An innovations approach for identifying linear, time invariant, stochastic vector processes

van Gent, A.R.

Award date:
1985

DEPARTMENT OF ELECTRICAL ENGINEERING
EINDHOVEN UNIVERSITY OF TECHNOLOGY
Group Measurement and Control

AN INNOVATIONS APPROACH FOR IDENTIFYING LINEAR, TIME INVARIANT, STOCHASTIC VECTOR PROCESSES

by Bert van Gent

This report is submitted in fulfillment of the requirements for the degree of electrical engineer (M.Sc.) at the Eindhoven University of Technology. The work was carried out from May 1982 until May 1983 in charge of Prof. dr. ir. P. Eykhoff under supervision of dr. ir. A.A.H. Damen and dr. ir. A.J.W. van den Boom

Summary

In this report an innovations approach is presented to identify stochastic linear, time invariant, Gaussian distributed, time discrete vector processes. The innovations representation is obtained from the estimated covariance by way of an approximated realization of the estimated covariance and by then solving the algebraic Riccati equation.
The Riccati equation is solved by means of the computation of the stable subspace of the associated pencil.
The presented algorithm is tested by a number of simulations.
Also the algorithm is used to identify a physical process, viz. dunes in a sand flume.

Samenvatting

In dit verslag wordt het 'innovations' model gebruikt om stochastische, lineaire, tijd invariante, tijd discrete, Gaussisch verdeelde processen te identificeren. Op grond van de geschatte covariantie informatie kunnen we het'innovations' model bepalen, via de benaderde realisatie voor de geschatte covariantie,en het daarna oplossen van de algebraïsche Riccati vergelijking .
De Riccati vergelijking wordt opgelost door het berekenen van de stabiele ruimte behorende bij het geassocieerde 'pencil'.
Het gepresenteerde algoritme wordt getest d.m.v. van een aantal simulaties.
Bovendien wordt het algoritme gebruikt om duinen in een zandgoot te identificeren.

Contents

Seperate report:

User's manual for the stochastic identification programs

Chapter 0

INTRODUCTION

System identification is a field in science that is concerned with the study of systems. A system is a collection of things or parts working together in a regular relation. Based on observations and a priori knowledge one tries to obtain more insight into the considered system by means of a model. For our purpose we define a model as representation of the essential aspects of the systems (cf. Eykhoff (1974)). Generally the a priori knowledge may be just sufficient for the determination of the modelstructure, but the estimation of the parameters has to be based on the input and output data. In this report we restrict ourselves to study stochastic processes, i.e. processes from which we can only measure the output data. A stochastic process is schematically depicted in figure 1.1. A definition of such a process is:

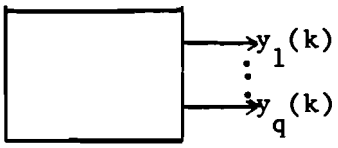


Figure 1.1 A stochastic vector process

Definition (cf. Doob (1953))

A stochastic vector process y(k) is a family of random variables with the time k as argument.

In this report we consider time discrete, stochastic processes with the following properties:

- linear
- $\{y(k)\}$ has a Gaussian distribution
- time invariant
- there exists a finite dimensional model.

The statistical properties of the considered process are represented by the set of all probability densities (or the corresponding distribution functions). Because we are dealing with Gaussian processes the covariance knowledge provides sufficient information and the process can be modeled by the innovations representation (Anderson (1979)):

$$x(t+1) = Ax(t) + Ke(t) \qquad (0.1)$$

$$y(t) = Cx(t) + e(t) \qquad (0.2)$$

with
- $y(t)$ the q-output vector
- $x(t)$ the n-state vector
- $e(t)$ the q-white noise vector (innovations process)
- A nxn system matrix
- K nxq Kalman gain matrix
- C qxn output matrix
- Q qxq covariance matrix of the innovations process, $E\{e(t)e^T(t)\}=Q$
- q the number of outputs
- n the dimension of the process

$\{A,C,K,Q\}$ is called the innovations representation of the process $\{y(t)\}$ corresponding with equations (0.1) and (0.2)

If the exact knowledge of the covariance is available we are able to compute an innovations approach and this is called the stochastic realization problem and is stated in chapter 1. Generally the covariance information is not available and we have to estimate the covariance with a limited number of samples. The effect of a limited number of samples on the estimated covariance is studied in chapter 2. When we have to estimate the covariance of a system from the output data and try to find a model, then this will be called the stochastic identification problem.

Both in the stochastic identification and in the stochastic realization one has to solve a highly nonlinear matrix equation: the algebraic Riccati equation. In chapter 3 is presented a numerical suitable method to solve this equation. With the theory presented in chapter 1 and 3 we are able to compute an innovations approach, but it is by no means clear that this is a good identification method of stochastic processes. The reliability of this algorithm, when the covariance has to be estimated, is checked in chapter 4.

A description of the computer programs can be found in a separate report.

Chapter 1

STOCHASTIC REALIZATION

1.1 Introduction

In this chapter we present an object that has been extensively studied in the last ten years: the stochastic realization. We limited ourselves to wide sense stationary,purely nondeterministic processes $\{y(k)\}$. This does simplify the discussion considerably. We will try to summerize the main results which are necessary for the discussion of stochastic realization. We consider rational time series $\{y(k)\}$, that is $\{y(k)\}$ can be modelled by a finite dimensional state space model. A special kind of state space model is the innovations representation. This representation was first introduced by Kailath in 1968. The advantage of this model is that it is useful in a wide variety of fields such as least square estimation, representation of stochastic processes and identification of stochastic processes.

This chapter is organized as follows: in section 1.2 the class of linear models which we study is given. In section 1.3 the dimension of the process is defined,and Hankel and Page matrices are presented with their application for finding an approximate realization of the covariance. Next is presented the set of all state space models which have the same covariance functions. In section 1.5 an algorithm is given to find a realization of the considered stochastic process and finally in section 1.6 the conclusion is given.

1.2 Rational time series and the state space representation

We shall be concerned with a q-vector rational time series:

$$y(t)= \begin{bmatrix} y_1(t) \\ y_2(t) \\ \vdots \\ y_q(t) \end{bmatrix} \quad t=1,..,N \qquad (1.1)$$

A rational time series is defined as:

Definition 1.1 (Faurre (1976))

A stochastic process y(t) is said to be rational if it can be modeled as the output of a linear finite dimensional system with a white noise input.

Later we shall explain what is meant by a finite dimensional process and actually define the dimension of a process.
The state space representation for a time discrete series y(t) is a model of the form:

$$x(t+1) = A\,x(t) + B\,u(t) \qquad (1.2)$$

$$y(t) = C\,x(t) + D\,v(t) \qquad (1.3)$$

$$\text{with } E\{x(0)\}=0,\ E\{x(0).v^T(t)\}=0, E\{x(0).u^T(t)\}=0$$

$$E\left\{ \begin{bmatrix} u(t) \\ v(t) \end{bmatrix} . \begin{bmatrix} u^T(\tau), v^T(\tau) \end{bmatrix} \right\} = \begin{bmatrix} \Sigma & S \\ S^T & M \end{bmatrix} . \delta(t-\tau) \geq 0 \qquad (1.4)$$

The output q-vector y(t) is a linear combination of a related n-vector process $\{x(t)\}$, the state vector, plus additive, zero mean, Gaussian white noise with constant covariance matrix M. M is of course a nonnegative symmetric matrix.
The vector process $\{x(t)\}$ is assumed to be a Markov process, with A and B constant matrices. Σ is a full rank covariance matrix of the zero mean Gaussian white noise u(t). The input p-vector u(t) and the q-vector v(t), respectively the input-noise and the output-noise, are mutually correlated with:

$S=\{u(t).v^T(\tau)\}.\delta(t-\tau)$.

The q-vector process $\{y(t)\}$ is a linear projection of an n-vector Markov process.
We will first define the covariance of a stationary process $\{y(t)\}$:

$$R(k) = E\{ y(t+k).y^T(t)\} \tag{1.5}$$

Further on we assume that y(t) is purely nondeterministic in the sense that :

$$\lim_{k\to\infty} R(k) \to 0 \tag{1.6}$$

Together with an observability condition on $\{A,C\}$ this implies that x(t) is stationary and purely nondeterministic and thus:

$$E\{x(t).x^T(t)\} = P \tag{1.7}$$

where P is a symmetric nonnegative matrix. Stationarity implies that P is the solution of the Lyapunov equation (see Anderson (1979)):

$$P - A\,P\,A^T = B\,\Sigma\,B^T \tag{1.8}$$

A necessary and sufficient condition for the existense of a nonnegative solution is that A is asymptotically stable; this implies that A has all its eigenvalues inside the unit circle.

Now we will derive an expression for the covariance R(k). From the state space equations we can derive an expression for the output y(t+k):

$$y(t+k) = C\,A^k\,x(t) + \sum_{i=0}^{k-1} C\,A^i\,B\,u(t+k-i-1) + Dv(t+k) \tag{1.9}$$

By using equation (1.4) and the white noise properties of respectively v(t) and u(t) we can derive the next expression

for the covariance for k>0:

$$R(k) = C\,A^{k}\,P\,C^{T} + C\,A^{k-1}\,B\,S\,D^{T} \tag{1.10}$$

$$R(k) = C\,A^{k-1}\left[A\,P\,C^{T}+B\,S\,D^{T}\right] = C\,A^{k-1}\,L \tag{1.11}$$

With $L = A\,P\,C^{T} + B\,S\,D^{T}$ (1.12)

For k<0:

$$R(k)=E\{y(t-\|k\|).y^{T}(t)\}=E\{y(t).y^{T}(t+\|k\|)\}=R^{T}(-\|k\|) \tag{1.13}$$

Immediately follows that :

$$R(k) = L^{T}\,(A^{T})^{\|k\|-1}\,C^{T} \qquad k<0 \tag{1.14}$$

And finally we obtain for k=0:

$$R(0) = E\{y(0).y^{T}(0)\} = C\,P\,C^{T} + D\,M\,D^{T} \tag{1.15}$$

We can conclude that the covariance of a rational stationary time series can be written as:

$$R(k) = C\,A^{k-1}\,L\,1(k) + L^{T}\,(A^{T})^{\|k\|-1}\,C^{T}1(-k) + R(0)\,\delta(k)$$

$$\text{with } 1(k) = 1-\delta(k) \tag{1.16}$$

The spectrum $\Phi(z)$ of y(t), the z-transform of the covariance R(k) is:

$$\phi(z) = \sum_{-\infty}^{+\infty} R(k)\,z^{-k} = C\,(z\,A-I)^{-1}L + L^{T}(z^{-1}\,I-A^{T})^{-1}C^{T} + R(0) \tag{1.17}$$

One can prove that all infomation that is needed to find a stochastic realization of a Gaussian distributed process , is the knowledge of the covariance (see Gevers and Wouters (1978))

We have assumed that $x(0), \{v(t)\}, \{u(t)\}$ are Gaussian.Then the knowledge of the mean and the covariance is sufficient to deduce the density function of any order. In the non Gaussian case, knowledge of the mean and the covariance does not provide

complete information about the higher order moments, so the covariance information is not sufficient.

1.3 Dimension of the rational time series and the realization of the covariance

In this section we will give another definition of a rational time series and simultaneously define the dimension of the process by means of the Hilbert space. Later another definition is given by means of the indefinite Hankel matrix.
The Hilbert space is defined as the orthogonal projection of the "future" $\{y(t+k),k=0,1,.,\infty\}$ on the past $\{y(t+k),k=-1,.,-\infty\}$. Because we limited ourselves to zero mean Gaussian processes orthogonal projections (denoted by /) and conditional expectations are the same, see Hajdasinski (1980 ,p51).
The orthogonal projection from the "future" on the past is denoted by:

$$\chi = \{y(t+k),k=0,1,.,\infty\}/\{y(t+m),m=-1,.,-\infty\}$$

Definition 1.2

A stationary time series is said to be rational if and only if dim χ =n, where n is finite and is called dimension of the process.

$$\dim\{y(t)\}= n \quad \leftrightarrow \dim \chi = n \tag{1.18}$$

Such a time series can be represented by a state space model, as shown in the previous section, but we can give another equivalent definition of the dimension of $\{y(t)\}$.

Now we will introduce a matrix with a special structure: the Hankel matrix. This Hankel matrix H is composed by covariance matrices and has the shape:

$$H(j)= \begin{bmatrix} R(1) & R(2) & R(3). & . & . & R(j) \\ R(2) & R(3) & . & . & . & . \\ . & & & & & \\ R(j) & . & . & & . & R(NOM) \end{bmatrix} \qquad (1.19)$$

with NOM =2.j-1 (NOM: number of used covariance matrices)

Now we can give another equivalent definition of the dimension of $\{y(t)\}$ (see Gevers and Wouters (1978)).

Definition 1.3

The stationary process $\{y(t)\}$ is a process of dimension n if and only if the rank of the infinite Hankel matrix $H(\infty)$ is n.

$$\dim\{y(t)\} = n \leftrightarrow \operatorname{rank}(H(\infty)) = n \qquad (1.20)$$

Definition 1.2 and 1.3 are depicted in figure 1.1

$$\begin{array}{ccc} \dim\{y(t)\} = n & \leftrightarrow & \dim \chi = n \\ \updownarrow & & \updownarrow \\ \operatorname{rank}(H(\infty)) & & \end{array}$$

Figure 1.1

From the deterministic realization problem we know,that we can involve the Hankel matrix to find a realization $\{A,L,C\}$ of the covariance. Before we give an essential theorem to find a realization by means of the Hankel approach, several definitions must be introduced .

Definition 1.4

Any polynomial f(z):

$$f(z) = z^k - a_1 \cdot z^{k-1} - \quad . \quad . \quad -a_{k-1} \cdot z - a_k \cdot z^0 \qquad (1.21)$$

for which holds f(A) = 0, is called an annihilating polynomial of matrix A.

Definition 1.5

The polynomial f(z) of the smallest degree s, non equal to zero, fulfilling the above given definition is called the minimal polynomial.

Theorem 1.1 (see Van den Hof (1982),Gevers and Wouters (1978))

The sequence of covariance matrices R(k),k=1,2,.... has a finite dimensional realization $\{A,L,C\}$ if and only if there is an integer s and constants a_i such that :

$$R(s+j) = \sum_{i=1}^{s} a_i . R(s+j-i) \text{ for all } j>0 \quad (1.22)$$

One can derive another powerful result,based on (1.22), and this result plays very important role in the realization theory.

Theorem 1.2 (Gevers and Wouters (1978))

R(k), k=1,...NOM (with NOM>2.s-1) has a finite dimensional realization $\{A,L,C\}$ with A(nxn)-matrix,(A,L) controllable and (C,A) observable if and only if rank(H(s))=n and rank(H(s+j))=n for all j>0

There are several algorithms to find a realization $\{A,L,C\}$ of the covariance from the Hankel matrix. A comparison of these methods can be found in Van den Hof (1982).
Futher theorem 1.2 implies that the dimension of the state vector is equal to the dimension of $\{y(t)\}$.

In 1982 Damen and Van den Hof introduced another approach to find a realization of covariance matrices, namely by means of the Page matrix. The Page matrix is defined as follows:

$$P(\eta,\mu) = \begin{bmatrix} R(1) & R(2) & \cdots & R(\mu) \\ R(\mu+1) & & & \\ \vdots & & & \\ R((\eta-1)\mu+1) & \cdots & & R(\eta\mu) \end{bmatrix} \qquad (1.23)$$

When we compare the Hankel and Page matrix, then we see that the covariance matrix with time shift k appears only once in the Page matrix in contrast with the Hankel matrix. Other properties and realization algorithms by means of the Page matrix can be found in Van den Hof (1982). Without futher comment we give a similar theorem for the Page matrix as for the Hankel matrix.

Theorem 1.3

If R(k), k=1,2,...NOM (NOM$\geq s^2$) has a finite dimensional realization $\{A,L,C\}$ with A nxn-matrix, (A,L) controllable, (A,C) and (A^{μ},C) observable then is rank P(s,s)=n and rank (P(s+i,s+j))=n for all i,j ≥ 0

It is not clear whether rank(P(s+i),(s+j)) for all i,j is a sufficient condition; this is still a point of investigation. An advantage of the Page matrix over the Hankel matrix are the smaller dimensions, when taking the same number of covariance matrices into account.
Anticipating the use of both approaches in the stochastic identification, i.e. the covariance has to be estimated, we can notice that there is an equally balanced filtering in the Page matrix.
Finally we can depict the main results of this section in figure 1.2.

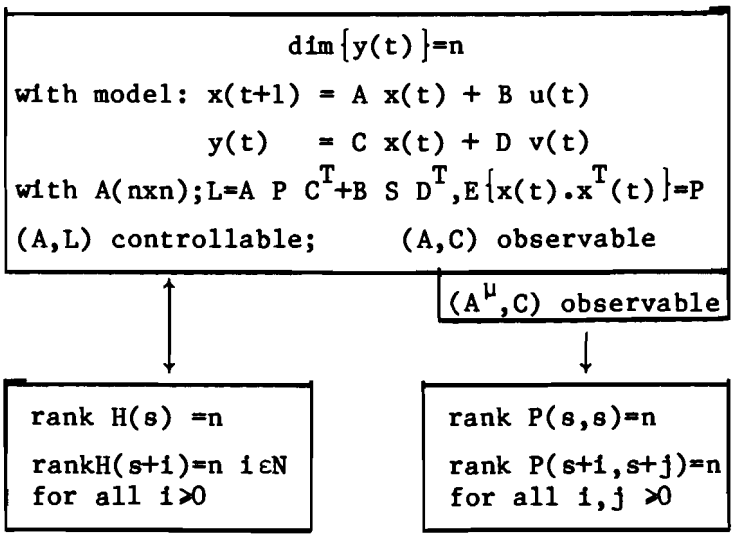


Figure 1.2

1.4 The set of all state space representations

In this section we want to point out that several state space models may generate the same rational time series $\{y(t)\}$ or more precisely ,processes that have the same covariance functions R(k). These models will be covariance-equivalent realizations.

Definition 1.6

We define the set $\{P\}$ of all state space representations associated with the given covariance R(k), or equivalently with $\{A,L,C,R(0)\}$, as the set of all symmetrical positive definite matrices $P=P^T>0$ satisfying:

$$P - A P A^T = B \Sigma B^T \tag{1.8}$$

$$L - A P C^T = B S D^T \tag{1.12}$$

$$R(0) - C P C^T = D M D^T \tag{1.15}$$

$$\begin{bmatrix} B \Sigma B^T & B S D^T \\ D S^T B^T & D M D^T \end{bmatrix} \geq 0 \tag{1.24}$$

To find a state space representation, assuming that $\{A,L,C\}$ are known, one has to solve the linear equations (1.8),(1.12) and (1.15) with the highly nonlinear constraint (1.24). Before we give a necessary and sufficient condition ,assuming that the set $\{P\}$ is non void, we have to introduce the indefinite symmetrical Toeplitz matrix Tp and we have to give a definition of the positive realness of the covariance R(k). The Toeplitz matrix Tp has the form:

$$Tp = \begin{bmatrix} R(0) & R(1) & R(2) & \cdots \\ R^T(1) & R(0) & R(1) & \\ R^T(2) & R^T(1) & & \\ \vdots & \vdots & & \end{bmatrix} \qquad (1.25)$$

Definition 1.6 (Faurre (1976))

The covariance R(k) is positive real if and only if for any sequence w(i):

$$\sum_{i,j} w^T(i).R(i-j).w(j) \geq 0 \qquad (1.26)$$

A sequence of matrices is a sequence of covariance matrices if and only if the sequence of matrices is positive real. This is a sufficient and necessary condition so that the set $\{P\}$ is non void (see Faurre (1976)). In the literature this is known as the positive real lemma.

Positive Real Lemma

R(k) is positive real if and only if the associated set P is non void.

$$Tp \geq 0 \quad \leftrightarrow \text{ the set } \{P\} \text{ is non void}$$

The set $\{P\}$ is a closed bounded convex set and we can define respectively a maximal point P^+ and a minimal point P^-.

Definition 1.8

P^+ is a maximal point of the set P i.e. $P \leq P^+$ for all elements of the set $\{P\}$.

Definition 1.9

P^- is a minimal point of the set P i.e. $P \geq P^-$ for all elements of the set $\{P\}$.

In section 1.2 we have presented a state space representation of a rational time series. Without loss of general validity of the state space model, i.e. we can represent a spectral equivalent process, we can replace the state space equations (1.2) and (1.3) by:

$$x(t+1) = A\,x(t) + u_1(t) \tag{1.27}$$

$$y(t) = C\,x(t) + v_1(t) \tag{1.28}$$

where $u_1(t)$ is an n-vector and $v_1(t)$ a q-vector stationary Gaussian white noise process with covariance:

$$E\left\{\begin{vmatrix} u_1 \\ v_1 \end{vmatrix} \begin{vmatrix} u_1^T & v_1^T \end{vmatrix}\right\} = \begin{bmatrix} B\,\Sigma\,B^T & B\,S\,D^T \\ D\,S^T\,B^T & D\,M\,D^T \end{bmatrix} = \begin{bmatrix} \Sigma_1 & S_1 \\ S_1^T & M_1 \end{bmatrix} \geq 0 \tag{1.29}$$

This means that the distribution matrix and the output matrix take into account the covariance matrices of respectively the input and the output noise.
Let $\{A,B,C,D,\Sigma,S,M\}$ be a realization of $\{y(t)\}$ then $\{T\,A\,T^{-1}, CT^{-1}, TB, D, \Sigma, S, M\}$ is also a realization of $\{y(t)\}$ with T any non singular matrix. This is simply a coordinate transformation in the state space model, but does not involve the driving noise covariances. However there are other equivalent realizations, which have not the same input-output map, but still yields the same covariance R(k). For a given $\{A,C\}$, which fixes the coordinate system for the state space,

any combination of $\{B,D,\Sigma,S,M\}$ which gives such L satisfy.
We can distinguish two categories:

(I) replace $\{B,\Sigma,S,M,D\}$ by $\{I,\Sigma_1,S_1,M_1,I\}$ according to equation (1.29). The non unique model is given by (1.27) and (1.28).

(II) there exists an equivalent <u>minimal innovations representation</u> :

$$z(t+1) = A\ z(t) + K\ e(t) \tag{1.30}$$

$$y(t) = C\ z(t) + e(t) \tag{1.31}$$

e(t) is a p-vector white noise process (innovations process) with $E\{e(t+\tau).e^T(t)\} = Q.\delta(\tau)$ and K the Kalman gain with $Q = R(0) - C\ P^-\ C^T$ (1.32)

and $K = (L - A\ P^-\ C^T)(R(0) - C\ P^-\ C^T)^{-1}$ (1.33)

with P^- smallest positive definite solution of the algebraic Riccati equation:

$$P = A.P.A^T+(L-A.P.C^T).(R(0)-C.P.C^T)^{-1}.(L^T-C.P.A^T) \tag{1.34}$$

The solution P corresponds with the minimal solution of the set P of equations (1.8),(1.12),(1.15) and (1.24).
The corresponding covariance matrix of the input- and output noise is:

$$\begin{bmatrix} KQK^T & KQ \\ QK^T & Q \end{bmatrix} = \begin{bmatrix} KQ^{\frac{1}{2}} \\ Q^{\frac{1}{2}} \end{bmatrix} \begin{bmatrix} Q^{\frac{1}{2}}K^T & Q^{\frac{1}{2}} \end{bmatrix} \begin{matrix} \uparrow \\ q \\ \downarrow \end{matrix} \tag{1.35}$$

Remark: the innovation e(t) is defined as:

$$e(t) = y(t) - \hat{y}(t|t-1) \tag{1.36}$$

with $\hat{y}(t|t-1)$ the least square estimate of y(t), based on the past.

The innovations approach is considered among others in Anderson (1979), in which the equivalence between the innovations model and the state space model as well as their relation by means of the Riccati equation is stated.
Both representations mentioned above give a reduction of the total number of coefficients, if we compare these representations with the original model, given by equations (1.2) and (1.3).

For our stochastic realization we are interested in the innovations representation, because it has advantages in the least squares estimation, such as filtering and prediction, and the number of coefficients is less than in other covariance equivalence models. In the last part of this section we give some properties of the innovations representation.

Properties:

Given a minimal state space realization $\{A,B,C,D,\Sigma,S,M\}$ for a q-vector process $\{y(t)\}$, there always exists a covariance equivalent minimal innovations realization $\{A,C,K,Q\}$ such that:

- The input and output noise respectively u(t) and v(t) are both replaced by the same q-vector white noise e(t).
- The covariance of the state of the innovations model is always smaller than or equal to the covariance of the state of the original model and is denoted by P^-.
- The innovations representations is causally invertible, i.e. there exists a filter:

$$z(t+1) = (A-K\,C)\,z(t) + K\,y(t) \qquad (1.37)$$

$$e(t) = y(t) - C\,z(t) \qquad (1.38)$$

The above mentioned properties can be found in Anderson (1979).

1.5 The state space realization algorithm

In this section we will present an algorithm to find a state space model for a rational time series $\{y(t)\}$. We assume that the covariance is known. In the previous section we have shown that such a time series can be represented by an innovations model. In figure 1.3 the realization algorithm is depicted. A realization $\{A,L,C\}$ of the covariance can be found by involving a Hankel or a Page matrix. The realization algorithms

by means of the Hankel or Page approach can be found in Van den Hof (1982).

The next step in our algorithm is to solve the algebraic Riccati equation and to compute the Kalman gain and the covariance of the innovations process. A reliable method to solve this equation is given in chapter 3. Finally we have a finite dimensional state space representation of the rational time series $\{y(t)\}$: the innovations representations.

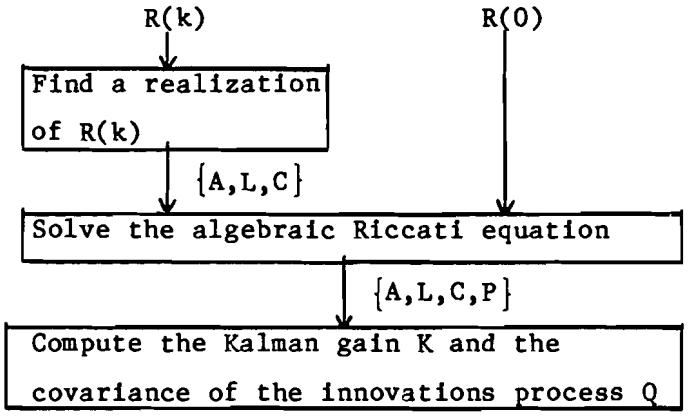


Figure 1.3 Stochastic realization algorithm

1.6 Conclusion

It is by no means clear that the realization algorithm, presented in this chapter and based on the exact knowledge of the covariance, is also a good identification algorithm, based on the covariance estimated from a finite number of samples. Before we check the reliability of this algorithm as an identification algorithm, we consider first some properties of the estimated covariance in the next chapter.

Chapter 2

SOME NOTES ON THE ESTIMATED COVARIANCE

2.1 Introduction

In this chapter we will investigate what is the consequence of a limited number of samples for calculating the covariance of an output signal of a time discrete process. We also examine the second order crossmoment in channel number and in time shift of the covariance function. These subjects are studied, because we want to have an impression of the second order crossmoment of the estimated covariance matrix function, as it highly influences: - the choice of the order of the realization
- the ultimate estimated realization.

This is the final goal of this study.
For this purpose Hankel or Page matrix is involved, these are composed of the covariance block matrices corresponding to the various time shifts. As we will use the estimated covariance instead of the exact covariance in the realization algorithm, it is important to know how the noise on these block-matrices is correlated.
In section 2.2 an expression will be derived for the second order crossmoment in channel number of the estimated covariance and in section 2.3 an expression will be derived for the second order crossmoment in channel number and in time shift of the estimated covariance. These expressions are functions of the unknown exact covariance. In section 2.4 we replace in these expressions the unknown exact covariance by the estimated covariance, that is we generate output series of a first order single output process, and calculate the second order crossmoment of the estimated covariance from the output data. We compare the expressions of the covariance of the estimated covariance calculated with respectively the exact covariance and the estimated covariance with each other in section 2.4.

2.2 The second order crossmoment in channel number of the estimated covariance

The covariance matrix function of an output signal of a time invariant vector process will only be a function of the time shift.We define the covariance function of a time discrete vector process as:

$$R(k) = E\{ y(t+k).y^T(t)\} \qquad \text{for all } t \qquad (2.1)$$

with y(k) a q-vector.
Because of the assumption that y is a purely non deterministic vector and that $E\{y(t)\}=0$, it follows that :

$$\lim_{k\to\infty} \|R(k)\| \to 0$$

Other properties of the covariance matrix are:

- $R(k)=R^T(-k)$
- R(0) is positive definite

An unbiased estimator of the covariance-matrix function is:

$$\tilde{R}(k)= \frac{1}{(N-k)} \sum_{i=1}^{N-k} y(i+k).y(i)^T \qquad (2.2)$$

$$\text{i.e. } E\{\tilde{R}(k)\} = R(k)$$

where N is the number of sample vectors.
The second order crossmoment in channel number of the estimated covariance is given by:

$$\begin{aligned} Var(\tilde{R}(k))_{i,j,l,m} &= E[(\tilde{R}_{i,j}(k)-E\{\tilde{R}_{i,j}(k)\})(\tilde{R}_{l,m}(k)-E\{\tilde{R}_{l,m}(k)\})] \\ &= E\{\tilde{R}_{i,j}(k).\tilde{R}_{l.m}(k)\} - R_{i,j}(k).R_{l,m}(k) \qquad (2.3) \end{aligned}$$

with $R_{i,j}(k)$: (i,j) element of the matrix R(k)

We see that the $Var(\tilde{R}(k))$ is a function of the unknown R(k).

The second order crossmoment in channel number for a single output process will be called variance of the estimated covariance. In this section we will first derive an expression for the variance of the estimated covariance of a single output Gaussian process, because this gives more insight. At the end

of this section the second order crossmoment of the estimated covariance is given for a q-vector Gaussian process.
The variance of the estimated covariance of a single output process is:

$$\mathrm{Var}(\tilde{R}(k)) = E\{\tilde{R}^2(k)\} - R^2(k) \qquad (2.4)$$

Substitution of (2.2) in (2.4) leads to:

$$\mathrm{Var}(\tilde{R}(k)) = \frac{1}{(N-k)^2} \sum_{i=1}^{N-k} \sum_{j=1}^{N-k} E\{y(i)y(j)y(i+k)y(j+k)\} - R^2(k) \qquad (2.5)$$

Based on the assumption that x is a normally distributed variable,one can write (cf. Laning and Battin (1956)):

$$E\{x_1x_2x_3x_4\} = E\{x_1x_2\}.E\{x_3x_4\} + E\{x_1x_3\}.E\{x_2x_4\} + E\{x_1x_4\}.E\{x_2x_3\} \qquad (2.6)$$

When we apply (2.6) in (2.5) we finally get the expression:

$$\mathrm{Var}(\tilde{R}(k)) = \frac{R^2(0)}{(N-k)} + \frac{R^2(k)}{(N-k)} + \frac{2}{(N-k)} \sum_{i=1}^{N-k-1} \left[1 - \frac{i}{(N-k)}\right].\left[R^2(i) + R(i+k).R(i-k)\right] \qquad (2.7)$$

The derivation of formula (2.7) is given in appendix A2.1 for a q-vector process including time shifts and in Van den Hof (1981) for a scalar process without time shifts.
If the covariance function has the form of a dirac function, the values of $|R(k)|$ for $k \neq 0$ will be much smaller than for R(0). Under this assumption an approximation of the variance of the covariance is:

$$\mathrm{Var}(\tilde{R}(k)) \approx \frac{R^2(0)}{(N-k)} \qquad (2.8)$$

If the covariance function has not a δ-form, formula (2.7) is quite complicated. We want to have a more handy expression for formula (2.7). To get this, we have to make approximations.

One can roughly recognize two areas for k:

$$\text{I} \quad k < \ell \tag{2.9 a}$$

$$\text{II} \quad k \gg \ell \tag{2.10a}$$

where ℓ has been defined by: if $k > \ell$ then $R(k) \ll R(0)$. Area I then corresponds to the decreasing part of the variance of the estimated covariance with an increasing k and area II with the increasing part. The approximations for the variance of the estimated covariance of the two areas are (in the case $\ell \ll N$):

$$\text{I} \quad \mathrm{Var}(\tilde{R}(k)) \approx \frac{1}{N} \left(C + R^2(k) + 2 \cdot \sum_{i=1}^{2\ell} R(i+k) \cdot R(i-k) \right)$$

$$\text{with } C \approx R^2(0) + 2 \sum_{i=1}^{\ell} R^2(i) \tag{2.9 b}$$

C : the power of the autocorrelation

$$\text{II} \quad \mathrm{Var}(\tilde{R}(k)) \approx \frac{C}{(N-k)} \tag{2.10b}$$

The derivation of formulas (2.9b and 2.10 b) is given in appendix A2.2.

Remark: if we want to make approximations, we need a criterion to determine ℓ quantitatively. The criterion chosen is:

$$\frac{\|R(k)\|}{\|R(0)\|} < \delta \quad \text{for all } k > \ell \tag{2.11}$$

The choice of δ is not arbitrary, that is if we choose δ too large the reliability of the approximations decreases. On the otherhand if we choose δ too small, we have accurate approximations, but the number of covariance values that must be taken into account increases.

Example: A first order single output process.

The state space equations of this process are:

$$x(k+1) = a\, x(k) + \xi(k) \tag{2.12a}$$

$$y(k) = x(k) \tag{2.12b}$$

with $E\{\xi(k).\xi(k+t)\}=\delta(t)$ and $|a|<1$ .The associated covariance function of the output is:

$$R(k) = \frac{1}{(1-a^2)} .a^{|k|} \qquad (2.12c)$$

The covariance is depicted in figure 2.1 for a=0.95. If we substitute the covariance of this process (2.12c) in the expressions for the approximated variance of the estimated covariance (respectively equation (2.9b) and (2.10b)), then we find the following expressions for the two areas:

$$\text{for } k=0: \quad Var(\tilde{R}(k)) = \frac{2}{N(1-a^2)^3}\left[1+ a^2- 2a^{2+2\ell}\right] \approx \frac{2(1+a^2)}{N(1-a^2)^3} \qquad (2.13a)$$

$$\text{I for } 0<k<\ell \quad Var(\tilde{R}(k)) = \frac{1}{N(1-a^2)^3}\left[1+a^2+a^{2k}(3-a^2+2k(1-a^2)) -2a^{2\ell}(1+a^{2\ell})\right]$$

$$\approx \frac{1}{N(1-a^2)^3}\left[1+a^2+a^{2k}(3-a^2+2k(1-a^2))\right] \qquad (2.13b)$$

$$\text{II for } k>>\ell \quad Var(\tilde{R}(k)) = \frac{1}{(N-k)(1-a^2)^3}\left[1+a^2-2a^{2+2\ell}\right] \approx \frac{1+a^2}{(N-k)(1-a^2)^3} \qquad (2.13c)$$

If k increases in expression (2.13b) then the variance of the estimated covariance decreases, which can easily be seen. But if k increases in expression (2.13c) then the variance of the estimated covariance increases, which is the result of the decreasing number of estimated covariance values that can be taken into account.

In figure 2.2 the exact variance of the estimated covariance together with the approximations of the variance of the estimated covariance are depicted for a=0.95 and N=1000, where the chosen value of δ=0.15, which leads to a corresponding value of ℓ=37. The derived approximations of the variance of the estimated covariance (equation (2.13)) are reliable with a suitable value of δ, as we can see from figure 2.2.

As mentioned before we will give an expression for the second order crossmoment in channel number of the estimated covariance of a q-vector Gaussian process:

$$\mathrm{Var}(\tilde{R}(k))_{i,j,l,m} = \frac{1}{(N-k)^2} \sum_{s=1}^{N-k-1} [R_{i,l}(s-n)\cdot R_{j,m}(s) + R_{i,l}(\ -s)\cdot R_{j,m}(-s) + R_{i,m}(s+k)\cdot R_{j,l}(\ s-k)) + R_{i,m}(-s+k)\cdot R_{j,l}(-s-k\ )] + \frac{1}{N-k}\cdot[R_{i,l}(\ 0)\cdot R_{j,m}(0) + R_{i,m}(k)\cdot R_{j,l}(\ -k)] \quad (2.14)$$

The derivation of this expression is given in appendix A2.1
Remark: A property of a normally distributed signal is that all moments greater than 2 can be expressed in the mean (first-moment) and the second moment. Thus the second order crossmoment in channel number of the estimated covariance can be expressed in the second order-moment (mean=0), as one can see in formula (2.14).

2.3 The second order crossmoment in channelnumber and in time shift of the estimated covariance

The covariance of the estimated covariance is defined for a vector process as:

$$\mathrm{Cov}(\tilde{R}_{i,j}(k),\tilde{R}_{l,m}(n+k)) = E\{(\tilde{R}_{i,j}(k) - E\{\tilde{R}_{i,j}(k)\})(\tilde{R}_{l,m}(n+k) - E\{\tilde{R}_{l,m}(n+k)\}) \quad (2.15)$$

The covariance of the estimated covariance for a Gaussian process is (see appendix A2.1):

$$\mathrm{Cov}\{\tilde{R}_{i,j}(k).\tilde{R}_{l,m}(n+k)\} = \frac{1}{(N-k)^2}\sum_{s=1}^{N-k-1}\Big[R_{i,l}(s-n).R_{j,m}(s) + R_{i,l}(-s-n).R_{j,m}(-s) + R_{i,m}(s+k).R_{j,l}(s-n-k)) + R_{i,m}(-s+k).R_{j,l}(-s-k-n)\Big] + \frac{1}{N-k}.\Big[R_{i,l}(-n).R_{j,m}(0) + R_{i,m}(k).R_{j,l}(-n-k)\Big] \quad (2.16)$$

We see that that the second order moment in channel number and in time shift of the estimated covariance is also a function of the unknown covariance. The second order crossmoment in time shift for a scalar process will be called covariance of the estimated covariance. The covariance of the estimated covariance of a first order scalar process with a=0.95 and N=800 is depicted in figure 2.8 with the time shift as parameter.

2.4 Simulation results for respectively the covariance and the variance of the estimated covariance

In this section we assume that we can calculate $\tilde{R}(k)$ with a fixed number of N2 samples for each $\tilde{R}(k)$.In this way we do not use all available infomation. We need, see formula (2.7), $(N2+k_{max})$ covariance matrices to compute $\mathrm{var}(\tilde{R}(k))$,consequently we need $2.N2+k_{max}$ samples, with k_{max} the number of variances of the estimated covariances we want to compute. It also means that all N-k are replaced by N2 and consequently the approximated variance of the estimated covariance in area II becomes a constant and is equal to the power of the autocorrelation divided by N2,as we can see by comparing figure 2.2 with figure 2.5. If we use a constant number of samples to compute the covariance and the variance of the estimated covariance we have to replace N-k by N2 respectively in formulas (2.6),(2.8),(2.9),(2.10),(2.11),(2.13),(2.14),(2.17).

In practice we have to substitute the unkown R(k) by $\tilde{R}(k)$ in the right side of the expressions (2.14) and (2.16). If we want to see what is the consequence of this substitution, then we have to simulate a process,that is to generate output data and calculate $\tilde{R}(k)$ with these data . The simulations have been done for a first order, single output process with white Gaussian noise as input signal. Several noise genarators have been tested on their whiteness and their distribution function ;see appendix A2.3.

In figure 2.3 is depicted the estimated covariance of the white noise generator and also the two sided $99.8^0/o(3\sigma)$ confidence interval. The estimated covariance is calculated with equation (2.2) and the corresponding confidence interval with equation (2.8). In both equations (N-k) is replaced by N2. From figure 2.3 we can conclude that the used noise generator is "white". In figure 2.4 is depicted the covariance of the first order process and the estimated covariance of the simulated process. In figure 2.5 and figure 2.6 respectively is depicted, for two different values of δ of a first order process:

A) the variance of the estimated covariance calculated with a limited number of N2=800 samples (2.7).
B) the approximation of the variance of the estimated covariance for area I (2.13a).
C) the approximation of the variance of the estimated covariance for area II (2.13b).
D) the estimated variance of the estimated covariance, by using $\tilde{R}(k)$ instead of R(k) in equation (2.7)
E) the approximation for area I of the variance by using $\tilde{R}(k)$ instead of R(k) in equation (2.9b)
F) the approximation for area II of the variance by using $\tilde{R}(k)$ instead of R(k) in equation (2.10b)

In figure 2.5 δ is chosen relatively large, consequently ℓ is small, that is the contributions of relatively large covariance values are negligible. In contrast for figure 2.6 δ is chosen relatively small and the reliability of the approximations increases. In both cases δ is is chosen and ℓ and ℓb is computed according the criterion (2.11), with ℓb the number of values that is taken into account of $\tilde{R}(k)$. These values are shown in tabel 2.1.

Tabel 2.1

	δ=0.2	δ=0.4
ℓ	32	14
ℓb	269	18

We can notice that the difference between ℓ and ℓb corresponding with δ=0.2 is large. The reason is, if we consider nondeterministic,stable,zero mean processes: $R(k) \underset{k\to\infty}{\to} 0$, that ℓb corresponds with a statistical event. In other words we have chosen $\delta < 3\sqrt{\text{var}(\tilde{R}(k))}$

When we compare figure 2.5 and figure 2.6 we see that respectively curve E and D in figure 2.5 are good approximations of the var($\tilde{R}(k)$), computed with $\tilde{R}(k)$ and that curve E and F in figure 2.6 are good approximations of the var($\tilde{R}(k)$),computed with R(k). This holds,of course, in their defined regions. The latter is caused by the neglect of disturbances in the tail of $\tilde{R}(k)$. When we consider a nondeterministic,zero mean, processes, then we have to choose $\delta > 3.\sqrt{\text{var}(\tilde{R}(k))}$ to get a reliable approximation of var($\tilde{R}(k)$).

Unfortunally we do not know this quantity and this vicious circle is depicted in figure 2.11.

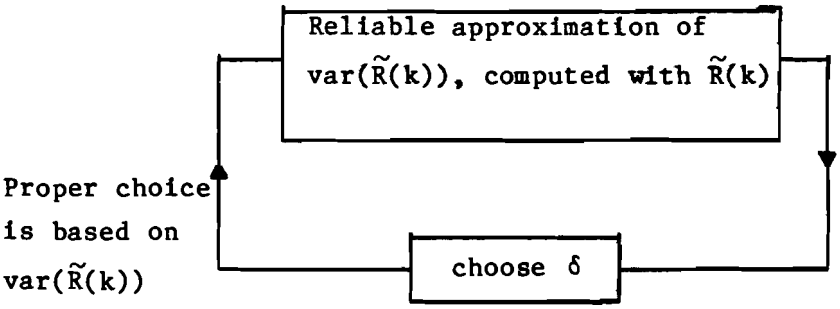


Figure 2.11 Vicious circle

If we calculate the standard-deviation from figure 2.5, $\sqrt{var(\tilde{R}(k))}$, we see that the difference between the estimated covariance and the covariance is less than 3 times the standard deviation (98.8 $^{0}/_{o}$ confidence interval).
In figure 2.8 we see that the errors introduced by calculating the covariance for time shifted signals with a limited number of samples are highly dependent. In other words: the noise on the estimated covariance is far from white.
In figure 2.9 is shown the estimated covariance of the estimated covariance (eq.(2.16),calculated with $\tilde{R}(k)$) for respectively the time shifts n=0 through 5.
The absolute value of the variance of the estimated covariance is for small k greater than for large k for a first order single output process; see figure 2.5,curve A. The relative value of the variance of the estimated covariance is for small k less than for large k; see figure 2.10.
If the stochastic process has finite dimensions, which is a common assumption in linear stochastic identification, we need a finite number of values of the estimated covariance for the determination of the dimension of the process and for factorizing the covariance function. If we use the values with the smallest time shift k and consider nondeterministic processes,then for a first order single output process we have shown that this estimated covariance has the smallest relative variance.

2.5 Conclusion

The noise on the estimated covariance is generally non white. The second order moment in channel number and in time shift of the estimated covariance of a Gaussian process can be expressed in the sum of products of the second order moment, which is generally unknow. If ,however, the number of samples is large enough we can find a reliable approximation of the variance and covariance of the estimated covariance. The accuracy of the estimated covariance is inversely proportional to the square root of the number of samples.

Chapter 3

A NUMERICALLY RELIABLE METHOD FOR SOLVING THE RICCATI EQUATION

3.1 Introduction

In this chapter a modern and robust method is presented for solving the algebraic Riccati-equation. The Riccati-equation appears in the following fields:

a) optimal control: continuous time case
b) optimal control: discrete time case
c) spectral factorization: continuous time case
d) spectral factorization: discrete time case

In the four fields mentioned above one has to solve different types of Riccati-equations. In the literature a lot of attention is paid for solving the algebraic Riccati equation associated with the optimal control problem (see Willems (1971),Vaughan (1970)).Generally speaking we can distinguish two categories of methods for solving these equations respectively an iterative method (see Faurre (1976)) and a direct method.

For modest size problems, say order 100 or less, with dense unstructured coefficient matrices the computional experience is, that a direct method is preferable;cf. Laub (1979). One can distinguish two direct methods for solving the algebraic Riccati-equation (A.R.E). These methods are the classical eigenvector approach and the real Schur approach. The latest approach is introduced by Laub in (1979). The use of the eigenvector approach is often highly unsatisfactory from a numerical point of view, whereas the Schur method is numerically better suitable to compute a basis for a certain subspace of interest in the problem.

In this chapter is given a direct method to solve the A.R.E. associated with the spectral factorization problem and with the covariance factorization problem in the discrete time case. The solution of the A.R.E is obtained by means of the computation of the deflating subspace of the pencil associated with the spectral factorization problem.
An iterative algorithm is derived by Faurre (1976) for computing the solution of the A.R.E. associated with the spectral factorization problem. But there is no indication how fast this algorithm converges. From experience we know, that if the associated pencil has generalized eigenvalues near the unit circle this iterative algorithm converges very slowly (see Van Zee (1981)).
This chapter is organized as follows: In section 3.2 two corresponding factorization problems are described with the dual Riccati-equation. The necessary linear algebra for solving the Riccati-equation by means of a direct method is given in section 3.3 . In section 3.4 the pencil associated with the Riccati-equation is derived . Some remarks on the developed subroutines are given in section 3.6.
Futher I want to thank Messrs. Van Schuppen (Mathematisch Centrum, Amsterdam) and Van Dooren (Philips Research Laboratory, Brussels) for their remarks and contributions.
In this chapter we will use the following abbreviations :

G.E.P. :generalized eigenvalue problem
D.A.R.E. :discrete algebraic Riccati equation

3.2 Two factorization problems with the same corresponding Riccati-equation.

In stochastic identification there are two ways to reach the final goal. The first method is to compute the innovations representation directly from the covariance of the output vector (see chapter 1) and the second method is by means of a

power spectrum matrix $\Phi(z)$ (Anderson (1979)). In both approaches one has to solve the same kind of Riccati-equation, if we assume that the covariance is stationary.
The first approach is explained in chapter 1.Once the realization triplet is known by means of an approximate realization of the covariance, then one has to solve the D.A.R.E. to compute the Kalman-gain and the covariance-matrix of the innovations process. The Riccati-equation which one has to solve is:

$$P = A\,P\,A^T + (L-A\,P\,C^T)\,[R(0)-C\,P\,C^T]^{-1}\,(L^T-C\,P\,A^T) \qquad (3.1)$$

In $\{A,L,C\}$, the realization triplet of the covariance sequence. $\{A,L\}$ is a completely controllable and $\{A,C\}$ is completely observable. The associated equations from which the Riccati-equation is derived,are (see chapter 1,but in the sequel we assume D=I):

$$P - A\,P\,A^T = \Sigma \qquad (3.2)$$

$$L - A\,P\,C^T = S \qquad (3.3)$$

$$R(0) - C\,P\,C^T = M \qquad (3.4)$$

$$\begin{bmatrix} \Sigma & S \\ S^T & M \end{bmatrix} \geq 0 \ , P>0 \qquad (3.5)$$

Which can be written into a matrix equation:

$$\begin{bmatrix} P-A\,P\,A^T & L-A\,P\,C^T \\ L^T-C\,PA^T & R(0)-C\,P\,C^T \end{bmatrix} = \begin{bmatrix} \Sigma & S \\ S^T & M \end{bmatrix} \geq 0 \qquad (3.6)$$

(1.35) implies that we can factorize (3.6) in:

$$\begin{bmatrix} \Sigma & S \\ S^T & M \end{bmatrix} = \begin{bmatrix} KQ^{\frac{1}{2}T} \\ Q^{\frac{1}{2}} \end{bmatrix} \underset{\leftarrow p \rightarrow \leftarrow n \rightarrow}{\begin{bmatrix} Q^{\frac{1}{2}}K^T & Q^{\frac{1}{2}} \end{bmatrix}} \begin{matrix} \uparrow \\ q \\ \downarrow \end{matrix} \qquad (3.7)$$

This leads to the rank condition:

$$\text{rank}\begin{bmatrix} P-A\,P\,A^T & L-A\,P\,C^T \\ L^T-C\,P\,A^T & R(0)-C\,P\,C^T \end{bmatrix} = q \qquad (3.8)$$

That there exists such a positive definite P follows directly from the positive real lemma (see chapter 1).

The second approach , the minimum phase spectral factorization, is a related problem and is treated in this chapter, because in the literature,see van Dooren (1981a,b), this problem is attacked. Moreover it can be helpful in other stochastic realization algorithms and is also used in circuit theory.
In the classical problem the power density matrix $\phi(z)$ is given and one seeks a transfer function W(z) with certain properties (e.g. mimimum phase). Assume that $\Phi(z)$ is given as a matrix of rational functions:

$$\Phi(z) = Z(z) + Z^T(z^{-1}) \qquad (3.9)$$

where Z(z) has only poles in $\|z\|<1$ and with $\{F,G,H,J\}$ is a minimal realization of Z(z):

$$Z(z) = H(z\,I-F)^{-1}\,G + J \qquad (3.10)$$

with $\{F,G\}$ completely controllable and $\{F,H\}$ observable.The spectral factorization of $\Phi(z)$ is:

$$\Phi(z) = W^T(z^{-1})\,W(z) \qquad (3.11)$$

The discrete algebraic Riccati-equation which one has to solve to compute the spectral factor W(z) is

$$P = F^T\,P\,F + (H^T-F^T\,P\,G)\,[J+J^T-G^T\,P\,G]^{-1}(H-G^T\,P\,F) \qquad (3.12)$$

See Denham (1975),Van Dooren (1981a). The equations associated with the Riccati equation writen in matrix form are:

$$\begin{bmatrix} P-F^TPF & H-F^TPG \\ H-G^TPF & J+J^T-G^TPG \end{bmatrix} = \begin{bmatrix} L^T \\ N^T \end{bmatrix} \begin{bmatrix} L & N \end{bmatrix} \quad \begin{matrix} \uparrow \\ q \\ \downarrow \end{matrix}$$

$\leftarrow p \rightarrow \leftarrow q \rightarrow$ (3.13a)

with rank $[\ L\ N\ ] = q$

The Riccati-equation (3.12)is equivalent with the rank condition:

$$\text{rank} \begin{bmatrix} P-F^TPF & H-F^TPG \\ H-G^TPF & J+J^T-G^TPG \end{bmatrix} = q \qquad (3.13b)$$

$$\text{with } W(z) = L\,(zI-F)^{-1}G + N \qquad (3.14)$$

In Denham (1975) it is proven that the smallest positive solution P of respectively (3.12) and (3.13a) , corresponds with a W(z), which has a stable inverse. By comparing the two approaches one sees that equations (3.8) and (3.14) are equivalent if we interchange respectively F by A^T, H by L^T and G by C^T. Thus (3.14) is the dual expression of (3.8).

3.3 A review of linear algebra.

The algebraic Riccati-equation is solved via the computation of the deflating subspace of the associated pencil. Before we derive the pencil in section 3.4 , we will give a few definitions and properties related to this direct method.

Definition I: (see Gantmacher (1959))
Consider the generalized eigenvalue problem:

$$X\,z = \lambda\,Y\,z \qquad (3.15)$$

The generalized eigenvalues are the roots of the generalized characteristic equation :

$$\det(X-\lambda\,Y)=0 \qquad (3.16)$$

For each generalized eigenvalue, the corresponding non-zero vector z satisfying (3.15) will be called a generalized eigenvector. If the multiplicity of the generalized eigenvalue λ is larger than the number of corresponding independent vectors z, then one can define the generalized principal vectors. Since we do not need this concept in the sequel, we do not go into further details.

In this chapter we assume that the dimensions of the matrices X,Y are (2nx2n), with n the dimension of the process.

Definition II

The function $f: C \rightarrow R^{2n \times 2n}$ and

$f(\lambda) = X-\lambda.Y$ is a called pencil.

In section 3.4 will be shown that if Z_{11}, Z_{21} and Λ satisfy:

$$X \begin{bmatrix} Z_{11} \\ Z_{21} \end{bmatrix} = Y \begin{bmatrix} Z_{11} \\ Z_{21} \end{bmatrix} \Lambda \qquad (3.17)$$

where Λ is in Jordan canonical form and the columns of $\begin{bmatrix} Z_{11} \\ Z_{21} \end{bmatrix}$ are the corresponding generalized eigenvectors.
Then $P=Z_{21} Z_{11}^{-1}$ is a solution of the Riccati-equation, with Z_{11}, Z_{21} elements of C^{nxn} and Λ an element of $\bar{C}=\{C^- \mid \|c\| < 1\}$

The construction of the matrices X,Y will also be discussed later in this chapter. The computation of P is reduced to computation of the generalized eigenvectors of the associated G.E.P.
The disadvantages of this method,however,are serious from a numerical point of view in the case of multiple or nearly multiple eigenvalues. The computation of eigenvectors is fraught with difficulties and the eigenvectors themselves are not needed. All that is needed is a basis for the eigenspace associated with the stable eigenvalues of the pencil; see Laub

(1979),van Dooren (1981a). A basis for this stable subspace can be found with the real Schur-vectors. The definition of this real Schur approach will be given later. This real Schur method has the advantage that the number of computations is less than that of the eigenvector approach and that it is numerically very attractive, because it uses only orthogonal transformations. The notation of the eigenvectors must therefore be extended to the notation of the deflating subspace of a regular pencil $X-\lambda Y$.

Definition III

The deflating subspace of the columns of χ of $X-\lambda Y$ is defined as any subspace satisfying: $\dim(Y\chi + X\chi) = \dim\chi$ (3.18)

where dim stands for dimension of the space spanned by the columns of.
The deflating subspace is in the case Y=I equal to the invariant subspace spanned by the columns of χ. The remaining question is how we can compute the deflating subspace. The answer is by the real Schur approach (see van Dooren (1981a) and Pappas (1980))

Theorem I (see Moler and Steward(1973))

There are unitary matrices Q and Z so that $\hat{X}=Q.X.Z$ and $\hat{Y}=Q.Y.Z$ are both upper triangular with X(2nx2n),Y(2nx2n),Q(2nx2n) and Z(2nx2n). The coefficients of these matrices belong to the set of complex numbers C. The form is:

$$Q\ (\lambda\ Y-X)\ Z=\lambda\ \hat{Y}\ -\ \hat{X}$$

$$\lambda\cdot\begin{bmatrix} \hat{y}_{11} & \cdots & \hat{y}_{1\ 2n} \\ \vdots & \ddots & \vdots \\ 0 & & \hat{y}_{2n2n} \end{bmatrix} - \begin{bmatrix} \hat{x}_{11} & \cdots & \hat{x}_{1\ 2n} \\ \vdots & \ddots & \vdots \\ 0 & & \hat{x}_{2n2n} \end{bmatrix} \qquad (3.19)$$

The ratios $\lambda_i=\hat{x}_{ii}/\hat{y}_{ii}$ are called the generalized eigenvalues.

Remark: when Y is singular the degree of the characteristic polynomial is less than n and there does not exists a complete set of eigenvalues for the problem

Theorem II

In the real case, that is $X,Y\epsilon R^{2nx2n}$, the decomposition of (3.19) also exists, but involves complex matrices Q,Z,$\hat{X}$ and $\hat{Y}$,

where the spectrum of $\Lambda(\hat{X},\hat{Y})$ contains complex elements. Under orthogonal real transformations Q and Z, λ Y-X can be transformed to a quasi upper triangular form:

$$Q'(\lambda\ Y-X)\ Z=\lambda\ \hat{Y}-\hat{X}=\lambda\begin{bmatrix} \hat{Y}_{11} & \cdots & \vdots \\ \vdots & \ddots & \vdots \\ 0 & & \hat{Y}_{kk} \end{bmatrix} - \begin{bmatrix} \hat{X}_{11} & \cdots & \vdots \\ \vdots & \ddots & \vdots \\ 0 & & \hat{X}_{kk} \end{bmatrix} \qquad (3.20)$$

where the diagonal pencils $\lambda\ \hat{Y}_{ii}-\hat{Y}_{ii}$ have sizes 1 or 2. If the size is 1, then $\Lambda(\hat{X}_{ii},\hat{Y}_{ii})$ is real (may be infinite). If the size is 2, then $\Lambda(\hat{X}_{ii},\hat{Y}_{ii})$ contains two complex conjugate numbers. (3.19) and (3.20) are called the generalized Schur decomposition.

Assumption :

Let:

$$X=\begin{bmatrix} F - G\,R(0)^{-1}H & 0 \\ -H^T\,R(0)^{-1}H & I \end{bmatrix} \qquad Y=\begin{bmatrix} I & -G\,R(0)^{-1}G^T \\ 0 & F^T-G^T\,R(0)^{-1}\,H \end{bmatrix} \tag{3.21}$$

In the last part of this section we will give a number of properties of the G.E.P. $X\,z=\lambda\,Y\,z$,when X,Y are defined by (3.21)

Theorem III

Consider the generalized eigenvalue problem $X\,z=\lambda\,Y\,z$, X and Y conform (3.21), then if λ is an eigenvalue, then $1/\lambda$ is also an eigenvalue with the same multiplicity (proof see Pappas(1980)).

Conjecture

Non of the eigenvalues of the generalized eigenvalue problem lies on the unit circle (further information can be found in section 3.4,in the the proof AI →AII)

Theorem IV

Consider the generalized eigenvalue problem $X\,z=\lambda\,Y\,z$. If 0 is an eigenvalue with multiplicity r,then there are 2n-r finite eigenvalues. We may say that the r missing eigenvalues are 'infinite' eigenvalues (or reciprocals of 0); see Pappas (1980).

We can arrange the eigenvalues of our problem in the following way :

$$\underbrace{0,0,0,\ldots 0}_{r},\underbrace{\lambda_{r+1},\ldots\ldots\lambda_n}_{n-r},\underbrace{\frac{1}{\lambda_n}\ldots\ldots\frac{1}{\lambda_{r+1}}}_{n-r},\underbrace{\infty,\infty,\infty\ldots\ldots\infty,\infty}_{r}$$

If X respectively Y are given by (3.21), then we have n stable eigenvalues, and we are able to compute an n-dimensional stable subspace.

3.4 Main conjecture

In this section we will give an important conjecture,which is fundamental for computing the solution of the Algebraic Riccati Equation (3.1) and the dual equation (3.12) by means of a direct method. This method to be used can be either an eigenvector approach or a Real Schur approach.

Main conjecture

Let the power density matrix of a rational time series be:

$$\Phi(z) = H\,(z\,I-F)^{-1}\,G + G^T\,(z^{-1}\,I-F^T)^{-1}H^T + R(0) \qquad (3.22)$$

$$R(0) = J + J^T > 0$$

with $\{F,G\}$ controllable and $\{F,H\}$ observable.

Then there exists a P element of R^{nxn} such that:

AI:
$$\begin{cases} 1)\ P=P^T>0 & (3.23)\\ 2)\ R(0)-G^T P G>0 & (3.24)\\ 3)\ P=F^TP F+(H^T-F^TP G)(R(0)-G^TP G)^{-1}(H^T-F^TP G)^T & (3.25)\end{cases}$$

if and only if there exist $Z_{11}, Z_{21}, \Lambda \in C^{nxn}$ and Λ in Jordan canonical form, such that:

AII:
$$\begin{cases} 1)\ Z_{11} \text{ is non singular} & (3.26)\\ 2)\ R(0)-G^T (Z_{21} Z_{11}^{-1}) G \text{ is non singular} & (3.27)\\ 3)\ \begin{bmatrix} F - G R(0)^{-1}H & 0\\ -H^TR(0)^{-1}H & I\end{bmatrix}\begin{bmatrix} Z_{11}\\ Z_{21}\end{bmatrix}=\begin{bmatrix} I & -GR(0)^{-1}G^T\\ 0 & (F-GR(0)^{-1}H)^T\end{bmatrix}\begin{bmatrix} Z_{11}\\ Z_{21}\end{bmatrix}\Lambda \\ \text{and } P= Z_{21} Z_{11}^{-1} & (3.28)\end{cases}$$

B: $P^- = Z_{21} Z_{11}^{-1}$ iff all eigenvalues associated with Λ are in C^-.

P^-: the minimal positive definite solution of the D.A.R.E.

$P^+ = Z_{21} Z_{11}^{-1}$ iff all eigenvalues associated with Λ are in $C^+=\{c^+ \mid \|c\| >1\}$

P^+: the maximum positive definite solution of the D.A.R.E.

Remark: An n-th order D.A.R.E associated with the discrete time optimal control problem has $\begin{bmatrix} 2n\\ n\end{bmatrix}$ solutions (see Laub (1979)), which expresses all possible combinations of n eigenvalues out of the total set of 2n, and one is interested in the unique positive definite solution. The D.A.R.E. associated with the time discrete spectral factorization problem has a set of positive definite solutions, and all possible combinations of the eigenvalues give a positive definite solution.

Proof AI → AII:

1-Let $R(0)>0$, $P=P^T$, $R(0)-G^T P\ G>0$ and $F_1= F-G\ R(0)^{-1}\ H$ (3.29)

Then expression (3.27) is equal to :

$$P=F_1^T\ P\ F_1+H^T(J+J^T)^{-1}H+F_1^T\ P\ G\ (J+J^T-G^T\ P\ G)^{-1}G^T\ P\ F_1 \quad (3.30)$$

The proof is given in appendix 3.1.

2- Let $F_2= F_1+ G\ (J+J^T-G^T\ P\ G)^{-1}\ G^T\ P\ F_1$

There exists a Z_{11} element of C^{nxn} non singular and a Λ element of C^{nxn} in Jordan canonical form, such that

$Z_{11}^{-1}F_2\ Z_{11}= \Lambda$

Let $Z_{21}=P\ Z_{11}$, hence $P=Z_{21}\ Z_{11}^{-1}$.

3- $F_1^T\ P\ F_2= F_1^T\ P\ F_1+ F_1^T\ P\ G\ (J+J^T-G^TP\ G)^{-1}G^TP\ F_1=P -H\ (J+J^T)^{-1}H$

referring to the second n-rows of (3.28):

$$F_1^T\ Z_{21}\ \Lambda = F_1^T\ Z_{21}\ Z_{11}^{-1}\ F_2\ Z_{11}=F_1^T\ P\ \ F_2\ Z_{11}=$$

$$= P\ Z_{11}-H^T(J^T+J\ )^{-1}H\ Z_{11}=Z_{21}-H^T\ (J^T+J\ )^{-1}H\ Z_{11}$$

referring to the first n-rows of (3.28):

$$F_1\ Z_{11}-Z_{11}\ \Lambda = (F_1-Z_{11}\ \Lambda\ Z_{11}^{-1}\ )\ Z_{11}=(F_1-\ F_2)\ Z_{11}=$$

$$=-G\ (J+J^T-G\ \ P\ G)^{-1}G^T\ P\ F_1\ Z_{11}\ \ =$$

$$=-G\ (J+J^T)^{-1}(J+J^T-G^TP\ G+G^TP\ G)\ (J+J^T-G^TP\ G)^{-1}G^TP\ F_1\ Z_{11}=$$

$$=-G\ (J+J^T)^{-1}G^TP\ F_1\ Z_{11}-G(J+J^T)^{-1}G^TPG(J+J^T\ -G\ P\ G)^{-1}G^TP\ F_1\ Z_{11}=$$

$$=-G\ (J+J^T)^{-1}\ G^TP(F_1+G\ (J+J^T-G^TP\ G)^{-1}G^TP\ F_1)\ Z_{11}=$$

$$=-G\ (J+J^T)^{-1}G^TP\ F_2\ Z_{11}=-G\ (J+J^T)^{-1}G^TZ_{21}\ Z_{11}^{-1}\ Z_{11}\ \Lambda\ Z_{11}^{-1}\ Z_{11}=$$

$$=-G\ (J+J^T)^{-1}\ G^TZ_{21}\ \Lambda$$

This proof implies AI → AII.

"Proof" AII → AI

4-Assume Z_{11}, Z_{21} and Λ with the indicated properties exist.

Then: $F_1 Z_{11}=Z_{11} \Lambda-G (J+J^T)^{-1} G^T Z_{21} \Lambda$ (from 3.28)

$-H^T (J+J^T)^{-1} H Z_{11}+Z_{21}=F_1^T Z_2 \Lambda$ (from 3.28)

Let $P=Z_{21} Z_{11}^{-1}$

And $P -F_1^T P F_1-H (J+J^T) H^T-F_1^T P G(J+J^T-G^T P G)^{-1}G^T P F_1 =$

$=Z_{21} Z_{11}^{-1}-F_1^T Z_{21} Z_{11}^{-1}(Z_{11} \Lambda Z_{11}^{-1}-G (J+J)^{-1}G^T Z_{21} \Lambda Z_{11}^{-1})$

$-F_1^T Z_{21} Z_{11}^{-1} G (J+J^T-G^T Z_{21} Z_{11}^{-1} G)^{-1}G^T Z_{21} Z_{11}^{-1} F_1$

$-Z_{21} Z_{11}^{-1}- F_1 Z_{21} \Lambda Z_{11}^{-1}$

by the above equalities.

$= F_1^T Z_{21} Z_{11}^{-1}G (J+J^T-G^T Z_{21} Z_{11}^{-1} G)^{-1}.$

$((J+J^T-G^T Z_{21} Z_{11}^{-1} G (J+J^T)^{-1} G^T Z_{21} \Lambda Z_{11}^{-1}-G^T Z_{21} Z_{11}^{-1} F_1)=$

$= F_1^T Z_{21} Z_{11} G (J+J -G^T Z_{21} Z_{11} G)^{-1}.$

$(G^T Z_{21} \Lambda Z_{11}^{-1}-G^T Z_{21} Z_{11}^{-1} G(J+J^T)^{-1}G^T Z_{21} \Lambda Z_{11}^{-1}-$

$G^T Z_{21} Z_{11}^{-1} (Z_{11} \Lambda Z_{11}^{-1}-G (J+J)^{-1}G^T Z_{21} \Lambda Z_{11}^{-1})) =0$

This shows that:

$P = F_1^T P F_1+H^T (J+J^T)^{-1} H +F_1^T P G(J+J^T-G^T P G)^{-1} G P F_1$

5- $P = P^T>0$, can not yet be proven

Remark:The D.A.R.E associated with the optimal control problem, see Pappas (1980), is simular with the D.A.R.E. assocated with the spectral factorization problem, except that there appears a minus sign in the factor which has to be inverted. This minus sign causes that ,instead of one unique positive solution, there is a set of positive definite solutions. The consequences of this minus sign are :in the first place we cannot use the proof of Pappas (1980) to show that $P=P^T>0$, because in the proof the property is used that the sum of two nonnegative matrices is nonnegative. In the spectral factorization case, however, there appears a minus sign and the difference if two nonnegative matrices can be negative

In Pappas (1980) there is also proven that none of the eigenvalues lies on the unit circle, but we can not use this proof on the same arguments as mentioned above.

6- 4,5 and the assumption that Z_{11}, Z_{21}, Λ satisfy 1,2,3 implies equations (3.25),(3.26) and (3.27).

"Proof B":

This conjecture is based on simular results of the A.R.E. associated with the continuous time spectral factorization problem. In Anderson (1973) it is suggested that this can be proven by using the bi-linear transformation from the s-domain to the z-domain, but it is not evident that all properties, which are valid for the continuous time case are invariant of this bi-linear transformation.

Consider equation (3.30) and exchange Z_{11} and Z_{21} and postmultiply by Λ^{-1} (Λ can have eigenvalues zero and so Λ^{-1} eigenvalues ∞) then we get:

$$\begin{bmatrix} I & -H^T(R(0)H \\ 0 & F-GR(0)^{-1}H \end{bmatrix} \begin{bmatrix} Z_{21} \\ Z_{11} \end{bmatrix} \Lambda^{-1} = \begin{bmatrix} (F-GR(0)^{-1}H)^T & 0 \\ -GR(0)^{-1}G^T & I \end{bmatrix} \begin{bmatrix} Z_{21} \\ Z_{11} \end{bmatrix} \tag{3.31}$$

By comparing equation (3.28) and (3.31) one sees that these equations are dual.This means that we may replace F by F^T, G by H^T and H by G^T.

If P is the solution of (3.28) then P^{-1} is the solution of (3.31),this solution will be called Π. The minimal solution of (3.28) P^- corresponds with the maximal solution of (3.31) $\Pi^+=(P^-)^{-1}$. Analoguosly one can find the minimal solution with $P^-=(\Pi^+)^{-1}$.

Thus if we replace Λ^{-1} by Λ_1, Z_{21} by W_{11} and Z_{11} by W_{21} and A by F, G by L and H by C in (3.31) then we get:

$$\begin{bmatrix} I & -C^T R(0)^{-1} C \\ 0 & A^T - L R(0)^{-1} C \end{bmatrix} \begin{bmatrix} W_{11} \\ W_{21} \end{bmatrix} \Lambda_1 = \begin{bmatrix} A^T - C R(0)^{-1} L^T & \bar{0} \\ -LR(0)^{-1} L^T & \bar{I} \end{bmatrix} \begin{bmatrix} W_{11} \\ W_{21} \end{bmatrix} \tag{3.32}$$

and $\bar{\Pi} = W_{21} \cdot W_{11}^{-1}$, Λ_1 corresponding to the stable eigenvalues. With Π the minimal solution of the D.A.R.E. associated with the covariance factorization problem (see equation (3.1)).

3.5 Implementations of the algorithm

The implementation of the real Schur vector approach consists of the computation of orthogonal matrices Q(2nx2n) and Z(2nx2n), which transform the matrix X to a real Schur form and Y to upper triangular form, in such a way that the diagonal blocks corresponding to the stable eigenvalues are in the upper left quarter of the matrices. The following sequence of subroutines:

QZHES, QZIT and QZVAL (see EISPACK)

reduces X to a real Schur form and Y to upper triangular form, but the order in which the eigenvalues appear is arbitrary. The separate stages are (see Moler and Stewart (1973)):

- QZHES: is a generalization of the Householder reduction of a single matrix to Hessenberg form, X is reduced to upper Hessenberg form and at the same time Y is reduced to upper triangular form.
- QZIT is a generalization of the Francis implicit double shift QR algorithm, X is reduced to quasi-triangular form while the triangular form of Y is maintained.
- QZVAL the quasi-triangular matrix is effectively reduced to triangular form and the eigenvalues extracted

The subroutine ORDER reorders the diagonal blocks in the decomposition, this subroutine is written by van Dooren,see van Dooren (1981). The sequence of subroutines:

QZHES,QZIT,QZVAL and ORDER

gives the desired result, that means the blocks corresponding with the stable eigenvalues are in the left quarter of the matrices X and Y.

Finally we have to solve the n-th order linear matrix equation:

$$P=Z_{21}\,Z_{11}^{-1} \quad \text{with } Z=\begin{bmatrix} Z_{11} & Z_{12} \\ Z_{21} & Z_{22} \end{bmatrix}$$

and P,Z_{11},Z_{12},Z_{21} and $Z_{22} \in R^{nxn}$

Any good lineair equation solver can be used for this purpose (see van Gent (1983)).

3.6 Conclusion

In this chapter an algorithm is discussed for solving the discrete Algebraic Riccati Equation, by means of a numerical stable method, if the matrices which have to be inverted are not or nearly singular. If these matrices are badly conditioned seriously numerical difficulties occur. There is presented a way in van Dooren (1981) to circumvent this by a trick. Unfortunately we have not been able to proof all aspects of the algorithm, but we expect is that the missing aspects can be proven.

The direct method of solving the D.A.R.E. associated with the covariance factorization problem by means of computing a stable subspace is used in the stochastic identification. The results obtained all have strengthened the conjecture given in section 3.4.

Chapter 4

STOCHASIC IDENTIFICATION

4.1 Introduction

In chapter 1 is discussed the stochastic realization problem by means of the innovations approach, based on the exact knowledge of the covariance, and in this chapter we will investigate the reliability of the realization algorithm, when the covariance has to be estimated.
The algorithm is based on the results given priviously, presented in chapter 1 and 3, and is therefore suited to identify time invariant, Gaussian distributed, rational time series. In section 4.2.1 the block diagram of the algorithm and some essential information of the algorithm is presented. To check the reliability of this algorithm a number of measures are introduced in section 4.2.2. Before we test the algorithm in the noisy case,due to the fact that the covariance has to be estimated, we have to make some remarks about the singular value decomposition as a tool to determine the order of the process and to obtain an approximate realization of the covariance sequence. These remarks are given in section 4.3.

We test the reliability of the algorithm with the measures introduced in section 4.2.2 and a number of processes,which are defined in section 4.4. In the first place, in section 4.5, the numerical accuracy and the stability of the algorithm is tested in the case where exact knowledge about the covariance is available. Secondly for the defined processes we generate vector time series $\{y(t)\}$ and afterwards we offer these time series to the identification algorithm. The advantage of simulating a process is that all assumed conditions are under control and can be fulfilled. Besides a comparison between the computed quantities and the known quantities are possible and statements on the reliability of the algorithm are possible, section 4.6.3.

The final goal of the development of this algorithm is to identify physical processes. In section 4.7, the algorithm is used to identify dunes in a flume of running water. A detailed process discription is given in section 4.7.1 and the identification results are presented and discussed in section 4.7.2 and 4.7.3 respectively.

4.2 Stochastic identification algorithm

4.2.1 Block diagram and essential information of the algorithm

The stochastic identification algorithm used is based on the previous theoretical results presented in chapter 1 and 3. The block diagram is depicted in figure 4.1a and in this section the separated blocks (subroutines) are stated.
In the stochastic identification problem the covariance is not avialable and we have to estimate the covariance of the vector process under study. We have to restrict ourselves to 3 channels and a maximum process dimension 6. These restrictions are necessary, because the algorithm is implemented on a PDP 11/60 computer with an available memory space for each user of 32k words.
For each time shift of the estimated covariance matrix there a fixed number of samples is used, i.e.:

$$R(k) = \frac{1}{N} \sum_{i=1}^{N} y(i+k).y^T(i) \qquad (2.1)$$

The exact covariance, of course only known in simulations, is computed by means of solving the time discrete Lyapunov equation:

$$P - A.P.A^T = B.\Sigma.B^T \qquad (1.8)$$

and the sequence of covariance matrices is given by:

$$R(k) = C.A^{k-1}.L \quad k>0 \qquad (1.11)$$

$$L = A.P.C^T + B.S.D^T \qquad (1.12)$$

The Lyapunov equation is solved by means of a direct method, i.e. a similar method as used to solve the Discrete Algebraic Riccati equation. However the pencil associated with the Lyapunov equation is less complicated. The associated pencil with the Lyapunov equation is (see appendix 4.2):

$$\lambda.\begin{bmatrix} I & 0 \\ 0 & A \end{bmatrix} - \begin{bmatrix} A^T & 0 \\ B\Sigma B^T & I \end{bmatrix} = \lambda Y - X \qquad (4.1)$$

with $P=Z_{21}.Z_{11}^{-1}$, where $\begin{bmatrix} Z_{11} \\ Z_{21} \end{bmatrix}$ is a basis associated with the stable subspace, i.e. the subspace corresponding with the stable generalized eigenvalues.

The exact and the estimated covariance sequence is offered to the subroutine, which computes the approximate realization of the sequence of estimated covariances. The realization can be computed by means of the Hankel or Page approach and the method is based, in both approaches, on the Ho-Kalman algorithm modified by Hajdasinski and Damen (Hajdasinski and Damen (1982)).

Remark: the maximum number of covariance matrices which is used to construct the Hankel or Page matrix is NOM=24; because the realization algorithm uses a shifted Hankel or Page matrix to compute the realization triplet $\{\hat{A},\hat{L},\hat{C}\}$,the total number of available covariance matrices must be NOM+1.

The last step to identify the vector process is to solve the Riccati equation, see chapter 3, and with the solution of this equation we can compute the Kalman gain K and the covariance of the innovations process Q. Finally we have the innovations representation of the available time series. The stochastic identification algorithm can be seen as a mapping of $\{y(k)\}$ to the innovations representation;see figure 4.2

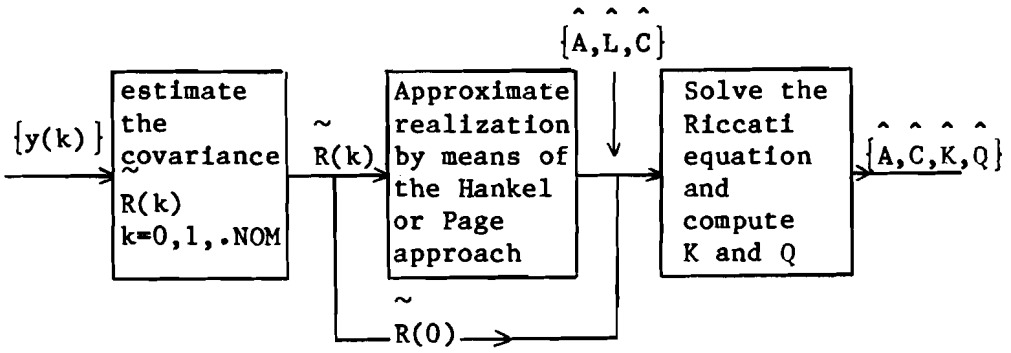


Figure 4.1b Stochastic identification algorithm

The algorithm admits the following choices:

- number of channels offered to the identification algorithm
- the Hankel or Page approach:
  - Hankel : number of used covariance matrices to construct the Hankel matrix.
  - Page : number of used covariance matrices to to construct the Page matrix and the block dimension of the Page matrix respectively η, μ.(see chapter 1).
- the order of the process, based on the singular value decomposition of the Hankel or Page matrix (see section 4.3).

One of the properties of the innovations representations is, that it is causally invertible (see chapter 1). That means that if we consider the vector y(t) as input of the causally invertible filter, then the output is a white noise process (innovations process e(t)), see figure 4.2.

Figure 4.2

This caussaly invertible filter is also called G.I.F.: General Innovations Filter, see a.o. Wouters and Gevers (1978). This G.I.F. is introduced to compute a measure of quality of the

identification algorithm, as will be shown in the next subsection.

4.2.2. Test measures

To evaluate the results of our identification algorithm a number of quantities are introduced. We define the quantities in the order of succession in the identification algorithm and if necessary we give the purpose of introducing this measure. The defined measures are :

-1 Relative error in the the covariance, time shift 0

$$RECO = \frac{\| \tilde{R}(0) - R(0) \|_E^2}{\| R(0) \|_E^2} \qquad (4.2)$$

This measure is of course only useful in simulations. This measure affects directly the accuracy of the computed innovations representations, because $\tilde{R}(0)$ is used directly in the D.A.R.E. to compute the Kalman gain and the covariance of the innovations.

-2 Relative error in the input covariance sequence

$$ERRO = \frac{\sum_{k=1}^{NOM} \| \tilde{R}(k) - R(k) \|_E^2}{\sum_{k=1}^{NOM} \| R(k) \|_E^2} \qquad (4.3)$$

where $\{R(k)\}$, k=1,2,..NOM is the sequence of exact covariance matrices and $\{\tilde{R}(k)\}$ k=1,2,..NOM the sequence of estimated covariance matrices.

-3 Relative error in the realized sequence of covariance matrices

$$ERR1 = \frac{\sum_{k=1}^{NOM} \|\hat{R}(k) - R(k)\|_E^2}{\sum_{k=1}^{NOM} \|R(k)\|_E^2} \qquad (4.4)$$

where $\{\hat{R}(k)\}$, k=1,2,...NOM is the realized sequence of covariance matrices $R(k) = \hat{C}.\hat{A}.^{k-1}.\hat{L}$, k>0

$\{\hat{A},\hat{L},\hat{C}\}$ are computed by the realization algorithm. This is a measure for the ability of the realization algorithm to fit a sequence of estimated covariance matrices to the exact sequence of covariance matrices.

When R(k) is not available R(k) is replaced by $\tilde{R}(k)$ in (4.4).

-4 The relative accuracy of the solution of the discrete algebraic Riccati equation

$$RARE = \frac{\|P - P_R\|_E^2}{\|P_L\|_E^2} \qquad (4.5)$$

where P_L is the solution of the D.A.R.E (equation 3.1) and P_R the right side of the D.A.R.E computed with P_L.
This measure is introduced to check the accuracy of the direct method for solving the D.A.R.E.

-5 The whiteness of the output vector of the G.I.F.(see figure 4.2)

$$WINV = \frac{\sum_{i=1}^{NOM} \|R_e(k)\|_E^2}{NOM \; \|R_e(0)\|_E^2} \qquad (4.6)$$

$$\text{with } R_e(k) = \frac{1}{N} . \sum_{i=1}^{N} e(i+k).e^T(i) \; ,$$

e(i) the output of the G.I.F.

This test quantity is constructed to have an indication of the whiteness of the output vector of the G.I.F. The whiteness of input noise and output noise is computed in a simular way and the results are given in section 4.5. These are computed to have a reference.

4.3 Singular value decomposition as a tool to determine the order of the process and the approximate realization of the estimated covariance

In contrast with chapter 1, where the exact covariance is available,we have to estimate the covariance sequence. This sequence appears to be disturbed with coloured noise, see chapter 2. Consequently the infinite Hankel and Page matrix have no finite dimension, as in the case when exact knowledge about the covariance is available. However our goal is to find a finite realization of the estimated covariance. To determine this finite dimension and the approximate realization we use the singular value decomposition of the Hankel or Page matrix. In the case where exact knowledge about the covariance is available the singular value decompostion of the Hankel matrix is (see Hajdasinski and Damen (1979) and Van den Hof (1982)):

$$H = W \, \Sigma_{svd} \, V^T \tag{4.7a}$$

$$H = W \begin{bmatrix} \Sigma_{svd,k} & 0 \\ 0 & 0 \end{bmatrix} V^T \tag{4.7b}$$

where H is a $g \times \ell$ -matrix

$\rho = \min(g, \ell)$

W: is a $g \times \rho$ -matrix, consisting of ρ orthonormal eigenvectors of HH^T.

V: is a $\ell \times \rho$ -matrix, consisting of ρ orthonormal eigenvectors of H^TH.

Σ_{svd}: is a $\rho x \rho$ diagonal matrix: $diag(\delta_1, \delta_2, \delta_3, \ldots \delta_\rho)$

and $\delta_1 > \delta_2 > \delta_3 > \ldots > \delta_\rho > 0$, and δ_i is the square root of an eigenvalue of $H^T H$, and is called singular value

Because W and V are matrices consisting of orthonormal vectors and because Σ_{svd} is a diagonal matrix, we can write:

$$\text{rank } H = \text{rank } \Sigma_{svd} \leq \rho \tag{4.8}$$

This means that the rank of Σ_{svd} is a direct measure of the rank of H. Thus if H contains sufficient covariance matrices then k=n, with k the number of nonzero singular values of the Hankel matrix.

In the case where we have to estimate the covariance,the singular value decompostion of the Hankel matrix is:

$$\tilde{H} = \tilde{W} \begin{bmatrix} \tilde{\delta}_1 & & & \\ & \tilde{\delta}_2 & & \\ & & \ddots & \\ & & & \tilde{\delta}_\rho \end{bmatrix} \tilde{V}^T \tag{4.9}$$

The singular values $\tilde{\delta}_{k+1} > \ldots > \tilde{\delta}_\rho > 0$ are caused by the noise, because in absence of the noise these values would be equal to zero. The singular values are depicted in figure 4.3, when exact knowledge of the covariance is available and in a noisy situation for a third order process.

The feasibility to distinct the singular values caused by noise from the singular values associated with the process depends on the noise level. If the noise level is sufficiently small, then a reliable order determination is possible,as there may be a clear jump in the singular values sequence

The approximate realization,based on the chosen rank k, is suboptimal. That is matrix H_k minimizes $\|H-H_k\|_E$, with H_k a general matrix consructed with singular values ,but this matrix lacks the Hankel structure.The singular value decomposition is also used to determine the approximate realization $\{\hat{A},\hat{L},\hat{C}\}$, but is not treated in this report and can be found in Hajdasinski and Damen (1979).

Similar results can be given for the Page approach. There are important differences compared with the Hankel approach. These differences are:

- the weighting of the covariance matrices.
- the properties of the approximate realization $\{\hat{A},\hat{L},\hat{C}\}$

Theoretical backgrounds and additional information can be found in Van den Hof (1982).

4.4 Definition of test systems

In order to test the accuracy and the stability of respectively the realization- and the identification algorithm we have to define a number of systems. We start to define a number of MO-systems:

MO-systems: (MO: multiple output)

We consider four MO-systems with 3 inputs and 2 outputs with dimension 4. The B,C and D matrices of the four MO systems are chosen to be the same. The B and C matrix are chosen according to Damen and Hajdasinski (1982).

$$B=\begin{bmatrix} 1. & 0. & 1. \\ 1. & .5 & .5 \\ 0. & 1. & -.5 \\ 0. & .5 & -1. \end{bmatrix} \quad C=\begin{bmatrix} .5 & 0. & 1. & 1. \\ 1. & -.5 & .5 & -.5 \end{bmatrix} \quad D=\begin{bmatrix} 1. & 0. \\ 0. & 1. \end{bmatrix} \quad * \qquad (4.10)$$

To be able to have a clear view on the eigenvalues of the respective systems, the A-matrices are chosen in a diagonal form:

SYS0: A0 = diag(0.4,0.3,0.2,0.1) (4.11a)

SYS1: A1 = diag(0.6,0.5,0.9,0.7) (4.11b)

SYS2: A2 = diag(0.9,0.8,0.85,0.3) (4.11c)

SYS3: A3 = diag(0.2,0.1,0.9,0.15) (4.11d)

* The channel dependency is included in the M_2 matrix instead of the D.

In every system a specific aspect of the capability of the realization (identification) algorithm is tested: a relatively fast decreasing impuls response (SYS0), a slowly decreasing impuls response (SYS1) and 3 closely postioned large eigenvalues (SYS2) and 3 small closely positioned eigenvalues (SYS3).

We have chosen for all MO systems as source,respectively channel independent Gausssian white noise $\{\Sigma_1,S_1,M_1\}$ and channel interdependent Gaussian white noise $\{\Sigma_2,S_2,M_2\}$*. The covariance matrices of the channel independent input- and output noise are $\Sigma_1=I,S_1=0$ and $M_1=I$. The covariance matrices of the channel interdependent input- output noise are:

$$\Sigma_2=\begin{bmatrix}1. & 0.5 & 0.5\\ 0.5 & 2. & 0.5\\ 0.5 & 0.5 & 3.0\end{bmatrix} \quad S_2=\begin{bmatrix}0.5 & 0.5\\ 0.5 & 0.5\\ 0.5 & 0.5\end{bmatrix} \quad M_2=\begin{bmatrix}1.0 & 0.5\\ 0.5 & 2.0\end{bmatrix}$$

$$\text{with} \quad \begin{bmatrix}\Sigma_2 & S_2\\ S_2 & M_2\end{bmatrix} > 0$$

The systems which are driven with the non-channel interdependent noise are denoted by SYS.-1 and with the channel interdependent noise by SYS.-2.

SO-systems (SO: single output)

There are used three SO-systems respectively SYS4,SYS5 and SYS6. SYS4 has dimension 1 and SYS5,SYS6 have dimension 2. The corresponding matrices are:

SYS4: $A=0.9$ $B=1.0$ $C=1.0$ $D=1.0$

SYS5: $A=\begin{bmatrix} 0.8 & -0.5 \\ 0.5 & 0.8 \end{bmatrix}$ $B=\begin{bmatrix} 1.0 \\ 0.25 \end{bmatrix}$ $C=\begin{bmatrix} 0.5 & 1.0 \end{bmatrix}$ $D=1.0$

the eigenvalues are: $0.8 \pm j0.5$

SYS6: $A=\begin{bmatrix} 1.35 & -0.45 \\ 1.35 & -0.45 \end{bmatrix}$ $B=\begin{bmatrix} 1.0 \\ 0.25 \end{bmatrix}$ $C=\begin{bmatrix} 0.5 & 1.0 \end{bmatrix}$ $D=1.0$

the eigenvalues are: 0.9 and 0.0

The covariance of the used input and output noise is chosen channel independent $\{\Sigma_3=1, S_3=0, M_3=1\}$. SYS6 is a special case, namely it has eigenvalues 0.9 and 0.0. The eigenvalue 0.0 corresponds with a pure time delay and the identification algorithm should ignore this pole. SYS4 and SYS5 are introduced to test the correct execution of the algorithm for SO processes.

4.5 The accuracy of the algorithm

In this section we offer an exact* sequence of covariance matrices R(k),k=1,2,...NOM, to the realization algorithm and compute a realization by means of Hankel or Page approach. From $\{\bar{A},\bar{L},\bar{C}\}$ and the exact* R(0) we compute the innovations representation. From this innovations representation we calculate the covariance sequence, $\bar{R}(k)$,k=1,2,..NOM,and compare this sequence with the original exact sequence. The difference between the covariance sequence,$\bar{R}(k)$,k=1,2,..NOM, calculated from the innovations representation and the exact covariance sequence R(k),k=1,2,..NOM, is a measure of the accuracy realization algorithm. The relative difference between these sequences is denoted by ERR2. ERR2 is:

* within an accuracy of $2\ 10^{-8}$

$$ERR2 = \frac{\| \sum_{i=1}^{NOM} \bar{R}(k) - R(k) \|_E^2}{\| \sum_{i=1}^{NOM} R(k) \|_E^2} \qquad (4.12)$$

ERR2 is computed for the defined MO- and SO-systems and given in tabel 4.1 , driven with channel independent noise and channel dependent noise.
Considering the results, we may say that generally the Hankel approach for finding a realization $\bar{R}(k)$,k=1,..NOM,is more accurate than the Page approach. That means ERR2 associated with the Hankel approach is smaller than ERR2 associated with the Page approach. This result is noticed previously by Van de Hof (1982),that is in noise "free sequences" the Hankel approach is generally preferable.
Finally we have performed the next experiment to get an indication of the numerical stability of the algorithm.

Experiment

We have computed a sequence of covariance matrices of SYS2, R(k), k=0,1,2...NOM, driven with channel independent noise, and compute a realization triplet,denoted by $\{\hat{A}_1,\hat{L}_1,\hat{C}_1\}$. R(0) and $\{\hat{A}_1,\hat{L}_1,\hat{C}_1\}$ are used to determine the innovations representation $\{\hat{A}_1,\hat{K}_1,\hat{C}_1,\hat{Q}_1\}$. With this innovations representation we are able to compute again a sequence of covariance matrices, denoted by $\bar{R}_1(k)$,k=0,1,.,NOM . The next step is to compute a realization of $\bar{R}_1(k)$, denoted by $\{\hat{A}_2,\hat{L}_2,\hat{C}_2\}$ and with $R_2(0)$ to compute the innovations $\{\hat{A}_2,\hat{C}_2,\hat{K}_2,\hat{Q}_2\}$ and again $\bar{R}_2(k)$. The experiment is

schematically depicted in figure 4.4. When we consider the results, we can see that there is a loss of one digit in the Page approach and a loss of two digits in the Hankel approach in 10 iterations. The executed experiment gives confidence in the reliability of the algorithm. However from this relatively small number of iterations and only one numerical example we are not able to conclude whether this algorithm is numerically stable or not. By stability is meant that the measure ERR2 is non globally increasing for an increasing number of iterations.

4.6 The reliability of the stochastic identification algorithm

4.6.1 Simulation results

In this section the reliability of the algorithm is checked with the previously described systems, that means by the four MO-systems and the three SO systems. With a specific system we generate six time series of respectively N=300,1000 and 10000 samples and each time series is caused respectively by channel independent and by channel dependent noise.

Gaussian white noise generators

The maximum number of noise channels is 6, respectively maximum 3 inputnoise and 3 output noise channels. These Gaussian white noise generators are 6 "GAUSS1" noise generators (see appendix 2.3) with 12 different initial values. The used values are shown in tabel 4.3.

The standard deviation of these Gaussian noise generators is 1.0 and they are channel independent. The desired channel dependency is achieved by premultiplying the original u_l and v_l as follows:

$$\begin{bmatrix} u_1(t) \\ v_1(t) \end{bmatrix} \text{ by R with } E\left\{\begin{bmatrix} u_1(t) \\ v_1(t) \end{bmatrix}\begin{bmatrix} u_1(t) & v_1(t) \end{bmatrix}\right\}=I$$

$$\text{with } R.R^T = \begin{bmatrix} \Sigma & S \\ S^T & M \end{bmatrix} > 0. \text{ and } E\left\{\begin{bmatrix} u(t) \\ v(t) \end{bmatrix}.\begin{bmatrix} u^T(t) & v^T(t) \end{bmatrix}\right\}=\begin{bmatrix} \Sigma & S \\ S^T & M \end{bmatrix}$$

We will give a measure of the whiteness of these noise generators (GAUSS1) by means of a measure defined in the previous section. Let:

$$w(t) = \begin{bmatrix} u(t) \\ v(t) \end{bmatrix}$$

We substitude this vector w(t) in equation (4.6) and compute WINV. WINV is depicted respectively for channel interdependent and channel independent noise in tabel 4.4.

Choices made in the identification algorithm

In the identification of all systems we make use of two different approaches, namely the Hankel and Page approach. The Hankel and Page matrix are constructed with 24(+1) covariance matrices,(+1 covariance matrix for the shifted version of the both matrices). The block dimension of the Page matrix is chosen 6 by 4. The dimension of the process is chosen according the previously known dimension of the simulated processes.

Results

The results are depicted in tabel 4.5 a through ℓ for MO- and SO-processes. For SYS2,driven with channel independent noise,is depicted the sequence of:

- exact covariance matrices R(k)
- estimated covariance matrices $\tilde{R}(k)$
- realized sequence of covariance matrices $\hat{R}(k)$, computed with $\{\hat{A},\hat{L},\hat{C}\}$,by means of the Page approach for N=10000

in figure 4.6a,b,c,d

4.6.2 Discussion of the results

In this subsection we will give a number of remarks and statements associated with the results given in section 4.5.1.

- The difference between the estimated covariance sequence and the exact covariance sequence must, from a statistical point of view, decrease with an increasing number of samples. We have introduced two measures to check this statement, namely RECO and ERRO. If we compare the values of RCDO computed with respectively N=300,1000,10000 samples the statement given above is confirmed,except in the case of SYS5,but from a statistical point of view this deviating result is admitted. If we execute this simulation with other initial value of the random generators, we achieve the decreasing pattern. The same remarks can be given about ERRO.

- The ability of the algorithm to find a realization of the estimated covariance sequence is presented by the measure ERR1. The Page approach is preferable in case of N=300 and 1000 samples. In other words: in case the covariance sequence is heavily noise corrupted, the realization by means of the Page approach is more accurate.This statement coincides with previously found results by Van den Hof (1982). However there is one difference: the noise on the sequence of covariance matrices is not white in contrast with the white noise corrupted experiments executed by Van de Hof (1982).
  In the case where we consider a sequence of covariance matrices with relatively small disturbances, actually the estimated sequence of covariance matrices with N=10000, in a number of cases,the Hankel approach is preferable. The Hankel approach preferable in the following cases: SYS1-2 with N=10000,SYS2-1 with N=10000 and SYS3-2 with N=10000
  At this moment we can not distiguish any regularity within this aspect.
- The quality of the innovations representations depends on $\tilde{R}(0)$ and $\tilde{R}(k)$,k=1,2,.,NOM.The realization $\{\hat{A},\hat{L},\hat{C}\}$ obtained by Hankel or Page approach is based on $\tilde{R}(k)$,k=1,2,.NOM and is

independent of $\tilde{R}(0)$. Consequently a more accurate realization does not automatically imply a better innovations representation of the process: e.g. SYS1-2 N=10000 tabel 4.5f, the realization with the Hankel matrix is more accurate than with the Page approach, in spite of the fact that the ultimate WINV, associated with the Hankel approach, is larger than WINV associated with the Page approach. The above mentioned situation can of course occur within limited bound, because both $\tilde{R}(0)$ and $\tilde{R}(k)$,k=1,2,..NOM are computed with the same number of samples and the accuracy of realization is proportionally to the noise level on $\tilde{R}(k)$ k=1,2,...,NOM.

- If we compare the quality of the identified systems, measured with WINV,the differences between the systems with the channel independent noise and the channel interdependent noise are relatively small, with a small preference of the channel independent driven processes. The reason is the proper choice of $\{\Sigma_1, S_1, M_1\}$, i.e. the noise is chosen to be balanced over the channels.

- The identification algorithm generally recognizes the large eigenvalues of the process. This is caused by the fact that the contributions of the small eigenvalues are not sufficiently detectable in the disturbed covariance sequences. This holds even in the case of N=10000 samples. We are able to compute the approximate covariance of the estimated covariance matrix (see equation (2.16)). Then we obtain an indication how many covariance matrices can be taken meaningfully into account,i.e. covariance matrices which are not completely disturbed by noise, to construct an approximate realization of the estimated covariance. This procedure is not performed and in every test the number of used covariance matrices is 24(+1). In the case of small eigenvalues this number will be much smaller than with large eigenvalues and consequently the approximate realization less accurate.

- In the case of SYS1-2, N=300 with the Page approach and SYS6-1,N=300 and N=10000 with the Hankel approach, an error in the execution of the algorithm has occured, namely the stable subspace, associated with the D.A.R.E, is of smaller dimension than the unstable subspace. The dimension of the stable subspace is by SYS1-2 3 and by SYS6-1 0. In chapter 1 is proven that if $T_p>0$, Toeplitz matrix constructed with covariance matrices, the set of solutions $\{P\}$ is non void. It is possible that the realized sequence of "covariance" matrices $\tilde{R}(k)$,k=1,.,NOM,does not satisfy the properties of a sequence of a covariance matrices (that $T_p>0$), because in the realization algorithm no restriction is embedded for the realization. In the case mentioned above the obtained solution $P<0$ and consequently the realized sequence is not a covariance sequence (see chapter 1).
  The estimated covariance sequence is always a covariance sequence, because it is computed with with a fixed number of samples for each time shift, see van Zee (1981)
  There are two possibilities to recognize this situation:
  1 to detect if there is a solution of the D.A.R.E and ,if there is a solution P,to check whether P is positive definite. If there is no solution or $P \leq 0$ then this implies that $T_p<0$ and consequently the sequence is not sequence of covariance matrices.
  2 To check before computing the solution of the Riccati equation that $T_p>0$ (the finite T_p). This can easily be verified by means of of a small number of computations;see appendix 4.1

- For all experiments RARE $< 10^{-28}$, that means that the accuracy of the solution is in all cases sufficiently, if we compare this with the accuracy of the approximate realization (ERR1).
- From figure 4.6a,b,c,d, where respectively the covariance sequence of $R(k),\tilde{R}(k)$ and $\hat{R}(k)$ of SYS2-1 has been shown, we can see that the realized covariance $\hat{R}(k)$ fits the estimated

covariance sequence $\tilde{R}(k)$, as already was shown in tabel 4.5; in figure 4.7 e,f,g,h is shown the covariance sequence of the output vector process of the G.I.F. (see figure 3.4). From this figure we can notice that the innovations process is white. We have obtained a good innovations model of the time series $\{y(t)\}$

- In SYS6-1 is embedded a pure time delay (eigenvalue 0). By a correct execution of the algorithm this second order process, should be reduced to a first order process. As mentioned in section 4.3 the singular values can be used to determine the order of the process. The simulations are executed with N=300,1000 and 10000 samples and with the Hankel as well as with the Page approach. The singular values of both approaches are represented in tabel 4.5ℓ. We can notice that with an increasing number of samples the gap between δ_1 and δ_2 increases ,see also figure 4.8a,b. In the case of N=10000 it is obviously a first order process. In the Hankel approach an error occurs, namely the dimension of the stable subspace associated with the D.A.R.E. is 0 and so the set $\{P\}$ is void.

Generally speaking,if $N \leqslant 10000$, we may conclude for these simulations that the identification algorithm is reliable and that the Page approach is preferable, because it is more accurate and occupies less memory space than the Hankel approach with the same number of covariance matrices. However there two weak links in the algorithm:

1 to compute a reliable estimation of the covariance matrices we need a large number of samples (N=10000).

2 it is not sure that the approximate realization of the estimated covariance fulfills the properties of a sequence of covariance matrices, i.e. $Tp \succ 0$.

4.7 Dunes in running water

4.7.1 Problem formulation

When we consider a riverbed of granular matter,e.g. sand, the surface of this bed will be deformed. This deformation process is caused by a number of forces of the running water and turbulences in the water. In figure 4.9a has been depicted a dune . The turbulence in the water causes that the angle $\alpha <$ angle β, see figure 4.9a.
The purpose is to find a reliable model associated with these dunes,see figure 4.9b, and the corresponding combination of physical measures of the process, such as velocity of the water,water height,diameter of the granular matter. If these corresponding physical properties are found, we hope to be able to predict the dune pattern associated with these properties. In this report we limit ourselves to find a reliable model and it is not our intention to study the physical dune process.

Laboratory circumstances

The dune-shaping procces is a very complicated process. To reduce the complexity of the process and to obtain accurate measurements a flume is constructed in the laboratory, namely in " Het Waterloopkundig Laboratorium De Voorst, Emmeloord".The dimensions of the sand flume, depicted in figure 4.10,are:

length: 45 m
variable width: 0.5 - 1.5 m
heigth: 1.30 m

The width is fixed over the entire length. The three dune traces are measured lengthwise by means of 3 instruments. These intruments are attached on a vehicle, which is moving lengthwise with a constant velocity. The three intruments measure the hight of the dunes every centimeter (lengthwise), with an accuracy of $0.5\ 10^{-3}$ m. The measuring circumstances are depicted in tabel 4.6.

4.7.2 Identification of the dune profiles

In previous reports,see Meertens (1981), van Gent (1981), Pfennings (1982) and Vaessen (1982), methods are discribed to obtain ARMA (Auto Regressive Moving Average) models based on one dune trace. In this report we are able to consider three dune traces simultaneously . As mentioned before, the identification algorithm computes first the sequence of estimated covariance matrices $\tilde{R}(k)$ of the dune process. This sequence of covariance matrices is depicted in figure 4.11. The separate pictures correspond the elements of the covariance matrix and the time shift k corresponds with the horizontal axes. When we compare these separate figures, we see that these figures have a simular shape, which means that the channels are highly correlated.

Our main aim is to find a model representing the low frequently aspects of the dune process, that is we are not interested in the ribs on the dunes. Instead of using all available data (3000 samples, as the dune height is measured every 0.01 m), we use 1 of every 10 samples (total number of resulting samples is 300, i.e. the dune height measured every 0.10 m). This data-processing affects directly the sequence of covariance matrices and the corresponding realization, because we can offer just 24 covariance matrices to the realization algorithm.

The realization is performed with respectively the Hankel and Page approach. The intermediate results of the identification algorithm are given in tabel 4.6b.In figure 4.12a,b the singular values of respectively the Hankel and Page matrix are depicted. If we consider these singular values: $\delta_1, \delta_2, \ldots, \delta_\rho$, than we can notice that there is a gap between δ_2 and $\delta_3$in both approaches . Based on this gap and on theoretical backgrounds (see a.o. Van den Hof (1982)),we assume that the dune process is a second order process. Element (1,1),respectively computed by the Hankel and Page approach, of a second order approximate realization of the covariance matrix sequence is depicted respectively in figure 4.13a and 4.13b together with the estimated sequence. The other elements of the covariance sequence of the estimated realization are not depicted, because the have a simular shape.

The next step is to compute the innovations model and to check the quality of the model. There are two possibilities to check the quality, by means of:

-1 the G.I.F.and to check the whiteness of e(t) by means of WINV

-2 a simulation of the dune process and compare these results with the measured dunes.

Both possibilities are performed and compared.

4.7.3 Discussion of the identification results of the dunes

The Hankel approach (executed with 300 samples) is more accurate than Page approach for finding an approximate realization of the estimated covariance in the dune identification. If we compare ERR1 of tabel 4.5(simulations performed with 300 samples) and ERR1 of tabel 4.6b,we see the opposite, namely the Page approach is more accurate than the Hankel. This probably is caused by the fact that the dune process is a nonlinear process.

In contrast with the above mentioned result the innovations model obtained by the Page approach is better. This statement is based on the associated measure WINV.
As mentioned before, the innovations filter depends on the realization of the covariance $\{\hat{A},\hat{L},\hat{C}\}$and the estimated covariance $\tilde{R}(0)$.
Obviously the combination of $\tilde{R}(0)$ and the estimated realization by means of the Page approach leads to a better innovations filter. In both approaches the whiteness of the G.I.F. is not sufficient, based on the $3\sqrt{Var(\tilde{R}(k))}$ (99.8 ^{0}/o two sided confidence interval, see figure 4.14a,b)
Because the whiteness of the innovations approach is not sufficient we have tried to choose a higher order. These attempts did not succeed, because then the realized sequence has not the properties of a covariance sequence and so the set $\{P\}$ is void (P>0).

We simulate the dune process with the found innovations process $\{\hat{A},\hat{C},K,Q\}$ found by respectively the Hankel or Page approach and a random sequence , see figure 4.15a and b. Then we can notice that the difference between both simulations is negligible and that there is a slight difference in the minimal and maximal values of the dunes. The fundamental frequency of the simulated dunes agrees with the measured dunes in both approaches.
The difference between the left- and the righthandside of a dune (see measured dunes figure 4.9a,b) is not present in the simulated dunes, because this is a nonlinear aspect and cannot be modeled by the innovations representation.
At this moment we can not say how good the innovations representation fits certain statistical properties of the dune process, because additional tests are necessary and we have to consider more different dune measurements.

Chapter 5

CONCLUSIONS

In this report two recently developed techniques are combined in one stochastic realization or identification algorithm,namely the approximate realization by means of the Hankel or Page approach and solving the algebraic Riccati equation by means of the stable subspace of the associated pencil.

The simulation tests performed have proven that respectively the stochastic realization and,generally,the stochastic identification algorithm is a reliable, efficient and useful algorithm to obtain an innovations representations of the output process $\{y(t)\}$. Concerning the accuracy of the estimated innovations model the Page method is generally preferable in the executed simulations. The number of output samples is ranging from 300-10000 . However in the realization phase the Hankel approach is more accurate.
The stochastic identification algorithm is also used to identify dunes in a sand flume, but the chosen second order innovations model does not satisfy, because firstly the output of the causally invertible filter is not white and secondly the nonlinear aspects, such as the difference between the left- and righthandside of the dunes, are not represented by the innovations approach.
In some simulations performed (with 300 samples) and if higher order models,associated with the dune process, are chosen the sequence of the estimated covariance matrix, time shift 0 and the realized sequence of covariance matrices, time shift 1 until NOM $(\tilde{R}(0),\hat{R}(1),\hat{R}(2)....\hat{R}(NOM)$ (NOM:the number of considered covariance matrices), is not a covariance sequence, i.e. does not fulfil the properties of a covariance sequence.

Suggestions for further investigation are:

- To avoid these false sequences a restriction must be embedded in the approximate realization algorithm of the covariance. (However it is possible, especially when R(0) is heavily noise corrupted,that no such a sequence can be found).These aspects are not studied, but are recommended to be investigated.

  Remark: it may be possible to use the results of chapter 2, i.e. the knowledge that noise on the covariance sequence is coloured.
- to model other physical processes by means of the stochastic identification algorithm (e.g. brain waves).
- to complete the proof of the direct method for solving the Riccati equation.

Applications of the stochastic identification algorithm are:

- to use the algorithm as an on-line predictor. In that case the covariance has to be computed recursively, which can be achieved by a slight modification of the algorithm;
- to use the algorithm in recursive estimators, o.a. the maximum likelihood estimator, to obtain the coefficients of the filter to 'whiten' the residuals.

LIST OF SYMBOLS

A : system matrix (nxn)

a : eigenvalue of a first order single output process

B : distribution matrix (nxp)

C : - output matrix (nxq)
- power of the autocorrelation

C^- : area inside the unit circle

C^+ : area outside the unit circle

Γ : matrix introduced to construct the desired covariance matrix. $\Gamma\ \Gamma^T$ is a covariance matrix

D : input-output matrix (qxq)

δ_i : singular value

δ : relative minimum of the covariance (in single output processes), that must taken into account

E : expected value operator $E\{\ \}$

e(t): innovations process

F : transfer matrix (nxn)

G : distribution matrix in the spectral factorization problem (nxp)

H : output matrix in the spectral factorization problem (qxn)

H(ρ): Hankel matrix contructed with ρ covariance matrices

I : identity matrix

i : general index

j : general index

K : Kalman gain (nxq)-matrix

k : -sample number
-time shift

L : distribution matrix in the realization of the covariance $\{A,L,C\}$

ℓ : the number of covariance matrices that must be taken into account

Λ : matrix in Jordan canonical form

λ : generalized eigenvalue

M : covariance matrix of the output white noise vector v(t)

μ : block dimension of the Page matrix ($\eta x \mu$)

N : - number of samples
- matrix of dimension (qxq)

N2 : number of samples (fixed for all covariance matrices)

n : system dimension

0 : zero matrix

P(ηxμ) : Page matrix

P : solution of the Riccati equation

P^- : smallest positive definite solution of Riccati equation

p : number of inputs

Π : solution of the algebraic Riccati equation

Q : orthogonal matrices, used in the real Schur approach

q : number of outputs

R(k): covariance matrix corresponding with time shift k (qxq)

r : number of zero and infinite eigenvalues of the generalized eigen value problem

S : mutual correlation matrix between input and output noise

s : - degree of the minimal polynomial
- general index

Σ_{svd}: diagonal matrix of singular values

Σ : covariance matrix of the white input noise u(t) (pxp)

σ : variance

T : similarity transformation matrix in the state space

t : time

u : input vector

V : right singular vector matrix

Φ(z): power density matrix

W : left singular vector matrix

W(z): spectral factor

w : abitrary vector

X : matrix of the generalized eigenvalue problem $Xz=\lambda Yz$ (2nx2n)

Y : matrix of the generalized eigenvalue problem $Xz=\lambda Yz$ (2nx2n)

Z : orthogonal matrix, the columns of Z spans a subspace corresponding with the genaralized eigenvalue problem.

z : - generalized eigenvectors
- time delay

$\hat{}$: - realized sequence

-orthogonally transformed matrices, corresponding to the Real Schur approach

$\sim$: - estimated sequence or noise corrupted sequence

$\bar{}$: computed sequence with the innovations model $\{\hat{A},\hat{C},\hat{K},\hat{Q}\}$

$\|A\|_E$: the Euclidian norm, $\sum_{i=1}^{n}\sum_{j=1}^{n} A_{i,j}^2$

Appendix A2.1

The derivation of the second order crossmoment of the estimated covariance

In this appendix we will derive an expression for the second order crossmoment in channel number and in time shift of the estimated covariance of Gaussian distributed signals. At the end of this section expressions will be given for the scalar case.

The second order crossmoment of the estimated covariance of a q-vector signal is denoted by:

$$\mathrm{Cov}\{\tilde{R}_{i,j}(k)\cdot\tilde{R}_{l,m}(n+k)\}=E[\{\tilde{R}_{i,j}(k)-R_{i,j}(k)\}\{\tilde{R}_{l,m}(n+k)-R_{l,m}(n+k)\}]$$

$$=E\{\tilde{R}_{i,j}(k)\cdot\tilde{R}_{l,m}(n+k)\}-R_{i,j}(k)\cdot R_{l,m}(n+k)$$

$$\text{with } \tilde{R}_{i,j}(k)=\frac{1}{N-k}\cdot\sum_{p=1}^{N-k}y_i(p+k)\cdot y_j(p),$$

with $y_i(k)$ being the i^{th} component of the vector $y(k)$,

and $E\{\tilde{R}_{i,j}(k)\}=R_{i,j}(k)=E\{y_i(p+k)\,y(p)\}$

this can be written as:

$$\mathrm{Cov}\{\tilde{R}_{i,j}(k)\cdot\tilde{R}_{l,m}(n+k)\}=E\{\frac{1}{(N-k)^2}\sum_{p=1}^{N-k}\sum_{q=1}^{N-k}y_i(p+k)\cdot y_j(p)\cdot y_l(q+n+k)y_m(q)\}$$

$$-R_{i,j}(k)\cdot R_{l,m}(n+k)$$

We have assumed that y(t) is Gaussian distributed and we can use the following property:

$$E\{y_1y_2y_3y_4\}=E\{y_1y_2\}\cdot E\{y_3y_4\}+E\{y_1y_3\}\cdot E\{y_2y_4\}+E\{y_2y_3\}\cdot E\{y_1y_4\}$$

We can rewrite the expression for the covariance of the estimated covariance by interchanging the summation and expectation operators as:

$$\mathrm{Cov}\{\tilde{R}_{i,j}(k)\cdot\tilde{R}_{l,m}(n+k)\}=$$

$$\frac{1}{(N-k)^2}\cdot\sum_{p=1}^{N-k}\sum_{q=1}^{N-k}(R_{i,l}(p-q-n)R_{j,m}(p-q)+R_{i,m}(p+k-q)\cdot R_{j,l}(p-q-n-k)$$

If we introduce s=p-q then the covariance of the estimated covariance is given by:

$$\mathrm{Cov}\{\tilde{R}_{i,j}(k).\tilde{R}_{l,m}(n+k)\} = \frac{1}{(N-k)^2}\sum_{s=1}^{N-k-1}\Big[R_{i,l}(s-n).R_{j,m}(s) + R_{i,l}(-s-n).R_{j,m}(-s) + R_{i,m}(s+k).R_{j,l}(s-n-k)) + R_{i,m}(-s+k).R_{j,l}(-s-k-n)\Big] + \frac{1}{N-k}.\big[R_{i,l}(-n).R_{j,m}(0) + R_{i,m}(k).R_{j,l}(-n-k)\big]$$

In the case n≠0 and if we cannot dispose of the real covariance , an estimated covariance may not be available as s+n may be greater than k.

Special cases:

- first case : n=0 :no fundamental changes
- second case: one channel :all i,l,j,m disappear
- third case : n=0 and one channel :

$$\mathrm{Cov}\{\tilde{R}(0).\tilde{R}(0)\} = \frac{2}{(N-k)^2}\sum_{s=1}^{N-k-1}(N-k-s).\big[R^2(s)+R^2(-s)+R(s+k).R(s-k)+R(-s+k).R(-s-k)\big] + \frac{1}{N-k}.(R^2(0)+R(k).R(-k))$$

R is an even function, consequently:

$$R(s) = R(-s)$$
$$R(-s-k) = R(s+k)$$
$$R(-s+k) = R(s-k)$$
$$R(k) = R(-k)$$

so we can write:

$$\mathrm{Cov}\{\tilde{R}(0).\tilde{R}(0)\} = \frac{2}{(N-k)^2}\sum_{s=1}^{N-k-1}(N-k-s).\big[R^2(s)+R(s+k).R(s-k)\big] + \frac{1}{N-k}.\big[R^2(0)+R^2(k)\big]$$

Appendix A2.2

Approximations of the variance of the estimated covariance

If one want to approximate the variance of the estimated covariance,one has to distinct two areas. The two areas are generally characterised by a roughly decreasing part (Area I) and an increasing part(Area I), as one can see in figure A2.2.1

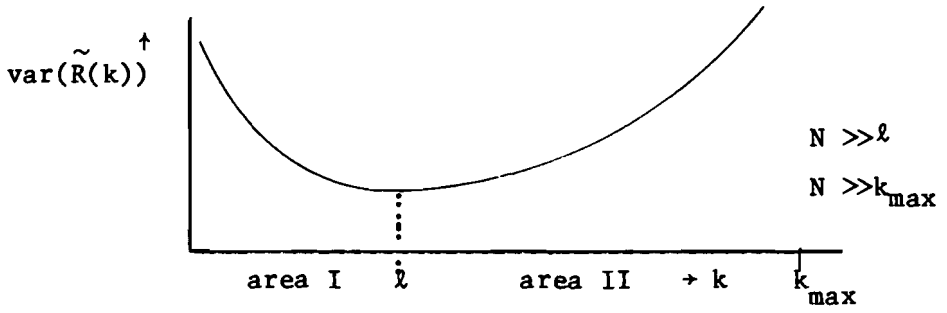


Figure A2.2.1

The variance of the estimated covariance is:

$$Var(\tilde{R}(k))= \frac{R^2(0)}{N-k} + \frac{R^2(k)}{N-k} + \frac{2}{N-k}\sum_{i=1}^{N-k}\left(1-\frac{i}{N-k}\right).(R^2(i)+R(i+k).R(i-k)) \qquad (A2.2.1)$$

I Approximations for the decreasing part

Under the assumption :i>ℓ,R(i) << R(0) we get:

$$Var(\tilde{R}(k)) = \frac{R^2(0)}{N-k} + \frac{R^2(k)}{N-k} + \frac{2}{N-k}\sum_{i=1}^{2\ell}\left(1-\frac{i}{N-k}\right).(R^2(i)+R(i+k).R(i-k)) \qquad (A2.2.2)$$

summation until 2.ℓ as we have indexed R(i-k)

If ℓ<<N then (N-k ~ N and ℓ/N <<1):

$$Var(\tilde{R}(k)) = \frac{R^2(0)}{N} + \frac{R^2(k)}{N} + \frac{2}{N}\sum_{i=1}^{2\ell}\left(1-\frac{i}{N}\right).(R^2(i)+R(i+k).R(i-k)) \qquad (A2.2.3)$$

$$Var(\tilde{R}(k)) = \frac{1}{N}\left(R^2(0)+R^2(k)+2.\sum_{i=1}^{\ell} R^2(i) +2.\sum_{i=1}^{2.\ell} R(i+k).R(i-k)\right) \qquad (A2.2.4)$$

$$\mathrm{Var}(\tilde{R}(k)) = \frac{1}{N}\,(C + R^2(k) + 2.\sum_{i=1}^{2\ell} R(i+k).R(i-k)) \qquad (A2.2.5)$$

with C : the 'power in the autocorrelation

$$C = R^2(0) + 2.\sum_{i=1}^{\ell} R^2(i) \qquad (A2.2.6)$$

C:the power in the autocorrelation

II Approximation for the increasing part of the estimated covariance

$k \gg \ell, k \ll N$

The same approximations can be made as in area I and on assumptions: $|R(k)| \ll R(0)$ for $k>1$ we may write:

$$\mathrm{Var}(\tilde{R}(k)) = \frac{C}{N-k} \qquad (A2.2.7)$$

Notice that again N-k has been used instead of N as k is approaching N now.

Appendix 2.3

Gaussian white noise generators

In this appendix we will investigate three Gaussian white noise generators on the following properties:

a) the probability distribution function
b) the whiteness by means of the autocorrelation function.

The three Gaussian white noise generators are :

a) GAUSS1.

   Available in the subroutine GAUSS1(SD). This subroutine takes 12 samples of an independent rectangularly distributed random generator RAN(I,J). RAN(I,J) is a FORTRAN IV library function to produce rectangularly distributed random numbers (see FORTRAN IV users guide)

b) GAUSS2.

   Available in the subroutine GAUSS2(X,R2). This subroutine produces white Gaussian distributed numbers by means of two independent rectangularly distributed random generators R1 and R2 (mean=0). Then one can obtain normally distributed random numbers N1 and N2 (mean=0,standard deviation=1) using the Box-Muller transformation,see Dahlquist (1974). This subroutine needs the subroutine RAND(IX). RAND(IX) produces rectangularly distributed white numbers.

c) GAUSS3

   Available in the subroutine GAUSS3(RAN1). The same as GAUSS2, but it needs the subroutine SRAN(IX) instead of RAND(IX) to produce independent rectangularly distributed numbers.

The subroutines are written in FORTRAN IV-plus. The subroutines are included in the User's manual for stochastic identification programs (see Van Gent (1983)). The tests are executed on a PDP 11/60 computer with a machine accuracy of 2.10^{-8}.

The autocorrelation of respectively white noise generators GAUSS1,GAUSS2 and GAUSS3 are depicted in figure A2.3.1,A2.3.2 and A2.3.3 There are no significant differences between the three autocorrelation functions. The correlation between two samples y(i) and y(k) (i≠k) is less than abs(0.1) (the two sided 99.8^0/o confidence interval).

In figure A2.3.4,A2.3.5 and A2.3.6 the distribution function of respectively random generator GAUSS1,GAUSS2 and GAUSS3 is depicted. These distribution functions are calculated with 1000 random numbers. Also in these figures the normal distribution function:

$$F(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-u^2/2}.du$$

is depicted as a reference distribution. F(x) is calculated with a Runga-Kutta method with an absolute accuracy of 10^{-5} . The difference between the desired normal and the generated distribution function by respectively random generator GAUSS1, GAUSS2 and GAUSS3 is negligible. The differences between figures A2.3.4,A2.3.5 A2.3.6 are incidental, because they are caused by the choosen initial value. There are presumable "bad" and "good" initial values of the random generators. This has not been investigated, but from experience we know that 0 is a bad value, because the first generated random number is always greater than three times the standard deviation, thus quite exceptional. An initial value of five digits is a "good"one.

Appendix 3.1

Proof belonging to chapter 3

In this appendix will be proven that:

if:

$\{F,G,H,R(0)\}$ satisfy equation (3.24) and

$\{F,G\}$ is controllable , $\{F,H\}$ is observable, $R(0)=J+J^T>0, P\epsilon R^{nxn}$

$P=P^T, R(0)-G^T P G$ and $F_1=F - G R(0)^{-1} H \epsilon R^{nxn}$

then: $P-F^T P F-(H -F P G)(R(0)-G P G)^{-1}(H^T-F^T P G)^T=0$

iff $P-F_1^T P F_1-H^T R(0)^{-1}-F_1^T P G (R(0)-G^T P G)^{-1}G^T P F_1=0$

Proof:

$$
\begin{aligned}
&F^T P F+(H^T-F^T P G) (R(0)-G^T P G)^{-1}(H^T-F^T P G)\\
&\quad -F_1^T P F_1-H^T R(0)^{-1} H-F_1^T P G (R(0)-G^T P G)^{-1}G^T P F_1\\
&=F^T P F + (H^T-F^T P G) (R(0)-G^T P G)^{-1} (H^T-F^T P G)^T\\
&\quad -(F-G R(0)^{-1} H)^T P (F-G R(0)^{-1}H) - H^T R(0)^{-1} H\\
&\quad -(F-G R(0)^{-1}H)^T P G (R(0)-G^T P G)^{-1}G^T P (F-G R(0)^{-1}H)\\
&=F^T P G\Big[R(0)^{-1}+\underbrace{(R(0)^{-1}-G P G)^{-1}G^T P G R(0)^{-1}-(R(0)^{-1}-G^T PG)^{-1}}_{=0,\ \text{see 1}}\Big] H\\
&\quad +H^T \Big[\underbrace{R(0)^{-1}+R(0)^{-1}G^T P G(R(0)-G^T P G R(0)^{-1}-(R(0)-G^T P G)^{-1}}_{=0,\ \text{see 1}}\Big] G^T PF\\
&\quad +H^T R(0)^{-1}\Big[\underbrace{R(0)+R(0) (R(0)-G^T P G)^{-1}R(0)-G^T P G-}\\
&\qquad \underbrace{-G^T P G(R(0)-G^T P G)^{-1}G^T P G}_{=0,\ \text{see 2}}\Big] R(0)^{-1}H=0
\end{aligned}
$$

because:

1- $R(0)^{-1}+(R(0)-G^T P G)^{-1}G^T P G R(0)^{-1}-(R(0)-G^T P G)^{-1}$

$=(R(0)-G^T P G)^{-1} ((R(0)-G^T P G) R(0)^{-1}+G^T P G R(0)^{-1}-I)=0,$

2- $-R(0)+R(0)(R(0)-G^T P G)^{-1}R(0)-G^T P G-G^T P G (R(0)-G^T P G)G^T PG$

$=-R(0)-G^T P G+R(0) (R(0)-G^T P G)^{-1}(R(0)-G^T P G)$

$+(R(0)-G^T P G) (R(0)-G^T P G)^{-1} G^T P G=0$ q.e.d.

Appendix 4.1

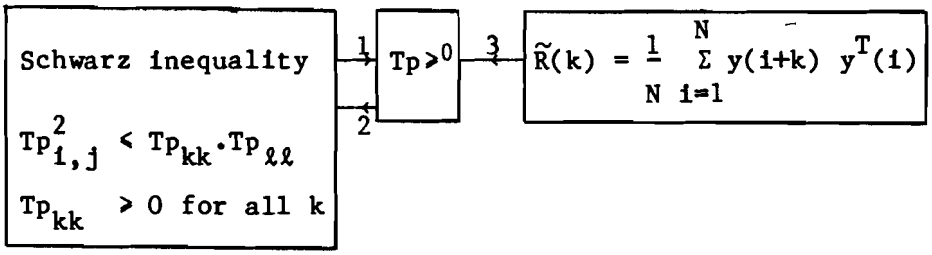
Algorithm to detect the positive real property of a sequence of matrices

In this appendix it is proven that the nonnegativity of the finite Toeplitz matrix is equivalent with a Schwartz inequality on the first q-rows or q-columns of this matrix ,which is important for the determination of the positive realness of the finite covariance sequence.

The symmetrical, finite Toeplitz matrix is defined as:

$$Tp = \begin{bmatrix} \tilde{R}(0) & \tilde{R}(1) & \tilde{R}(2) & \cdot \\ \tilde{R}^T(1) & \tilde{R}(0) & \tilde{R}(1) & \cdot \\ \cdot & & & \cdot \\ \cdot & & & \cdot \\ \cdot & \cdot & \cdot & \cdot \end{bmatrix} \begin{matrix} \uparrow \\ \\ qN \\ \\ \downarrow \end{matrix}$$

We need the following statements in the algorithm:



Implication 1

<u>If</u> $Tp_{ii} \geq 0$ and $|Tp_{ij}| \leq \sqrt{Tp_{ii} Tp_{jj}}$ for all $i,j \in (1,..Nq)$

<u>then</u> we have to prove that: $\sum_{i,j} Tp_{i,j}\, x_j\, x_i \geq 0$ for all $x_i \neq 0$

<u>Proof</u>

$$\left[\sum_{i=1}^{Nq} s_i \sqrt{Tp_{ii}}\, x_i \right]^2 \geq 0 \text{ for all combinations of } s_i = \pm 1$$

$$\rightarrow \sum_{i=1}^{Nq} Tp_{ii}x_i^2 + \sum_{\substack{i,j \\ i\neq j}}^{Nq} s_i s_j \sqrt{Tp_{ii}}\sqrt{Tp_{jj}}\, x_i x_j \geq 0$$

$$\rightarrow \sum_{i=1}^{Nq} Tp_{ii}x_i^2 \geq \left| \sum_{i\neq j}^{Nq} s_i s_j \sqrt{Tp_{ii}Tp_{jj}}\cdot x_i x_j \right| \quad \text{for all } s_i$$

The maximal value of the rightside of the equation given above is: $\sum_{i\neq j}^{Nq} \sqrt{Tp_{ii}Tp_{jj}}\,|x_i x_j|$ by putting $s_i = \text{sign}(x_i)$

$$\rightarrow \sum_{i=1}^{Nq} Tp_{ii}x_i^2 \geq \sum_{i\neq j}^{Nq} \sqrt{Tp_{ii}Tp_{jj}}\,|x_i x_j| \geq \sum_{i\neq j}^{Nq} |Tp_{ij}||x_i x_j|$$

$$\sum_{i=1}^{Nq} Tp_{ii}\, x_i^2 \geq \sum_{i\neq j}^{Nq} |Tp_{ij}|\,|x_i x_j|$$

Based on the inequality mentioned above, we can say:

$$\sum_{i=1}^{Nq} Tp_{ii}\, x_i^2 + \sum_{i\neq j}^{Nq} Tp_{ij}\, x_i x_j \geq 0 \qquad \text{q.e.d.}$$

Thus Tp is nonnegative.

Implication 2

See Veltkamp (1982,p 135)

Implication 3

See Van Zee (1982)

Remark: this implication is inserted , because this shows that the estimated covariance computed with a fixed number samples always implies that $Tp \geq 0$ or that the estimated sequence is a covariance sequence.

Algorithm

The diagonal elements of a Toeplitz matrix are the diagonal elements of the covariance matrix time shift 0 and it is sufficient if we compare these q-elements with other nondiagonal elements.
The first block row or column contains all elements, which appears in the Toeplitz matrix. If we compare the q-diagonal elements of R(0) with the first block column' (Nq^2-q elements) and if :

$$\left.\begin{array}{l} Tp_{i,i} \geq Tp_{i,j} \text{ for } i=1,..q,\ j=2,....Nq \\ \qquad\qquad\qquad\qquad\qquad\qquad i<j \end{array}\right] \rightarrow Tp \geq 0$$

($Tp_{i,i}$ is always ≥ 0, because it is a quadratic form)

' A column is chosen, because in most computers an matrix is stored column by column.

Appendix 4.2

The solution of the Lyapunov equation by means of a direct method

Given the discrete algebraic Riccati equation associated with the optimal control problem (cf. Van Dooren (1981a)):

$$P = F^T P F - F^T P G (R + G^T P G)^{-1} G^T P F + Q$$

with : $x(i+1) = F x(i) + G u(i)$ a controllable system
and find the control $u(i) = -K x(i)$ minimizing:

$$J = \sum_{i=0}^{\infty} [x^T(i) Q x(i) + u^T(i) R u(i)]$$

with $K = (R + G P G)^{-1} G^T P$

Q nxn-matrix ≥ 0

R mxm-matrix > 0

F nxn-matrix

G nxm-matrix

K mxn matrix, the Kalman gain

u m-vector (input vector)

x n-vector (state vector)

P nxn-matrix (positive definite solution of the algebraic Riccati equation)

(F,Q) observable

The Riccati equation can be solved by means of computing the stable deflating subspace of the pencil:

$$\lambda \begin{bmatrix} I & G R^{-1} G^T \\ 0 & F^T \end{bmatrix} - \begin{bmatrix} F & 0 \\ -Q & I \end{bmatrix}$$

with $\begin{bmatrix} Z_{11} \\ Z_{21} \end{bmatrix}$ a basis for the stable subspace ,i.e. the subspace associated with the eigenvalues inside the unit circle; see Van Dooren (1981a).

If we choose : $F^T=A$, $Q=B\ \Sigma\ B^T$, $R=I$ and $G=0$ then we fulfil the properties: $Q \geq 0$

$R > 0$

and we have to require that $(A^T, B\ \Sigma\ B^T)$ is observable.

The discrete algebraic Riccati equation becomes:

$$P = A\ P\ A^T + B\ \Sigma\ B^T$$

This equation is called Lyapunov equation and the associated pencil is:

$$\lambda \begin{bmatrix} I & 0 \\ 0 & A \end{bmatrix} - \begin{bmatrix} A^T & 0 \\ -B\ \Sigma\ B^T & I \end{bmatrix}$$

The solution is $P=Z_{21}Z_{11}^{-1}$ with $\begin{bmatrix} Z_{11} \\ Z_{21} \end{bmatrix}$ a basis for the stable subspace.

Tabel 4.1 Accuracy of the realization algorithm

System	Hankel/Page approach	Non channel interdependent noise driven ERR2	Channel interdependent noise driven ERR2
0	Page	$0.91\ 10^{-13}$	$0.19\ 10^{-13}$
0	Hankel	$0.24\ 10^{-13}$	$0.96\ 10^{-13}$
1	Page	$0.22\ 10^{-11}$	$0.53\ 10^{-11}$
1	Hankel	$0.15\ 10^{-11}$	$0.20\ 10^{-11}$
2	Page	$0.21\ 10^{-11}$	$0.55\ 10^{-11}$
2	Hankel	$0.19\ 10^{-11}$	$0.17\ 10^{-11}$
3	Page	$0.89\ 10^{-10}$	$0.22\ 10^{-10}$
3	Hankel	$0.97\ 10^{-11}$	$0.75\ 10^{-12}$
4	Page	$0.13\ 10^{-11}$	$0.50\ 10^{-12}$
4	Hankel	$0.24\ 10^{-15}$	$0.11\ 10^{-12}$
5	Page	$0.27\ 10^{-11}$	$0.79\ 10^{-11}$
5	Hankel	$0.75\ 10^{-12}$	$0.87\ 10^{-11}$
6	Page	-	-
6	Hankel	-	-

Tabel 4.2 Stability test of the realization algorithm

Iterations	Hankel approach ERR2	Page approach ERR2
1	$0.19\ 10^{-11}$	$0.21\ 10^{-11}$
2	$0.17\ 10^{-11}$	$0.17\ 10^{-11}$
3	$0.25\ 10^{-10}$	$0.45\ 10^{-11}$
4	$0.30\ 10^{-10}$	$0.66\ 10^{-11}$
5	$0.35\ 10^{-10}$	$0.11\ 10^{-10}$
6	$0.58\ 10^{-10}$	$0.10\ 10^{-10}$
7	$0.92\ 10^{-10}$	$0.76\ 10^{-11}$
8	$0.14\ 10^{-9}$	$0.45\ 10^{-11}$
9	$0.13\ 10^{-9}$	$0.49\ 10^{-11}$
10	$0.22\ 10^{-9}$	$0.38\ 10^{-10}$

Channel number	KX	KY
1	9584	5668
2	12589	4687
3	5685	2790
4	25797	12468
5	467	6256
6	3680	79

Tabel 4.3

Initial noise generator values.

Tabel 4.4

N	Non channel interdependent noise	Channel interdependent noise
300	$0.74\ 10^{-2}$	$0.32\ 10^{-2}$
1000	$0.63\ 10^{-3}$	$0.31\ 10^{-3}$
10000	$0.57\ 10^{-5}$	$0.29\ 10^{-5}$

Results stochastic identification A=diag(0.4,0.3,0.2,0.1)

System: SYS0-1

Hankel/Page approach	Number of samples	Eigenvalues realization	RECO	ERR0	ERR1	WINV
Page	300	0.83 0.40 ±j0.33 -0.29	$0.92\ 10^{-2}$	$0.44\ 10^{+1}$	$0.30\ 10^{-0}$	$0.18\ 10^{-2}$
Hankel	300	0.92 ±j0.36 0.95 0.259	$0.92\ 10^{-2}$	$0.44\ 10^{+1}$	$0.14\ 10^{+1}$	$0.18\ 10^{-2}$
Page	1000	0.75 ±j0.21 0.234 -.526	$0.31\ 10^{-2}$	$0.12\ 10^{+1}$	$0.98\ 10^{-1}$	$0.11\ 10^{-3}$
Hankel	1000	0.97 ±j0.34 0.50 -0.06	$0.31\ 10^{-2}$	$0.12\ 10^{+1}$	$0.20\ 10^{-0}$	$0.18\ 10^{-3}$
Page	10000	0.13 ±j0.10 0.51 0.29	$0.47\ 10^{-3}$	$0.74\ 10^{-1}$	$0.40\ 10^{-2}$	$0.71\ 10^{-6}$
Hankel	10000	0.27 ±j0.44 0.37 0.78	$0.47\ 10^{-3}$	$0.74\ 10^{-1}$	$0.10\ 10^{-1}$	$0.86\ 10^{-6}$

Tabel 4.5a

Results stochastic identification A=diag(0.4,0.3,0.2,0.1)

System:SYS0-2

Hankel/Page approach	Number of samples	Eigenvalues realization	RECO	ERR0	ERR1	WINV
Page	300	0.69 ±j0.36 0.45 -0.13	$0.33\ 10^{-2}$	$0.15\ 10^{+1}$	$0.87\ 10^{-1}$	$0.22\ 10^{-2}$
Hankel	300	0.85 ±j0.31 0.63 ±j0.65	$0.33\ 10^{-2}$	$0.15\ 10^{+1}$	$0.59\ 10^{-0}$	$0.23\ 10^{-1}$
Page	1000	0.52 ±j0.12 -0.46 0.32	$0.38\ 10^{-2}$	$0.25\ 10^{-0}$	$0.18\ 10^{-1}$	$0.68\ 10^{-4}$
Hankel	1000	0.76 ±j0.46 0.29 0.81	$0.38\ 10^{-2}$	$0.25\ 10^{-0}$	$0.53\ 10^{-1}$	$0.73\ 10^{-4}$
Page	10000	0.38 ±j0.24 0.0646 0.51	$0.72\ 10^{-4}$	$0.60\ 10^{-1}$	$0.12\ 10^{-1}$	$0.16\ 10^{-5}$
Hankel	10000	0.35 ±j0.21 0.85 ±j0.48	$0.72\ 10^{-4}$	$0.60\ 10^{-1}$	$0.26\ 10^{-1}$	$0.18\ 10^{-5}$

Tabel 4.5b

Results stochastic identification A=diag(0.6,0.5,0.9,0.7)

System:SYS1-1

Hankel/Page approach	Number of samples	Eigenvalues realization	RECO	ERR0	ERR1	WINV
Page	300	0.83 ±j0.24 0.36 0.74	$0.65\ 10^{-1}$	$0.55\ 10^{-0}$	$0.40\ 10^{-0}$	$0.16\ 10^{-2}$
Hankel	300	0.93 ±j0.33 0.85 0.86	$0.65\ 10^{-1}$	$0.55\ 10^{-0}$	$0.47\ 10^{-0}$	$0.33\ 10^{-2}$
Page	1000	0.60 0.85 1.1 1.02	$0.22\ 10^{-1}$	$0.96\ 10^{-1}$	$0.91\ 10^{-1}$	$0.13\ 10^{-3}$
Hankel	1000	0.94 ±j0.20 0.86 0.77	$0.22\ 10^{-1}$	$0.96\ 10^{-1}$	$0.86\ 10^{-1}$	$0.10\ 10^{-2}$
Page	10000	0.74 ±j0.12 0.90 0.49	$0.14\ 10^{-2}$	$0.56\ 10^{-2}$	$0.31\ 10^{-2}$	$0.86\ 10^{-6}$
Hankel	10000	0.77 ±j0.12 0.90 0.49	$0.14\ 10^{-2}$	$0.56\ 10^{-2}$	$0.51\ 10^{-2}$	$0.85\ 10^{-6}$

Tabel 4.5c

Results stochastic identification A=diag(0.6,0.5,0.9,0.7)

System:SYS1-2

Hankel/Page approach	Number of samples	Eigenvalues realization	RECO	ERR0	ERR1	WINV
Page	300	0.79 ±j0.18 0.81 0.98	$0.26\ 10^{-1}$	$0.40\ 10^{-0}$	$0.25\ 10^{-0}$	$0.20\ 10^{+2}$
Hankel	300	0.83 ±j0.29 0.86 ±j0.14	$0.26\ 10^{-1}$	$0.40\ 10^{-0}$	$0.37\ 10^{-0}$	$0.34\ 10^{-2}$
Page	1000	0.80 ±j0.13 0.45 0.86	$0.22\ 10^{-1}$	$0.68\ 10^{-1}$	$0.58\ 10^{-1}$	$0.81\ 10^{-4}$
Hankel	1000	0.84 ±j0.02 0.90 ±j0.25	$0.22\ 10^{-1}$	$0.70\ 10^{-1}$	$0.63\ 10^{-1}$	$0.18\ 10^{-3}$
Page	10000	0.53 ±j0.14 0.77 0.90	$0.13\ 10^{-2}$	$0.88\ 10^{-2}$	$0.81\ 10^{-2}$	$0.14\ 10^{-5}$
Hankel	10000	0.95 0.75 0.86 0.60	$0.13\ 10^{-2}$	$0.88\ 10^{-2}$	$0.79\ 10^{-2}$	$0.34\ 10^{-5}$

Tabel 4.5d

Error occurred in the execution of the stochastic identification algorithm. There is no solution P>0, i.e. the realized sequence is not a covariance sequence.

Results stochastic identification A=diag(0.9,0.8,0.85,0.3)

System:SYS2-1

Hankel/Page approach	Number of samples	Eigenvalues realization	RECO	ERR0	ERR1	WINV
Page	300	0.71 ±j0.40 0.94 0.57	$0.19\ 10^{-1}$	$0.38\ 10^{-0}$	$0.32\ 10^{-0}$	$0.16\ 10^{-2}$
Hankel	300	0.91 ±j0.33 0.94 0.99	$0.19\ 10^{-1}$	$0.38\ 10^{-0}$	$0.37\ 10^{-0}$	$0.32\ 10^{-2}$
Page	1000	0.91 ±j0.07 0.64 0.76	$0.44\ 10^{-2}$	$0.46\ 10^{-1}$	$0.38\ 10^{-1}$	$0.13\ 10^{-3}$
Hankel	1000	0.98 ±j0.39 0.91 0.82	$0.44\ 10^{-2}$	$0.46\ 10^{-1}$	$0.37\ 10^{-1}$	$0.20\ 10^{-2}$
Page	10000	0.51 ±j0.20 0.89 0.83	$0.59\ 10^{-3}$	$0.25\ 10^{-1}$	$0.20\ 10^{-2}$	$0.97\ 10^{-6}$
Hankel	10000	0.86 ±j0.18 0.89 0.83	$0.59\ 10^{-3}$	$0.25\ 10^{-1}$	$0.24\ 10^{-2}$	$0.14\ 10^{-5}$

Tabel 4.5e

Results stochastic identification A=diag(0.9,0.8,0.85,0.3)

System:SYS2-2

Hankel/Page approach	Number of samples	Eigenvalues realization	RECO	ERRO	ERR1	WINV
Page	300	0.90 ±j0.27 0.91 0.64	$0.70\ 10^{-2}$	$0.14\ 10^{-0}$	$0.11\ 10^{-0}$	$0.16\ 10^{-2}$
Hankel	300	0.92 ±j0.29 0.91 0.68	$0.70\ 10^{-2}$	$0.14\ 10^{-0}$	$0.13\ 10^{-0}$	$0.16\ 10^{-2}$
Page	1000	0.86 ±j0.06 0.96 0.36	$0.18\ 10^{-1}$	$0.47\ 10^{-2}$	$0.33\ 10^{-1}$	$0.72\ 10^{-4}$
Hankel	1000	0.99 ±j0.17 0.81 0.89	$0.18\ 10^{-1}$	$0.47\ 10^{-2}$	$0.40\ 10^{-1}$	$0.95\ 10^{-4}$
Page	10000	0.30 0.94 0.82 0.89	$0.39\ 10^{-2}$	$0.98\ 10^{-2}$	$0.10\ 10^{-1}$	$0.11\ 10^{-5}$
Hankel	10000	0.39 0.99 0.85 0.89	$0.39\ 10^{-2}$	$0.98\ 10^{-2}$	$0.94\ 10^{-2}$	$0.15\ 10^{-5}$

Tabel 4.5f

Results stochastic identification A=diag(0.2,0.1,0.9,0.15)

System:SYS3-1

Hankel/Page approach	Number of samples	Eigenvalues realization	RECO	ERRO	ERR1	WINV
Page	300	0.78 ±j0.17 -0.33 0.66	$0.73\ 10^{-1}$	$0.57\ 10^{-0}$	$0.34\ 10^{-0}$	$0.15\ 10^{-2}$
Hankel	300	0.98 ±j0.20 0.73 -0.03	$0.73\ 10^{-1}$	$0.57\ 10^{-0}$	$0.51\ 10^{-0}$	$0.16\ 10^{-2}$
Page	1000	0.91 ±j0.05 -0.05 0.63	$0.27\ 10^{-1}$	$0.74\ 10^{-1}$	$0.83\ 10^{-1}$	$0.13\ 10^{-3}$
Hankel	1000	0.81 ±j0.37 0.95 0.84	$0.27\ 10^{-1}$	$0.74\ 10^{-1}$	$0.63\ 10^{-1}$	$0.17\ 10^{-3}$
Page	10000	0.90 -0.03 0.40 0.51	$0.12\ 10^{-2}$	$0.44\ 10^{-2}$	$0.24\ 10^{-2}$	$0.87\ 10^{-6}$
Hankel	10000	0.25 ±j0.45 0.91 0.78	$0.12\ 10^{-2}$	$0.44\ 10^{-2}$	$0.39\ 10^{-2}$	$0.96\ 10^{-6}$

Tabel 4.5g

Results stochastic identification A=diag(0.2,0.1,0.9,0.15)

System:SYS3-2

Hankel/Page approach	Number of samples	Eigenvalues realization	RECO	ERRO	ERR1	WINV
Page	300	0.78 ±j0.32 0.06 0.77	$0.35\ 10^{-1}$	$0.41\ 10^{-0}$	$0.30\ 10^{-0}$	$0.21\ 10^{-2}$
Hankel	300	0.87 ±j0.12 0.76 ±j0.51	$0.35\ 10^{-1}$	$0.41\ 10^{-0}$	$0.39\ 10^{-0}$	$0.23\ 10^{-2}$
Page	1000	0.70 ±j0.16 0.87 0.15	$0.17\ 10^{-1}$	$0.95\ 10^{-1}$	$0.58\ 10^{-1}$	$0.73\ 10^{-4}$
Hankel	1000	0.79 ±j0.36 0.97 0.89	$0.17\ 10^{-1}$	$0.95\ 10^{-1}$	$0.54\ 10^{-1}$	$0.18\ 10^{-3}$
Page	10000	0.13 ±j0.07 0.89 0.83	$0.15\ 10^{-2}$	$0.86\ 10^{-2}$	$0.92\ 10^{-2}$	$0.17\ 10^{-5}$
Hankel	10000	0.58 ±j0.26 0.93 0.29	$0.15\ 10^{-2}$	$0.86\ 10^{-2}$	$0.83\ 10^{-2}$	$0.29\ 10^{-5}$

Tabel 4.5h

Results stochastic identification A=0.9

System:SYS4-1

Hankel/Page approach	Number of samples	Eigenvalues realization	RECO	ERRO	ERR1	WINV
Page	300	0.90	$0.94\ 10^{-1}$	$0.15\ 10^{-0}$	$0.13\ 10^{-0}$	$0.11\ 10^{-3}$
Hankel	300	0.91	$0.94\ 10^{-1}$	$0.15\ 10^{-0}$	$0.14\ 10^{-0}$	$0.15\ 10^{-3}$
Page	1000	0.89	$0.37\ 10^{-2}$	$0.12\ 10^{-1}$	$0.57\ 10^{-2}$	$0.44\ 10^{-4}$
Hankel	1000	0.89	$0.37\ 10^{-2}$	$0.12\ 10^{-1}$	$0.99\ 10^{-2}$	$0.95\ 10^{-4}$
Page	10000	0.90	$0.20\ 10^{-3}$	$0.31\ 10^{-2}$	$0.23\ 10^{-2}$	$0.24\ 10^{-6}$
Hankel	10000	0.89	$0.20\ 10^{-3}$	$0.31\ 10^{-2}$	$0.30\ 10^{-2}$	$0.33\ 10^{-6}$

Tabel 4.5i

Results stochastic identification Eigenvalues system matrix 0.8 ± j0.5

System:SYS5-1

Hankel/Page approach	Number of samples	Eigenvalues realization	RECO	ERRO	ERR1	WINV
Page	300	0.819±j0.498	$0.15\ 10^{-2}$	$0.25\ 10^{-2}$	$0.25\ 10^{-1}$	$0.56\ 10^{-3}$
Hankel	300	0.813±j0.497	$0.15\ 10^{-2}$	$0.25\ 10^{-2}$	$0.22\ 10^{-1}$	$0.18\ 10^{-3}$
Page	1000	0.800±j0.505	$0.21\ 10^{-1}$	$0.55\ 10^{-1}$	$0.48\ 10^{-1}$	$0.20\ 10^{-4}$
Hankel	1000	0.801±j0.505	$0.21\ 10^{-1}$	$0.55\ 10^{-1}$	$0.55\ 10^{-1}$	$0.21\ 10^{-4}$
Page	10000	0.805±j0.499	$0.32\ 10^{-3}$	$0.38\ 10^{-2}$	$0.35\ 10^{-2}$	$0.22\ 10^{-6}$
Hankel	10000	0.805±j0.502	$0.32\ 10^{-3}$	$0.38\ 10^{-2}$	$0.33\ 10^{-2}$	$0.97\ 10^{-6}$

Tabel 4.5j

SYS6-1 Results stochastic identification Eigenvalues (0.9,0.0)

Hankel/Page approach	Number of samples	eigenvalues realization	RECO	ERR0	ERR1	WINV
Page	300	0.908	$0.37\ 10^{-1}$	0.54	$0.59\ 10^{-1}$	$0.56\ 10^{-3}$
Hankel*	300	0.864	$0.37\ 10^{-1}$	0.54	$0.46\ 10^{-1}$	-
Page	1000	0.899	$0.34\ 10^{-1}$	$0.57\ 10^{-1}$	$0.45\ 10^{-1}$	$0.59\ 10^{-4}$
Hankel*	1000	0.893	$0.34\ 10^{-1}$	$0.57\ 10^{-1}$	$0.48\ 10^{-1}$	-
Page	10000	0.912	$0.75\ 10^{-2}$	$0.28\ 10^{-1}$	$0.26\ 10^{-1}$	$0.52\ 10^{-6}$
Hankel	10000	0.913	$0.75\ 10^{-2}$	$0.28\ 10^{-1}$	$0.26\ 10^{-1}$	$0.29\ 10^{-5}$

* Error occurred in the execution of the identification algorithm. There is no solution P>0, i.e. the realized sequence is not a covariance sequence.

Tabel 4.5k

Tabel 4.5ℓ Singular values of SYS6-1

Approach	Singular values		
	N=300	N=1000	N=10000
Page	51.577	49.367	46.514
	5.826	1.080	0.334
	0.329	0.184	0.050
	0.039	0.078	0.017
Hankel	93.000*	105.374*	102.038
	44.210	6.984	1.713
	1.937	1.937	0.234
	0.975	0.389	0.194

* Error occurred in the execution of the identification algorithm There is no solution P>0,i.e. the realized sequence is not a covariance sequence

Tabel 4.6a Laboratory circumstances:Measurement 39.2

-Flume width	0.50 m
-Measured water velocity	134.0 /sec.
-Sample distance	0.01 m
-Water height	0.765

Tabel 4.6b **Intermediate identification results of the dunes**

	Page	Hankel
ERR1	$0.18\ 10^{-1}$	$0.80\ 10^{-2}$
Eigenvalues	0.865±j0.318	0.862±j0.337
Singular values	3.064	5.548
	1.251	4.762
	0.119	0.298
	0.054	0.228
	0.040	0.168
	0.036	0.134
	0.012	0.104
	0.007	0.098
	0.006	0.060
	0.004	0.047
	0.003	0.039
	0.002	0.030
		ect.
RARE	$0.2\ 10^{-30}$	$0.6\ 10^{-31}$
WINV	$0.39\ 10^{-2}$	$0.12\ 10^{-1}$

24 covariance matrices used to construct the Hankel and Page matrix. The block dimension of the Page matrix is 6*4

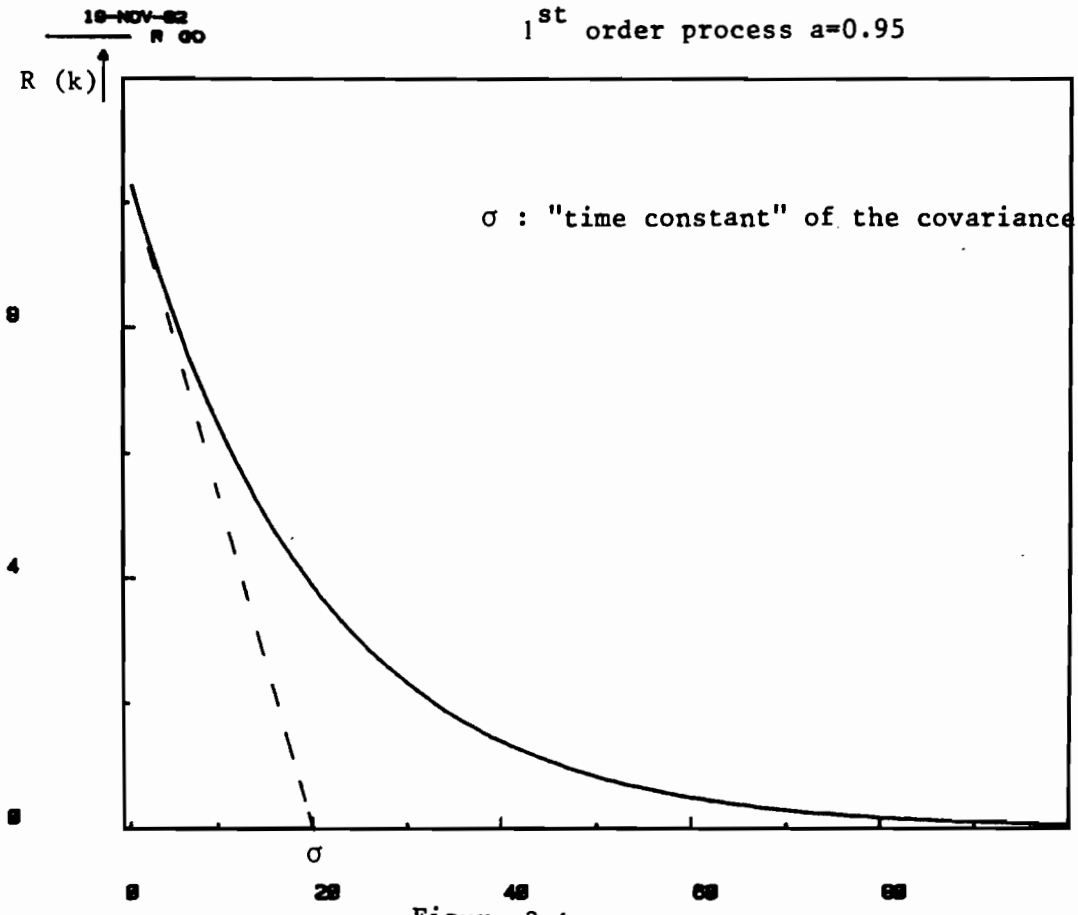


Figure 2.1
Covariance function of a first order process

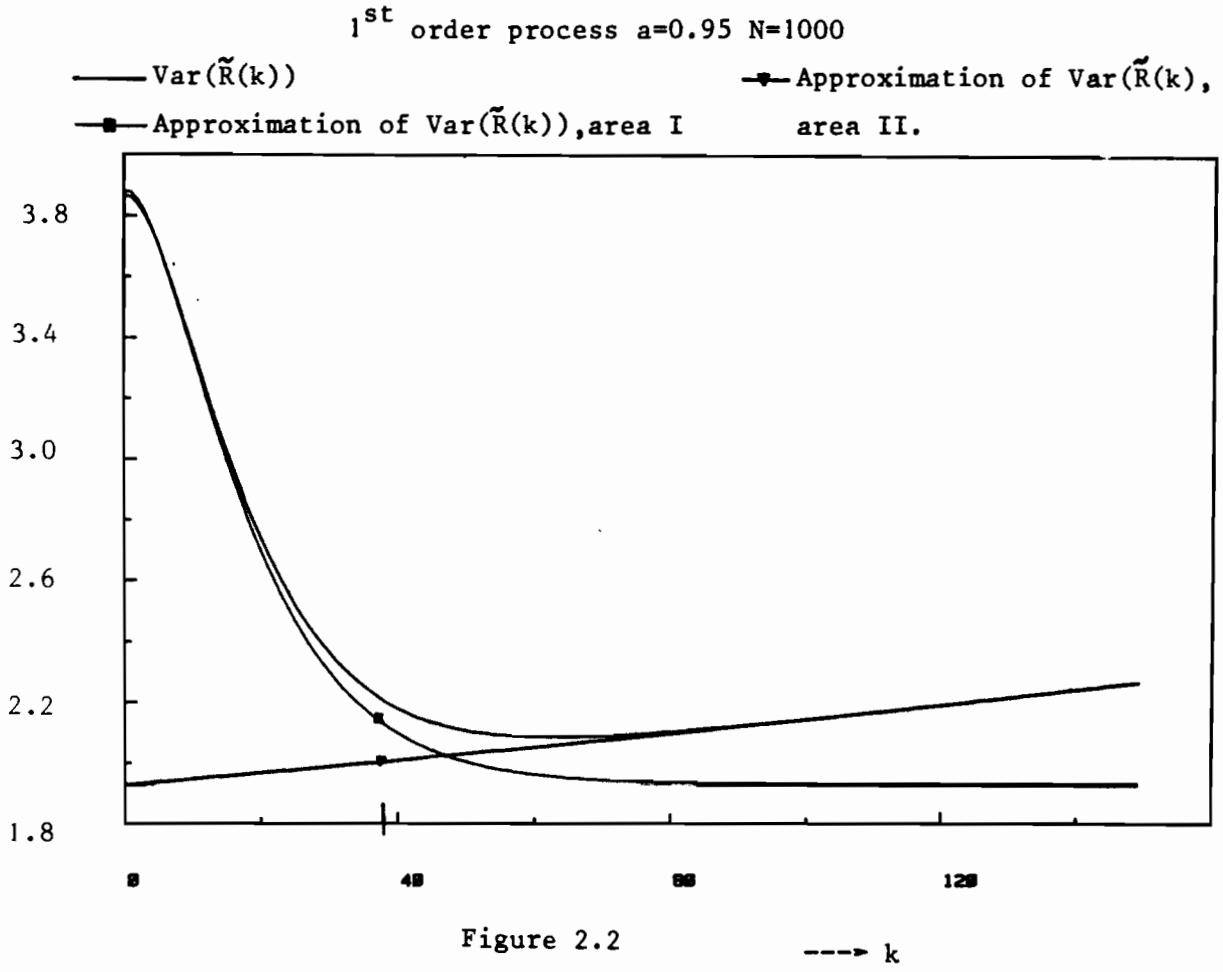

1st order process a=0.95 N=1000
Var(R̃(k))
Approximation of Var(R̃(k)), area I
Approximation of Var(R̃(k), area II.
3.8
3.4
3.0
2.6
2.2
1.8
0
40
80
120
---> k

Figure 2.2

$\tilde{R}_{\xi}(k)$: the covariance of the "white" noise generator 1, estimated with N=800 samples.

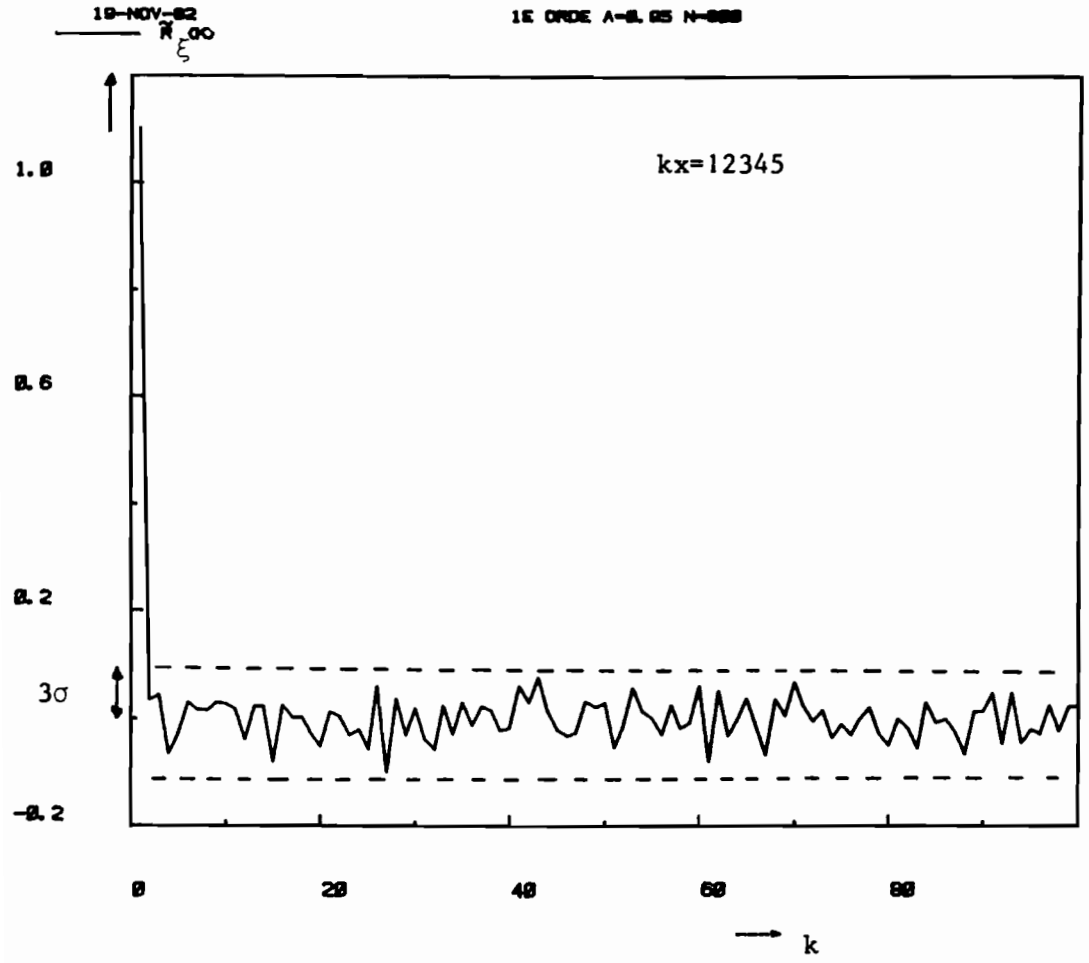
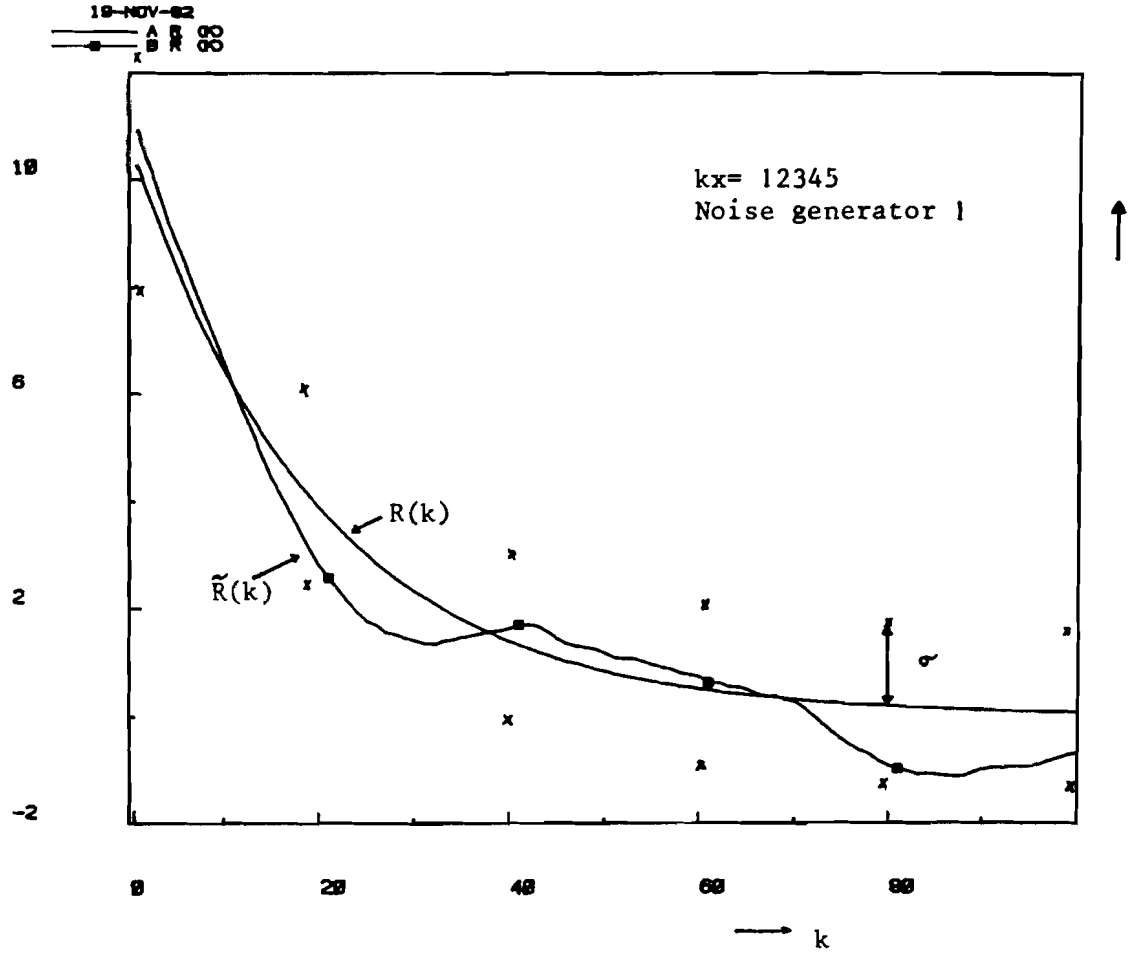


Figure 2.3 The covariance of the "white" noise generator.

A: R(k): the covariance
B: $\tilde{R}(k)$: the estimated covarian

1^{st} order process a=0.95 N=800



x=$\sqrt{\text{Var}(\tilde{R}(k))}$, computed from figure 2.5(A)

Figure 2.4 The covariance of a first order process.

A: the variance of the estimated covariance, calculated with R(k)
B: the approximated variance of the estimated covariance for area I, calculated with R(k).
C: the approximated variance of the estimated covariance for area II, calculated with R(k).
D: the variance of the estimated covariance, calculated with $\tilde{R}(k)$.
E: the approximated variance of the estimated covariance for area I, calculated with $\tilde{R}(k)$.
F: the approximated variance of the estimated covariance for area II, calculated with $\tilde{R}(k)$.

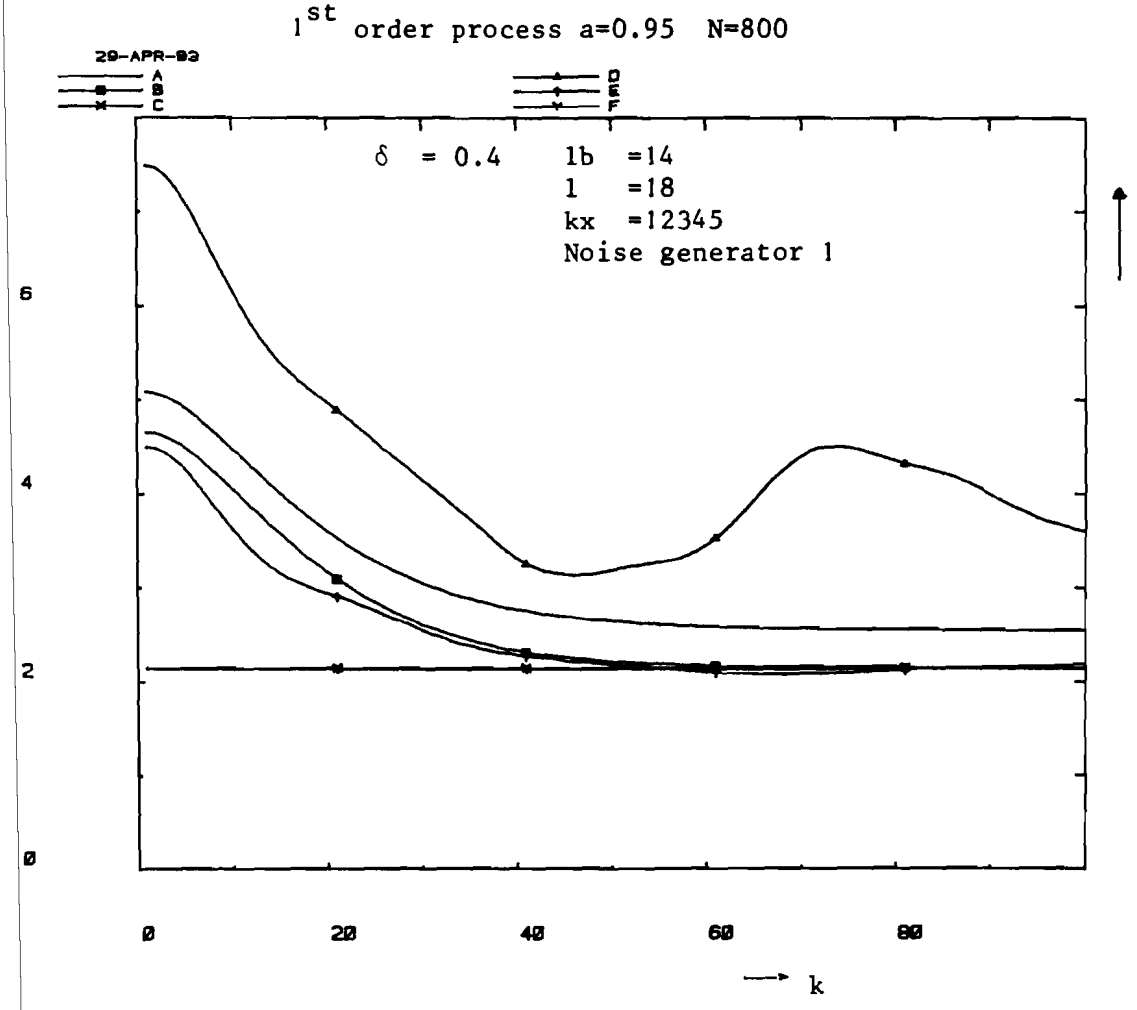


Figure 2.5 The variance of the estimated covariance

A : the variance of the estimated covariance, calculated with R(k)
B: the approximated variance of the estimated covariance for area I, calculated with R(k).
C: the approximated variance of the estimated covariance for area II calculated with R(k).
D: the variance of the estimated covariance, calculated with $\tilde{R}(k)$
E: the approximated variance of the estimated covariance for area I, calculated with $\tilde{R}(k)$.
F: the approximated variance estimated covariance for area II calculated with $\tilde{R}(k)$.

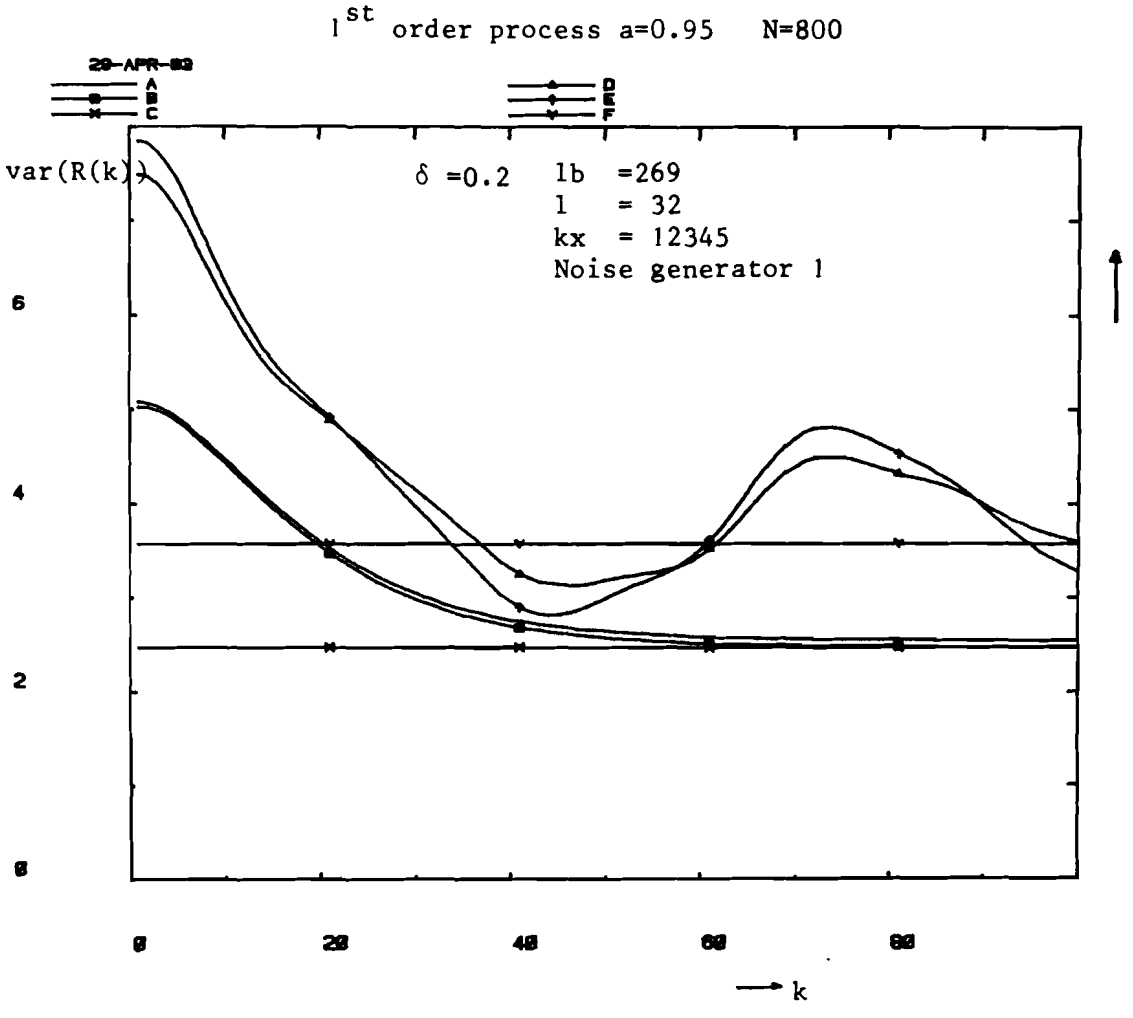


Figure 2.6 The variance of the estimated covariance.

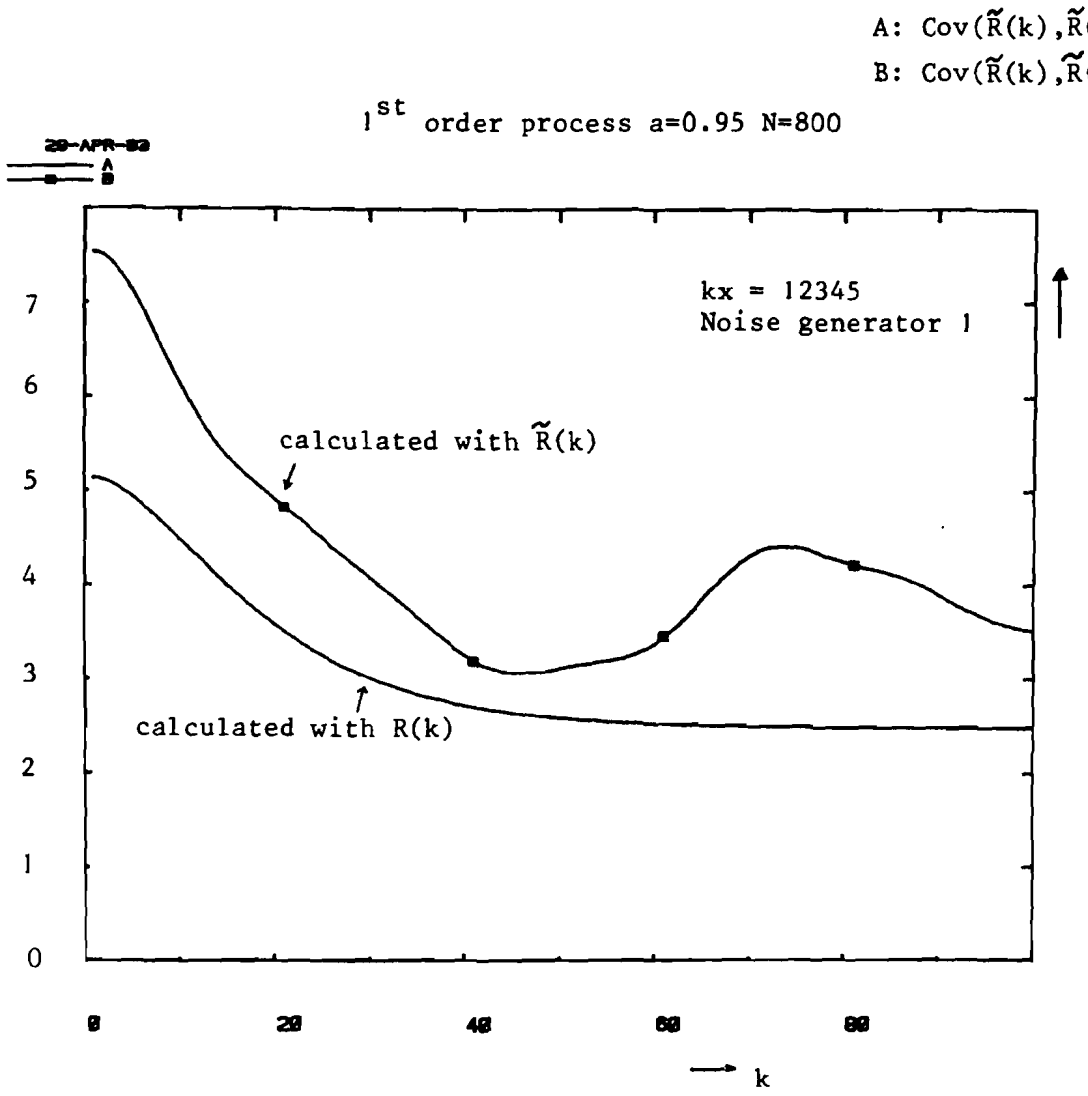


Figure 2.7 The covariance of the estimated covariance

A: Cov($\tilde{R}(k),\tilde{R}(k+1)$)
B: Cov($\tilde{R}(k),\tilde{R}(k+2)$)
C: Cov($\tilde{R}(k),\tilde{R}(k+2)$)
D: COV($\tilde{R}(k),\tilde{R}(k+3)$)
E: Cov($\tilde{R}(k),\tilde{R}(k+4)$)
F: Cov($\tilde{R}(k),\tilde{R}(k+5)$)

1^{st} order process a=0.95 N=800

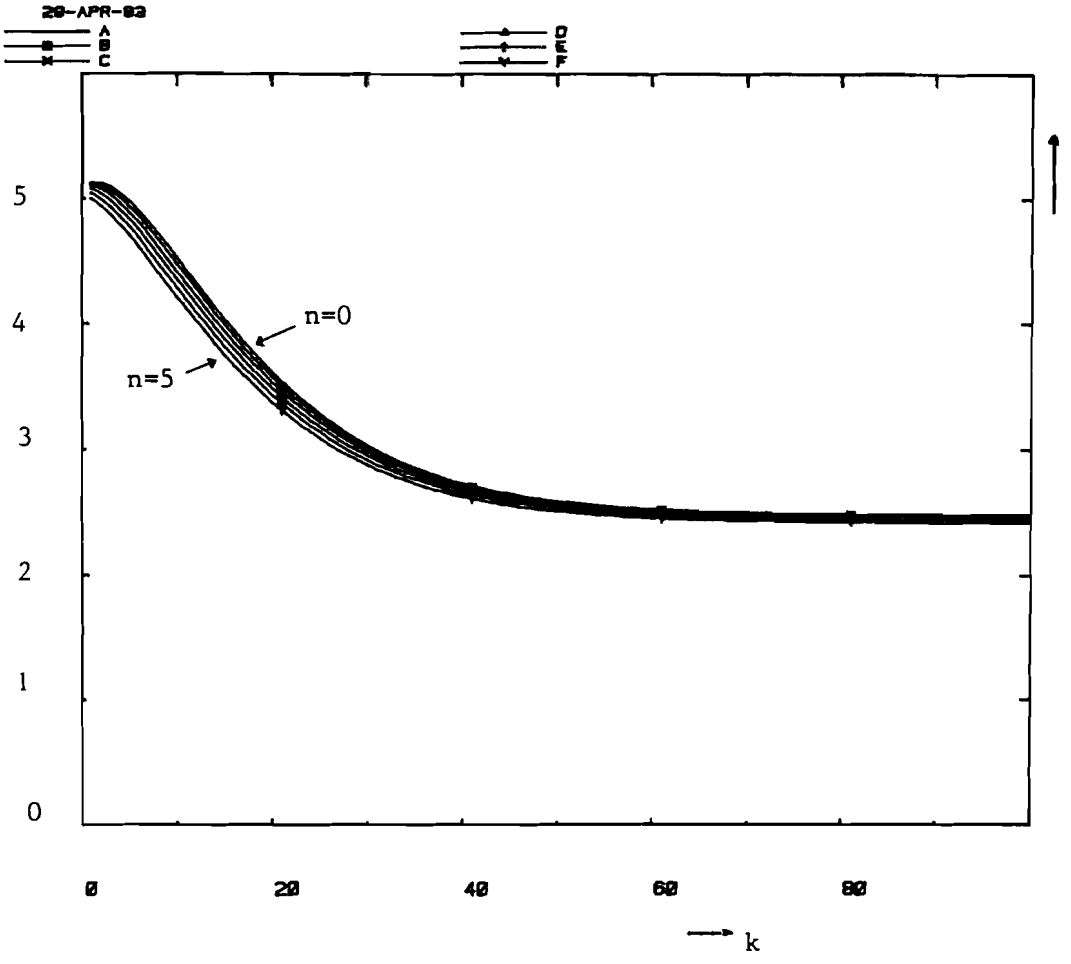


Figure 2.8 The covariance of the estimated covariance, calculated with R(k)

A: Cov($\tilde{R}$(k),$\tilde{R}$(k+0))
B: Cov($\tilde{R}$(k),$\tilde{R}$(k+1))
C: Cov($\tilde{R}$(k),$\tilde{R}$(k+2))
D: Cov($\tilde{R}$(k),$\tilde{R}$(k+3))
E: Cov($\tilde{R}$(k),$\tilde{R}$(k+4))
F: Cov($\tilde{R}$(k),$\tilde{R}$(k+5))

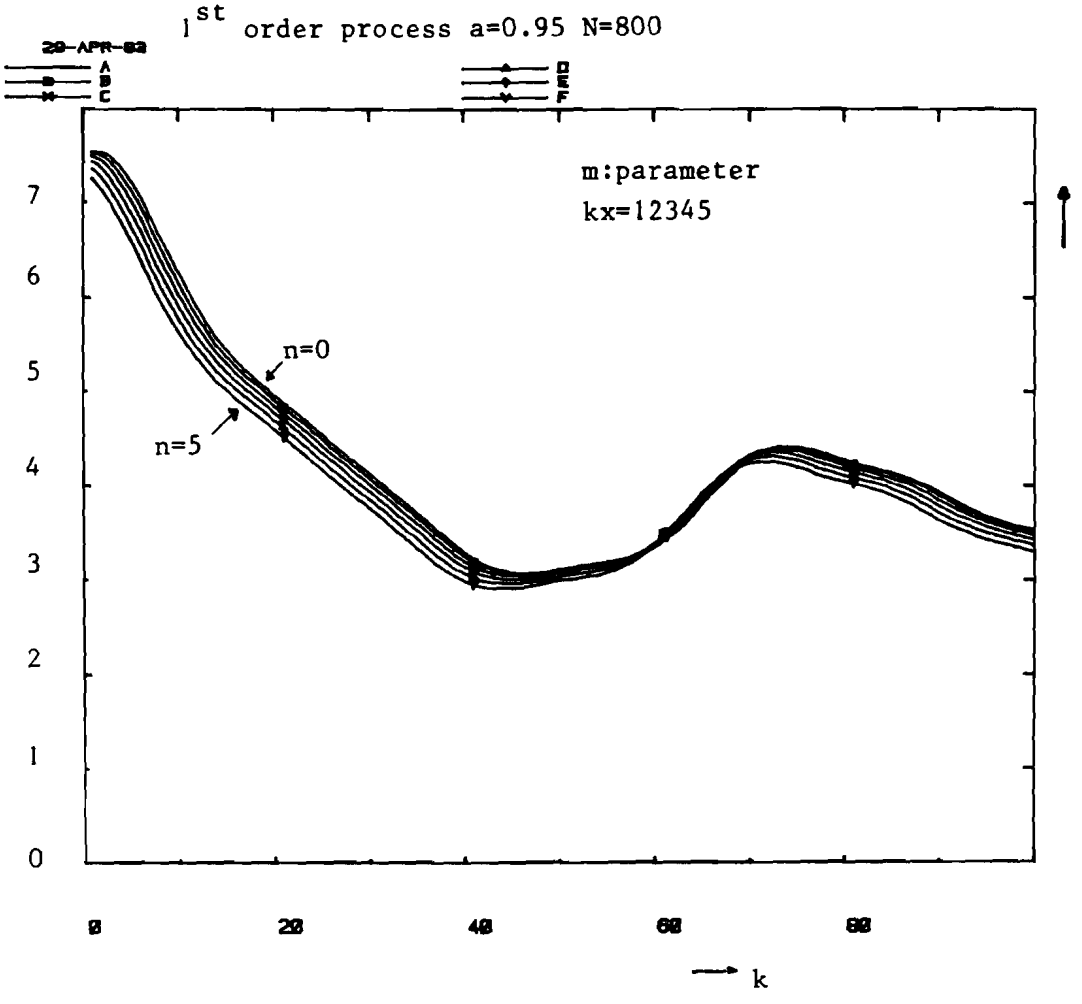


Figure 2.9 The covariance of the estimated covariance, calculated with $\tilde{R}$(k).

A: the relative variance of the estimated covariance:

$$\mathrm{var}(\tilde{R}(k))/\ R(k)$$

B: the relative variance of the estimated covariance:

$$\mathrm{var}(\tilde{R}(k))/\ \tilde{R}(k)$$

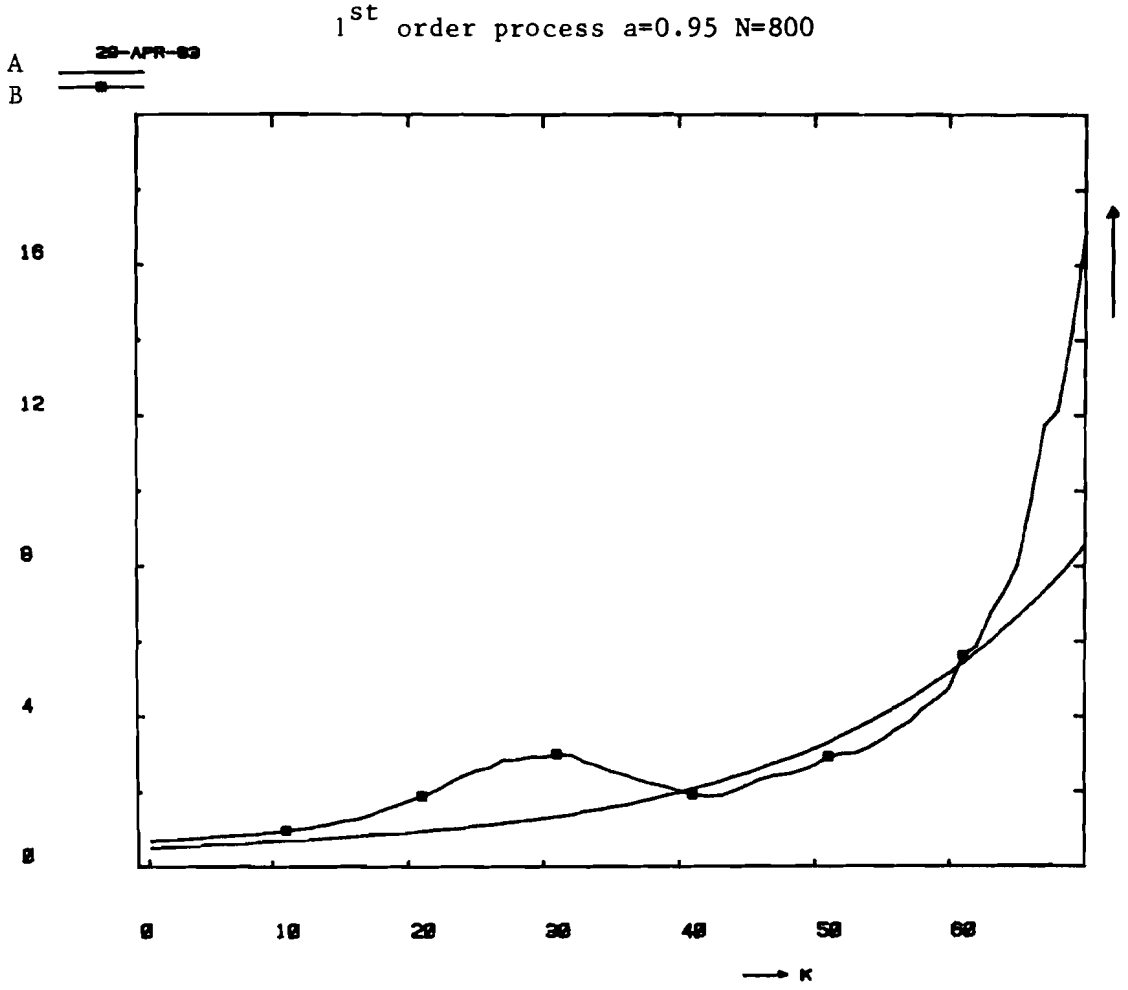


Figure 2.10 The relative variance of the estimated covariance

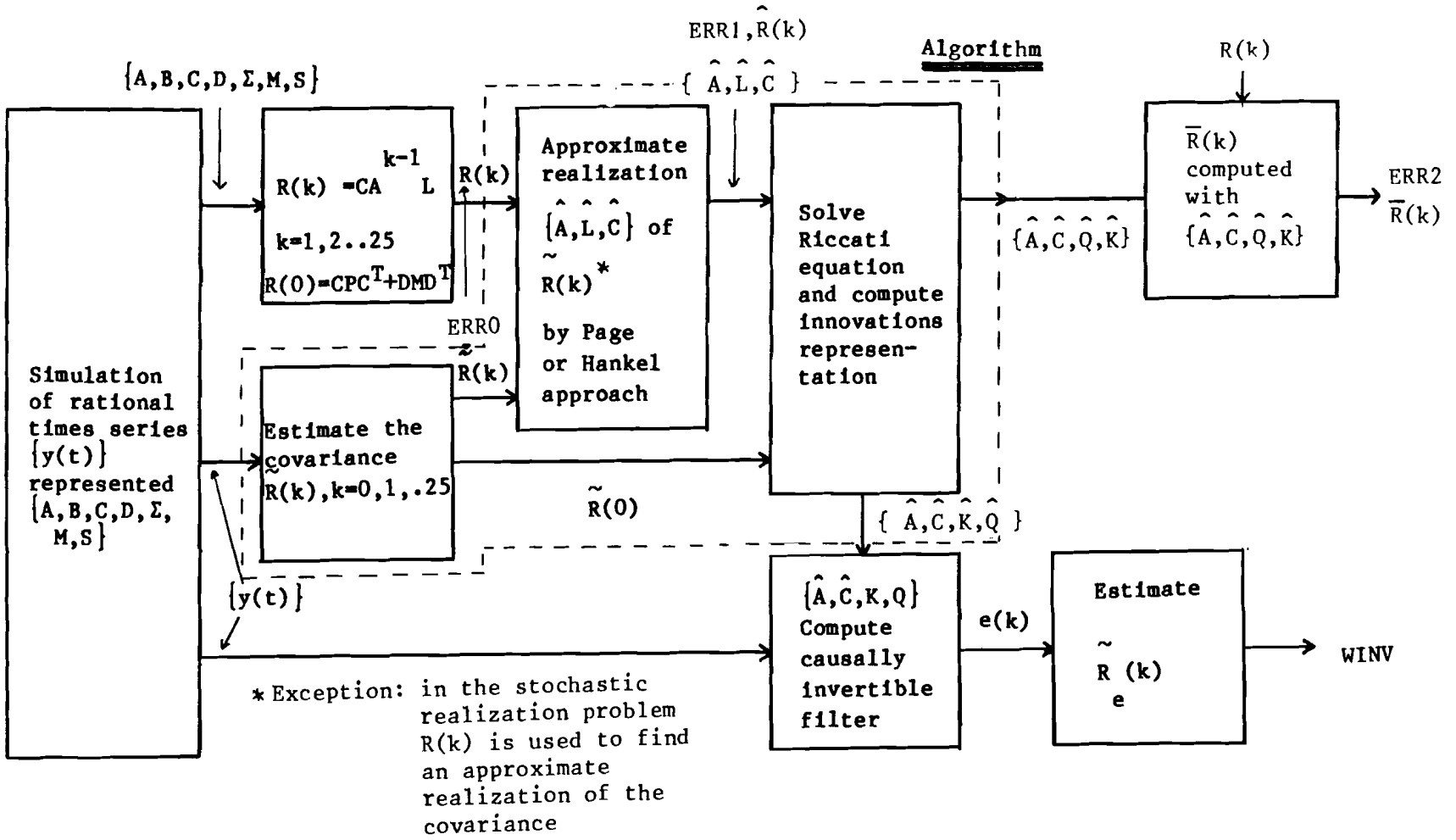


Figure 4.1a Block diagram of stochastic identification algorithm and corresponding tests

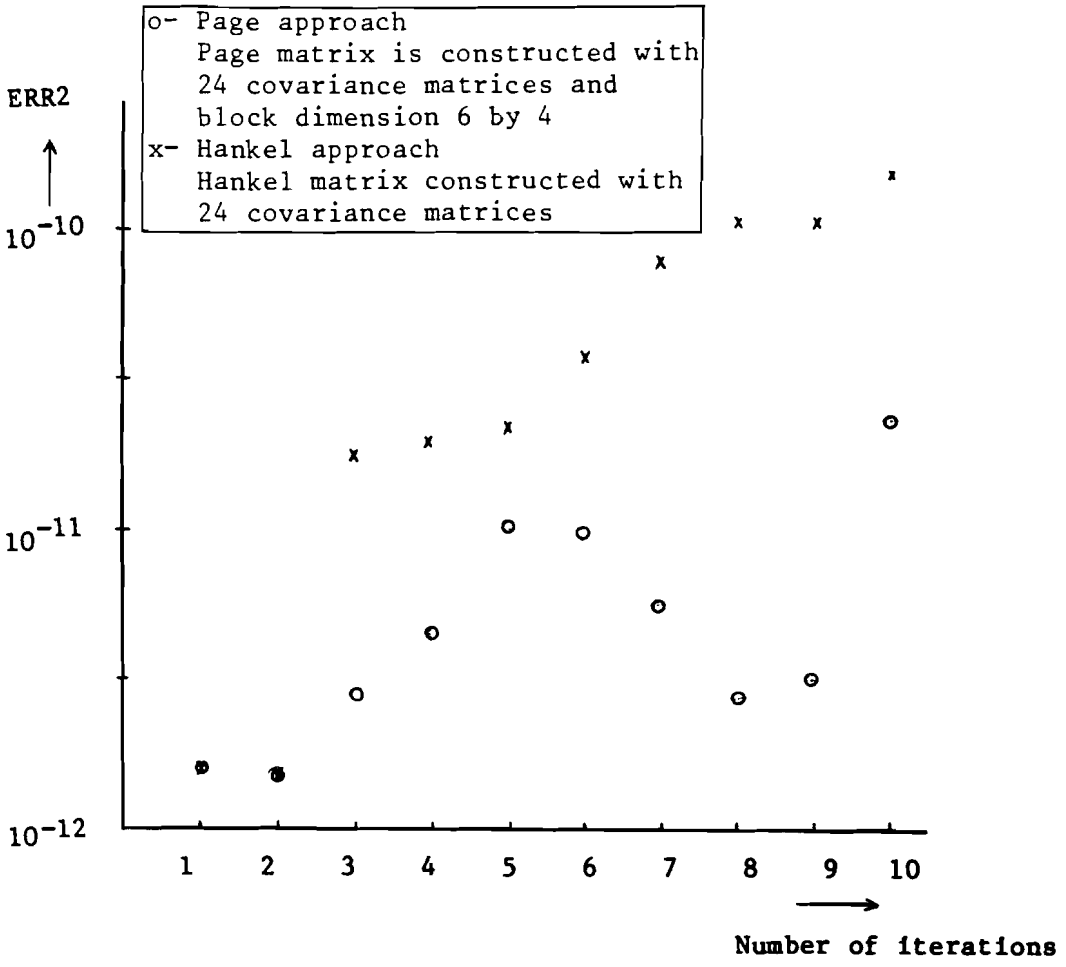


Figure 4.5 Stability test (SYS2)

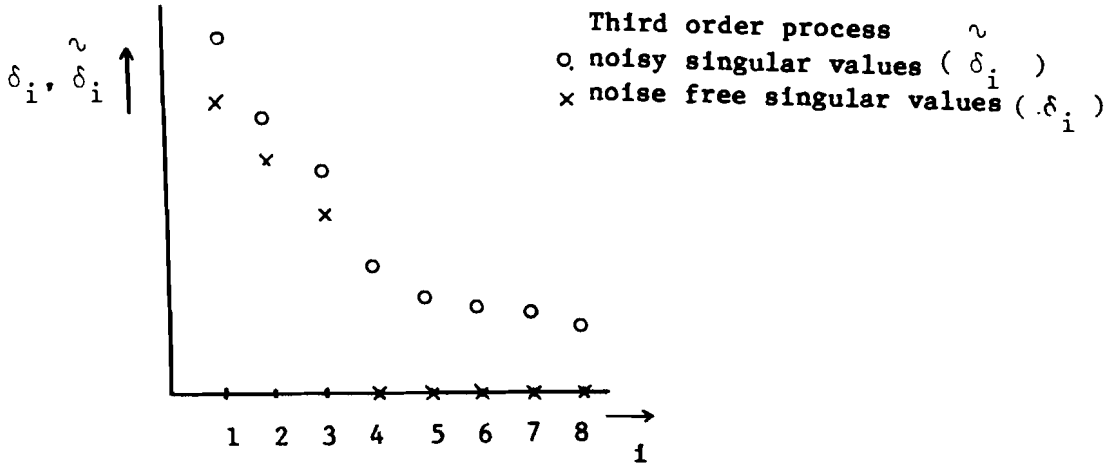


Figure 4.3 Singular values of a third order process in a noise free-and a noisy case.

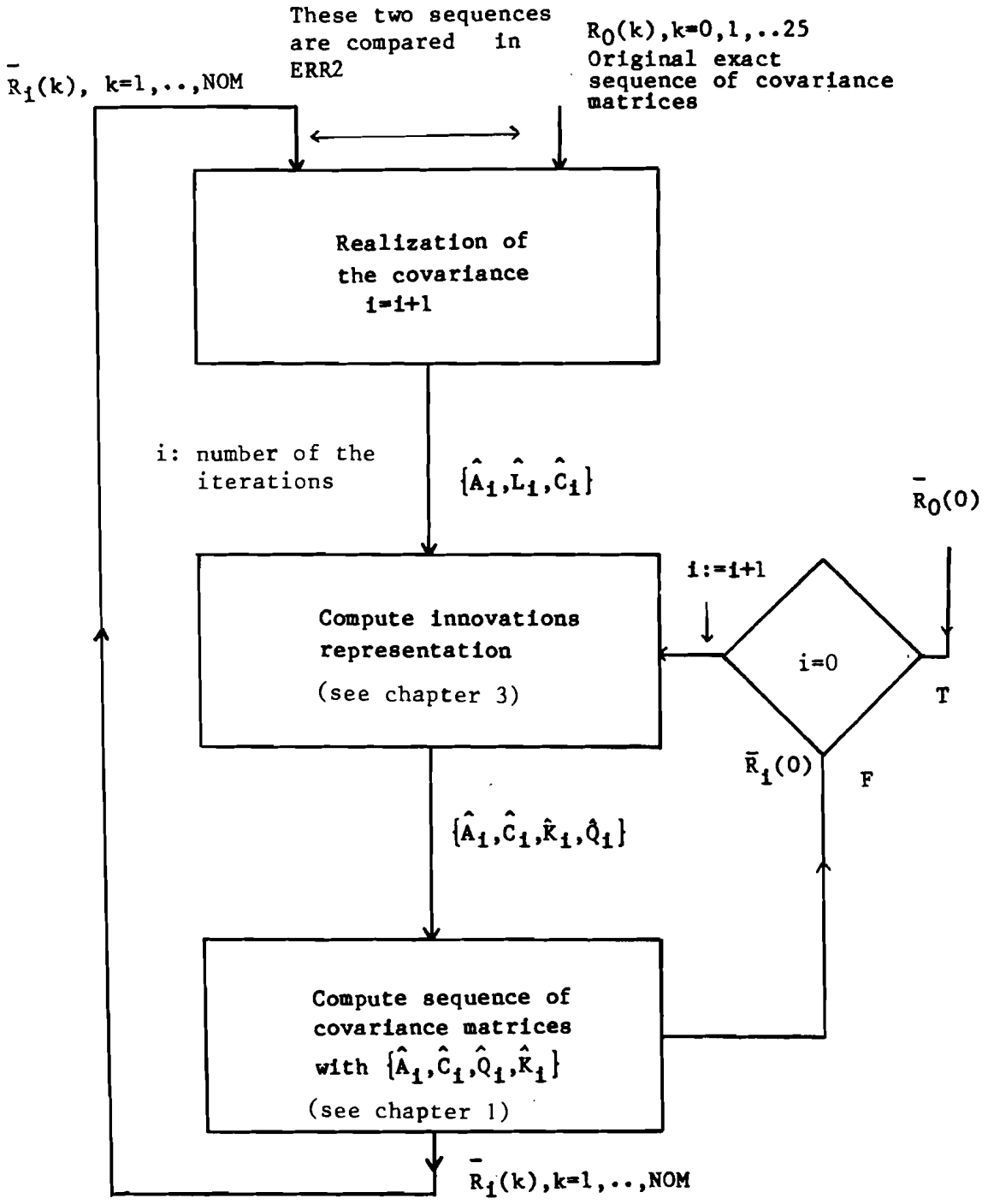


Figure 4.4 Stability test schematically depicted.

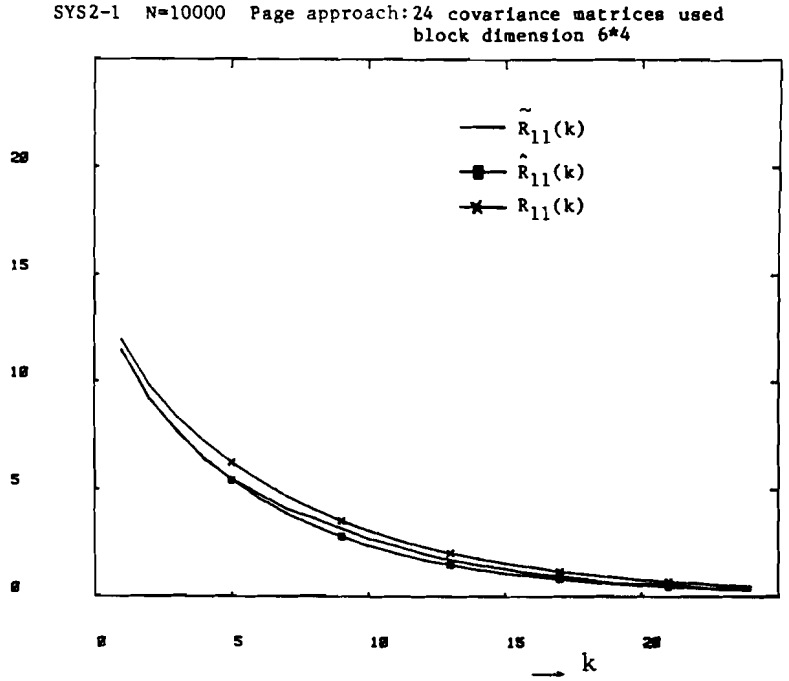


Figure 4.6a Element (1,1) of the covariance sequence of SYS2-1

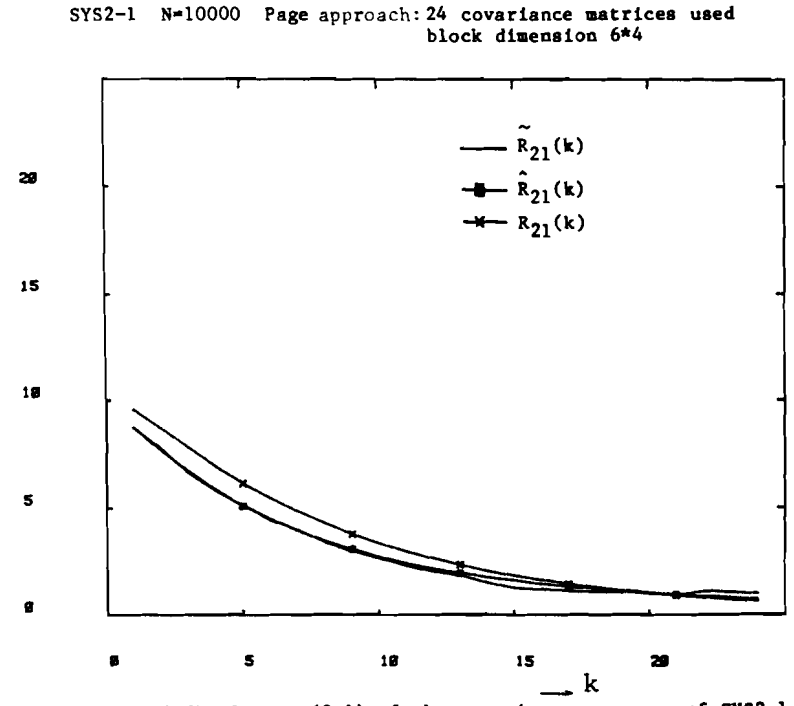


Figure 4.6b Element (2,1) of the covariance sequence of SYS2-1

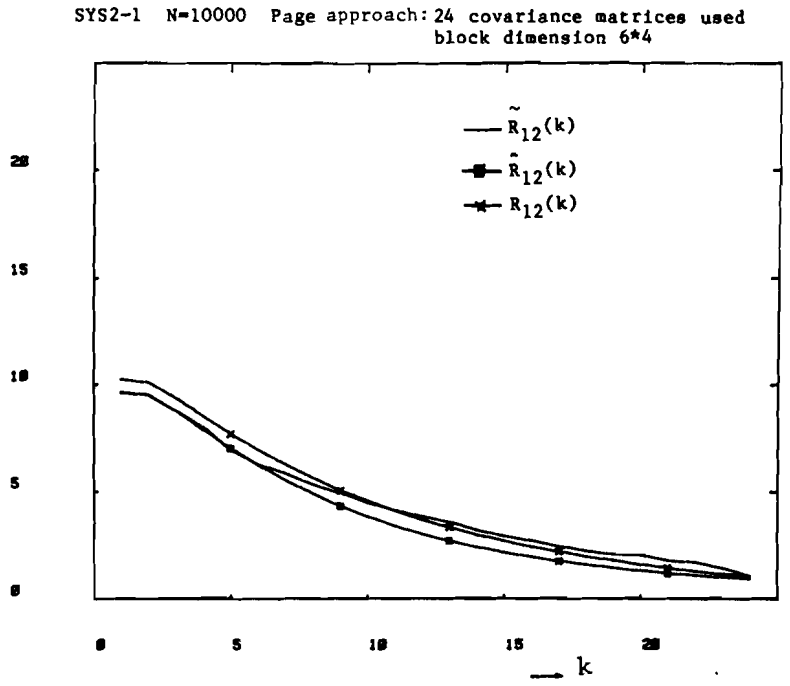


Figure 4.6c Element (1,2) of the covariance sequence of SYS2-1

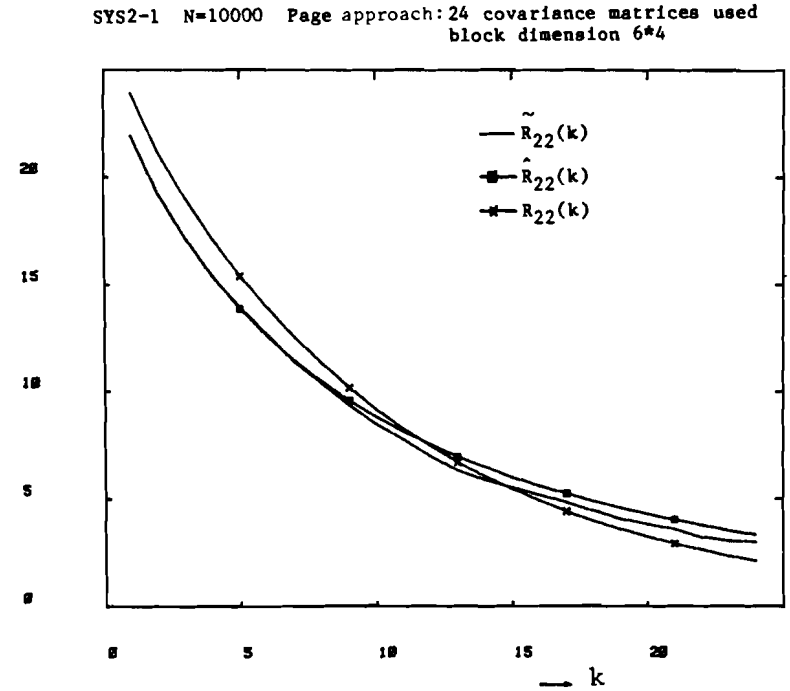


Figure 4.6d Element (2,2) of the covariance sequence of SYS2-1

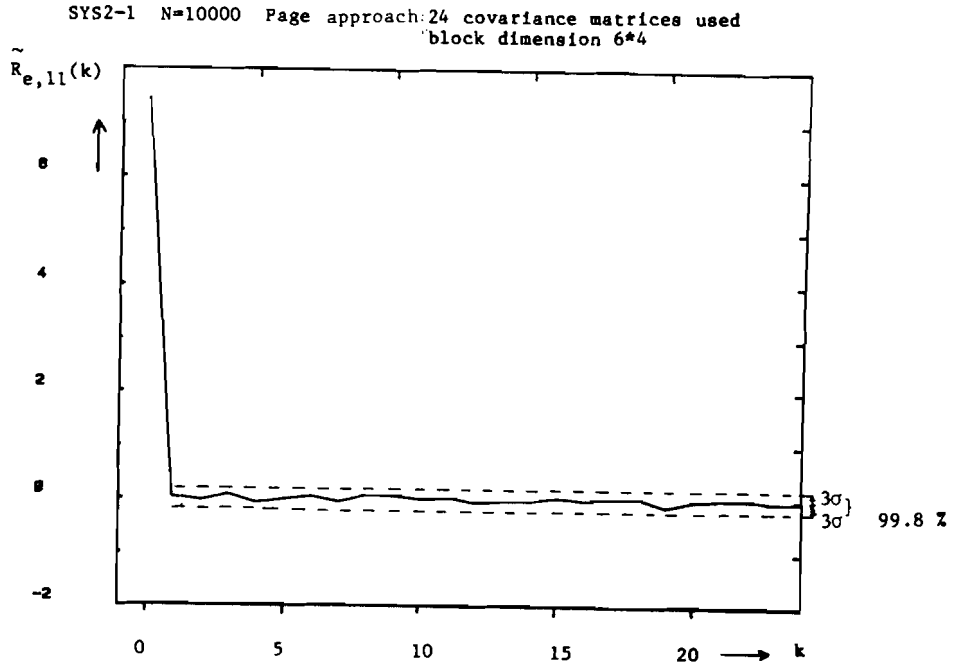


Figure 4.7a Element (1,1) of the covariance sequence of the innovations process of SYS2-1

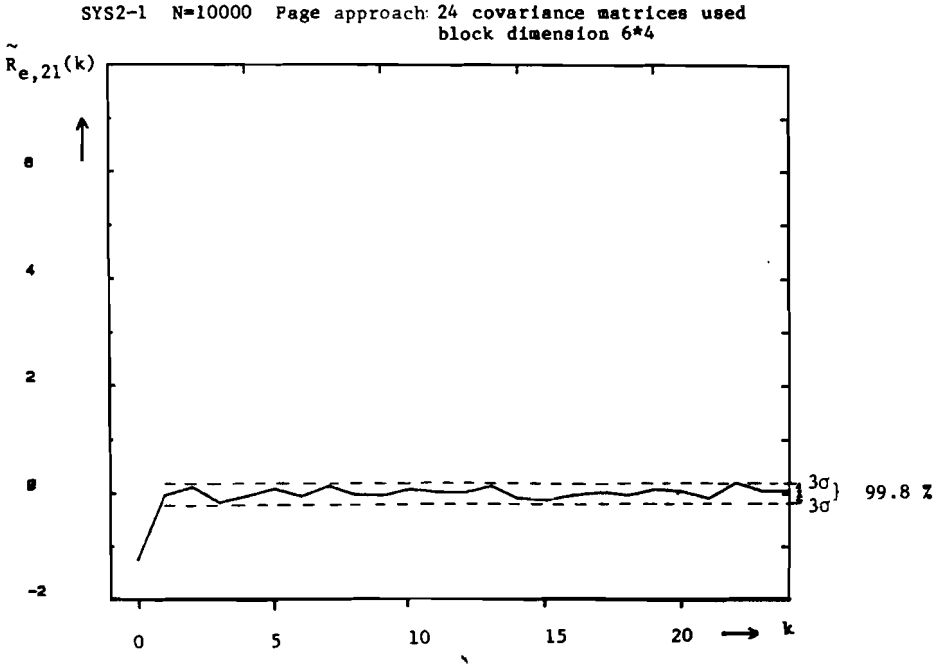


Figure 4.7b Element (2,1) of the covariance sequence of the innovations process of SYS2-1

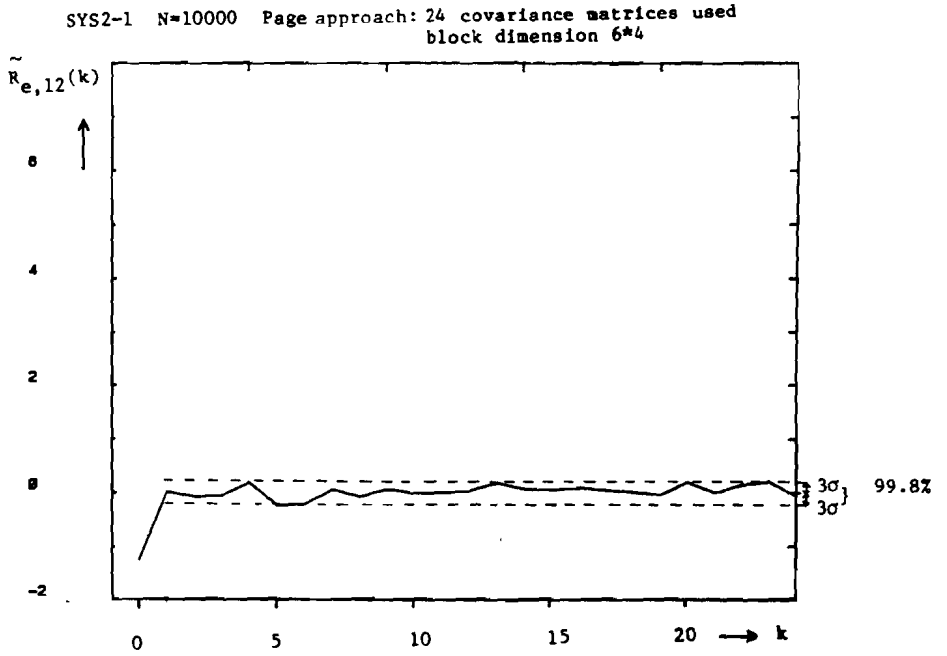


Figure 4.7c Element (2,1) of the covariance sequence of the innovations process of SYS2-1

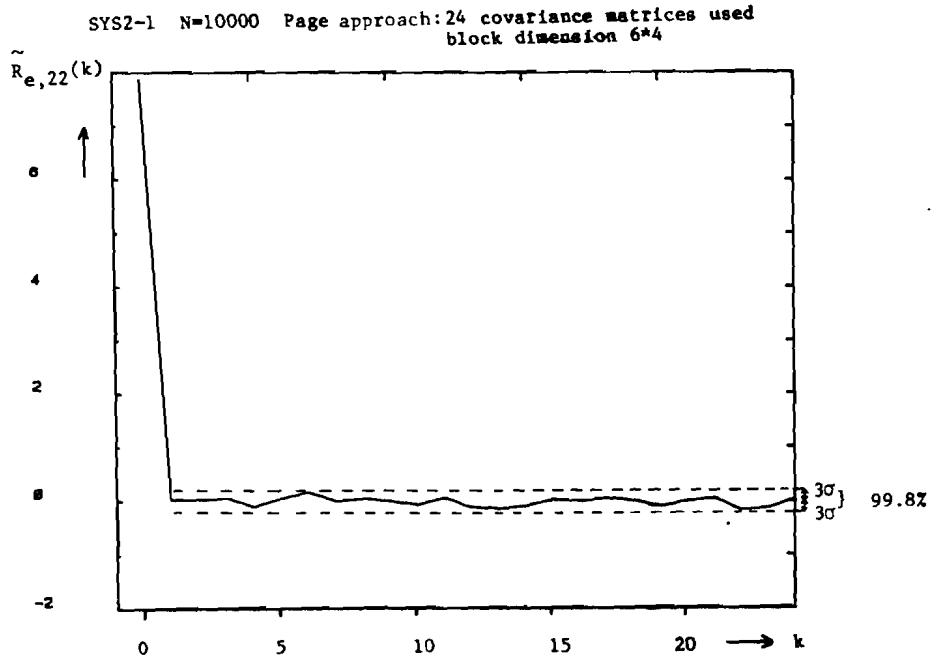


Figure 4.7d Element (2,2) of the covariance sequence of the innovations process of SYS2-1

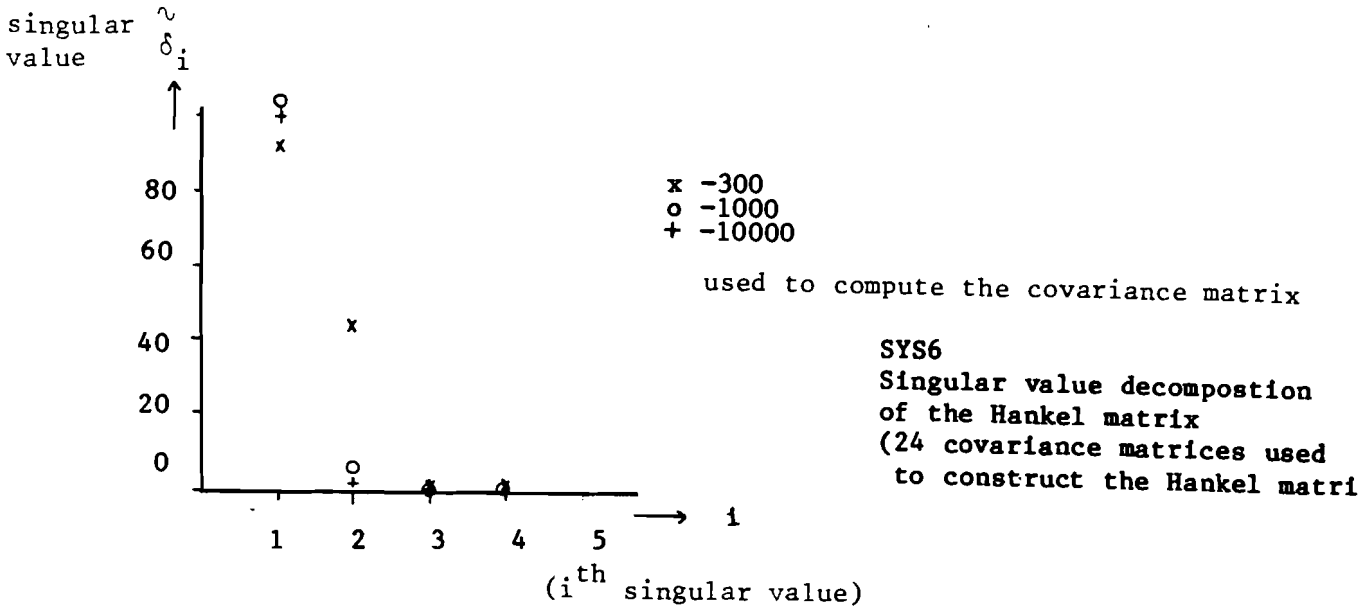


Figure 4.8a Singular value decomposition of the Hankel matrix SYS6-1

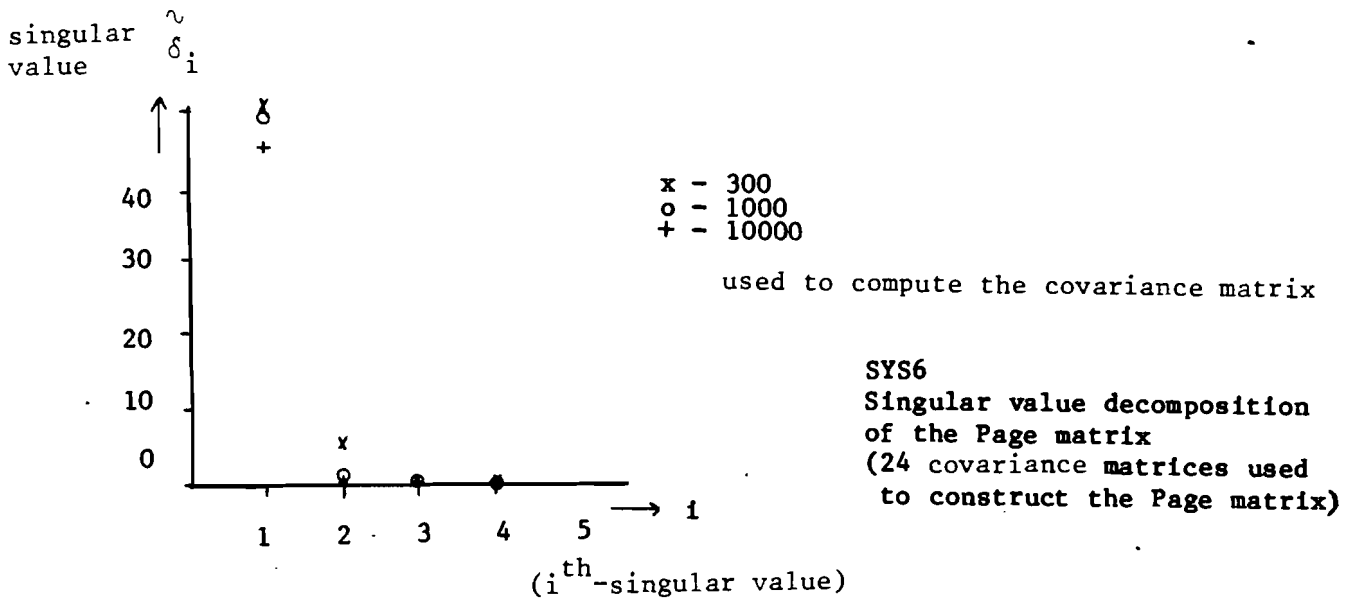
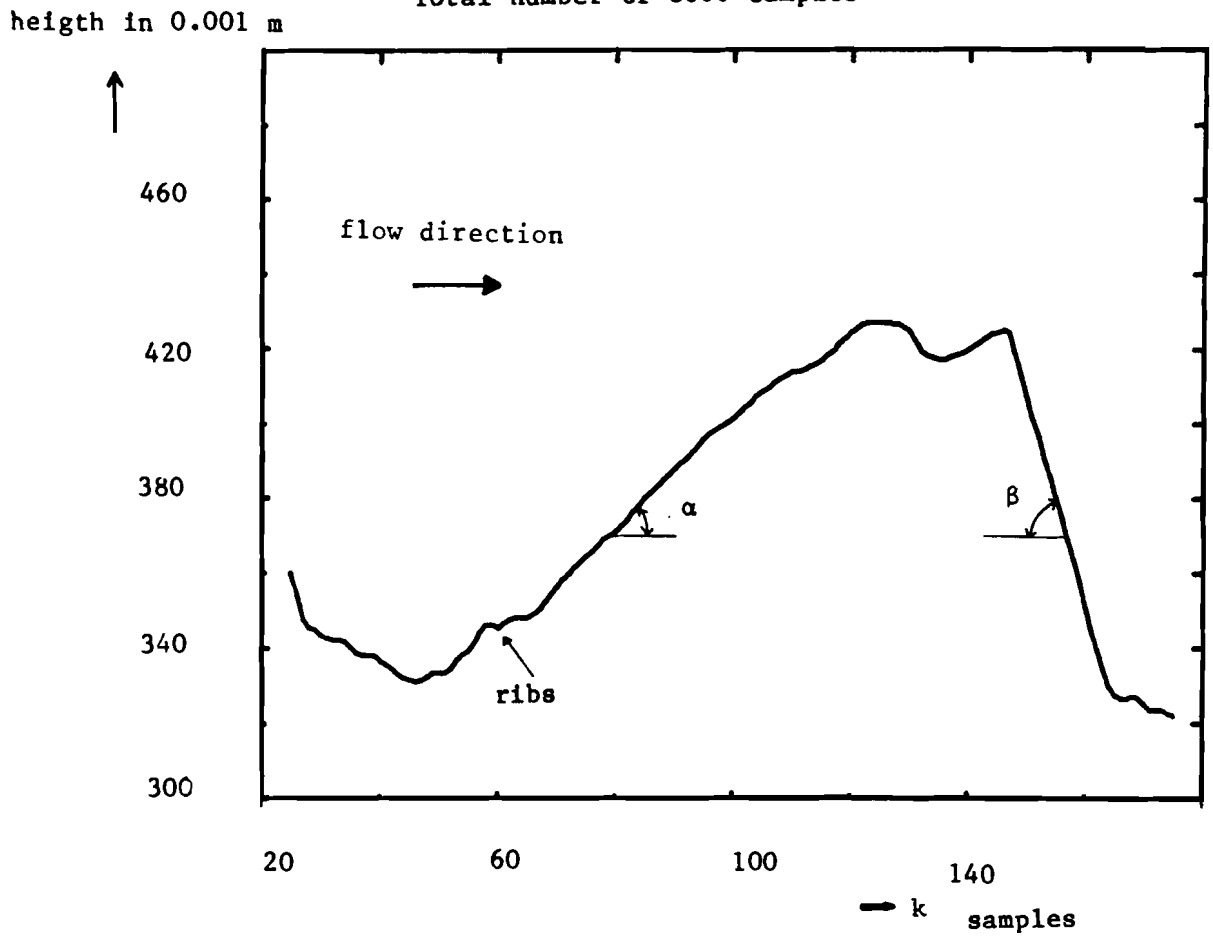


Figure 4.8b Singular value decomposition of the Page matrix SYS6-1

Dune measurement 39.2 (see section 4.7.1: Laboratory circumstances)



The sample distance is 0.01 m

Figure 4.9a Dune (measurement 39.2)

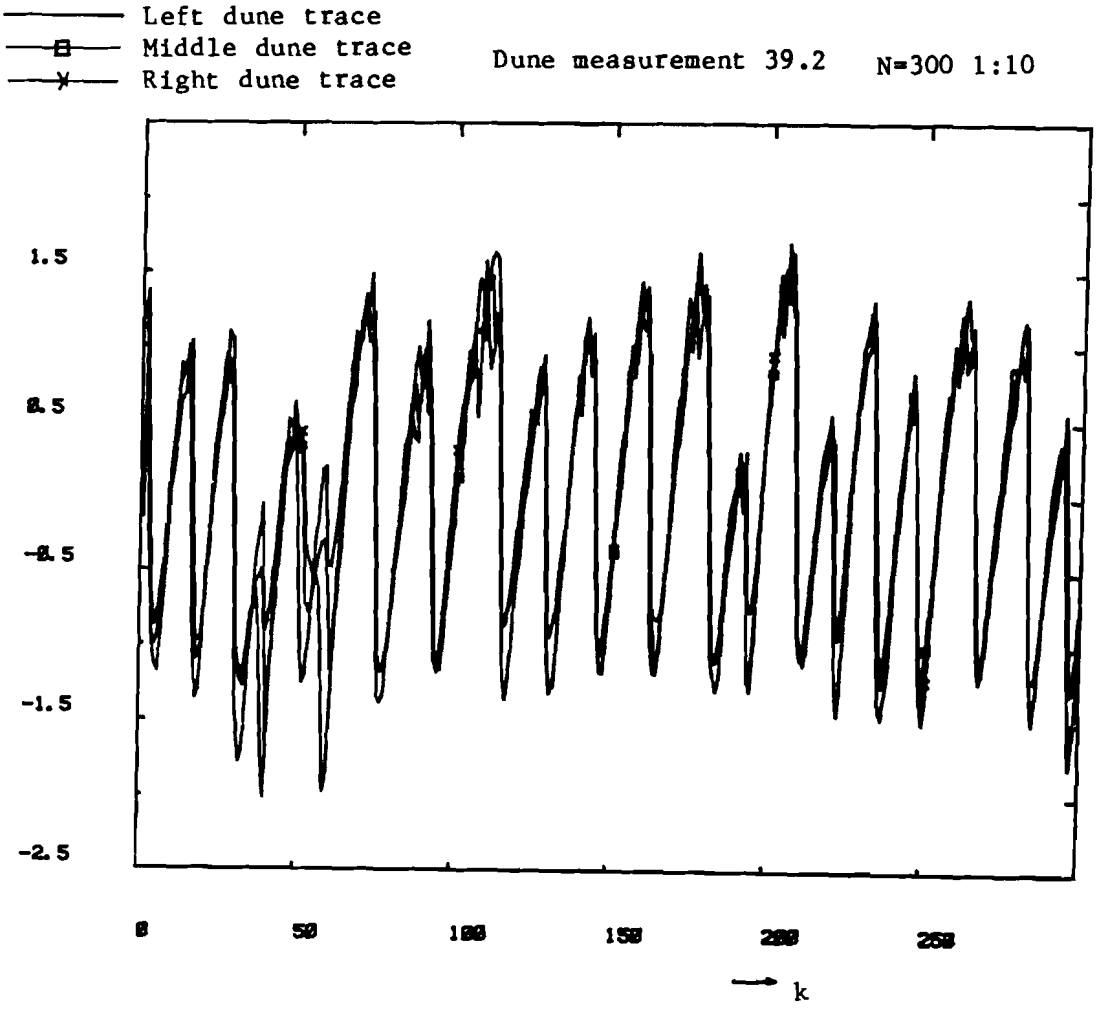


Figure 4.9b Dune measurement 39.2,

without bias.

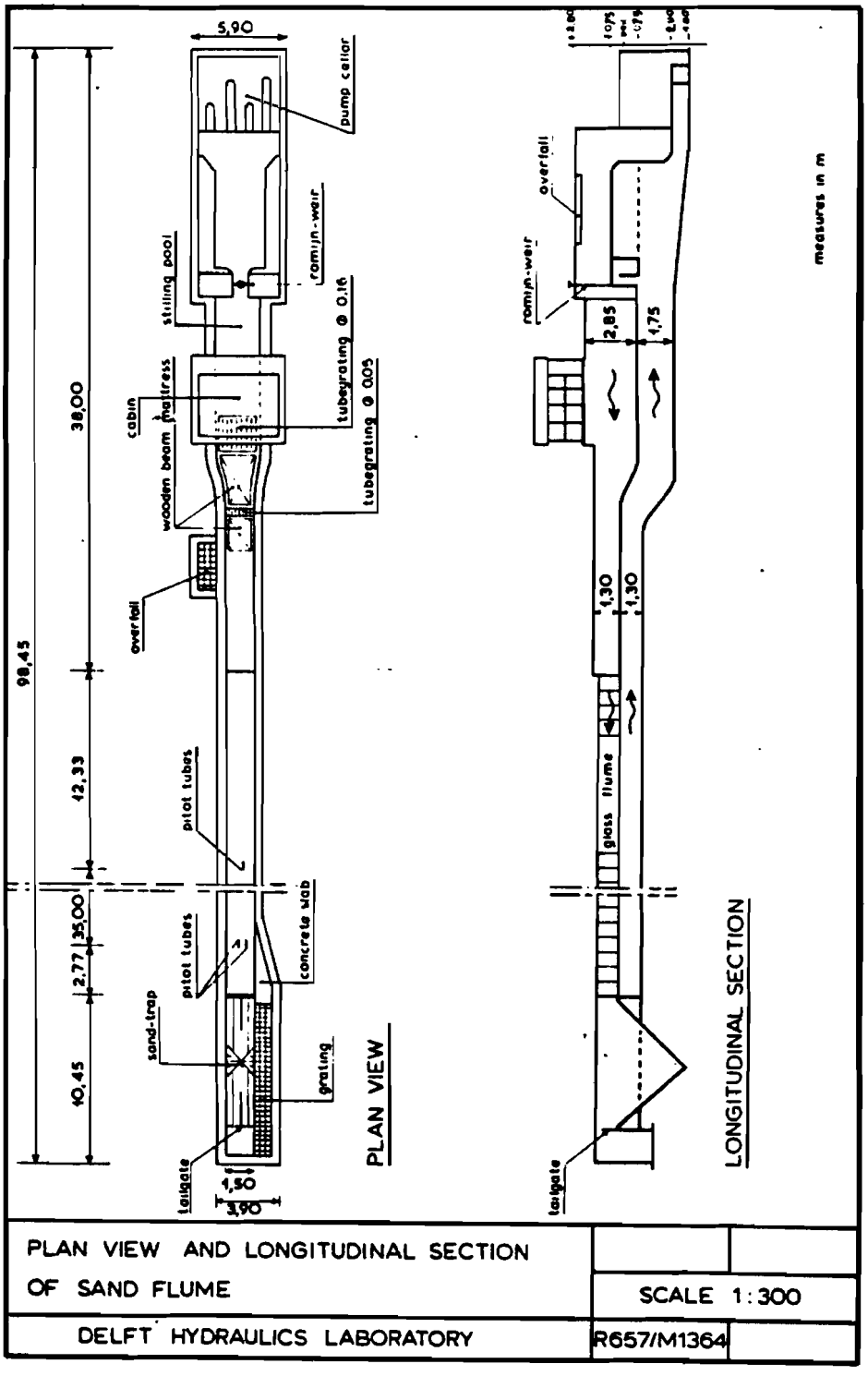
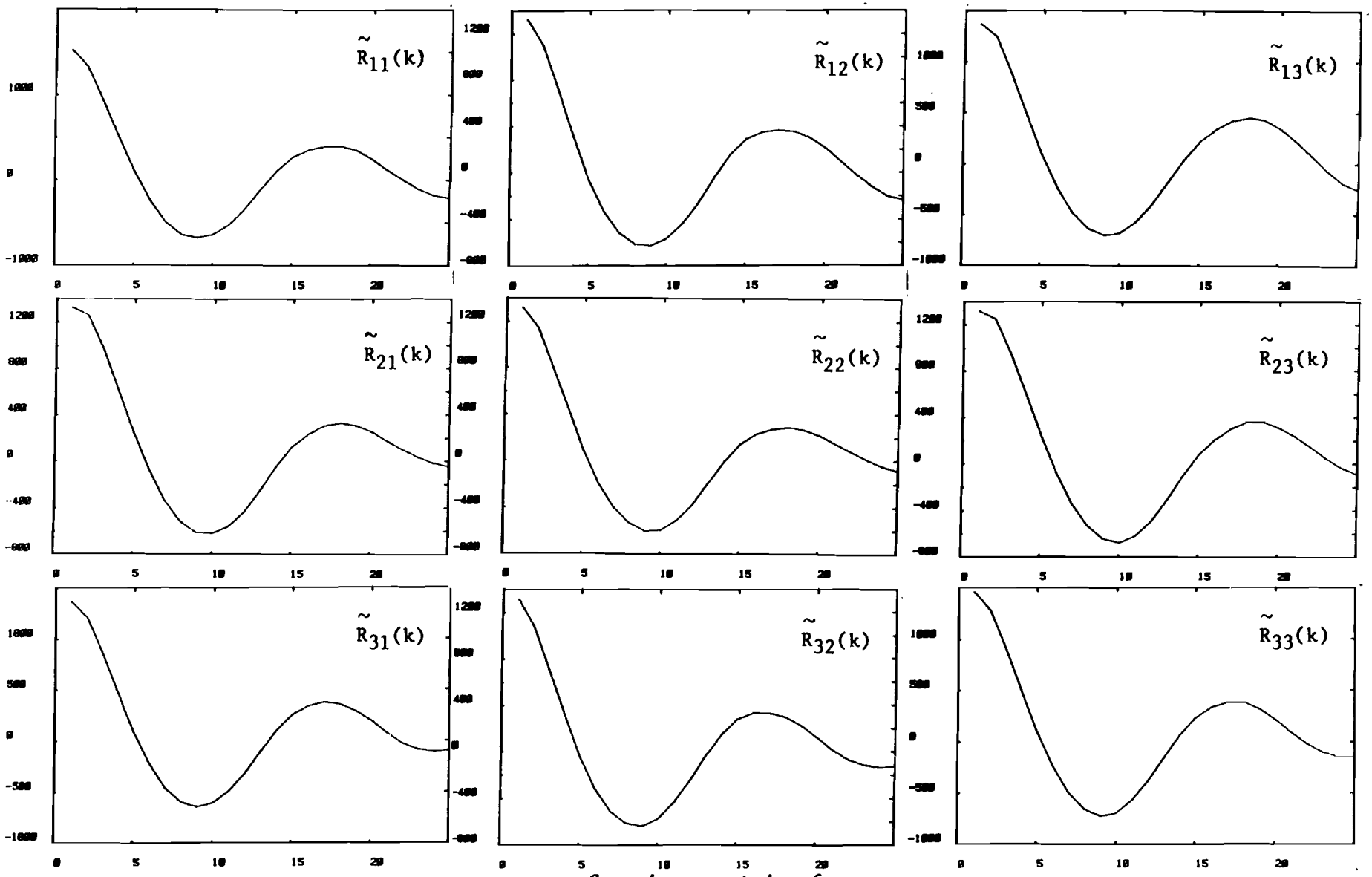
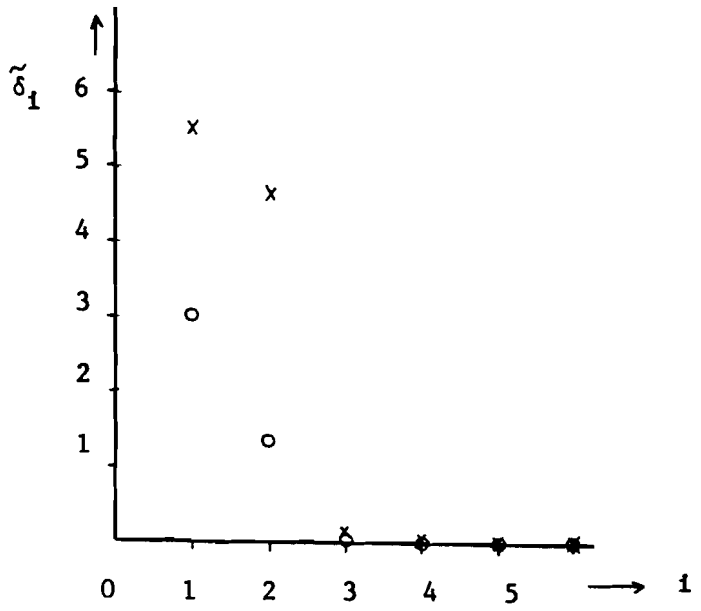


Figure 4.10 Sand flume



Covariance matrix of
the 3 dune traces

Figure 4.11



o Page approach,
24 covariance matrices
used to construct the
Page matrix.
Block dimension 6*4

x Hankel approach,
24 covariance matrices
used to construct
Hankel matrix

Figure 4.12 Singular values of the Hankel and Page matrix, corresponding with the dune identification (39.2)

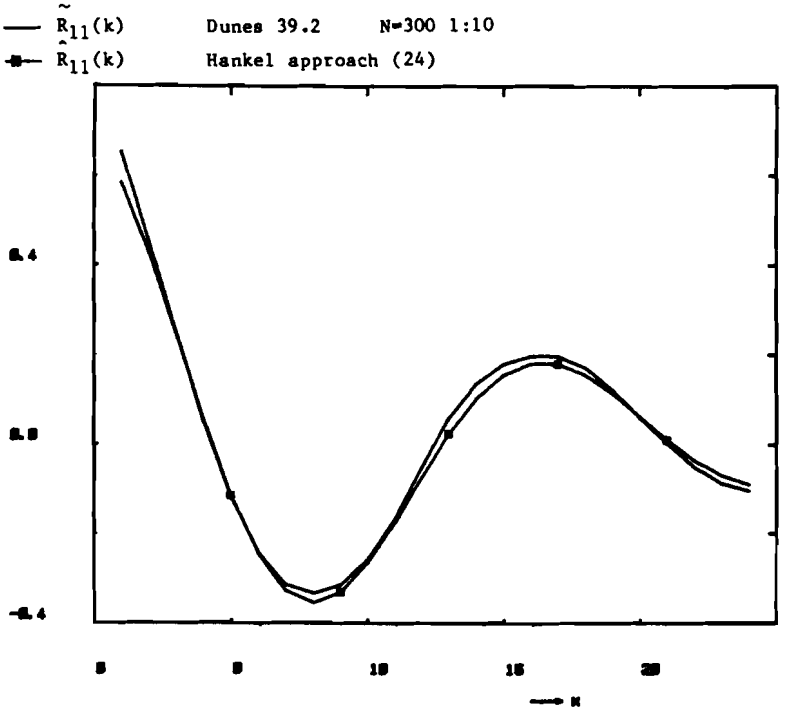


Figure 4.13a Element (1,1) of the covariance sequence associated with the dunes (39.2)

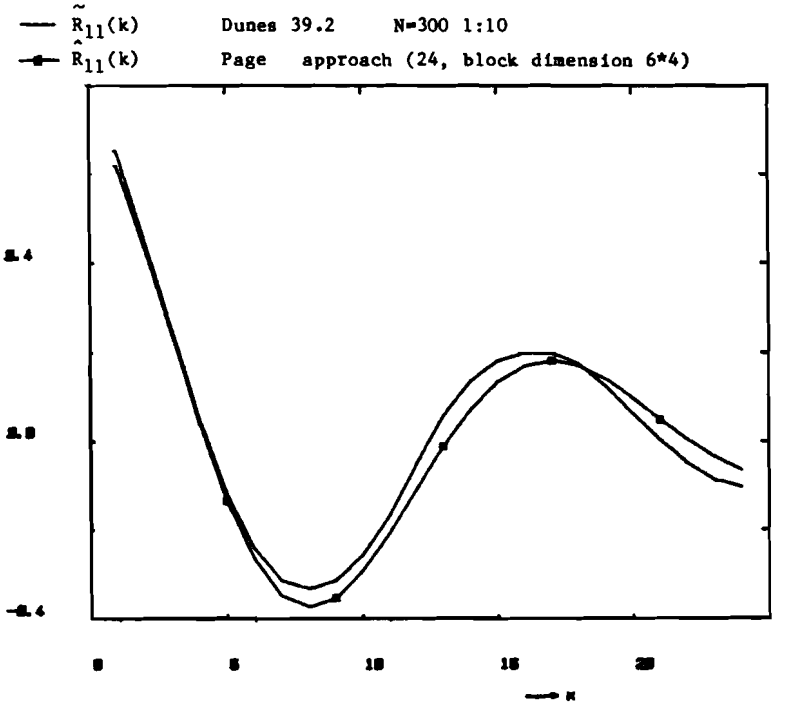


Figure 4.13b Element (1,1) of the covariance sequence associated with the dunes (39.2)

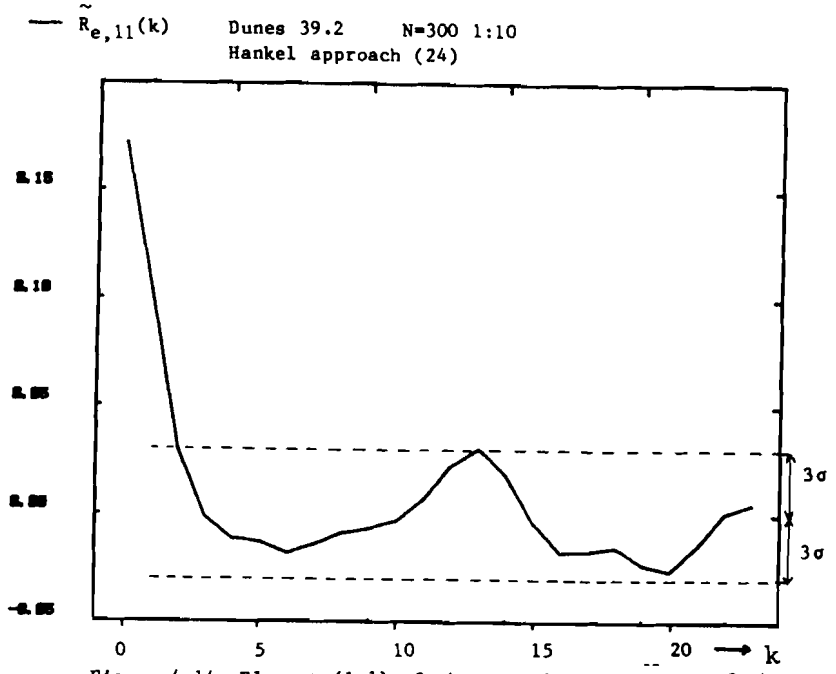


Figure 4.14a Element (1,1) of the covariance sequence of the innovations process (Hankel approach)

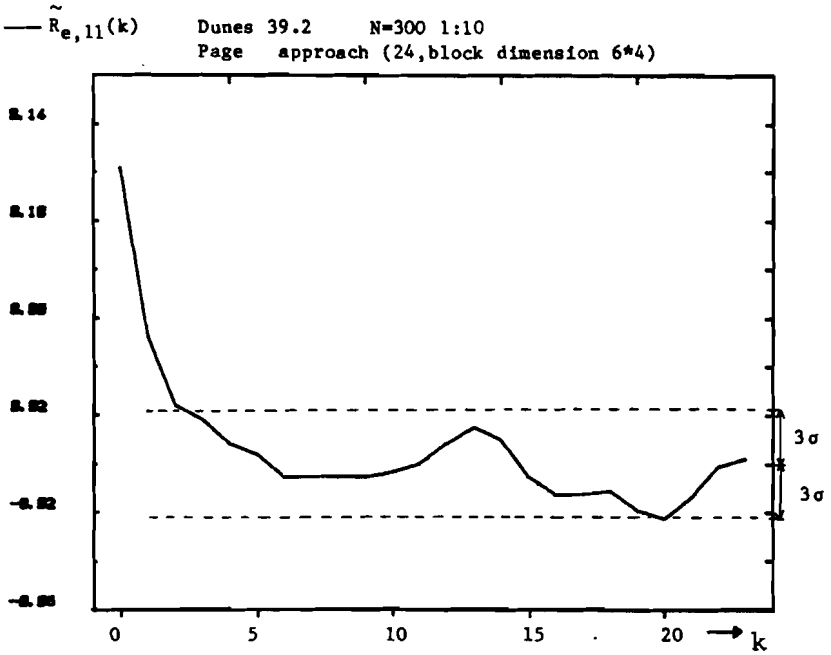


Figure 4.14b Element (1,1) of the covariance sequence of the innovations process (Page approach)

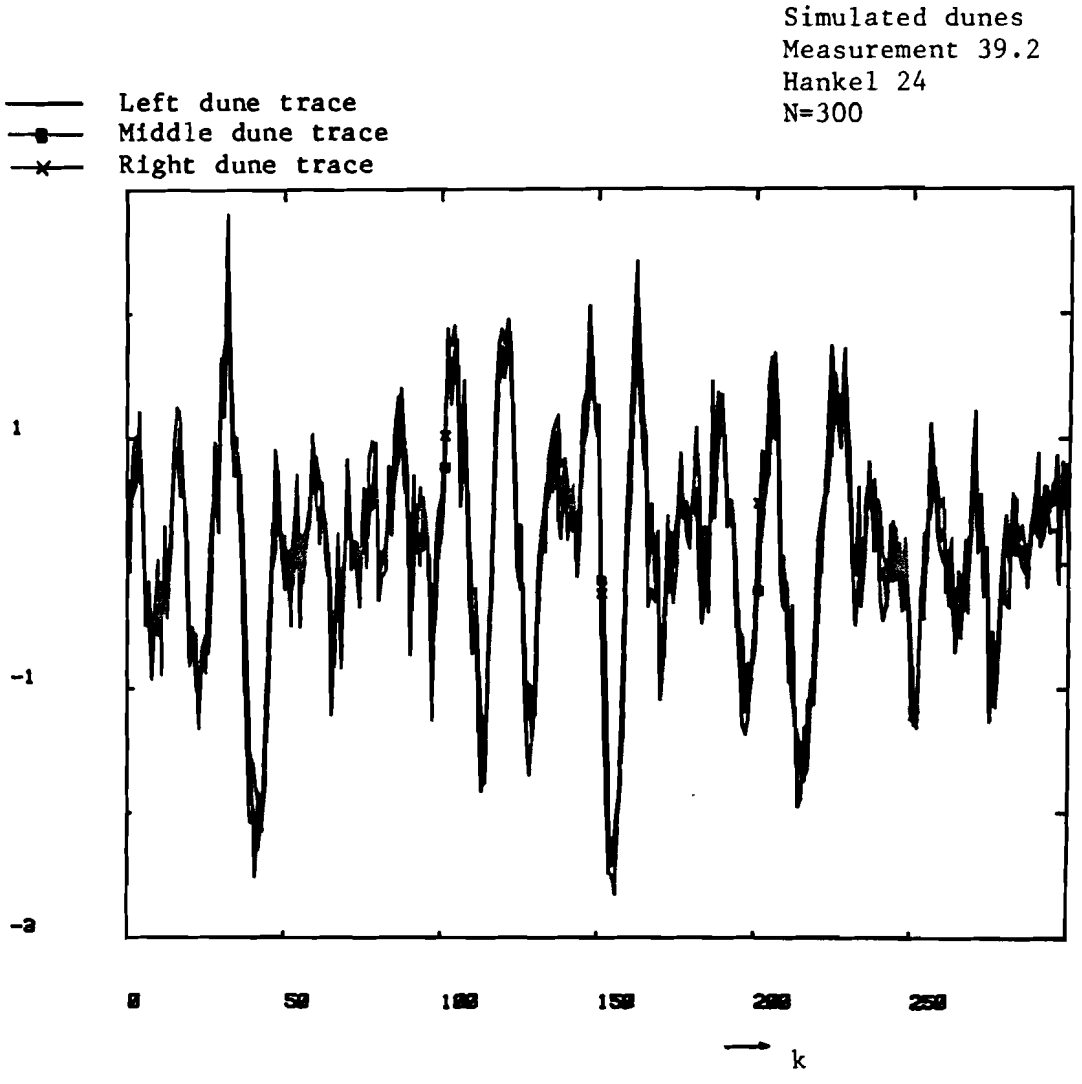


Figure 4.15a Simulated dunes (Hankel appraoch)

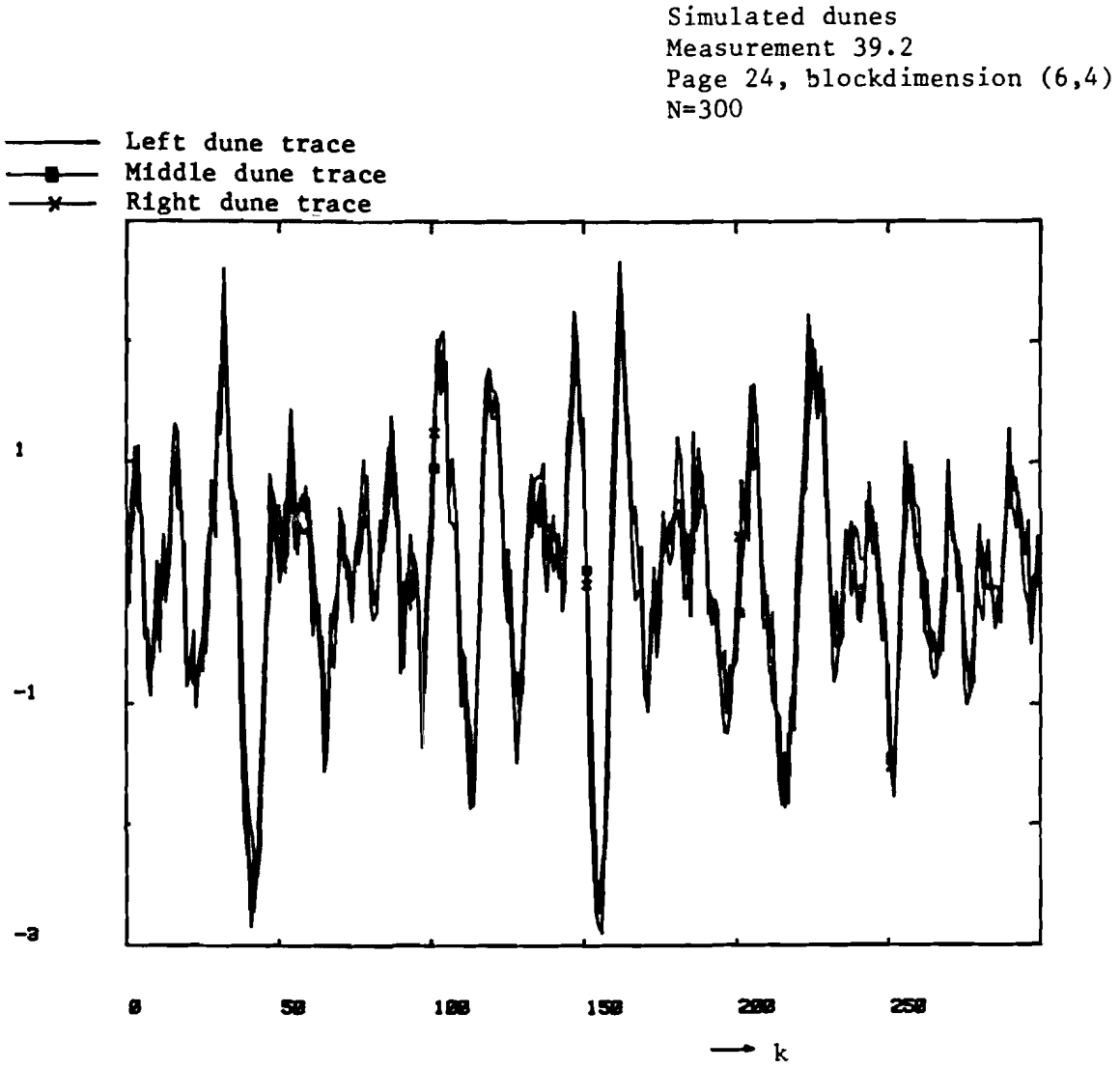


Figure 4.15b Simulated dunes (Page approach)

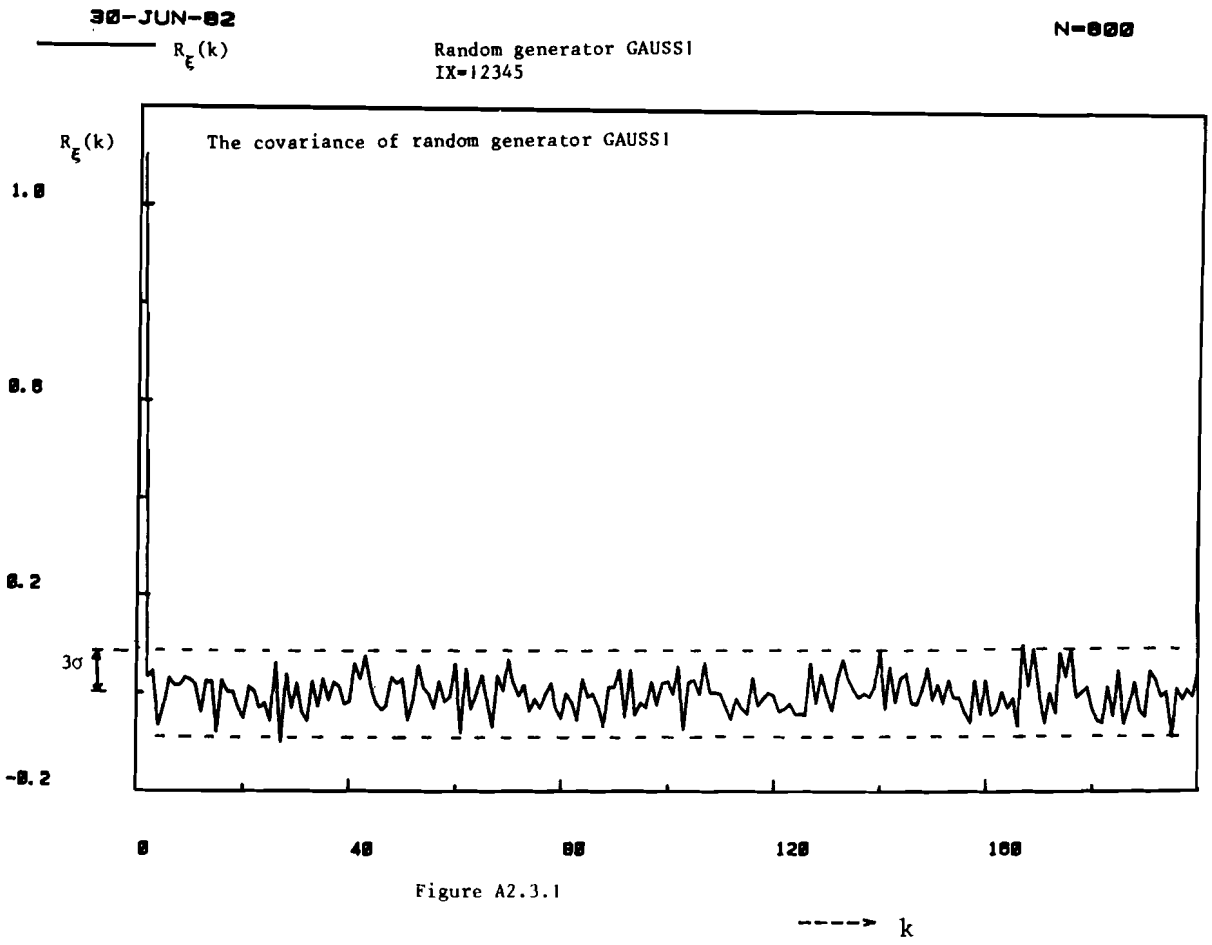

30-JUN-82
N=800
$R_\xi(k)$
Random generator GAUSS1
IX=12345
$R_\xi(k)$
The covariance of random generator GAUSS1
1.0
0.6
0.2
-0.2
3σ
0
40
80
120
160
k

Figure A2.3.1

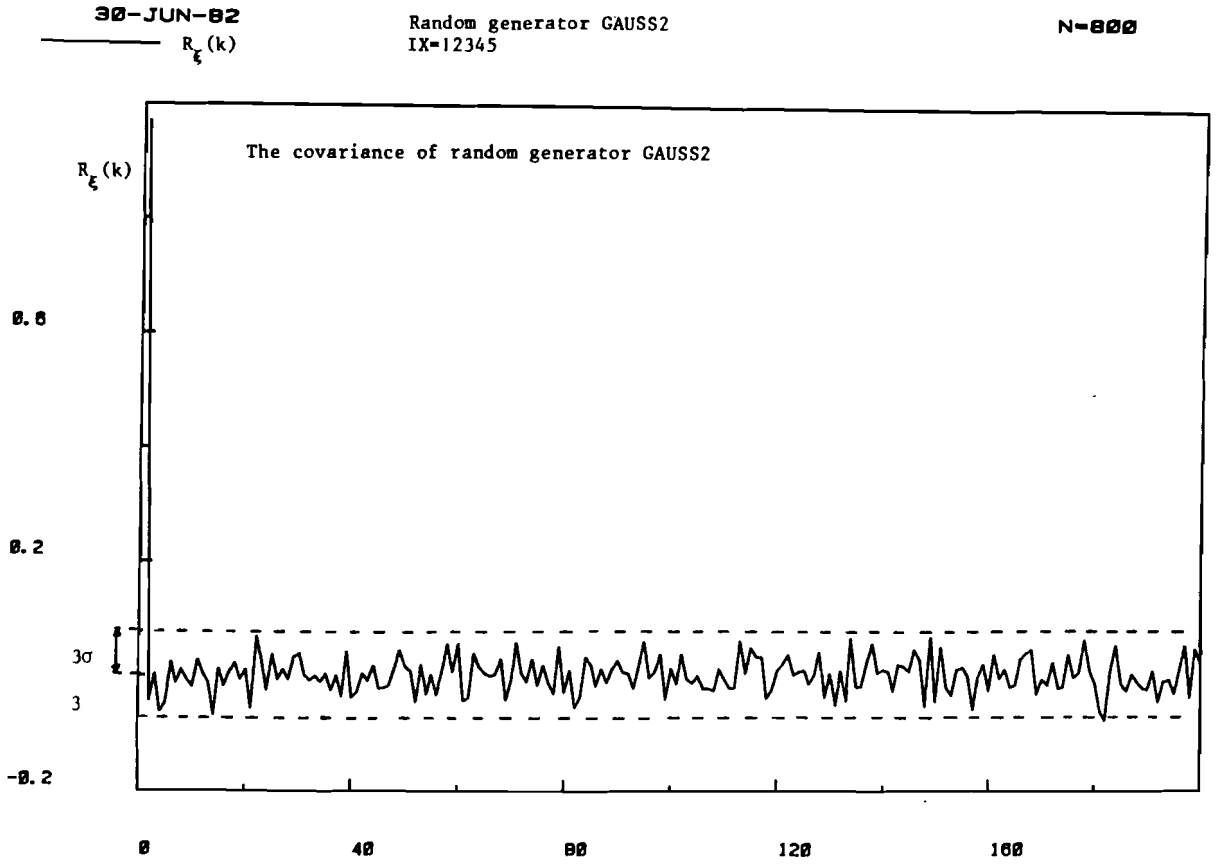

30-JUN-82
$R_\xi(k)$
Random generator GAUSS2
IX=12345
N=800
The covariance of random generator GAUSS2
$R_\xi(k)$
0.8
0.2
3σ
3
-0.2
0
40
80
120
160
k

Figure A2.3.2

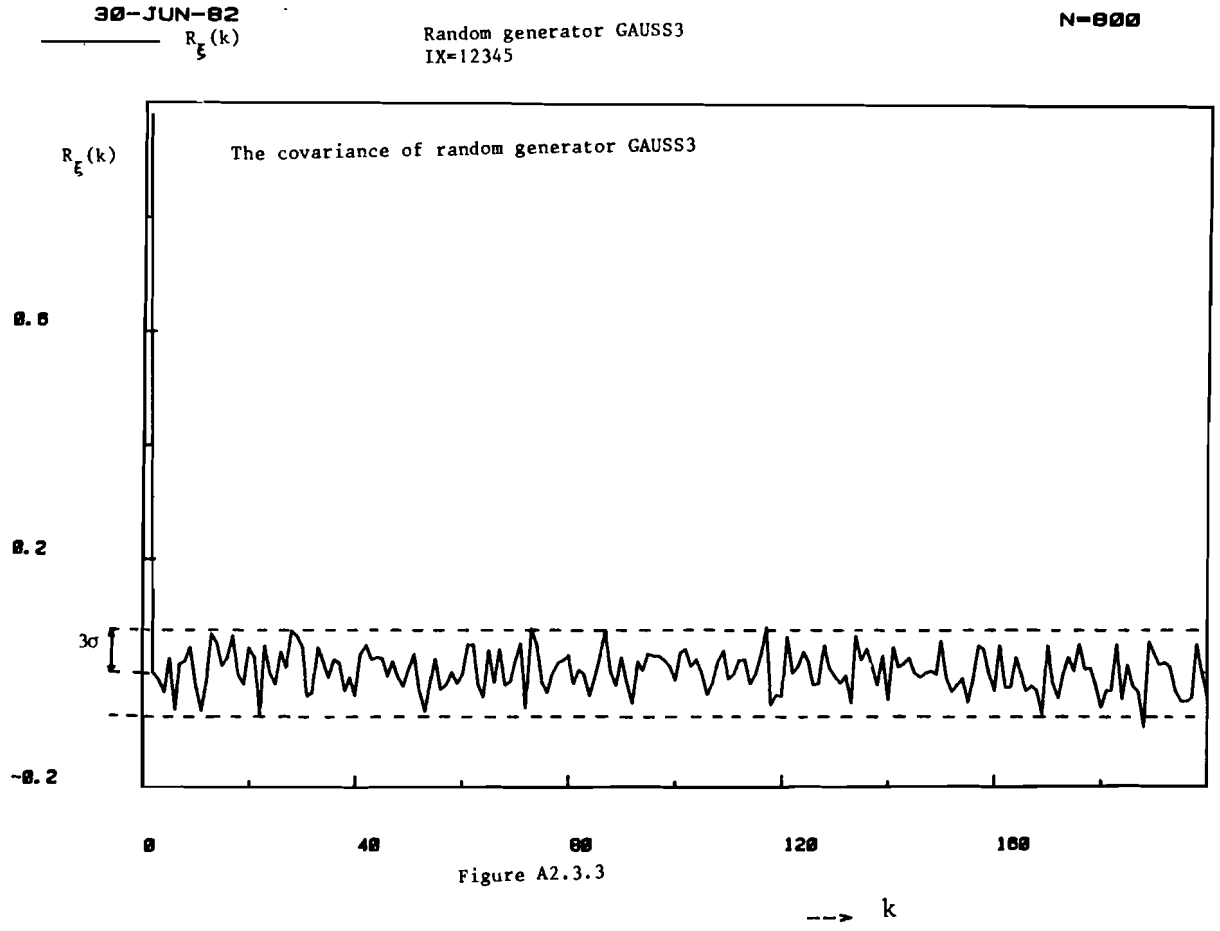

30-JUN-82
$R_\xi(k)$
Random generator GAUSS3
IX=12345
N=800
$R_\xi(k)$
The covariance of random generator GAUSS3
0.6
0.2
3σ
-0.2
0
40
80
120
160
--> k

Figure A2.3.3

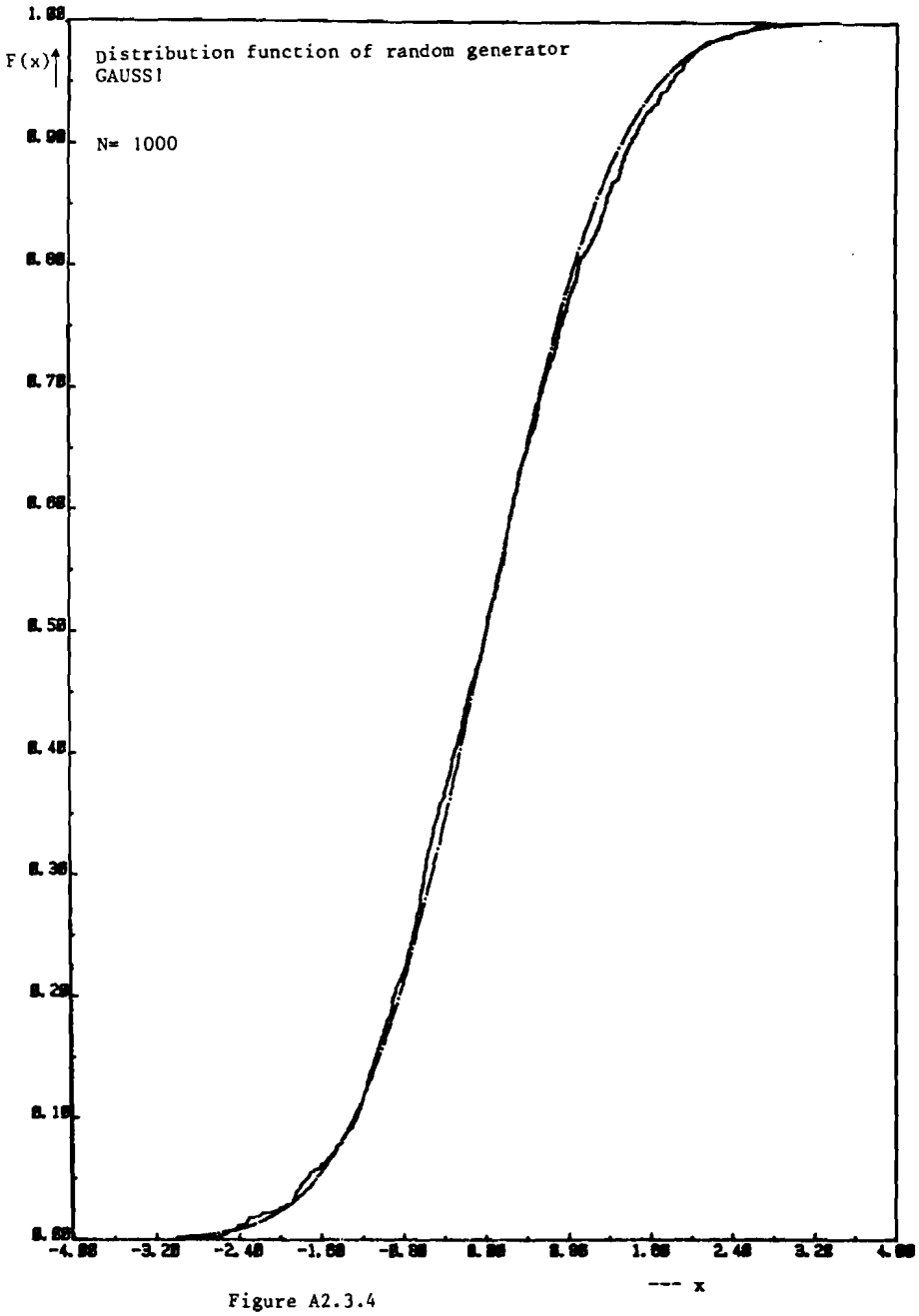

Distribution function of random generator
GAUSS1
N= 1000
F(x)
--- x

Figure A2.3.4

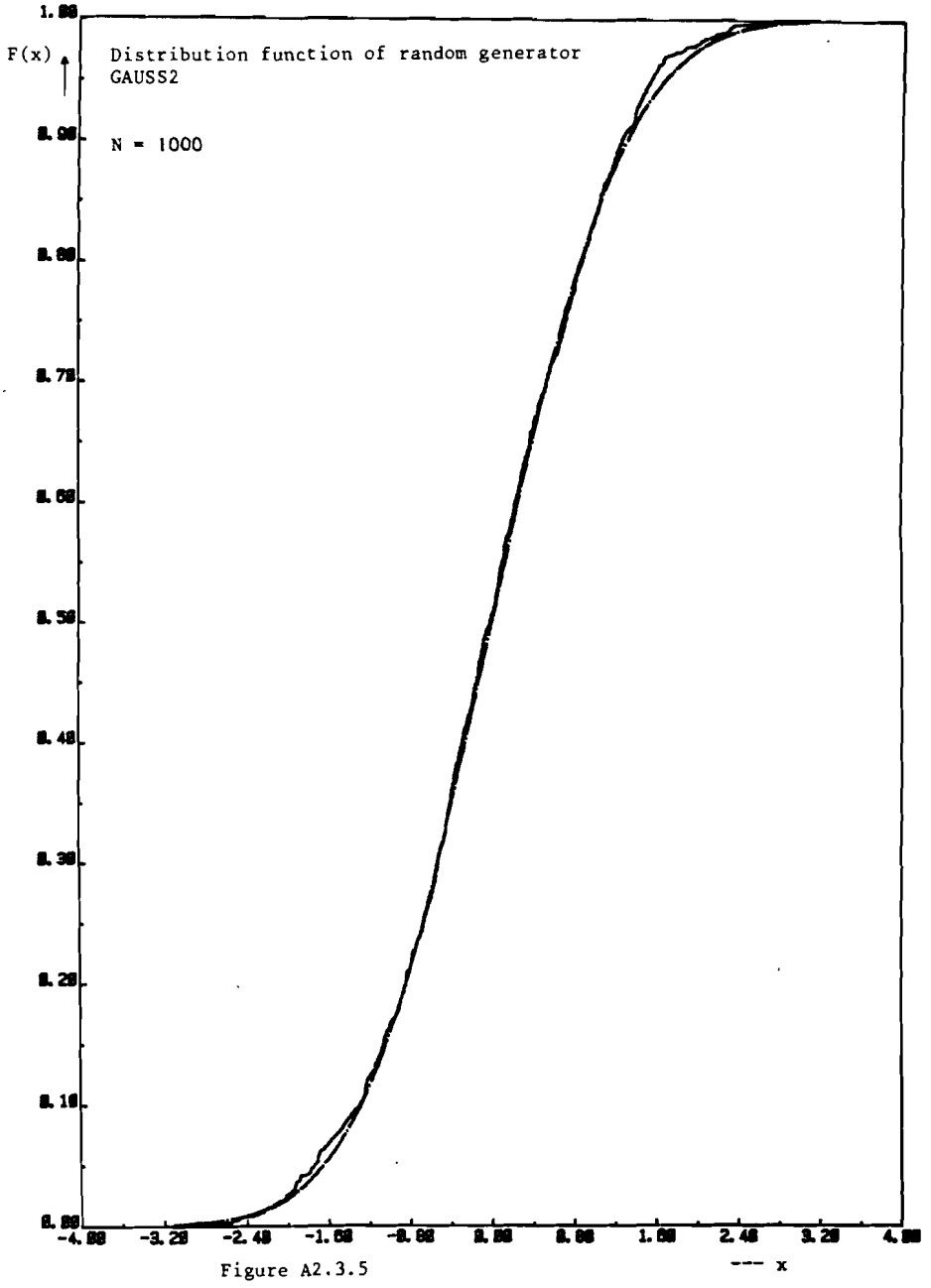

F(x)
Distribution function of random generator
GAUSS2
N = 1000
1.00
0.90
0.80
0.70
0.60
0.50
0.40
0.30
0.20
0.10
0.00
-4.00
-3.20
-2.40
-1.60
-0.80
0.00
0.80
1.60
2.40
3.20
4.00
--- x

Figure A2.3.5

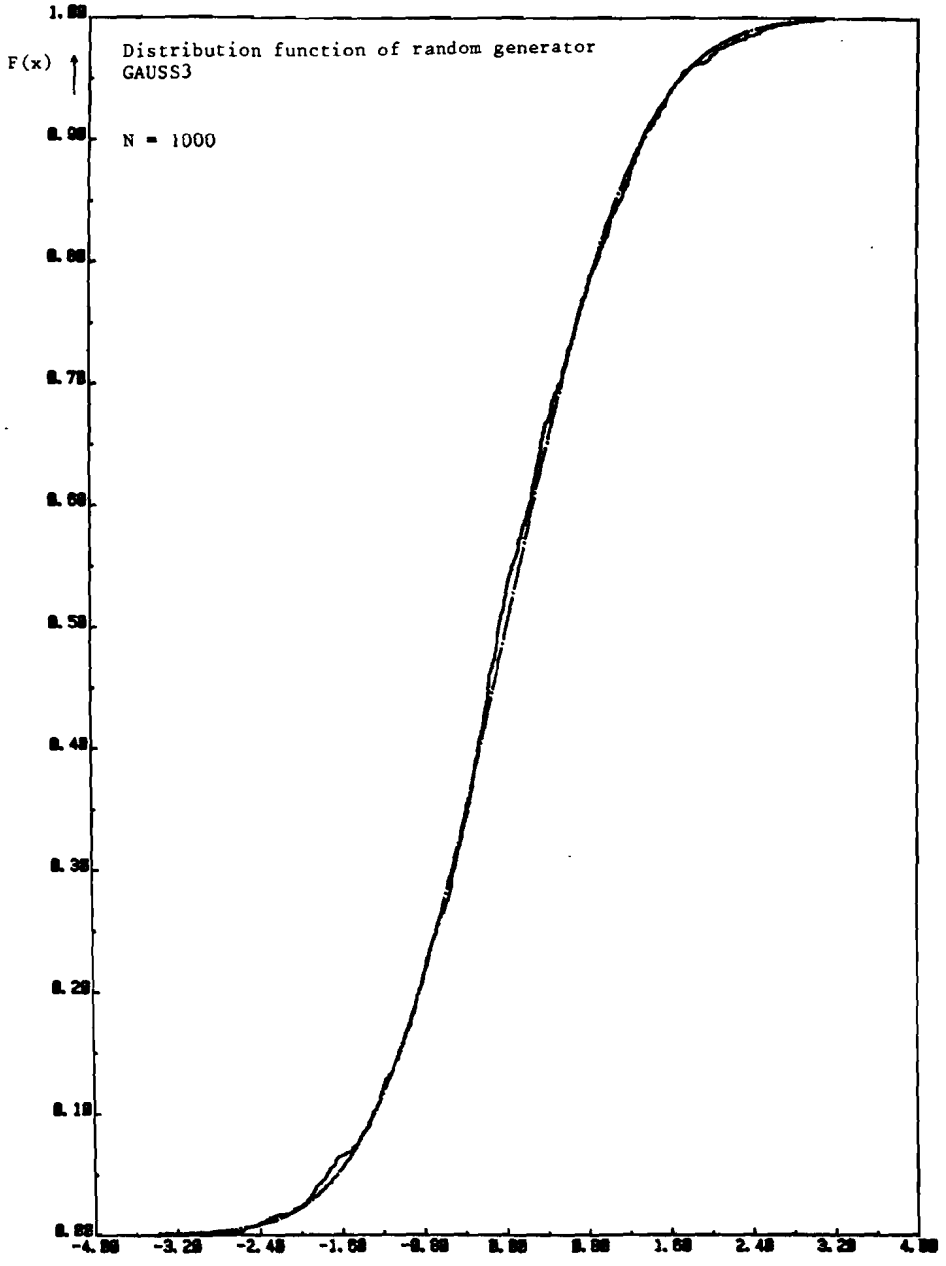

Distribution function of random generator
GAUSS3
N = 1000
F(x)
x

Figure A2.3.6

References

- Anderson,B.D.O. and J.Moore (1979)
  Optimal filtering.
  New York:Prentice Hall.
- Dahlquist,G,A. Bjorck,A. and Anderson (1979)
  Numerical methods.
  Englewood Cliffs:Prentice Hal.
- Damen, A.A.H.,P.M.J. Van den Hof and A.K. Hajdasinski (1982)
  Approximate realization based upon an alternitive to the
  Hankel matrix: the Page matrix.
  Sytems and Control Letters,Vol. II,p202-208
- Denham,M.J. (1976)
  On the factorization of discrete time rational
  spectral density matrices.
  IEEE Trans. Autom. Control,AC-20 ,p 535-537
- Dooren,P.van (1981a)
  A generalized eigenvalue approach for solving
  Riccati equations.
  SIAM. J. Stat.and Sci. Comp.,2 ,p121-135
- Dooren,P.v.(1981b)
  A new numerical method for factorization and
  related problems in system theory.
  In: Circuit and Design,Proceedings of the 1981
  European conference on circuit theory and design,
  The Hague, The Netherlands,
  Ed. R.Boite and P. Dewilde,p.538-544
- Eykhoff, P. (1974)
  Sytem identification: parameter and state estimation.
  London:Wiley and Sons.
- Faurre,P.L.(1976)
  Stochastic relization algorithms
  In: System identification: advances and case
  studies.
  Ed. by R.K. Mehra and D.G. Lainiotes,New York,
  Academic Press, p.1-25

- Gantmacher,F.R.(1959)
The theory of matrices, Vol. I and II.
New York: Chelsea.
- Gent,B. van (1981)
Orde-testen voor een lineair ARMAX-model voor duinen in rivierbeddingen
Department of electrical engineering,
Eindhoven University of Technology (in Dutch).
- Gent,B. van (1983)
User's manual for the stochastic identification programs.
Department of electrical engineering,
Eindhoven University of Technology.
- Gevers,M and W.R.E.Wouters (1978)
An innovation approach to the discrete-time stochastic realization problem,
Journal A,19 ,p 90-110
- Hajdasinski, A.L. (1980)
Linear multivariabel systems. Preliminary problems in mathematical description,modelling and identification.
Department of electrical engineering,
Eindhoven University of Technology;
TH-Report 80-E-106
- Hajdasinski,A.K. and A.A.H. Damen (1979)
Realization of the Markov parameter sequence using the singular value decomposition of the Hankel matrix.
Department of electrical engineering,
Eindhoven University of Technology.
- Hof,P. van den (1981)
Inventory control: a cognitive human operator model.
Department of electrical engineering,
Eindhoven University of Technology,
Group Mesurement and Control,stageverslag
- Hof,P. van den (1982)
Approximate realization of noisy linear systems: the Hankel and Page matrix approach
Department of electrical engineering,
Eindhoven University of Thechnology,
Group Measurement and Control,M.Sc. Thesis.

- Hitz,K.L. and B.D.O.Anderson (1972)
An iterative method of computing the limiting solution of the matrix Riccati differential equation.
Proc. IEEE,119 ,p.1402-1406 .
- Lanning,J.H. and R.H. Battin (1956)
Random processes in automatic control.
New York: Mc Graw-Hill .
- Laub,A.J.(1979)
A Schur method for solving algebraic Riccati equations.
IEEE Trans. Aut.Control,AC-24 ,p 913-921
- Meertens, J.B (1981)
Het karakteriseren van duinen in rivierbeddingen m.b.v. het ARMA model en polen en nulpunten beelden.
Department of electrical engineering,
Eindhoven University of Technology (in Dutch).
- Moler,C.B.and G.W.Steward (1973)
An algorithm for generalized matrix eigenvalue problems.
SIAM. J. Num.Anal.,10 ,No.2.
- Pappas,T,A.J.Laub and N.R.J.Sandel (1980)
On the numerical solution of the discrete time algebraic Riccati equation.
IEEE Trans. Autom. Control,AC-25 ,p 631-641
- Pfennings J.M.F.G.(1982)
Simulaties van duinen in rivierbeddingen
Department of electrical engineering,
Eindhoven University of Technology (in Dutch).
- Vaesen J.H.P.M.(1982)
Een karakterisatie van duinen in rivier beddingen m.b.v. simulaties.
Department of electrical engineering,
Eindhoven University of Technology (in Dutch).

- Vaughn,D.R. (1970)
A nonrecursive algebraic solution for the discrete Riccati equation.
IEEE Trans. Autom. Control,AC-15,p 597-599.
- Veltkamp,G.W. (1982)
Matrixtheorie.
Department of electrical engineering,
Eindhoven University of Technology (in Dutch)
Dictaatnr. 2.287.
- Willems,J.C. (1971)
Least squares stationary optimal control and algebraic Riccati equation,
IEEE Trans. Autom. Control,AC-16 .
- Wouters,W.R.E. and G.Gevers (1978)
An innovations approach to the discrete-time linear least-squares estimation problem,
Journal A.,19 ,p 8-20
- Zee,G.A. van (1981)
System identification for multivariable control.
Dr. Diss., Delft University of Technology,
Delft University Press