MASTER

Approximate realization of noisy linear systems : the Hankel and Page matrix approach

van den Hof, P.M.J.

Award date:
1983

4470

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

APPROXIMATE REALIZATION OF NOISY LINEAR SYSTEMS: THE HANKEL AND PAGE MATRIX APPROACH

by Paul Van den Hof

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 Jan. until Dec. 1982 in charge of Prof. dr. ir. P. Eykhoff under supervision of dr. ir. A.A.H. Damen and dr. ir. A.K. Hajdasinski

"Stuckness shouldn't be avoided.
It's the predecessor of all real
understanding."

Robert M. Pirsig in "Zen and the
Art of Motorcycle Maintenance".

SUMMARY

The Ho-Kalman algorithm creates a minimum realization of a linear, time invariant system, when given a sufficiently long series of deterministic Markov parameters. However if such a "truncated" series of Markov parameters has been disturbed with noise, an approximating Hankel matrix has to be constructed for applying the realization algorithm. This approximating Hankel matrix has either the improper rank, or it lacks the Hankel structure. Furthermore the Markov parameters are not processed with a constant weighting factor, which implies that the noise filtering is inadequate.
In this report an alternative matrix is introduced and investigated: the Page matrix. This matrix is much smaller than the Hankel matrix, which offers the advantage of a considerable reduction in computation. It is shown that the method using this Page matrix might be better suited for handling noisy Markov parameters. The Page matrix approach however still does not provide an optimal solution to the approximate realization problem.
The two approaches are compared theoretically and their practical performance is tested in a set of simulations.

SAMENVATTING

Het Ho-Kalman algoritme geeft een minimale realisatie van een lineair tijdinvariant systeem, wanneer een reeks Markovparameters is gegeven. Wanneer echter een dergelijke eindige reeks Markovparameters met ruis gestoord is moet er een benaderende Hankel matrix gekonstrueerd worden om het realisatie algoritme toe te passen. Deze benaderende Hankel matrix heeft ofwel de verkeerde rang, ofwel geen Hankel struktuur. Bovendien worden de Markovparameters niet gewogen met eenzelfde weegfaktor, wat betekent dat de ruisfiltering niet adekwaat is.

In dit verslag wordt een nieuwe matrix geïntroduceerd en onderzocht: de Pagina matrix. Deze matrix is veel kleiner dan de Hankel matrix, wat een aanzienlijke vermindering betekent van de benodigde rekentijd.

Er wordt beschreven dat deze methode beter geschikt zou kunnen zijn voor het behandelen van ruisgestoorde Markovparameters. De Pagina matrix benadering geeft echter ook geen optimale oplossing van het benaderde realisatieprobleem.

De twee benaderingen worden theoretisch vergeleken en hun eigenschappen worden getest in een serie simulaties.

CONTENTS

INTRODUCTION

System theory is the field of science that is concerned with the study of 'systems' in the widest sense.
It is not simple to formulate a general definition of a system. Any part of physical reality that we consider to be belonging together or to comprise cause and effect symptoms, we may consider as a system. In engineering studies mainly technical systems are considered and the purpose of system theory is to find mathematical ways of describing the behaviour of a system.

The description of the behaviour of a system happens by way of a model. It is an abstract set of relations between variables, occuring in the system, and it has to describe the system in all its characteristic properties; see also Eykhoff (1974). These relations can be differential equations, algebraic equations, difference equations etc. according to the choice of the kind of model, which in turn is, among other things, dependent on the system.
A schematical way of representing a model of a system with multiple inputs and multiple outputs (MIMO-system) is given in figure -0.1-.

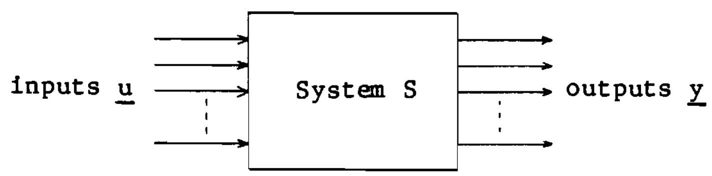


fig-0.1- General MIMO model

In general there will not be enough a priori information about a (technical) system to create directly a mathematical model that describes its behaviour in an exact way. Therefore information has to be drawn from the observation of the system, in more technical terms: on measurements of variables that occur in the system.
These variables can be divided into two groups: a set of input variables $u_1, u_2, \ldots u_p$ that are supposed to represent the excitation of the system, and a set of output variables $y_1, y_2, \ldots, y_q$, representing the response of the system to the excitation.

The problem of system identification now is to find a mathematical description of the relations between the inputs and outputs, resulting in a general description of the behaviour of the system. Information that is used in this identification problem can be a priori information on the system, and measurements of input-variables and output-variables.
The results of system identification are used in many different fields: from diagnostic purposes in biomedical and econometric situations to control purposes in engineering situations.

For MIMO-systems several kinds of models are available from the system theory. For this moment we mention transfer function matrix models, state space models and the Hankel model. A comparison is made in Hajdasinski (1980).
In this report we deal only with discrete-time systems that are linear and time-invariant, and our attention is focussed on the state space models.
In a state space model the behaviour of such a system is described by means of state variables $\underline{x}$ in the following way:

$$\begin{aligned} \underline{x}(k+1) &= A\,\underline{x}(k) + B\,\underline{u}(k) \\ \underline{y}(k) &= C\,\underline{x}(k) + D\,\underline{u}(k) \end{aligned} \tag{0.1}$$

where $\underline{x}(k)$: the $(n\times 1)$-state vector $\underline{x}$ at time instant k,
$\underline{u}(k)$: the $(p\times 1)$-input vector $\underline{u}$ at time instant k,
$\underline{y}(k)$: the $(q\times 1)$-output vector $\underline{y}$ at time instant k,
A : $(n\times n)$-System matrix, where n is called the dimension of the state space or dimension of the system,
B : $(n\times p)$-Distribution matrix,
C : $(q\times n)$-Output matrix,
D : $(q\times p)$-Input-output matrix.
and p,q : resp. the number of inputs and outputs of the system.

The set of matrices $\{A,B,C,D\}$ is called a realization of the system described by eq.(0.1), and it defines completely the external dynamical behaviour of the system.

The main subject of this report is the realization of noisy systems. This concerns the problem of finding a realization $\{A,B,C,D\}$, with a certain dimension n, based upon measurements on the system that are contaminated with noise.
In this report these noisy measurements are considered to be a series of multivariable impulse response matrices, the so called Markov parameters M_k. Based upon this series of Markov parameters $\{M_k\}, k=0,\ldots,L+1$ the purpose is to find a realization $\{A,B,C,D\}$ in such a way that the model described by eq.(0.1) generates a series of Markov parameters that is, in some sense, as close as possible to the exact impulse response of the system.

The application of this technique can be found in the analysis of several subjects in system theory as well as in other fields:

- The approximation problem: looking for a realization $\{A,B,C,D\}$ that describes the system in some optimal and unique way.
- The model reduction problem: in control situations it is very often necessary to approximate a model of high order with a lower order one. Under this restriction a new realization has to be found.
- The design of digital filters: very often filters of reduced order have to be designed to replace the theoretically required infinite order ones.

The study, as reported here, is directed to the approximation problem.

Some assumptions that are taken as a starting point are stated next:

- We deal with linear, time-invariant, discrete-time, stable systems; since the Markov parameter at time instant 0 equals matrix D, this matrix is available direct from the sequence of Markov parameters. For this reason no attention will be paid to this matrix in this report.
- A series of noise corrupted Markov parameters of the system is available. The Markov parameters are generated by a real system and this generated series is disturbed with additive noise, as sketched in fig-0.2-.

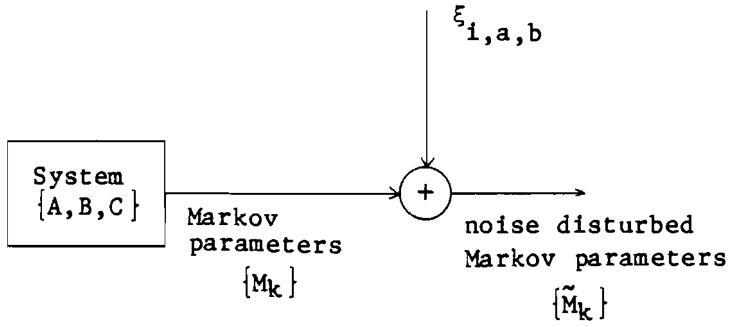


fig-0.2- Creation of noise disturbed Markov parameters

- The noise sample $\xi_{i,a,b}$ is considered to be chosen from SWAYING noise.
  SWAYING noise is defined as follows:

$$\tilde{M}_i\,(a,b) = M_i(a,b) + \xi_{iab} \tag{0.2}$$

where $M_i(a,b)$ is element a,b in matrix M_i and ξ_{iab} is the corresponding additive noise. This noise is assumed to be stationary (S), white (W) (zero mean), additive (A), signal-independent (Y), inter-independent (among channels) (I), with non-changing global variance σ^2 (NG):

$$E\{\xi_{iab}\} = 0 \qquad \forall\, i,a,b \in N \tag{0.3}$$

$$E\{\xi_{iab}\xi_{jcd}\} = \begin{cases} 0 & i,a,b \neq j,c,d \\ \sigma^2 & i,a,b = j,c,d \end{cases} \tag{0.4}$$

The assumptions about the noise on the Markov parameters are quite heavy. In fact it is the least complicated way of considering a noise disturbance. This assumption should be considered as a starting point for the analysis of the problem as described above. More sophisticated ways of noise influence (e.g. coloured noise), would complicate too heavily the analysis at this stage of the project. In later studies this aspect can be dealt with.
On the other hand assuming SWAYING noise on the Markov parameters is not an exceptional assumption when we take into account that these parameters can be direct results of measurements; the noise

is then supposed to represent measurement errors.

In chapter 1 an approach to the realization problem of noisy systems is given that is mainly used in the literature: the adapted Ho-Kalman algorithm. In chapter 2 an alternative approach is introduced. Chapter 3 gives a formal proof of a theorem that is given in the previous chapter. This is quite a detailed proof and the unsuspecting reader should be aware of this. This chapter can be skipped without loosing crucial understanding of the report. Chapter 4 gives a comparison of the two methods as presented in the first two chapters, and in chapter 5 results of simulations are given. The report ends with some conclusions.
A description of computer programs that are used for the simulation is given in a separate report.

CHAPTER 1: HANKEL MATRIX APPROACH TO THE REALIZATION PROBLEM

1.1 PRELIMINARIES

1.1.1 Introduction

The problem of realization of the Markov parameter sequence of a system, given some noisy measurements or estimations of these parameters, has very often been treated by using a Hankel matrix. A Hankel matrix consisting of Markov parameters is a matrix that is constructed in the following way:

$$H[j] = \begin{bmatrix} M_1 & M_2 & M_3 & \cdot & \cdot & M_j \\ M_2 & M_3 & M_4 & \cdot & \cdot & M_{j+1} \\ \cdot & \cdot & & \cdot & & \cdot \\ \cdot & \cdot & & & \cdot & \cdot \\ M_j & M_{j+1} & M_{j+2} & \cdot & \cdot & M_L \end{bmatrix} \qquad (1.1)$$

$L=2j-1$

Ho and Kalman were the first who introduced an algorithm that constructs a realization $\{A,B,C\}$ of a linear, time-invariant, state-space model, given a noise-free Hankel matrix of the system of sufficient dimensions.
In addition to this approach, many publications have been written about the realization of Markov parameter sequences with respect to the approximation problem. Both in the deterministic case and in the noisy case all presented methods use Ho-Kalman-like algorithms and the Hankel matrix. See Silverman (1971); Zeiger and McEwen (1974); Kung (1978); v.Zee and Bosgra (1979); Hajdasinski and Damen (1979); Damen and Hajdasinski (1982); Staar, Vandewalle and Wemans (1981a,1981b).
In the deterministic case the Ho-Kalman algorithm leads to an exact solution of the approximation problem: an exact and minimal realization $\{A,B,C\}$ of the system.

From this realization the sequence of Markov parameters can be found, by way of the relation

$$M_k = C\,A^{k-1}\,B \qquad k \geq 1 \tag{1.2}$$

The Hankel matrix now can be reconstructed in an exact way.

Contrary to the model reduction problem, where a model of the system is known and a reduced order model has to be found, in the approximation problem the noisy case is of crucial importance.
The Ho-Kalman algorithm, and related algorithms as mentioned above, in fact are only functioning correctly in the deterministic case.
In the noisy case, however, the correctness of these algorithms become less self-evident.
In the first place the realization will never be exact, but always an approximation of the system; this will become clear in future sections. The way how to approximate the system, and how to deal with the noise now become crucial questions.
In the second place the term "minimal realization" becomes a subjective one, that is not fixed anymore.
Also this feature will be dealt with in the next sections.

1.1.2. Definitions and theorems

Before the properties of the Ho-Kalman algorithm are discussed in more detail, some definitions have to be given and some theorems have to be mentioned.
They can be found e.g. in Hajdasinski and Damen (1979), and Ho and Kalman (1966).

Definition 1: Any polynomial $f(z)$

$$f(z) = z^k - a_1 z^{k-1} - a_2 z^{k-2} - \dots - a_{k-1} z - a_k z^0 \tag{1.3}$$

for which holds $f(A) = [0]$ is called an annihilating polynomial of the matrix A.

Definition 2: The polynomial $f(z)$ of the smallest degree r, nonequal to zero, fulfilling definition 1 is called the minimal polynomial.
r is called the realizability index.

Theorem 1: The sequence of Markov parameters $\{M_k\}$, $k=1,2,\ldots$ has a finite dimensional realization $\{A,B,C\}$ if and only if there are an integer r and constants a_i such that

$$M_{r+j} = \sum_{i=1}^{r} a_i M_{r+j-i} \qquad \text{for all } j>0 \tag{1.4}$$

the minimum value of r is called the realizability index.

Theorem 2: If $\{M_k\}$ has a finite dimensional realization then

$$n = \text{rank } H[r] \tag{1.5}$$

where n is the minimal dimension of the state-space, and $H[r]$ is the Hankel matrix as described in the previous section.

Because of the linear dependence between the Markov parameters it follows that

$$n = \text{rank } H[r] = \text{rank } H[r+k] \quad \text{for all } k>0 \tag{1.6}$$

The most essential aspect of the Hankel matrix approach is the following equivalence relation:

$$\begin{tikzcd} \boxed{\begin{array}{l}\{M_k\}, k=1,..,L \text{ has a finite dimensional} \\ \text{realization with minimum dimension } n\end{array}} \arrow[d, shift right=2ex, "\text{necessary}"'] \\ \boxed{\text{rank } H[r] = n} \arrow[u, shift right=2ex, "\text{sufficient}"'] \end{tikzcd} \tag{1.7}$$

where $H[r]$ is built up from $2r-1$ Markov parameters ($L>2r-1$).

The necessary condition for this relation is given by theorem 2; the sufficient condition is shown by Ho and Kalman in applying their algorithm.

1.2 HO-KALMAN AND RELATED ALGORITHMS

1.2.1. The Ho-Kalman algorithm

When given a series of Markov parameters $\{M_k\}$, $k=1,..L+1$, having a finite dimensional realization, with the aid of eq.(1.2) the Hankel matrix $H[r]$ can be decomposed in the following way:

$$H[r] = \begin{bmatrix} C \\ CA \\ CA^2 \\ \cdot \\ \cdot \\ CA^{r-1} \end{bmatrix} \cdot \begin{bmatrix} B & AB & A^2B & \cdot & \cdot & A^{r-1}B \end{bmatrix} \qquad (1.8)$$

$$= \Gamma[r] \cdot \Delta[r]$$

where $\Gamma[r]$ is the r-observability matrix and
$\Delta[r]$ the r-controllability matrix.

We assume that $\{A,B,C\}$ forms a minimum realization of the given sequence of dimension n.
Because of the rank condition rank $H[r] = n$ both matrices $\Gamma[r]$ and $\Delta[r]$ have to have at least rank n. However because of their dimensions the rank of both matrices can not exceed n, and therefore there has to hold:

$$\text{rank}\ \Gamma[r] = \text{rank}\ \Delta[r] = n \qquad (1.9)$$

Both matrices therefore will have full rank.

Now, any such decomposition of $H[r]$ into two matrices with full rank n will show the same structure as in eq.(1.8) with some minimum realization $\{A^*, B^*, C^*\}$, because:

$$H = \Gamma^* \Delta^* = \Gamma \Delta \quad ^{1)} \tag{1.10}$$

$$\rightarrow \quad \Delta^* = (\Gamma^*)^{+} \Gamma \Delta = T \Delta \tag{1.11}$$

where + stands for pseudo inverse and T is non-singular as Γ and Γ^* are of full rank n. Analogously we may put:

$$\Gamma^* = \Gamma \Delta (\Delta^*)^{+} = \Gamma S \tag{1.12}$$

Finally, substitution of (1.11) and (1.12) into (1.10) leads to:

$$H = \Gamma \Delta = \Gamma S T \Delta \rightarrow \tag{1.13}$$

$$\Gamma^{+} \Gamma S T \Delta \Delta^{+} = I \rightarrow S T = I \rightarrow S = T^{-1} \tag{1.14}$$

So the equivalence transformation can be defined as:

$$\begin{array}{ll} \Delta^* = T \Delta & \Gamma^* = \Gamma T^{-1} \\ B^* = T B & C^* = C T^{-1} \\ A^* = T A T^{-1} & \end{array} \tag{1.15}$$

The complete set of all possible (Γ, Δ) together then produces the complete equivalence class of the system under study.
The triplet (A,B,C) can be obtained from (Γ, Δ) as follows: the matrices B and C can be recognized as the first blocks in Δ and Γ respectively. In order to obtain matrix A, we need a shifted matrix, which we indicate by an arrow. A vertically pointing arrow indicates a shift of one block row, whereas a horizontally pointing arrow denotes a shift of one block column. From this it is clear that

$$H\uparrow = \overleftarrow{H} = \Gamma A \Delta \tag{1.16}$$

---

[1]) the indication of the dimension of the matrices will be omitted if it is clear what is meant.

where

$$H[j]\uparrow = \overset{\leftarrow}{H}[j] = \begin{bmatrix} M_2 & M_3 & M_4 & \cdot & \cdot & M_{j+1} \\ M_3 & M_4 & M_5 & \cdot & \cdot & M_{j+2} \\ \cdot & \cdot & & \cdot & & \cdot \\ \cdot & \cdot & & & \cdot & \cdot \\ M_{j+1} & M_{j+2} & M_{j+3} & \cdot & \cdot & M_{L+1} \end{bmatrix} \qquad (1.17)$$

As Γ and Δ have a maximum rank n, we may write:

$$A = \Gamma^{+} \, H\uparrow \, \Delta^{+} \qquad (1.18)$$

Note that we needed the extra Markov parameter M_{L+1} to construct the shifted Hankel matrix.

Although Ho and Kalman choose their Hankel matrix in a block squared way, like the original definition of a Hankel matrix, this is not strictly required for applying the algorithm. With respect to this a sufficient condition for the validity of the algorithm is that a Hankel-like matrix $H[i,j]$ is constructed for which there exist matrices $\Gamma[i]$ and $\Delta[j]$ in such a way that

$$\text{rank } \Gamma[i] = \text{rank } \Delta[j] = n \qquad (1.19)$$

Minimal values of i and j for which eq.(1.19) holds are called the observability index (α) resp. controllability index (β).
We define:

$$s = \max(\alpha, \beta) \qquad (1.20)$$

The problem of finding a realization $\{A,B,C\}$ now is reduced to the problem of finding a decomposition of H into two full rank matrices with rank n. This decomposition can be done by singular value decomposition (S.V.D.), as will be explained in section 1.4.

1.2.2. Silverman's algorithm

For the complete derivation and description of Silverman's algorithm we refer to Silverman (1971).
For this moment the essential differences with the Ho-Kalman algorithm will be pointed out.

- Silverman's algorithm does not require a decomposition of the Hankel matrix in full rank matrices.

- The essential matrix around which the algorithm is designed is a submatrix of the Hankel matrix, consisting of the first n independent rows in H. This means that only a small part of the available sequence of Markov parameters is used: such a part that the dimensions of the mentioned submatrix are just large enough to construct a realization.

- For determining matrix A, a comparable shift operation as in the Ho-Kalman algorithm is used.The matrices B and C are evalated based upon linear operations on submatrices of H.

In case of deterministic Markov parameters Silverman's algorithm, like the Ho-Kalman algorithm,leads to a minimum realization. In case of noise corrupted Markov parameters however, its properties are quite different, as will be pointed out in section 1.5 and chapter 4.

1.2.3. Kung's method

The third algorithm presented is based on an idea introduced by Kung (1978). It is a modification of the Ho-Kalman algorithm and it differs from this algorithm only in the determination of matrix A.
Matrices B and C are evaluated in the same way, after decomposition of the Hankel matrix.
The evaluation of matrix A however is not based on the shifting property of the Hankel matrix (eq.(1.16)), but on properties of shifted observability and controllability matrices.

Considering the structure of the i-observability and j-controllability matrix Γ_i and Δ_j, we can write

$$\Gamma\!\uparrow\, = \Gamma A \qquad \text{or} \qquad A\Delta = \overset{\leftarrow}{\Delta} \tag{1.21}$$

Consequently this yields:

$$A = \Gamma^{+}\Gamma^{\uparrow} \qquad \text{or} \qquad A = \overset{\leftarrow}{\Delta}\,\Delta^{+} \tag{1.22}$$

In order to construct the shifted matrices $\Gamma\!\uparrow$ or $\overset{\leftarrow}{\Delta}$ having the same dimensions as Γ or Δ, we lack information on what to insert in the latter blocks of Γ or Δ during this operation. Therefore we have to apply a Γ and a Δ with reduced dimensions in (1.21) and (1.22). Reduction is accomplished by omitting the last block.

When matrices $\Gamma[i-1]$ and $\Delta[j-1]$ of reduced dimension are used to construct shifted matrices

$$\Gamma[i-1]\!\uparrow \text{ or } \overset{\leftarrow}{\Delta}[j-1]$$

one has to take into account that the rank condition now has to hold also for the reduced dimension matrices:

$$\text{rank } \Gamma[i-1] = \text{rank } \Delta[j-1] = n \tag{1.23}$$

Only if this condition is fulfilled a system matrix A of dimension n will be found.
Consequently the conditions on the integers i and j become:

$$\begin{aligned} i &> \alpha+1 \\ j &> \beta+1 \end{aligned} \tag{1.24}$$

and therefore the minimum dimensions of the required Hankel matrix are higher than in case of the Ho-Kalman algorithm.
For determination of A only one shift operation is required: either $\Gamma\!\uparrow$ or $\overset{\leftarrow}{\Delta}$ is sufficient for evaluating A. Therefore the equivalent conditions (1.23) and (1.24) have to hold for either i or j, corresponding with the choice of $\Gamma\!\uparrow$ or $\overset{\leftarrow}{\Delta}$.

1.3 DETERMINISTIC VERSUS NOISY SITUATION

1.3.1 Deterministic case

As was mentioned in the previous section, the Ho-Kalman as well as the Silverman and Kung algorithm lead to an exact and minimum realization $\{A,B,C\}$ of the system under study.
The dimension of this realization is the minimum dimension of the state-space, the minimum number of state-variables that is needed to describe the system in a complete way. According to theorem 2 in section 1.1.2 this minimal dimension n can directly be derived from the Hankel matrix via the relation (1.6).
So, the size of the Hankel matrix is not really important in this deterministic case, as long as it is greater than or equal to the size of $H[r]$ (or $H[r+1]$ for Kung's method in some specific situations). Under this condition the rank of the Hankel matrix will always equal n, and the result of the algorithms will be a realization with minimal dimension n.
It is quite understandable that, from a certain minimum level, it is no use to increase the size of the Hankel matrix by taking more Markov parameters into account; the linear dependence between the Markov parameters remains the same, and no extra information is added.

From all these remarks it has to be clear that it is a general assumption that the given sequence of Markov parameters $\{M_k\}$ $k=1,2,...$ has in fact a finite dimensional realization. In other words: there can be found a triplet of constant matrices $\{A,B,C\}$ in such a way that $CA^{k-1}B = M_k$, $k=1,2,...$

1.3.2. Noisy case

The noisy situation will be defined as the case where there does not exist anymore a finite dimensional realization that reconstructs $\{M_k\}$,$k=1,2,..$ in an exact way. Then with increasing L (the number of Markov parameters) the rank of the Hankel matrix will always increase. This can be caused by the stochastic nature of the system or by the procedure of gathering the Markov parameters, when noise occurs in the measuring or estimation procedure.

At this moment, and in the remaining part of this report, we will assume that the given Markov parameters can be represented as a deterministic part to which a stochastic part has been added (see Introduction).
The problem of reconstructing the sequence of Markov parameters with a triplet $\{A,B,C\}$ in an exact way, now becomes a problem of approximating the sequence with a sequence that is generated by a realization of some finite dimension.
In other words: a finite dimensional (n) realization has to be found that generates a sequence of Markov parameters, which sequence is, in some way, as close as possible to the given sequence (of length $L \gg n$) of noise disturbed Markov parameters.
In the formulation of this problem, the problem of order determination is directly contained. Contrary to the deterministic case, where the dimension of the minimal realization is directly encountered in the Hankel matrix, now the rank of the Hankel matrix has to be reduced to create a realization of restricted dimension.
According to this rank reduction of the Hankel matrix, the same problem as mentioned above can be stated in another way:

> given a Hankel matrix with noisy Markov parameters, find a Hankel matrix with reduced rank in which the Markov parameters are, in some defined way, as close as possible to the Markov parameters of the original Hankel matrix.

The reduction of the rank of the Hankel matrix can be regarded as a noise filter that operates on the Hankel matrix. It filters out the Markov parameters in such a way that $\{M_k\}$ $k=1,2,\ldots$ can be reconstructed with a finite dimensional realization,as both the Hankel structure and the rank being n are necessary and sufficient conditions.

The rank reduction of the original Hankel matrix has to result into an approximating Hankel matrix. With respect to the realization algorithm this approximating Hankel matrix would perform the same role as the deterministic Hankel matrix in the deterministic situation. Unfortunately until now such an approximating Hankel matrix of proper structure and rank has not been found for a Euclidian norm (see also section 1.4). Only suboptimal solutions have been studied pragmatically. These multiple suboptimal solutions are caused by the several possibilities for applying the formulae of section 1.2.1. in order to find a realization. These possibilities are compared in Damen and Hajdasinski (1982). For the shifted Hankel matrix we may either use the original H (v. Zee and Bosgra (1979)), or the approximating Hankel matrix (Hajdasinski and Damen (1979)), or the method of Kung may be used (Kung (1978)).

Now we can state that there are three functions that have to be performed by the algorithms based on the Hankel matrix:

1. order determination
2. noise filtering
3. realization

In the deterministic case there is only one function: realization.

To investigate the properties of the Ho-Kalman and related algorithms in the case of noisy Markov parameters, the performance of the three functions as mentioned has to be analysed.
For investigating these functions the singular value decomposition of the Hankel matrix will be introduced in the next section. This singular value decomposition performs a decomposition of the Hankel matrix into full rank matrices.There also will be given a description of the performance of the three functions.

1.4 SINGULAR VALUE DECOMPOSITION AS A TOOL IN THE HO-KALMAN ALGORITHM

1.4.1 Singular value decomposition

The singular value decomposition of an arbitrary matrix H is given by (see Hajdasinski and Damen(1979)):

$$H = W \Sigma V^T \tag{1.25}$$

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 is a $\rho \times \rho$ diagonal matrix: $\mathrm{diag}(\delta_1, \delta_2, \ldots, \delta_\rho)$

$\delta_1 \geq \delta_2 \geq \ldots \geq \delta_\rho \geq 0$

and δ_i is the square root of an eigenvalue of H^TH (or HH^T), and is called a singular value.

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

$$\mathrm{rank}\ H = \mathrm{rank}\ \Sigma \leq \rho \tag{1.26}$$

This means that the rank of Σ is a direct measure for the rank of H. In other words: the number of singular values unequal zero reflects the rank of H.

Because of the fact that all singular values unequal zero are positive, the singular value decomposition can also be written as:

$$H = W \Sigma^{\frac{1}{2}} \Sigma^{\frac{1}{2}} V^T \tag{1.27}$$

where

$$\Sigma^{\frac{1}{2}} = \begin{bmatrix} \delta_1^{\frac{1}{2}} & 0 & \cdot & \cdot & & \cdot \\ & \cdot & & & \emptyset & \\ & & \delta_n^{\frac{1}{2}} & & & \cdot \\ & \emptyset & & 0 & & \\ & & & & \cdot & \\ & & & & & 0 \end{bmatrix}_{(\rho\times\rho)} \qquad (1.28)$$

where n = rank H.

The decomposition $H = \Gamma \cdot \Delta$ where $\Gamma = W\,\Sigma^{\frac{1}{2}}$
and $\Delta = \Sigma^{\frac{1}{2}}V^T$ (1.29)

now performs a decomposition of H into two full rank matrices of rank n. This is exactly a decomposition that is required in the Ho-Kalman algorithm for finding an observality and controllability matrix. A realization based on the decomposition (1.29) we call a balanced realization because of the balanced distribution of Σ over Γ and Δ. Formally the (unique) balanced realization is only obtained in case of infinite Hankel matrices (see Silverman and Bettayeb (1980), and Moore (1981)).

1.4.2. Noise filtering

The optimal noise filter would be accomplished by finding a Hankel matrix H_k of reduced rank $k<\rho$ that is, in some defined way, as close as possible to the original Hankel matrix H. As mentioned in section 1.3.2 this is not yet found. However, if we skip the condition of Hankel structure, the solution is suboptimal but easy and straightforward.

The singular value decomposition supplies us with a tool for this rank reduction of H:

In Van der Kam and Damen (1978) it is proved that for a given required dimension k, the matrix H_k with rank k that minimizes $\|H_k-H\|$ is given by:

$$H_k = U \cdot \left[\begin{array}{c|c} \Sigma_k & \emptyset \\ \hline \emptyset & \emptyset \end{array}\right] \cdot V^T \qquad (1.30)$$

where $\Sigma_k = \mathrm{diag}(\delta_1, \delta_2, .., \delta_k)$
and $k<\rho$.

This result holds for two choices of the matrix-norm $\|H_k-H\|$.
Both for the Euclidian norm, defined as:

$$\|G\|_E = \{\sum_{i,j} G_{ij}^2\}^{\frac{1}{2}} \qquad (1.31)$$

as for the spectral-norm, defined as:

$$\|G\|_s = \max_{\underline{x}\neq\underline{0}} \frac{\|G\underline{x}\|_E}{\|G\|_E} \qquad (1.32)$$

(G = general matrix)

the result of Van der Kam and Damen is valid.
In this report we will focus our attention on the Euclidian matrix norm.

Another way of writing $\|H\|_E$, that corresponds to the definition (1.31), and that will also be used in the sequel, is:

$$\|H\|_E^2 = tr(H^TH) = \sum_{j=1}^{\rho} \delta_j^2 \qquad (1.33)$$

According to the second way of writing $\|H\|_E$, it holds that

$$\|H_k - H\|_E^2 = \sum_{j=k+1}^{\rho} \delta_j^2 \qquad (1.34)$$

Now, what do these results mean for the practical situation?
Given a noisy Hankel matrix of full rank ρ. This rank can be very large, according to the number of Markov parameters that is taken into account. To find a realization of dimension k the rank of H has to be reduced to k.
This matrix with reduced rank can be taken as H_k: putting the smallest ρ-k singular values of H to zero, thus minimizing $\|H_k-H\|_E$.
It should be noticed that minimizing $\|H_k-H\|_E$ is just one choice for a possible criterion. Several other choices for such a criterion are possible.

Putting the ρ-k smallest singular values of H equal to zero can be regarded as the noise filtering.The deterministic part of the system is supposed to have a k-dimensional realization and therefore also k singular values unequal zero. The singular values $\delta_{k+1},\ldots,\delta_p$ are considered to be caused by the noise, because in absence of noise these values would equal zero. In fig-1.4.1- this is elucidated for an example with k=n=3.

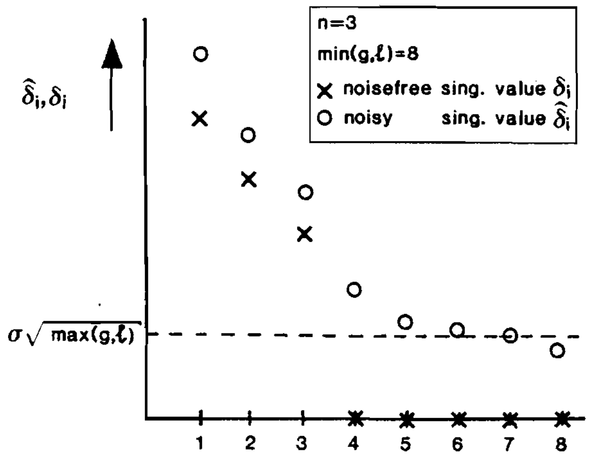


Fig-1.4.1- Singular values of a 3^{rd} order system in the deterministic and in the noisy case.

More attention to the influence of noise on singular values is paid in section 2.4.

The choice of the minimization criterion as mentioned above gives very good opportunities for performing the rank reduction by way of the singular value decomposition.

However one facet of this procedure should be noticed:

Minimizing $\|H_k - H\|^2_E$ means minimizing

$$\sum_{i,j} \{[H_k]_{i,j} - [H]_{i,j}\}^2$$

It is an overall fit of the Hankel matrix where all elements of this matrix are weighted with the same factor. Because of the special structure of the Hankel matrix this means that the first and the last Markov parameter, that appear only once in the Hankel matrix, are weighted with a factor that is much smaller than that of the Markov parameters in the middle of the "time range".

In fact, during the rank reduction procedure a window is used with an isosceles characteristic as given in fig-1.4.2- (in case of a block square Hankel matrix).

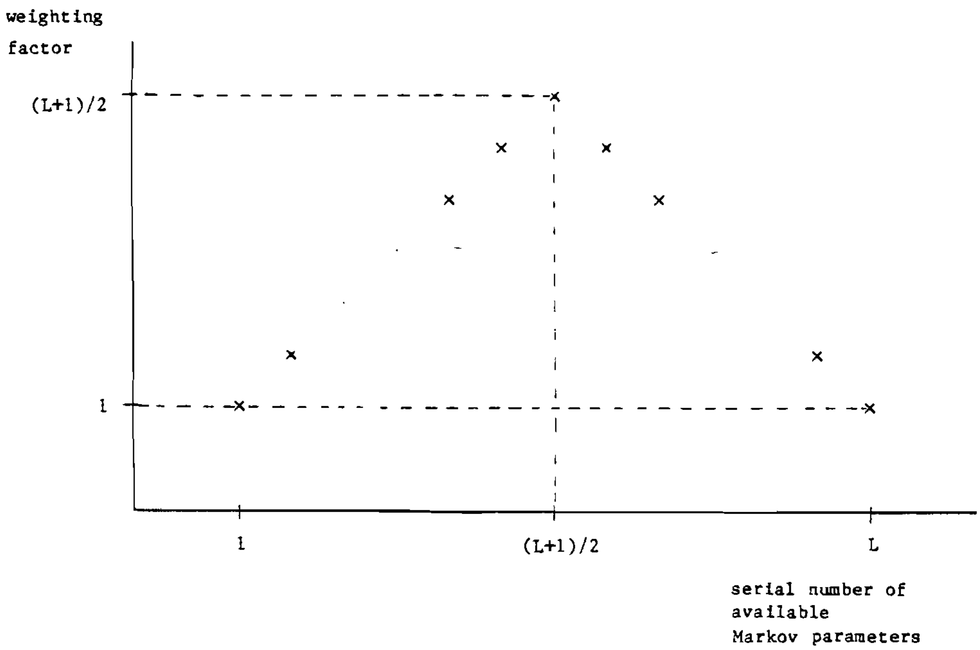


Fig-1.4.2- Weighting of the Markov parameters in the least squares approximation of the Hankel matrix.

If we assume that all Markov parameters are disturbed with SWAYING noise, as defined in the introduction, this property, as described above, will result in a nonequally balanced approximation of the realization of the system.

1.4.3. Order determination

The problem of order determination is concerned with the choice of k. Of course this k can be chosen a priori as an imposed restriction.

k also can be chosen according to the result of the singular value decomposition. The error between H_k and H:

$$\|H_k - H\|_E^2 = \sum_{j=k+1}^{\rho} \delta_j^2 \qquad \text{is dependent on k.}$$

So the absolute or the relative error that is made choosing some particular k, can be used to decide for some k.

The singular values of H, and specially the way in which they

decrease, are used for this choice. In the deterministic case this is clear: the first n singular values are unequal zero, and the following ones are zero. The gap between δ_n and δ_{n+1} indicates that the rank of the approximating Hankel matrix, and consequently the dimension of the realization should be chosen n.
In the noisy situation all available singular values will be unequal zero, and the gap between $\tilde{\delta}_n$ and $\tilde{\delta}_{n+1}$ becomes far less clear (see also fig-1.4.1-). A situation where $\tilde{\delta}_n/\tilde{\delta}_{n+1}$ is small can be a good indication for choosing the correct dimension n of the system.
To obtain such a situation, δ_n should not be very small; an optimum situation in this respect is a situation where the signal energy in H is equally distributed over $\delta_1{}^2,..,\delta_n{}^2$, and consequently all deterministic singular values are on the same level.
Also properties of the noise in the Hankel matrix can be used for determination of k. If we are able to predict the influence of the noise on the singular values $\tilde{\delta}_1,..\tilde{\delta}_\rho$, we are able to introduce an estimated noise level on singular values. Singular values below this level can be considered to be caused by the noise, and values above this level as noise disturbed deterministic singular values. An analysis of the influence of noise on singular values will be given in section 2.4 and 4.3.
In analysing these noise properties one meets the problem of the Hankel structure again. Because of the repetition of Markov parameters in the Hankel matrix, the noise contributions in this matrix are not independent.

1.4.4. Realization

Given a Hankel matrix with rank n and consisting of L Markov parameters, the Ho-Kalman algorithm will always result in an n-dimensional realization of the L Markov parameters in the Hankel matrix. This result is stated by eq.(1.7).
After noise filtering, however, when the Hankel matrix H is transformed to H_k, a problem arises. In general H_k will not be a real Hankel matrix with a block symmetric structure. Therefore in H_k there is not anymore a unique sequence of Markov parameters. This violation of the Hankel structure of H_k will be an extra

problem. For because of this the Ho-Kalman algorithm can not supply an exact realization of H_k; the sufficient condition of equivalence relation (1.7) is not fulfilled. Therefore in the application of formulae (1.16)-(1.18) on H_k, though H_k is not Hankel, a second approximation step is introduced in the algorithm.

1.5 CONCLUDING REMARKS

The Hankel matrix algorithms, as presented in section 1.2, create a minimum realization of a linear time-invariant system, when given a series of deterministic Markov parameters. In fact these algorithms are all designed for applying them to this deterministic case.

When a "truncated" series of Markov parameters has been disturbed with noise, the algorithms have to be modified: an approximating Hankel matrix has to be constructed for applying the realization algorithm.

Zeiger and McEwen (1974) introduced the use of singular value decomposition as a tool for executing the required decomposition of the Hankel matrix, as well as for the construction of the approximating Hankel matrix.

Yet both the Ho-Kalman algorithm and its modification by Kung, can be applied to the approximating Hankel matrix. However this approximating Hankel matrix has either an improper rank, or it lacks the Hankel structure.

Furthermore the Markov parameters in the Hankel matrix are not processed with a constant weighting factor during this approximation step. In case of the Ho-Kalman algorithm and Kung's modified version, an isosceles triangular weighting function in fact is applied (see fig-1.4.2-).

In this respect applying Silverman's algorithm to the approximating Hankel matrix will even be worse. Apart from the fact that its realization is not balanced, the applied weighting function is even less optimal then in the previous situation: only the first Markov parameters are used in the Hankel-submatrix. Moreover not all elements in these Markov parameters are necessarily used, and all information in the neglected remaining Markov parameters is lost. Therefore an application of Silverman's algorithm to noisy Markov parameters, in contrast with the other two algorithms, has never been proposed to our knowledge.

With respect to the approximation of a Hankel matrix with a Hankel matrix of reduced rank an alternative method is known from Silverman and Bettayeb (1980). However this approach, based on the work of Adamjan, Arov and Krein considers only infinite Hankel matrices and takes the spectral norm as an error-criterion, which is not relevant for our application.
This will be elucidated briefly: the use of the spectral norm as an error criterion between H and H_k is based on the argument that the highest eigenvalue of the residual between H and H_k should be minimized. This argument is initiated by the model reduction problem: a certain known model has to be approximated with a model of reduced order. The external behaviour of the two models should be as close as possible. The Hankel matrix now is looked upon as an operator on input signals in order to produce outputs:

$$Y[+j \| k] = H[j] \, U[-j \| k] \qquad (1.35)$$

where $U[-j \| k]$ is an array of j input vectors $\underline{u}_{k-1},\ldots,\underline{u}_{k-j}$
and $Y[+j \| k]$ an array of j output vectors $\underline{y}_k,\ldots,\underline{y}_{k+j-1}$

In this respect the spectral norm is a very useful tool, because it indicates the extreme output of the system for an arbitrary input.
In the subject under study however, the crucial aspect of the approximate realization problem is the identification of the system. The system has to be identified and the resulting array $\{\hat{M}_k\}$ is supposed to represent the essential characteristics of the model, and not just an output signal. In addition to this the Euclidian norm is a reasonable choice, because it considers a least squares fit on the complete array of Markov parameters. Moreover in case of SWAYING noise a Maximum Likelihood Estimate will require the Euclidian norm.

The special block symmetric structure of the Hankel matrix is shown to be a disadvantage in the noise filtering. However in the realization part the block symmetric structure of a Hankel matrix will guarantee that an exact reconstruction of the Hankel matrix can be found (see eq.(1.7)).

In an attempt to overcome the problem of the block symmetric structure a new approach to the realization problem has been studied. In the next section the approach with a Page matrix will be introduced.

CHAPTER 2: INTRODUCTION OF THE PAGE MATRIX

2.1. INTRODUCTION

In an attempt to find a new approach to the realization problem that overcomes the problems of the existing methods, a new matrix is introduced that represents the Markov parameters in a simple way.

$$P[\eta,\mu] = \begin{bmatrix} M_1 & M_2 & M_3 & \cdot & \cdot & M_\mu \\ M_{\mu+1} & \cdot & \cdot & & \cdot & M_{2\mu} \\ \cdot & & & \cdot & & \cdot \\ \cdot & & & \cdot & & \cdot \\ M_{(\eta-1)\mu+1} & & \cdot & & \cdot & M_{\eta\mu} \end{bmatrix} \qquad (2.1)$$

see also Damen, Van den Hof and Hajdasinski (1982b)

If it is possible to use this Page matrix for the functions described in the previous section, it will probably lead to better results of the approximation.
The usefulness of this Page matrix has to be proved, and especially the question has to be answered whether the same algorithms that are used for the Hankel matrix can be applied to the Page matrix.
In the Page matrix all Markov parameters appear only once, which means that, when reducing the rank with the aid of singular value decomposition, there is an equally balanced filtering over the parameters.
Moreover, a Page matrix of reduced rank provides a unique sequence of Markov parameters, contrary to the approximating Hankel matrix H_k of the previous chapter.
A third advantage of the Page matrix over the Hankel matrix is the smaller dimensions, when taking the same number of Markov parameters into account. This directly leads to faster calculations.

In the purely deterministic case, the Page matrix can be decomposed in a manner similar to the Hankel matrix:

$$P = \begin{bmatrix} C \\ CA^{\mu} \\ \cdot \\ \cdot \\ CA^{(\eta-1)\mu} \end{bmatrix} \cdot \begin{bmatrix} B & AB & A^2B & \cdot & \cdot & A^{\mu-1}B \end{bmatrix} \qquad (2.2)$$

$$P = \Gamma_{\mu}[\eta] \cdot \Delta[\mu] \qquad (2.3)$$

Whereas a Hankel matrix is the product of an observability matrix Γ and a controllability matrix Δ of the system $\{A,B,C\}$, a Page matrix is the product of an observability matrix of $\{A^{\mu},C\}$ and a controllability matrix of $\{A,B\}$.

When ordering the Markov parameters in the column direction in stead of the row direction, another matrix (the "Chinese" Page matrix) occurs, which is a product of the observability matrix of $\{A,C\}$ and the controllability matrix of $\{A^{\eta},B\}$.

In order to use the above Page matrix for decomposition and for realization of the deterministic system, a <u>necessary</u> condition for achieving a correct decomposition as in eq.(2.2) which leads to a minimum realization $\{A,B,C\}$ will be

$$\text{rank } P = n \qquad (2.4)$$

In order to fulfill this condition, it is necessary and sufficient that both the observability matrix and the controllability matrix of P have rank n. Therefore μ and η have to fulfill the condition $\mu \geq \beta$, $\eta \geq \alpha$.
The formal proof of the preservation of the rank in P is given in chapter 3, together with the necessary and sufficient conditions for the existence of a minimum realization $\{A,B,C\}$ via the Ho-Kalman algorithm applied to P.

Anticipating on the results of chapter 3, the crucial theorem in this context will be stated next:

<u>Theorem</u>:
If the dimensions of the Page matrix for a system, which has a minimum realization with dimension n, are chosen large enough: we take $\mu, \eta \geq n$ as a sufficient condition, and if (C, A^{μ}) is a completely observable couple, it holds that

$$\text{rank } P = n \tag{2.5}$$

and any decomposition in $\Gamma_{\mu}[\eta]$ and $\Delta[\mu]$ of minimum dimension n will lead to a minimum realization, according to the realization methods as presented in section 1.2.

This means that the matrices B and C can be determined based on the decomposition of P as stated in eq.(2.3) and that the matrix A can be determined by way of a shifting operation:

$$\overset{\leftarrow}{P} = \Gamma_{\mu}[\eta] \cdot A \cdot \Delta \tag{2.6}$$

where

$$\overset{\leftarrow}{P} = \begin{bmatrix} M_2 & M_3 & M_4 & \cdots & M_{\mu+1} \\ M_{\mu+2} & \cdot & \cdot & \cdot & M_{2\mu+1} \\ \cdot & & \cdot & & \cdot \\ \cdot & & \cdot & & \cdot \\ \cdot & & \cdot & & \cdot \\ M_{(\eta-1)\mu+2} & & \cdot & \cdot & M_{\eta\mu+1} \end{bmatrix} \tag{2.7}$$

In the same way as in section 1.2. the method of Kung can be applied with a shifting operation in the observability or controllability matrix.
An analysis of the consequences of the above mentioned theorem is given in section 2.2.

Whereas the block square Hankel matrix has got fixed dimensions, determined by the number of Markov parameters L, the dimensions of the Page matrix can be chosen within some range, as long as $\eta \cdot \mu = L$.
The special properties of the Page matrix according to this feature is analysed in section 2.3

In the situation that we assume the Markov parameters to be disturbed by independent noise, we can see that the noise in the Page matrix is completely independent. This might facilitate the analysis of the influence of noise, a task which has been quite difficult in case of the Hankel matrix. Attention will be paid to this subject in section 2.4.

2.2 DETERMINISTIC AND NOISY SITUATION

For a finite sequence of arbitrary Markov parameters $\{M_k\}$, $k=1,..,L$ a finite dimensional realization with some dimension n can always be found that reconstructs this sequence in an exact way. We can define the distinction between the <u>deterministic</u> and the <u>noisy</u> situation: in the deterministic situation the value of n will remain constant from a certain level on if we increase L; in the noisy case n will always increase with increasing L.
Having this in mind, based on the theorem in the previous section we can state (corresponding to eq.(1.7)):

$$\begin{tikzcd} \boxed{\begin{array}{l} \{M_k\},k=1,..,L \text{ has a finite dimensional} \\ \text{realization with minimum dimension } n \\ \text{and } (C,A^{\mu}) \text{ is an observable pair} \end{array}} \arrow[d, shift right=2ex, "\text{necessary}"'] \\ \boxed{\text{rank } P = n} \arrow[u, shift right=2ex, "\text{sufficient}"'] \end{tikzcd} \qquad (2.8)$$

where P is chosen sufficiently large.

The necessary condition of eq.(2.8) is out of discussion: both in the deterministic and in the noisy situation a sequence of Markov parameters with a finite dimensional realization and minimum dimension n will always lead to a Page matrix with rank n, under the conditions that: -the dimensions of P are chosen large enough
-(C,A^{μ}) is an observable pair.

The sufficient condition of eq.(2.8) gives rise to more problems. It can be stated as follows: given an arbitrary Page matrix with rank P = n; is it possible now, in all cases, to find a realization with minimum dimension n in such a way that the given Page matrix can be reconstructed in an exact way?
Ho and Kalman have proved the correctness of this statement for the Hankel matrix in their algorithm (1966). In this proof the block symmetric structure of the Hankel matrix was used.
The theorem in section 2.1. about the rank of the Page matrix states that it is possible to find an n dimensional realization of the system if there exists one. In fact nothing is said about the existence of such a realization when solely given that rank P =n.

By way of an example it can be shown that this condition indeed is not sufficient:
Assume a SISO system with:
$\{M_k\}_{,k=1,\ldots} = 1,1,1,1,0.5,0.25,0.125,0.0675,\ldots$
The impulse response can be described by a 4^{th} order system:

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.5 \end{bmatrix} \qquad \begin{array}{l} B = \begin{bmatrix} 1 & 1 & 1 & 1 \end{bmatrix}^T \\ \\ C = \begin{bmatrix} -1 & -2 & -4 & 8 \end{bmatrix} \end{array}$$

Physically this system can be interpreted as a first order SISO system in series with a time delay of three time instants.
If we fill a Page matrix with 8 of the given Markov parameters, for $\mu > 2$ the rank of P will never exceed 2. However it is impossible to find a correct realization of dimension 2.
In this example the couple (C,A^{μ}) is not an observable pair; at this moment it is not clear whether this failing observability condition (which of course is not known beforehand) is the only possible situation for the non sufficiency of the rank condition rank P = n in eq.(2.8).

Because of this lacking sufficient condition also in Page matrix algorithms a second non-exact approximation step is introduced: apart from the rank reduction, the realization also becomes an approximation.
In chapter 4 more attention will be paid to this feature.

2.3 CHOICE OF THE DIMENSIONS OF THE PAGE MATRIX

2.3.1 Introduction

In case of a block square Hankel matrix, a choice of the number of Markov parameters L that is taken into account, directly determines the dimensions of the Hankel matrix:

$$H[\gamma] = \begin{bmatrix} M_1 & M_2 & M_3 & \cdot & \cdot & M_\gamma \\ M_2 & M_3 & M_4 & \cdot & \cdot & M_{\gamma+1} \\ \cdot & \cdot & & \cdot & & \cdot \\ \cdot & \cdot & & & \cdot & \cdot \\ M_\gamma & M_{\gamma+1} & M_{\gamma+2} & \cdot & \cdot & M_L \end{bmatrix}_{L=2\gamma-1} \qquad (2.9)$$

Although it is not strictly necessary to choose the Hankel matrix in a block square way, this has very often been the case in the literature: Ho and Kalman (1966); Hajdasinski and Damen (1979); Damen and Hajdasinski(1982); Niederlinski and Hajdasinski (1979). In Silverman (1971), Kung (1978) and Staar, Vandewalle and Wemans (1981) a more general Hankel matrix is employed. To our knowledge an analysis of how to choose the dimensions of the Hankel matrix, if not chosen block square, has never been given.

In order to be able also to construct Hankel matrices containing an even number of Markov parameters, an extra block row will be added to the Hankel matrix in such situations.

Because the Page matrix is not restricted to a special block symmetric structure, the dimensions of the matrix can be chosen in many different ways.

Given the number of Markov parameters L, the choice of the block dimensions of P: $\eta \times \mu$ is restricted to $\eta . \mu = L$.

However the choice for η and μ has its influence on the results of the approximation algorithm. In the next paragraphs these influences will be analysed and criteria will be given for the choice of η and μ.

2.3.2 Restrictions given by the structure of P

Suppose the dimensions of the Page matrix are $h \times m$,

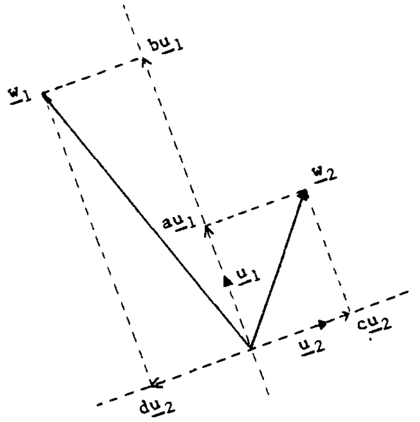
$$\text{where } h = \eta . q \qquad (2.10)$$
$$\text{and } m = \mu . p .$$

1. The first restriction on h and m comes directly from the theorem in section 2.1.
   To preserve the right rank of P:
   $$h > \alpha . q$$
   $$m > \beta . p \qquad (2.11)$$
   are necessary and sufficient conditions for using the Page matrix in the realization algorithm.

2. In case of a stable system the second block row in the Page matrix will contain elements with (much) smaller values than the elements in the first block row. This is caused by the fact that the impulse response of a stable system will tend to zero.
   This phenomenon may be harmful to the distribution of the "energy" in the Page matrix over the singular values after singular value decomposition of P.
   To make this clear, a further consideration of the properties of singular value decomposition is required.

   <u>Example</u>

   Assume a Page matrix consisting of two rows, that can be represented by two vectors $\underline{w}$ (see fig-2.3.1-).
   The singular value decomposition looks for a direction $\underline{u}_1$ in the space spanned by $\underline{w}_1,\underline{w}_2$, in such a way that the sum of squared projections of $\underline{w}_1$ and $\underline{w}_2$ on $\underline{u}_1$ becomes maximal.
   Next these projections are subtracted from the vectors $\underline{w}_1$ and $\underline{w}_2$ and with the remaining vectors the same procedure is repeated. For each procedure step the maximal squared sum of projections results in a value of a $\delta^2{}_i$.



$$\delta_1^2 = a^2 + b^2$$
$$\delta_2^2 = c^2 + d^2$$
$$\|\underline{u}_1\| = 1$$
$$\|\underline{u}_2\| = 1$$

fig-2.3.1- Singular value decomposition of a 2×2 Page matrix.

From the example it can be seen that if the ratio of the amplitudes $\|\underline{w}_1\|/\|\underline{w}_2\|$ becomes bigger, the first singular value will increase, and the second one will decrease.
Converting this remark to a general Page matrix, we can conclude that an increase of the timeshift of μ steps between the first and the second block row of P (the so-called "jump") creates an increase of the first q singular values, at the cost of a decrease of the higher ordered ones. This same feature appears when the columns of P are considered: a gap after the first p singular values is introduced which, however, is smaller because of the smaller jump (only 1 shift).
In the deterministic case this "driving a gap" between δ_q and δ_{q+1}, will have no influence on the final realization. In the noisy case, however, where the singular values are disturbed by noise, and where the order of the system has to be estimated according to these noisy singular values, it is an effect that could be harmful to the accuracy. This holds even more because, as far as the noise is concerned, there will be no jump, as the noise level does not decrease with increasing time.
For this reason the "jump" μ has to be minimized.

2.3.3. Optimal dimensions with respect to the noise

Assume that the elements of the Page matrix are all disturbed with SWAYING noise. This means that neither the noise on distinctive Markov parameters, nor the noise on elements within one Markov

parameter are correlated.
The total noise energy within the Page matrix can be written as

$N_P = \|\tilde{P} - P\|_E^2$ where $\tilde{P}$ is the Page matrix disturbed with noise.

According to eq.(1.32):

$$N_P = \sum_{i=1}^{\min(h,m)} \delta_{i,N}^2 \qquad (2.12)$$

where $\delta_{i,N}$ is a singular value of the noise matrix $(\tilde{P}-P) = \Xi_P$.

Because of the dimensions of $\tilde{P}$: $h{\times}m$, rank $\tilde{P}$ will be equal to $\min(h,m)$.

Now there can be written:

$$E\{N_P\} = E\{\|\Xi_P\|_E^2\} = E\{tr(\Xi_P\Xi_P^T)\} = h \cdot m \cdot \sigma^2 \qquad (2.13)$$

In the next section 2.4 the following will be proved:
When we consider a series of noise matrices Ξ_P and for each matrix we determine the product

$\tilde{P}\,\tilde{P}^T$ (if $h{\leqslant}m$) or $\tilde{P}^T\,\tilde{P}$ (if $m{\leqslant}h$), the expectation of this matrix product (the Grammian) can be written as:

$$\begin{aligned} E\{\tilde{P}\,\tilde{P}^T\} &= U\,(D^2 + \sigma^2 \cdot m \cdot I_h)\,U^T \qquad (h{\leqslant}m) \\ E\{\tilde{P}^T\,\tilde{P}\} &= V\,(D^2 + \sigma^2 \cdot h \cdot I_m)\,V^T \qquad (m{\leqslant}h) \end{aligned} \qquad (2.14)$$

If we choose the correct product to analyse, according to m and h, we can state that within the expectation of this product the singular values can be written as:

$$d_i^2 = \delta_i^2 + \sigma^2 .\ \max(h,m) \qquad (2.15)$$

The first n singular values, which we will select for the realization, have been corrupted by noise. Again according to the expectation of the Grammian matrix of $\tilde{P}$, this deterioration can be quantified by eq.(2.15). As it is desired to decrease the influence of the noise to a minimum level, (2.15) has to be minimized for $i{\leqslant}n$.

Minimizing $\max(h,m)$ is the same as choosing the Page matrix $\tilde{P}$ as square as possible; not in block-dimensions but in real dimensions.

Summarizing this result it can be said that, if we consider a fixed number of Markov parameters L, $\tilde{P}$ has to be chosen in such a way that the rank of $\tilde{P}$ is maximum. ($\tilde{P}$ as square as possible). Then the number of singular values is maximum, and because the noise energy in expectation matrices is equally distributed over all singular values, the influence of the noise on each singular value separately, is minimum.
For the noise filtering this result means that the noise energy that is filtered out when choosing the reduced rank n, is maximum.

2.3.4 Remarks

The results of this section will briefly be summarized:
There have been given 3 criteria for the choice of the dimensions of P, all three proposed from a different point of view:

1. $h > \alpha.q$, $m > \beta.p$ (section 2.3.2) (2.16)

   This is a necessary condition to ensure the Page matrix to have the correct rank even in the deterministic situation.
   It leads to $\eta > \alpha$, $\mu > \beta$.
   Because $n > s$ in practice often $\eta, \mu > n$ (2.17)
   is taken as a sufficient condition for the rank property of P.
   More attention will be paid to this subject in section 4.3.

2. m minimum (section 2.3.2) (2.18)

   This criterion has been generated as an optimal dimension for the behaviour of deterministic singular values.

3. max(h,m) minimum (section 2.3.3) (2.19)

   The analysis of the noisy singular values and the effect of the noise filtering has been the basis for this criterion.

Taking these criteria together, we do not know a best solution for dim(P) at this moment. There are several actions in the procedure, working at the same time. Their positive/negative influences have to be analysed. Notice that the number of Markov parameters L has been assumed to be a given number. In practice this is a variable that can be chosen freely, of course keeping in mind the specific application and its necessary restrictions.
In a later chapter this L will be the subject of discussion (chapter 4).

2.4 INFLUENCE OF NOISE ON SINGULAR VALUES

2.4.1. Introduction

Assume a Page matrix $\tilde{P}$ that is disturbed by SWAYING noise with variance σ^2.
We can write $\tilde{P} = P + \Xi_P$, where Ξ_P is the matrix containing all noise samples:

$$\Xi_P = \begin{bmatrix} \xi(1,1). & . & . \ \xi(1,m) \\ . & & . \\ \xi(h,1). & . & . \ \xi(h,m) \end{bmatrix} \tag{2.20}$$

and where $\xi(i,j)$ are independent noise samples.
In this section 2.4 the influence of Ξ_P on singular values will be analysed.
$\tilde{P}$ can be written as: $\tilde{P} = \tilde{W}\,\tilde{\Sigma}\,\tilde{V}^T$ (2.21)

We would like to know the influence of Ξ_P on $\tilde{\Sigma}$ and consequently on the singular values $\tilde{\delta}_i$ in $\tilde{\Sigma}$.
However up till now it has only been possible to find an expression for the influence of Ξ_P on the Grammian matrix of $\tilde{P}$ in expectation.

2.4.2 Analysis of the Grammian matrices of $\tilde{P}$

For the Page matrix $\tilde{P}$ two products can be analysed in order to recognize the influence of the noise on singular values:

$$\begin{aligned} \tilde{P}^T\tilde{P} &= P^TP + \Xi_P^T\Xi_P + \Xi_P^TP + P^T\Xi_P \\ \tilde{P}\tilde{P}^T &= PP^T + \Xi_P\Xi_P^T + \Xi_PP^T + P\Xi_P^T \end{aligned} \tag{2.22}$$

With eq.(2.21) there can be written:

$$\begin{aligned} \tilde{P}^T\tilde{P} &= \tilde{V}\,\tilde{\Sigma}^2\,\tilde{V}^T \\ \tilde{P}\tilde{P}^T &= \tilde{W}\,\tilde{\Sigma}^2\,\tilde{W}^T \end{aligned} \tag{2.23}$$

Because of the character of the noise, as a result of which Ξ_P and Ξ_P^T are matrices with zero expectation, the two last terms in both equations (2.22) will be zero in expectation.

With the singular value decomposition of P, there can be written:

$$E\{\tilde{P}^T\tilde{P}\} = V\Sigma^2V^T + E\{\Xi_p^T\Xi_p\}$$
$$E\{\tilde{P}\tilde{P}^T\} = W\Sigma^2W^T + E\{\Xi_p\Xi_p^T\} \qquad (2.24)$$

This expectation is taken with respect to an ensemble of noise matrices Ξ_p and therefore with respect to an ensemble of Page matrices $\tilde{P}$.

$\tilde{P}^T\tilde{P}$ is an m×m matrix

$\tilde{P}\tilde{P}^T$ is an h×h matrix.

To recognize the influence of Ξ_p on singular values,it is prefered to analyse thât Grammian matrix of $\tilde{P}$ where the dimensions of the noise part in (2.24) coincide with the dimensions of Σ^2. As the dimensions of Σ^2 are min(h,m)×min(h,m), we will analyse

$$\tilde{P}^T\tilde{P} \quad \text{if } m<h,$$
$$\text{and} \quad \tilde{P}\tilde{P}^T \quad \text{if } h<m.$$

In the remaining part of this section the situation h<m will be assumed. The results in the dual case m<h will be similar.

The matrix $\Xi_p\Xi_p^T$ can be written as:

$$\Xi_p\Xi_p^T = \begin{bmatrix} \sum_{i=1}^{m} \xi^2(1,i) & \cdots & \sum_{i=1}^{m} \xi(1,i)\xi(h,i) \\ \vdots & & \vdots \\ \sum_{i=1}^{m} \xi(h,i)\xi(1,i) & \cdots & \sum_{i=1}^{m} \xi^2(h,i) \end{bmatrix} \qquad (2.25)$$

Because of the character of the noise the diagonal terms in this product matrix will follow a χ^2-distribution. With respect to the off-diagonal terms only conclusions can be drawn when we consider the expectation of this matrix:

$$E\left\{\sum_{i=1}^{m} \xi(j,i)\xi(k,i)_{\substack{j,k<h\\ j\neq k}}\right\} = 0 \qquad (2.26)$$

For the product matrix $\Xi_p\Xi_p^T$ we now can write:

$$(\Xi_p\Xi_p^T)_{ii} = \chi_m^2\sigma^2 \quad , \quad i<h \qquad (2.27)$$

$$E\{(\Xi_p\Xi_p^T)_{ij}\} = 0 \quad , \quad i\neq j,\ i,j<h \qquad (2.28)$$

These two equations lead to:

$$E\{\Xi_p \Xi_p^T\} = E\{\chi_m^2\} \cdot \sigma^2 \cdot I_h$$

$$= m \cdot \sigma^2 \cdot I_h \qquad (2.29)$$

Referring back to eq.(2.24) there can be seen that the matrix $E\{\Xi_p \Xi_p^T\}$ has the same dimensions and the same structure (diagonality) as the squared singular value matrix Σ^2.
Because $E\{\Xi_p \Xi_p^T\}$ is a diagonal matrix with constant diagonal elements, and because W is an orthonormal, square matrix, there can be written:

$$\begin{aligned} E\{\tilde{P}\tilde{P}^T\} &= W\Sigma^2 W^T + W\, m\sigma^2 I_h\, W^T \\ &= W\,(\Sigma^2 + m\sigma^2 I_h)\, W^T \end{aligned} \qquad (2.30)$$

So there exists a matrix $\tilde{D}^2$ in the expectation of $\tilde{P}\tilde{P}^T$ with squared singular values, that can be written as:

$$E\{\tilde{P}\tilde{P}^T\} = WD^2W^T \qquad \text{where } D^2 = \Sigma^2 + m\sigma^2 I_h \qquad (2.31)$$

and consequently for all squared singular values $d_i{}^2$ in D^2:

$$d_i^2 = \delta_i^2 + m\,\sigma^2 \qquad (2.32)$$

When the assumption m>h is left out, the last two equations can be generalized to:

$$D^2 = \Sigma^2 + \sigma^2 . \max(h,m) . I_{\min(h,m)} \qquad (2.33)$$

$$d_i^2 = \delta_i^2 + \sigma^2 . \max(h,m) \qquad (2.34)$$

From these results it can be seen that the number of singular values of the Grammian matrix under consideration equals min(h,m) and that each singular value represents the energy of max(h,m) noise sources. Moreover, the equation above makes clear that the disturbance of the singular values is additive in the squared sense.
However it should be noted, as mentioned in the introduction, that these singular values do not represent the singular values of $\tilde{P}$.

A comparable study in case of a Hankel matrix is given in section 4.3.

CHAPTER 3: RANK CONDITION OF THE PAGE MATRIX

3.1 INTRODUCTION

This chapter is devoted to the central theorem stated in section 2.1, where the idea of a Page matrix was introduced. The theorem is made concrete by deriving all conditions under which the rank of the deterministic Page matrix equals the dimension of the minimum realization.

It will be shown that complete observability of the system $\{C,A^{\mu}\}$ given a complete observable system $\{C,A\}$ is a necessary and sufficient condition for the validity of rank P = n.

This replaces the problem to complete observability of $\{C,A^{\mu}\}$. Sufficient conditions for the observability of $\{C,A^{\mu}\}$ given observability of $\{C,A\}$ are given by Hautus (1969). Necessary conditions however appear to be less severe. Both conditions will be derived in this chapter. These results were published in a similar way in Damen, Van den Hof and Hajdasinki (1982a)

3.2 PROBLEM REDUCTION TO THE OBSERVABILITY OF SYSTEM $\{A^{\mu},B,C\}$

Assuming a sequence of deterministic Markov parameters $\{M_k\}_{k=1,..,L+1}$ with a finite dimensional realization $\{A,B,C\}$, the Hankel matrix can be written as:

$$H = \Gamma[\gamma] \cdot \Delta[\gamma] \quad \dim(H)=g\times\ell \tag{3.1}$$

where

$$\Gamma[\gamma] = \begin{bmatrix} C \\ CA \\ CA^2 \\ \cdot \\ \cdot \\ CA^{\gamma-1} \end{bmatrix} \quad \Delta[\gamma] = [\ B \quad AB \quad A^2B \quad \cdot \quad \cdot \quad A^{\gamma-1}B\ ] \tag{3.2}$$

$$\dim(\Gamma)=g\times n$$
$$\dim(\Delta)=n\times\ell$$

where $\gamma=(L+1)/2$.

If γ is large enough, the γ-observability matrix $\Gamma[\gamma]$ and the γ-controllability matrix $\Delta[\gamma]$ have full rank n. The limit which we will use is $\gamma \geq n$ though theoretically the condition $\gamma \geq r$ is sufficient where r is the realizability index, or even a non block square Hankel matrix could be used (see section 1.2.1).
The Hankel matrix H transforms the space R_ℓ into R_g via R_n and this is made explicit by means of Δ $(R_\ell \rightarrow R_n)$ and Γ $(R_n \rightarrow R_g)$. A lot of freedom, however, is left by not defining the base in R_n. All possibilities together define the equivalence class which is invariant under the following equivalence transformation: If the following two sets correspond to the same Hankel matrix:

$$(\Gamma,\Delta,A,B,C) \xleftrightarrow{\text{equivalent in H}} (\Gamma^*,\Delta^*,A^*,B^*,C^*)$$

then the equivalence transformation is given by:

$$\begin{array}{ll} \Gamma = \Gamma^* \, T^{-1} & \Delta = T \, \Delta^* \\ C = C^* \, T^{-1} & B = T \, B^* \\ A = T \; A^* \, T^{-1} & \quad \text{where T nonsingular,} \dim(T)=n\times n \end{array} \tag{3.3}$$

A numerically stable solution for Γ and Δ can be found from a singular value decomposition of H :

$$H = W \, \Sigma \, V^T$$

and consequently:

$$H = (W\Sigma^{\frac{1}{2}})(\Sigma^{\frac{1}{2}}V^T) = \Gamma \, \Delta \tag{3.4}$$

so that we may choose: $\Gamma = W \, \Sigma^{\frac{1}{2}}$

$$\Delta = \Sigma^{\frac{1}{2}} \, V^T \tag{3.5}$$

(Because of the orthonormality of W and V, both matrices $(W\Sigma^{\frac{1}{2}})$ and $(\Sigma^{\frac{1}{2}}V^T)$ will have rank n.)
Based on this decomposition of H into two matrices of rank n, a minimum realization $\{A,B,C\}$ can be found (see Zeiger and McEwen (1974) and Damen and Hajdasinski (1982)) by:

$$C = E_q^{\gamma q} \; W \quad \Sigma^{\frac{1}{2}} \tag{3.6}$$

$$B = \Sigma^{\frac{1}{2}} \; V^T \; E^p_{\gamma p} \tag{3.7}$$

$$A = \Sigma^{-\frac{1}{2}} \, W^T \; \overleftarrow{H} \quad V \quad \Sigma^{-\frac{1}{2}} \tag{3.8}$$

where

$$E_q^{\gamma q} = [\; I_q \quad \emptyset_q \quad \emptyset_q \quad \cdot \quad \cdot \quad \emptyset_q \,] \tag{3.9}$$

$$E^p_{\gamma p} = [\; I_p \quad \emptyset_p \quad \emptyset_p \quad \cdot \quad \cdot \quad \emptyset_p \,]^T \tag{3.10}$$

and the shifted Hankel matrix:

$$\overleftarrow{H} = \Gamma \cdot A \cdot \Delta = \begin{bmatrix} M_2 & M_3 & M_4 & \cdot & \cdot & M_{\gamma+1} \\ M_3 & M_4 & \cdot & \cdot & \cdot & M_{\gamma+2} \\ M_4 & \cdot & & & & \cdot \\ \cdot & \cdot & & & & \cdot \\ \cdot & \cdot & & & & \cdot \\ M_{\gamma+1} & M_{\gamma+2} & \cdot & \cdot & \cdot & M_{2\gamma} \end{bmatrix} \tag{3.11}$$

Whether this same algorithm for finding a minimum realization may be applied to the Page matrix is dependent on the fact whether the rank of P equals the minimum dimension of the system n.
The Page matrix may be decomposed as:

$$P = \Gamma_\mu[\eta] \cdot \Delta[\mu] \qquad \dim(P) = h \times m \tag{3.12}$$

where

$$\Gamma_\mu[\eta] = \begin{bmatrix} C \\ CA^\mu \\ CA^{2\mu} \\ \cdot \\ \cdot \\ CA^{(\eta-1)\mu} \end{bmatrix} \qquad \begin{aligned} &\Delta[\mu] = [\; B \quad AB \quad A^2B \quad \cdot \quad \cdot \quad A^{\mu-1}B \;] \\ &\dim(\Gamma_\mu[\eta]) = h \times n \\ &\dim(\Delta[\mu]) = n \times m \end{aligned} \tag{3.13}$$

It is obvious, that <u>if</u> $\eta > n$ and $\mu > n$ and <u>both</u> the system $\{A,B,C\}$ is completely observable and controllable <u>and the system $\{A^\mu,B,C\}$ is completely observable</u>, the <u>same</u> equivalence class can be defined for P as for H, with exactly the <u>same</u> equivalence transformation. The crucial condition is that the system $\{A^\mu,B,C\}$ is completely observable,because then the η-observability matrix $\Gamma_\mu[\eta]$ has full rank n.

(In the dual case of the Chinese Page matrix, the controllability of $\{A^\eta,B,C\}$ is required to assure the full rank of $\Delta_\eta[\mu]$).

It can be proved that the rank $\Gamma_\mu[\eta]$=n (which will be the hardest job),then the following can be stated about the rank of P; A general rule for the rank of a product of matrices XY is

rank XY $\leq$ min(rank X,rank Y). The Sylvester's inequality (see Gantmacher (1959) p.66) is an extension of this relation: it states that for the rank of the product of rectangular matrices X and Y of dimensions resp. $m\times n$ and $n\times q$, it holds that:

rank X + rank Y - n $\leq$ rank(XY) $\leq$ min(rank X,rank Y)

Applying this inequality to the Page matrix (eq.(3.12)),this leads to

$$n \leq \text{rank } P \leq n \tag{3.14}$$

And as the result :

$$\text{rank } P = n \tag{3.15}$$

If rank $\Gamma_\mu[\eta]$=n and consequently rank P=n, a singular value decomposition of P leads to a decomposition of P into two matrices each with rank n. From this decomposition a minimum realization can be found in the same way as for the Hankel matrix, i.c. using the Ho-Kalman algorithm. As for the Hankel matrix all the subsequent steps are invariant under the equivalence transformation (with nonsingular matrix T).

In the remaining part of this chapter the condition rank P=n will be a subject of study. As a summary it can be stated that if we can prove that rank $\Gamma_\mu[\eta]$=n, this immediately leads to the required rank condition for P. This feature will be discussed in the next section, where we will just indicate all exceptions for which rank $\Gamma_\mu[\eta] \neq n$.

3.3 OBSERVABILITY OF $\{A,B,C\}$

3.3.1. Introduction

Our task is to prove that

$$\text{rank} \begin{bmatrix} C \\ CA^{\mu} \\ CA^{2\mu} \\ \cdot \\ \cdot \\ CA^{(\eta-1)\mu} \end{bmatrix} = n \tag{3.16}$$

given a minimum realization $\{A,B,C\}$. If $\{A,B,C\}$ is a minimum realization, it is proved by Ho and Kalman (1966) that

$$\text{rank} \begin{bmatrix} C \\ CA \\ \cdot \\ \cdot \\ CA^{\eta-1} \end{bmatrix} = n \tag{3.17}$$

in other words : $\{A,B,C\}$ is a completely observable realization. Rank $\Gamma_{\mu}[\eta]$=n can only be assured if $\eta \geq n$. This is a similar condition as the condition $\mu \geq n$ for the extended controllability matrix $\Delta_{\eta}[\eta]$.
The condition rank $\Gamma_{\mu}[\eta]$=n corresponds to the statement that $\{A^{\mu},B,C\}$ is a completely observable system. The final formulation of the problem now becomes:

Given a completely observable system $\{A,B,C\}$; under which conditions is the system $\{A^{\mu},B,C\}$ also completely observable for any $\mu > 1$.

To deal with the problem of complete observability of systems a congruent definition of this feature will be introduced, taking into account the structure of the matrix A, and more specifically its eigenvalues.
Such a definition is given by Chen and Desoer (1968) in the first instance for complete controllability. Their theorem will be stated here and the proof of the theorem will be given along the same lines as they did. This will be necessary for applying the theorem

to the situation of the system $\{A^{\mu},B,C\}$. For this purpose the Jordan canonical form is necessary, which will be defined next.

3.3.2. Jordan canonical form

Consider a system with ν different eigenvalues. The system can always be represented in its Jordan canonical form, in which there are no two Jordan blocks associated with the same eigenvalue. This Jordan form can be written as follows:

$$A_J = \underset{(n\times n)}{\begin{bmatrix} A_1 & 0 & \cdot & \cdot & \cdot & 0 \\ 0 & A_2 & \cdot & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & & & \cdot \\ \cdot & \cdot & & & \cdot & \cdot \\ 0 & \cdot & & \cdot & \cdot & A_\nu \end{bmatrix}}, \quad C_J = \underset{(q\times n)}{[\, C_1 \; C_2 \; \cdot \; \cdot \; C_\nu \,]} \tag{3.18}$$

where A_i and C_i will denote all Jordan blocks associated with eigenvalue λ_i.

Every Jordan block can be represented by a numer of Jordan cages, again ordered in a block diagonal way:

$$A_i = \underset{(n_i\times n_i)}{\begin{bmatrix} A_{i1} & & & \emptyset & \\ & A_{i2} & & & \\ & & \cdot & & \\ & \emptyset & & \cdot & \\ & & & & A_{ir(i)} \end{bmatrix}}, \quad C_i = \underset{(q\times n)}{[\, c_{i1} \; c_{i2} \; \cdot \; \cdot \; c_{ir(i)} \,]} \tag{3.19}$$

With every eigenvalue λ_i there are associated r(i) Jordan cages. This Jordan cages have the following form:

$$A_{ij} = \underset{(n_{ij}\times n_{ij})}{\begin{bmatrix} \lambda_i & 1 & 0 & \cdot & \cdot & \cdot & 0 \\ 0 & \lambda_i & 1 & 0 & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & 0 & \cdot & & \cdot & \cdot & 1 \\ 0 & \cdot & \cdot & \cdot & & \cdot & \lambda_i \end{bmatrix}}, \quad C_{ij} = \underset{(q\times n_{ij})}{[\, \underline{c}_{1ij} \; \underline{c}_{2ij} \cdot \cdot \underline{c}_{nij} \,]} \tag{3.20}$$

where $\underline{c}_{1ij}$ and $\underline{c}_{nij}$ are the first and the last column of C_{ij}, respectively.

The matrix A in the Jordan canonical form is completely defined by all numbers $\lambda_i \in \mathbb{C}$,and n_{ij},j=1..r(i), i=1...ν.

λ_i, i=1..ν define the values of the diagonal elements,

n_{ij}, i=1...ν,j=1...r(i) define the structure of matrix A

$$\sum_{j=1}^{r(i)} n_{ij} = \text{the multiplicity of "pole" } \lambda_i$$

For the remainder of this chapter we assume,that $\{A,B,C\}$ has been brought into a Jordan canonical form $\{A_J,B_J,C_J\}$, so that the index J will be dropped.

3.3.3. A criterion for observability of $\{A,B,C\}$

The theorem of Chen and Desoer (1968) now states the following: The system $\{A,B,C\}$ is completely observable if and only if the condition E holds, where

E: for each i=1,2,...ν, the set of r(i) q-dimensional column vectors $\underline{c}_{1i1},\underline{c}_{1i2},\cdots\underline{c}_{1ir(i)}$ is a linearly independent set.

Note that these are the columns of C corresponding to <u>that</u> (first) state in each cage, which is independent on <u>all</u> other states in that cage.

In the dual case of controllability ,those rows of B have to form an independent set which correspond to <u>that</u> (last) state in each cage, which is independent on (but influencing) all other states in that cage. (see Chen and Desoer (1968))

For poles with multiplicity one (single poles) the set consists of just one element (row or column) and independence then means, that this is not a zero vector. At least one input should influence the corresponding state or in the dual case at least from one output one should be able to observe the corresponding state.

To prove this theorem there has to be demonstrated that this property of A and C fits with the definition of Kalman:

$$\text{rank } \Gamma[n] = n \tag{3.21}$$

For the clearness of this text there will be defined:

$$\Gamma(A,C,n) = \begin{bmatrix} C \\ CA \\ CA^2 \\ \cdot \\ \cdot \\ CA^{n-1} \end{bmatrix} \tag{3.22}$$

To prove the equivalence relation

$$\text{condition } E \leftrightarrow \text{rank } \Gamma(A,C,n) = n \tag{3.23}$$

we will need three assertions that will be stated first:

<u>Assertion 1:</u> For any integer n, any complex λ, and any matrices A and C having proper dimensions

$$\text{rank } \Gamma(A,C,n) = \text{rank } \Gamma(A-\lambda I,C,n) \tag{3.24}$$

Proof : $C(A-\lambda I)^i$ can always be written as a linear combination of CA^i, $CA^{i-1} \cdot \cdot \cdot$, C, and therefore it holds that $\Gamma(A-\lambda I,C,n)$ is obtainable from $\Gamma(A,C,n)$ by applying to it a sequence of elementary row operations. More specifically there exists a nonsingular matrix Q(hxh) in such a way that

$$\Gamma(A-\lambda I,C,n) = Q\Gamma(A,C,n) \tag{3.25}$$

According to the Sylvester inequality (Gantmacher (1959) p.18) it can be written:

$$\text{rank}\Gamma(A,C,n)+h-h \leq \text{rank}\Gamma(A-\lambda I,C,n) \leq \text{rank}\Gamma(A,C,n) \tag{3.26}$$

Therefore it follows that eq. (3.24) holds.
In words : the rank of the observability matrix will not change when all diagonal elements of A are increased or decreased with the same constant $\lambda \in \mathbb{C}$ (which simply means a change of the origin in the complex z-plane).

<u>Assertion 2:</u> rank $\Gamma(A_i,C_i,s)$ = rank $\Gamma(A_i,C_i,\overline{n_i})$ for all $s > \overline{n_i}$, where $\overline{n_i} = \max\{ n_{ij}, j=1,2,\ldots r(i)\}$.

Proof: Because of the special Jordan structure of A_i it follows that

$$(A_i - \lambda_i I)^{\overline{n_i}} = \emptyset \tag{3.27}$$

and therefore $\quad (A_i - \lambda_i I)^s = \emptyset \quad$ for all $s > \overline{n_i}$ $\quad$ (3.28)

Because of assertion 1:

$$\text{rank } \Gamma(A_i,C_i,s) = \text{rank } \Gamma(A_i - \lambda_i I,C_i,s) \tag{3.29}$$

With the result above this leads to

$$\text{rank } \Gamma(A_i,C_i,s) = \text{rank } \Gamma(A_i,C_i,n_i) \quad \text{for all } s > \overline{n_i} \tag{3.30}$$

In words: a Jordan canonical form of a system with one distinct eigenvalue, forms an observability matrix for which holds that the number of matrix products that has to be taken into account to come to the maximal rank of Γ, is determined by the dimension of the largest Jordan cage.

<u>Assertion 3:</u> If there exist a nonzero n-dimensional column vector $\underline{q}$ such that $\Gamma(A,C,n)\underline{q} = \underline{0}$, then for any complex λ, $\Gamma(A-\lambda I,C,n)\underline{q} = \underline{0}$

Proof: From assertion 1 it follows that

$$\Gamma(A-\lambda I,C,n) = Q\Gamma(A,C,n) \tag{3.31}$$

consequently if $\Gamma(A,C,n)\underline{q} = \underline{0}$, then for any $\lambda \in \mathbb{C}$

$$\Gamma(A-\lambda I,C,n)\underline{q} = \underline{0} \tag{3.32}$$

For the proof of the theorem, as stated at the beginning of this section, some more resources will be needed ;a schematical way of representing $\Gamma(A,C,n)$ is given next.

Because of the block diagonal structure of A, $\Gamma(A,C,n)$ can be written as:

$$\Gamma(A,C,n) = \left[\Gamma(A_1,C_1,n) \quad \Gamma(A_2,C_2,n) \quad \cdot \quad \cdot \quad \cdot \quad \Gamma(A_\nu,C_\nu,n)\right] =$$
$$= \left[\Gamma(A_{11},C_{11},n) \quad \cdot \quad \cdot \quad \cdot \quad \Gamma(A_{\nu r(\nu)},C_{\nu r(\nu)},n)\right] \quad (3.33)$$

When we take the Jordan block associated with eigenvalue λ_k, then for the Jordan cage with index j these can be written.

$$A_{kj} - \lambda_k I = \underset{(n_{kj}\times n_{kj})}{\begin{bmatrix} 0 & 1 & 0 & \cdot & \cdot & 0 \\ 0 & 0 & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & 1 \\ 0 & \cdot & \cdot & \cdot & \cdot & 0 \end{bmatrix}} \quad (3.34)$$

which leads to

$$\Gamma(A_{kj}-\lambda_k I,C_{kj},n) = \begin{bmatrix} \underline{c}_{1kj} & \underline{c}_{2kj} & \cdot & \cdot & \cdot & \underline{c}_{nkj} \\ 0 & \underline{c}_{1kj} & \cdot & \cdot & \cdot & \underline{c}_{(n-1)kj} \\ \cdot & & \cdot & & & \cdot \\ \cdot & \cdot & \cdot & & \cdot & \cdot \\ 0 & \cdot & \cdot & & \cdot & \underline{c}_{1kj} \\ 0 & \cdot & \cdot & & \cdot & 0 \\ \cdot & \cdot & \cdot & & \cdot & \cdot \\ 0 & \cdot & \cdot & & \cdot & 0 \end{bmatrix} \quad (3.35)$$

Now all instruments are available to prove the theorem.

3.3.4. Proof of the theorem of Chen and Desoer

There is going to be proved that rank $\Gamma(A,C,n) = n$ if and only if condition E, as described in section 3.2.3., holds.

1. Necessary condition: $(\rightarrow)$

Suppose rank $\Gamma(A,C,n) = n$ and E does not hold for some $i=k$. This means that the set

$\underline{c}_{1k1},\underline{c}_{1k2},\ldots,\underline{c}_{1kr(k)}$ is a linearly dependent set.

Now consider the way of writing $\Gamma(A,C,n)$ and $\Gamma(A_k-\lambda_k I,C_k,n)$ as in eq.(3.33) and (3.35). If the given set of $\underline{c}$-vectors is linearly dependent, then there exists a linear combination of columns of $\Gamma(A-\lambda_k I,C,n)$ that is linear dependent. From this it follows that rank $\Gamma(A-\lambda_k I,C,n) < n$ and , by assertion 1, that rank $\Gamma(A,C,n) < n$. This is in contradiction with the hypothesis, and therefore for rank $\Gamma(A,C,n) = n$, a necessary condition is given by E.

2. Sufficient condition:($\leftarrow$)

This proof will be done in two steps. First it will be demonstrated that if E holds, this leads to:

$$\text{rank } \Gamma(A_i,C_i,n) = n_i \qquad \text{for all } i \tag{3.36}$$

Given this condition it will be proved that rank $\Gamma(A,C,n) = n$.

a. If E holds then rank $\Gamma(A_i,C_i,n) = n_i$ for all i.

Proof: From assertion 1 it follows that

$$\text{rank } \Gamma(A_i,C_i,n) = \text{rank } \Gamma(A_i- \lambda_i I,C_i,n) \tag{3.37}$$

With assertion 2 there can be written :

$$\text{rank } \Gamma(A_i,C_i,n) = \text{rank } \Gamma(A_i- \lambda_i I,C_i,\overline{n_i}) \tag{3.38}$$

Now suppose that rank $\Gamma(A_i- \lambda_i I,C_i,\overline{n_i}) < n_i$. The considered matrix has dimensions $\overline{n_i}q \times n_i$ (q= number of outputs); if its rank is smaller then n_i, a linear combination of the n_i column vectors of Γ can be brought to zero.

In other words:

There exists a nonzero column vector $\underline{q}$ in such a way that

$$\Gamma(A_i- \lambda_i I,C_i,\overline{n_i})\underline{q} = \underline{0} \tag{3.39}$$

The matrix $\Gamma(A_i- \lambda_i I,C_i,\overline{n_i})$ can be written as (in the example $\overline{n_i} = n_{1k}$)

$$\left[\begin{array}{cccc:cccc:c} \underline{c}_{1k1} & \underline{c}_{2k1} & \cdots & \underline{c}_{nk1} & \underline{c}_{1k2} & \cdots & \underline{c}_{nk2} & & \\ & \ddots & & \vdots & & \ddots & \vdots & & \\ & & \ddots & \vdots & & & \underline{c}_{1k2} & & \cdots \\ & \emptyset & & \vdots & & & 0 & & \\ & & & \underline{c}_{1k1} & 0 & \cdots & 0 & & \end{array}\right] \qquad (3.40)$$

Because $\underline{c}_{1k1}, \underline{c}_{1k2}, \ldots, \underline{c}_{1kr(k)}$ is a linearly independent set, it can be seen that $\Gamma(A_i-\lambda_i I, C_i, n_i)\cdot\underline{q}$ can only be $\underline{0}$ if $\underline{q} = \underline{0}$.
This is in contradiction with the assumption, and therefore it must hold that

$$\text{rank } \Gamma(A_i-\lambda_i I, C_i, \overline{n_i}) = n_i \qquad (3.41)$$

Because $n \geq \overline{n_i}$, and with assertions 1. and 2. it follows that

$$\text{rank } \Gamma(A_i, C_i, n) = n_i \quad \text{for all } i \qquad (3.42)$$

b. If E holds then all columns of $\Gamma(A_i,C_i,n)$ and $\Gamma(A_j,C_j,n), i \neq j$, will be independent.

Proof: Suppose

$$\text{rank } [\ \Gamma(A_1,C_1,n) \| \Gamma(A_2,C_2,n)\ ] < n_1 + n_2 \qquad (3.43)$$

Then there exists a n_1+n_2-dimensional column vector

$$\begin{matrix}\underline{q}_1 \\ \underline{q}_2\end{matrix} \neq \underline{0}, \text{ and } [\Gamma(A_1,C_1,n) \| \Gamma(A_2,C_2,n)] \begin{matrix}\underline{q}_1 \\ \underline{q}_2\end{matrix} = \underline{0} \qquad (3.44)$$

With assertion 3 it follows then:

$$[\ \Gamma(A_1,C_1,n) \| \Gamma(A_2,C_2,n)\ ] \begin{matrix}\underline{q}_1 \\ \underline{q}\end{matrix} = \underline{0} \qquad (3.45)$$

$$(A_2-\lambda_2 I)^s = \emptyset \text{ for all } s \geq \overline{n_2} \qquad (3.46)$$

Then it follows for the $(n-\overline{n_2}) \times n_1$ left under part of the matrix equation (3.45):

$$\begin{bmatrix} C_1(A_1-\lambda_2 I)^{\overline{n_2}} \\ \vdots \\ C_1(A_1-\lambda_2 I)^{n-1} \end{bmatrix} \cdot \underline{q}_1 = \underline{0} \qquad (3.47)$$

In other words:

$$\Gamma(A_1-\lambda_2 I,C\ ,n-\overline{n_2})(A_1-\lambda_2 I)^{\overline{n_2}}\cdot\underline{q}_1 = \underline{0} \qquad (3.48)$$

Since $\lambda_1 \neq \lambda_2$, rank $(A_1-\lambda_2 I)^{\overline{n_2}} = n_1$ (3.49)

rank $\Gamma(A_1-\lambda_2 I,C_1,n-\overline{n_2})$ = rank $\Gamma(A_1,C_1,n-\overline{n_2})$ =

= rank $\Gamma(A_1,C_1,\overline{n_1})$ because $\overline{n_1} < n-\overline{n_2}$ (3.50)

With equation (3.42) it has been proved that rank $\Gamma(A_i,C_i,\overline{n_i}) = n_i$, so rank $\Gamma(A_1,C_1,\overline{n_1}) = n_1$ (3.51)

Equations (3.49) and (3.50) together with the Sylvester inequality now show that equation (3.48) is a vector equation with a $n_1 \times n_1$-coefficient matrix of full rank n_1.

As a result equation (3.48) can only be fulfilled if $\underline{q}_1=\underline{0}$.

Along the same lines with substitution of λ_1 in equation (3.45) it follows that $\underline{q}_2=\underline{0}$.

Then $\underline{q}_1=\underline{0}$ and $\underline{q}_2=\underline{0}$ and this is in contradiction with the hypothesis.

Therefore:

$$\text{rank}\left[\Gamma(A_i,C_i,n)\,\|\,\Gamma(A_j,C_j,n)\right] = n_i+n_j \quad (\text{if } \lambda_i \neq \lambda_j) \qquad (3.52)$$

RESULT:

When we consider the representaion of A and A_i in equation(3.18) and (3.19), it can be seen that $n = \sum_{i=1}^{r(i)} n_i$.

Because of the fact that every value of i is associated with a different λ_i ,it follows that

$$\text{rank } \Gamma(A,C,n) = n \qquad (3.53)$$

With this result the theorem as stated in section 3.2.3 has been proved.

3.3.5. Remarks

With the given criterion for complete observability it is much more easy to analyse the observability of a system with a more physical understanding than with the definition of Kalman, at least if the Jordan canonical form of the system is known.
Because of our purely theoretical interest in the definition at this moment, this criterion is very suitable.
In the next section it will be demonstrated under which circumstances $\{A^{\mu},B,C\}$ is completely observable, given $\{A,B,C\}$ is completely observable. The special structure of the matrix A in the Jordan canonical form will be a great help in this task.

3.4. OBSERVABILITY OF $\{A^{\mu},B,C\}$

3.4.1. Jordan structure of A^{μ}.

Our goal is to find all possible situations where the system $\{A^{\mu},B,C\}$ is nonobservable, whereas the system $\{A,B,C\}$ is completely observable.
In these cases, and only in these cases, the Ho-Kalman algorithm applied to the Page matrix will not lead to a minimum realization.
The criterion for complete observability, as introduced by Chen and Desoer (1968) and described in the previous section, can quite easily be applied to the new system $\{A^{\mu},B,C\}$.
When A is assumed to be in the Jordan canonical form, as described in equation (3.18), A^{μ} can be found by raising each Jordan cage to the power μ.
It is known (see Gantmacher (1968), p.154) that a Jordan cage associated with eigenvalue λ_i when raising it to the power μ can be written in the next form:

if $\mu > n_{ij}$

$$A_{ij}^{\mu} = \begin{bmatrix} \lambda_i^{\mu} & \mu\lambda_i^{\mu-1} & (\mu-1)\lambda_i^{\mu-2} & \cdot & \cdot & \cdot & (\mu-n_{ij}+2)\lambda_i^{\mu-n_{ij}+1} \\ 0 & \lambda_i^{\mu} & \mu\lambda_i^{\mu-1} & \cdot & \cdot & \cdot & (\mu-n_{ij}+3)\lambda_i^{\mu-n_{ij}+2} \\ 0 & 0 & \lambda_i^{\mu} & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & \cdot & \cdot & \cdot & \cdot & \cdot & \lambda_i^{\mu} \end{bmatrix} \quad (3.54a)$$

and if $\mu < n_{ij}$ there appears at least one 1 or in general:

$$A_{ij}^{\mu} = \begin{bmatrix} \lambda_i^{\mu} & \mu\lambda_i^{\mu-1} & (\mu-1)\lambda_i^{\mu-2} & \cdot & 2\lambda_i & 1 & \cdot & \cdot & \cdot & 0 \\ 0 & \lambda_i^{\mu} & \mu\lambda_i^{\mu-1} & \cdot & 0 & 2\lambda_i & 1 & 0 & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \lambda_i^{\mu} \end{bmatrix} \quad (3.54b)$$

In order to find out under which condition the system $\{A^{\mu},B,C\}$ is completely observable it has to be investigated which properties of the matrix A are used in the criterion of observability in section 3.2.3.

If the criterion is also applicable to A-matrices with a structure as in equation (3.54) then no situations of nonobservability of $\{A^{\mu},B,C\}$ will occur.

For assertions 1. and 3. (section 3.2.3.) no restrictions on the matrix A are made. They hold for any matrix A, and therefore they are also applicable to matrix A^{μ}.

Assertion 2. (section 3.2.3.) assumes that for all i :

$$(A_i - \lambda_i I)^{\overline{n_i}} = \emptyset \qquad (3.55)$$

$$\text{with } \overline{n_i} = \max_j n_{ij}$$

This assertion not only holds for a Jordan cage ,but for any right upper matrix with equal diagonal elements λ_i and dimension less than $\overline{n_i}$.

Because A^{μ}_{ij} fulfills this condition, assertion 2. will also remain valid.
In the proof of the theorem itself (section 3.2.4), apart from the three assertions, only use has been made of the fact that matrix $\Gamma(A_i-\lambda_i I, C_i, \overline{n_i})$ could be written as:

$$\begin{bmatrix}
\underline{c}_{1k1} & \underline{x} & \underline{x} & \cdot & \underline{x} & \underline{c}_{1k2} & \underline{x} & \cdot & \underline{x} & \cdot \\
\underline{0} & \underline{c}_{1k1} & \underline{x} & \cdot & \underline{x} & \underline{0} & \underline{c}_{1k2} & \cdot & \underline{x} & \cdot \\
\underline{0} & \underline{0} & \underline{c}_{1k1} & \cdot & \underline{x} & \underline{0} & \underline{0} & \cdot & \underline{x} & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot & \underline{0} & \cdot & \cdot & \underline{c}_{1k2} & \cdot \\
\underline{0} & \cdot & \cdot & \cdot & \underline{c}_{1k1} & \underline{0} & \cdot & \cdot & \underline{0} & \cdot
\end{bmatrix}$$

where $\underline{x}$ denotes any column vector, which is irrelevant for the proof. In case we deal with cages in the form of (3.54) we get:

$$\begin{bmatrix}
\underline{c}_{1k1} & \underline{x} & \underline{x} & \cdot & \underline{x} & \underline{c}_{1k2} & \underline{x} & \cdot & \underline{x} & \cdot \\
\underline{0} & 2\lambda_i\underline{c}_{1k1} & \underline{x} & \cdot & \cdot & \underline{0} & 2\lambda_i\underline{c}_{1k2} & \cdot & \cdot & \cdot \\
\underline{0} & \underline{0} & 2\lambda_i\underline{c}_{1k1} & \cdot & \cdot & \underline{0} & \underline{0} & 2\lambda_i\underline{c}_{1k2} & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot & \underline{0} & \cdot & \cdot & 2\lambda_i\underline{c}_{1k2} & \cdot \\
\underline{0} & \cdot & \cdot & \cdot & 2\lambda_i\underline{c}_{1k1} & \underline{0} & \cdot & \cdot & \underline{0} & \cdot
\end{bmatrix}$$

and now also $\Gamma(A_i-\lambda_i I, C_i, \overline{n_i})\cdot\underline{q} = \underline{0}$ will always give $\underline{q}=\underline{0}$ unless $\lambda_i=0$.

The situation where $\lambda_i = 0$ gives rise to an exception, which we will discuss in the next section.
We may say now that the criterion for complete observability can also be applied to the system $\{A^{\mu},B,C\}$, that's to say, the system $\{A^{\mu},B,C\}$ is completely observable (controllable) if the system $\{A,B,C\}$ is completely observable (controllable), apart from two exceptions, which <u>may</u> disturb this:

1. The Jordan structure has been changed because $\lambda_i^\mu = \lambda_j^\mu$ for some $i \neq j$

2. $\lambda_i = 0$ for some i

In section 3.2.2. we have seen that the Jordan structure of the matrix A is completely defined by all numbers n_{ij} $i=1,\nu$ $j=1,r(i)$

From the remarks in this section the conclusion can be drawn that the complete observability of $\{A^\mu,B,C\}$ for sure is preserved if all numbers n_{ij} of A and A^μ are the same. In this situation applying the criterion of observability to A,C and A^μ,C will lead to exactly the same results.
Problems may arise when the Jordan structure of A and A^μ are not the same. This will be dealt with in the next section.

3.4.2. Situations of nonobservability

The Jordan cage structure of A^μ and A will not be the same in the following situations

a) $\lambda_i^\mu = \lambda_j^\mu$ for some $i \neq j$, while $\lambda_i \neq \lambda_j$ (3.56)

b) $\lambda_i = 0$ for some i (3.57)

Ad a) In this case two originally different eigenvalues of A will be transformed into two equal eigenvalues of A^μ. This means that the number of distinct eigenvalues ν in A is decreased for A^μ by one to $\nu-1$, and that two Jordan blocks are linked up into one block.
For complete observability of the system the set of r(i) q-dimensional column vectors $\underline{c}_{1i1},\underline{c}_{1i2},\dots\dots,\underline{c}_{1ir(i)}$ has to be a linearly independent set. When two Jordan blocks are linked up, r(i) increases and the condition of independence of the relevant column vectors of C has to be investigated again.
The independence of the new set of r(i) vectors is the only criterion for observability of $\{A^\mu,B,C\}$.

Ad b) From $\lambda_i=0$ it follows that $\lambda_i^k = 0$ for all k.
In case μ is sufficiently large ($\mu > n_{ij}$) we are dealing with the situation as in equation (3.45a). This leads to a matrix A^μ that is completely filled with zero's: $A^\mu = \emptyset$.

What originally were r(i) Jordan cages ,each with dimensions n_{ij}, j=1,r(i), now become $n_i = \sum_{j=1}^{r(i)} n_{ij}$ Jordan cages of length 1.

In other words all zero eigenvalues of A become noncommon zero eigenvalues of A^μ.

For observability the criterion of independence of the set of r(i) vectors of C becomes a criterion of independence of n_i column vectors in case of A^μ.

This new criterion has to be tested again. It is clear that the criterion in the latter case can only be fulfilled if $q > n_i$.

It should be noted that if $\mu > n$ is taken as an assumption the condition $\mu > n_{ij}$ always will be fulfilled.

In case μ would be chosen smaller than n_{ij} the matrix A^μ_{ij} becomes:

$$A^\mu_{ij} = \begin{bmatrix} 0 & 0 & \cdot & \cdot & 0 & 1 & 0 & \cdot & \cdot & \cdot & 0 \\ 0 & 0 & \cdot & \cdot & 0 & 0 & 1 & 0 & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 1 & 0 \\ 0 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 0 & 1 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 0 \end{bmatrix} \begin{array}{l} \\ \\ \\ \\ \\ \left.\begin{array}{l} \\ \\ \\ \end{array}\right\} w \end{array}$$

Here the latter w rows indicate the w quasi noncommon poles we have got in the system A^μ. This gives rise to a different Jordan form again, where we get w cages in stead of originally one. This can be accomplished by a suitable transformation matrix T, which just interchanges some states. Once this has been done, the criterion of Chen and Desoer may be applied again, which extends the relevant set with the columns of C corresponding to the indicated w states.

Example: $n_{ij} = 4$

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \qquad \mu = 2 \rightarrow A^\mu = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \qquad w = 2$$

interchange state 2 and 3:

$$T = T^{-1} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \rightarrow \quad A^{\mu} := \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

In both situations a) and b) the complete observability of $\{A^{\mu},B,C\}$ is determined by the (in)dependence of some specific column vectors of C.

3.4.3. Replacement of poles in the z-domain

In this section we want to make clear in which situations the problem of nonobservability of $\{A^{\mu},B,C\}$ arises.
As mentioned in the previous section two situatiions can be distinguished:

a) $\lambda_i^{\mu} = \lambda_j^{\mu}$,while $\lambda_i \neq \lambda_j$ for $i \neq j$

b) $\lambda_i = 0$ for some i

Situation b) is very clear: a single pole in z = 0 will cause no problems because the Jordan structure of A^{μ} remains the same as the one for A.
Two or more common poles in z = 0 may cause nonobservability depending on C ,because they are transformed top noncommon poles for A^{μ} ; these noncommon poles can be non-distinguishable.

Situation a) may occur e.g. when $\lambda_i = -\lambda_j$. Then for all even μ holds:
$\lambda_i^{\mu} = (-\lambda_j)^{\mu} = \lambda_j^{\mu} > 0$ as shown in Fig-3.4.1-

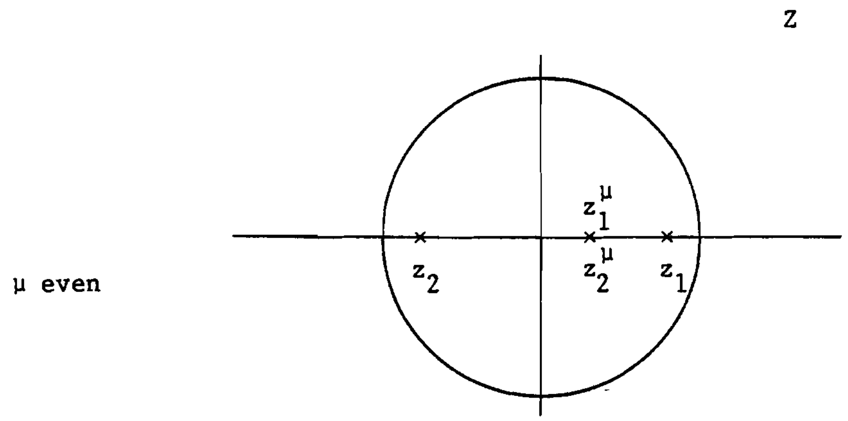


Fig-3.4.1- Poles in A and A^μ when $\lambda_i = -\lambda_j$ and μ even

A second possibility is that λ_i and λ_j are a complex conjugated pair $z_{1,2} = re^{\pm j\phi}$ and $\mu\phi = k\pi$,k ε **Z** (see Fig-3.2.).

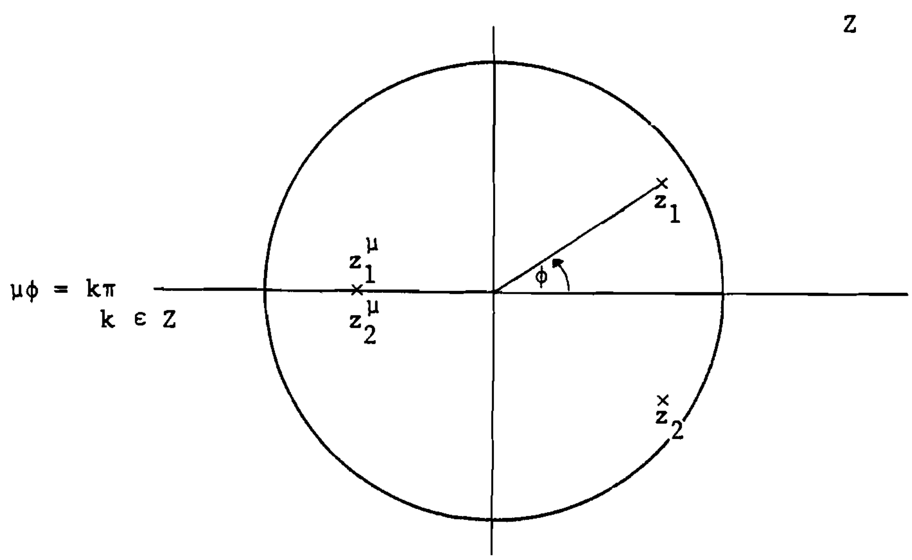


Fig-3.4.2- Poles in A and A^μ when λ_i and λ_j complex conjugate and $\mu\phi=k\pi$

And evidently all combinations of the two examples may occur. Notice that the final conclusion on observability of $\{A^\mu,B,C\}$ can only be drawn when knowing C. The described situations are the exclusive possibilities in which nonobservability might happen. Besides, situation a) can sometimes be avoided if all poles are in the right half plane (high sampling rate) and by choosing μ small enough. Situation b) may be eliminated by excluding delays and finite responses.

3.5. REMARKS

For purpose of clearity the results of this chapter are briefly stated again:

Given an observable system $\{A,B,C\}$.

The system $\{A^\mu,B,C\}$ is completely observable, if condition F does not hold:

F 1)for some i: $\lambda_i = 0$ and $r(i) > 1$ or
for some $i \# j$: $\lambda_i^\mu = \lambda_j^\mu$

and 2) for each i the set of r(i) q-dimensional column vectors $\underline{c}_{1i1}, \underline{c}_{1i2}, \ldots, \underline{c}_{1ir(i)}$ of the rearranged Jordan form for A^μ ,which are the columns of C corresponding to the independent states, is a linearly dependent set.

assuming that λ_i and λ_j are distinct eigenvalues of A.

If $\{A^\mu,B,C\}$ is completely observable then the Ho-Kalman algorithm applied to the Page matrix will lead to a minimum realization.

It should be noted that condition F, as stated above ,will generally not hold in the noisy case, because it lays quite heavy restrictions on the positions of the poles of A.

In general one can state that the situations in which the Page matrix algorithm will not give a minimum realization appear seldom. Nevertheless in case $\lambda_i^\mu \approx \lambda_j^\mu$ and/or $\lambda_i \approx 0$, this might be a cause that the resulting solution is ill-conditioned.

A dual situation arises when the Chinese page matrix is subject of study. In stead of the observability of the system, the controllability now is the critical feature. Then in stead of the matrix C matrix B is more essential. This dual approach can be a good alternative if $\{A^\mu,B,C\}$ is an nonobservable but controllable system. The Ho-Kalman algorithm applied to the Chinese page matrix then will lead to a minimum realization.

CHAPTER 4: CONSIDERATIONS WITH RESPECT TO THE TWO APPROACHES

4.1 THEORETICAL COMPARISON OF THE PRESENTED METHODS

4.1.1 Deterministic situation

In general a system with a finite dimension has an impulse response of infinite length. The characteristic feature of the deterministic situation is that from a given series of Markov parameters of finite length, the infinite sequence can be generated. In other words: complete information about the system is present in a finite number of Markov parameters. This is a direct result of theorem 1 in section 1.1.2.
The complete information is given by $\{M_k\}$,k=1,..r and $\{a_i\}$, i=1,..r, where $-a_i$ is the i^{th} coefficient of the minimal polynomial, and r is the realizability index. With these two arrays it is possible to generate the infinite sequence of Markov parameters, a unique representation of the considered system.
This approach can be recognized in both the Hankel matrix and Page matrix algorithms: when for both matrices the methods are applied using a shifted H or P, the total number of Markov parameters that is used equals 2r. The information in this sequence is sufficient for obtaining the information mentioned above: when given $\{M_k\}$,k=1,..r, the coefficients a_i,i=1,..r can be derived from the Markov parameters $\{M_k\}$,k=r+1,..2r as a result of a set of r.p.q linear equations with r unknowns.
If the rank of a Hankel matrix fulfills the condition:

$$\text{rank } H_{r+N} = \text{rank } H_r \quad \text{for all } N>0$$

then the structure of the Hankel matrix implicitly guarantees the existence of a finite dimensional realization. If in the Hankel matrix the addition of an extra block row does not change its rank, the block row can be written as a linear combination of previous ones, corresponding to eq.(1.4). A finite dimensional realization then will always exist (compare eq.(1.7)), and the r.p.q equations with r unknowns, as mentioned above, will be non conflicting.

In the Page matrix this block structure is not available. The rank condition for the Page matrix now does not automatically lead to a linear relation between the Markov parameters as in eq.(1.4). It is proven in chapter 3 that if a finite dimensional realization exists, it can be found by using a Page matrix algorithm, apart from some exceptional cases. However the rank condition alone is not a sufficient condition for the existence of such a realization.

Apparently for Page matrix algorithms, this lacking condition has been exchanged for a remarkable reduction in computation, if an equal number of Markov parameters is taken into account. If L Markov parameters are available, the size of the Page matrix will be $L.p.q$ and the size of the Hankel matrix $p.q.L^2/4$. Consequently the reduction of the computational effort is considerable.

For the method of Kung comparable remarks can be made with respect to the attainment of the required information. In this method no shifted version of H or P is used. To compensate this, the rank condition as stated in chapter 1 has to hold also for the observability or controllability matrix of reduced dimensions (see eq.-(1.23)). This means that the Markov parameter M_{2r} has to be incorporated before in the original Hankel or Page matrix. Therefore in these algorithms the same amount of information is used as in the situation above.

4.1.2 Noisy situation

When noise is added to the Markov parameters, the infinite sequence will never have a finite dimensional realization. With the results of the previous section and chapters 1 and 2 we now can sketch the following diagram:

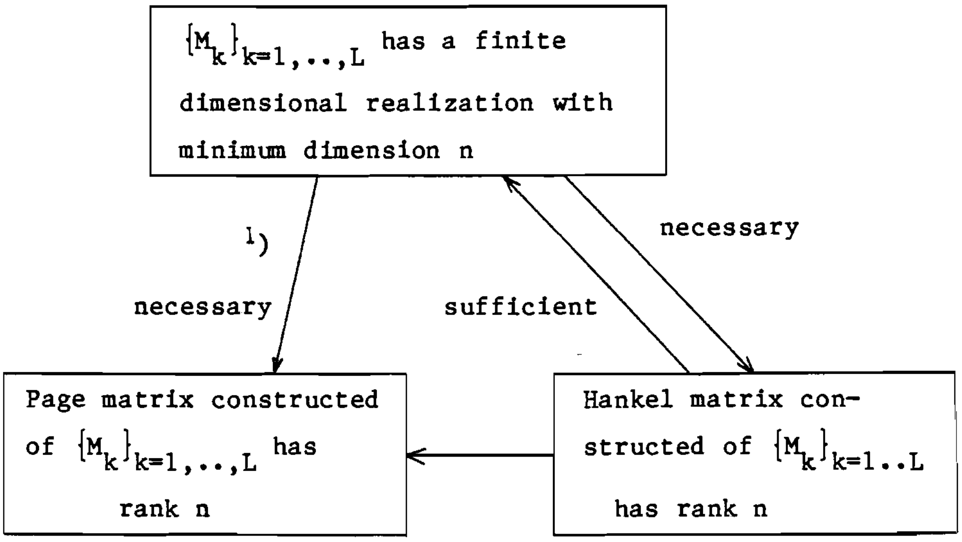


Fig-4.1.1- Necessary and sufficient conditions for the existence of an n-dimensional realization.

1) The conditions for the validity of this statement are given in chapter 3.

We will now point our attention to the consequence of the relations in the above diagram for the application of the different algorithms. As was mentioned in section 1.4 the singular value decomposition has been used to obtain a Page or Hankel matrix of reduced rank n, by creating an array of modified,or filtered Markov parameters. In fact this singular value decomposition fulfills a number of tasks: a decomposition of P or H into two full rank matrices, a tool for determining the dimension of the realization, and moreover a least squares approximation of P or H. This last feature will work out differently for the algorithms presented in chapter 1.
For the Ho-Kalman algorithm the Markov parameters that constitute the Hankel matrix are weighted by an isosceles triangular function. This is due to the fact that, depending on their index, the Markov parameters appear more frequently in the Hankel matrix. Nevertheless this noise adapted Ho-Kalman algorithm weights all available Markov parameters at least with a nonzero weighting factor. If we were to use the algorithm suggeted by Silverman (see sect.1.2.2) we would select a full rank submatrix from the Hankel

matrix which is just big enough to calculate a realization. Because a strictly limited part of the Markov sequence is used, this is quite inappropriate for the noisy cases (no redundancy): the basis, thus obtained for the realization, may be quite ill conditioned numerically.
For the algorithms based on Kung's method, a shifted Hankel and Page matrix is not required. Possibly this is favourable in the noisy case, as here the last Markov parameter is also taken into account during the singular value decomposition of H or P.
For the Page matrix all Markov parameters are weighted once; a least squares fit on P is directly a least squares fit on the complete sequence of Markov parameters.

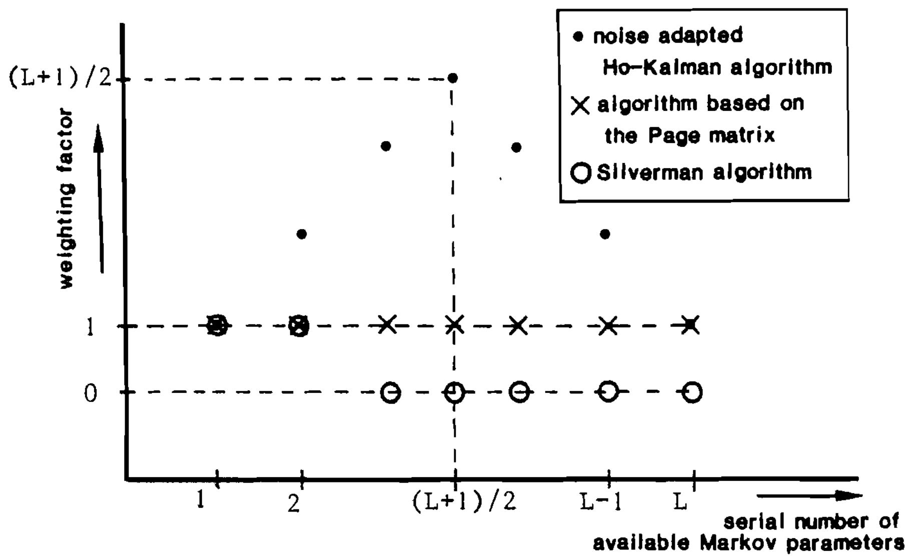


Fig-4.1.2- Weighting of Markov parameters in different algorithms

These statements and their presentations in fig-4.1.2- show that the Ho-Kalman and Silverman approach introduce non optimal approximations with respect to this rank reduction, or noise filtering of H, as far as the weighting is concerned.

A second non optimality of the Hankel approach is given by the fact that the reduced rank matrix H_k, constructed by setting a number of smallest singular values to zero, is no real but an approximating Hankel matrix. This means that formally the existence of an n-dimensional realization of the noise filtered Markov parameters in H_k is not guaranteed. Moreover a unique sequence of Markov parameters can not be obtained from H_k.

As a final remark on these methods we note that statistical considerations are difficult to make, as the noise on the entries in the Hankel matrix is not independent, but exactly the same noise data appear frequently in several entries.

Whether the Page matrix approach, with its lacking sufficient condition for the rank, or the Hankel matrix approach, with its non optimal noise filtering and lacking Hankel structure of H_k, will give better results in practical situations has to be tested in simulations.
This will be the subject of the next chapter.

4.2 SOME REMARKS WITH RESPECT TO THE NUMBER OF MARKOV PARAMETERS

In the previous chapters several remarks have been made concerning the dimensions of the Page and Hankel matrices. Because of the structure of both matrices the dimensions and the number of Markov parameters are directly related to eachother.
From the deterministic situation some minimal dimensions of P and H can be considered:

$$\text{for H:}\quad \begin{array}{llll} 1. & \gamma \geq \alpha, & \zeta \geq \beta & \text{leading to } L \geq \alpha+\beta-1 \\ 2. & \gamma \geq n-q+1, & \zeta \geq n-p+1 & \text{leading to } L \geq 2n-p-q+1 \\ 3. & \gamma \geq n, & \zeta \geq n & \text{leading to } L \geq 2n-1 \end{array} \qquad (4.1)$$

$$\text{for P:}\quad \begin{array}{llll} 1. & \eta \geq \alpha, & \mu \geq \beta & \text{leading to } L \geq \alpha.\beta \\ 2. & \eta \geq n-q+1, & \mu \geq n-p+1 & \text{leading to } L \geq (n-q+1)(n-p+1) \\ 3. & \eta \geq n, & \mu \geq n & \text{leading to } L \geq n^2 \end{array} \qquad (4.2)$$

Each of the bounds 1., 2. and 3. is equivalent for H and P. All three are stated from a specific approach.
Bound 1. is a theoretical required minimum level. This corresponds directly with the definitions of α and β. Bound 3. is a sufficient condition and quite a broad margin. It is an often used boundary because it is directly related to the dimension of the system, which is a variable that is more suitable than the, often unknown, observability and controllability indices.

Bound 2. is an intermediate approach. For this condition, that also can be found in Silverman (1971), it is assumed that the set of inputs as well as the set of outputs of the system is internally and statically independent. It is based on a nice selection of the basis in the state space. An upper bound for α and β can be given then.

It can easily be seen that the minimum number of Markov parameters required in the Hankel matrix is smaller than that required in the Page matrix.

For the limits mentioned above it is assumed that Kung's algorithm is not applied. For this algorithm a stronger condition on the dimensions of P and H has to be fulfilled, in order to guarantee the correct rank of a shifted observability or controllability matrix. When the controllability matrix Δ is used to determine the matrix A, the conditions (4.1) and (4.2) have to be fulfilled for ζ-1 and μ-1.

This leads to the following conditions <u>in case of Kung's algorithm</u>:

for H: 1. $\zeta > \beta+1$ leading to $L > \alpha+\beta$
2. $\zeta > n-p+2$ leading to $L > 2n-p-q+2$ (4.3)
3. $\zeta > n+1$ leading to $L > 2n$

for P: 1. $\zeta > \beta+1$ leading to $L > \alpha(\beta+1)$
2. $\zeta > n-p+2$ leading to $L > (n-q+1)(n-p+2)$ (4.4)
3. $\zeta > n+1$ leading to $L > n^2+n$

Of course, in the noisy case, at least these same conditions have to be fulfilled. However, it is only possible to find the conditions if the deterministic system, or its dimensions are known. In the noisy situation the linear dependency between the Markov parameters over a finite time range is lost. Therefore a long sequence of parameters is required, compared with the above mentioned boundaries, in order to obtain enough redundancy to eliminate the influence of the noise to a reduced level.

If L increases the total amount of signal energy in both a Hankel and a Page matrix will increase. Also the total noise energy in these matrices will increase, because of the addition of new noise terms. The noise has been chosen to be of an absolute character: not related to some value of Markov parameters, but fixed with a

constant variance. Because we consider stable systems, the relative disturbance of Markov parameters $\{M_k\}$ by noise will become heavier with increasing k. From a certain value of k, taking an extra Markov parameter M_{k+1} into account will increase the noise energy in the Hankel or Page matrix more than the signal energy. This will probably have its harmful influences on the accuracy of the realization results.
At all events this will be disadvantageous for the distribution of singular values. Because of the increasing energy, the level of singular values will grow, and the relative gap between $\tilde{\delta}_n$ and $\tilde{\delta}_{n+1}$ will become smaller.
Following this argument it should be possible to find an optimal value of L depending, however, on the specific system chosen. Systems with a slowly decreasing impulse response would have a relatively high L_{opt} and systems with a fastly decreasing M_k a relatively low L_{opt}.
A detailed analysis of the possible appearance of an optimal value of L and its behaviour has not been worked out yet.

4.3 NOTES ON THE NOISE IN THE HANKEL MATRIX

In this section a short analysis will be given of the noise in the Hankel matrix, corresponding to section 2.4. where this has been done for the Page matrix.
Consider a noise disturbed Hankel matrix with dimensions $g \times l = \gamma . q \times \zeta . p$, where γ, ζ are the block dimensions, and consider a matrix Ξ_H with the same dimensions, containing all noise samples on H. The noise is again assumed to be SWAYING noise.
Because we only consider systems with p>q, in many cases g will be the smallest dimension of H, and therefore also the rank of Ξ_H.

$$\Xi_H \Xi_H^T = \begin{bmatrix} N_1 & \cdot & \cdot & N_\zeta \\ \cdot & & & \cdot \\ \cdot & & & \cdot \\ N_\gamma & \cdot & \cdot & N_{\gamma+\zeta-1} \end{bmatrix} \cdot \begin{bmatrix} N_1^T & \cdot & \cdot & N_\gamma^T \\ \cdot & & & \cdot \\ N_\zeta^T & \cdot & \cdot & N_{\gamma+\zeta-1}^T \end{bmatrix} \qquad (4.5)$$

where N_i is a q×p-matrix of independent noise samples on the Markov parameters.

$$\Xi_H \Xi_H^T = \begin{bmatrix} \sum_{i=1}^{\zeta} N_i N_i^T & \cdot & \cdot & \sum_{i=1}^{\zeta} N_i N_{i+\gamma-1}^T \\ \vdots & & & \vdots \\ \sum_{i=1}^{\zeta} N_{i+\gamma-1} N_i^T & \cdot & \cdot & \sum_{i=\gamma}^{\zeta} N_{i+\gamma-1} N_{i+\gamma-1}^T \end{bmatrix} \qquad (4.6)$$

Because of the character of the noise:

$E\{\Xi_H \Xi_H^T\}$ is a diagonal matrix, containing ζ diagonal submatrices that can be written as

$$E\ \{\sum_{i=1}^{\zeta} N_i N_i^T\} = \zeta . p . \sigma^2 . I_q \qquad (4.7)$$

As a result of this,

$E\{\Xi_H \Xi_H^T\}$ is a diagonal matrix that can be written as:

$$E\{\Xi_H \Xi_H^T\} = 1 . \sigma^2 . I_g \qquad (4.8)$$

Because we consider 1 to be max(g,1), this result is similar to the result in case of the Page matrix. However in this case the diagonal elements of $E\{\Xi_H \Xi_H^T\}$, are not independent of eachother. In eq.(4.6) it can be seen that on the diagonal of $\Xi_H \Xi_H^T$ terms appear at the same time in different positions.

CHAPTER 5: RESULTS OF SIMULATIONS

5.1 INTRODUCTION

Statements have been made as a result of theoretical analysis of the problem of approximate realization. In order to test these statements simulations have to be done. In simulations all external influences on the (simulated) process can be controlled and all boundary conditions can be chosen in such a way that the result of the simulations is well suited for drawing conclusions on the properties of the theory.
In our situation one of the external factors is the noise corruption of the series of Markov parameters. As was indicated in the introduction of this report, the noise contribution on these Markov parameters is chosen to be of a special character (SWAYING noise, see Introduction).

In the simulations described in this chapter different algorithms will be tested. A theoretical description of these algorithms can be found in chapters 1 and 2. The algorithms will be applied to different systems. MIMO systems as well as SISO systems will be used as test-systems.

Two programs have been written to execute the simulations described in this chapter: MARK2 and STATEX. The first one creates a series of Markov parameters of a chosen system and disturbs this series, on demand, with SWAYING noise. The second one calculates an approximate realization of this series of Markov parameters by applying to it one of the available algorithms.
These two programs are described in detail in a separate report: "User's manual for Hankel and Page matrix approximate realization programs".

In section 5.2 a detailed description of the simulations and their processing will be given, and all choices with respect to these simulations will be elucidated. Next the results of simulations will be reported: in the deterministic case (section 5.3), in the noisy MIMO case (section 5.4) and in the noisy SISO case (section 5.5). Finally in section 5.6 some concluding remarks are stated.

5.2 DESCRIPTION OF SIMULATIONS AND PROCESSING

5.2.1 Block diagram, chosen algorithms and systems

The algorithms applied to the simulations are based on the methods described in chapters 1 and 2. In order to clarify the chosen abbreviations of the algorithms and to introduce the chosen methods a list of the algorithms is given next:

method nr.	abbrev.	explanation
1	REHA	Realization based on the Hankel matrix; Ho-Kalman algorithm modified by Damen/ Hajdasinski
2	REHAK	Realization based on the Hankel matrix; method according to Kung.
3	REPAH	Realization based on the Page matrix; Ho-Kalman algorithm modified by Damen/ Hajdasinski.
4	REPAK	Realization based on the Page matrix; method according to Kung.
5	REPHH	Realization based on a combination of Page and Hankel matrix. Before applying the Hankel realization algorithm an extra noise filtering step with the Page matrix is performed. The method is chosen according to Damen/Hajdasinski.

This fifth algorithm is applied in order to be able to test the combination of Page noise filtering and Hankel realization. In this method first a Page matrix is constructed and the rank of this matrix is reduced. With the resulting Page matrix of reduced rank a series of noise filtered Markov parameters is defined. With these Markov parameters a Hankel matrix is constructed and the original Hankel realization algorithm is applied, an extra rank reduction step included.

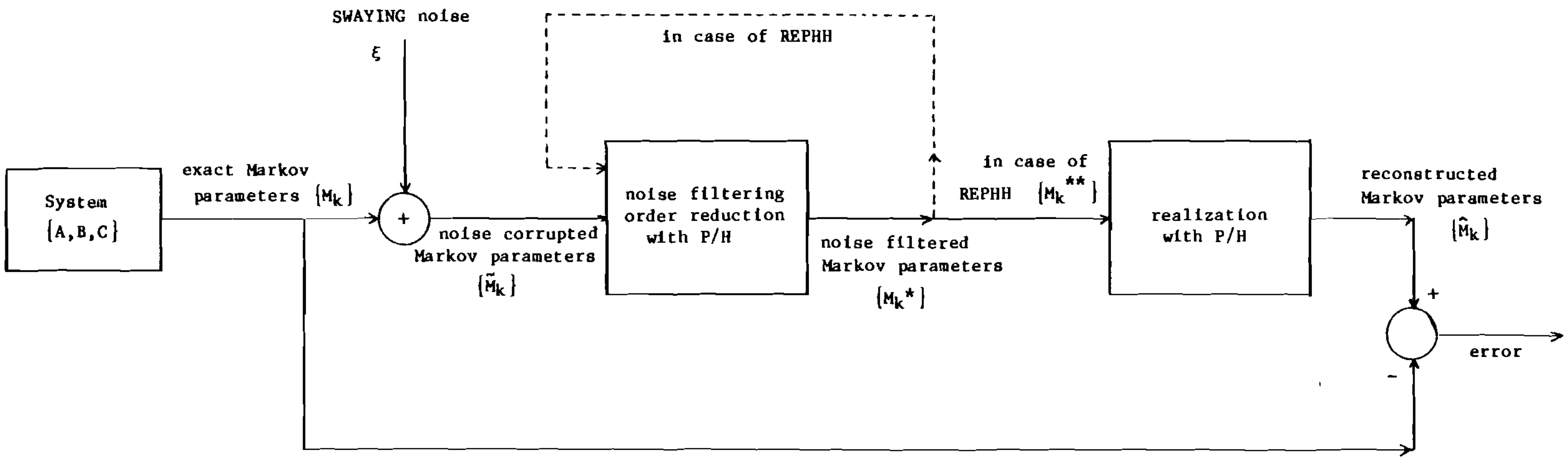


Fig-5.2.1- Block diagram of the simulation and processing

In the sequel of this chapter the abbreviations given above will be used to indicate the applied algorithms.
A schematic representation of the simulations is given in fig-5.2.1-. The result of the noise filtering and order reduction step in the given block diagram is a series of noise filtered Markov parameters. In case of a Page matrix algorithm such a series is uniquely defined. In case of a Hankel algorithm however this is not the case. To be able yet to draw conclusions upon the properties of the noise filtering of the Hankel matrix algorithms, a series of noise filtered Markov parameters is constructed. This construction is performed by averaging the available elements in the Hankel matrix corresponding with the same Hankel position.

The exact Markov parameters are generated for one of the systems chosen for simulation, according to the formula

$$M_k = C\,A^{k-1}\,B \qquad k \geq 1 \tag{5.1}$$

4 MIMO systems and 2 SISO systems have been chosen to test the realization algorithms. These chosen systems will be explained next.

<u>MIMO systems</u>

The B and C matrices of the four MIMO systems are chosen the same for all systems, corresponding to Damen and Hajdasinski (1982). All four systems are 3 input, 2 output systems with dimension 4.

$$B = \begin{bmatrix} 1. & 0. & 1. \\ -1. & .5 & .5 \\ 0. & 1. & -.5 \\ 0. & .5 & -1. \end{bmatrix} \qquad C = \begin{bmatrix} .5 & 0. & 1. & 1. \\ 1. & -.5 & .5 & -.5 \end{bmatrix} \tag{5.2}$$

As mentioned in the introduction of this report, the matrix D is chosen to be $\emptyset$.
To be able to have a clear view on the eigenvalues of the respective systems, the A matrices are chosen in a diagonal way:

$$\begin{aligned} &\text{SYS0: } A_0 = \text{diag}(0.4,0.3,0.2,0.1) \\ &\text{SYS1: } A_1 = \text{diag}(0.6,0.5,0.9,0.7) \\ &\text{SYS2: } A_2 = \text{diag}(0.9,0.8,0.85,0.3) \\ &\text{SYS3: } A_3 = \text{diag}(0.2,0.1,0.9,0.15) \end{aligned} \tag{5.3}$$

In every system a specific aspect of the capability of the realization algorithms is tested: a relatively fast decreasing impulse response (SYS0), slow decreasing impulse response (SYS1), and 3 closely positioned eigenvalues large (SYS2) or small (SYS3). SYS2 and SYS3 are specially chosen to test the properties of the different realization algorithms in recognizing these closely positioned eigenvalues when the system is noise disturbed.

SISO systems

Two SISO systems are used, SYS4 and SYS5; SYS4 having dimension 1, SYS5 dimension 2. The matrices:

$$\begin{aligned} &\text{SYS4: } A_4 = 0.9 && B_4 = 1. && C_4 = 0.5 \\ &\text{SYS5: } A_5 = \mathrm{diag}(0.9,0.8) && B_5 = \begin{bmatrix} 1. \\ 0.5 \end{bmatrix} && C_5 = \begin{bmatrix} 0.5 & 1. \end{bmatrix} \end{aligned} \tag{5.4}$$

These SISO systems are incorporated in order to test a first order situation. Because we consider systems with linear statically independent outputs, the number of outputs of a system should not exceed the dimension of the system. Therefore for a first order system a single output system is required.
Another reason for using SISO systems is the possibility to test situations where a large number of Markov parameters is considered in the approximation procedure. Because of a restriction on the available memory space of the PDP 11/60 computer used, it is a tremendous job in the MIMO situation to increase this number of Markov parameters to a value higher than 24.

5.2.2 Choices of external quantities

In this section more detailed information on the executed simulations will be presented. Names of variables used in the rest of this chapter will be explained.

NOM is the number of Markov parameters used to construct the Hankel or Page matrix. It corresponds to L in the previous chapters. In the algorithms based on Kung's method this NOM also is the

totally required number of Markov parameters for the realization. In the other algorithms one extra Markov parameter is required for creating a shifted Hankel or Page matrix.
In the noisy MIMO situations the values of NOM are chosen: 8,12, 16,20,24. In the noisy SISO situation NOM=64.
If we apply the conditions of the dimensions of H and P from eq.-(4.1) - (4.4) and consider approach 2) we can conclude that the minimum number of Markov parameters is given by the values in the next table:

	Minimum NOM required	Mimimum total number of Markov parameters required
REHA	4	5
REHAK	5	5
REPAH	6	7
REPAK	9	9
REPHH	6	7

Fig-5.2.2- Minimum number of Markov parameters for MIMO systems

This means that for NOM=8 care has to be taken with respect to the algorithm REPAK. The required rank condition is not fulfilled in this situation.
In the deterministic MIMO case also experiments have been executed with Page and Hankel matrix of equal size: Page matrix NOM=16, Hankel matrix NOM=7.

In case of a block square Hankel matrix the number of Markov parameters in such a matrix is always odd (see eq.(1.1)). With respect to the choices of NOM the use of a block square Hankel matrix would lead to the use of different numbers of Markov parameters in the Page and Hankel algorithms. For this reason one extra block row is added to the Hankel matrix, in such a way that it is possible to construct H with even NOM.
The dimensions of the Page matrices are chosen corresponding to the theoretical aspects mentioned in chapter 2: P is chosen as square as possible. In the MIMO situation this gives the next results:

NOM	PAGE		HANKEL	
	block dimensions	real dimensions	block dimensions	real dimensions
	$\eta\times\mu$	$h\times m$	$\gamma\times\zeta$	$g\times l$
8	4×2	8×6	5×4	10×12
12	4×3	8×9	7×6	14×18
16	4×4	8×12	9×8	18×24
20	5×4	10×12	11×10	22×30
24	6×4	12×12	13×12	26×36

Fig-5.2.3- Block and real dimensions of Page and Hankel matrix.

In the SISO situation with NOM=64, block dimensions and real dimensions are equal: P = (8×8).

The choice for the rank n of the realization, which has to be made during the rank reduction of P or H, is fixed to the known dimension of the system. At this stage of the project the result of choosing the dimension of the system not correctly will not be analysed. In the MIMO situation this means n=4, for SYS4 n=1 and for SYS5 n=2.

With respect to algorithms based on Kung's method, either the observability or the controllability matrix can be used for estimation of the matrix A. In all situations the controllability matrix is used for this purpose.

With respect to Page matrix algorithms both the Page and the Chinese Page matrix can be used for the approximation algorithm.
In all situations only the Page matrix is used, in order to limit the total number of experiments. The difference between these two approaches lies in the position of the "jump" in singular values (see also sect.2.3.2). In the Page matrix this "jump" appears after δ_q, in the Chinese Page matrix after δ_p. Because p>q, in the latter case more eigenvalues will reveale themselves in the first significant singular values before the "jump". Therefore the Chinese Page matrix may show better results.

For all MIMO systems experiments have been run with two different noise levels, apart from the deterministic case when no noise is added. These noise levels with variances $\sigma_1{}^2$ and $\sigma_2{}^2$ are chosen in such a way that one level (σ_1) can be considered as relatively low noise, and one level (σ_2) as relatively high noise.
The values σ_1 and σ_2 are chosen based on the singular values of a deterministic Page matrix, constructed with 16 Markov parameters. The squared sum of these singular values represents the energy content in 16 deterministic Markov parameters. According to statements in chapter 2 with respect to the noise influence on singular values (see eq.(2.35)), it can be stated that the expression

$$\delta_i^2(N) = \sigma_i^2 \cdot \max(h,m) \tag{5.5}$$

is a measure for the noise influence on deterministic squared singular values for different noise levels indicated by i. In addition to the existing gap in deterministic singular values of P between δ_2 and δ_3, the variances of the noise will be chosen in such a way that for $\sigma=\sigma_1$ the expression (5.5) will be in the region of $\delta_3{}^2$ and $\delta_4{}^2$ and for $\sigma=\sigma_2$ it will be in between the regions of $\delta_2{}^2$ and $\delta_3{}^2$.

	$\delta_1{}^2$	$\delta_2{}^2$	$\delta_3{}^2$	$\delta_4{}^2$	$\delta_1{}^2(N)$	$\delta_2{}^2(N)$	$\sigma_1{}^2$	$\sigma_2{}^2$
SYS0	3.80	3.55	$.3 \cdot 10^{-4}$	$.1 \cdot 10^{-5}$	$.3 \cdot 10^{-5}$	.75	$.3 \cdot 10^{-6}$	$.6 \cdot 10^{-1}$
SYS1	12.47	5.01	1.00	$.1 \cdot 10^{-2}$	$.8 \cdot 10^{-2}$	.75	$.6 \cdot 10^{-3}$	$.6 \cdot 10^{-1}$
SYS2	13.98	5.96	0.12	$.5 \cdot 10^{-1}$	$.3 \cdot 10^{-1}$	.75	$.3 \cdot 10^{-2}$	$.6 \cdot 10^{-1}$
SYS3	9.45	3.63	0.20	$<.3 \cdot 10^{-6}$	$.8 \cdot 10^{-2}$	.75	$.6 \cdot 10^{-3}$	$.6 \cdot 10^{-1}$

Fig-5.2.4- Variances of the chosen noise series related to deterministic singular values of P; NOM=16; MIMO systems.

In case of the SISO systems the same procedure is followed; NOM=64 is taken as the reference situation in the deterministic case. In case of SYS4 also two noise levels are chosen: one (σ_1) below the level of $\delta_1{}^2$ and one (σ_2) above this level. In case of SYS5, where n=2, three levels are chosen.

	δ_1^2	δ_2^2	$\delta_1^2(N)$	$\delta_2^2(N)$	$\delta_3^2(N)$	σ_1^2	σ_2^2	σ_3^2
SYS4	1.32	—	.98	2.	—	.12	.25	—
SYS5	3.80	$.1\cdot10^{-2}$	$.6\cdot10^{-3}$	.98	3.92	$.7\cdot10^{-4}$	.12	.49

Fig-5.2.5- Variances of the chosen noise series related to deterministic singular values of the Page matrix; NOM=64, SISO case.

In general one series of noise samples is used, generated by starting the available noise generator with starting values:

KX = 9740 KY = 2254

(see also description of programs in separate report).
For this series different values of σ can be chosen.
In some specific situations (SYS0,σ_2 and all SISO systems) three runs with three different noise series have been executed. The two additional runs have been generated by:

KX = 1926 KY = 2159 and

KX = 3228 KY = 1154

When averaged results over 3 noise series are presented, this is clearly marked.

5.2.3 The error measures

To evaluate the results of the simulations different measures will be defined to indicate the results of the most essential steps in the algorithms. To understand the output of the simulations these measures are defined next:

1. Relative error in input Markov parameter sequence:

$$\text{ERRO} = \frac{\sum_{k=1}^{NOM} \| \tilde{M}_k - M_k \|_E^2}{\sum_{k=1}^{NOM} \| M_k \|_E^2} \tag{5.6}$$

where $\{M_k\}_{k=1,..,NOM}$ is the deterministic array of Markov parameters.

2. Average error in input Markov parameter sequence:

$$\text{AVERO} = \frac{\sum_{k=1}^{N} \| \tilde{M}_k - M_k \|_E^2}{p \cdot q \cdot N} \tag{5.7}$$

This is an average error per element of each Markov parameter matrix, always averaged over a fixed number of Markov parameters. In the MIMO situation N=25; in the SISO situation N=65.

3. Relative error in noise filtered Markov parameters:

$$\text{ERRF} = \frac{\sum_{k=1}^{NOM} \| M_k^* - M_k \|_E^2}{\sum_{k=1}^{NOM} \| M_k \|_E^2} \tag{5.8}$$

$\{M_k^*\}_{k=1,..,NOM}$ is the array of Markov parameters appearing in the rank reduced Page matrix. In case of a Hankel algorithm it is an artificially created series by averaging corresponding elements in the Hankel matrix of reduced rank. In case of the combined Page/Hankel algorithm the Markov parameters are filtered twice: first by the Page matrix, secondly by the Hankel matrix.

4. The noise reduction factor:

$$\text{NRF} = \frac{\text{ERRF}}{\text{ERRO}} \tag{5.9}$$

This is a measure for the quality of the noise filter: it indicates which part of the original noise energy on the Markov parameters remains after noise filtering.

5. Relative error in the realization part of the algorithm:

$$\text{ERRB} = \frac{\sum_{k=1}^{NOM} \| \hat{M}_k - M_k^* \|_E^2}{\sum_{k=1}^{NOM} \| M_k \|_E^2} \tag{5.10}$$

where $\{\hat{M}_k\}_{k=1,..,NOM}$ is the array of reconstructed Markov parameters from the resulting realization $\{\hat{A},\hat{B},\hat{C}\}$. ERRB gives an indication for the exactness of the realization of noise filtered Markov parameters.

6. Overall relative error:

$$\mathrm{ERR1} = \frac{\sum_{k=1}^{NOM} \| \hat{M}_k - M_k \|_E^2}{\sum_{k=1}^{NOM} \| M_k \|_E^2} \qquad (5.11)$$

This is a measure for the ability of the algorithms to fit a series of Markov parameters to the series of deterministic Markov parameters.

7. Overall average error:

$$\mathrm{AVER1} = \frac{\sum_{k=1}^{N} \| \hat{M}_k - M_k \|_E^2}{p \cdot q \cdot N} \qquad (5.12)$$

Again in MIMO cases is chosen N=25, in SISO cases N=65.

5.3 RESULTS IN THE DETERMINISTIC CASE

5.3.1 Introduction

In the deterministic case a series of Markov parameters is generated by one of the systems chosen, and without any disturbance this series is supplied to the five algorithms as described in section 5.2.1.

A sufficient condition for the correct dimensions of the Hankel and Page matrix is that $\gamma,\mu,\eta > n$. Taking this as an assumption both matrices contain at least n^2 Markov block matrices. In all chosen MIMO systems n equals 4, and the minimum number of Markov block matrices equals 16 under the condition stated above. Because of the different structure of P and H however, a block dimension of 4×4 in the two situations will lead to a different number of Markov parameters that is contained in the matrix: in the Hankel matrix NOM=7 and in the Page matrix NOM=16. We can distinguish between the situations where we consider the dimensions of P and H to be the same, and the situation where we consider the number of Markov parameters to be the same. In the deterministic case above sufficient numbers (differing for P and H) seem appropriate as an

abundancy does not give extra information. In the noisy case the number of Markov parameters used is a direct measure for the information used, so this should be fixed. For the MIMO systems these two cases are treated resp. in 5.3.2 and 5.3.3.
For the SISO systems a fixed number of parameters is assumed. This is dealt with in section 5.3.4.

5.3.2 MIMO systems; Hankel and Page matrix of equal dimensions

For the next results the dimensions of P and H are chosen equally large. This means for both matrices block dimensions 4×4 and real dimensions 8×12. The Hankel matrix is constructed from 7 Markov parameters, the Page matrix from 16.
First the singular values of both matrices for the four MIMO systems are given:

Deterministic singular values

		δ_1	δ_2	δ_3	δ_4
SYS0	H	2.083	1.920	0.073	0.028
	P	1.949	1.883	0.005	0.001
SYS1	H	4.742	2.740	0.310	0.043
	P	3.531	2.239	0.310	0.029
SYS2	H	5.029	2.971	0.504	0.123
	P	3.739	2.441	0.345	0.219
SYS3	H	3.935	1.934	0.552	0.035
	P	3.074	1.906	0.447	0.000

Fig-5.3.1- Singular values of H and P for MIMO systems. H and P of equal size; deterministic situation; H: NOM=7, P: NOM=16. μ=4, n=4, p=3, q=2.

In order to use the decrease in singular values for determination of the dimension of the chosen realization, the best situation would be an equal distribution of the energy in the matrix over all 4 singular values.

Because of the stability of the chosen systems and the structure of P and H, the total energy given by:

$$S = \sum_{i=1}^{4} \delta_i^2$$

will be higher in the Hankel matrix than in the Page matrix. This is very clear in fig-5.3.1-.

Another remark that should be made is the relatively high gap between δ_2 and δ_3 and between δ_3 and δ_4. For the cause of this gap we will refer to the example given in section 2.3.2 concerning the singular value decomposition. The second block rows in H and P are shifted in time with respect to the first block rows, resp. 1 and μ time instants. Because we deal only with stable systems, the energy in these rows will become smaller. This not only happens with respect to the rows but also with respect to the columns in H and P. However in this case the shift is equal for H and P: 1 time instant. Therefore the high "jump" between the first and the second block in P causes the Page matrix to show a higher gap between δ_q and δ_{q+1} than in the Hankel matrix. The "jump" between the first and second block column is equal for H and P, and therefore H and P will roughly show the same size of gap between δ_p and δ_{p+1}.

Especially in the cases SYS0 and SYS3 where very small eigenvalues appear in A the smallest singular values of P and H become very small; δ_4 of SYS3 even smaller than $0.5 \cdot 10^{-3}$. This effect is the heaviest in this case because the influence of small eigenvalues has almost died out in the second block row of P. In general it can be stated that if the number of small eigenvalues exceeds q, only q eigenvalues will reveale themselves in the first q singular values of relevant weight; the rest will become very small. Note however that a definition of "small" eigenvalue has not been given, because this phenomenon is related to different properties of the system, like the other eigenvalues, the sampling time and the noise level.
In case of SYS1 and SYS2, where relatively high eigenvalues appear, the distribution of singular values is less worse for Page matrix situations. For SYS2 the distribution in case of P is even better than for H.

Overall average error AVER1

	REHA	REHAK	REPAH	REPAK	REPHH
SYS0	$.16\cdot10^{-13}$	$.10\cdot10^{-13}$	$.18\cdot10^{-13}$	$.12\cdot10^{-13}$	$.39\cdot10^{-13}$
SYS1	$.95\cdot10^{-13}$	$.21\cdot10^{-12}$	$.50\cdot10^{-11}$	$.13\cdot10^{-12}$	$.44\cdot10^{-12}$
SYS2	$.18\cdot10^{-12}$	$.12\cdot10^{-12}$	$.52\cdot10^{-13}$	$.59\cdot10^{-12}$	$.92\cdot10^{-13}$
SYS3	$.27\cdot10^{-13}$	$.22\cdot10^{-12}$	$.32\cdot10^{-12}$	$.21\cdot10^{-12}$	$.22\cdot10^{-12}$

Fig-5.3.2- Overall average error for 5 algorithms applied to the MIMO systems; H and P of equal size; deterministic situation.

Fig-5.3.2- shows the results of AVER1 for all 5 algorithms. Only in case of SYS2, of which we have mentioned that the distribution of singular values was more equable for P than for H, a Page matrix algorithm shows the best results. Although the number of Markov parameters in the Hankel matrix is smaller, these Hankel algorithms show a smaller error at the end than the Page algorithms for the other systems. Note however that all values are very small and in the same range.

The realized eigenvalues, i.e. the eigenvalues of the resulting realization, appear to be exact for all systems and for all algorithms.

5.3.3 MIMO systems; Hankel and Page matrix with equal number of Markov parameters (16)

For this test the Hankel and Page matrix were both constructed from 16 Markov parameters (NOM=16). The Page matrix therefore has the same dimensions as in the previous section ($\eta\times\mu=4\times4$, $h\times m=8\times12$), but the Hankel matrix now has block dimensions $(\gamma+1)\times\gamma=9\times8$ and real dimensions $g\times l=18\times24$.

Deterministic singular values:

	δ_1	δ_2	δ_3	δ_4
SYS0	2.084	1.920	0.073	0.028
SYS1	6.357	2.854	0.554	0.077
SYS2	7.112	3.411	0.678	0.337
SYS3	5.579	1.971	0.672	0.039

Fig-5.3.3- Singular values of H for MIMO systems; NOM=16; deterministic situation.

Because of the larger dimensions of the Hankel matrix the total energy has increased. For the systems with small eigenvalues (SYS0 and SYS3) the increase in singular values is small because there is only a very small amount of energy in the added Markov parameters M_8 - M_{16}. The total picture of distribution of energy over δ_1-δ_4 remains the same.

Overall average error AVER1

	REHA	REHAK	REPAH	REPAK	REPHH
SYS0	$.28 \cdot 10^{-13}$	$.13 \cdot 10^{-13}$	$.18 \cdot 10^{-13}$	$.12 \cdot 10^{-13}$	$.39 \cdot 10^{-13}$
SYS1	$.96 \cdot 10^{-13}$	$.10 \cdot 10^{-12}$	$.50 \cdot 10^{-11}$	$.13 \cdot 10^{-12}$	$.44 \cdot 10^{-12}$
SYS2	$.59 \cdot 10^{-12}$	$.49 \cdot 10^{-13}$	$.52 \cdot 10^{-13}$	$.59 \cdot 10^{-12}$	$.92 \cdot 10^{-13}$
SYS3	$.22 \cdot 10^{-12}$	$.41 \cdot 10^{-13}$	$.32 \cdot 10^{-12}$	$.21 \cdot 10^{-12}$	$.22 \cdot 10^{-12}$

Fig-5.3.4- Overall average error for 5 algorithms applied to MIMO systems; NOM=16; deterministic situation.

In general the changes in these results for Hankel matrix algorithms are small. The results for Page matrix algorithms are the same as presented in fig-5.3.2-. The general picture of this error measure remains the same when NOM is chosen equal for H and P. The fact that the Page matrix algorithms in this situation require fewer mathematical operations than the Hankel matrix algorithms to obtain a realization, does not cause the first algorithms to show smaller overall errors.
The eigenvalues found for the given situation are exact for all systems and algorithms, within the precision of the mantissa.

5.3.4 SISO systems

In the simulations of SISO systems always a fixed number of Markov parameters has been used: NOM=64.
For the dimensions of H and P this means: H = (33×32) and P = (8×8); block dimensions and real dimensions are the same in this situation.

Deterministic singular values:

		δ_1	δ_2
SYS4	H	2.629	
	P	1.147	
SYS5	H	3.899	0.119
	P	1.948	0.032

Fig-5.3.5- Singular values of H and P for SISO systems; NOM=64; deterministic situation.

The singular values of SYS5 show the same behaviour as elucidated in the previous section: because of p=q=1 the gap now appears between δ_1 and δ_2 and δ_2/δ_1 is higher in case of P.

Overall relative error:

	REHA	REHAK	REPAH	REPAK	REPHH
SYS4	$.49 \cdot 10^{-13}$	$.47 \cdot 10^{-12}$	$.68 \cdot 10^{-12}$	$.16 \cdot 10^{-12}$	$.33 \cdot 10^{-11}$
SYS5	$.33 \cdot 10^{-11}$	$.40 \cdot 10^{-12}$	$.45 \cdot 10^{-11}$	$.11 \cdot 10^{-11}$	$.86 \cdot 10^{-12}$

Fig-5.3.6- Overall relative error for 5 algorithms applied to SISO systems; NOM=64; deterministic situation.

In fig-5.3.6- the overall relative error is given. The results are all very close to each other and very small.

Also in this SISO situation the realized eigenvalues are exact within the precision of the mantissa.

5.4 RESULTS IN THE NOISY MIMO CASE

In this section the results of the simulations of noisy MIMO systems will be presented successively. The complete set of results has been given in fig-5.4.1- up to fig-5.4.29-. For ease of survey these figures are added in an appendix. In this section the most remarkable results will be discussed.

1. Positions of singular values (fig-5.4.2- up to fig-5.4.9-)

As expected it can been seen that in all cases the level of singular values is higher in the Hankel matrix than in the Page matrix. This is caused by the larger dimensions of this first matrix. However what is important is the decrease of singular values and especially the gap between $\tilde{\delta}_4$ and $\tilde{\delta}_5$. The gap between these singular values should lead to a choice of the dimension of the realization of 4 (the exact dimension of the system).
For $\sigma=\sigma_2$ it can be seen that hardly any difference can be made between the decrease in singular values in the two situations H and P. The singular values in H are on a higher level, but the distribution is exactly the same, especially with respect to $\tilde{\delta}_4$ and $\tilde{\delta}_5$. In case of SYS2 the Page matrix shows a small advantage in this respect. This coincides with the results in the deterministic situation (see fig-5.3.1-).
In case of SYS0 and SYS3 where the deterministic singular value δ_4 in the Page matrix became very small, this drawback is more or less compensated by the fact that $\tilde{\delta}_5$ for the Page matrix is smaller than for the Hankel matrix. Especially for $\sigma=\sigma_2$ the gap between $\tilde{\delta}_4$ and $\tilde{\delta}_5$ in the Page matrix is a somewhat better.
For situations of very small disturbances (SYS0,σ_1) the mentioned drawback is still there.

For all test cases it holds that for large NOM the relative gap between $\tilde{\delta}_4$ and $\tilde{\delta}_5$ becomes smaller. Probably for every situation there can be found an optimum NOM for which holds that increasing this NOM will introduce more noise energy than signal energy, and therefore only will deteriorate the results.

2. Noise reduction factor (fig-5.4.12- and -5.4.13-)

A short outline of the results is given in fig-5.4.30- and fig-5.4.31-.

NRF; MIMO systems; $\sigma=\sigma_1$; NOM=16

$\sigma = \sigma_1$	H	P	P/H
SYS0	0.26	0.79	0.27
SYS1	0.32	0.78	0.40
SYS2	0.29	0.83	0.29
SYS3	0.29	0.76	0.31

Fig-5.4.30- Noise reduction factor for Hankel, Page and combined algorithms; MIMO systems; $\sigma=\sigma_1$; NOM=16.

NRF; MIMO systems; $\sigma=\sigma_2$; NOM=16

$\sigma = \sigma_2$	H	P	P/H
SYS0	0.47	0.84	0.36
SYS1	0.33	0.88	0.40
SYS2	0.38	0.86	0.39
SYS3	0.40	0.85	0.38

Fig-5.4.31- Noise reduction factor for Hankel, Page and combined algorithms; MIMO systems; $\sigma=\sigma_1$; NOM=16.

The most remarkable result is the relatively high value of NRF for Page matrix algorithms. In comparison with the Hankel matrix, noise reduction with the Page matrix is far less effective. In spite of the equally balanced weighting of Markov parameters during this noise filtering, the total reduction of noise energy is much smaller. Apparently the missing sufficient condition for the existence of an n-dimensional realization when rank $\tilde{P}$=n, causes a heavier disturbance of the optimum realization than the non equally balancedness of the Hankel filtering, and the lacking Hankel structure of the approximating H.

The condition rank(H)=n is much heavier than rank(P)=n as the Hankel matrix is much bigger than the Page matrix, so this condition decreases the degrees of freedom more drastically.

In some cases the double noise filtering of the combined algorithm P/H gives better results than in the Hankel case. Especially for $\sigma=\sigma_2$, where in 70 percent of the situations P/H gives better results, apparently prefiltering of the Markov parameters in a Page matrix leads to a result that improves the result of the Hankel matrix algorithms.

In section 2.4 an analysis has been given of the influence of noise on singular values of the Grammian matrix of $\tilde{P}$ in expectation. The result was:

$$d_i^2 = \delta_i^2 + \sigma^2 . \max(h,m) \qquad i=1,..,\min(h,m)$$

In order to test if these squared singular values have any predictive value with respect to the singular values of $\tilde{P}$, a theoretically derived noise reduction factor for the Page matrix can be given.

If $d^2{}_i$ would represent squared singular values of $\tilde{P}$, then the deleted noise energy during the noise filtering would be:

$$[\min(h,m)-n] \,.\, \sigma^2.\max(h,m)$$
$$= h.m.\sigma^2 - n.\sigma^2.\max(h,m)$$

where the first term represents the original appearing noise energy and the second term the remaining noise energy after noise filtering.

Now there can be written:

$$\mathrm{NRF} = \frac{n.\sigma^2.\max(h,m)}{h.m.\sigma^2} = \frac{n}{\min(h,m)}$$

According to fig-5.2.2. for NOM=16, min(h,m)=8, and NRF would be 4/8 = 0.50.

The given results show that this value of NRF is never reached in Page algorithms. Moreover the results of NRF appear to be dependent on σ. Apparently it is premature to conclude that the expression for the singular values of the Grammian matrices of $\tilde{P}$ does give us straightforward information on the properties of the singular values of $\tilde{P}$.

3. Error in the realization part (fig-5.4.14- and -5.4.15-)

This error ERRB indicates the exactness of the resulting realization when a Page matrix or an approximating Hankel matrix of reduced rank is given.

When analysing the results represented in the figures mentioned above, it is very clear that in Hankel matrix algorithms this error is much smaller than in Page matrix algorithms.

For $\sigma=\sigma_1$ these differences are about a factor 10-20 and for SYSO, where the disturbance of noise is very small, even a factor 10^3. This can be explained by the fact that for a very low noise level the approximating Hankel matrix apparently will approximate a real Hankel matrix very close, and consequently the existence of an exact n-dimensional realization will become closer. For the Page matrix the non validity of this sufficient condition will remain and apparently has a much heavier disturbing influence than the non exactness in case of the Hankel matrix.

For $\sigma=\sigma_2$ the differences become smaller but the superiority of the Hankel matrix algorithms remains.

In many cases the combined algorithm gives the best results, but at first sight these situations appear quite randomly.

4. Overall errors (fig-5.4.16 up to .19-, -5.4.28-, -5.4.29-)

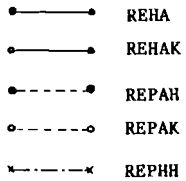
For $\sigma=\sigma_1$ the results of ERR1 and AVER1 are quite similar to the results of ERRB.

The difference between Hankel and Page algorithms is quite large. In many situations the Page algorithms even show a higher relative error than the error ERRO at the input.

This means that the total noise energy in the array of Markov parameters has been increased during the algorithm in order to find an n-dimensional realization. In Hankel matrix algorithms this never happens.

The combined algorithm shows better results in some cases, especially when NOM is small (8-12). The results of ERR1 and AVER1 show the same trend.

Results of ERR1; MIMO systems; $\sigma=\sigma_2$



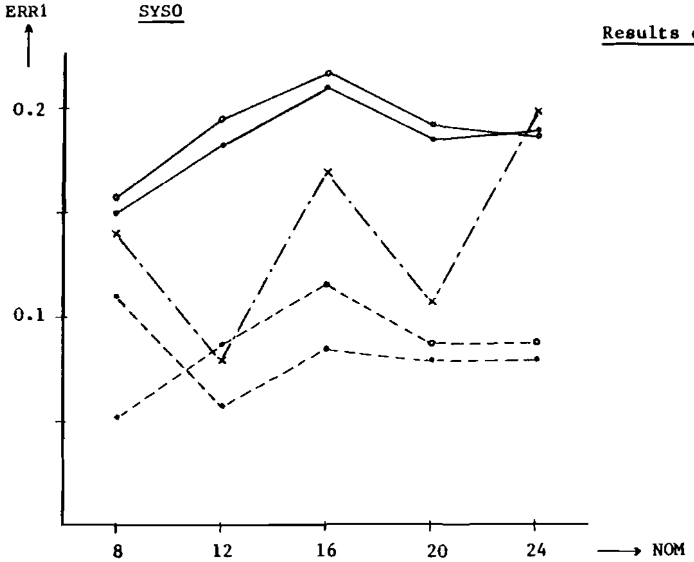


Fig-5.4.32a- Overall relative error in case of SYS0; $\sigma=\sigma_2$; NOM=8-24.

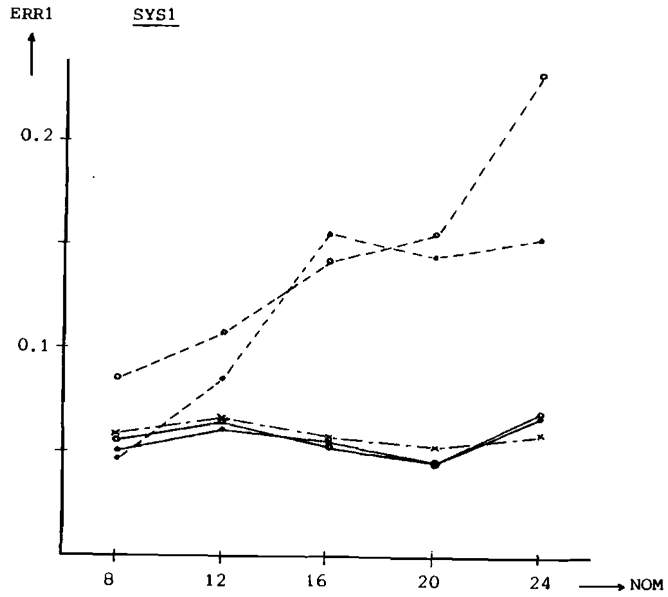


Fig-5.4.32b- Overall relative error in case of SYS1; $\sigma=\sigma_2$; NOM=8-24.

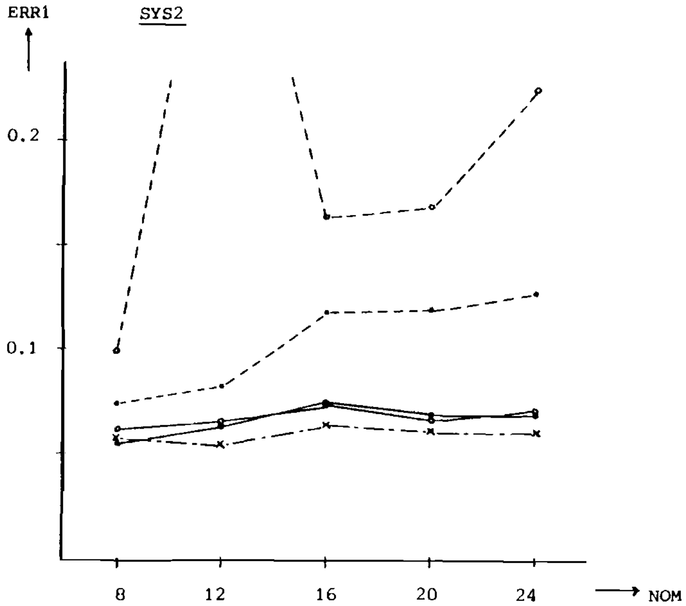


Fig-5.4.32c- Overall relative error in case of SYS2; $\sigma=\sigma_2$; NOM=8-24.

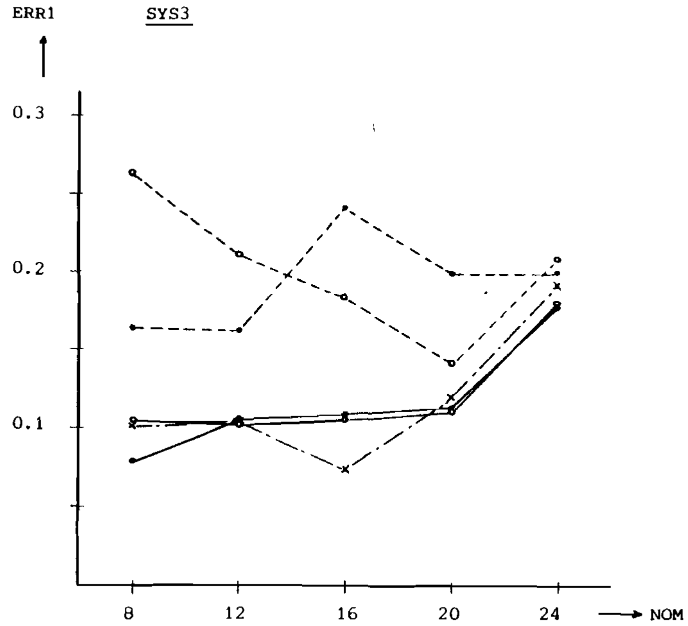


Fig-5.4.32d- Overall relative error in case of SYS3; $\sigma=\sigma_2$; NOM=8-24.

For higher noise contributions the differences become much smaller and the character of the results changes. The results of ERR1 are schematically presented in fig-5.4.32-.
In the given figures the following features can be noticed:

For SYS0 the results of the Page matrix algorithms are better than those of the Hankel matrix algorithms.

For SYS2 the combined algorithm shows the best results and for SYS1 and SYS3 the Hankel and combined algorithms are very close together, with better results than the Page matrix algorithms.
In case of SYS1 quite a strange effect happens: the chosen noise series acts for this specific system as a kind of dependent noise. This appears from the singular values of the noisy Page matrix where the squared sum of noisy singular values appears to be smaller than the squared sum of deterministic singular values.
The accidentally chosen starting value of the noise generator causes the noise to be dependent on the Markov parameters in case of SYS1. This may cause the increase of ERR1 for SYS1 with increasing NOM. Because of the large eigenvalues of SYS1 we would rather expect these results to become smaller with increasing NOM.

In case of SYS2, NOM=12 an instability occurs in the algorithm REPAK. This is caused by the fact that in this Kung algorithm for NOM=12 the system matrix A is approximated based on a very small controllability matrix:

$$A = \overleftarrow{\Delta} \cdot \Delta^{+} \qquad \text{with } \dim(\Delta) = n \times 6 = 4 \times 6.$$

Some "unfortunate" noise samples now easily can cause the matrix A to have unstable eigenvalues, and therefore to find a series of approximated Markov parameters which differs quite a lot from the original series. Apart from this effect also the non sufficiency of the rank condition can be the cause of this instability.
It should be noted that for NOM=8 the algorithm REPAK uses a controllability matrix with dimensions 4×3. This means that the given matrix is too small for approximating the correct A matrix of dimension 4. Therefore in these situations the REPAK algorithm does not fulfill the correct rank condition.

The results of AVER1 are quite similar to the results of ERR1. Because of the remarkable better results of the Page matrix algorithms in case of SYS0, σ_2 two extra runs of this situation have been executed with two different noise series. The average results have been calculated (fig-5.4.28- and fig-5.4.29-).
For ERR1 these results are given schematically in fig-5.4.33-:

Results of ERR1; SYS0; $\sigma=\sigma_2$; averaged over 3 noise series

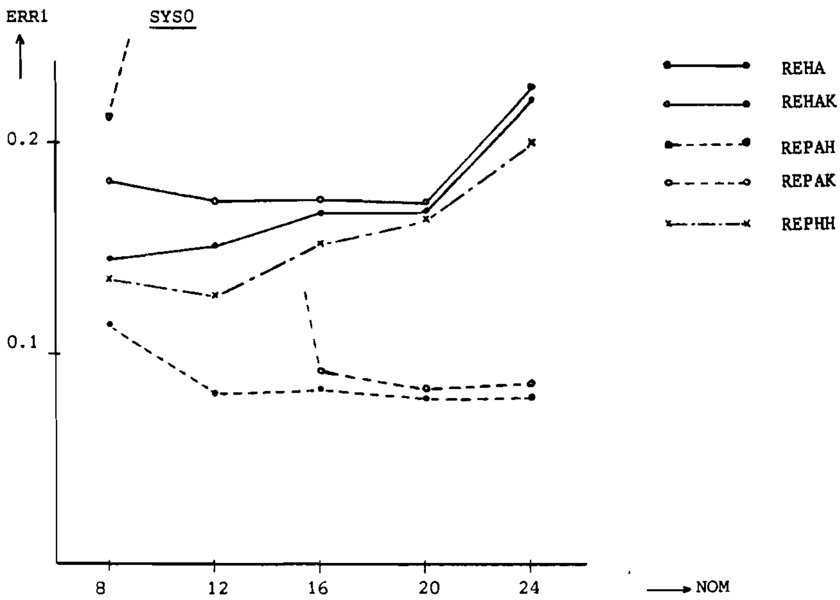


Fig-5.4.33- Overall relative error in case of SYS0; $\sigma=\sigma_2$; NOM=8-24.
ERR1 averaged over 3 noise series.

The trend as noted in fig-5.4.32a- remains the same: Page matrix algorithms give better results than Hankel matrix algorithms. Now in this situation also an unstability of REPAK with NOM=12 occurs.

5. Realized eigenvalues (fig-5.4.20- up to fig-5.4.27-)

For small σ the realized eigenvalues can be fairly well compared with the original ones of the system. Although in some situations there appear complex eigenvalues, the imaginary part is always much smaller than the real part. The Page matrix algorithms have problems with the realization of the smallest eigenvalue λ_4 of SYS0; in case of SYS3 all algorithms fail in realizing this smallest one. This is caused by the fact that the number of small eigenvalues exceeds q (see also section 5.3.2).

For $\sigma=\sigma_2$ at most two eigenvalues are comparable with the exact ones. Any comparison between the different algorithms is very hard. In all cases (as well σ_1 and σ_2) it can be notified that in the situation REPAK, NOM=12 the rank condition is not fulfilled: the smallest eigenvalue λ_4 is always very small. The information on this eigenvalue has already been lost during the algorithm.

5.5 RESULTS IN THE NOISY SISO CASE

The complete set of results are given in fig-5.5.1- up to fig-5.5.9-. Again for ease of survey these figures are given in the appendix.

Apart from the realized eigenvalues, all results discussed below are averaged over 3 noise series with different starting values of the noise generator, as mentioned in section 5.2.2.

1. Positions of singular values (fig-5.5.2- and fig-5.5.3-).
Apart from the higher level of singular values in the Hankel matrix, it is hard to say anything about differences between the two matrices. The decrease in singular values shows about the same trend, although the relative gap between $\tilde{\delta}_{n+1}$ and $\tilde{\delta}_n$ is a little bit larger in case of P.

2. Noise reduction factor (fig-5.5.5-).
With respect to H and P the results of NRF show the same behaviour as in the MIMO case. The noise reduction factor is a lot higher in the Page matrix than in the two other situations. For middle or high level noise the combined algorithm shows the best results. It can be concluded that for high level noise the role of the ill working window in the Hankel matrix becomes more important. An extra filtering with a Page matrix then will give the best results.

3. Error in the realization part (fig-5.5.6-).
In general the combined algorithm shows the best results in these experiments. For SYS4 P shows better results than H, for SYS5 this is the other way around. The algorithm REPAK again shows an unstable result; this happened in one of the executed runs. The cause of this unstability is not exactly clear.

4. Overall error (fig-5.5.7- and fig-5.5.8-).

Except the situation of SYS5 with low noise ($\sigma=\sigma_1$), in all other experiments the Page matrix algorithms show the best overall results. In case of SYS5 and high noise ($\sigma=\sigma_2$) the difference in ERR1 even reaches almost a factor 2.

The results of AVER1 show the same trend; this could be expected because NOM is fixed to 64 and N, the number of Markov parameters over which AVER1 is averaged, is fixed to 65.

Although the two approximations, noise filtering and approximate realization, show worse results in case of the Page matrix, the overall errors show the best results.

However, care has to be taken by extrapolating these results to a general statement. The dynamics of the specifically chosen system may play a role in this respect.

5. Realized eigenvalues (fig-5.5.9-).

In case of SYS4 only REPHH in case of $\sigma=\sigma_1$ can realize an eigenvalue that is comparable with the exact one. In case of SYS5, $\sigma=\sigma_1$, the Page matrix algorithms have problems with the realization of λ_2. In all other situations the realized eigenvalues are hardly comparable with the exact ones.

5.6 REMARKS

The main results of this chapter are briefly summarized here. Because of the total amount of experiments the number of noise series for each experiment in the MIMO case had to be restricted to one. In one situation (SYS0, $\sigma=\sigma_2$) three different noise series were applied, but this did not change the results essentially. From the results of the simulations it can be concluded that some of the theoretically very attractive properties of the Page matrix algorithms appear to work out less positively in practice. Both Page and Hankel matrix algorithms do not result in an optimal approximation. However the theoretical non-optimality of Page matrix algorithms appears to be more severe in many practical situations.

For the chosen MIMO systems the Hankel matrix algorithms show better results than the Page matrix matrix algorithms, in noise filtering, approximate realization, as well as in overall fit. This superiority also holds for the combined Page/Hankel algorithm, that in some cases improves the results of the Hankel algorithms.
In one situation (SYSO, $\sigma=\sigma_2$) the Page matrix algorithms show better overall results.
It should be stressed that in general the Page matrix algorithms provide these results with very little computational effort compared to the Hankel matrix algorithms; this happens especially if a large number of Markov parameters is used.
In the SISO situation with a high number of Markov parameters and a heavy noise disturbance the Page matrix algorithms show better overall results, although the noise filtering and realization are worse.
In both test situations MIMO and SISO the algorithm REPAK showed cases of unstable results for NOM=12.

Based on these results in general cases the Page matrix approach can be advised in three cases:

1. If the signal to noise ratio is bad, due to low sample frequency (small eigenvalues) and/or high noise level.
2. For SISO systems and many Markov parameters are available.
3. If computation time and memory space cause problems.

CONCLUSIONS

The purpose of this project has been to investigate whether it is possible to develop an approximate realization method that overcomes the problems of the existing methods, like the adapted Ho-Kalman algorithm. Within this scope the Page matrix approach has been introduced and analysed.
From the comparison of these two approaches, in theory and in simulation results, the following conclusions can be drawn:

- In case of deterministic Markov parameters an exact realization of the system will be obtained with the Page matrix algorithm, under the condition that such a realization with dimension n (=rank(P)) exists, and that delay lines are not involved.

- The Page matrix algorithm, just like the already available algorithms, does not accomplish an optimal solution for the approximate realization problem when Markov parameters are disturbed with SWAYING noise. This non-optimality is caused by the fact that rank(P)=n is not a sufficient condition for the existence of an n-dimensional realization of the Markov parameters in P.

- When taking into account a fixed number of Markov parameters, the size of the Page matrix is smaller than the size of the Hankel matrix. This implies a considerable reduction in computation time.

- During the noise filtering of the Page matrix (by omitting the non relevant singular values), there is a constant weighting factor for the Markov parameters. This noise filtering provides us directly with a set of unique Markov parameters in the reduced rank Page matrix of rank n, contrary to the situation for the Hankel matrix.

- With respect to the noise filtering, the results of the simulations show that the non optimality of the Page matrix approach is more severe than the non optimality of the Hankel matrix algorithms. Although the Markov parameters are weighted with a constant factor, the noise energy that is filtered out during the rank reduction is far less than in case of the Hankel matrix.

- In case of systems disturbed by low level noise the overall results of the Hankel matrix approach surpass the results of the Page matrix algorithm. A combined algorithm of Page and Hankel noise filtering and Hankel realization gives comparable results as the Hankel matrix approach, and even is superior in several situations.

- The Page matrix approach gave better results in two cases. In the first case if the signal to noise ratio is bad, due to low sample frequency and high noise level (SYSO,σ_2). In the second case of SISO systems disturbed by high level noise when a long sequence of Markov parameters is available.

LIST OF SYMBOLS

α	Observability index
A	System matrix, $(n \times n)$
β	Controllability index
B	Distribution matrix, $(n \times p)$
C	Output matrix, $(q \times n)$
D	Input-output matrix, $(q \times p)$
δ	Singular value
$\delta(N)$	Singular value caused by noise
$\Delta[\mu]$	Controllability matrix consisting of μ block matrices
g	Number of rows in the Hankel matrix
$\Gamma[\eta]$	Observability matrix consisting of η block matrices
γ	Number of block rows in the Hankel matrix
h	Number of rows in the page matrix
$H[\gamma,\zeta]$	Hankel matrix, $(g \times l) = (\gamma.q \times \zeta.p)$
$\tilde{H}$	Noise disturbed Hankel matrix
H_k	Approximating Hankel matrix of reduced rank
ℓ	Number of columns in the Hankel matrix
L	Number of Markov parameters
λ	Eigenvalue
m	Number of columns in the page matrix
μ	Number of block columns in the page matrix
M_i	Markov parameter with index i
$\tilde{M}_i$	Markov parameter disturbed by noise
M_i^*	Noise filtered Markov parameter
$\hat{M}_i$	Reconstructed Markov parameter
n	Rank of the system
η	Number of block rows in the page matrix
N	Matrix of noise samples
p	Number of inputs
$P[\eta,\mu]$	Page matrix, $(h \times m) = (\eta.q \times \mu.p)$
$\tilde{P}$	Page matrix disturbed with noise
q	Number of outputs
r	Realizability index
ρ	Number of singular values (smallest dimension) of a matrix
s	Maximum of α and β
σ^2	Variance of the noise
Σ	Diagonal matrix of singular values
ξ	Sample from a zero mean Gaussian (SWAYING) noise source
Ξ_H	Matrix of noise samples in case of the Hankel matrix
Ξ_P	Matrix of noise samples in case of the page matrix
V	Matrix of right singular vectors
W	Matrix of left singular vectors
ζ	Number of block columns in the Hankel matrix

ERRO, AVERO; MIMO systems.

ERRO AVERO	NOM	$\sigma = \sigma_1$		$\sigma = \sigma_2$	
		ERRO	AVERO	ERRO	AVERO
SYS0	8	$.15 \cdot 10^{-5}$	$.28 \cdot 10^{-6}$	0.370	$.70 \cdot 10^{-1}$
	12	$.24 \cdot 10^{-5}$	$.28 \cdot 10^{-6}$	0.598	$.70 \cdot 10^{-1}$
	16	$.34 \cdot 10^{-5}$	$.28 \cdot 10^{-6}$	0.862	$.70 \cdot 10^{-1}$
	20	$.43 \cdot 10^{-5}$	$.28 \cdot 10^{-6}$	1.082	$.70 \cdot 10^{-1}$
	24	$.55 \cdot 10^{-5}$	$.28 \cdot 10^{-6}$	1.380	$.70 \cdot 10^{-1}$
SYS1	8	$.17 \cdot 10^{-2}$	$.70 \cdot 10^{-3}$	0.167	$.70 \cdot 10^{-1}$
	12	$.26 \cdot 10^{-2}$	$.70 \cdot 10^{-3}$	0.255	$.70 \cdot 10^{-1}$
	16	$.36 \cdot 10^{-2}$	$.70 \cdot 10^{-3}$	0.360	$.70 \cdot 10^{-1}$
	20	$.45 \cdot 10^{-2}$	$.70 \cdot 10^{-3}$	0.448	$.70 \cdot 10^{-1}$
	24	$.57 \cdot 10^{-2}$	$.70 \cdot 10^{-3}$	0.569	$.70 \cdot 10^{-1}$
SYS2	8	$.60 \cdot 10^{-2}$	$.28 \cdot 10^{-2}$	0.150	$.70 \cdot 10^{-1}$
	12	$.90 \cdot 10^{-2}$	$.28 \cdot 10^{-2}$	0.225	$.70 \cdot 10^{-1}$
	16	$.13 \cdot 10^{-1}$	$.28 \cdot 10^{-2}$	0.315	$.70 \cdot 10^{-1}$
	20	$.16 \cdot 10^{-1}$	$.28 \cdot 10^{-2}$	0.390	$.70 \cdot 10^{-1}$
	24	$.20 \cdot 10^{-1}$	$.28 \cdot 10^{-2}$	0.495	$.70 \cdot 10^{-1}$
SYS3	8	$.23 \cdot 10^{-2}$	$.70 \cdot 10^{-3}$	0.225	$.70 \cdot 10^{-1}$
	12	$.34 \cdot 10^{-2}$	$.70 \cdot 10^{-3}$	0.340	$.70 \cdot 10^{-1}$
	16	$.48 \cdot 10^{-2}$	$.70 \cdot 10^{-3}$	0.477	$.70 \cdot 10^{-1}$
	20	$.59 \cdot 10^{-2}$	$.70 \cdot 10^{-3}$	0.591	$.70 \cdot 10^{-1}$
	24	$.75 \cdot 10^{-2}$	$.70 \cdot 10^{-3}$	0.750	$.70 \cdot 10^{-1}$

Fig-5.4.1- Relative and average error in input Markov parameter sequence; MIMO systems; $\sigma=\sigma_1,\sigma_2$; NOM=8-24.

Singular values of SYS0; $\sigma=\sigma_1$.

$\sigma=\sigma_1$	NOM	$\tilde{\delta}_1$	$\tilde{\delta}_2$	$\tilde{\delta}_3$	$\tilde{\delta}_4$	$\tilde{\delta}_5$
H	8	2.084	1.919	0.073	0.028	0.002
	12	2.084	1.919	0.073	0.029	0.003
	16	2.084	1.919	0.073	0.029	0.004
	20	2.084	1.919	0.073	0.029	0.004
	24	2.084	1.919	0.073	0.029	0.005
P	8	1.955	1.875	0.026	0.014	0.002
	12	1.950	1.881	0.011	0.006	0.002
	16	1.949	1.882	0.005	0.002	0.002
	20	1.949	1.882	0.005	0.002	0.002
	24	1.949	1.882	0.005	0.003	0.002

Fig-5.4.2- Singular values of the noise disturbed Hankel and Page matrix in case of SYS0; $\sigma=\sigma_1$; NOM=8-24.

Singular values of SYS0; $\sigma=\sigma_2$.

$\sigma=\sigma_2$	NOM	$\tilde{\delta}_1$	$\tilde{\delta}_2$	$\tilde{\delta}_3$	$\tilde{\delta}_4$	$\tilde{\delta}_5$
H	8	2.165	2.011	1.174	0.943	0.866
	12	2.425	2.138	1.560	1.453	1.154
	16	2.531	2.313	1.967	1.853	1.838
	20	2.679	2.520	2.224	2.141	2.065
	24	2.966	2.748	2.459	2.395	2.371
P	8	1.943	1.673	0.904	0.568	0.519
	12	1.977	1.884	1.065	0.749	0.594
	16	2.171	1.950	1.100	0.810	0.674
	20	2.260	1.984	1.177	0.934	0.811
	24	2.284	2.007	1.281	1.183	1.083

Fig-5.4.3- Singular values of the noise disturbed Hankel and Page matrix in case of SYS0; $\sigma=\sigma_2$; NOM=8-24.

Singular values of SYS1; $\sigma=\sigma_1$

$\sigma=\sigma_1$	NOM	$\tilde{\delta}_1$	$\tilde{\delta}_2$	$\tilde{\delta}_3$	$\tilde{\delta}_4$	$\tilde{\delta}_5$
H	8	4.934	2.756	0.344	0.126	0.113
	12	5.795	2.821	0.448	0.166	0.152
	16	6.320	2.850	0.495	0.197	0.190
	20	6.645	2.867	0.544	0.223	0.218
	24	6.836	2.878	0.617	0.249	0.239
P	8	3.308	2.200	0.254	0.084	0.049
	12	3.447	2.209	0.307	0.103	0.082
	16	3.484	2.224	0.339	0.103	0.079
	20	3.501	2.224	0.364	0.107	0.102
	24	3.510	2.225	0.366	0.118	0.112

Fig-5.4.4 Singular values of the noise disturbed Hankel and Page matrix in case of SYS1; $\sigma=\sigma_1$; NOM=8-24.

Singular values of SYS1; $\sigma=\sigma_2$

$\sigma=\sigma_2$	NOM	$\tilde{\delta}_1$	$\tilde{\delta}_2$	$\tilde{\delta}_3$	$\tilde{\delta}_4$	$\tilde{\delta}_5$
H	8	4.439	2.905	1.179	1.129	0.871
	12	5.596	3.037	1.629	1.558	1.339
	16	6.159	3.071	1.980	1.908	1.854
	20	6.404	3.142	2.251	2.188	2.076
	24	6.463	3.301	2.483	2.405	2.303
P	8	2.967	2.153	0.918	0.552	0.417
	12	3.229	2.247	1.133	0.815	0.527
	16	3.296	2.330	1.281	0.756	0.678
	20	3.303	2.341	1.377	1.008	0.817
	24	3.315	2.392	1.414	1.161	1.072

Fig-5.4.5- Singular values of the noise disturbed Hankel and Page matrix in case of SYS1; $\sigma=\sigma_2$; NOM=8-24.

Singular values of SYS2; $\sigma=\sigma_1$

$\sigma=\sigma_1$	NOM	$\tilde{\delta}_1$	$\tilde{\delta}_2$	$\tilde{\delta}_3$	$\tilde{\delta}_4$	$\tilde{\delta}_5$
H	8	5.368	2.931	0.637	0.254	0.225
	12	6.466	3.233	0.737	0.312	0.299
	16	7.137	3.394	0.780	0.452	0.393
	20	7.571	3.489	0.793	0.592	0.446
	24	7.821	3.527	0.829	0.676	0.498
P	8	3.477	2.317	0.405	0.168	0.097
	12	3.669	2.355	0.471	0.220	0.141
	16	3.723	2.391	0.408	0.271	0.189
	20	3.752	2.397	0.454	0.299	0.192
	24	3.757	2.400	0.464	0.309	0.219

Fig-5.4.6- Singular values of the noise disturbed Hankel and Page matrix in case of SYS2; $\sigma=\sigma_1$; NOM=8-24.

Singular values of SYS2; $\sigma=\sigma_2$

$\sigma=\sigma_2$	NOM	$\tilde{\delta}_1$	$\tilde{\delta}_2$	$\tilde{\delta}_3$	$\tilde{\delta}_4$	$\tilde{\delta}_5$
H	8	5.542	2.527	1.343	1.126	0.975
	12	6.774	3.202	1.716	1.553	1.387
	16	7.397	3.570	2.003	1.979	1.859
	20	7.841	3.733	2.264	2.207	2.135
	24	8.033	3.896	2.495	2.428	2.397
P	8	3.533	2.126	0.916	0.528	0.426
	12	3.819	2.188	1.155	0.861	0.593
	16	3.796	2.457	1.330	0.885	0.687
	20	3.816	2.540	1.440	1.104	0.761
	24	3.822	2.546	1.478	1.216	1.071

Fig-5.4.7- Singular values of the noise disturbed Hankel and Page matrix in case of SYS2; $\sigma=\sigma_2$; NOM=8-24.

Singular values of SYS3; $\sigma=\sigma_1$

$\sigma=\sigma_1$	NOM	$\tilde{\delta}_1$	$\tilde{\delta}_2$	$\tilde{\delta}_3$	$\tilde{\delta}_4$	$\tilde{\delta}_5$
H	8	4.112	1.910	0.598	0.112	0.095
	12	4.973	1.929	0.662	0.152	0.149
	16	5.529	1.941	0.683	0.193	0.185
	20	5.884	1.947	0.688	0.223	0.214
	24	6.099	1.951	0.710	0.247	0.233
P	8	2.848	1.837	0.425	0.081	0.047
	12	2.977	1.865	0.492	0.093	0.067
	16	3.022	1.885	0.479	0.093	0.081
	20	3.043	1.885	0.492	0.111	0.096
	24	3.052	1.886	0.498	0.128	0.111

Fig-5.4.8- Singular values of the noise disturbed Hankel and Page matrix in case of SYS3; $\sigma=\sigma_1$; NOM=8-24.

Singular values of SYS3; $\sigma=\sigma_2$

$\sigma=\sigma_2$	NOM	$\tilde{\delta}_1$	$\tilde{\delta}_2$	$\tilde{\delta}_3$	$\tilde{\delta}_4$	$\tilde{\delta}_5$
H	8	3.607	1.954	1.288	1.100	0.845
	12	4.730	2.119	1.701	1.485	1.341
	16	5.273	2.257	1.981	1.901	1.838
	20	5.508	2.450	2.234	2.160	2.102
	24	5.582	2.667	2.472	2.381	2.248
P	8	2.549	1.704	0.894	0.533	0.441
	12	2.771	1.781	1.247	0.841	0.628
	16	2.836	1.967	1.324	0.771	0.742
	20	2.844	1.983	1.420	1.054	0.773
	24	2.850	2.012	1.437	1.207	1.086

Fig-5.4.9- Singular values of the noise disturbed Hankel and Page matrix in case of SYS3; $\sigma=\sigma_2$; NOM=8-24.

ERRF; MIMO systems; $\sigma = \sigma_1$

$\sigma = \sigma_1$	NOM	H	P	P/H
SYS0	8	$.76 \cdot 10^{-6}$	$.11 \cdot 10^{-5}$	$.76 \cdot 10^{-6}$
	12	$.82 \cdot 10^{-6}$	$.17 \cdot 10^{-5}$	$.81 \cdot 10^{-6}$
	16	$.89 \cdot 10^{-6}$	$.27 \cdot 10^{-5}$	$.91 \cdot 10^{-6}$
	20	$.94 \cdot 10^{-6}$	$.31 \cdot 10^{-5}$	$.94 \cdot 10^{-6}$
	24	$.99 \cdot 10^{-6}$	$.35 \cdot 10^{-5}$	$.97 \cdot 10^{-6}$
SYS1	8	$.93 \cdot 10^{-3}$	$.14 \cdot 10^{-2}$	$.92 \cdot 10^{-3}$
	12	$.11 \cdot 10^{-2}$	$.20 \cdot 10^{-2}$	$.98 \cdot 10^{-3}$
	16	$.12 \cdot 10^{-2}$	$.28 \cdot 10^{-2}$	$.14 \cdot 10^{-2}$
	20	$.14 \cdot 10^{-2}$	$.33 \cdot 10^{-2}$	$.17 \cdot 10^{-2}$
	24	$.17 \cdot 10^{-2}$	$.36 \cdot 10^{-2}$	$.16 \cdot 10^{-2}$
SYS2	8	$.27 \cdot 10^{-2}$	$.58 \cdot 10^{-2}$	$.33 \cdot 10^{-2}$
	12	$.44 \cdot 10^{-2}$	$.76 \cdot 10^{-2}$	$.41 \cdot 10^{-2}$
	16	$.37 \cdot 10^{-2}$	$.10 \cdot 10^{-1}$	$.36 \cdot 10^{-2}$
	20	$.41 \cdot 10^{-2}$	$.12 \cdot 10^{-1}$	$.43 \cdot 10^{-2}$
	24	$.41 \cdot 10^{-2}$	$.14 \cdot 10^{-1}$	$.42 \cdot 10^{-2}$
SYS3	8	$.13 \cdot 10^{-2}$	$.20 \cdot 10^{-2}$	$.11 \cdot 10^{-2}$
	12	$.14 \cdot 10^{-2}$	$.27 \cdot 10^{-2}$	$.13 \cdot 10^{-2}$
	16	$.14 \cdot 10^{-2}$	$.36 \cdot 10^{-2}$	$.15 \cdot 10^{-2}$
	20	$.15 \cdot 10^{-2}$	$.42 \cdot 10^{-2}$	$.14 \cdot 10^{-2}$
	24	$.18 \cdot 10^{-2}$	$.47 \cdot 10^{-2}$	$.16 \cdot 10^{-2}$

Fig-5.4.10- Relative error in noise filtered Markov parameters for Hankel, Page and combined algorithms; MIMO systems; $\sigma = \sigma_1$; NOM=8-24.

ERRF; MIMO systems; $\sigma = \sigma_2$

$\sigma = \sigma_2$	NOM	H	P	P/H
SYS0	8	0.246	0.329	0.211
	12	0.331	0.542	0.263
	16	0.401	0.728	0.314
	20	0.379	0.847	0.323
	24	0.399	0.998	0.334
SYS1	8	0.109	0.152	0.105
	12	0.132	0.234	0.133
	16	0.120	0.317	0.144
	20	0.113	0.357	0.148
	24	0.147	0.414	0.142
SYS2	8	0.086	0.147	0.084
	12	0.122	0.198	0.113
	16	0.121	0.272	0.123
	20	0.120	0.315	0.107
	24	0.133	0.348	0.108
SYS3	8	0.140	0.205	0.134
	12	0.201	0.302	0.178
	16	0.193	0.405	0.182
	20	0.216	0.473	0.233
	24	0.270	0.529	0.276

Fig-5.4.11- Relative error in noise filtered Markov parameters for Hankel, Page and combined algorithms; MIMO systems; $\sigma = \sigma_2$; NOM=8-24.

NRF; MIMO systems; $\sigma = \sigma_1$

$\sigma = \sigma_1$	NOM	H	P	P/H
	8	0.51	0.76	0.51
	12	0.34	0.71	0.34
SYS0	16	0.26	0.79	0.27
	20	0.22	0.71	0.22
	24	0.18	0.63	0.18
	8	0.56	0.86	0.55
	12	0.42	0.80	0.38
SYS1	16	0.32	0.78	0.40
	20	0.30	0.73	0.38
	24	0.30	0.63	0.29
	8	0.44	0.97	0.54
	12	0.49	0.84	0.45
SYS2	16	0.29	0.83	0.29
	20	0.26	0.79	0.28
	24	0.20	0.69	0.21
	8	0.56	0.87	0.50
	12	0.40	0.79	0.39
SYS3	16	0.29	0.76	0.31
	20	0.25	0.72	0.23
	24	0.24	0.62	0.22

Fig-5.4.12- Noise reduction factor for Hankel, Page and combined algorithms; MIMO systems; $\sigma = \sigma_1$; NOM=8-24.

NRF; MIMO systems; $\sigma = \sigma_2$

$\sigma = \sigma_2$	NOM	H	P	P/H
	8	0.66	0.89	0.57
	12	0.55	0.91	0.44
SYS0	16	0.47	0.84	0.36
	20	0.35	0.78	0.30
	24	0.29	0.72	0.24
	8	0.65	0.91	0.63
	12	0.52	0.92	0.52
SYS1	16	0.33	0.88	0.40
	20	0.25	0.80	0.33
	24	0.26	0.73	0.25
	8	0.57	0.98	0.56
	12	0.54	0.88	0.50
SYS2	16	0.38	0.86	0.39
	20	0.31	0.81	0.27
	24	0.27	0.70	0.22
	8	0.62	0.91	0.60
	12	0.59	0.89	0.52
SYS3	16	0.40	0.85	0.38
	20	0.37	0.80	0.39
	24	0.36	0.71	0.37

Fig-5.4.13- Noise reduction factor for Hankel, Page and combined algorithms; MIMO systems; $\sigma = \sigma_1$; NOM=8-24.

ERRB; MIMO systems; $\sigma=\sigma_1$

$\sigma=\sigma_1$	NOM	REHA	REHAK	REPAH	REPAK	REPHH
SYS0	8	$.13\cdot10^{-6}$	$.13\cdot10^{-6}$	$.28\cdot10^{-5}$	$.16\cdot10^{-3}$	$.13\cdot10^{-6}$
	12	$.18\cdot10^{-6}$	$.18\cdot10^{-6}$	$.45\cdot10^{-5}$	$.42\cdot10^{-5}$	$.18\cdot10^{-6}$
	16	$.25\cdot10^{-6}$	$.25\cdot10^{-6}$	$.24\cdot10^{-3}$	$.19\cdot10^{-3}$	$.26\cdot10^{-6}$
	20	$.30\cdot10^{-6}$	$.30\cdot10^{-6}$	$.31\cdot10^{-3}$	$.26\cdot10^{-3}$	$.29\cdot10^{-6}$
	24	$.35\cdot10^{-6}$	$.35\cdot10^{-6}$	$.27\cdot10^{-3}$	$.26\cdot10^{-3}$	$.32\cdot10^{-6}$
SYS1	8	$.29\cdot10^{-3}$	$.28\cdot10^{-3}$	$.21\cdot10^{-2}$	$.24\cdot10^{-2}$	$.34\cdot10^{-3}$
	12	$.52\cdot10^{-3}$	$.45\cdot10^{-3}$	$.37\cdot10^{-2}$	$.89\cdot10^{-2}$	$.39\cdot10^{-3}$
	16	$.48\cdot10^{-3}$	$.47\cdot10^{-3}$	$.36\cdot10^{-2}$	$.47\cdot10^{-2}$	$.69\cdot10^{-3}$
	20	$.52\cdot10^{-3}$	$.52\cdot10^{-3}$	$.44\cdot10^{-2}$	$.51\cdot10^{-2}$	$.67\cdot10^{-3}$
	24	$.76\cdot10^{-3}$	$.73\cdot10^{-3}$	$.60\cdot10^{-2}$	$.61\cdot10^{-2}$	$.57\cdot10^{-3}$
SYS2	8	$.14\cdot10^{-2}$	$.15\cdot10^{-2}$	$.76\cdot10^{-2}$	$.23\cdot10^{-1}$	$.12\cdot10^{-2}$
	12	$.23\cdot10^{-2}$	$.22\cdot10^{-2}$	$.25\cdot10^{-1}$	$.55\cdot10^{-1}$	$.19\cdot10^{-2}$
	16	$.31\cdot10^{-2}$	$.27\cdot10^{-2}$	$.25\cdot10^{-1}$	$.39\cdot10^{-1}$	$.16\cdot10^{-2}$
	20	$.19\cdot10^{-2}$	$.16\cdot10^{-2}$	$.28\cdot10^{-1}$	$.42\cdot10^{-1}$	$.15\cdot10^{-2}$
	24	$.13\cdot10^{-2}$	$.13\cdot10^{-2}$	$.29\cdot10^{-1}$	$.46\cdot10^{-1}$	$.13\cdot10^{-2}$
SYS3	8	$.56\cdot10^{-3}$	$.76\cdot10^{-3}$	$.22\cdot10^{-2}$	$.19\cdot10^{-1}$	$.44\cdot10^{-3}$
	12	$.70\cdot10^{-3}$	$.63\cdot10^{-3}$	$.68\cdot10^{-2}$	$.75\cdot10^{-2}$	$.60\cdot10^{-3}$
	16	$.84\cdot10^{-3}$	$.83\cdot10^{-3}$	$.46\cdot10^{-2}$	$.40\cdot10^{-2}$	$.96\cdot10^{-3}$
	20	$.89\cdot10^{-3}$	$.89\cdot10^{-3}$	$.71\cdot10^{-2}$	$.51\cdot10^{-2}$	$.80\cdot10^{-3}$
	24	$.98\cdot10^{-3}$	$.94\cdot10^{-3}$	$.64\cdot10^{-2}$	$.66\cdot10^{-2}$	$.78\cdot10^{-3}$

Fig-5.4.14- Relative error in the realization part of the applied algorithms; MIMO systems; $\sigma=\sigma_1$; NOM=8-24.

ERRB; MIMO systems; $\sigma=\sigma_2$

$\sigma=\sigma_2$	NOM	REHA	REHAK	REPAH	REPAK	REPHH
SYS0	8	0.126	0.128	0.180	0.273	$.86\cdot10^{-1}$
	12	0.197	0.193	0.415	0.426	0.193
	16	0.279	0.279	0.457	0.463	0.206
	20	0.303	0.309	0.523	0.521	0.219
	24	0.237	0.235	0.578	0.599	0.239
SYS1	8	$.62\cdot10^{-1}$	$.62\cdot10^{-1}$	0.136	0.181	$.67\cdot10^{-1}$
	12	$.90\cdot10^{-1}$	$.87\cdot10^{-1}$	0.243	0.282	$.88\cdot10^{-1}$
	16	$.77\cdot10^{-1}$	$.77\cdot10^{-1}$	0.343	0.371	1.000
	20	$.76\cdot10^{-1}$	$.76\cdot10^{-1}$	0.365	0.438	$.96\cdot10^{-1}$
	24	$.97\cdot10^{-1}$	$.97\cdot10^{-1}$	0.395	0.580	$.81\cdot10^{-1}$
SYS2	8	$.33\cdot10^{-1}$	$.36\cdot10^{-1}$	0.126	0.156	$.29\cdot10^{-1}$
	12	$.80\cdot10^{-1}$	$.74\cdot10^{-1}$	0.229	0.468	$.65\cdot10^{-1}$
	16	$.52\cdot10^{-1}$	$.52\cdot10^{-1}$	0.285	0.364	$.64\cdot10^{-1}$
	20	$.60\cdot10^{-1}$	$.60\cdot10^{-1}$	0.308	0.383	$.59\cdot10^{-1}$
	24	$.74\cdot10^{-1}$	$.74\cdot10^{-1}$	0.324	0.475	$.52\cdot10^{-1}$
SYS3	8	0.102	0.142	0.264	0.365	$.77\cdot10^{-1}$
	12	0.164	0.152	0.369	0.414	0.134
	16	0.143	0.141	0.455	0.410	0.150
	20	0.169	0.169	0.458	0.434	0.157
	24	0.180	0.183	0.471	0.628	0.180

Fig-5.4.15- Relative error in the realization part of the applied algorithms; MIMO systems; $\sigma=\sigma_2$; NOM=8-24.

ERR1; MIMO systems; $\sigma=\sigma_1$

$\sigma=\sigma_1$	NOM	REHA	REHAK	REPAH	REPAK	REPHH
SYS0	8	$.63\cdot10^{-6}$	$.63\cdot10^{-6}$	$.28\cdot10^{-5}$	$.16\cdot10^{-3}$	$.64\cdot10^{-6}$
	12	$.64\cdot10^{-6}$	$.64\cdot10^{-6}$	$.29\cdot10^{-5}$	$.27\cdot10^{-5}$	$.64\cdot10^{-6}$
	16	$.64\cdot10^{-6}$	$.64\cdot10^{-6}$	$.24\cdot10^{-3}$	$.18\cdot10^{-3}$	$.66\cdot10^{-6}$
	20	$.64\cdot10^{-6}$	$.64\cdot10^{-6}$	$.31\cdot10^{-3}$	$.26\cdot10^{-3}$	$.67\cdot10^{-6}$
	24	$.63\cdot10^{-6}$	$.63\cdot10^{-6}$	$.27\cdot10^{-3}$	$.26\cdot10^{-3}$	$.65\cdot10^{-6}$
SYS1	8	$.78\cdot10^{-3}$	$.81\cdot10^{-3}$	$.17\cdot10^{-2}$	$.23\cdot10^{-2}$	$.78\cdot10^{-3}$
	12	$.74\cdot10^{-3}$	$.66\cdot10^{-3}$	$.23\cdot10^{-2}$	$.74\cdot10^{-2}$	$.81\cdot10^{-3}$
	16	$.88\cdot10^{-3}$	$.85\cdot10^{-3}$	$.21\cdot10^{-2}$	$.27\cdot10^{-2}$	$.91\cdot10^{-3}$
	20	$.87\cdot10^{-3}$	$.84\cdot10^{-3}$	$.28\cdot10^{-2}$	$.27\cdot10^{-2}$	$.98\cdot10^{-3}$
	24	$.85\cdot10^{-3}$	$.88\cdot10^{-3}$	$.41\cdot10^{-2}$	$.38\cdot10^{-2}$	$.10\cdot10^{-2}$
SYS2	8	$.29\cdot10^{-2}$	$.31\cdot10^{-2}$	$.85\cdot10^{-2}$	$.18\cdot10^{-1}$	$.34\cdot10^{-2}$
	12	$.54\cdot10^{-2}$	$.55\cdot10^{-2}$	$.20\cdot10^{-1}$	$.50\cdot10^{-1}$	$.53\cdot10^{-2}$
	16	$.34\cdot10^{-2}$	$.32\cdot10^{-2}$	$.16\cdot10^{-1}$	$.28\cdot10^{-1}$	$.24\cdot10^{-2}$
	20	$.28\cdot10^{-2}$	$.25\cdot10^{-2}$	$.17\cdot10^{-1}$	$.28\cdot10^{-1}$	$.29\cdot10^{-2}$
	24	$.23\cdot10^{-2}$	$.24\cdot10^{-2}$	$.17\cdot10^{-1}$	$.31\cdot10^{-1}$	$.26\cdot10^{-2}$
SYS3	8	$.92\cdot10^{-3}$	$.11\cdot10^{-2}$	$.17\cdot10^{-2}$	$.19\cdot10^{-1}$	$.92\cdot10^{-3}$
	12	$.76\cdot10^{-3}$	$.74\cdot10^{-3}$	$.46\cdot10^{-2}$	$.55\cdot10^{-2}$	$.72\cdot10^{-3}$
	16	$.75\cdot10^{-3}$	$.73\cdot10^{-3}$	$.24\cdot10^{-2}$	$.13\cdot10^{-2}$	$.77\cdot10^{-3}$
	20	$.76\cdot10^{-3}$	$.74\cdot10^{-3}$	$.47\cdot10^{-2}$	$.21\cdot10^{-2}$	$.76\cdot10^{-3}$
	24	$.73\cdot10^{-3}$	$.76\cdot10^{-3}$	$.35\cdot10^{-2}$	$.24\cdot10^{-2}$	$.86\cdot10^{-3}$

Fig-5.4.16- Overall relative error; MIMO systems; $\sigma=\sigma_1$; NOM=8-24.

ERR1; MIMO systems; $\sigma=\sigma_2$

$\sigma=\sigma_2$	NOM	REHA	REHAK	REPAH	REPAK	REPHH
SYS0	8	0.150	0.158	0.110	$.52\cdot10^{-1}$	0.141
	12	0.182	0.195	$.57\cdot10^{-1}$	$.87\cdot10^{-1}$	$.79\cdot10^{-1}$
	16	0.210	0.217	$.84\cdot10^{-1}$	0.116	0.169
	20	0.185	0.192	$.79\cdot10^{-1}$	$.87\cdot10^{-1}$	0.107
	24	0.190	0.188	$.81\cdot10^{-1}$	$.88\cdot10^{-1}$	0.199
SYS1	8	$.50\cdot10^{-1}$	$.55\cdot10^{-1}$	$.47\cdot10^{-1}$	$.86\cdot10^{-1}$	$.57\cdot10^{-1}$
	12	$.61\cdot10^{-1}$	$.64\cdot10^{-1}$	$.85\cdot10^{-1}$	0.107	$.65\cdot10^{-1}$
	16	$.54\cdot10^{-1}$	$.54\cdot10^{-1}$	0.155	0.142	$.57\cdot10^{-1}$
	20	$.45\cdot10^{-1}$	$.45\cdot10^{-1}$	0.145	0.156	$.53\cdot10^{-1}$
	24	$.68\cdot10^{-1}$	$.68\cdot10^{-1}$	0.153	0.233	$.59\cdot10^{-1}$
SYS2	8	$.56\cdot10^{-1}$	$.62\cdot10^{-1}$	$.74\cdot10^{-1}$	$.99\cdot10^{-1}$	$.58\cdot10^{-1}$
	12	$.64\cdot10^{-1}$	$.66\cdot10^{-1}$	$.83\cdot10^{-1}$	0.360	$.55\cdot10^{-1}$
	16	$.75\cdot10^{-1}$	$.75\cdot10^{-1}$	0.118	0.164	$.64\cdot10^{-1}$
	20	$.69\cdot10^{-1}$	$.68\cdot10^{-1}$	0.120	0.169	$.62\cdot10^{-1}$
	24	$.70\cdot10^{-1}$	$.71\cdot10^{-1}$	0.128	0.225	$.62\cdot10^{-1}$
SYS3	8	$.79\cdot10^{-1}$	0.105	0.164	0.264	0.102
	12	0.105	0.103	0.163	0.211	0.104
	16	0.107	0.105	0.241	0.183	$.73\cdot10^{-1}$
	20	0.114	0.113	0.200	0.143	0.120
	24	0.178	0.180	0.200	0.209	0.193

Fig-5.4.17- Overall relative error; MIMO systems; $\sigma=\sigma_2$; NOM=8-24.

AVER1; MIMO systems; $\sigma=\sigma_1$

$\sigma = \sigma_1$	NOM	REHA	REHAK	REPAH	REPAK	REPHH
SYS0	8	$.31\cdot10^{-7}$	$.31\cdot10^{-7}$	$.14\cdot10^{-6}$	$.77\cdot10^{-5}$	$.31\cdot10^{-7}$
	12	$.32\cdot10^{-7}$	$.32\cdot10^{-7}$	$.14\cdot10^{-6}$	$.13\cdot10^{-6}$	$.31\cdot10^{-7}$
	16	$.31\cdot10^{-7}$	$.31\cdot10^{-7}$	$.12\cdot10^{-4}$	$.90\cdot10^{-5}$	$.32\cdot10^{-7}$
	20	$.31\cdot10^{-7}$	$.31\cdot10^{-7}$	$.15\cdot10^{-4}$	$.13\cdot10^{-4}$	$.33\cdot10^{-7}$
	24	$.31\cdot10^{-7}$	$.31\cdot10^{-7}$	$.13\cdot10^{-4}$	$.13\cdot10^{-4}$	$.32\cdot10^{-7}$
SYS1	8	$.25\cdot10^{-3}$	$.38\cdot10^{-3}$	$.61\cdot10^{-2}$	$.15\cdot10^{-2}$	$.20\cdot10^{-3}$
	12	$.15\cdot10^{-3}$	$.94\cdot10^{-4}$	$.78\cdot10^{-3}$	$.18\cdot10^{-2}$	$.17\cdot10^{-3}$
	16	$.13\cdot10^{-3}$	$.12\cdot10^{-3}$	$.35\cdot10^{-3}$	$.40\cdot10^{-3}$	$.13\cdot10^{-3}$
	20	$.11\cdot10^{-3}$	$.11\cdot10^{-3}$	$.39\cdot10^{-3}$	$.35\cdot10^{-3}$	$.12\cdot10^{-3}$
	24	$.10\cdot10^{-3}$	$.11\cdot10^{-3}$	$.50\cdot10^{-3}$	$.47\cdot10^{-3}$	$.12\cdot10^{-3}$
SYS2	8	$.97\cdot10^{-3}$	$.96\cdot10^{-3}$	$.84\cdot10^{-2}$	$.45\cdot10^{-2}$	$.12\cdot10^{-2}$
	12	$.11\cdot10^{-2}$	$.11\cdot10^{-2}$	$.90\cdot10^{-2}$	$.12\cdot10^{-1}$	$.11\cdot10^{-2}$
	16	$.52\cdot10^{-3}$	$.50\cdot10^{-3}$	$.24\cdot10^{-2}$	$.44\cdot10^{-2}$	$.43\cdot10^{-3}$
	20	$.49\cdot10^{-3}$	$.45\cdot10^{-3}$	$.24\cdot10^{-2}$	$.41\cdot10^{-2}$	$.52\cdot10^{-3}$
	24	$.33\cdot10^{-3}$	$.34\cdot10^{-3}$	$.24\cdot10^{-2}$	$.42\cdot10^{-2}$	$.37\cdot10^{-3}$
SYS3	8	$.85\cdot10^{-4}$	$.10\cdot10^{-3}$	$.17\cdot10^{-2}$	$.57\cdot10^{-2}$	$.11\cdot10^{-3}$
	12	$.67\cdot10^{-4}$	$.58\cdot10^{-4}$	$.83\cdot10^{-3}$	$.80\cdot10^{-3}$	$.63\cdot10^{-4}$
	16	$.68\cdot10^{-4}$	$.66\cdot10^{-4}$	$.25\cdot10^{-3}$	$.13\cdot10^{-3}$	$.69\cdot10^{-4}$
	20	$.71\cdot10^{-4}$	$.69\cdot10^{-4}$	$.47\cdot10^{-3}$	$.20\cdot10^{-3}$	$.70\cdot10^{-4}$
	24	$.66\cdot10^{-4}$	$.69\cdot10^{-4}$	$.32\cdot10^{-3}$	$.22\cdot10^{-3}$	$.78\cdot10^{-4}$

Fig-5.4.18- Overall average error; MIMO systems; $\sigma=\sigma_1$; NOM=8-24.

AVER1; MIMO systems; $\sigma=\sigma_2$

$\sigma = \sigma_2$	NOM	REHA	REHAK	REPAH	REPAK	REPHH
SYS0	8	$.74\cdot10^{-2}$	$.78\cdot10^{-2}$	$.54\cdot10^{-2}$	$.26\cdot10^{-2}$	$.69\cdot10^{-2}$
	12	$.89\cdot10^{-2}$	$.95\cdot10^{-2}$	$.28\cdot10^{-2}$	$.43\cdot10^{-2}$	$.39\cdot10^{-2}$
	16	$.10\cdot10^{-1}$	$.11\cdot10^{-1}$	$.41\cdot10^{-2}$	$.57\cdot10^{-2}$	$.83\cdot10^{-2}$
	20	$.90\cdot10^{-2}$	$.94\cdot10^{-2}$	$.39\cdot10^{-2}$	$.43\cdot10^{-2}$	$.53\cdot10^{-2}$
	24	$.93\cdot10^{-2}$	$.92\cdot10^{-2}$	$.40\cdot10^{-2}$	$.43\cdot10^{-2}$	$.97\cdot10^{-2}$
SYS1	8	$.81\cdot10^{-2}$	$.77\cdot10^{-2}$	$.76\cdot10^{-2}$	$.14\cdot10^{-1}$	$.81\cdot10^{-2}$
	12	$.76\cdot10^{-2}$	$.84\cdot10^{-2}$	$.12\cdot10^{-1}$	$.15\cdot10^{-1}$	$.81\cdot10^{-2}$
	16	$.65\cdot10^{-2}$	$.65\cdot10^{-2}$	$.20\cdot10^{-1}$	$.18\cdot10^{-1}$	$.68\cdot10^{-2}$
	20	$.54\cdot10^{-2}$	$.54\cdot10^{-2}$	$.18\cdot10^{-1}$	$.19\cdot10^{-1}$	$.66\cdot10^{-2}$
	24	$.81\cdot10^{-2}$	$.81\cdot10^{-2}$	$.18\cdot10^{-1}$	$.28\cdot10^{-1}$	$.70\cdot10^{-2}$
SYS2	8	$.13\cdot10^{-1}$	$.20\cdot10^{-1}$	$.11\cdot10^{-1}$	$.20\cdot10^{-1}$	$.11\cdot10^{-1}$
	12	$.90\cdot10^{-2}$	$.95\cdot10^{-2}$	$.14\cdot10^{-1}$	0.181	$.84\cdot10^{-2}$
	16	$.11\cdot10^{-1}$	$.11\cdot10^{-1}$	$.17\cdot10^{-1}$	$.31\cdot10^{-1}$	$.93\cdot10^{-2}$
	20	$.98\cdot10^{-2}$	$.96\cdot10^{-2}$	$.17\cdot10^{-1}$	$.26\cdot10^{-1}$	$.90\cdot10^{-2}$
	24	$.96\cdot10^{-2}$	$.98\cdot10^{-2}$	$.18\cdot10^{-1}$	$.31\cdot10^{-1}$	$.85\cdot10^{-2}$
SYS3	8	$.75\cdot10^{-2}$	$.11\cdot10^{-1}$	$.19\cdot10^{-1}$	$.31\cdot10^{-1}$	$.94\cdot10^{-2}$
	12	$.92\cdot10^{-2}$	$.98\cdot10^{-2}$	$.18\cdot10^{-1}$	$.22\cdot10^{-1}$	$.92\cdot10^{-2}$
	16	$.96\cdot10^{-2}$	$.94\cdot10^{-2}$	$.23\cdot10^{-1}$	$.18\cdot10^{-1}$	$.66\cdot10^{-2}$
	20	$.10\cdot10^{-1}$	$.10\cdot10^{-1}$	$.18\cdot10^{-1}$	$.13\cdot10^{-1}$	$.11\cdot10^{-1}$
	24	$.16\cdot10^{-1}$	$.16\cdot10^{-1}$	$.18\cdot10^{-1}$	$.19\cdot10^{-1}$	$.17\cdot10^{-1}$

Fig-5.4.19- Overall average error; MIMO systems; $\sigma=\sigma_2$; NOM=8-24.

Eigenvalues of SYSO; $\sigma=\sigma_1$

$\sigma=\sigma_1$	NOM	λ_1	λ_2	λ_3	λ_4
REHA	8	0.397	0.267	0.215	0.122
	12	0.397	0.268	0.215	0.122
	16	0.397	0.268	0.214	0.121
	20	0.397	0.269	0.213	0.121
	24	0.398	0.272	0.213	0.120
REHAK	8	0.397	0.268	0.215	0.122
	12	0.397	0.268	0.215	0.122
	16	0.397	0.268	0.214	0.121
	20	0.397	0.269	0.213	0.121
	24	0.398	0.272	0.213	0.120
REPAH	8	0.386	0.164	0.225	$\pm$ j 0.116
	12	0.439	0.351	0.187	0.047
	16	0.369	$\pm$ j 0.024	0.211	-0.124
	20	0.353	$\pm$ j 0.031	0.227	-0.111
	24	0.354	$\pm$ j 0.053	0.237	-0.091
REPAK	8	0.357	0.278	0.134	$-.61 \cdot 10^{-6}$
	12	0.427	0.348	0.184	$.38 \cdot 10^{-1}$
	16	0.374	$\pm$ j 0.031	0.212	-0.228
	20	0.361	$\pm$ j 0.047	0.235	-0.104
	24	0.369	$\pm$ j 0.064	0.236	-0.065
REPHH	8	0.397	0.267	0.215	0.122
	12	0.397	0.268	0.214	0.122
	16	0.397	0.261	0.218	0.124
	20	0.397	0.260	0.219	0.124
	24	0.397	0.266	0.217	0.122

Fig-5.4.20- Realized eigenvalues of SYSO; $\sigma=\sigma_1$; NOM=8-24.

Eigenvalues of SYSO; $\sigma=\sigma_2$

$\sigma=\sigma_2$	NOM	λ_1	λ_2	λ_3	λ_4
REHA	8	0.571	-0.348	-0.043	$\pm$ j 0.527
	12	0.505	-0.609	-0.020	$\pm$ j 0.670
	16	0.574	-0.623	0.019	$\pm$ j 0.657
	20	0.650	-0.644	-0.027	$\pm$ j 0.637
	24	0.521	-0.944	0.193	$\pm$ j 0.599
REHAK	8	0.580	-0.355	-0.058	$\pm$ j 0.526
	12	0.494	-0.706	0.007	$\pm$ j 0.676
	16	0.580	-0.669	0.024	$\pm$ j 0.665
	20	0.674	-0.674	-0.027	$\pm$ j 0.629
	24	0.509	-0.963	0.201	$\pm$ j 0.597
REPAH	8	0.526	-0.111	-0.177	$\pm$ j 0.505
	12	0.282	$\pm$ j 0.107	-0.196	$\pm$ j 0.073
	16	0.595	-0.207	0.038	$\pm$ j 0.441
	20	0.478	0.257	-0.264	$\pm$ j 0.172
	24	0.699	0.277	-0.115	$\pm$ j 0.156
REPAK	8	0.349	-0.283	$.76 \cdot 10^{-3}$	$.15 \cdot 10^{-4}$
	12	0.338	$\pm$ j 0.084	-0.690	-0.062
	16	0.674	-0.245	0.028	$\pm$ j 0.460
	20	0.629	0.198	-0.094	$\pm$ j 0.056
	24	0.877	0.245	-0.045	$\pm$ j 0.036
REPHH	8	0.560	-0.294	-0.140	$\pm$ j 0.530
	12	0.564	-0.088	-0.321	$\pm$ j 0.652
	16	0.624	0.416	-0.119	$\pm$ j 0.485
	20	0.512	-0.115	0.116	$\pm$ j 0.859
	24	0.641	$\pm$ j 0.018	0.230	$\pm$ j 0.659

Fig-5.4.21- Realized eigenvalues of SYSO; $\sigma=\sigma_2$; NOM=8-24.

Eigenvalues of SYS1; $\sigma=\sigma_1$

$\sigma=\sigma_1$	NOM	λ_1	λ_2	λ_3	λ_4
REHA	8	0.875	0.718	0.581	0.227
	12	0.882	0.698	0.577	-0.221
	16	0.888	0.703	0.577	-0.107
	20	0.890	0.722	0.595	-0.372
	24	0.899	0.736	0.598	-0.952
REHAK	8	0.869	0.715	0.575	0.316
	12	0.890	0.705	0.578	-0.218
	16	0.889	0.704	0.578	-0.136
	20	0.891	0.723	0.595	-0.394
	24	0.898	0.735	0.599	-0.973
REPAH	8	0.813	0.618	1.05	-0.155
	12	0.857	0.671	0.570	-0.263
	16	0.879	0.691	0.578	-0.182
	20	0.877	0.693	0.599	-0.113
	24	0.867	0.668	0.557	-0.261
REPAK	8	0.839	0.644	0.330	0.000
	12	0.845	0.701	0.575	-0.243
	16	0.881	0.702	0.578	-0.282
	20	0.886	0.718	0.605	-0.238
	24	0.868	0.649	0.565	-0.210
REPHH	8	0.881	0.713	0.592	0.054
	12	0.881	0.683	0.578	-0.389
	16	0.889	0.709	0.587	-0.770
	20	0.893	0.726	0.605	-0.976
	24	0.901	0.742	0.605	-0.975

Fig-5.4.22- Realized Eigenvalues of SYS1; $\sigma=\sigma_1$; NOM=8-24.

Eigenvalues of SYS1; $\sigma=\sigma_2$

$\sigma=\sigma_2$	NOM	λ_1	λ_2	λ_3	λ_4
REHA	8	0.878	0.736	-0.106	± j 0.643
	12	0.889	0.649	-0.206	± j 0.772
	16	0.886	0.650	-0.152	± j 0.687
	20	0.873	0.613	-0.156	± j 0.474
	24	0.867	-0.934	0.358	± j 0.192
REHAK	8	0.865	0.735	-0.118	± j 0.638
	12	0.904	0.646	-0.197	± j 0.784
	16	0.889	0.653	-0.158	± j 0.706
	20	0.874	0.623	-0.164	± j 0.472
	24	0.867	-0.954	0.357	± j 0.202
REPAH	8	0.782	± j 0.044	-0.363	0.097
	12	0.809	0.617	-0.346	± j 0.296
	16	0.753	0.656	-0.294	± j 0.422
	20	0.781	0.583	0.207	-0.244
	24	0.784	0.590	-0.044	± j 0.169
REPAK	8	0.828	-0.627	0.490	$-.10\cdot10^{-6}$
	12	0.784	0.683	-0.656	-0.342
	16	0.914	0.626	-0.338	± j 0.381
	20	0.927	0.582	0.318	-0.364
	24	0.948	0.603	0.095	-0.126
REPHH	8	0.879	0.745	-0.160	± j 0.490
	12	0.889	0.679	-0.321	± j 0.800
	16	0.883	-0.823	0.665	± j 0.129
	20	0.859	± j 0.055	0.637	-0.894
	24	0.869	-0.980	0.519	0.303

Fig-5.4.23- Realized eigenvalues of SYS1; $\sigma=\sigma_2$; NOM=8-24.

Eigenvalues of SYS2; $\sigma=\sigma_1$

$\sigma=\sigma_1$	NOM	λ_1	λ_2	λ_3	λ_4
REHA	8	0.887	0.871	0.396	0.213
	12	0.881	0.860	0.067	$\pm j$ 0.147
	16	0.886	$\pm j$ 0.022	0.856	0.229
	20	0.888	$\pm j$ 0.047	0.842	0.191
	24	0.881	$\pm j$ 0.027	0.851	0.202
REHAK	8	0.890	0.870	0.403	0.248
	12	0.876	0.871	0.086	$\pm j$ 0.176
	16	0.891	$\pm j$ 0.023	0.873	0.232
	20	0.893	$\pm j$ 0.044	0.845	0.191
	24	0.881	$\pm j$ 0.029	0.850	0.202
REPAH	8	0.966	0.888	0.352	-0.279
	12	0.879	$\pm j$ 0.100	0.772	0.138
	16	0.886	0.042	0.714	$\pm j$ 0.047
	20	0.883	0.049	0.731	$\pm j$ 0.043
	24	0.879	0.074	0.769	$\pm j$ 0.048
REPAK	8	0.900	0.630	0.141	$.30 \cdot 10^{-7}$
	12	0.931	0.173	0.625	$\pm j$ 0.114
	16	0.898	0.747	0.585	0.071
	20	0.899	0.757	0.655	0.106
	24	0.895	0.126	0.761	$\pm j$ 0.045
REPHH	8	0.887	0.875	0.251	-0.311
	12	0.874	$\pm j$ 0.011	0.228	-0.647
	16	0.898	$\pm j$ 0.031	0.840	0.216
	20	0.899	$\pm j$ 0.044	0.852	0.192
	24	0.879	$\pm j$ 0.035	0.854	0.203

Fig-5.4.24- Realized eigenvalues of SYS2; $\sigma=\sigma_1$; NOM=8-24.

Eigenvalues of SYS2; $\sigma=\sigma_2$

$\sigma=\sigma_2$	NOM	λ_1	λ_2	λ_3	λ_4
REHA	8	0.938	0.852	-0.058	$\pm j$ 0.596
	12	0.858	$\pm j$ 0.025	-0.130	$\pm j$ 0.733
	16	0.890	$\pm j$ 0.052	0.003	$\pm j$ 0.625
	20	0.874	$\pm j$ 0.056	-0.238	$\pm j$ 0.404
	24	0.855	$\pm j$ 0.046	0.089	-0.931
REHAK	8	0.957	0.848	-0.071	$\pm j$ 0.546
	12	0.871	$\pm j$ 0.039	-0.112	$\pm j$ 0.745
	16	0.890	$\pm j$ 0.053	$.27 \cdot 10^{-3}$	$\pm j$ 0.634
	20	0.875	$\pm j$ 0.055	-0.241	$\pm j$ 0.400
	24	0.857	$\pm j$ 0.048	0.085	-0.949
REPAH	8	0.894	0.460	-0.137	$\pm j$ 0.278
	12	0.834	0.683	-0.059	$\pm j$ 0.352
	16	0.865	0.512	-0.189	$\pm j$ 0.287
	20	0.845	0.587	-0.024	$\pm j$ 0.046
	24	0.815	0.659	0.212	-0.057
REPAK	8	0.805	0.428	-0.418	$.60 \cdot 10^{-7}$
	12	0.996	0.566	-0.404	$\pm j$ 0.520
	16	0.948	0.564	-0.194	$\pm j$ 0.282
	20	0.940	0.594	0.284	-0.235
	24	0.930	0.670	0.397	-0.144
REPHH	8	0.927	0.841	-0.053	$\pm j$ 0.562
	12	0.868	$\pm j$ 0.049	-0.310	$\pm j$ 0.396
	16	0.880	$\pm j$ 0.054	-0.239	$\pm j$ 0.267
	20	0.886	$\pm j$ 0.068	-0.520	0.080
	24	0.882	$\pm j$ 0.068	0.126	-0.983

Fig-5.4.25- Realized eigenvalues of SYS2; $\sigma=\sigma_2$; NOM=8-24.

Eigenvalues of SYS3; $\sigma=\sigma_1$

$\sigma=\sigma_1$	NOM	λ_1	λ_2	λ_3	λ_4
REHA	8	0.903	0.230	0.123	-0.309
	12	0.900	0.182	0.028	-0.304
	16	0.899	0.182	0.063	-0.236
	20	0.897	0.184	0.099	-0.535
	24	0.898	0.184	0.106	-0.931
REHAK	8	0.900	0.233	0.130	-0.355
	12	0.903	0.182	0.041	-0.330
	16	0.900	0.182	0.069	-0.270
	20	0.897	0.183	0.099	-0.558
	24	0.898	0.184	0.107	-0.952
REPAH	8	0.934	0.263	0.184	-0.272
	12	0.874	0.165	0.059	-0.051
	16	0.893	0.182	0.141	-0.316
	20	0.884	0.259	0.171	-0.033
	24	0.890	0.051	0.188	$\pm j$ 0.020
REPAK	8	0.804	0.185	-0.040	$.32\cdot10^{-7}$
	12	0.882	0.106	0.159	$\pm j$ 0.085
	16	0.906	-0.449	0.179	$\pm j$ 0.011
	20	0.892	0.176	0.219	$\pm j$ 0.089
	24	0.907	0.178	0.152	$\pm j$ 0.143
REPHH	8	0.907	0.214	0.139	-0.632
	12	0.902	0.187	0.115	-0.611
	16	0.900	0.181	0.108	-0.951
	20	0.897	0.183	0.096	-0.760
	24	0.898	0.185	0.088	-0.954

Fig-5.4.26- Realized eigenvalues of SYS3; $\sigma=\sigma_1$; NOM=8-24.

Eigenvalues of SYS3; $\sigma=\sigma_2$

$\sigma=\sigma_2$	NOM	λ_1	λ_2	λ_3	λ_4
REHA	8	0.897	0.408	-0.073	$\pm j$ 0.443
	12	0.910	-0.002	-0.206	$\pm j$ 0.612
	16	0.911	-0.157	0.009	$\pm j$ 0.552
	20	0.880	-0.389	0.276	$\pm j$ 0.219
	24	0.872	-0.877	0.186	$\pm j$ 0.687
REHAK	8	0.838	0.392	-0.073	$\pm j$ 0.300
	12	0.931	-0.076	-0.184	$\pm j$ 0.592
	16	0.914	-0.191	0.006	$\pm j$ 0.537
	20	0.883	-0.387	0.280	$\pm j$ 0.122
	24	0.872	-0.899	0.188	$\pm j$ 0.691
REPAH	8	0.814	0.400	-0.300	$\pm j$ 0.148
	12	0.747	0.231	-0.142	$\pm j$ 0.114
	16	0.606	$\pm j$ 0.089	0.117	-0.391
	20	0.740	0.313	0.173	-0.377
	24	0.789	-0.251	0.187	$\pm j$ 0.130
REPAK	8	-0.873	0.391	0.213	$.10\cdot10^{-6}$
	12	0.645	-0.494	0.245	$\pm j$ 0.133
	16	0.747	0.535	0.181	-0.419
	20	0.823	0.411	0.220	-0.381
	24	0.956	-0.203	0.271	$\pm j$ 0.067
REPHH	8	0.904	0.141	-0.179	$\pm j$ 0.615
	12	0.917	0.204	-0.280	$\pm j$ 0.683
	16	0.920	0.655	0.170	-0.678
	20	0.861	$\pm j$ 0.061	0.157	-0.840
	24	0.879	-0.933	0.194	$\pm j$ 0.695

Fig-5.4.27- Realized eigenvalues of SYS3; $\sigma=\sigma_2$; NOM=8-24.

ERR1; SYS0; $\sigma=\sigma_2$; averaged results

NOM	REHA	REHAK	REPAH	REPAK	REPHH
8	0.145	0.182	0.113	0.212	0.135
12	0.151	0.172	$.81\cdot10^{-1}$	$.19\cdot10^{+2}$	0.127
16	0.167	0.173	$.84\cdot10^{-1}$	$.92\cdot10^{-1}$	0.152
20	0.168	0.171	$.78\cdot10^{-1}$	$.83\cdot10^{-1}$	0.164
24	0.221	0.227	$.80\cdot10^{-1}$	$.86\cdot10^{-1}$	0.201

Fig-5.4.28- Overall relative error in case of SYS0 and $\sigma=\sigma_2$; ERR1 averaged over 3 noise series; NOM=8-24.

AVER1; SYS0; $\sigma=\sigma_2$; averaged results

NOM	REHA	REHAK	REPAH	REPAK	REPHH
8	$.81\cdot10^{-2}$	$.21\cdot10^{-1}$	$.55\cdot10^{-2}$	$.10\cdot10^{-1}$	$.75\cdot10^{-2}$
12	$.84\cdot10^{-2}$	$.11\cdot10^{-1}$	$.44\cdot10^{-2}$	$.48\cdot10^{+4}$	$.76\cdot10^{-2}$
16	$.83\cdot10^{-2}$	$.86\cdot10^{-2}$	$.41\cdot10^{-2}$	$.45\cdot10^{-2}$	$.75\cdot10^{-2}$
20	$.82\cdot10^{-2}$	$.84\cdot10^{-2}$	$.38\cdot10^{-2}$	$.41\cdot10^{-2}$	$.80\cdot10^{-2}$
24	$.11\cdot10^{-1}$	$.11\cdot10^{-1}$	$.39\cdot10^{-2}$	$.42\cdot10^{-2}$	$.98\cdot10^{-2}$

Fig-5.4.29- Overall relative error in case of SYS0 and $\sigma=\sigma_2$; AVER1 averaged over 3 noise series; 5 algoriothms; NOM=8-24.

ERRO, AVERO; SISO systems

	σ	ERRO	AVERO
SYS4	σ_1	$.59\cdot10^{+1}$	0.120
	σ_2	$.12\cdot10^{+2}$	0.244
SYS5	σ_1	$.12\cdot10^{-2}$	$.71\cdot10^{-4}$
	σ_2	$.20\cdot10^{+1}$	0.120
	σ_3	$.82\cdot10^{+1}$	0.479

Fig-5.5.1- Relative and average error in input Markov parameter sequence; SISO systems; $\sigma=\sigma_1,\sigma_2,\sigma_3$; NOM=64.

Singular values of SYS4; averaged results

SYS4		$\tilde{\delta}_1$	$\tilde{\delta}_2$	$\tilde{\delta}_3$	$\tilde{\delta}_4$	$\tilde{\delta}_5$
σ_1	H	3.913	3.794	3.467	3.226	3.046
	P	1.903	1.656	1.347	0.929	0.731
σ_2	H	5.555	5.360	4.703	4.559	4.310
	P	2.604	2.235	1.930	1.319	1.002

Fig-5.5.2- Singular values of the noise disturbed Hankel and Page matrix in case of SYS4; $\sigma=\sigma_1,\sigma_2$; NOM=64; results averaged over 3 noise series.

Singular values of SYS5; averaged results

SYS5		$\tilde{\delta}_1$	$\tilde{\delta}_2$	$\tilde{\delta}_3$	$\tilde{\delta}_4$	$\tilde{\delta}_5$
σ_1	H	3.901	0.127	0.094	0.091	0.078
	P	1.949	0.047	0.033	0.030	0.018
σ_2	H	4.554	3.831	3.694	3.224	3.071
	P	2.393	1.689	1.334	0.987	0.800
σ_3	H	7.785	7.518	6.641	6.433	6.034
	P	3.729	3.182	2.690	1.884	1.443

Fig-5.5.3- Singular values of the noise disturbed Hankel and Page matrix in case of SYS5; $\sigma=\sigma_1,\sigma_2,\sigma_3$; NOM=64; results are averaged over 3 noise series.

ERRF; SISO systems

ERRF	σ	H	P	P/H
SYS4	σ_1	1.40	2.71	1.27
	σ_2	1.80	5.04	1.54
SYS5	σ_1	$.53\cdot10^{-4}$	$.53\cdot10^{-3}$	$.49\cdot10^{-4}$
	σ_2	1.38	1.31	0.17
	σ_3	2.71	5.51	1.84

Fig-5.5.4- Relative error in noise filtered Markov parameters for Hankel, Page and combined algorithms; SISO systems; $\sigma=\sigma_1,\sigma_2,\sigma_3$; NOM=64.

NRF; SISO systems; averaged results

NRF	σ	H	P	P/H
SYS4	σ_1	0.15	0.42	0.17
	σ_2	0.14	0.40	0.11
SYS5	σ_1	0.06	0.58	0.06
	σ_2	0.30	0.60	0.11
	σ_3	0.29	0.61	0.21

Fig-5.5.5- Noise reduction factor for Hankel, Page and combined algorithms; SISO systems; $\sigma=\sigma_1,\sigma_2,\sigma_3$; NOM=64; NRF averaged over 3 noise series.

ERRB; SISO systems; averaged results

ERRB		REHA	REHAK	REPAH	REPAK	REPHH
SYS4	σ_1	0.91	0.91	0.88	0.88	0.68
	σ_2	1.00	1.00	0.96	0.96	0.73
SYS5	σ_1	$.50\cdot10^{-4}$	$.62\cdot10^{-4}$	$.25\cdot10^{-2}$	$.50\cdot10^{+7}$	$.67\cdot10^{-4}$
	σ_2	0.35	0.35	0.71	1.19	0.18
	σ_3	0.83	0.84	0.92	0.92	0.51

Fig-5.5.6- Relative error in the realization part of the applied algorithms; SISO systems; $\sigma=\sigma_1,\sigma_2,\sigma_3$; NOM=64; ERRB averaged over 3 noise series.

ERR1; SISO systems; averaged results

ERR1		REHA	REHAK	REPAH	REPAK	REPHH
SYS4	σ_1	0.91	0.91	0.74	0.73	0.71
	σ_2	0.99	0.99	0.81	0.82	0.81
SYS5	σ_1	$.54 \cdot 10^{-4}$	$.54 \cdot 10^{-4}$	$.20 \cdot 10^{-2}$	$.50 \cdot 10^{+7}$	$.76 \cdot 10^{-4}$
	σ_2	0.40	0.40	0.32	1.40	0.12
	σ_3	1.20	1.20	0.63	0.72	0.89

Fig-5.5.7- Overall relative error; SISO systems; $\sigma=\sigma_1,\sigma_2,\sigma_3$; NOM=64; ERR1 averaged over 3 noise series.

AVER1; SISO systems; averaged results

AVER1		REHA	REHAK	REPAH	REPAK	REPHH
SYS4	σ_1	$.18 \cdot 10^{-1}$	$.18 \cdot 10^{-1}$	$.15 \cdot 10^{-1}$	$.15 \cdot 10^{-1}$	$.14 \cdot 10^{-1}$
	σ_2	$.20 \cdot 10^{-1}$	$.20 \cdot 10^{-1}$	$.16 \cdot 10^{-1}$	$.17 \cdot 10^{-1}$	$.16 \cdot 10^{-1}$
SYS5	σ_1	$.32 \cdot 10^{-5}$	$.32 \cdot 10^{-5}$	$.12 \cdot 10^{-3}$	$.45 \cdot 10^{+6}$	$.45 \cdot 10^{-5}$
	σ_2	$.23 \cdot 10^{-1}$	$.23 \cdot 10^{-1}$	$.18 \cdot 10^{-1}$	$.83 \cdot 10^{-1}$	$.69 \cdot 10^{-2}$
	σ_3	$.70 \cdot 10^{-1}$	$.71 \cdot 10^{-1}$	$.37 \cdot 10^{-1}$	$.42 \cdot 10^{-1}$	$.52 \cdot 10^{-1}$

Fig-5.5.8- Overall average error; SISO systems; $\sigma=\sigma_1,\sigma_2,\sigma_3$; NOM=64; AVER1 averaged over 3 noise series.

Eigenvalues of SISO systems

		REHA	REHAK	REPAH	REPAK	REPHH
SYS4	σ_1	-0.653	-0.653	0.159	0.167	0.777
	σ_2	-0.654	-0.654	-0.075	-0.098	-0.457
SYS5	σ_1	0.887 0.751	0.887 0.751	0.857 0.138	0.857 0.002	0.890 0.771
	σ_2	-0.599 $\pm$ j 0.766	-0.598 $\pm$ j 0.769	0.649 -0.587	0.650 -0.617	0.711 -0.381
	σ_3	-0.627 $\pm$ j 0.768	-0.626 $\pm$ j 0.770	0.505 -0.513	0.539 -0.523	-0.395 $\pm$ j 0.877

Fig-5.5.9- Realized eigenvalues of SISO systems; $\sigma=\sigma_1,\sigma_2,\sigma_3$; NOM=64.

REFERENCES.

- Chen, C.T. and C.A. Desoer (1968)
  A PROOF OF CONTROLLABILITY OF JORDAN FORM STATE EQUATIONS.
  IEEE Trans. Autom. Contr., AC-13(1968), no.2, p.195.

- Damen, A.A.H. and A.K. Hajdasinski (1982)
  PRACTICAL TESTS WITH DIFFERENT APPROXIMATE REALIZATIONS BASED ON THE SINGULAR VALUE DECOMPOSITION OF THE HANKEL MATRIX.
  Prepr./Proc. of the 6^{th} IFAC Symp. on Ident. and Syst. Param. Estim., Washington,1982.

- Damen, A.A.H., P.M.J. Van den Hof and A.K. Hajdasinski (1982a)
  THE PAGE MATRIX: AN EXCELLENT TOOL FOR NOISE FILTERING OF MARKOV PARAMETERS, ORDER TESTING AND REALIZATION.
  EUT-Report 82-E-127, Dept. Elec. Eng., Eindhoven University of Technology, 1982.

- Damen, A.A.H., P.M.J. Van den Hof and A.K. Hajdasinski (1982b)
  APPROXIMATE REALIZATION BASED UPON AN ALTERNATIVE TO THE HANKEL MATRIX: THE PAGE MATRIX.
  Systems and Control Letters, Vol.II (1982), No.4.

- Eykhoff, P. (1974)
  SYSTEM IDENTIFICATION: PARAMETER AND STATE ESTIMATION.
  London: Wiley and Sons, 1974.

- Gantmacher, F.R. (1959)
  THE THEORY OF MATRICES, VOL.1.
  New York: Chelsea, 1959.

- Hajdasinski, A.K. and A.A.H. Damen (1979)
  REALIZATION OF THE MARKOV PARAMETER SEQUENCES USING THE SINGULAR VALUE DECOMPOSITION OF THE HANKEL MATRIX.
  Dep. of Electr. Eng., Eindhoven University of Technology, 1979;
  TH-Report 79-E-95.

- Hajdasinski, A.K. (1980)
LINEAR MULTIVARIABLE SYSTEMS. Preliminary Problems in Mathematical Description, Modelling and Identification.
Dep. of Electr. Eng., Eindhoven University of Technology, 1980; TH-Report 80-E-106.

- Hautus, M.L.J. (1969)
CONTROLLABILITY AND OBSERVABILITY CONDITIONS OF LINEAR AUTONOMOUS SYSTEMS.
Proc. Kon. Akad. Wet. Ser. A, vol.72 (= Indag. Math., Vol. 31), (1969). p.444-448.

- Ho, L.B. and R.E. Kalman (1966)
EFFECTIVE CONSTRUCTION OF LINEAR STATE-VARIABLE MODELS FROM INPUT-OUTPUT FUNCTIONS.
Regelungstechnik, Vol.14(1966), p.545-548.

- Kam, J.J. van der and A.A.H. Damen (1978)
OBSERVABILITY OF ELECTRICAL HEART ACTIVITY STUDIED WITH THE SINGULAR VALUE DECOMPOSITION.
Dep. of Electr.Eng., Eindhoven University of Technology, 1978; TH-Report 78-E-81.

- Kung, S. (1978)
A NEW IDENTIFICATION AND MODEL REDUCTION ALGORITHM VIA SINGULAR VALUE DECOMPOSITIONS.
12-th Asilomar Conference on Circuits, Systems and Computers, Nov.6-8, 1978, Pacific Grove, California.

- Moore, B.C. (1981)
PRINCIPLE COMPONENT ANALYSIS IN LINEAR SYSTEMS: CONTROLLABILITY, OBSERVABILITY AND MODEL REDUCTION.
IEEE Trans. Autom. Control, AC-26(1981), no.1. p.17-32.

- Niederlinski, A. and A.K. Hajdasinski (1979)
MULTIVARIABLE SYSTEM IDENTIFICATION - A SURVEY.
Prepr./Proc. 5-th IFAC Symp. on Ident. and Syst. Param. Estim. Darmstadt, 1979, p.43-76.

- Silverman, L.M. (1971)
REALIZATION OF LINEAR DYNAMICAL SYSTEMS.
IEEE Trans. Autom. Control, AC-16(1971), no.6, p.554-567.

- Silverman, L.M. and M. Bettayeb (1980)
OPTIMAL APPROXIMATION OF LINEAR SYSTEMS.
Proc. 1980 Joint Autom. Contr. Conf., San Francisco, CA, Paper FA8.A, Aug.1980.

- Staar, J., J. Vandewalle and M. Wemans (1981a)
A GENERAL CLASS OF NUMERICALLY RELIABLE ALGORITHMS FOR THE REALIZATION OF TRUNCATED IMPULSE RESPONSES.
Prepr. 8-th IFAC World Congress,
Kyoto, 1981, Vol.I, p.13-18.

- Staar, J., J. Vandewalle and M. Wemans (1981b)
REALIZATION OF TRUNCATED IMPULSE RESPONSE SEQUENCES WITH PRESCRIBED UNCERTAINTY.
Prepr. 8-th IFAC World Congress,
Kyoto, 1981, Vol.I, p.7-12.

- Zee, G.A. van and O.H. Bosgra (1979)
THE USE OF REALIZATION THEORY IN THE ROBUST IDENTIFICATION OF MULTIVARIABLE SYSTEMS.
Prepr./Proc. 5-th IFAC Symp. Ident. and Syst. Param. Estim.
Darmstadt, 1979, p.477-484.

- Zeiger, H.P. and J. McEwen (1974)
APPROXIMATE LINEAR REALIZATIONS OF GIVEN DIMENSION VIA HO'S ALGORITHM.
IEEE Trans. Autom. Control, AC-19(1974), p.153.