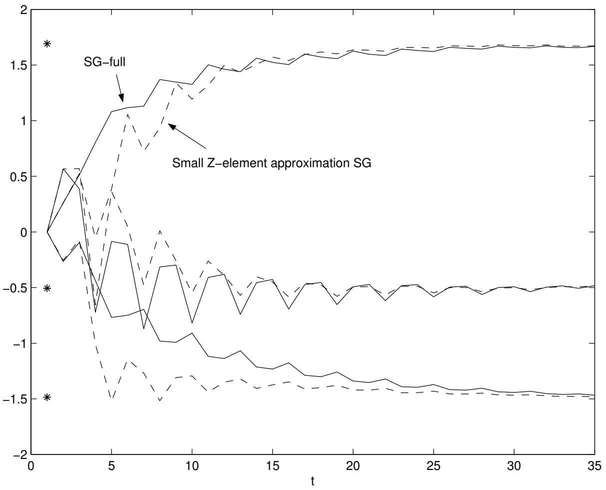

Fig. 4. Comparison between the parameter estimates using the full SG-algorithm and the small matrix elements approximation. Asterisks give the true values of the parameters.

An elementwise representation of the parameter updates and the related Riccati equation is utilized to decrease the computational load for the case of periodic regressor at the price of a small decrease in the identification algorithm performance.

REFERENCES

H. Akçay and N. At. Membership set identification with periodic inputs and orthonormal regressors. *Signal Processing*, 86(12):3778–3786, 2006.

W. J. Beyn and L. Elsner. Infinite products and paracontracting matrices. *The Electronic Journal of Linear Algebra*, 2:1–8, 1997.

S. Bittanti, P. Bolzern, and M. Campi. Convergence and exponential convergence of identification algorithms with directional forgetting factor. *Automatica*, 26(5): 929–932, 1990.

L. Cao and H. M. Schwartz. Analysis of the Kalman filter based estimation algorithm: An orthogonal decomposition approach. *Automatica*, 40:5–19, 2004.

I. Daubechies and J. C. Langarias. Sets of matrices all finite products of which converge. *Linear Algebra and Its Applications*, 161:227–263, 1992.

I. Daubechies and J. C. Langarias. Corrigendum/addendum to: Sets of matrices all finite products of which converge. *Linear Algebra and Its Applications*, 327:69–83, 2001.

M. Evestedt and A. Medvedev. Stationary behavior of an anti-windup scheme for recursive parameter estimation under lack of excitation. *Automatica*, 42:151–157, January 2006a.

M. Evestedt and A. Medvedev. Model-based slopping monitoring by change detection. In *Proceedings of IEEE Conference on Control Applications*, Munich, Germany, 4-6 October 2006b.

M. Evestedt, A. Medvedev, and T. Wigren. Windup properties of recursive parameter estimation algorithms in acoustic echo cancellation. In *Proceedings of the 16th IFAC World Congress*, Prague, Czech Republic, July 2005.

S. L. Gay. An efficient, fast converging adaptive filter for network echo cancellation. In *Proceedings of Asilomar Conference*, Monterey, USA, November 1998.

T. Hägglund. *New Estimation Techniques for Adaptive Control.* Lund Institute of Technology, Sweden, 1983.

D. J. Hartfiel. *Nonhomogeneous Matrix Products.* World Scientific Publishing, 2002.

L. Ljung and S. Gunnarsson. Adaptation and tracking in system identification - A survey. *Automatica*, 26(1): 7–21, 1990.

A. Medvedev. Stability of a Riccati equation arising in recursive parameter estimation under lack of excitation. *IEEE Transactions on Automatic Control*, 49(12):2275–2280, December 2004.

A. Medvedev and M. Evestedt. Elementwise decoupling of the riccati equation in the SG-algorithm with convergence estimates. In *Submitted to the 17th IFAC World Congress*, Seoul, Korea, July 2008.

S. Mokino. Exponentially weighted stepsize NLMS adaptive filter based on the statistics of a room impulse response. *IEEE Transactions on Speech and Audio Processing*, 1(1):101–108, January 1993.

C. Olsson. Active vibration control of multibody systems. Ph. D. Thesis Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology: 59, Uppsala University, 2005.

P. Ramos, R. Torrubia, A. López, A. Salinas, and E. Masgrau. Step size bound of the sequential partial update LMS algorithm with periodic input signals. *EURASIP Journal on Audio, Speech, and Music Processing*, 2007, 2007.

M. Rupp and J. Cezanne. Robustness conditions of the LMS algorithm with time-variant matrix step-size. *Signal Processing*, 80:1787–1794, 2000.

T. Söderström and P. Stoica. *System Identification.* Prentice Hall, 1989.

B. Stenlund and F. Gustafsson. Avoiding windup in recursive parameter estimation. In *Preprints of Reglermöte 2002*, pages 148–153, Linköping, Sweden, May 2002.

T. Wigren. Fast converging and low complexity adaptive filtering using an averaged Kalman filter. *IEEE Transactions on Signal Processing*, 46(2):515–518, February 1998.

Appendix A. PROOF OF PROPOSITION 4

After the transformation of (5) by $\bar{\theta}(t) = T^T(t)\hat{\theta}(t)$ the system matrix becomes $\bar{A}_t^{-1}(\cdot)$. Furthermore, the input matrix is transformed as

$$T^T K(t) = \begin{bmatrix} \xi_1^{1T} \\ \vdots \\ \xi_1^{nT} \end{bmatrix} \frac{P(t-1)\varphi(t)}{r + \rho_i \xi_1^{i\,T} P(t-1)\xi_1^i} = \frac{1}{||\varphi(t)||_2} D_i(P(t-1))$$

Thus the transformed system equations can be written in terms of individual elements as

$$\bar{\theta}_k(t) = \bar{\theta}_k(t-1) - D_i^k(P(t-1)) \left(\bar{\theta}_i(t-1) + \frac{1}{||\varphi(t)||_2} y(t) \right)$$

or, for the special case of $k = i$

$$\bar{\theta}_i(t) = (1 - D_i^i(P(t-1))\bar{\theta}_i(t-1) + D_i^i(P(t-1)) \times \frac{1}{||\varphi(t)||_2} y(t)$$

The above equation together with (15) yields the desired result.