MASTER

Matrix fraction descriptions of linear systems and their use in the instrumental variable parameter estimation method

van Helmont, J.B.

Award date:
1984

4783

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

MATRIX FRACTION DESCRIPTIONS OF LINEAR SYSTEMS AND THEIR USE IN THE INSTRUMENTAL VARIABLE PARAMETER ESTIMATION METHOD.

by Hans van Helmont

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 Oct. 1982 until Dec. 1983
in charge of Prof. dr. ir. P. Eykhoff
under supervision of dr. ir. A.J.W. van den Boom
and dr. ir. A.K. Hajdasinski

Contents

A multi-input multi-output system can be described with a rational transfer matrix, or alternatively as a socalled ARMA-model with the help of two polynomial matrices. The transfer matrix is equal to one polynomial matrix, premultiplied by the inverse of the other one.
The pair of polynomial matrices is called a matrix fraction description of the transfer matrix. With the help of equivalence relations, we can transform one of the polynomial matrices to a form, in which the degrees of the polynomials obey certain requirements. Several of these uniquely parametrized forms have been described along with the construction of these special forms from a given polynomial matrix.
The relation of the matrix fraction description to state-space models has been studied and a transformation from a matrix fraction description to a state-space model has been described.
The matrix fractions describe a model type for which parameter estimations have been performed. Based on this model an instrumental variable estimation procedure has been developed, which is related to the least-squares estimation procedure, but which gives unbiased parameter estimates in the presence of noise. By using this particular method there is no need to model the noise, which means a reduction of the complexity of the estimation procedure in comparison with other methods.
The usefulness of the matrix fraction description for transfer functions and the performance of the instrumental variable estimation procedure have been tested in a number of simulations.

Chapter 1: INTRODUCTION

In describing the behaviour of linear, time-invariant systems the state-space models are frequently used.
For single-input single-output systems (SISO) another approach to describe the dynamic behaviour is with the help of transfer functions.
In stead of giving a description in the time domain (as the state-space approach does), the transfer function gives a frequency-domain representation. While the extension of the state-space method for MIMO-systems is straightforward, the extension of the frequency domain approach to MIMO systems encounters various difficulties. One of the difficulties of the extension is (for both methods) the nonuniqueness of the description.
The so called matrix fraction description (MFD) of the transfer matrix can be considered as an extension of the SISO - ARMA model to the multivariable case. The MFD method is based on the decomposition of the transfer matrix (which in general is a rational matrix) into polynomial matrices,one of which is inverted.
Obviously the study of MFDs draws heavily on the theory of polynomial matrices, as e.g. the study of the state-space approach draws on the theory of real matrices.
Besides giving a useful tool in the study of MIMO-systems, the MFD approach has found applications for various areas of study, e.g. linear state variable feedback and realization of linear systems.

The aim of this investigation was, first of all, to give a description of the different forms of MFDs that are given in the literature, and to study the theory on which the matrix fractions are based, i.c. the theory of polynomial matrices. The relation and the transformation between state-space models and certain types of MFDs are given also. The second important part of this work is concerned with the use of MFDs in estimation procedures. A number of simulations were made to estimate the parameters of MFD models using an instrumental variable technique.

Presentation

The algebraic terms encountered in this work can be found in most textbooks on algebra. E.g. Birkhoff and MacLane.
Systems are defined over an arbitrary field F whose algebraic closure is denoted $\bar{F}$. The algebraic closure of a field is the set of all elements of F, plus the zeros of all polynomials that have coefficients in F.
In most cases we have F=R, the real numbers.
The notation $F[z]$ indicates the ring of polynomials in the indeterminate z. Likewise the notation F(z) indicates the field of rational functions in the indeterminate z.
Polynomials and rational functions will be assumed to have coefficients in F.
The notation a(z)/b(z) means that the polynomial a(z) divides the polynomial b(z). For example this means that z/(z-1)z indicates that the polynomial z is a factor of the polynomial z(z-1).
The matrix $[0]$ denotes a null-matrix of appropriate size,and can be found e.g. in composite matrices as $[A : 0]$.
The matrix I represents the identity matrix.

Chapter 2: MATRIX FRACTION DESCRIPTIONS

In this chapter we will define matrix fractions of transfer matrices and we will give the relation to other model forms. T(z) will be the transfer function matrix of a system having q outputs and p inputs, and it gives the relation between sequences of input-signals $\{u(k)\}$ and sequences of output-signals $\{y(k)\}$:

$$y(k) = T(z).u(k) \tag{2.1}$$

where y(k) represents the output- and u(k) the input-vector:

$$\begin{aligned} y^T(k) &= [y_1(k) \quad y_2(k) \quad \dots \quad y_q(k)] \\ u^T(k) &= [u_1(k) \quad u_2(k) \quad \dots \quad u_p(k)] \end{aligned} \tag{2.2}$$

and the matrix T(z) has elements $T_{ij}(z)$ which are rational functions over the field F.

A matrix fraction description of a qxp matrix T(z) of rational functions over an arbitrary field is a pair of polynomial matrices (matrices over elements of F[z]) (P(z),Q(z)) of dimension qxq and qxp respectively, satisfying

$$T(z) = P^{-1}(z)Q(z). \tag{2.3}$$

where we have:

$$P(z)= [p_{ij}(z)],\ Q(z) = [q_{ij}(z)], \text{and } p_{ij}(z), q_{ij}(z) \in F[z] \text{ and}$$

$$p_{ij}(z)= p_{ij,\nu_{ij}+1}z^{\nu_{ij}} + p_{ij,\nu_{ij}}z^{\nu_{ij}-1} +\dots+ p_{ij,2}z + p_{ij,1} \tag{2.4}$$

$$q_{ij}(z)= q_{ij,\nu_{ij}+1}z^{\nu_{ij}} + q_{ij,\nu_{ij}}z^{\nu_{ij}-1} +\dots+ q_{ij,2}z + q_{ij,1} \tag{2.5}$$

To be more precise, the form (2.3) is referred to as a left matrix fraction. The right MFD can be defined in an analogous way.

We will concentrate ourselves upon the left MFD because of its more direct relation to the ARMA-model for which parameter-estimations are performed.

The degree of a MFD is defined to be equal to the degree of determinant of the P(z) matrix.
The definition of a MFD is quite analogous to the description of the transfer function of a SISO system as a quotient of two polynomials. According to this the P(z)-matrix is sometimes referred to as the denominator-matrix , and Q(z) as the numerator-matrix.

In this report only strictly proper transfer matrices will be considered. A strictly proper transfer matrix is a matrix in which the degree of the numerator polynomial of each entry is less than the degree of its corresponding denominator polynomial.This has as a consequence that there is no direct link between input and output , i.e. all inputs are delayed at least one sample before appearing in the output.
Similarly, a proper transfer matrix has the degrees of the denominator polynomials equal to or higher than the degrees of the numerator polynomials, and therefore a direct coupling is possible between input and output . A direct coupling implies that an input signal at time k has an effect on the output y(k) at the same instant. (Wolovich, 1974)
Another way of writing the input - output relation (2.1) is in the time domain, with z^{-1} as the unitary delay operator:

$$P(z)y(k) = Q(z)u(k) \tag{2.6}$$

The timeshift operator z works as follows:

$$z.y(k) = y(k+1) \tag{2.7}$$

$$z^{-1}y(k) = y(k-1)$$

The notation P(z) instead of $P(z^{-1})$ has been adopted for convenience only. In places where the polynomials are used to calculate time-shifts, the equation will be multiplied by z^{-n} on both sides using a suitable n, and thus only time delays are considered to be part of the system.

The determinant of P(z) has to be non-zero.
If det P(z) = 0 , this means that there is a linear dependence between 2 or more expressions with y-components on the left side of equation

(2.1). In that case a unique solution for $y(k)$ for a given sequence $u(k)$ cannot exist.

In this place we may hint at the relation of the MFD model and the ARMA-model. If we write:

$$Y(z) = T(z)U(z) \ ,$$

where $Y(z)$ and $U(z)$ are the z-transforms of $y(k)$ and $u(k)$ we have :

$$Y(z) = P^{-1}(z)Q(z)U(z) \qquad \text{or,}$$

$$P(z)Y(z) = Q(z)U(z). \tag{2.8}$$

The relation between this last equation and (2.6) is found by substituting the z-transform $Y(z)$ and $U(z)$:

$$P(z).\left[\sum_{k=0}^{\infty} y(k)z^{-k}\right] = Q(z).\left[\sum_{k=0}^{\infty} u(k)z^{-k}\right] \tag{2.9}$$

and equating coefficients of powers of z.

A very simple way of generating a MFD is the following:
write:

$$T(z) = Q(z)/d(z) \tag{2.10}$$

where $d(z)$ is the least common multiple of the denominator polynomials of the entries in the rational matrix $T(z)$. The MFD is thus $(d(z)I, Q(z))$ since

$$T(z) = (d(z)I)^{-1}.Q(z) \ . \tag{2.11}$$

Chapter 3: POLYNOMIAL MATRICES

As the treatment of matrix fractions is largely algebraical, we will need a basic knowledge of a number of properties of polynomial matrices. In this paragraph some definitions are given as well as some accompanying examples.
The purpose of this chapter is to give the most important definitions, along with some illustrations of their use. These definitions are given in a somewhat condensed form, but their understanding is essential for the study of matrix fractions. For a more elaborate treatment of the theory of real and polynomial matrices, the reader may consult e.g. Kailath (1980), MacDuffee (1956), Gantmacher (1959), or Lancaster e.a. (1982). A comprehensive list of references can be found in an article by Barnett (1973).

3.1 : Unimodular matrices

We will frequently encounter the term elementary row operations, or elementary column operations. These operations are defined as follows:
There are 3 possible elementary row operations:

1) The first is to exchange two rows of a polynomial matrix.
2) The second is to multiply a row with a constant.
3) The third is to exchange a row by itself plus any row multiplied by any polynomial.

For elementary column operations these definitions are analogous.
The application of elementary row operations can be described as a premultiplication of the polynomial matrix by a unimodular matrix.(Wolovich, 1974)
A unimodular matrix is a polynomial matrix having a determinant which is not a function of z and is therefore an constant.
The important fact about multiplication of a polynomial matrix P(z) with a unimodular matrix U(z) is that the degree of the determinant of U(z).P(z) is equal to the degree of det P(z):

$$\det\left[U(z)P(z)\right] = \det U(z).\det P(z) = \text{constant}.\ \det P(z)$$

The second interesting property of unimodular matrices is that their inverse is also unimodular. This can be understood by writing:

$$U^{-1}(z) = \frac{U^{+}(z)}{\det U(z)} = \frac{U^{+}(z)}{\text{constant}} \quad ,$$

and $U^{+}(z)$ is the adjugate matrix of $U(z)$.

3.2 Common divisors and coprimeness

We will investigate if two polynomial matrices have common factors (which will be called divisors) and how these common factors can be calculated. 3 tests will be given to detect common factors which are not unimodular. The cancellation of these factors from the MFD $(P(z),Q(z))$ leads to a reduction of the degrees of the polynomials in $P(z)$ and $Q(z)$. This cancellation is exactly what is called pole-zero cancellation in the transfer function of SISO - systems.

$P(z)$ is a polynomial left divisor of $Q(z)$ if there is a polynomial matrix $X(z)$ such that

$$Q(z) = P(z).X(z) \tag{3.1}$$

$Q(z)$ is also called a right multiple of $P(z)$.
The notions of right divisor and left multiple are defined analogously: $X(z)$ is a right divisor of $Q(z)$ and $Q(z)$ is a left multiple of $X(z)$.

Example 3.1

Consider the matrix $\begin{vmatrix} z^2 & 0 \\ z & z \end{vmatrix}$.

A left divisor of this matrix is: $\begin{vmatrix} z & 0 \\ 0 & z \end{vmatrix}$, because we can write: $\begin{vmatrix} z^2 & 0 \\ z & z \end{vmatrix} = \begin{vmatrix} z & 0 \\ 0 & z \end{vmatrix} \begin{vmatrix} z & 0 \\ 1 & 1 \end{vmatrix}$

A greatest common left divisor (gcld) of two polynomial matrices having the same number of rows is a polynomial left divisor of the matrices that is also a right multiple of any other common left divisor of the two matrices.
Thus if L(z) is a gcld of P(z) and Q(z) and

$$P(z) = W(z)P^*(z), \quad Q(z) = W(z)Q^*(z)$$

for some polynomial matrices $P^*(z)$, $Q^*(z)$ and W(z) then there is a polynomial matrix $W^*(z)$ with

$$L(z) = W(z)W^*(z). \tag{3.2}$$

If (P(z),Q(z)) is a MFD of T(z), then $(P^*(z),Q^*(z))$ is also a MFD of T(z) because $0 \neq \det P(z) = \det P^*(z) . \det W(z)$ so that both $P^*(z)$ and W(z) are nonsingular.

The adjective "greatest" indicates that a gcld is a common left divisor which has the degree of its determinant equal to the maximum degree of determinants of all common left divisors.
If a greatest common left divisor of two polynomial matrices is a unimodular matrix, the two matrices will be called left coprime.
In this case the degree of the determinant of P(z) cannot be reduced by cancelling common left divisors in P(z) and Q(z), since det U(z) is a constant.

Note the fact that we did not talk about "the" gcld of two polynomial matrices. This is because gcld's are not unique.
This nonuniqueness can be demonstrated as follows:
If $L_1(z)$ is a gcld of P(z) and Q(z) then $L_1(z).U(z)$ is also a gcld of P(z) and Q(z) because it fulfils the requirements mentioned above, with W^* replaced by $W^*(z)U(z)$.

Example 3.2

Find the greatest common left divisor of the two following matrices:
$\begin{vmatrix} z^2 & 0 \\ z & z \end{vmatrix}$ and $\begin{vmatrix} z & 0 \\ 0 & z \end{vmatrix}$. Obviously a common left divisor is: $\begin{vmatrix} z & 0 \\ 0 & z \end{vmatrix}$

Elimination of this cld results in matrices:
$\begin{vmatrix} z & 0 \\ 1 & 1 \end{vmatrix}$ and $\begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix}$. We conclude that the cld is a greatest common left divisor because no polynomial matrices can be a divisor of the identity matrix.

3.3 Tests for coprimeness

Next we will describe how these greatest common left divisors can be found, and we will give 3 tests to see if these gcld's are unimodular or not.

Our first approach constructs the gcld of two polynomial matrices, and after calculating the degree of its determinant we can tell if the two matrices are coprime.
Consider two polynomial matrices P(z) and Q(z) of dimensions qxq and qxp resp. If the composite matrix $[P(z):Q(z)]$ is reduced to a form $[T_1(z):\ 0\ ]$ where $T_1(z)$ is a qxq lower left triangular matrix and the null matrix $:\ 0\ ]$ has dimensions qxp, then T_1 is a gcld of (P(z),Q(z)).
The reduction can be represented as follows with the help of a unimodular matrix V(z):

$$[P(z):Q(z)].V(z) = [T_1(z):\ 0\ ]$$

Wolovich (1974) proved that a lower left triangular form exists for every polynomial matrix.
We may write this as:

$$[P(z):Q(z)] \begin{vmatrix} V_1(z) & V_2(z) \\ V_3(z) & V_4(z) \end{vmatrix} = [T_1(z)\ :\ 0\ ] \qquad (3.3)$$

where $V_1(z)$ and $V_4(z)$ have dimensions qxq and qxp respectively so that

$$P(z)V_1(z) + Q(z)V_3(z) = T_1(z)$$

Now if W(z) is any left divisor of P(z) and Q(z), then

$$W(z)\ (A^*(z)V_1(z) + B^*(z)V_3(z)) = T_1(z) \qquad (3.4)$$

which shows that $T_1(z)$ is a right multiple of W(z).
If X(z) is the polynomial inverse of the unimodular matrix V(z) then:

$$\left[P(z)\ :\ Q(z)\right] = \left[T_1(z)\ :\ 0\ \right] \begin{vmatrix} X_1(z) & X_2(z) \\ X_3(z) & X_4(z) \end{vmatrix} \qquad (3.5)$$

giving

$$\begin{aligned} P(z) &= T_1(z)X_1(z) \\ Q(z) &= T_1(z)X_2(z) \end{aligned} \qquad (3.6)$$

So that $T_1(z)$ is indeed a left divisor of P(z) and Q(z).
To see why $T_1(z)$ is a gcld we note that (from eq. (3.4)) $T_1(z)$ is a right multiple of any left divisor.
The practical use of this approach may be quite tedious because of the numerical operations it involves.

There is an alternative test for coprimeness which is relatively simple. The matrices P(z) and Q(z) are left coprime iff the rank of the composite matrix $[P(z):Q(z)]$ is q for every $z \in \bar{F}$, the algebraic closure of the field F.(Dickinson, 1974)
Calculation of the rank of a polynomial matrix is relatively simple. If it is found however, that the matrices are not coprime, then the calculation of the gcld still remains to be done.

Example 3.3

Consider the matrices from the previous example. If we calculate the rank of the composite matrix $\begin{vmatrix} z^2 & 0 & z & 0 \\ z & z & 0 & z \end{vmatrix}$ for $z = 0$

it find that it is 0 (the composite matrix is a nullmatrix in this case). And so, because $0 < 2$ ($= q$), the matrices are not coprime. After elimination of the greatest common left factor we can construct the composite matrix : $\begin{vmatrix} z & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 \end{vmatrix}$.

This matrix has rank two for all possible values of z.

A third possibility to test for coprimeness is offered by calculating the Smith form (which will be described in a later chapter) of the composite matrix $[P(z):Q(z)]$.
The Smith form of $[P(z):Q(z)]$ is $[I_q: 0]$ if P(z) and Q(z) are coprime. Where I_q is the qxq identity matrix. This will be treated further in the chapter about the Smith form.(Kailath 1982)

A left MFD in which the matrices P(z) and Q(z) are left coprime will be called irreducible if they have only unimodular common left factors. We can see that all irreducible MFDs of T(z) are closely related:
if $(P_1(z),Q_1(z))$ and $(P_2(z),Q_2(z))$ are 2 irreducible left MFDs of T(z) then

$$\begin{aligned} P_1(z) &= V(z)P_2(z) \\ Q_1(z) &= V(z)Q_2(z) \end{aligned} \tag{3.7}$$

for some unimodular matrix V(z).
Because V(z) is unimodular we find that

$$\deg\det P_1(z) = \deg\det V(z).P_2(z) = \deg\det P_2(z) \tag{3.8}$$

From the above we can be sure that an irreducible MFD has minimal degree. The fact of irreducibility of two polynomial matrices P(z) and Q(z) tells us that the only common divisors of P(z) and Q(z) are unimodular, and therefore the degree of det P(z) cannot be reduced by eleminating a common divisor.

3.4 Coprimeness of reciprocal polynomial matrices

There remains one point about coprimeness which may not be noticed immediately.

If two polynomial matrices P(z) and Q(z) are coprime, it does not follow that the reciprocal polynomials $z^nP(z^{-1})$ and $z^nQ(z^{-1})$ are coprime also. Instead we have the next property.

Consider two polynomial matrices:

$$P(z)= P_{n+1}z^n+\ldots+P_2z+P_1 \text{ and } Q(z) = Q_{n+1}z^n+\ldots+Q_2z+Q_1 \tag{3.9}$$

and their reciprocal polynomial matrices as defined already.

P(z) has dimensions qxq and Q(z) dimensions qxp.

It can be proven (see Stoica and Söderström, 1982) that the following statements are equivalent:

1) P(z) and Q(z) are left coprime and rank $[P_{n+1} : Q_{n+1}]= q$.

2) $z^nP(z^{-1})$ and $z^nQ(z^{-1})$ are left coprime and rank $[P_1 : Q_1] = q$.

For our use of polynomials in z^{-1} there is no need to worry because the matrices with these polynomials are calculated in a different way (see chapter 8).

3.5 The construction of right MFDs from left MFDs

The construction of a right irreducible MFD from a left MFD will be shown next.

Starting again with equation (3.3):

We have : $[P(z):Q(z)] \begin{vmatrix} V_1(z) & V_2(z) \\ V_3(z) & V_4(z) \end{vmatrix} = [T_1(z): 0]$

When $P(z)$ is nonsingular (as we will assume always) then $V_4(z)$ will be nonsingular and we derive from equation (3.3)

$$P^{-1}(z)Q(z)=V_2(z)V_4^{-1}(z). \tag{3.10}$$

Furthermore $(V_2(z),V_4(z))$ will be an irreducible right MFD. (See Kailath, 1982)

3.6 Nonuniqueness of MFDs

The nonuniqueness of the decomposition of $T(z)$ as a MFD is easily demonstrated as follows. If $(P_1(z),Q_1(z))$ is a MFD of $T(z)$ then for every nonsingular qxq polynomial matrix $L(z)$ we find that

$$(L(z).P_1(z),L(z)Q_1(z)) = (P_2(z),Q_2(z)) \tag{3.11}$$

is also a MFD of $T(z)$:

$$(L(z).P_1(z))^{-1}.L(z)Q_1(z) = P_1^{-1}(z)L^{-1}(z).L(z)Q_1(z)$$
$$= P_1^{-1}(z).Q_1(z)$$

Now we may ask ourselves if there is a way to find a MFD that is the simplest in some sense. The deletion of common polynomial matrix factors gives us the possibility to simplify the MFD.
After cancellation of a number of common factors the only common factors are unimodular matrices. At this point we have an MFD of least possible degree.

Example 3.4

Consider the rational matrix $T(z)= \begin{vmatrix} \dfrac{z}{z+1} & \dfrac{1}{z+2} \\ \dfrac{z+3}{z+2} & \dfrac{2z}{z+2} \end{vmatrix}$ (3.12)

Now we want to construct a MFD for $T(z)$. The simplest way to do this is to compute the least common multiple of the denominators of $T(z)$. The least common multiple is: $d(z)= (z+1)(z+2)$. We can construct a MFD directly as follows:

$$T(z)= P^{-1}(z).Q(z) \text{ where} \tag{3.13}$$

$$P(z)=\begin{vmatrix} (z+1)(z+2) & 0 \\ 0 & (z+1)(z+2) \end{vmatrix}, \quad Q(z)=\begin{vmatrix} z(z+2) & z+1 \\ (z+3)(z+1) & 2z(z+1) \end{vmatrix}$$

The degree of this MFD is equal to 4.
This particular MFD is not irreducible however, as can be shown.
It can be easily seen that the matrix $U(z)=\begin{vmatrix} 1 & 0 \\ 0 & z+1 \end{vmatrix}$
is a common left factor in both matrices:

$$P(z)= U(z). \begin{vmatrix} (z+1)(z+2) & 0 \\ 0 & z+2 \end{vmatrix} = U(z).P_1(z) \quad \text{and}$$

$$Q(z)= U(z). \begin{vmatrix} z(z+2) & z+1 \\ z+3 & 2z \end{vmatrix} = U(z).Q_1(z) \tag{3.14}$$

To see if $P_1(z)=U^{-1}(z)P(z)$ and $Q_1(z)=U^{-1}(z)Q(z)$ are coprime, we perform the rank test: first we form the composite matrix: $C(z)=[P_1(z) : Q_1(z)]$ and check its rank for all values of z. Obviously as the rank of C(z) calculated for z=-1 and z=-2 is equal to two we can be sure that the rank of C(z) is 2 for all values of z. This means that the matrices $P_1(z)$ and $Q_1(z)$ are left coprime and that the corresponding MFD $(P_1(z),Q_1(z))$ is irreducible. Calculating the degree of the MFD we find that its value has been reduced to 3.

3.7 Row proper matrices

A row proper matrix is a square polynomial matrix of which the sum of the degrees of its rows equals the degree of its determinant. E.g. every diagonal polynomial matrix is row proper.
The importance of row properness of the P(z) matrix in a MFD will become clear at the end of the section on equivalent MFDs.
In the future we will talk about standard MFDs when the P(z) matrix is row proper. This is done because the term "proper" is already used in relation with transfer matrices (chapter 1).

Taking a closer look at $[P(z)]_h$: the high order coefficient matrix of $P(z)$, it will be nonsingular in case $P(z)$ is row proper.
$[P(z)]_h$ is the real matrix containing the coefficients of the terms having row-degree in each row of $P(z)$.
For example if $P(z) = \begin{vmatrix} z^2+z & z \\ 2z & 3z \end{vmatrix}$ then $[P(z)]_h = \begin{vmatrix} 1 & 0 \\ 2 & 3 \end{vmatrix}$

(3.15)

We see that the row degrees are 2 (first row) and 1. The degree of det $P(z)$ is 3, and because $1 + 2 = 3$ we conclude that the matrix $P(z)$ is row proper.

It is easy to show that if $P(z)$ is row proper, then its inverse is a proper matrix. This follows by noting that $(P(z),I)$ is a standard irreducible MFD of $P^{-1}(z)$ when $P(z)$ is row proper. (Dickinson, 1974)
Furthermore if a row proper matrix $P_1(z)$ has a right divisor $P_2(z)$ which is also row proper then the degrees of the left factor are less than.the respective row degrees of $P_1(z)$.

Now we have the following interesting property:
every nonsingular polynomial matrix $P(z)$ can be written as

$$P(z)= U(z)\cdot P_r(z) \tag{3.16}$$

where $P_r(z)$ is row proper and $U(z)$ is unimodular. The construction of the matrix $U(z)$ is described next.
Let $P(z)$ be a nonsingular polynomial matrix: k_i is the ith row degree.

Consider the matrix:

$$P^c(z)= \text{diag}\{z^{k_1}, z^{k_2}, \ \ldots \ , z^{k_q}\}.[P(z)]_h \tag{3.17}$$

and suppose that det $P^c(z) = 0$, i.e. $P^c(z)$ is a singular matrix.
This implies that the q rowvectors: $P_i^c(z)$ are linearly dependent over $F[z]$, the ring of polynomials, and in particular that

$$\sum_{i=1}^{q} p_i(z).P_i^c(z) = 0 \tag{3.18}$$

for two or more nonzero monomials $p_i(z)$, for $i=1,2,..,q$ (N.B. monomials are polynomials that consist of only one term.) We can change at least one monomial into a 1, by dividing the previous equation by a nonzero monomial of lowest degree.

$$\sum_{i=1}^{q} (p_i(z)/p_k(z)).P_i^c(z) = \sum_{i=1}^{q} \tilde{p}_i(z).P_i^c(z) = 0 \qquad (3.19)$$

where $\tilde{p}_j(z)=1$. The replacement of row j of $P^c(z)$ by

$\sum_{i=1}^{q} \tilde{p}_i(z).P_i^c(z)$ is analogous to premultiplication of

$P^c(z)$ by the unimodular matrix $U_1(z)$, where $U_1(z)$ is the identity matrix with a suitably changed j-th row:

$$U_1(z) = \begin{vmatrix} 1 & 0 & . & 0 & 0 \\ 0 & 1 & & & \\ & . & & & \\ \tilde{p}_1(z) & \tilde{p}_2(z) & \ldots\ \tilde{p}_j(z) & \ldots & \tilde{p}_q(z) \\ & & & & \\ 0 & 0 & & . & 1 \end{vmatrix} \qquad (3.20)$$

where at least one monomial $\tilde{p}_i(z)$, $i \neq j$, is nonzero. Clearly if $P^c(z)$ is premultiplied by $U_1(z)$, the j-th row of $P^c(z)$ will be identically zero. Now if

$$\bar{P}(z)=U_1(z).P(z) \qquad (3.21)$$

$\bar{P}(z)$ is identical to $P(z)$ except for a changed j-th row. More specifically, the degree of the j-th row of $\bar{P}(z)$ is strictly less than the degree of the j-th row of $P(z)$. $\bar{P}(z)$ is therefore a new candidate of lower degree for a

row proper matrix.

Example 3.5

As an illustration of the algorithm we will bring the matrix

$$P(z)=\begin{vmatrix} z^2-3 & 1 & 2z \\ 4z+2 & 2 & 0 \\ -z^2 & z+3 & -3z+2 \end{vmatrix} \quad \text{to row proper form.} \qquad (3.22)$$

First we determine $P^c(z)$:

$$P^c(z)=\operatorname{diag}\{z^2,z,z^2\}.\begin{vmatrix} 1 & 0 & 0 \\ 4 & 0 & 0 \\ -1 & 0 & 0 \end{vmatrix} = \begin{vmatrix} z^2 & 0 & 0 \\ 4z & 0 & 0 \\ -z^2 & 0 & 0 \end{vmatrix} \qquad (3.23)$$

Because det $P(z)= 6z^3+44z^2+28z-16\neq 0$, $P(z)$ is nonsingular.
However $P(z)$ is not row proper since det $P^c(z)=0$.
To achieve row proper form for $P(z)$ we have to determine 3 monomials $p_i(z)$ which satisfy relation (3.).
It is easily checked that $p_1(z)=1$, $p_2(z)=z, p_3(z)=5$ fulfill the requirements.
As $p_1(z)$ is a monomial of least degree, we divide the other monomials by $p_1(z)$, and we have $\tilde{p}_1(z)=1, \tilde{p}_2(z)=z, \tilde{p}_3(z)=5$.
Now a new candidate for a row proper matrix is:

$$\bar{P}_1(z)= U_1(z).P(z)= \begin{vmatrix} 1 & z & 5 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix} .P(z)= \qquad (3.24)$$

$$= \begin{vmatrix} 2z-3 & 7z+16 & -13z+10 \\ 4z+2 & 2 & 0 \\ -z^2 & z+3 & -3z+10 \end{vmatrix}$$

If we calculate det $\bar{P}_1^c(z)$ we find that $\bar{P}_1(z)$ is not row proper.
Repeating the above procedure we find a second unimodular matrix $U_2(z)$ such that:

$$\bar{P}_2(z) = U_2(z).\bar{P}_1(z) = \begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & \frac{z}{4} & 1 \end{vmatrix} .\bar{P}_1(z) =$$

$$= \begin{vmatrix} 2z-3 & 7z+16 & -13z+10 \\ 4z+2 & 2 & 0 \\ 2z & \frac{3}{2}z+3 & -3z+10 \end{vmatrix} \qquad (3.25)$$

Calculating det $\bar{P}_2^c(z)$ we find that it equals $-6z^2$ and therefore we conclude that $P_2(z)$ is a row proper matrix.

3.8 Equivalence of polynomial matrices and MFDs

In this chapter we will be concerned with the relations among MFDs. It was already discovered in section that different MFDs can lead to the same transfer matrix.
At the end of this chapter, a very interesting property is given, which has consequences for the use of MFDs in identification procedures.

If we have a set of polynomial matrices, we can define an equivalence relation on this set. This relation can be described with a function f from the original set, say O , to a set, say S : $f: O \to S$.
Two elements X and Y are equivalent if the function f transforms them into the same element of S:

$$(X\ E\ Y);\ X,Y \in O \to f(X)=f(Y).$$

The function f is called an invariant for the equivalence relation.
For pairs of polynomial matrices $(P(z),Q(z))$ (MFDs) an equivalence relation can be defined as follows:
two pairs $(P_1(z),Q_1(z))$ and $(P_2(z),Q_2(z))$ are equivalent if and only if $P_1^{-1}(z).Q_1(z) = P_2^{-1}(z).Q_2(z)$.

If we define the equivalence relation on the set of coprime polynomial matrix-pairs then there is a is a nonsingular unimodular matrix $L(z)$ for which:

$$\begin{aligned} P_2(z) &= L(z).P_1(z) \\ Q_2(z) &= L(z).Q_1(z) \end{aligned} \qquad (3.26)$$

The equivalence relation for polynomial matrices which is useful for the left MFD, is the left unitary equivalence. Two polynomial matrices $P_1(z)$ and $P_2(z)$ are left unitary equivalent if a nonsinglar unimodular matrix L(z) can be found such that $P_1(z)=L(z).P_2(z)$.
One can choose a certain matrix $\Xi(z)$ for which it holds that

$$\Xi(z)=f(P_1(z))=f(P_2(z)).$$

Such a matrix $\Xi(z)$ can then act as a representative for all matrices equivalent to $P_1(z)$.For identification procedures this means that we can choose a convenient form among all possible equivalent forms.
When we speak about equivalent MFDs we will obviously refer to pairs of polynomial matrices $(P_1(z),Q_1(z))$ and $(P_2(z),Q_2(z))$ that are equivalent. This means that equivalent MFDs give the same transfer function.
The concept of equivalent MFDs means that we have some freedom in describing a certain transfer function and also that we can change from one equivalent description to another as it pleases us. This can be useful in identification procedures.

All elements that are equivalent in a set are said to belong to the same equivalence class.
Choosing a unique element as a representative for a whole equivalence class leads to the definition of canonical forms for a certain equivalence relation. There exist various definitions of the term canonical in the literature. The relations between these definitions are still a matter of investigation. (See Tse, Weinert, Anton, 1973; Denham, 1974). Still, these unique forms can be useful in identification procedures and will therefore be included in the remaining discussion. If a unique form for an MFD is desired, we can choose, which of the two matrices in the pair will be brought to a standard form. The structure of the other matrix cannot be chosen freely because the resulting MFD has to be equivalent to the original one:
given an irreducible MFD (P(z),Q(z)) we choose a standard form for P(z): P*(z).

$$P^*(z)=L(z).P(z).$$

Then the equivalent MFD with P(z) in the desired form will be

$(P^*(z),Q^*(z))=(L(z).P(z),L(z).Q(z))$.
$L(z)$ is a unimodular polynomial matrix.
Note the link between equivalent left MFDs and left unitary equivalence of polynomial matrices.
The choice of a standard form is made by appointing the degrees of the polynomials $p_{ij}(z)$. One of the essential limitations on these degrees is that deg det $P(z) = n$, and n equals the dimension of the associated minimal state-space model (except in case of the diagonal parametrization). In estimation practice the value of n may not be known behorehand and several values of n can be investigated for their adequateness.
For a more rigorous treatment of canonical forms the reader is referred to any textbook on algebra (e.g. Birkhoff and MacLane).
Canonical forms for matrix fractions are: the form of Guidorzi and Beghelli, the Hermite form, the echelon form, and the reversed echelon form. In these forms a particular structure is given to the denominator matrix. These forms will be discussed in detail, along with the algorithm to construct these forms. The canonicity of the other forms is still under study.
The equivalence of matrix fractions has its counterpart for state space descriptions. For a state-space description the equivalent triples (F,G,H) are related by a transformation of the basis of the state-space. The relation between state-space and matrix fraction descriptions has been extensively described by Antoulas (1981).

3.8.1 Sets of row degrees

A key property of standard irreducible MFDs is given next.
Consider all standard irreducible MFDs $(P(z),Q(z))$ of a proper rational matrix $T(z)$. Define $V_d(z)$ as the vector space spanned by the highest order coefficient (row-)vectors of the rows of $P(z)$ having degree less than or equal to d.
Then 1) the sets of row degrees of all standard irreducible MFDs are the same,
2) $V_d(P(z))$ is the same for each $P(z)$ for every d.
(Dickinson, 1974)

These statements can be proved as follows:
As shown previously any 2 standard irreducible MFDs are connected by a left unimodular matrix. Since each MFD has a nonsingular highest order coefficient matrix, a row of the first, say p_1 , having degree d_1, must be a linear combination over $F[z]$ of the rows of $P_2(z)$ having degree less than or equal to d_1. Also the sum of the row degrees of both
matrices are equal. If the sets of row degrees are not equal then the highest order coefficient matrix of one of the matrices must be singular, which is a contradiction.
Therefore these sets are equal and consequently the dimensions of the vector spaces $V_d(P_1(z))$ and $V_d(P_2(z))$ are the same.
However the highest order coefficient row vector of a row of $P_1(z)$, say $[p_1]_h$ is also a linear combination of the highest order row vectors of the rows of $P_2(z)$ having degree less than or equal to d_1.
Therefore the vector spaces $V_d(P_1(z)$ and $V_d(P_2(z))$ must also be the same.

For our discussion on equivalent forms for MFDs this means, that the forms of Guidorzi and the echelon form must have the same set of row degrees, while this set is not necessarily equal to the set of row degrees of the hermite form.

3.9 The Smith-McMillan representation

The Smith-McMillan representation of a rational matrix illustrates one way of obtaining an irreducible MFD directly, without the need to extract common left divisors of a preliminary MFD.
First the Smith-form of a polynomial matrix is presented.
In the subsequent section we will describe the Smith-McMillan form of a rational matrix.

3.9.1 The Smith-form of a polynomial matrix.

Any polynomial matrix P(z) can be expressed as

$$P(z) = U_1(z).\Lambda(z).U_2(z) \qquad (3.27)$$

$P(z) = [\ p_{ij}(z)\ ]$, and $U_1(z)$, $U_2(z)$ are unimodular matrices. $\Lambda(z)$ is called the Smith form and is a diagonal matrix with the same dimensions as $P(z)$:

$$\Lambda(z) = \left|\begin{array}{cccccccc} \mu_1(z) & & & & & & & \\ & \mu_2(z) & & & & & & \\ & & \cdot & & & & & \\ & & & \cdot & & & & \\ & & & & \mu_r(z) & & & \\ & & & & & 0 & & \\ & & & & & & \cdot & \\ & & & & & & & 0 \end{array}\right| \tag{3.28}$$

where r is the rank of the matrix P(z) and the diagonal entries of $\Lambda(z)$ are monic polynomials satisfying:

$$\mu_i / \mu_{i+1} \qquad i \leq r-1 \tag{3.29}$$

This implies that $\deg \mu_r \leq \deg \mu_{r-1} \leq \dots \leq \deg \mu_1$, and also that the greatest common divisor of μ_i and μ_{i+1} equals μ_i (Kailath, 1980). Define $\Delta_0(z) = 1$ and $\Delta_i(z)$ as the monic greatest common divisor of the i by i minors of P(z). The invariant polynomials μ_i can

then be computed as

$$\mu_i = \frac{\Delta_i(z)}{\Delta_{i-1}(z)} \qquad i \leq r \tag{3.30}$$

and this shows that $\Lambda(z)$ is uniquely specified. (See e.g. MacDuffee (1956).

*Minors are the determinants of a i by i submatrices that are formed by deleting a number of rows and a number of columns in a matrix.

3.9.2 The Smith-McMillan form of a rational matrix

Any rational matrix $T(z) = [T_{ij}(z)] = \left| \dfrac{p_{ij}(z)}{q_{ij}(z)} \right|$ can be represented as

$$T(z) = U_1(z).\Pi^{-1}(z).E(z).U_2(z) \tag{3.31}$$

where the product $\Pi^{-1}(z).E(z)$ is called the Smith-McMillan form, and $U_1(z)$ and $U_2(z)$ are unimodular matrices .
$\Pi(z)$ and $E(z)$ are diagonal qxp and qxq polynomial matrices:

$$E(z) = \begin{vmatrix} e_1(z) & & & & & \\ & \cdot & & & & \\ & & \cdot & & & \\ & & & e_r(z) & & \\ & & & & 0 & & \\ & & & & & \cdot & \\ & & & & & & 0 \end{vmatrix}$$

$$\Pi(z) = \begin{vmatrix} \pi_1(z) & & & & & \\ & \cdot & & & & \\ & & \cdot & & & \\ & & & \pi_r(z) & & \\ & & & & 1 & & \\ & & & & & \cdot & \\ & & & & & & 1 \end{vmatrix} \tag{3.32}$$

The polynomials in these matrices must satisfy:

$$\begin{aligned} & e_i \ / \ e_{i+1} \\ & \pi_{i+1} / \ \pi_i \\ & e_i \text{ and } \pi_i \text{ are relatively prime polynomials.} \end{aligned} \tag{3.33}$$

The use of the Smith-McMillan form may become apparent in the following. Let d(z) be the monic least common multiple of the denominators of the entries of T(z). Then d(z).T(z) is a polynomial matrix with Smith form $\Lambda(z)$:

$$d(z).T(z) = U_1(z).\Lambda(z).U_2(z) \tag{3.34}$$

Dividing both sides by d(z) and reducing the elements in $\Lambda(z)/d(z)$ to lowest degree leads to the identification of E(z) as the matrix of numerators, and $\Pi(z)$ as the matrix of denominators with additional ones added if $\Lambda(z)$ is not of full rank.

The fact that $d(z) = \pi_1(z)$ may be noted as follows.
We may write:

$$T(z) = U_1(z) \cdot \begin{vmatrix} \frac{\mu_1(z)}{d(z)} & & & & & \\ & \cdot & & & & \\ & & \cdot & & & \\ & & & \frac{\mu_r(z)}{d(z)} & & \\ & & & & 0 & \\ & & & & & \cdot \\ & & & & & & 0 \end{vmatrix} \cdot U_2(z)$$

$$= U_1(z) \cdot \begin{vmatrix} \frac{e_1(z)}{\pi_1(z)} & & & & & \\ & \cdot & & & & \\ & & \cdot & & & \\ & & & \frac{e_r(z)}{\pi_r(z)} & & \\ & & & & 0 & \\ & & & & & \cdot \\ & & & & & & 0 \end{vmatrix} \cdot U_2(z) \qquad (3.35)$$

If $d(z)$ were not equal to $\pi_1(z)$ then this would imply that every element of $\Lambda(z)$ was divisible by a factor of $d(z)$. There would then be a monic least common denominator for the elements of $T(z)$ with lower degree than $d(z)$ contrary to assumption. In other words

$$\pi_1(z) \cdot T(z) = U_1(z) \cdot \Lambda(z) \cdot U_2(z) \qquad (3.36)$$

must be a polynomial matrix so that $\pi_1(z)$ must be a common multiple of denominators of $T(z)$.

Now, returning to our original aim, a left MFD can be obtained very easily as follows:

$$T(z) = (P(z), Q(z))$$

$$= (\Pi(z) \cdot U_1^{-1},\ E(z) \cdot U_2(z)) \qquad (3.37)$$

This is an irreducible MFD because the relative primeness of e_i and π_i guarantees the existence of diagonal polynomial matrices $X_1(z)$ and $Y_2(z)$ such that

$$E(z) \cdot X_1(z) + \Pi(z) \cdot Y_1(z) = I \qquad (3.38)$$

Taking $X(z) = U_2^{-1}(z).X_1(z)$

$Y(z) = U_1(z).Y_1(z)$

we have $P(z).Y(z) + Q(z).X(z) = I$ which shows that the MFD is irreducible.

There is no guarantee that this MFD will be row proper, as a simple example will demonstrate.

Suppose $\Pi(z) = \begin{vmatrix} z^2 & 0 \\ 0 & z \end{vmatrix}$ and $U^{-1}(z) = \begin{vmatrix} 1 & 0 \\ z & 1 \end{vmatrix}$

then $\Pi(z).U^{-1}(z) = \begin{vmatrix} z^2 & 0 \\ z^2 & z \end{vmatrix}$ and this matrix is not row proper.

Example 3.6

Let us calculate the Smith form of the composite matrix

$$[P(z):Q(z)] = \begin{vmatrix} (z+1)(z+2) & 0 & z(z+2) & z+1 \\ 0 & (z+1)(z+2) & (z+1)(a+3) & 2z(z+1) \end{vmatrix}$$

which is made of the matrices P(z) and Q(z) of example 1.

By definition $\Delta_0=1$.

There are 8 minors of dimension 1x1 (namely the elements of the composite matrix), and their greatest common divisor is 1.

Therefore $\Delta_1=1$.

There are 6 minors of dimension 2x2, formed by taking together the columns 1+2, 1+3, 1+4, 2+3, 2+4, 3+4. Columns 2+3 give a minor $z(z+2)^2(z+1)$ and columns 3+4 a minor $(z+1)(z^3-z^2-2z-3)$.

Their greatest common divisor is z+1, so $\Delta_2=z+1$.

The invariant polynomials are:

$$\mu_1= \Delta_1 / \Delta_0 = 1$$

$$\mu_2 = \Delta_2 \;/\; \Delta_1 = z+1$$

and $\Lambda(z) = \begin{vmatrix} 1 & 0 & 0 & 0 \\ 0 & z+1 & 0 & 0 \end{vmatrix}$

It is interesting to calculate the invariant polynomials for the matrix $[P_1(z):Q_1(z)]$, where $P_1(z)$ and $Q_1(z)$ are the coprime matrices from example 1:

$\Delta_0 = 1,\ \Delta_1 = 1,\ \Delta_2 = 1,$

and $\mu_1 = 1$, $\mu_2 = 1$, which gives a matrix $\Lambda(z)$:

$$\Lambda(z) = \begin{vmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{vmatrix} = [\ I : 0\ ]$$

This is exactly what we would have expected for two coprime matrices.

Chapter 4 : UNIQUE FORMS FOR MFDs

As seen previously there is an equivalence class of irreducible MFDs of a fixed rational T(z). Even the restriction to standard irreducible MFDs does not result in a unique representative of this class (for left unitary equivalence). Unique forms for irreducible MFDs will be of primary interest because of their correspondence to useful canonical state-space forms.

We will describe several unique equivalent forms for MFDs:

1 the Hermite form
2 the echelon form
3 the reversed echelon form
4 the numerator form
5 the denominator form
6 Guidorzi's form

In addition there are some forms for MFDs which can be used for identification and estimation purposes that are not uniquely identifiable but which can still deliver useful estimates of system-parameters under the right conditions. These forms are the fully parametrized form and the diagonal form.
The fully parametrized form (see Stoica and Söderström, 1982) has all degrees of the polynomials equal to n (see equation (3.9))
The already given diagonal parametrization can be constructed if the transfer matrix T(z) is known. However, this form cannot be constructed with the help of elementary row operations only, and therefore this form is not equivalent in the sense of premultiplication with a unimodular matrix.
When we talk about a special form of an MFD we mean that the P(z) matrix has a special form, and the matrix Q(z) is appropriately adjusted. The only exception is the numerator form where the Q(z) matrix has a special form and the matrix P(z) is in a less structured form.

4.1 The Hermite canonical form

The Hermite form $(P_H(z),Q_H(z))$ is obtained by taking an irreducible MFD $(P(z),Q(z))$ and applying left elementary row operations that bring $P(z)$ to its Hermite form: $P_H(z)$. The same operations applied to $Q(z)$ then give $Q_H(z)$.
$P_H(z)$ is a unique lower triangular polynomial matrix and has monic entries on the diagonal which are of higer degree than any other polynomials in their respective columns.

$$\begin{vmatrix} p_{11}(z) & 0 & \cdot & \cdot & \cdot & 0 \\ p_{21}(z) & p_{22}(z) & \cdot & \cdot & \cdot & 0 \\ \cdot & & \cdot & & & 0 \\ \cdot & & & \cdot & & \\ p_{q1}(z) & p_{q2}(z) & & \cdot & & p_{qq}(z) \end{vmatrix} \qquad (4.1)$$

Note that $P_H(z)$ is column proper so that a controllable state-space realization can be obtained as indicated in the previous section. (Wolovich, 1974)

4.1.1 Construction of the column Hermite form

If the last column of $P(z)$ is not identically zero, we can choose a polynomial of least degree from its elements, and by a permutation of the rows, make it the q,q element.
Then we apply the division-theorem theorem to every element in the last column. This means that we write every nonzero element $p_{iq}(z)$ as follows:

$$p_{iq}(z) = p_{qq}(z) \cdot s_{iq}(z) + r_{iq}(z) \qquad (4.2)$$

where either $r_{iq}(z)=0$ or $\text{degree}(r_{iq}(z)) < \text{degree}(p_{qq}(z))$
We then subtract from each nonzero i-th row, the last row

multiplied by $s_{iq}(z)$. If not all the remainders $r_{iq}(z)$ are zero, we choose one of least degree and make it the q,q element by a permutation of the rows. Since the degree of the q,q element is finite, this process must end at some stage. In particular when all the remaining elements of the last column are identically zero.

We next consider column q-1 of this altered matrix and ignoring the last column for the moment, apply the above procedure to the elements beginning with column q-1 and row q-1. In this way we zero all of the elements above the (q-1,q-1) entry. If the (q,q-1) entry is of equal or higher degree than the (q-1,q-1) element, the division algorithm can be employed to reduce the (q,q-1) element, to the remainder term, associated with the division of the (q,q-1) element by the (q-1,q-1) entry or to zero if both elements are scalars.

Continuing in this manner, we eventually reduce P(z) to the Hermite form, which will be lower left triangular.

Example 4.1

Consider the polynomial matrix
$$P(z) = \begin{vmatrix} z^2 & z+1 \\ z^3+2z+1 & z^3+2z^2-1 \end{vmatrix} \tag{4.3}$$

The row degrees are $k_1 = 2$, $k_2 = 3$. The degree of det P(z) is 5 so P(z) is row proper.

Following the rules for the construction of the Hermite form we find a sequence of unimodular matrices with which we can calculate $P_H(z)$:

$$\begin{aligned} P_H(z) &= \begin{vmatrix} -1 & 0 \\ 0 & 1 \end{vmatrix} \begin{vmatrix} 1 & -z^2-z+1 \\ 0 & 1 \end{vmatrix} \begin{vmatrix} 0 & 1 \\ 1 & 0 \end{vmatrix} \cdot P(z) = \\ &= \begin{vmatrix} z^4-z^2-2z-1 & 0 \\ z^2 & z+1 \end{vmatrix} \end{aligned} \tag{4.4}$$

$P_H(z)$ has row degrees $k_1 = 4$, $k_2 = 2$, while deg det $P_H(z)=5$.

Obviously $P_H(z)$ is not row proper as was already remarked previously. Also the set of row degrees of P(z) and of $P_H(z)$ are not identical.

4.2 The echelon form

A second unique form under left unimodular equivalence is the echelon form, which was introduced by Popov. (Dickinson,Kailath, Morf, 1974). We denote this form by $(P_E(z),Q_E(z))$.
The echelon form is specified bvy the following requirements.
There is a set of pivot indices $\{s_i\}$, $i \leqslant q$ defined so that

$P_E(z) = [p_{ij}(z)]$ with

$$\begin{array}{ll} 1 & k_1 \geqslant k_2 \geqslant \ldots \geqslant k_q \\ 2 & \text{degree } p_{is_i}(z) = k_i \\ 3 & p_{is_i}(z) \text{ is a monic polynomial} \\ 4 & \text{degree } p_{is_j}(z) < k_j \quad \text{for } i \neq j \\ 5 & \text{degree } p_{ij}(z) < k_i \quad \text{for } j > s_i \\ 6 & \text{if } k_i = k_j \text{ and } i > j \text{ then } s_i > s_j . \end{array} \tag{4.5}$$

In other words, $P_E(z)$ is a row proper matrix with row degrees $\{k_i\}$. The polynomial $p_{is_i}(z)$ is monic and has a degree that is higher than that of any other polynomial in column s_i, and of any polynomial to its right in row i. Finally if two rows have the same degree, the pivot index of the first is smaller than that of the second. This means that any proper rational matrix T(z) has a unique MFD $(P_E(z),Q_E(z))$ which is standard and irreducible and has $P_E(z)$ in echelon form.
The construction of the echelon form can be accomplished analogously to the construction of Guidorzis form (q.v.)

Example 4.2

If we start with the same polynomial matrix P(z) as in example 4.1 we can calculate its equivalent row echelon form.
Because P(z) is already row proper, we can order the row degrees and check if all conditions are fulfilled.

The pivot indices are: $s_1 = 2$, $s_2 = 1$.
The echelon form is:

$$P_E(z) = \begin{vmatrix} 1 & -z \\ 0 & 1 \end{vmatrix} \begin{vmatrix} 0 & 1 \\ 1 & 0 \end{vmatrix} . P(z) = \begin{vmatrix} 2z+1 & z^3-z^2+z-1 \\ z^2 & z+1 \end{vmatrix} \qquad (4.6)$$

The row degrees are $k_1 = 3$ and $k_2 = 2$, while deg det $P_E(z) = 5$.

4.3 The reversed echelon form

A related unique form is the reversed echelon form denoted $(P_R(z),Q_R(z))$ which is defined only for irredubible MFDs whose denominator satisfies det $P(0) \neq 0$.
This condition corresponds to transfer functions whose minimal realizations have no pole at the origin and hence for which the state matrix is nonsingular.
The transformation to reversed echelon form can be done in two steps. First the echelon MFD is obtained. Then elementary operations using pivot indices defined by the lowest degree coefficient matrix (which is nonsingular because det $P_E(0) \neq 0$) are performed to reduce the "reversed" matrix

$$\tilde{P}(D) = \text{diag } (D^{k_1}, D^{k_2}, \ldots, D^{k_q}).P_E(D^{-1}) \qquad (4.7)$$

to echelon form, say

$\tilde{P}_E(D)$.
Then $P_R(z)$ is defined as

$$P_R(z) = \text{diag } (z^{k_1}, \ldots, z^{k_q}).\tilde{P}_E(z^{-1}). \qquad (4.8)$$

Notice that $(P_R(z),Q_R(z))$ obtained by this procedure is a standard irreducible MFD.

Example 4.3

To construct the reversed echelon form we start with the echelon form and calculate $\tilde{P}(D)$.

Note that $P(0)= \begin{vmatrix} 0 & 1 \\ 1 & -1 \end{vmatrix}$.

$$\tilde{P}(D) = \begin{vmatrix} D^3 & 0 \\ 0 & D^2 \end{vmatrix} \begin{vmatrix} 2D^{-1}+1 & D^{-3}-D^{-2}+D^{-1}-1 \\ D^{-2} & D^{-1}+1 \end{vmatrix} = \begin{vmatrix} D^3+2D^2 & -D^3+D^2-D+1 \\ 1 & D^2+D \end{vmatrix}$$

$$\tilde{P}_E(D)= \begin{vmatrix} 1 & -2 \\ 0 & 1 \end{vmatrix} \begin{vmatrix} 1 & D \\ 0 & 1 \end{vmatrix} .\tilde{P}(D) = \begin{vmatrix} D^3+2D^2+D-2 & -3D+1 \\ 1 & D^2+D \end{vmatrix}$$

$$P_R(z)= \begin{vmatrix} z^3 & 0 \\ 0 & z^2 \end{vmatrix} \begin{vmatrix} z^{-3}+2z^{-2}+z^{-1}-2 & -3z^{-1}+1 \\ 1 & z^{-2}+z^{-1} \end{vmatrix}$$

$$= \begin{vmatrix} -2z^3+z^2+2z+1 & z^3-3z^2 \\ z^2 & z+1 \end{vmatrix} \qquad (4.9)$$

The row degrees are $k_1= 3$ and $k_2= 2$.

4.4 The numerator form

As indicated by Dickinson (1974) linear state variable feedback alters only the denominator of any standard MFD of the system transfer function. Therefore to study properties that remain invariant under linear state variable feedback it will be useful to have a MFD in which the numerator matrix has a unique form. When Q(z) has rank less then p, a completely unique MFD cannot be obtained by Q(z) alone. However a unique numerator can always be obtained and since the application of the result will be in the analysis of feedback, the nonuniqueness in the denominator can be eliminated in other ways.

T(z) is a qxp proper rational matrix that has a standard irreducible MFD ($P_N(z)$, $Q_N(z)$) with Q_N in a uniquely determined form: the numerator form.

The row degrees of $P_N(z)$ are : $k_1 \geq \ldots \geq k_q$.

Then there are 2 indices c_j and g_j satisfying the properties given below, provided the jth row of Q_N is not zero.

$$1)\quad \left.\frac{q_{ij}(z)}{z^k}\right|_{z=0} = 0 \qquad 0 \leq k \leq g_j,\ i \leq q$$

$$2)\quad \left.\frac{q_{ij}(z)}{z^{g_j}}\right|_{z=0} = \begin{cases} 0 & i > c_j \\ 1 & i = c_j \\ \text{arbitrary} & i < c_j \end{cases} \tag{4.10}$$

$$3)\quad \left.\frac{q_{c_i,j}(z)}{z^{k+g_i}}\right|_{z=0} = 0 \qquad i > j \text{ and } 0 \leq k \leq k_j - k_i$$

4 if $k_i = k_j$ and $i < j$, then $c_i < c_j$.

Finally, all zero rows in $Q_N(z)$ associated with the same row degree follow the nonzero rows associated with the same k_j, and $Q_N(z) = V(z) \cdot Q_E(z)$, for unimodular $V(z)$ whose j,ith entry has degree max $(0, k_j - k_i)$.

This can be proved as follows.

Because $P_N(z)$ is row proper we can construct $(P_N(z), Q_N(z))$ from $(P_E(z), Q_E(z))$ in a unique way, while preserving the row degrees for $P_E(z)$.

The requirements 1 - 4 uniquely specify this construction as follows. The lowest power of z in the last row of $Q_E(z)$ is normalized to have coefficient one in the last column that it appears . This is column c_q and the polynomial has its lowest power of z equal to g_q. By elementary row operations which preserve the row degrees of $P_E(z)$ the polynomials above $q_{c_q,q}(z)$ say $q_{c_q,j}(z)$ have their coefficients of z^l, $g_q \leq l \leq g_q + k_j - k_q$ nulled.

This procedure is then repeated for each row of $Q_E(z)$ from the bottom to the top. Row interchanges are made when $k_i = k_j$ to order the pivot columns .

Notice that the pivot columns are not necessarily distinct. The degree preserving elementary row operations implicit in U(z) can be applied to any qxp polynomial matrix X(z) with row degrees $k_i \leq w_i$, $i \leq q$, for some ordered set of integers $w_1 \leq \ldots \leq w_q$.

Another case of special interest occurs when the polynomial matrix P(z) has a full rank lowest coefficient P(0), and the integers are the degrees of the rows of P(z). In this case $P_N(z)$ can also be found by taking

$$\tilde{P}(D) = \text{diag}\ (D^{k_1}, D^{k_2}, \ldots, D^{k_q}).P(D^{-1}).$$

finding its echelon form $\tilde{P}_E(D)$ and then noting that

$$P_N(z) = \text{diag}\ (z^{k_1}, \ldots, z^{k_q}).\tilde{P}_E(1/z). \qquad (4.11)$$

which is therefore unique. It is this structure that motivated this definition of the numerator form rather than one introduced by Wolovich (1974). $\tilde{P}(D)$ corresponds to the reversed order polynomial matrix formed from the coefficients of P(z).

4.5 The denominator form

In particular the degree preserving elementary row operations from the previous section can be applied to $P_E(z)$ itself with $w_i=k_i$. The resulting MFD is $(P_D(z),Q_D(z))$, the denominator form. (Dickinson, 1974)

Example 4.4

Since the operations that bring the P(z) matrix to the denominator form are similar to the operations that bring about the numerator form of Q(z), we will give only one example for both forms.
We start again with the echelon form $P_E(z)$ of example 4.2:

$$P_E(z) = \begin{vmatrix} 2z+1 & z^3-z^2+z-1 \\ z^2 & z+1 \end{vmatrix}$$

The lowest power of z in the last (second) row of $P_E(z)$ is zero: $g_2=0$, and it appears in the second column: $c_2=$. Our next move is to zero certain powers of z: z^1 in the polynomials above $p_{22}(z)$. The degree 1 of these terms has to obey the following relation: $g_2 = 0 \leq 1 \leq 1 = g_2 + k_1 - k_2$, Therefore terms with z^0 and z have to be zeroed.

The result is:

$$P_D(z) = \begin{vmatrix} 1 & 1-2z \\ 0 & 1 \end{vmatrix} . P_E(z) = \begin{vmatrix} -2z^3+z^2+2z+1 & z^3-3z \\ z^2 & z+1 \end{vmatrix} \qquad (4.12)$$

The next operation would be to search in the q-1 st row of $P_E(z)$ for the lowest power of z. Because there are no rows above the first row of $P_E(z)$ however we can stop here.

It can now be understood why the limitation $g_q < l \leq g_q + k_j - k_i$ was introduced. This condition takes care that the row degrees of $P_E(z)$ do not change.

It becomes clear also after zeroing terms in a column, above a certain row i, that in the rows above row i, the lowest power of z is either lower than g_i or higher than $g_i + k_j - k_i$ $(k_j > k_i)$.

Working through all the rows, the number of possible values for g_i thus steadily decreases.

4.6 Guidorzi's form

When the irreducible MFD is brought to Guidorzis form (Beghelli and Guidorzi, 1976) then the numerator and denominator matrices have elements $p_{ij}(z)$ and $q_{ij}(z)$ which have the following form:

$$\begin{aligned}
p_{ii}(z) &= z^{\nu_{ii}} - a_{ii,\nu_{ii}} z^{\nu_{ii}-1} - \dots - a_{ii,2} z - a_{ii,1} \\
p_{ij}(z) &= -a_{ij,\nu_{ij}} z^{\nu_{ij}-1} - \dots - a_{ij,2} z - a_{ij,1} \\
q_{ij}(z) &= \beta_{(\nu_{11}+..+\nu_{ii}),j} z^{\nu_{ii}-1} + .. + \\
&\quad \beta_{(\nu_{11}+..+\nu_{i-1,i-1}+2),j} z + \\
&\quad \beta_{(\nu_{11}+..+\nu_{i-1,i-1}+1),j}
\end{aligned} \qquad (4.13)$$

The integers ν_{ii}, a_{ijk} and β_{ij} are related to the search for independent vectors in the observability matrix (F;H) and the expression of the dependent vectors in terms of selected independent ones.

The matrix elements $p_{ij}(z)$ and $q_{ij}(z)$ have to obey certain relations.

These relations are:

a) degree $p_{ii}(z)$ > degree $p_{ji}(z)$ $\qquad i \neq j$

b) degree $p_{ii}(z)$ > degree $p_{ij}(z)$ $\qquad j > i$

c) degree $p_{ii}(z)$ $\geq$ degree $p_{ij}(z)$ $\qquad j < i$ $\qquad$ (4.14)

d) degree $p_{ii}(z)$ > degree $q_{ij}(z)$

In words this means that for P(z): in each column the diagonal element has highest degree while in one row degrees of polynomials to the left of the diagonal element are at most equal to the degree of the diagonal element. To the right of the diagonal elements the polynomials have strictly lower degrees. Also, the row degrees of Q(z) are always strictly less than the corresponding column degrees of P(z).

4.6.1 The construction of Guidorzis canonical form

To achieve the canonical form of the MFD Guidorzi and Beghelli (1976) elaborated the following algorithm. Again it is noted that the denominator matrix of the MFD will have a particular structure after application of the algorithm. Performing the same sequence of operations on Q(z), that was performed on P(z), will complete the canonical MFD.

The first step is to bring P(z) to row proper form. This can be achieved (for every polynomial matrix - see Wolovich) by premultiplication with a suitable unimodular matrix M1. Now the row degrees of $P_1(z)=M_1(z).P(z)$ become the ordered set of Kronecker invariants associated to the pair $\{F,H\}$ of every observable state-space realization $\{F,G,H\}$ of (P(z),Q(z)). These indices remain unchanged in the subsequent steps.

Because $P_1(z)$ is row proper, we are able to place, in every row of $P_1(z)$, the polynomials, whose degree equals the row degree, on the main diagonal. This is achieved by exchanging rows, and can be described by premultiplication of $P_1(z)$ with real matrix M_2.

Achieving the row condition (4.14b) comes next. The entries $p_{m-1,m}$, $p_{m-2,m}, \dots, p_{1,m}, p_{m-3,m-1}, \dots, p_{1,m-1}, \dots, p_{1,2}$ are tested in the given order with respect to row condition (4.14b).
It is necessary to consider two situations: the first occurs when an off-diagonal element has a degree less than the degree of the diagonal element. In this case no operation is performed. The second case occurs when an off-diagonal element has a degree equal to the row degree.
If $\deg p_{ij}(z) = \nu_{ij} > \deg p_{jj}(z) = \nu_{jj}$, the degree of $p_{ij}(z)$ is lowered by subtracting from the ith row of P(z) the jth row multiplied by $\alpha\, z^{\nu_{ij}-\nu_{jj}}$, where α is the ratio of the maximal degree coefficients in $p_{ij}(z)$ and $p_{jj}(z)$.
If $\deg p_{ij}(z)= \nu_{ij} < \deg p_{jj}(z)= \nu_{jj}$ we replace the jth row by the ith row, and replace the ith row by the difference of the jth row and the ith row multiplied by $\alpha . z^{\nu_{jj}-\nu_{ij}}$, where α is the ratio of the maximal degree coefficients in $p_{jj}(z)$ and $p_{ij}(z)$. This last operation does not alter the row condition (4.14b) that was achieved for previously tested polynomials. Note that the row degrees are put in a potentially different order. This part of the algorithm can be described as the premultiplication of $P_2(z)$ with a unimodular matrix

$M_3(z)$: $\quad P_3(z)=M_3(z).P_2(z)$.

Now let us turn to the column conditon (4.14b) for the right upper triangular part of P(z). The polynomials are tested in the previously mentioned order.
If $\nu_{ij}<\nu_{jj}$ no operation is performed.
If $\nu_{ij}>\nu_{jj}$ the degree of $p_{ij}(z)$ is lowered by subtracting from the ith row of P(z), the jth row mulitiplied by $\alpha . z^{\nu_{ij}-\nu_{jj}}$, where α is the ratio of the maximal degree coefficients in $p_{ij}(z)$ and $p_{jj}(z)$.
After this operation the next polynomial in the given sequence is tested also if condition (4.14a) with respect to $p_{ij}(z)$ is not fulfilled.

Clearly, the degree of $\{p_{ij}(z) - \alpha . z^{\nu_{ij}-\nu_{jj}} . p_{jj}(z)$ is at most $(\nu_{ij}-1)$, but this can still be greater than ν_{jj}. The entire step must be repeated until all the polynomials in the

right upper triangular part of P(z) satisfy condition (4.14a). Again the operation from this step leaves the conditions attained for previously tested polynomials unchanged.
We can describe this step as the premultiplication of P(z) with a unimodular matrix $P_4(z)=M_4(z).P_3(z)$
To achieve column condition (4.14a) for the lower left triangular part , we consider the sequence of entries:
$p_{21}(z),p_{31}(z),\dots,p_{m1}(z),p_{32}(z),\dots,p_{m2}(z),\dots,p_{m,m-1}(z)$.
The mentioned polynomials will be tested in the given order.
If $\nu_{ij}<\nu_{jj}$ no operation is performed.
If $\nu_{ij}\geq\nu_{jj}$ the degree of $p_{ij}(z)$ is lowered by subtracting from the ith row of P(z), the jth row multiplied by $\alpha.z^{\nu_{ij}-\nu_{jj}}$ where α is the ratio of the maximal degree coefficients of $p_{ij}(z)$ and $p_{jj}(z)$. After this operation the next polynomial in the sequence must be tested also if condition (4.14a) with respect to $p_{ij}(z)$ and $p_{jj}(z)$ is not fulfilled. This entire step must be repeated until condition (5.10a) is met for the lower left triangular part of P(z). This step does not change all the conditions obtained in previous steps. It can be decribed as premultiplication of $P_4(z)$ with a unimodular matrix:

$$P_5(z)=M_5(z).P_4(z).$$

Finally the polynomials on the main diagonal of $P_5(z)$ are adjusted so that they become monic polynomials. This is simply done by dividing the ith row by the coefficients of $z^{\nu_{ii}}$. This can be translated as premultiplication of $P_5(z)$ with a real matrix:

$$P_6(z)=M_6(z).P_5(z).$$

The overall transformation from P(z) to the canonical form $P_6(z)$ is thus:

$$P_6(z)=M_6.M_5(z).M_4(z).M_3(z).M_2.M_1(z).$$

Now we can also calculate the corresponding numerator matrix

$$Q_6(z)=M_6.M_5(z).M_4(z).M_3(z).M_2.M_1(z).$$

Chapter 5 : STANDARD MFDS AND STATE-SPACE REALIZATIONS

It is possible to construct a minimal state-space realization of $T(z)$ from a given irreducible MFD, if $T(z)$ is a proper rational matrix (Dickinson,1974). To achieve this, the MFD must first be brought to a form in which the denominator matrix $P(z)$ is row proper.

In the second paragraph the relation of some unique MFDs with canonical state-space models is treated.

5.1 The construction of a minimal state-space model

If $T(z)$ is a strictly proper rational matrix having a standard MFD with degree n, there is a simple construction for obtaining an observable state-space realization of $T(z)$.

Let the row degrees of $P(z)$ be $k_1 > \ldots > k_q \quad 0$, $\sum k_i = n$.

Then $P(z)$ can be written:

$$P(z) = \left[\text{diag}\,(z^{k_1}, z^{k_2}, \ldots, z^{k_q}) + \Psi(z).P^*\right].\left[P(z)_h\right] \qquad (5.1)$$

$$\text{where } \Psi(z) = \left[\ \text{block diag}\ (\ z^{k_i - 1}, \ldots, z, 1)\right] \qquad (5.2)$$

so $\Psi(z)$ has diagonal blocks of dimension $1 x k_i$.

For example, if $k_1=2$, $k_2=4$ and $k_3=3$, we have:

$$\Psi(z) = \begin{vmatrix} z & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & z^3 & z^2 & z & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & z^2 & z & 1 \end{vmatrix}$$

P^* is an nxq matrix of coefficients.

Because $T(z)$ is strictly proper, the row degrees of $Q(z)$ will be strictly less than the row degrees of $P(z)$, so that $Q(z)$ may be written as

$$Q(z) = \Psi(z).\ Q^*, \qquad (5.3)$$

as above and Q^* is an nxp matrix of coefficients. With this notation it can be verified by direct calculation that an n-dimensional realization of the transfer function T(z) is

given by the triple (F,G,H) where

$$
\begin{aligned}
F &= \left[\, F_0 - P^*.H_0 \right] \\
G &= Q^* \\
H &= [P(z)_h]^{-1}.H_0
\end{aligned}
\qquad (5.4)
$$

and $F_0 = \left[\text{ block diag} \begin{vmatrix} 0 & 1 & 0 & & 0 \\ 0 & & 1 & & . \\ . & & & . & \\ 0 & . & . & . & 1 \\ 0 & . & . & . & 0 \end{vmatrix} \right]$ (5.5)

and the diagonal blocks have dimensions $k_i x k_i$.

$$H_0 = \left[\text{ block diag } (\, 1 \; 0 \, ... \, 0 \,) \right] \qquad (5.6)$$

and the diagonal blocks have dimensions $1xk_i$.
It is clear (from the structure of H_0) that this realization is observable. Also, the irreducibility of the MFD (P(z),Q(z)) ensures the controllability of the realization (F,G,H). Thus follows the important fact that the degree of det P(z), in an irreducible MFD of T(z), is the dimension of a minimal realization of T(z).
We can observe that in the SISO case, where the transfer function is a quotient of 2 prime polynomials, the order of a minimal realization equals the degree of the denominator polynomial.
The realization is said to be in observer form.
Analogously we can derive a controller form for a right standard irreducible MFD of T(z).
Note that we obtain a different observer form for every standard irreducible MFD of T(z).
If T(z) is proper but not strictly proper, let

$$E^* = \lim_{z \to \infty} T(z) \qquad (5.7)$$

and let (P(z),Q(z)) be a standard MFD of (T(z) - E*). Then an observer form realization of T(z) is given by (F_0, G_0, H_0, E^*) which describes the state equations:

$$
\begin{aligned}
x(k+1) &= F_0.x(k) + G_0.u(k) \\
y(k) &= H_0.x(k) + E^*.u(k)
\end{aligned}
\qquad (5.8)
$$

Example 5.1

Consider the transfer function $T(z)=\dfrac{z+2}{z^2+z+1}=\dfrac{q(z)}{p(z)}$, (5.9)

for a system with one input and one output.

The row degree is in this case : $k_1 = 2$.

First we write p(z) according to (5.1), and find that

$$p(z)=[\, z^2 + [z \quad 1]\cdot\begin{vmatrix} 1 \\ 1 \end{vmatrix}\,]\cdot[1] \text{ , and therefore:} \tag{5.10}$$

$$\psi(z)=[z \quad 1] \text{ , } p^*=\begin{vmatrix} 1 \\ 1 \end{vmatrix} \text{ and } [p(z)]_h=1 \text{ .} \tag{5.11}$$

Using these matrices we can calculate q*:

$$q(z)=[z\ 1]\cdot\begin{vmatrix} 1 \\ 2 \end{vmatrix} \text{ , so } q^*=\begin{vmatrix} 1 \\ 2 \end{vmatrix} \text{ .} \tag{5.12}$$

A minimal state-space realization is given by:

$$F=\begin{vmatrix} -1 & -1 \\ -1 & 0 \end{vmatrix},\quad G=\begin{vmatrix} 1 \\ 2 \end{vmatrix},\quad H=[1\ 0]. \tag{5.13}$$

This can be verified by calculating $H(zI-F)^{-1}G$. $(=T(z))$

A somewhat different construction of a state-space realization starts with the Guidorzi form of the MFD. This construction will be given in the following section. This construction can be reversed to obtain a standard MFD from a given observable state-space description.

Wolovich (1974) describes the construction of a MFD from an observa state-space realization, along the same line as is treated here. Onble difference is, that Wolovich starts with a triangular matrix H in e stead of a block diagonal one.

5.2 The relation of unique MFD to canonical state-space models

Dickinson et al. (1974) showed the correspondence of the first 3 unique MFDs to canonical state-space descriptions. The relation was

obtained by selecting certain rows of the observability matrix as basis vectors for the state-space. The selection of the rows was accomplished according to the entries of the matrix $P_E(z)$, $P_H(z)$ and $P_R(z)$.
Finally the relation of Guidorzi's form to a canonical state space will be described.

First we will consider the echelon form.
Using the observer form realization of the transfer function, we obtain

$$H(zI - F)^{-1}G = P_E^{-1}(z)Q_E(z) \tag{5.14}$$

and because $Q_E(z) = \Psi(z)G$, the input matrix G can be cancelled from the equation :

$$H(zI - F)^{-1} = P_E^{-1}(z)\ \Psi(z) \tag{5.15}$$

$$P_E(z)H(zI - F)^{-1} = \Psi(z) \tag{5.16}$$

Now writing $P_E(z)$ in terms of its coefficients gives:

$$P_E(z) = \sum_{i=0}^{k_1} P_i z^{k_1 - i} \tag{5.17}$$

(bearing in mind that $k_1 = \{\max(k_i)\}$) and by equating coefficients of z^{-1} on both sides of (5.16) we obtain:

$$\left[\ P_0\ P_1 \dots P_{k_1}\ \right] \left| \begin{array}{l} HF^{k_1} \\ HF^{k_1 - 1} \\ \vdots \\ H \end{array} \right| = [\ 0\ ] \tag{5.18}$$

Here use has been made of the fact that

$$(zI-F)^{-1} = z^{-1}(I-z^{-1}.F)^{-1} = z^{-1}\left[I+z^{-1}.F + z^{-2}.F^2 + \dots .\right] \tag{5.19}$$

and of the fact that the eigenvalues of F will be smaller than 1. Thus the entries of $P_E(z)$ express linear dependence relationships among the rows of the observability matrix of any minimal realization of $T(z)$.
The special structure of $P_E(z)$ corresponds to a special set of

linear dependence relations. Consider row i of $P_E(z)$.
The highest degree coefficient vector weights
$h_{s_i} F^{k_i}$ with a coefficient of 1, where s_i is the pivot index in the definition of the echelon form. Other vectors included in the relation are $h_l F^{k_i}$ for $l < s_i$ and $h_{s_l} F^j$ for $0 \leq j \leq k_l$, and $l \leq q$, because the polynomial $p_{s_l i}(z)$ has a lower degree then k_l.
The set of linear dependence relationships in (5.18) can be compactly written as:

$$h_{s_i} p_{ij}(F) = 0 \qquad\qquad j \leq q \qquad\qquad (5.20)$$

This is exactly the same set of dependence relationships that would be obtained by searching in lexicographic order for linearly independent vectors in the observability matrix of (F,H) in the style of Luenberger and Popov. This establishes the connection between the echelon MFD and one state-space construction for canonical forms (Denham, 1974). We may note that the integers $\{k_i\}$ are the Kronecker invariants associated with (F,H).

Next the Hermite MFD can be connected to a second common procedure for searching the rows of the observability matrix for independent vectors.
This scheme rearranges the observability matrix as:

$$\tilde{\Phi}^T = [\ h_1^T \quad F^T h_1^T .. (F^{n-1})^T h_1^T \quad h_2^T \ ..... \ (F^{n-1})^T h_p^T \ ] \qquad (5.21)$$

and then searches $\tilde{\Phi}$ in lexicographic order. To establish this connection we cannot use the observer form construction because $P_H(z)$ is not row proper. However,by the equivalence of both matrix fractions,

$$P_H(z) = U(z) P_E(z), \qquad Q_H(z) = U(z) Q_E(z) \qquad\qquad (5.22)$$

for some unimodular matrix U(z).
Using this fact in (5.14) and (5.15) gives the counterpart to

$$P_H(z) H(zI-F)^{-1} = \Psi(z).U(z) \qquad\qquad (5.23)$$

and the same procedure as in the echelon form case applies. Since $P_H(z)$ is lower triangular with monic diagonal elements of higher degree than other elements in their respective columns, it does indeed produce the linear dependence relations obtained by searching in (5.21). The independent vectors obtained by this procedure can then be used to provide a basis for the state-space. The matrix F with respect to this basis consists of triangularly coupled companion matrices. This was also one of the schemes described by Luenberger. (Denham,1974 ; Luenberger, 1967)

The reversed echelon form again provides a set of linear dependence relations, but this time the matrix P_{k_1} in (5.18) weights the various rows of H with 1. However by successively post-multiplying (5.18) by F^{-k_1+1}, F^{-k_2+1}, ...F^{-k_p+1}, it becomes clear that $P_R(z)$ gives the same choice of basis vectors as the echelon approach applied to the pair $(F^{-1}, F^{-1}H)$. Of course these basis vectors may be used as a basis for the original state-space, giving a block companion form whose coefficients are the same as the coefficients in the output matrix.
In the scalar case, this form was given by Chidambara (1971). It may be regarded as a "constructibility" canonical form.

A very simple construction of an observable state-space description is possible once we have a MFD of the system transfer function in Guidorzi's form.(Beghelli and Guidorzi ,1976)
By inspection we can read from the MFD the values of the parameters $\nu_{ii}, \nu_{ij}, a_{ijk}, \beta_{ij}$ and these are all that is needed.
First we construct the auxiliary matrix M with the use of the parameters a_{ijk} as follows.
The diagonal blocks of M and the off-diagonal blocks are as follows:

$$M_{ii} = \begin{vmatrix} -a_{ii,2} & -a_{ii,3} & \cdots & -a_{ii,\nu_{ii}} & 1 \\ -a_{ii,3} & \cdots & & 1 & 0 \\ \cdot & & \cdot & & \cdot \\ -a_{ii,\nu_{ii}} & 1 & & & \cdot \\ 1 & 0 & \cdot & \cdot & 0 \end{vmatrix} \qquad (5.24)$$

$$\text{and } M_{ij} = \begin{vmatrix} -a_{ij,\nu_{ij}} & \cdots\cdots & -a_{ij,\nu_{iq}} & 0 \\ \cdot & & 0 & 0 \\ \cdot & \cdot & & \cdot \\ -a_{ij,\nu_{ij}} & \cdot & & \cdot \\ 0 & & & \cdot \\ \cdot & & & \\ 0 & \cdot \quad \cdot & \cdot & 0 \end{vmatrix} \tag{5.25}$$

M_{ii} has dimension $\nu_{ii} x \nu_{ii}$ and M_{ij} has dimensions $\nu_{ii} x \nu_{ij}$.
Note that M is always nonsingular , because of its structure, and the fact that det M = 1.

Now we construct the matrix $B^* = \begin{vmatrix} \beta_{11} & \beta_{1p} \\ \beta_{n1} & \beta_{np} \end{vmatrix}$ (5.26)

Now it is possible to calculate the input matrix G as

$$G = M^{-1}.B^* \tag{5.27}$$

The system matrix F can be written by inspection of the matrix P(z)

$F = [F_{ij}]$ and

$$F_{ii} = \begin{vmatrix} 0 & 1 & & \cdot \\ \cdot & & 1 & \cdot \\ 0 & & & 1 \\ a_{ii,1} & \cdots & & a_{ii,\nu_i} \end{vmatrix} \qquad F_{ij} = \begin{vmatrix} 0 & \cdots & 0 & 0 \cdot 0 \\ \cdot & & \cdot & \cdot\ \cdot\ \cdot \\ 0 & & 0 & 0 \cdot 0 \\ a_{ij,1} & \cdots & a_{ij,\nu_{ij}} & 0 \cdot 0 \end{vmatrix} \tag{5.28}$$

Finally the matrix H is constructed from knowledge of the parameters ν_{ii}:

$$H = \begin{vmatrix} 1 & 0 & \cdot\cdot & 0 & & \cdot\cdot & 0 & & \cdot\cdot & & & 0 \\ 0 & 0 & \cdot\cdot & 1 & 0 & \cdot\cdot & 0 & & \cdot\cdot & & & 0 \\ 0 & & \cdot\cdot & 0 & & \cdot\cdot & 1 & 0 & \cdot\cdot & & & 0 \\ \vdots & & & \vdots & & & \vdots & & \vdots & & & \vdots \\ 0 & & \cdot\cdot & 0 & & \cdot\cdot & 0 & \cdot\cdot & 1 & 0 & \cdot\cdot & 0 \end{vmatrix} \tag{5.29}$$

The H-matrix consists of the first, ν_{11}+1-st, ..., $\nu_{11}+..+\nu_{m-1m-1}$+1st rows of the identity matrix.

Chapter 6 : The instrumental variable method for parameter estimation

The instrumental variable technique (IV) has been given much attention since the early 1960's.
The IV-approach to parameter estimation of time-series analysis was developed initially as a modification to the linear least squares regression solution. In particular it was needed for those cases where asymptotic estimation bias proved to be a problem.
It has the remarkable property that there is no need to model the noise. This means that the total number of parameters is less than in cases where the noise model has to be estimated also.
The term instrumental variable is not very sharply defined and therefore a variety of estimation procedures is comprised in the "IV-method".
The method has its origins in the statistical literature, where it has reveived considerable attention. From there it has found its way into control engineering.

Probably the first contribution on IV to the control literature was made by Joseph et al. (1961) although it was not explicitly called an IV method. From then on more attention was paid to the use and study of IV techniques. As a consequence, recursive estimation algorithms were developed in the field of discrete time-series models, by Mayne (1967) and Wong and Polak (1967). A related technique, known as the "tally principle", was discussed by Peterka and Smuk (1969). Rowe (1969) suggested a so-called "bootstrap"* method in a multivariable version to estimate the statistical properties of the noise from the IV residuals. Young (1971) presented the recursive approximate maximum likelihood method to estimate the coefficients of an ARMA-model of the noise.
These early publications were mainly concerned with single-input single-output systems. Although quite useful for these SISO systems, the theory had to be supplemented and extended to the multivariable case to be of use in more practical situations like in industrial processes.

A more recent discussion of various IV techniques is by e.g. Jakeman

(1979), who compares several IV procedures for multivariable systems. Stoica and Söderström (1982) study the consistency and optimality of IV methods for multivariable systems. Other contributions are by Kashyap and Nasburg (1974), Jakeman and Young (1979), Sinha and Caines.

* A bootstrap-method alternates the estimation of process parameters and noise parameters.

(1977), Gauthier and Landau (1978), and Rowe (1970).
A number of variants are derived by De Keyser (1979) from a general set of equations.
In contrast to the others Stoica and Söderström (1983) present a rather general and coherent treatment of instrumental variable theory.

6.1 The multivariable time series model

As a basis for our estimation procedure we use a multivariable discrete time-invariant linear system which is described by:

$$P^*(z^{-1})x(k) = Q^*(z^{-1})u(k) \tag{6.1}$$

$$y(k) = x(k) + \xi(k) \tag{6.2}$$

Here we have vectors:

$$\begin{aligned} x^T(k) &= [x_1(k) \quad x_2(k) \quad \dots \quad x_q(k)] \\ u^T(k) &= [u_1(k) \quad u_2(k) \quad \dots \quad u_p(k)] \\ y^T(k) &= [y_1(k) \quad y_2(k) \quad \dots \quad y_q(k)] \\ \xi^T(k) &= [\xi_1(k) \quad \xi_2(k) \quad \dots \quad \xi_q(k)] \end{aligned} \tag{6.3}$$

The argument $k=1,2,\dots,N$ denotes the sampling moment and z^{-1} is the backward shift operator. $x(k)$ is a q-dimensional vector representing the noisefree system output at time k, which is inherently unobservable. $u(k)$ is the p-dimensional input vector, which is assumed to be known exactly. $y(k)$ is a q-dimensional output vector, which will be considered as the superposition of the noisefree system output $x(k)$ and a noise signal $\xi(k)$ (the output error), representing the effects of all unmeasurable disturbances and measurement noise on the system. N is the number of available data points.

The qxq matrix $P^*(z^{-1})$ and the qxp matrix $Q^*(z^{-1})$ are related to the P(z) and Q(z) matrix from equation (2.6) in the following way. Take row i from P(z) and Q(z) and multiply its elements by z^{-k_i}, where k_i is the degree of the monic polynomial in row i from P(z). Repeat this for all rows of P(z) and Q(z).
The result can be written compactly as:

$$\begin{aligned} P^*(z^{-1}) &= \text{diag}\,[z^{-k_1}, z^{-k_2}, \ldots, z^{-k_q}].P(z) \\ Q^*(z^{-1}) &= \text{diag}\,[z^{-k_1}, z^{-k_2}, \ldots, z^{-k_q}].Q(z) \end{aligned} \tag{6.4}$$

In case the degrees of the monic polynomials equal the row degrees (which means that no polynomial in a row has degree that is higher than that of the monic polynomial), the model can be constructed exclusively from time-delays.
An exception is given by the Hermite parametrization where the monic polynomial does not necessarily have row degree. In this case interpretation of the products like in (6.4) is more difficult.
We will use only the notation of $P^*(z^{-1})$ however, even if this notation is not accurate for the Hermite parametrisation (due to the occurence of terms z, z^2, etc.)

Remark: if the polynomial $p_{ij}(z)$ of P(z) is the monic polynomial in row i, then by definition the coefficient of $z^{\nu_{ij}}$ is a 1.

This means that in $P^*(z^{-1})$ this polynomial becomes $z^{-k_i}.p_{ij}(z)$ and the coefficient of z^0 in this product becomes a 1, and this term corresponds to the lowest power of z^{-1}. Therefore it is not correct to call $z^{-k_i}.p_{ij}(z)$ monic .
Remark 2: it can easily be verified that the coefficients appearing

in $P(z)$ are also describing $P^*(z^{-1})$, since only the powers of the indeterminate z are altered.

In most publications (e.g. Jakeman(1979), Jakeman and Young(1979), Stoica and Söderström (1982)) an ARMA-model is used to create filtered white noise and disturb the system output $x(k)$ with it. For our present purpose the noise filtering is not taken into account, but can be implemented without great difficulty. In the simulations that were made, $\xi(k)$ is always a Gaussian-distributed white noise signal.

For application of the IV-method it is essential that the model is linear-in-the-parameters : in fact this means that we can write for all unique models as given in chapter 4 :

$$P^*(z^{-1})x(k) = Q^*(z^{-1})u(k)$$

$$x^m(k) = [I-P^*(z^{-1})T^T]x_m(k) + Q^*(z^{-1})u(k)$$

$$= \Omega^T(k).\Theta \qquad (6.5)$$

where Θ is a $n_\Theta x1$ vector containing all n_Θ parameters of the matrices $P^*(z^{-1})$ and $Q^*(z^{-1})$, and $\Omega(k)$ is a $n_\Theta xq$ matrix containing linear combinations of time-shifted input- and noise free system output-signals, and T is the permutation matrix.

The vector $x^m(k)$ represents the noise free system output with an ordering of the components $x_i(k)$ of the vector $x(k)$ which differs potentially from equation (6.1). The ordering depends on which polynomial is monic in a certain row of $P(z)$. E.g. if in row 1 the monic polynomial is in the third column, then the first component of $x^m(k)$ will be $x_3(k)$, because $x_3(k)$ can be written as a linear combination (determined by the polynomials $p_{ij}(z)$ and $q_{ij}(z)$) of other inputs and outputs. In this case the scalar matrix T has a 1 in the third column, and zeros in the remaining places of the first row. Thus the matrix T is obtained as a permutation of the identity matrix.

For those parametrizations, for which the monic polynomials appear as diagonal elements of the matrix $P(z)$, we have $T = I$, and $x^m(k)=x(k)$.

For the echelon parametrisation, T depends on the relation between the degrees of the monic polynomial.

Similarly to (6.5) we can write:

$$y^m(k) = \tilde{\Omega}^T(k).\Theta + \eta(k) \tag{6.6}$$

where $\tilde{\Omega}^T(k)$ now contains delayed noisy outputs y(k) in stead of x(k), and $\eta(k)$ represents the residuals vector, or equation error. The relation between $\eta(k)$ and $\xi(k)$ can be found by combining equations (6.1) and (6.5). The vector $y^m(k)$ is obtained by the same permutation that reorders the components of x(k).
From (6.1) it follows that:

$$y(k) = P^{*-1}(z^{-1}).Q^*(z^{-1}).u(k) + \xi(k)$$

$$P^*(z^{-1})y(k) = Q^*(z^{-1})u(k) + P^*(z^{-1})\xi(k)$$

$$y^m(k) = \tilde{\Omega}^T(k)\Theta + \eta(k) \tag{6.7}$$

From this equation we can see that the residuals are a filtered version of the noise $\xi(k)$, and that they depend on the model that is used (expressed in the matrix $P^*(z^{-1})$).

A number of requirements have to be met in order to allow successful estimations:

1) The system must be asymptotically stable and the transfer matrix $P^{*-1}(z^{-1})Q^*(z^{-1})$ must be strictly proper.

2) The input signal u(k) is a stationary signal of sufficiently high order of complexity. If the order of the signal is too low, the information content is insufficient to enable accurate parameter estimation, due to a lack of degrees of freedom.
A step signal is an example of such an unsuited input. This is why we use a Gaussian white noise signal as input.

3) The noise $\eta(k)$ is a zero mean, stationary stochastical process. It is uncorrelated with the input signal, and can be described as:

$$E\{\eta_i(k)\} = 0 \qquad \text{(E is the expectation operator)} \tag{6.8}$$

$$E\{\eta_i(k)u_j^T(l)\} = 0 \qquad \text{for all values of i and j} \tag{6.9}$$

The noise process can be modelled as :

$$\eta(k) = H(z^{-1})e(k) \tag{6.10}$$

where $\{e(k)\}$ is a sequence of independent and identically distributed random vectors with zero mean, and nonsingular covariance matrix Λ. $H(z^{-1})$ is an asymptotically stable filter. (See Anderson and Moore (1979), Stoica and Söderström (1983))

4) The matrix polynomial coefficients $p_{ij,m}$ and $q_{ij,m}$ in the matrices $P^*(z^{-1})$ and $Q^*(z^{-1})$ are the parameters that describe the system. Thus our error is linear in the parameters.

As an example for equation (6.6) suppose q=1, p=2, and

$$P(z) = z^2+p_2z+p_1 \;, \quad Q(z)= [q_{11,2}z+q_{11,1} \quad q_{12,2}z+q_{12,1}]$$

Then we have :

$$P^*(z^{-1}) = 1 + p_2z^{-1} + p_1z^{-2}$$

$$Q^*(z^{-1}) = [q_{11,2}z^{-1} + q_{11,1}z^{-2} \quad q_{12,2}z^{-1} + q_{12,1}z^{-2}]$$

$$x(k)=x_1(k) \;, \qquad y(k)=y_1(k) \;, \qquad u(k)=\begin{vmatrix} u_1(k) \\ u_2(k) \end{vmatrix}$$

The system equation (6.1) becomes:

$$[z^2+p_2z+p_1]x_1(k) = [q_{11,2}z+q_{11,1} \qquad q_{12,2}z+q_{12,1}]\begin{vmatrix} u_1(k) \\ u_2(k) \end{vmatrix}$$

Multiplying both sides of the last equation by z^{-2} (considering only time-delays), and performing the timeshift operation we have:

$$x_1(k)+p_2x_1(k-1)+p_1x_1(k-2) = q_{11,2}u_1(k-1)+q_{11,1}u_1(k-2)+ q_{12,2}u_2(k-1)+q_{12,1}u_2(k-2)$$

This can be written in a form corresponding to (6.5)

$$x_1(k) = \begin{vmatrix} -x_1(k-2) \\ -x_1(k-1) \\ u_1(k-2) \\ u_1(k-1) \\ u_2(k-2) \\ u_2(k-1) \end{vmatrix}^T \cdot \begin{vmatrix} p_1 \\ p_2 \\ q_{11,1} \\ q_{11,2} \\ q_{12,1} \\ q_{12,2} \end{vmatrix} = \Omega^T(k) \cdot \Theta$$

Note that an arbitrary polynomial of degree d has d+1 unknown parameters (coefficients), while structurally monic polynomials in the matrices
P(z) and Q(z) have d unknown parameters.

For the estimation procedure it is of practical importance that the parametrisation is such that we can write every equation i from the set of equations of (6.6) as:

$$y_j^m(k) = \tilde{\Omega}_i^T(k)\Theta_i + \eta_i(k) \qquad (i=1,2,\ldots,q) \qquad (6.11)$$

In this equation $\tilde{\Omega}_i(k)$ is a n_{θ_i} x1 vector containing delayed noisy outputs and inputs, whereas Θ_i contains the parameters of the polynomials of the i-th row of $P^*(z^{-1})$ and the i-th row of $Q^*(z^{-1})$, so.every output can be described by an independent parameter vector.

Note the occurence of output y_j^m in equation (6.11)
For some model parametrisations (e.g. the Hermite form) the i-th row gives an expression for $y_i(k)$ of the type (6.12), while for others (e.g. the echelon form) it depends on the requirements of the specific parametrisation , which output y_j^m is associated to equation i in a way as described by (6.11). In fact, the structurally monic polynomial $p_{ij}(z)$ in row i of P(z) determines which output y_j^m appears in (6.11).
In the section describing the various parametrisations the reader may verify that there is one and only one monic polynomial in every row of P(z), and that they occur in different columns.

Writing (6.6) in terms of (6.11) we find that

$$\tilde{\Omega}^T(k) = \begin{vmatrix} \tilde{\Omega}_1^T(k) & & & \\ & \tilde{\Omega}_2^T(k) & & \\ & & \cdot & \\ & & & \tilde{\Omega}_q^T(k) \end{vmatrix} \quad \text{and } \Theta = \begin{vmatrix} \Theta_1 \\ \Theta_2 \\ \cdot \\ \Theta_q \end{vmatrix} \tag{6.12}$$

Because we need the expression (6.11) for the estimation algorithm we take a detailed look at it.

Consider the i-th row of P(z). The following equation is associated to it:

$$[p_{i1}(z) \;\cdots\; p_{iq}(z)] \begin{vmatrix} x_1(k) \\ \cdot \\ x_q(k) \end{vmatrix} = [q_{i1}(z) \;\cdots\; q_{ip}(z)] \begin{vmatrix} u_1(k) \\ \cdot \\ u_p(k) \end{vmatrix} \tag{6.13}$$

The left hand side of this expression is:

$p_{i1}(z)x_1(k) + p_{i2}(z)x_2(k) + \dots + p_{iq}(z)x_q(k)$

Assume that $p_{ij}(z)$ is monic. This polynomial has degree ν_{ij}.

Substituting the polynomials we have:

$$\begin{aligned} &\{p_{i1,\nu_{i1}+1} z^{\nu_{i1}} + \dots + p_{i1,2} z + p_{i1,1}\} x_1(k) + \dots \\ &\qquad \{1.z^{\nu_{ij}} + \dots + p_{ij,2} z + p_{ij,1}\} x_j(k) + \dots \\ &\{p_{iq,\nu_{iq}+1} z^{\nu_{iq}} + \dots + p_{iq,2} z + p_{iq,1}\} x_q(k) \end{aligned} \tag{6.14}$$

Keep in mind that for the Hermite form there is no direct relation between ν_{il} $(l \neq j)$ and the degree of the diagonal element of P(z) : ν_{ii}. For the other parametrisations we have that $\nu_{il} < \nu_{ij}$ depending on the value of j with respect to i.

For models for which the row degrees equal the degree of the monic polynomials, the degrees of the polynomials $q_{ij}(z)$ are strictly less than the degree of the monic polynomial: ν_{ij}.

The right hand side of (6.11) then becomes:

$$\{ q_{i1,\nu_{ij}} z^{\nu_{ij}-1} + \dots + q_{i1,2} z + q_{i1,1}\} u_1(k) + \dots +$$

$$\{q_{ip,\nu_{ij}} z^{\nu_{ij}-1} + \dots + q_{ip,2} z + q_{ip,1}\} u_p(k) \qquad (6.15)$$

Multiplying both sides of the equation by $z^{-\nu_{ij}}$ and applying the time shift operation we have the following equation:

$$p_{i1,\nu_{i1}+1} x_1(k+\nu_{i1}-\nu_{ij}) + \dots + p_{i1,2} x_1(k+1-\nu_{ij}) + p_{i1,1} x_1(k-\nu_{ij}) +$$
$$+ \dots +$$
$$x_j(k) + \quad \dots \quad + p_{ij,2} x_j(k+1-\nu_{ij}) + p_{ij,1} x_j(k-\nu_{ij}) +$$
$$+ \dots +$$
$$p_{iq,\nu_{iq}+1} x_q(k+\nu_{iq}-\nu_{ij}) + \dots + p_{iq,2} x_q(k+1-\nu_{ij}) + p_{iq,1} x_q(k-\nu_{ij}) =$$
$$q_{i1,\nu_{ij}} u_1(k-1) + \dots + \quad q_{i1,2} u_1(k+1-\nu_{ij}) + q_{i1,1} u_1(k-\nu_{ij}) +$$
$$+ \dots +$$
$$q_{ip,\nu_{ij}} u_p(k-1) + \dots + q_{ip,2} u_p(k+1-\nu_{ij}) + q_{ip,1} u_p(k-\nu_{ij}) \qquad (6.16)$$

Now we are able to write $y_j^m(k)$ in a form corresponding to (6.11), where $y^m(k)$ replaces $x(k)$ (in (6.16)) and residuals $\eta(k)$ are also added.

Expression (6.13) gives the general formula for the noisefree model. From this expression we can derive de vectors $\tilde{\Omega}_i$ and Θ_i:

$$\tilde{\Omega}_i(k) = \begin{vmatrix} -y_1(k+\nu_{i1}-\nu_{ij}) \\ -y_1(k+\nu_{i1}-1-\nu_{ij}) \\ \cdot \\ -y_1(k-\nu_{ij}) \\ -y_2(k+\nu_{i2}-\nu_{ij}) \\ \cdot \\ \cdot \\ -y_{j-1}(k-\nu_{ij}) \\ -y_j^m(k-1) \\ \cdot \\ -y_q(k+\nu_{iq}-\nu_{ij}) \\ \cdot \\ -y_q(k-\nu_{ij}) \\ u_1(k-1) \\ u_1(k-2) \\ \cdot \\ u_1(k-\nu_{ij}) \\ u_2(k-1) \\ \cdot \\ \cdot \\ u_p(k-1) \\ \cdot \\ u_p(k-\nu_{ij}) \end{vmatrix} \qquad \theta_i = \begin{vmatrix} p_{i1,\nu_{i1}+1} \\ p_{i1,\nu_{i1}} \\ \cdot \\ p_{i1,1} \\ p_{i2,\nu_{i2}+1} \\ \cdot \\ \cdot \\ p_{i(j-1),1} \\ p_{ij,\nu_{ij}} \\ \cdot \\ p_{iq,\nu_{iq}+1} \\ \cdot \\ p_{iq,1} \\ q_{i1,\nu_{ij}} \\ q_{i1,\nu_{ij}-1} \\ \cdot \\ q_{i1,1} \\ q_{i2,\nu_{ij}} \\ \cdot \\ \cdot \\ q_{ip,\nu_{ij}} \\ \cdot \\ q_{ip,1} \end{vmatrix} \tag{6.17}$$

Of course the term $y_j^m(k)$ is absent from (6.17), because it appears on the left hand side of (6.12).

For the Hermite form it may also occur that $\nu_{il} > \nu_{ij}$, $l \neq j$, which implies that $y_j(k)$ is a function of $y_i(k+s)$, where $s=1,2,..,\nu_{il}-\nu_{ij}$, $l \neq j$.
In other words, $y_j(k)$ is a function of future values of $y_i(k)$.
For the (off line) estimation algorithm this is no problem because we have all samples available on record. For the other forms, we have: $\nu_{il} < \nu_{ij}$ for $l \neq j$.

6.2 The instrumental variable method

We start with the system equation

$$y(k) = \tilde{\Omega}^T(k)\Theta + \eta(k)$$

Let us introduce the "instrumental variable matrix" Z(k) of dimensions $n\Theta xq$. This matrix must fulfill two conditions:

1) $E\{Z(k).\eta(k)\}=0$ (6.18)

This condition tells us that the instrumental variables have to be correlated with the in- and output signals of the system. This condition is necessary to obtain a consistent estimate for Θ_t.

A consistent estimator Θ is such that $\lim_{N\to\infty} \Theta = \Theta_t$, where Θ_t is the "true" parameter vector, which has to belong to the model set.

The use of these conditions will become apparent, if we premultiply (6.6) by Z(k) and perform the expectation operation:

$$\begin{aligned} E\{Z(k)y^m(k)\} &= E\{Z(k)\tilde{\Omega}^T(k)\Theta_t + Z(k)\eta(k)\} \\ &= E\{Z(k)\tilde{\Omega}^T(k)\Theta_t\} + E\{Z(k)\eta(k)\} \\ &= E\{Z(k)\tilde{\Omega}^T(k)\}\Theta_t \end{aligned} \qquad (6.20)$$

Of course we can only approach the expectation operation by averaging over a (limited) number of samples, which gives

$$1/N \sum_{k=1}^{N} Z(k)y^m(k) = 1/N \left\{ \sum_{k=1}^{N} Z(k)\tilde{\Omega}^T(k). \right\}\hat{\Theta} \qquad (6.21)$$

and we find a solution for $\hat{\Theta}$:

$$\hat{\Theta} = \left\{ \sum_{k=1}^{N} Z(k)\tilde{\Omega}^T(k) \right\}^{-1} \left\{ \sum_{k=1}^{N} Z(k)y^m(k) \right\} \qquad (6.22)$$

In this last step we have used condition 2) for the matrix inversion.

It is possible to separate this solution in q independent parameter-vectors $\hat{\Theta}_i$ which can be computed (by using (6.9)) independently as the solution of

$$\hat{\Theta}_i = \left\{ \sum_{k=1}^{N} Z_i(k)\tilde{\Omega}_i^T(k) \right\}^{-1} \left\{ \sum_{k=1}^{N} Z_i(k)y_j^m(k) \right\} \qquad (6.23)$$

$$i=1,2,\dots,q$$

$Z_i(k)$ is the IV-matrix of dimensions $n_{\Theta_i} xq$.

The relation between Z(k) and $Z_i(k)$ is:

$$Z(k) = \begin{vmatrix} Z_1(k) & & & \\ & Z_2(k) & & \\ & & \cdot & \\ & & & Z_q(k) \end{vmatrix} \qquad (6.24)$$

The advantage of (6.23) over (6.22) lies in the fact that the computation of (6.23) can be faster, and the total number of arithmetical operations is smaller. In general the number of arithmetic operations is proportional to the cube of the number of unknown parameters (see Stoica and Söderström (1982)).
So provided the vectors Θ_i have approximately the same dimensions the computation of (6.23) involves approximately a fraction $\frac{q\,(n\Theta_i)^3}{(q.n\Theta_i)^3}$ or $1/q^2$ of the operations involved in (6.22).

It is possible to modify the IV estimate into an iterative scheme, which might give rise to more accurate parameter estimates. Extended IV-procedures have been given by e.g. Stoica and Soderstrom (1983).
One of the alternatives is to apply prefiltering of the data.
Then we have :

$$Z_p(k) = Z(k).F(z^{-1}) \tag{6.25}$$

where $F(z^{-1})$ represents the asymptotically stable filter.
A second possibility is the use of a "longer" Z(k) matrix, i.e. with a number of rows that is greater than the number of parameters to be computed. The result is that the system of n_Θ equations (6.22) becomes overdetermined, and a solution Θ can be found in a least squares sense.
Therefore the general expression for the extended IV-estimator is:

$$\Theta = \arg\min_{\Theta} \left\| \left\{ \sum_{k=1}^{N} Z(k)F(z^{-1})\tilde{\Omega}^T(k) \right\}\Theta - \left\{ \sum_{k=1}^{N} Z(k)F(z^{-1})y(k) \right\} \right\|_{\Gamma}^{2} \tag{6.27}$$

where $\|x\|_\Gamma^2 = x^T\Gamma x$, and Γ is a positive definite weighting matrix, and arg denotes the argument of a function (e.g. arg f(x) = x).

The consistency conditions for this extended estimator become:

1a) $E\{Z(k)F(z^{-1})\eta(k)\} = 0$ (6.27)

2a) $E\{Z(k)F(z^{-1})\tilde{\Omega}^T(k)\}$ exists and has rank $n\Theta$ (full rank)

6.3 Optimal IV and the choice of instruments

In this section we will describe the entries (instruments) of the $Z(k)$ matrix. In principle, any signal, that fulfills the two consistency conditions will be a good one. This does not mean, however, that every IV-variant gives equally accurate results. Therefore it is necessary to search for optimal instruments. Stoica and Soderstrom (1983) have developed expressions for optimal IV matrices, in terms of the noise covariance matrix, the noise filter and the matrix Ω which contains inputs and noise free outputs of the system. The optimality is formulated as a lower bound for the covariance matrix of the parameter estimates, which is reached for two special IV variants.

For completeness' sake the results of these authors will be cited, but they will not be investigated further.
Consider the estimate $\hat{\theta}$ given by (6.27). If the system is given by (6.11) and the requirements 1)-4) and consistency conditions (6.18) and (6.19) are fulfilled, then $\hat{\theta}$ is ,under mild conditions asymptotically Gaussian distributed:

$$\sqrt{N}.(\hat{\theta}-\theta_t) \to N(0,P_{iv}) \tag{6.28}$$

with the covariance matrix P_{iv} given by

$$P_{iv} = (R^T\Gamma R)^{-1}R^T\Gamma.E\{[\sum_{i=0}^{\infty} Z(k+i)K_i].\Lambda.[\sum_{j=0}^{\infty} K_j^T Z^T(k+j)]\}.\Gamma R(R^T\Gamma R)^{-1}) \tag{6.29}$$

In this expression, R is given by

$$R = E\{Z(k)F(z^{-1})\Omega^T(k)\} \tag{6.30}$$

and the sequence $\{K_i\}_{i=0}^{\infty}$ is defined by

$$\sum_{i=0}^{\infty} K_i z^i = F(z)H(z) \tag{6.31}$$

The lower bound for the covariance matrix can be attained if the IV-matrix $Z(k)$, the prefilter $F(z^{-1})$ (cf.(6.25)),and the weighting

matrix Γ (cf (6.27)) are appropriately chosen.
The matrix $H(z^{-1})$ appears in the modelling of the residuals. (cf. (6.7))
This lower bound is given by:

$$P_{iv}^{opt} = \{E[H(z^{-1})^{-1}\Omega^T(k)]^T \Lambda^{-1} [H(z^{-1})^{-1}\Omega^T(k)]\}^{-1} \qquad (6.32)$$

Stoica and Söderström give two solutions for which the optimal parameter covariance is attained.
The first solution is as follows:

$$Z_1(k) = [\Lambda^{-1}H(z^{-1})^{-1}\Omega^T(k)]^T$$
$$F_1(z^{-1}) = H(z^{-1})^{-1}, \quad \Gamma_1 = I \qquad (6.33)$$

and the parameter vector estimator becomes:

$$\Theta_1 = [\sum_{k=1}^{N} Z_1(k)H(z^{-1})^{-1}\tilde{\Omega}^T(k)]^{-1} \cdot [\sum_{k=1}^{N} Z_1(k)H(z^{-1})^{-1}y(k)]$$

A second solution is:

$$Z_2(k) = [H(z)^{-T}\Lambda^{-1}H(z^{-1})^{-1}\Omega^T]^T$$
$$F_2(z^{-1}) = I \quad , \quad \Gamma_2 = I \qquad (6.34)$$

and the second parameter vector estimator becomes:

$$\Theta_2 = [\sum_{k=1}^{N} Z_2(k)\tilde{\Omega}^T(k)]^{-1} \cdot [\sum_{k=1}^{N} Z_2(k)y(k)]$$

As can be seen, the important point in these solutions is, that both require the knowledge of the noisefree system output x(k) as well as of the noise autocorrelation Λ, and $H(z^{-1})$.
This has been noted in other instances in the literature, e.g. by Jakeman and Young (1979) who derive an optimal IV, based on the approximate maximum likelihood method. In the practical situation we know nothing about the system and the noise when we start an estimation procedure. This means that a theoretically optimal IV matrix cannot be constructed if the system is not known exactly.
The best we can do to approach this, is to generate an approximation of x(k) by an auxiliary model, as soon as we have our first parameter vector estimate $\hat{\Theta}$. The parameters of this model are the parameters that were estimated previously. These auxiliary model

outputs are then used in the IV matrix.
Now we can calculate new estimates (and hopefully better ones) and repeat this process until the parameter values have sufficiently converged. Thus we have an iterative estimation process.

At the start of this procedure the auxiliary model does not yet exist. Therefore we have to devise a method to calculate our first estimate of the parameter vector.

Of course, the most straightforward method is to use a least squares estimator, which will probably give a result, that is good enough to start the iteration process, as long as the noise power is small enough.
A different solution is to use a different Z(k) matrix. In stead of the outputs x(k) we can use the delayed inputs u(k). Care has to be taken that the matrix Z(k) does not contain identical elements.
In our simulations this first stage of every estimation has been a least squares estimation , which is simple and adequate as long as there is not too much noise involved.

Our next problem is how to fill the IV vector with instruments.
In our simulations the parameter vector θ_i determines the size of the IV vector $Z_i(k)$. Moreover, the parameter vector θ_i consists of two parts: one part contains the parameters of the polynomials $p_{ij}(z)$, and the second part contains the parameters of the polynomials $q_{ij}(z)$. The IV vector $Z_i(k)$ is built up similarly.

$$Z_i(k) = \begin{vmatrix} -\zeta_1(k+\nu_{i1}-k_i) \\ -\zeta_1(k+\nu_{i1}-1-k_i) \\ \cdot \\ -\zeta_1(k-k_i) \\ -\zeta_2(k+\nu_{i2}-k_i) \\ \cdot \\ -\zeta_2(k-k_i) \\ \cdot \\ \cdot \\ -\zeta_q(k+\nu_{iq}-k_i) \\ \cdot \\ -\zeta_q(k-k_i) \\ u_1(k-1) \\ \cdot \\ u_1(k-k_i) \\ u_2(k-1) \\ \cdot \\ \cdot \\ u_p(k-1) \\ \cdot \\ u_p(k-k_i) \end{vmatrix} \qquad (6.35)$$

In (6.36) $\zeta_i^{\cdot}(k)$ is the i-th output of the auxiliary model.
Note the similarity between $Z_i(k)$ and $\Omega_i(k)$.
This way of choosing instruments has much similarity with the IV that has been derived by Jakeman and Young (1979)on the basis of their approximate maximum likelihood method.

For the first iteration in which a least squares estimate $\hat{\Theta}_{i,ls}$ is determined, we can use the $\Omega_i(k)$ vector in place of $Z_i(k)$ to obtain:

$$1/N \sum_{k=1}^{N} \tilde{\Omega}_i(k)y_j^m(k) = 1/N \sum_{k=1}^{N} \tilde{\Omega}_i(k)\tilde{\Omega}_i^T(k) \cdot \hat{\Theta}_{i,ls} \qquad (6.36)$$

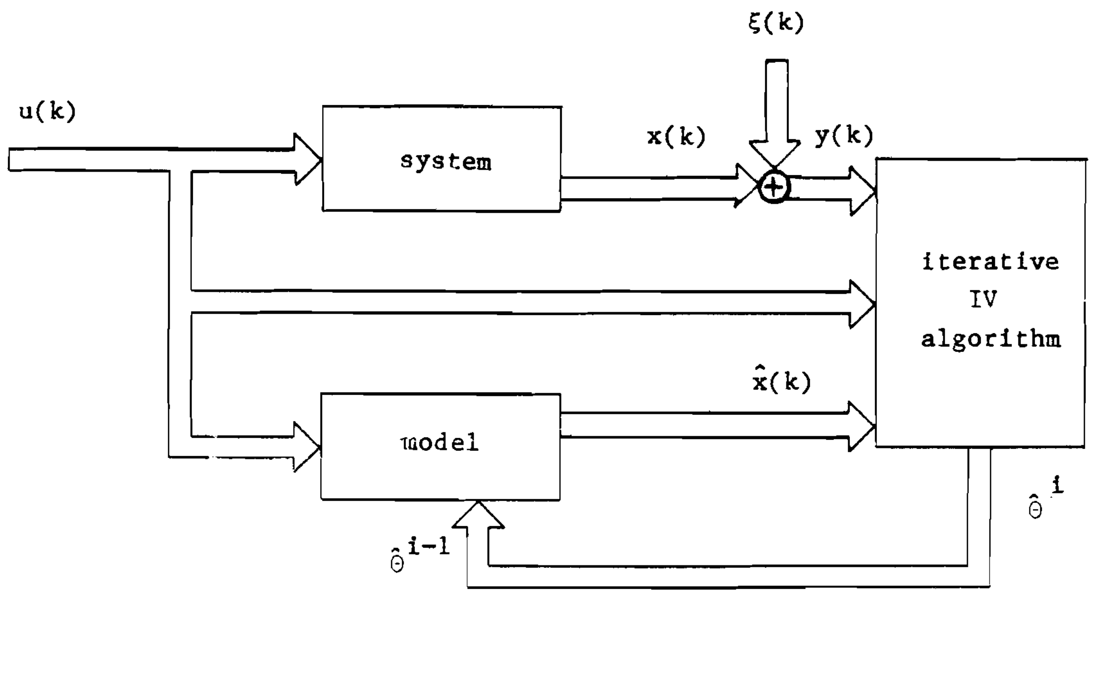
and we find:

$$\hat{\Theta}_{i,ls} = \{ \sum_{k=1}^{N} \tilde{\Omega}_i \tilde{\Omega}_i^T \}^{-1} \{ \sum_{k=1}^{N} \tilde{\Omega}_i(k)y_j^m(k) \} \qquad (6.37)$$

This least squares estimate $\hat{\Theta}_{i,ls}$ is now used for calculating the model outputs $\zeta_i(k)$ that will serve as instruments in the second iteration.

In the last equation the similarity between the least squares and the IV estimator becomes most apparent. From this expression it can also be understood that the instrumental variable technique is a modification of the least squares approach.

Schematically we can depict the estimation process as follows:



Chapter 7 : The experiments

In order to evaluate the usefulness of different MFDs for estimation purposes, some numerical test were performed.
These tests consisted of simulating a MIMO system and trying to estimate the parameters of an MFD-type model from the input- and output signals.
The advantage of simulating is that we can be certain that the signals over which we dispose, are derived from a linear system. Therefore we can also be certain that a linear model will be able to represent our system.
In practical situations (e.g. in identification of industrial processes) some nonlinearity can always be expected, and a linear model will only give the input- output behaviour of the system under consideration, to some extent.
Apart from this, real measurements will always involve noise (e.g. measurement noise or environmental noise). Therefore we have added noise to the output signals of our system, in well controlled amounts, and with carefully chosen properties. The effects of various quantities of noise on the estimates will be studied.

A program has been developed to calculate the parameters of a MIMO input- output model with a particular (MFD-) structure, using input- and output sequences from a simulated system.
The estimation was performed using an instrumental variable technique, which has been described already.
A detailed description of this program can be found in a separate report.

Before we make an attempt to interpret the results of the simulations a word of caution must be spoken. Parameter-estimations were performed for only one simulated system, and it is not clear what the results will be if a different system will be used, e.g. when the number of in- and outputs is increased.
Also, use was made of only one noise sequence for every output for each estimation. This means that no general conclusions can be drawn about the statistical properties of the estimators nor about the IV-estimation procedure.

The results of the experiments are presented in appendix 1.
All calculations were performed in double precision, which corresponds to a machine accuracy of 10^{-32}.

7.1 Structural identification

The only thing that is left before starting an esimation process, is to determine the structure of the model that will be used.
With the structure of the model in case of matrix fractions is meant the set of row degrees of the matrix P(z). If this set is known, and the type of model is chosen, the number of parameters, who are to be estimated is determioned.

A variety of order tests have been proposed in the literature.
E.g. Wellstead (197) introduced his product moment matrix, while Stoica and Söderström (1982) suggest a procedure which is related to one of the consistency conditions for the IV method.

For the parametrisation according to Guidorzis form, there exists a structural identification procedure, developed by Guidorzi(1982).
An implementation of this procedure has been made by J. Meertens (1983).
As we have seen, the sets of row degrees of P(z) for all standard irreducible MFDs are identical. This means that in principle Guidorzis structural identification can be used to determine this set. The structure for a particular MFD is than given by a permutation of this set.

For the MFDs which do not have a row proper P(z) matrix we can remark, that the requirement of irreducibility of the MFD gives the value of deg det P(z), which equals the dimension of the associated minimal state space realization.

At this point we are ready to calculate the IV-estimate (6.22), as we have the equation with which to calculate it, we have given a description of the instrumental variables,we have (in principle) a dataset to our disposal, and we assume that the structure of the model is already determined.

It should be noted that if the outputs are numbered in a different way, the structure of the model will generally be different. We have still the same model order though.
The consequence of this, is that the total number of parameters of the matrices P(z) and Q(z) depends on the ordering of the outputs.

7.2 The simulated system

To generate input and output sequences we used the following 3 input,2 output state-space system:

$$\begin{aligned} x(k+1) &= Fx(k) + Gu(k) \\ y(k) &= Hx(k) + \xi(k) \end{aligned} \qquad (7.1)$$

The state vector x(k) consists of 4 components: $x_1(k)$, .., $x_4(k)$. and accordingly our system order equals 4.
The input vector u(k) has 3 components: $u_1(k)$, $u_2(k)$ and $u_3(k)$.
We used exclusively 3 zero mean, Gaussian distributed,white noise signals as inputs, which are uncorrelated. The signal power of the 3 inputs was identical. As remarked before, the choice of white noise is made, because this type of signal has sufficient degrees of freedom to allow calculation of all our unknown parameters.
The output vector y(k) has two components $y_1(k)$ and $y_2(k)$ as well as the noise vector $\xi(k)$: $\xi_1(k)$ and $\xi_2(k)$.
We chose a zero mean, Gaussian distributed white noise signal to disturb the output, because this type of noise is expected to have a less deleterious effect on the estimations as coloured noise has. It was applied in such a way as to have equal signal to noise ratios on both outputs. The 5 noise generators were constructed to give signals that were uncorrelated with the other signals. All signals contain 300 samples, which should be sufficient to estimate all parameters.

The system matrix F is chosen as:

$$F = \begin{vmatrix} 0.7 & 0 & 0 & 0 \\ 0 & 0.6 & 0 & 0 \\ 0 & 0 & 0.2 & 0 \\ 0 & 0 & 0 & 0.1 \end{vmatrix} \qquad (7.2)$$

The diagonal elements of F represent its eigenvalues (and the poles of the system), and we can interpret the values of 0.7 and 0.6 as belonging to medium fast decreasing impuls responses , while the values 0.2 and 0.1 belong to fast decaying impuls responses.
The matrices G and H are chosen in such a way, that these 4 eigenvalues can be observed equally well from the output y(k).
This is important because our system will then behave as a 4th order system that can be modelled with a 4th order model.
We have

$$G = \begin{vmatrix} 1 & 0 & 1 \\ -1 & 0.5 & 0.5 \\ 0 & 1 & -0.5 \\ 0 & 0.5 & -1 \end{vmatrix} \qquad H = \begin{vmatrix} 0.5 & 0 & 1 & 1 \\ 1 & -0.5 & 0.5 & -0.5 \end{vmatrix} \qquad (7.3)$$

7.3 The model types

We estimated the parameters of the following MFD-models with the help of the in- and output signals of our simulated system:

1) The diagonal MFD, which has the P(z) matrix in diagonal form. The degrees of the monic polynomials in row 1 and 2 of P(z) were chosen in the following pairs: (3,3), (3,2), (3,1), (2,2), and (1,3). The choice of these degrees determines all polynomials in the P(z) and Q(z) matrices as far as the number of unknown parameters of the polynomials is concerned.
2) The hermite MFD, which has the P(z) matrix in hermite form. The degrees of the monic polynomials in row 1 and 2 of P(z) were chosen in pairs: (3,2), (3,1), (2,2) and (1,3).
3) The echelon form of P(z) with degrees of monic polynomials

(3,1), (2,2) and (1,3).
These sets of degrees of monic polynomials were tried, to be able to choose a structure which could be regarded as giving an optimal model. This was done because the estimation program did not incorporate a structural identification part.

The number of unknown parameters can be calculated if the degrees of the monic polynomials are known. This can be illustrated for the echelon form with degrees (1,3).
First, the degrees are ordered according to their size, which results in a 3 for row degree 1, and a 1 for row degree 2. This means that all polynomials in the first row of Q(z) have degree 2 (corresponding to 3 unknown parameters), and all polynomials in the second row of Q(z) have degree 0 (corresponding to 1 unknown parameter). The requirements specified in the section on the unique forms determine that the polynomial $p_{12}(z)$ has degree 0, and $p_{21}(z)$ has degree 1.

For the hermite and the echelon form combinations of degrees were taken , whose sum equals 4. Remember that the dimension of our simulated system is 4, and therefore the given combinations of degrees should contain the combination corresponding to an adequate model. (See the section on the relation between MFDs and state-space models.)

For the diagonal form combinations of degrees were tried that have a sum greater than 4, because this parametrisation is not uniquely identifiable. This means that we have to consider model structures (the degrees of the monic polynomials) who are apparently oversized. Looking backwards it would have been interesting to see the results of pairs of degrees : (4,4), (4,3), and (3,4). Unfortunately these combinations were not considered during the experiments.
To understand why these combinations could give a succesful model, remember that the system has 4 poles, and therefore the maximum necessary degree of the polynomials of P(z) is 4.

7.4 Reconstruction errors and correlation functions.

The parameters of the tested models could not be compared directly with the parameters of the simulated system, due to their differing types. Instead, the estimated model parameters in the noisefree case served as a reference.

Parameter- estimations were performed for the following signal to noise ratios :

1) no noise added (only round-off errors due to finite accuracy of the computations)
2) 75 dB
3) 55 dB
4) 35 dB
5) 15 dB
6) -5 dB

The signal to noise ratio for output i was calculated in the following way:

$$S/N\ (i) = \frac{\sum\limits_{k=1}^{N} \{x_i(k)\}^2}{\sum\limits_{k=1}^{N} \{y_i(k) - x_i(k)\}^2} \qquad i=1,2 \qquad (7.4)$$

The cited dB-values are obtained by taking the $\log_{10}$ of the ratio and multiplying it by 10.
In this expression N is the number of available samples, $y_i(k)$ is the noise free system output, and $y_i(k)$ is the noise corrupted system output. The output error $\xi(k)$ is :

$$\xi_i(k) = y_i(k) - x_i(k) \qquad (7.5)$$

After performing the parameter-estimation we have a model with which we can calculate an estimate of the system output : $\hat{x}_i(k)$, using the (noisefree) system input and the model parameters.
This enables us to calculate an estimate for the noise signal.
We simply have :

$$\hat{\xi}_i(k) = y_i(k) - \hat{x}_i(k) \tag{7.6}$$

which is the estimated output error.
We will now define the reconstruction error for output i as the ratio:

$$RE\ (i) = \frac{\sum_{k=1}^{N} \{y_i(k) - \hat{x}_i(k)\}^2}{\sum_{k=1}^{N} \{y_i(k)\}^2} = \frac{\sum_{k=1}^{N} \{\hat{\xi}_i(k)\}^2}{\sum_{k=1}^{N} \{y_i(k)\}^2} \tag{7.7}$$

This reconstruction error can be interpreted as the estimated-noise to signal ratio.
If the estimated parameters give a correct representation of the system we have: $\hat{x}_i(k) = x_i(k)$ and $\hat{\xi}_i(k) = \xi_i(k)$
As can be expected, there is a relation between the signal to noise ratio and the reconstruction error :

$$\frac{1}{RE\ (i)} = \frac{\sum_{k=1}^{N} \{y_i(k)\}^2}{\sum_{k=1}^{N} \{\hat{\xi}_i(k)\}^2} \approx \frac{\sum_{k=1}^{N} \{x_i(k) + \xi_i(k)\}^2}{\sum_{k=1}^{N} \{\xi_i(k)\}^2} . \tag{7.8}$$

The better the noise power is estimated, the more these ratios approach each other.
As the noisefree system output y_i and the output error ξ_i are uncorrelated the expression $1/N. \sum_{k=1}^{N} x_i(k)\xi_i(k)$ approaches zero for N large enough.
In this case we have:

$$\frac{1}{RE\ (i)} \approx \frac{\sum_{k=1}^{N} \{x_i(k)\}^2}{\sum_{k=1}^{N} \{\xi_i(k)\}^2} + 1 = S/N\ (i) + 1 \tag{7.9}$$

Keeping in mind the restrictions involved in this relation, we have a tool to interpret or to evaluate the estimation results. A second method to judge the results is derived from one of the consistency requirements for the instrumental variable method, which says that the residuals and the instruments should be uncorrelated.
We calculated the correlation between the estimated output error $\hat{\xi}_i(k+\tau)$ and the inputs $u_j(k)$ for time shifts τ from -20 to +20 samples. The correlation is calculated as:

$$corr(i,j,\tau) = \frac{\sum_{k=1}^{N'} \hat{\xi}_i(k+\tau)u_j(k)}{N' \cdot \frac{1}{N'} \sum_{k=1}^{N'} \{u_j(k)\}^2} \qquad \begin{array}{l} i=1,2 \\ j=1,2,3 \end{array} \tag{7.10}$$

In this expression N' is the number of samples that is used in the calculation of the sum. The value of N' was kept constant for all values of τ, for each model-structure.
If the estimated output error is not correlated with the inputs, then we can safely assume that the residuals (which are a filtered version of the output error) and the instruments (which are a filtered version of the input signals) are uncorrelated also.
The power of the input signals is taken into account, such that input signals of varying amplitude between cases (but still the same for all inputs) will result in identical values of corr(i,j,τ).
An absolute threshold for the correlation was not found, but we can compare among different models and modelstructures.

Graphs were made of the correlation functions and some of these are presented in an appendix 2.
As a means of characterising a correlation function with a single value, we calculate the norm of the correlation as:

$$\{\text{NORM } (i,j)\}^2 = \sum_{\tau=-20}^{20} \{\text{corr}(i,j,\tau)\}^2 \qquad \begin{matrix} i=1,2 \\ j=1,2,3 \end{matrix} \tag{7.11}$$

It is expected that the norm of the correlation will be small if it is related to a good model of the system. For models, which are not able to represent the system dynamics in an accurate way, we expect greater values of the norm.
The norm of the correlation was represented in single precision, which corresponds to a machine accuracy of 10^{-16}.

7.5 Convergence of the iterative IV-procedure.

In the literature there are no theoretical studies of the convergence properties of iterative IV-estimators . Therefore we cannot predict the behaviour of the estimates in the course of the iterations.
For the noisefree case, however, it is expected that the convergence is almost instantly (expressed in numbers of iterations) if the model is adequate. This is because the first iteration step is a least squares step, and if no noise is present the estimation should give unbiased results.
When noise is added to the system output, the least squares estimates will deviate from the correct values, and it is hoped that the iterative IV-process will result in correct parameter values.
For unadequate models, and for increasing output noise, we will just have to wait and see what happens. E.g. it can occur that the models become instable. The system outputs will cause numerical overflow in the computer and the estimation is abruptly stopped.
To give an indication of the rate of convergence in a particular case, the number of iterations is stated and also the maximum change in any parameter of the model during the last iteration. This maximum change is given as : $(\Delta \text{ par})_{max}$.

In the first estimations the only stop-criterion used, was if the

value of $(\Delta\ par)_{max}$ was less than 10^{-11}.
In the echelon (2,2) case with a S/N ratio of -5 dB this resulted in 57 iterations. Subsequently a maximum of 31 was posed on the number of iterations.
These limits were later reduced to 10^{-6} for $(\Delta\ par)_{max}$ and 25 for the number of iterations.

7.6 Results of the experiments

In this last section the results of the experiments are presented. The graphs and numerical results are found in appendices 1 and 2. The number of experiments was rather limited because the development of the estimation program required most of the available time. This means that we cannot derive any rocksteady conclusions. The results can be used, however, as a source of inspiration for more systematic studies.

7.6.1 Results for the echelon model

Looking at the noisefree case first (table 1.1 - 1.2), we see that the (2,2) structure gives clearly the best model, if we compare the reconstruction errors and the norms of the correlations. Both quantities are essentially equal to zero, remembering the computational accuracy of 10^{-32} for the reconstruction errors and 10^{-16} for the correlations.
This means that the model with structure (2,2) gives a good representation of the simulated system.
For the (1,3) and (3,1) structures the reconstruction errors and the norm of the correlations is significantly different from zero.
We can also observe that in these 2 cases the parameters had to go through several iterations before converging.
For the (2,2) model the parameters were estimated in the presence of noise. (Table 1.3 - 1.4)
The observations that we can make in these situations are that with increasing noise power, the reconstruction errors, the norms of

correlations, and the deviation of parameters from their noisefree reference values increases.
The rate of convergence decreases with increasing noise.
We see that the results for S/N ratios of 75 and 55 dB are very similar. An explanation of this fact is hard to find.
The correlation functions will be treated separately after the hermite and diagonal parametrizations.

Finally, we find that the relation between reconstruction error and signal to noise ratio holds excellently for the -5 dB case:
S/N = 0.3 $\quad$ ($\equiv$ -5dB)
1/(1+0.3) = 0.77 $\approx$ RE for the (2,2) model.
In the 15 dB case the relation holds also more or less.
For higher signal to noise ratios there is less conformity between 1/RE and S/N+1.
In general the RE increases for lower S/N ratios as is predicted by (7.9)

7.6.2 Results of the hermite parametrization

The results for the Hermite parametrization are presented in tables 1.5 - 1.8. .
The reconstruction errors for the noisefree case (table 1.5) indicate that the model with structure (3,1) gives a satisfactory result. As could be expected from section 3.8.1 the set of row degrees of the Hermite model is not equal to the set of row degrees of the best echelon model (which has a structure (2,2)).
The Hermite (2,2) and (1,3) models give reconstruction errors which are significantly different from zero.
The number of iterations is two, and the largest parameterchange during the second iteration is less than 10^{-10}. This indicates that the least squares step gives almost exact parametervalues for this model.
The (3,2) structure leads to dependencies in the linear set of equations from which the parameter vector has to be calculated. This is indicated by "instable model" in table 1.5.
Remembering that the simulated system has order 4, (and that the

degree of the Hermite (3,2) model is 5), this is just what could be expected.
Due to an error in the subroutine, which takes care of the Hermite parametrization, the (2,2) model was (erroneously) regarded as the best model. For this model parameterestimations were performed with various amounts of noise. After the error was detected, it appeared that only the results for the models with structure (3,2) and (3,1) were not correct. After the correction of the program, these two models were tested in the noise free case, and the results are presented in this report. There was no time left to subject the , now best, model with structure (3,1) to parameterestimations in the presence of noise.
The same general observations can be made as in the echelon-case: with increasing noise power, the reconstruction error rises along with the norm of the correlations, while the rate of convergence decreases and the parameters move away from the noisefree values.

7.6.3 Results for the diagonal model

As was mentioned earlier, the diagonal parametrization is not a uniquely identifiable form, that is also not equivalent under premultiplication by a unimodular matrix.
The ease with which it is parametrized made it impossible, however, to resist it in the experiments.
Examining the RE in the noisefree case, we see that the structure (3,3) gives the best results. For smaller structures (e.g. (3,1)) the RE is larger for both outputs.
Note that the RE for output 1 in the (3,3) and (3,1) cases and the RE for output 2 in the (3,2) and (2,2) cases are slightly different. These values should pair wise be equal, because parameter vectors belonging to output 1 and 2 are estimated independently. The difference is probably due to a change in the program with which the system was simulated.
The convergence rates are quite similar for all tested model structures.
Of course, the parameter estimation, RE and correlations behave in the same way as in the echelon and Hermite case, if output noise is added.

7.6.4 General observations and correlation functions

Comparing the reconstruction error for output 1 and 2 we see that (except in the echelon (1,3), hermite (1,3) and diagonal (1,3) cases) the first output has a (sometimes much) better value. There is no obvious reason for this tendency. We have to remember that all result are based on only 1 noise sequence and 1 input sequence, so it might s just a matter of chance. be

Examinination of the correlation functions reveals that their visual pattern follows a strange evolution when more output noise is added. Although the input signal is a random signal, and the output noise is also, the correlation function can have a remarkably "non-random" behaviour. This is the case for noisefree outputs and for S/N ratios 75 dB and 55 dB. Suddenly the situation changes when the S/N ratio of becomes 35 dB and less. In these graphs we see the expected random behaviour of the correlations.

An explanation of this effect might be that for low S/N ratios the estimate $\hat{\xi}(k)$ of the output noise consists for the greater part of the real output noise which is uncorrelated with the input signals For high S/N ratios we believe that the estimated output error $\hat{\xi}(k)$ is caused largely by the difference of the parameters from their "real" values.

Therefore the estimated output error is more or less correlated with the input signal for high S/N ratios.

A point which is difficult to understand, is why the correlation function undergoes a very marked change at a S/N ratio of 35 dB.

An overall examination reveals that the correlation functions of all tested models show more or less the same behaviour, which gives ample opportunity for further research.

Chapter 8 : Final remarks

As it is mostly the case at the end of a study, the work has raised at least as many questions as it has answered.
Some of these questions will be given below.

- Although right MFDs and left MFDs are connected as described in chapter 3, the description of unique forms can not be found in a straightforward way (by transposition of P(z) and Q(z) matrices) in every case. The connection between the descriptions of unique forms for both types of MFDs (some authors describe left MFDs, others right MFDs) remain to be established for all unique forms.

- There seems to exist a large body of literature concerning MFDs, outside the more renowned magazines on measurement and control theory. These articles, however, have mysteriously escaped my view.

- Guidorzi has developed a unique MFD, and also a structural identification procedure for this MFD. Although in case of standard irredurcible MFDs, the sets of row degrees are related some forms (e.g. the hermite form) cannot make a direct use of this structural identification. For the hermite form (and others: diagonal and fully parametrized form) the development of order tests would be useful.

- In this report use was made of a basic IV-estimation procedure. This procedure can be refined e.g. by prefiltering of data or the use of an IV-vector with more entries than the number of parameters to be estimated with it.
  Also, our first estimation step was a least squares estimation. This first step can be executed as in IV-form. In this case, the instruments have to be chosen carefully , as the absence of parameter-estimates precludes the use of a model to generate instruments.

- As a way to evaluate our estimates, we calculated the correlation between estimated output error and input signals. The consistency requirement for the IV-method is given in terms of the correlation between residuals and instruments. This calculation can probably be made without too great difficulty, and if it is done the interpretation of the results is more direct.

-The form of the calculated correlations between estimated output error and input signals has some quite unexpected properties for high S/N ratios. A possible explanation is that the observed peak is due to a difference between the estimated model parameters and their true values.

-The simulations started with a state-space model to generate input- and output-signals. The estimated parameters of the MFD model could therefore not be compared directly with the coefficients appearing in the original system. Therefore it could be instructive to generate signals with a MFD system and then perform parameter estimations .

-From the theory it is expected that the parameter estimation gives equivalent MFDs (within the limit of computational accuracy).
The hermite and echelon MFD should be able to represent our 4th order system with a MFD of degree 4. Only the echelon (2,2) succeeds in this. Why do the hermite (3,1), (2,2) and (1,3) forms give such unsatisfactory results?

-Even if the number of constant digits after some iteration step is only 6, the reconstruction error for 1 or both outputs can be as small as 10^{-28}. Does this mean that, even if the parameter values have not converged to their ultimate ability, the resulting model gives an accurate representation of the investigated system?
In other words: what is the use of yet another iteration step?

-The results of the parameter estimations for the S/N ratios of 75 and 55 dB are very similar. What can be the reason of this?

-If the number of outputs and inputs of the simulated system is increased the peculiarities of the various types of MFDs might be made more clear.

-The estimation program has been developed to do all calculations in double precision. What will happen if all the calculations are carried out in single precision? Will the calculations be numerically stable?

-During the final discussions concerning this work it appeared that the degrees of the polynomials $q(z)$ were too low in the case of a MFD

Conclusions

The purpose of this study was to give a summary of some of the literature on matrix fractions, and to develop a computer program which performs parameters-estimations for MIMO MFD models.
From this study the following conclusions can be drawn:

-Algorithms have been developed to parametrize some MFDs if the model structure is given; the estimations performed indicate that this type of model might be useful.

-An instrumental variable estimation procedure has been developed which appears to be able to estimate modelparameters even in extremely noisy cases (S/N ratio of -5 dB).

-The correlations between estimated output error and input signals (which is related to a consistency condition for the IV-method) shows a quite unexpected behaviour for high S/N ratios.

-The diagonal MFD that results from an estimation procedure is likely to have a higher degree than e.g. the resulting echelon MFD.

-Some system outputs are reconstructed very well by the model whereas other outputs are reconstructed very badly even in the noisefree case. This effect is caused by the independent estimation of parameters for every equation of the set of input-output equations.

List of symbols

a	coefficient in Guidorzi's form ($a_{ij,k}$)
arg	argument of a function
block diag	block diagonal matrix
corr(i,j,τ)	correlation of estimated noise on output i with input j for a time delay of τ samples
diag	diagonal matrix
deg	degree of a polynomial (deg p(z))
det	determinant of a matrix (det A)
Δ_i	ixi minor of a matrix
ε	element of ($x \varepsilon X$)
e	stochastic variable (e(k))
E	expectation operator ($E\{x\}$)
	equivalence indicator (X E Y)
f	function ($f: x \rightarrow f(x)$)
F	system matrix ($\{F,G,H\}$)
$F(z^{-1})$	filter (transfer matrix)
G	input distribution matrix ($\{F,G,H\}$)
Γ	weighting matrix
H	output distribution matrix ($\{F,G,H\}$)
$H(z^{-1})$	filter (transfer matrix)
$\eta(k)$	residual vector
$\Theta^{(i)}$	parameter vector resulting from i iterations.
Θ	parameter vector
Θ_i	parameter vector belonging to one equation of $P(z)y(k)=Q(z)u(k)$
Θ_t	true parameter vector
I	identity matrix
i	general index
j	general index
k	sample number
k_i	row degree
l	general index
$\Lambda(z)$	Smith form of a polynomial matrix
M	matrix of coefficients (Guidorzi)
μ_i	invariant polynomial (Smith form)
Norm(i,j)	norm of correlation corr(i,j,τ), τ= -20,..,+20

N	number of samples
N'	number of samples used in a calculation
n	system order, dimension
ν_{ij}	degree of polynomial $p_{ij}(z)$
$\Omega(k)$	matrix of delayed input and output vectors
$\tilde{\Omega}(k)$	matrix of delayed input and noisy output vectors
$[0]$	zero matrix
p	number of inputs
$P(z)$	MIMO autoregressive polynomial matrix
	appearing in matrix fraction $(P(z),Q(z))$
$P^*(z^{-1})$	matrix derived from $P(z)$
$\pi(z)$	polynomial in Smith-McMillan form $(\pi_i(z))$
$\Pi(z)$	diagonal polynomial matrix appearing in Smith-McMillan form
P_{iv}	covariance matrix of parameter estimates for IV method
$[P(z)]_h$	high order coefficient matrix
$P_i(z)$	i-th row of matrix $P(z)$
$P^c(z)$	matrix defined with $[P(z)]_h$
$p_{ij}(z)$	polynomial element of $P(z)$
$q_{ij}(z)$	polynomial element of $Q(z)$
$Q(z)$	MIMO moving average polynomial matrix
	appearing in matrix fraction $(P(z),Q(z))$
$Q^*(z^{-1})$	matrix derived from $Q(z)$
q	number of outputs
s_i	pivot index (e.g. in echelon MFD)
T	permutation of identity matrix
$T(z)$	transfer function
	transpose of a matrix
τ	time delay (number of samples)
$U(z)$	unimodular polynomial matrix
$u(k)$	input vector
$V(z)$	unimodular polynomial matrix
$V_d(z)$	vector space spanned by the highest order coefficient rowvectors of the rows of $P(z)$, having degrees less than or equal to d
$x(k)$	noisefree system output vector
	state vector
$x^m(k)$	permutated noisefree system output vector
X	set of elements

$x^m(k)$	estimated system output vector
$X(z)$	general polynomial matrix
$Y(z)$	general polynomial matrix
$y(k)$	noise-corrupted system output vector
$y^m(k)$	permutated noise-corrupted system output vector
$\hat{y}(k)$	estimated system output vector
z	time advance operator
z^{-1}	time delay operator
$Z(k)$	matrix of instrumental variables
$Z_i(k)$	vector containing instrumental variables
$\Psi(z)$	blockdiagonal polynomial matrix
$\xi(k)$	output error vector
$\hat{\xi}(k)$	estimated output error vector
$\{x(k)\}$	signal sequence

Echelon parametrisation (noisefree case)

	structure (3,1)	(2,2)	(1,3)
Reconstruction Error			
output 1	.356 E -26	.422 E -29	.568 E -2
output 2	.523 E - 2	.118 E -28	.319 E -5
Norm of correlation			
Norm (1,1)	.1826 E -14	.7449 E -16	.4266 E -2
Norm (1,2)	.4816 E -14	.1318 E -15	.5440 E -2
Norm (1,3)	.6859 E -14	.2504 E -15	.7143 E -2
Norm (2,1)	.3352 E - 2	.1489 E -14	.1036 E -3
Norm (2,2)	.8050 E - 2	.3067 E -15	.1600 E -3
Norm (2,3)	.1106 E - 1	.5628 E -15	.2477 E -3
Number of iterations	6	2	4
$(\Delta \text{ par})_{max}$	.233 E - 6	.929 E -13	.603 E -6

Table 1.1

Parametervalues of the echelon parametrisation (noisefree case).

	Structure (3,1)	(2,2)	(1,3)
$p_{11,1}$	-.1400 E -1	.9925 E -1	-.5455 E -2
$p_{11,2}$	.2300 E 0	-.8094 E 0	-
$p_{11,3}$	-.1000 E 1	-	-
$p_{12,1}$	.5630 E -13	.6792 E -1	-.5455 E -1
$p_{12,2}$	-	-.1132 E 0	.5655 E 0
$p_{12,3}$	-	-	-.1391 E 1
$q_{11,1}$	.1000 E -1	-.2245 E 0	.1273 E 0
$q_{11,2}$	-.1500 E 0	.5000 E 0	-.1086 E 1
$q_{11,3}$	.5000 E 0	-	.1500 E 1
$q_{12,1}$	.1400 E 0	-.9642 E 0	.3182 E -1
$q_{12,2}$	-.1250 E 1	.1500 E 1	-.7500 E -1
$q_{12,3}$	.1500 E 1	-	-.4608 E -14
$q_{13,1}$	-.1650 E 0	.8462 E 0	.1955 E 0
$q_{13,2}$	.1150 E 1	-.1000 E 1	-.8409 E 0
$q_{13,3}$	-.1000 E 1	-	.1000 E 1
$p_{21,1}$	-.5585 E -1	.4340 E -1	-.1282 E 0
$p_{21,2}$	.3415 E 0	-.2075 E 0	-
$p_{21,3}$	-	-	-
$p_{22,1}$	-.7335 E 0	.1143 E 0	-.2323 E 0
$p_{22,2}$	-	-.7906 E 0	-
$q_{21,1}$	.1673 E 1	-.2896 E 0	.5129 E 0
$q_{21,2}$	-	.1500 E 1	-
$q_{22,1}$	.5123 E 0	-.3863 E 0	.1492 E 1
$q_{22,2}$	-	-.2164 E -15	-
$q_{23,1}$	.6476 E 0	-.3302 E -1	-.9914 E 0
$q_{23,2}$	-	.1000 E 1	-

Table 1.2

Echelon parametrisation
Structure (2,2)

	S/N ratio noisefree	75 dB	55 dB	35 dB	15 dB	-5 dB
Reconstruction Error						
output 1	.422 E -29	.587 E -4	.602 E -4	.333 E -3	.280 E -3	.731 E 0
output 2	.118 E -28	.527 E -3	.525 E -3	.790 E -3	.307 E -1	.766 E 0
Norm of correlation						
Norm (1,1)	.7449 E -16	.2529 E -3	.2694 E -3	.7950 E -3	.7520 E -2	.7993 E -1
Norm (1,2)	.1318 E -15	.2537 E -3	.2716 E -3	.7576 E -3	.7068 E -2	.7590 E -1
Norm (1,3)	.2504 E -15	.2943 E -3	.3028 E -3	.7920 E -3	.7475 E -2	.7172 E -1
Norm (2,1)	.1489 E -14	.1116 E -3	.1626 E -3	.1084 E -3	.1034 E -1	.1045 E 0
Norm (2,2)	.3067 E -15	.3675 E -3	.3998 E -3	.1455 E -3	.1375 E -1	.1360 E 0
Norm (2,3)	.5628 E -15	.4052 E -3	.3995 E -3	.1229 E -3	.1280 E -1	.1148 E 0
Number of iterations	2	4	9	11	22	57
$(\Delta\ \mathrm{par})_{max}$	.929 E -13	-.438 E -7	-.197 E -12	-.163 E -11	-.476 E -11	-.941 E -11

Table 1.3

Parametervalues of the echelon parametrisation
Structure (2,2)

	S/N ratio noisefree	75 dB	55 dB	35 dB	15 dB	-5 dB
$p_{11,1}$	.9925 E -1	.9967 E -1	.9979 E -1	.1009 E 0	.1027 E 0	.3300 E 0
$p_{11,2}$	-.8094 E 0	-.8109 E 0	-.8103 E 0	-.8038 E 0	-.6884 E 0	-.8083 E 0
$p_{12,1}$	.6792 E -1	.7081 E -1	.6998 E -1	.6066 E -1	-.1475 E 0	-.3700 E 0
$p_{12,2}$	-.1132 E 0	-.1172 E 0	-.1162 E 0	-.1049 E 0	.1633 E 0	.4303 E 0
$q_{11,1}$	-.2245 E 0	-.2309 E 0	-.2292 E 0	-.2104 E 0	.2338 E 0	.3991 E 0
$q_{11,2}$	.5000 E 0	.4970 E 0	.4970 E 0	.4965 E 0	.4857 E 0	.2223 E 0
$q_{12,1}$	-.9642 E 0	-.9660 E 0	-.9648 E 0	-.9521 E 0	-.7530 E 0	-.7005 E 0
$q_{12,2}$	.1500 E 1	.1498 E 1	.1498 E 1	.1499 E 1	.1535 E 1	.2092 E 1
$q_{13,1}$	.8462 E 0	.8439 E 0	.8443 E 0	.8481 E 0	.9779 E 0	.1011 E 1
$q_{13,2}$	-.1000 E 1	-.9979 E 0	-.9979 E 0	-.9972 E 0	-.9878 E 0	-.8588 E 0
$p_{21,1}$	.4340 E -1	.4344 E -1	.4365 E -1	.4543 E -1	.2662 E -1	.2401 E 0
$p_{21,2}$	-.2075 E 0	-.2077 E 0	-.2080 E 0	-.2096 E 0	-.2157 E -2	-.2707 E 0
$p_{22,1}$	.1143 E 0	.1145 E 0	.1144 E 0	.1117 E 0	-.2667 E 0	-.6515 E 0
$p_{12,2}$	-.7906 E 0	-.7908 E 0	-.7905 E 0	-.7857 E 0	-.2842 E 0	.2908 E 0
$q_{21,1}$	-.2896 E 0	-.2899 E 0	-.2894 E 0	-.2791 E 0	.6095 E 0	.1597 E 1
$q_{21,2}$	.1500 E 1	.1500 E 1	.1500 E 1	.1498 E 1	.1477 E 1	.1036 E 1
$q_{22,1}$	-.3863 E 0	-.3865 E 0	-.3869 E 0	-.3878 E 0	-.7052 E -1	-.4368 E 0
$q_{22,2}$	.2164 E -15	-.1429 E -3	-.3150 E -3	-.1964 E -2	.3774 E -1	.8485 E 0
$q_{23,1}$	-.3302 E -1	-.3308 E -1	-.3240 E -1	-.2533 E -1	.2503 E 0	.6689 E 0
$q_{23,2}$	.1000 E 1	.1000 E 1	.9997 E 0	.9960 E 0	.9648 E 0	.7784 E 0

Table 1.4

Hermite parametrisation (noisefree case)

	structure			
	(3,2)	(3,1)	(2,2)	(1,3)
Reconstruction Error				
output 1	instable model	.192 E -26	.212 E - 3	.118 E 0
output 2		.281 E -24	.211 E - 4	.704 E - 4
Norm of correlation				
Norm (1,1)		.4496 E -13	.9307 E - 3	.2896 E - 1
Norm (1,2)		.6282 E -13	.1332 E - 2	.1434 E - 1
Norm (1,3)		.4168 E -13	.9285 E - 3	.3426 E - 1
Norm (2,1)		.5340 E -12	.4059 E - 3	.9657 E - 3
Norm (2,2)		.1021 E -11	.5177 E - 3	.4019 E - 3
Norm (2,3)		.1003 E -11	.3087 E - 3	.1157 E - 2
Number of iterations		2	4	4
$(\Delta \text{ par})_{max}$		-.358 E -11	.577 E - 7	.373 E - 6

Table 1.5

Parametervalues of the Hermite parametrisation. (Noisefree case)

	Structure			
	(3,2)	(3,1)	(2,2)	(1,3)
$p_{11,1}$	instable model	-.1400 E -1	.1010 E 0	-.1467 E 0
$p_{11,2}$		.2300 E 0	-.8337 E 0	-
$p_{11,3}$		-.1000 E 1	-	-
$q_{11,1}$		.1000 E -1	-.6784 E -1	.5647 E 0
$q_{11,2}$		-.1500 E 0	.5000 E 0	-
$q_{11,3}$		.5000 E 0	-	-
$q_{12,1}$		.1400 E 0	-.1001 E 1	.1531 E 0
$q_{12,2}$		-.1250 E 0	.1503 E 1	-
$q_{12,3}$		.1500 E 1	-	-
$q_{13,1}$		-.1650 E 0	.9826 E 0	-.9909 E 0
$q_{13,2}$		.1150 E 1	-.1001 E 1	-
$q_{13,3}$		-.1000 E 1	-	-
$p_{21,1}$		-.[illegible] E 0	.4340 E -1	-.5455 E -2
$p_{21,2}$		.7150 E 1	-.2075 E 0	-
$p_{21,3}$		-.8833 E 1	-	-
$p_{22,1}$		-.6000 E 0	.1143 E 0	-.5455 E -1
$p_{22,2}$		-	-.7906 E 0	.5655 E 0
$p_{22,3}$		-	-	-.1391 E 0
$q_{21,1}$		.1983 E 1	-.2896 E 0	.1273 E 0
$q_{21,2}$		-.4417 E 1	.1500 E 1	-.1086 E 1
$q_{21,3}$		.2057 E -13	-	.1500 E 1
$q_{22,1}$		.8517 E 1	-.3863 E 0	.3182 E -1
$q_{22,2}$		-.1325 E 2	-.6485 E -15	-.7500 E -1
$q_{22,3}$		-.6565 E -13	-	-.3409 E -14
$q_{23,1}$		-.7475 E 1	-.3302 E -1	.1955 E 0
$q_{23,2}$		.8833 E 1	.1000 E 1	-.8409 E 0
$q_{23,3}$		-.3827 E -13	-	.1000 E 1

Table 1.6

Hermite parametrisation
Structure (2,2)

	S/N ratio noisefree	75 dB	55 dB	35 dB	15 dB	-5 dB
Reconstruction Error						
output 1	.212 E -3	.278 E -3	.281 E -3	.566 E -3	.283 E -1	.723 E 0
output 2	.211 E -4	.547 E -3	.546 E -3	.808 E -3	.303 E -1	.772 E 0
Norm of correlation						
Norm (1,1)	.9307 E -3	.1103 E -2	.1106 E -2	.1335 E -2	.7625 E -2	.7772 E -1
Norm (1,2)	.1332 E -2	.1304 E -2	.1301 E -2	.1437 E -2	.7067 E -2	.6945 E -1
Norm (1,3)	.9285 E -3	.1064 E -2	.1070 E -2	.1304 E -2	.7091 E -2	.7204 E -1
Norm (2,1)	.4059 E -3	.4992 E -3	.5239 E -3	.1240 E -2	.1043 E -1	.1053 E 0
Norm (2,2)	.5177 E -3	.4962 E -3	.5108 E -3	.1446 E -2	.1371 E -1	.1355 E 0
Norm (2,3)	.3087 E -3	.5543 E -3	.5554 E -3	.1266 E -2	.1218 E -1	.1174 E 0
Number of iterations	4	5	10	11	26	31
$(\Delta\ par)_{max}$	.577 E -7	-.299 E -7	-.480 E -13	-.138 E -11	-.435 E -11	-.783 E -3

Table 1.7

Parametervalues of the Hermite parametrisation
Structure (2,2)

	S/N ratio noisefree	75 dB	55 dB	35 dB	15 dB	-5 dB
$p_{11,1}$	.1010 E 0	.1013 E 0	.1015 E 0	.1039 E 0	.1271 E 0	.2827 E 0
$p_{11,2}$	-.8337 E 0	-.8340 E 0	-.8342 E 0	-.8360 E 0	-.8531 E 0	-.7822 E 0
$q_{11,1}$	-.6784 E -1	-.6761 E -1	-.6784 E -1	-.7016 E -1	-.9278 E -1	-.1697 E 0
$q_{11,2}$	.5000 E 0	.4962 E 0	.4962 E 0	.4954 E 0	.4872 E 0	.3641 E 0
$q_{12,1}$	-.1001 E 1	-.1001 E 1	-.1001 E 1	-.1001 E 1	-.9961 E 0	-.5683 E 0
$q_{12,2}$	.1503 E 1	.1500 E 1	.1500 E 1	.1502 E 1	.1515 E 1	.1629 E 1
$q_{13,1}$	.9826 E 0	.9833 E 0	.9833 E 0	.9838 E 0	.9875 E 0	.7988 E 0
$q_{13,2}$	-.1001 E 1	-.9988 E 0	-.9987 E 0	-.9978 E 0	-.9886 E 0	-.9085 E 0
$p_{21,1}$	.4340 E -1	.4344 E -1	.4361 E -1	.4508 E -1	.3862 E -1	.1587 E 0
$p_{21,2}$	-.2075 E 0	-.2076 E 0	-.2077 E 0	-.2072 E 0	-.8628 E -1	.3435 E -1
$p_{22,1}$	.1143 E 0	.1145 E 0	.1144 E 0	.1113 E 0	-.2221 E 0	-.5988 E 0
$p_{22,2}$	-.7906 E 0	-.7908 E 0	-.7905 E 0	-.7859 E 0	-.3235 E 0	.1818 E 0
$q_{21,1}$	-.2896 E 0	-.2899 E 0	-.2892 E 0	-.2781 E 0	.5077 E 0	.1620 E 0
$q_{21,2}$	.1500 E 1	.1500 E 1	.1500 E 1	.1498 E 1	.1476 E 1	.1109 E 1
$q_{22,1}$	-.3863 E 0	-.3864 E 0	-.3864 E 0	-.3841 E 0	-.1957 E 0	.6322 E -1
$q_{22,2}$	-.6485 E -15	-.1431 E -3	-.3167 E -3	-.1981 E -2	.3274 E -1	.7549 E 0
$q_{23,1}$	-.3302 E -1	-.3311 E -1	-.3272 E -1	-.2793 E -1	.2969 E 0	.3189 E 0
$q_{23,2}$	.1000 E 1	.1000 E 1	.9997 E 0	.9961 E 0	.9649 E 0	.7423 E 0

Table 1.8

Diagonal parametrisation (noisefree case)

	structure (3,3)	(3,2)	(3,1)	(2,2)	(1,3)
Reconstruction Error					
output 1	.233 E -26	.233 E -26	.169 E -27	.212 E - 3	.118 E 0
output 2	.774 E - 4	.112 E - 2	.436 E - 2	.827 E - 3	.774 E - 4
Norm of correlation					
Norm (1,1)	.3124 E -14	.3124 E -14	.8463 E -15	.9307 E - 3	.2896 E - 1
Norm (1,2)	.4359 E -14	.4359 E -14	.1178 E -14	.1332 E - 2	.1434 E - 1
Norm (1,3)	.2890 E -15	.2890 E -15	.7834 E -15	.9285 E - 3	.3426 E - 1
Norm (2,1)	.3670 E - 3	.1922 E - 2	.4823 E - 2	.1864 E - 2	.3670 E - 3
Norm (2,2)	.8649 E - 3	.3315 E - 2	.7556 E - 2	.3232 E - 2	.8649 E - 3
Norm (2,3)	.1143 E - 2	.4316 E - 2	.8165 E - 2	.4253 E - 2	.1143 E - 2
Number of iterations	7	7	4	7	7
$(\Delta\ \mathrm{par})_{max}$	.18 E - 6	-.49 E - 6	-.27 E - 6	-.41 E - 6	.18 E - 6

Table 1.9

Parametervalues of the diagonal parametrisation. (Noisefree case)

	Structure (3,3)	(3,2)	(3,1)	(2,2)	(1,3)
$p_{11,1}$	-.1400 E -1	-.1400 E -1	-.1400 E -1	.1010 E 0	-.1467 E 0
$p_{11,2}$	.2300 E 0	.2300 E 0	.2300 E 0	-.8337 E 0	-
$p_{11,3}$	-.1000 E 1	-.1000 E 1	-.1000 E 1	-	-
$q_{11,1}$	.1000 E 1	.1000 E 1	.1000 E 1	-.6783 E -1	.5647 E 0
$q_{11,2}$	-.1500 E 0	-.1500 E 0	-.1500 E 0	.5000 E 0	-
$q_{11,3}$	.5000 E 0	.5000 E 0	.5000 E 0	-	-
$q_{12,1}$	.1400 E 0	.1400 E 0	.1400 E 0	-.1001 E 1	.1531 E 1
$q_{12,2}$	-.1250 E 1	-.1250 E 1	-.1250 E 1	.1503 E 1	-
$q_{12,3}$	.1500 E 1	.1500 E 1	.1500 E 1	-	-
$q_{13,1}$	-.1650 E 0	-.1650 E 0	-.1650 E 0	.9826 E 0	-.9909 E 0
$q_{13,2}$	.1150 E 1	.1150 E 1	.1150 E 1	-.1001 E 1	-
$q_{13,3}$	-.1000 E 1	-.1000 E 1	-.1000 E 1	-	-
$p_{21,1}$	-.4881 E -1	-.2308 E -1	-.6593 E 0	-.2246 E -1	-.4881 E -1
$p_{21,2}$	.4082 E 0	-.6494 E 0	-	-.6504 E 0	.4082 E 0
$p_{21,3}$	-.1177 E 1	-	-	-	-.1177 E 1
$q_{21,1}$	.1061 E 0	.2832 E -1	.1493 E 1	.2686 E -1	.1061 E 0
$q_{21,2}$	-.7645 E 0	.1505 E 1	-	.1505 E 1	-.7645 E 0
$q_{21,3}$	.1500 E 1	-	-	-	.1500 E 1
$q_{22,1}$	.1538 E -1	-.7840 E -1	-.6814 E -2	-.7839 E -1	.1538 E -1
$q_{22,2}$	-.7488 E -1	.8569 E -4	-	.4192 E -4	-.7488 E -1
$q_{22,3}$	.9332 E -4	-	-	-	.9332 E -4
$q_{23,1}$	.1562 E 0	-.9735 E -1	.9914 E 0	-.9837 E -1	.1562 E 0
$q_{23,2}$	-.6262 E 0	.9982 E 0	-	.9981 E 0	-.6262 E 0
$q_{23,3}$	.1000 E 1	-	-	-	.1000 E 1

Table 1.10

Diagonal parametrisation
Structure (3,3)

	S/N ratio noisefree	75 dB	55 dB	35 dB	15 dB
Reconstruction Error					
output 1	.233 E -26	.503 E -4	.511 E -4	.327 E -3	instable
output 2	.774 E - 4	.328 E -3	.338 E -3	.779 E -3	model
Norm of correlation					
Norm (1,1)	.3124 E -14	.1186 E -3	.1351 E -3	.8409 E -3	
Norm (1,2)	.4359 E -14	.1168 E -3	.1308 E -3	.6694 E -3	
Norm (1,3)	.2890 E -15	.6726 E -4	.1003 E -3	.6930 E -3	
Norm (2,1)	.3670 E - 3	.5146 E -3	.5497 E -3	.1465 E -2	
Norm (2,2)	.8649 E - 3	.9406 E -3	.1023 E -3	.2167 E -2	
Norm (2,3)	.1143 E - 2	.1128 E -2	.1173 E -2	.2063 E -1	
Number of iterations	7	7	7	9	
$(\Delta\ \text{par})_{max}$	-.18 E - 6	-.184 E -6	-.398 E -6	-.136 E -6	

Table 1.11

Parametervalues of the diagonal parametrisation
Structure (3,3)

	S/N ratio noisefree	75 dB	55 dB	35 dB	15 dB
$p_{11,1}$	-.1400 E -1	-.1377 E -1	-.1505 E -1	-.2498 E -1	instable
$p_{11,2}$	.2300 E 0	.2278 E 0	.2375 E 0	.3092 E 0	model
$p_{11,3}$	-.1000 E 1	-.9973 E 0	-.1009 E 1	-.1094 E 1	
$q_{11,1}$	.1000 E 1	.9778 E -2	.1041 E -1	.1479 E -1	
$q_{11,2}$	-.1500 E 0	-.1481 E 0	-.1540 E 0	-.1974 E 0	
$q_{11,3}$	.5000 E 0	.5001 E 0	.5000 E 0	.4983 E 0	
$q_{12,1}$	.1400 E 0	.1373 E 0	.1486 E 0	.2312 E 0	
$q_{12,2}$	-.1250 E 1	-.1245 E 1	-.1263 E 1	-.1388 E 1	
$q_{12,3}$	.1500 E 1	.1500 E 1	.1500 E 1	.1501 E 1	
$q_{13,1}$	-.1650 E 0	-.1624 E 0	-.1735 E 0	-.2543 E 0	
$q_{13,2}$	.1150 E 1	.1147 E 1	.1158 E 1	.1242 E 1	
$q_{13,3}$	-.1000 E 1	-.1000 E 1	-.1000 E 1	-.9990 E 0	
$p_{22,1}$	-.4881 E -1	-.4816 E -1	-.4478 E -1	-.1049 E -1	
$p_{22,2}$	.4082 E 0	.4080 E 0	.3994 E 0	.2828 E 0	
$p_{22,3}$	-.1177 E 1	-.1178 E 1	-.1173 E 1	-.1080 E 1	
$q_{21,1}$	.1061 E 0	.1064 E 0	.9603 E -1	.8771 E -2	
$q_{21,2}$	-.7648 E 0	-.7670 E 0	-.7590 E 0	-.6141 E 0	
$q_{21,3}$	.1500 E 1	.1500 E 1	.1500 E 1	.1499 E 1	
$q_{22,1}$	.1538 E -1	.1549 E -1	.1495 E -1	.6395 E -2	
$q_{22,2}$	-.7488 E -1	-.7519 E -1	-.7501 E -1	-.7386 E -1	
$q_{22,3}$	.9332 E -4	.2291 E -3	-.3321 E -3	-.4516 E -2	
$q_{23,1}$	.1562 E 0	.1554 E 0	.1497 E 0	.8697 E -1	
$q_{23,2}$	-.6262 E 0	-.6270 E 0	-.6218 E 0	-.5267 E 0	
$q_{23,3}$	.1000 E 1	.1000 E 1	.9998 E 0	.9953 E 0	

Table 1.12

Diagonal parametrisation
Structure (2,2)

	S/N ratio noisefree	75 dB	55 dB	35 dB	15 dB	-5 dB
Reconstruction Error						
output 1	.212 E -3	.278 E -3	.281 E -3	.566 E -3	.283 E -1	.723 E 0
output 2	.827 E -3	.145 E -2	.146 E -2	.184 E -2	.314 E -1	.786 E 0
Norm of correlation						
Norm (1,1)	.9307 E -3	.1103 E -2	.1106 E -2	.1335 E -2	.7625 E -2	.7772 E -1
Norm (1,2)	.1332 E -2	.1304 E -2	.1301 E -2	.1437 E -2	.7067 E -2	.6945 E -1
Norm (1,3)	.9285 E -3	.1064 E -2	.1070 E -2	.1304 E -2	.7091 E -2	.7204 E -1
Norm (2,1)	.1864 E -2	.1845 E -2	.1850 E -2	.2138 E -2	.1061 E -1	.1086 E 0
Norm (2,2)	.3232 E -2	.3397 E -2	.3423 E -2	.3894 E -2	.1471 E -1	.1389 E 0
Norm (2,3)	.4253 E -2	.4311 E -2	.4350 E -2	.1813 E -2	.1275 E -1	.1168 E 0
Number of iterations	7	7	7	5	13	31
$(\Delta\ par)_{max}$	-.410 E -6	-.388 E -6	.291 E -6	.823 E -6	-.172 E -11	.420 E -8

Table 1.13

Parametervalues of the diagonal parametrisation
Structure (2,2)

	S/N ratio noisefree	75 dB	55 dB	35 dB	15 dB	-5 dB
$p_{11,1}$	.1010 E 0	.1013 E 0	.1015 E 0	.1038 E 0	.1271 E 0	.2827 E 0
$p_{11,2}$	-.8337 E 0	-.8340 E 0	-.8342 E 0	-.8360 E 0	-.8531 E 0	-.7822 E 0
$q_{11,1}$	-.6783 E -1	-.6760 E -1	-.6784 E -1	-.7016 E -1	-.9278 E -1	-.1697 E 0
$q_{11,2}$	.5000 E 0	.4962 E 0	.4961 E 0	.4954 E 0	.4872 E 0	.3641 E 0
$q_{12,1}$	-.1001 E 1	-.1001 E 1	-.1001 E 1	-.1001 E 1	-.9961 E 0	-.5683 E 0
$q_{12,2}$	.1503 E 1	.1500 E 1	.1500 E 1	.1502 E 1	.1515 E 1	.1629 E 1
$q_{13,1}$	.9826 E 0	.9833 E 0	.9833 E 0	.9838 E 0	.9875 E 0	.7988 E 0
$q_{13,2}$	-.1001 E 1	-.9988 E 0	-.9987 E 0	-.9978 E 0	-.9886 E 0	-.9085 E 0
$p_{22,1}$	-.2246 E -1	-.2267 E -1	-.2375 E -1	-.3598 E -1	-.3303 E 0	-.5974 E 0
$p_{22,2}$	-.6504 E 0	-.6500 E 0	-.6483 E 0	-.6290 E 0	-.1804 E 0	.2543 E 0
$q_{21,1}$	.2686 E -1	.2751 E -1	.3046 E -1	.6319 E -1	.7654 E 0	.1694 E 1
$q_{21,2}$	.1505 E 1	.1504 E 1	.1504 E 1	.1503 E 1	.1479 E 1	.1186 E 1
$q_{22,1}$	-.7839 E -1	-.7832 E -1	-.7820 E -1	-.7695 E -1	-.6964 E -1	.2099 E -1
$q_{22,2}$	.4192 E -4	-.1768 E -3	-.1930 E -3	-.1379 E -3	.4689 E -1	.7981 E 0
$q_{23,1}$	-.9837 E -1	-.9794 E -1	-.9623 E -1	-.7690 E -1	.3492 E 0	.4226 E 0
$q_{23,2}$	.9981 E 0	.9983 E 0	.9979 E 0	.9941 E 0	.9635 E 0	.7655 E 0

Table 1.14

Diagonal parametrisation
Structure (3,1)

	S/N ratio noisefree	75 dB	55 dB	35 dB	15 dB
Reconstruction Error					
output 1	.169 E -27	.503 E -4	.511 E -4	.327 E -3	instable
output 2	.436 E - 2	.457 E -2	.456 E -2	.472 E -2	model
Norm of correlation					
Norm (1,1)	.8463 E -15	.1186 E -3	.1351 E -3	.8409 E -3	
Norm (1,2)	.1178 E -14	.1168 E -3	.1308 E -3	.6694 E -3	
Norm (1,3)	.7834 E -15	.6726 E -4	.1003 E -3	.6930 E -3	
Norm (2,1)	.4823 E - 2	.4842 E -2	.4843 E -2	.4965 E -2	
Norm (2,2)	.7556 E - 2	.7544 E -2	.7543 E -2	.7637 E -2	
Norm (2,3)	.8165 E - 2	.8181 E -2	.8171 E -2	.8141 E -2	
Number of iterations	4	4	5	7	
$(\Delta \text{ par})_{max}$	-.27 E - 6	-.254 E -6	.195 E -7	-.933 E -7	

Table 1.15

Parametervalues of the diagonal parametrisation
Structure (3,1)

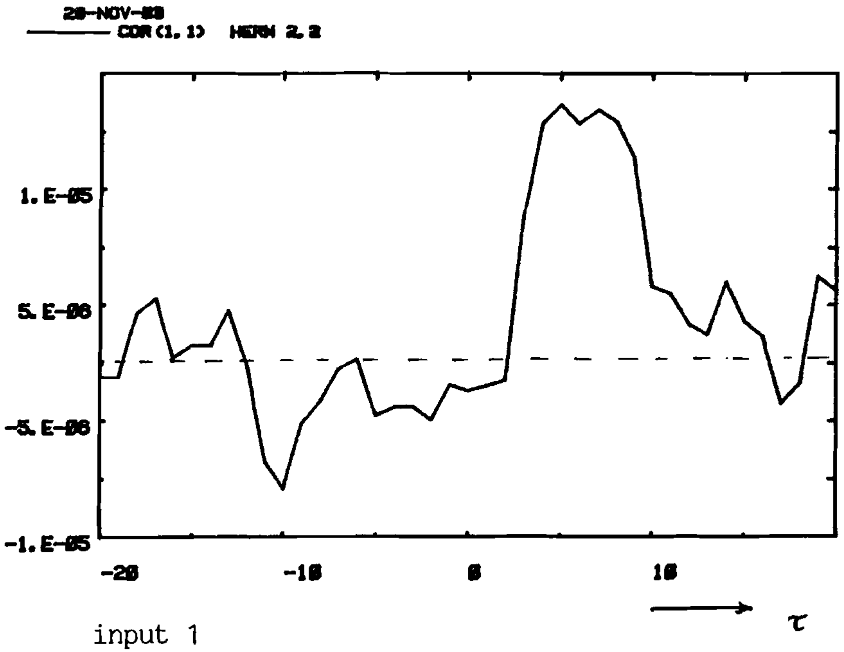
	S/N ratio noisefree	75 dB	55 dB	35 dB	15 dB
$p_{11,1}$	-.1400 E -1	-.1377 E -1	-.1505 E -1	-.2498 E -1	instable
$p_{11,2}$	.2300 E 0	.2278 E 0	.2375 E 0	.3092 E 0	model
$p_{11,3}$	-.1000 E 1	-.9973 E 0	-.1009 E 1	-.1094 E 1	
$q_{11,1}$	.1000 E 1	.9778 E -2	.1041 E -1	.1479 E -1	
$q_{11,2}$	-.1500 E 0	-.1481 E 0	-.1540 E 0	-.1974 E 0	
$q_{11,3}$	.5000 E 0	.5001 E 0	.5000 E 0	.4983 E 0	
$q_{12,1}$	.1400 E 0	.1373 E 0	.1486 E 0	.2312 E 0	
$q_{12,2}$	-.1250 E 1	-.1245 E 1	-.1263 E 1	-.1388 E 1	
$q_{12,3}$	.1500 E 1	.1500 E 1	.1500 E 1	.1501 E 1	
$q_{13,1}$	-.1650 E 0	-.1624 E 0	-.1735 E 0	-.2543 E 0	
$q_{13,2}$	.1150 E 1	.1147 E 1	.1158 E 1	.1242 E 1	
$q_{13,3}$	-.1000 E 1	-.1000 E 1	-.1000 E 1	-.9990 E 0	
p_{211}	-.6593 E 0	-.6593 E 0	-.6594 E 0	-.6602 E 0	
q_{211}	.1493 E 1	.1493 E 1	.1493 E 1	.1492 E 0	
q_{221}	-.6814 E -2	-.6813 E -2	-.6858 E -2	-.7312 E -2	
q_{231}	.9914 E 0	.9914 E 0	.9910 E 0	.9871 E 0	

Table 1.16

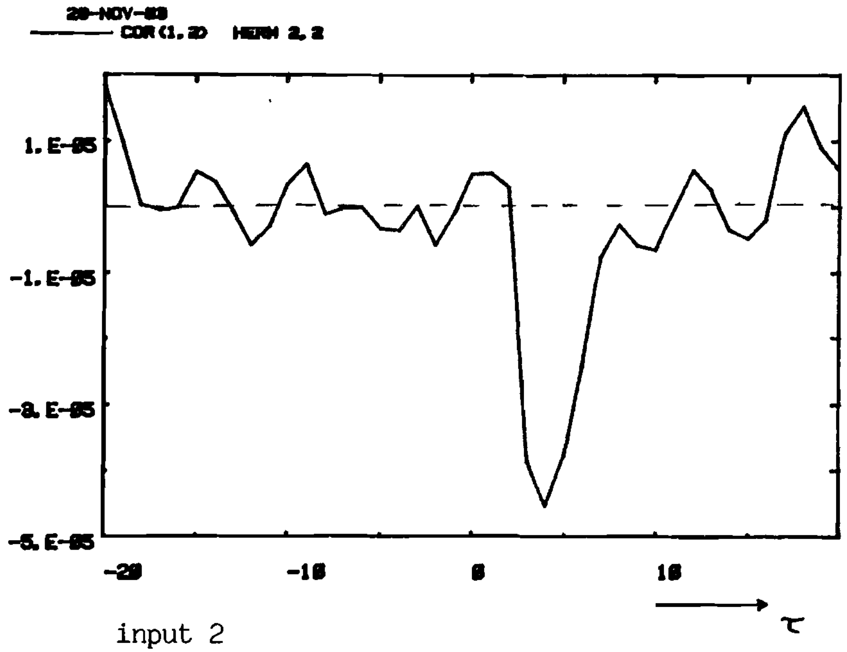
APPENDIX 2

In this appendix the graphical results of the parameter estimations are presented. In the figures the normed correlation functions are given.

Fig. 2.1 contains corr(1,i,τ), i=1,2,3 for S/N = infinite.
Fig. 2.2 contains corr(2,i,τ), i=1,2,3 for S/N = infinite.
Fig. 2.3 contains corr(1,i,τ), i=1,2,3 for S/N = 75 dB.
Fig. 2.4 contains corr(2,i,τ), i=1,2,3 for S/N = 75 dB.
Fig. 2.5 contains corr(1,i,τ), i=1,2,3 for S/N = 55 dB.
Fig. 2.6 contains corr(2,i,τ), i=1,2,3 for S/N = 55 dB.
Fig. 2.7 contains corr(1,i,τ), i=1,2,3 for S/N = 35 dB.
Fig. 2.8 contains corr(2,i,τ), i=1,2,3 for S/N = 35 dB.
Fig. 2.9 contains corr(1,i,τ), i=1,2,3 for S/N = 15 dB.
Fig. 2.10 contains corr(2,i,τ), i=1,2,3 for S/N = 15 dB.
Fig. 2.11 contains corr(1,i,τ), i=1,2,3 for S/N = -5 dB.
Fig. 2.12 contains corr(2,i,τ), i=1,2,3 for S/N = -5 dB.



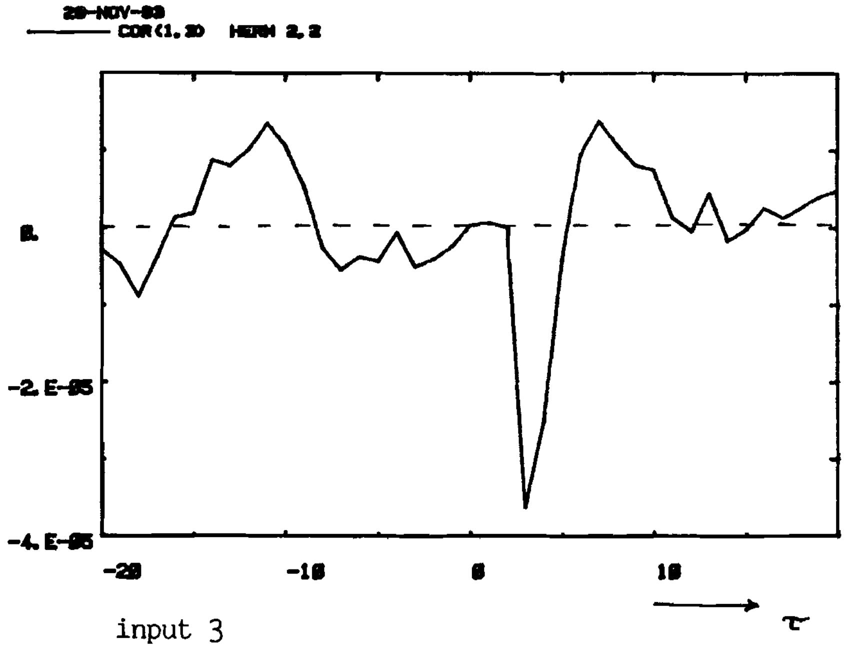
input 1



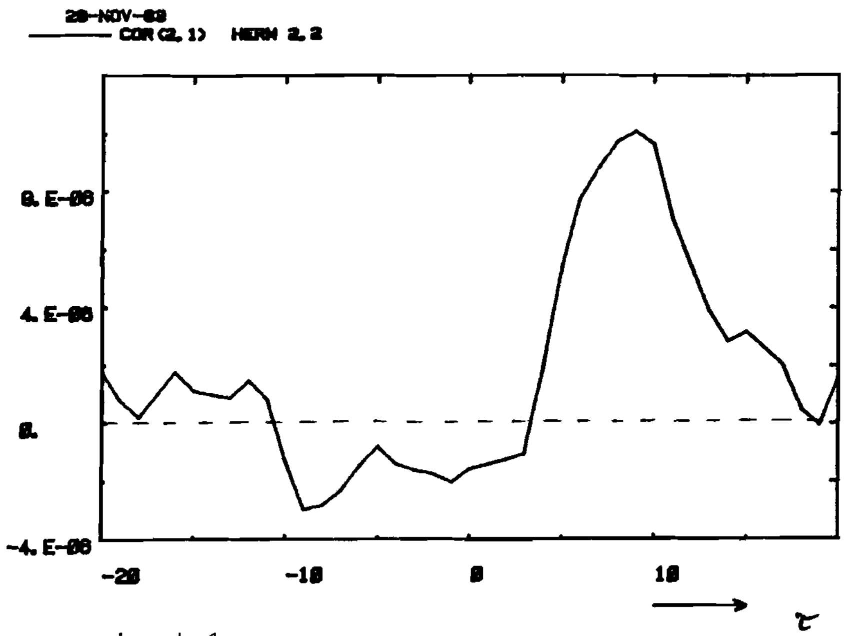
input 2

Fig. 2.1

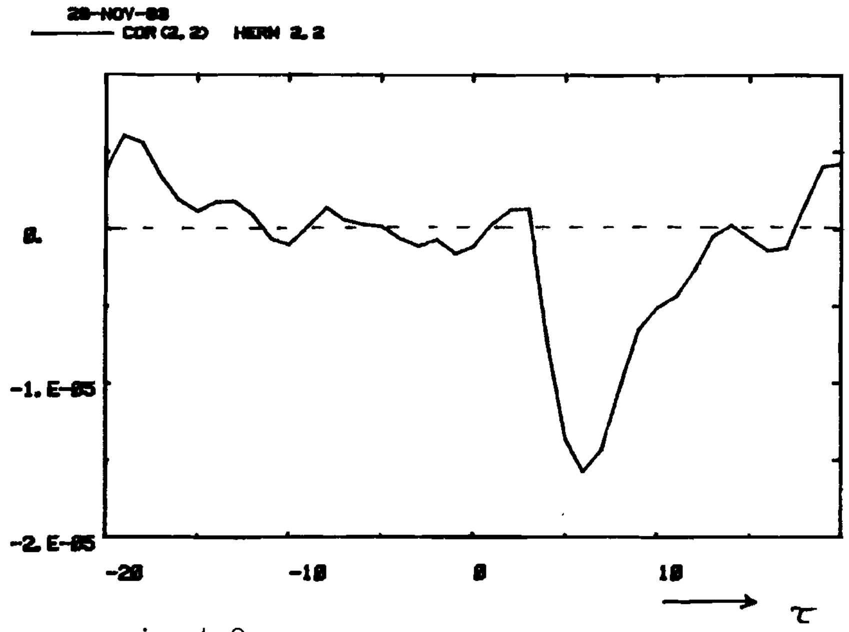
Correlation between estimated output error and inputs. Hermite parametrisation, structure (2,2); output 1. S/N = infinite. The vertical axis has to be multiplied by 256.



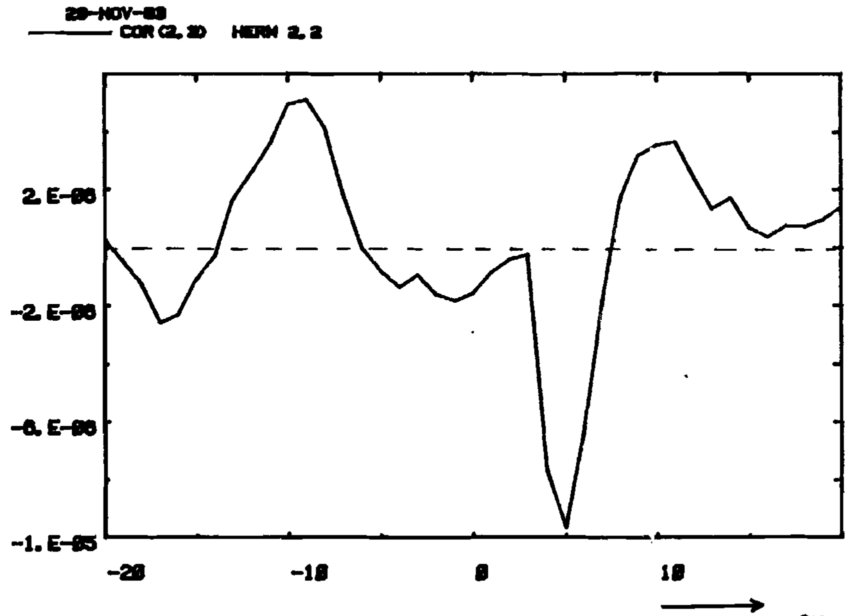
input 3



input 1



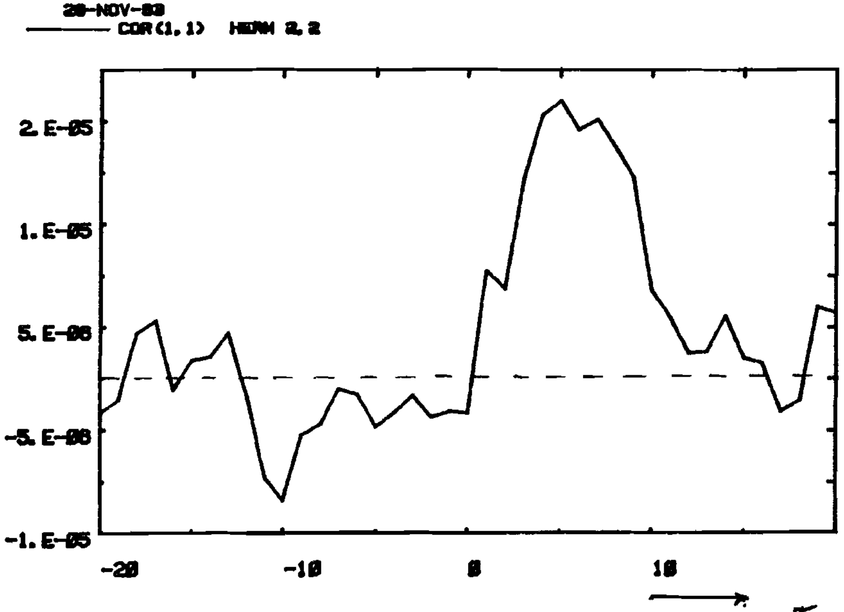
input 2



input 3

Fig 2.2

Correlation between estimated output error and inputs. Hermite parametrisation, structure (2,2); output 2. S/N = infinite. The vertical axis has to be multiplied by 256.

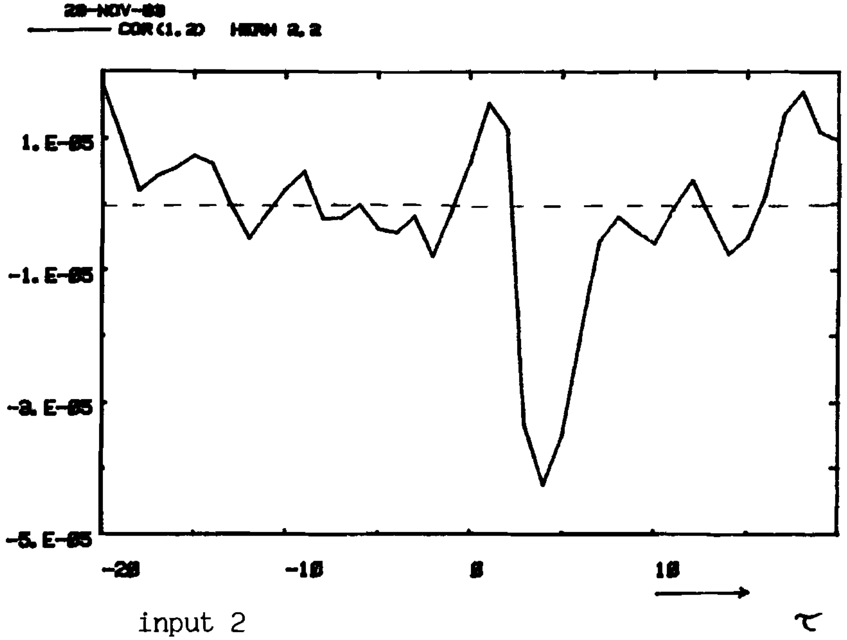


input 1

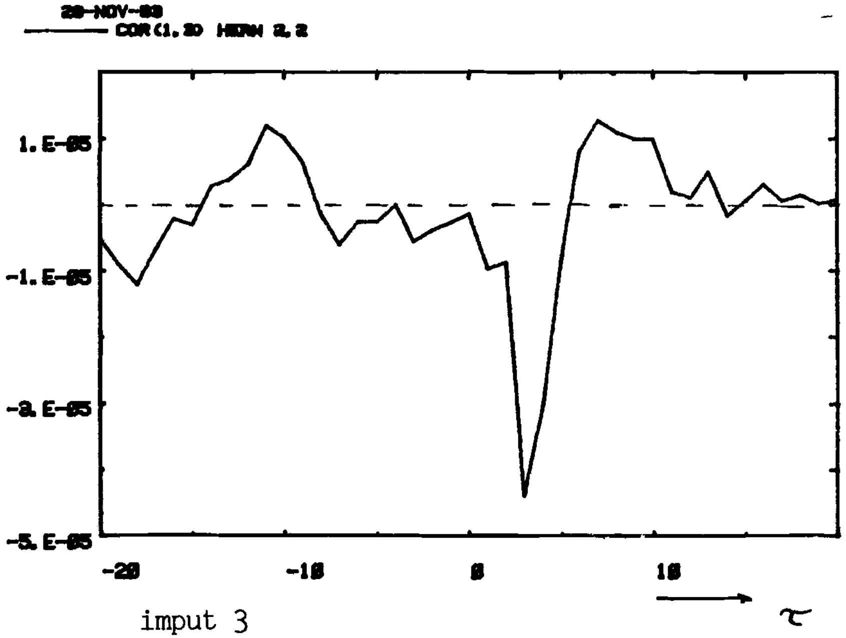
Fig 2.3

Correlation between estimated
output error and inputs.
Structure (2,2); output 1.
S/N = 75 dB.
The vertical axis has to be multiplied
by 256.

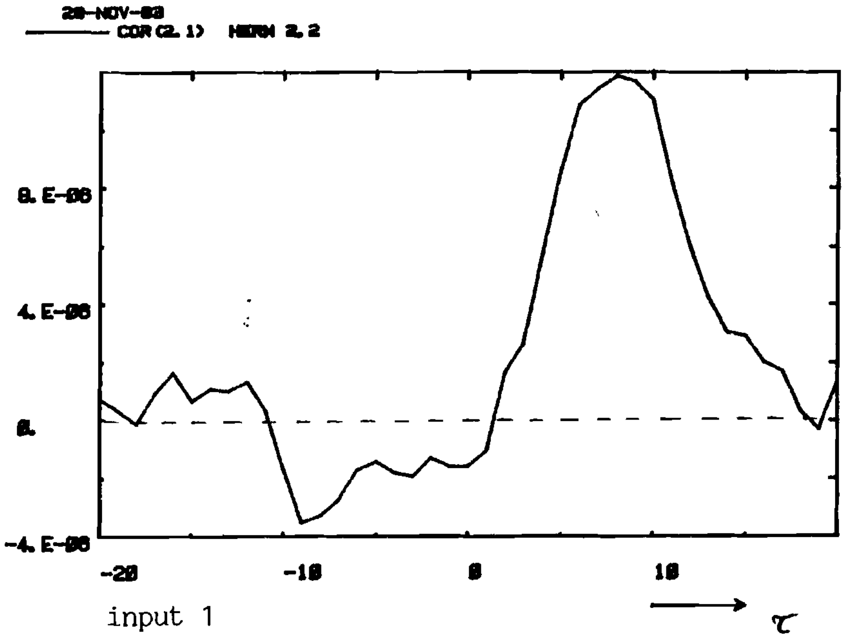
Hermite parametrization.



input 2



imput 3



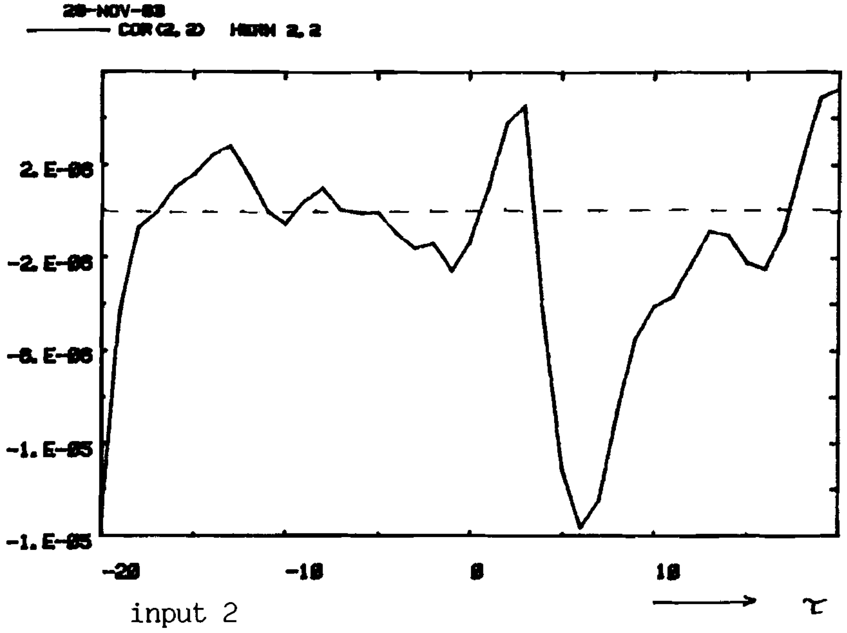


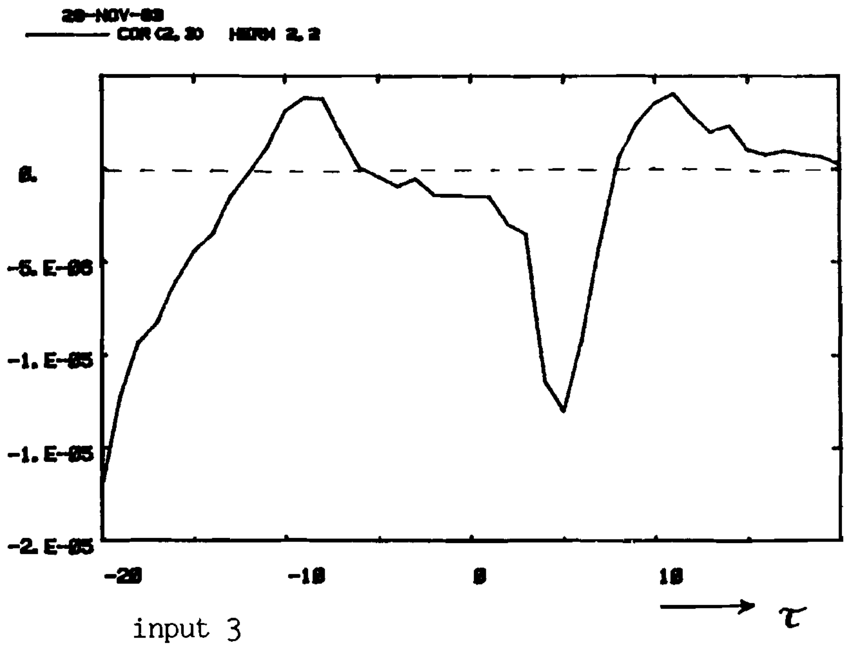
Fig 2.4

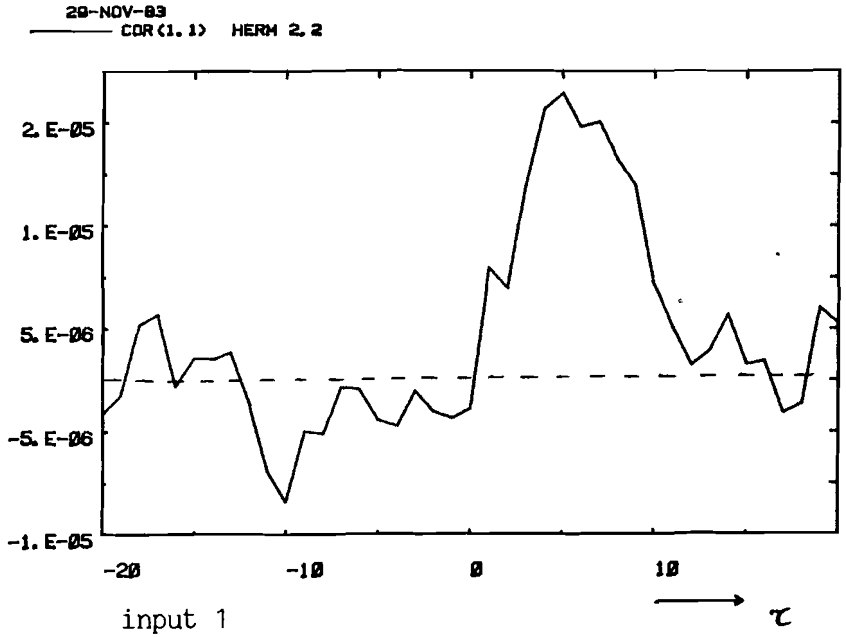
Correlation between estimated output error and inputs.
Structure (2,2); output 2.

Hermite parametrisation.

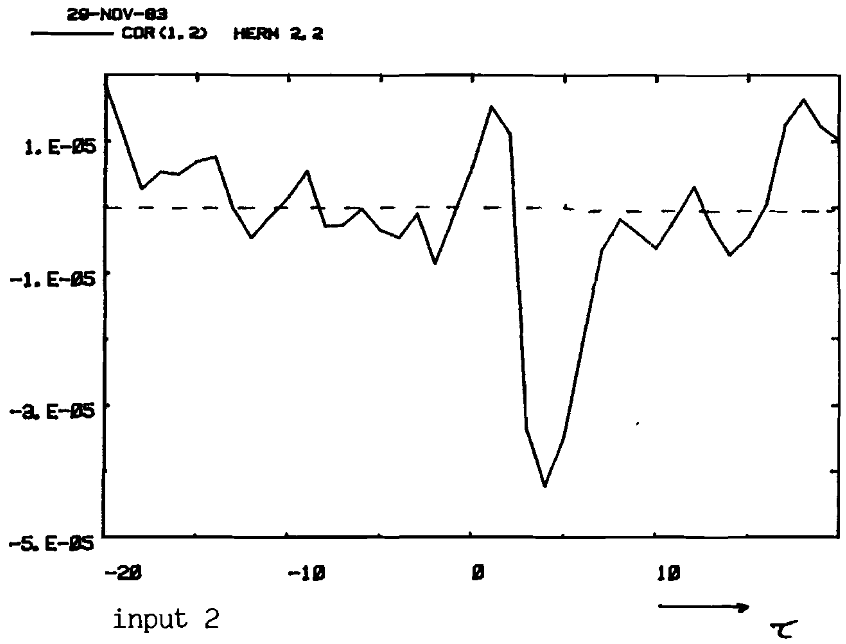
S/N = 75 dB.

The vertical axis has to be multiplied by 256.





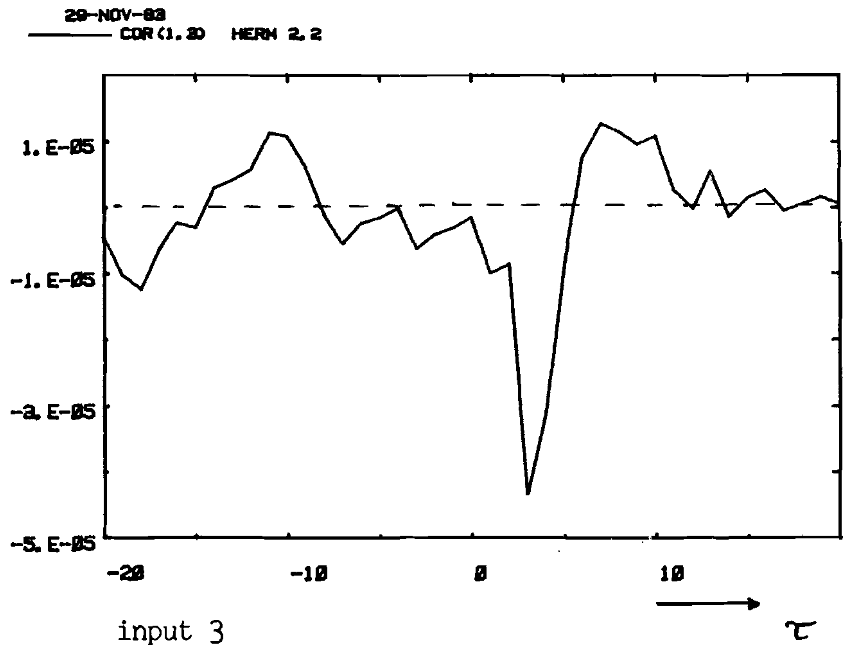
input 1



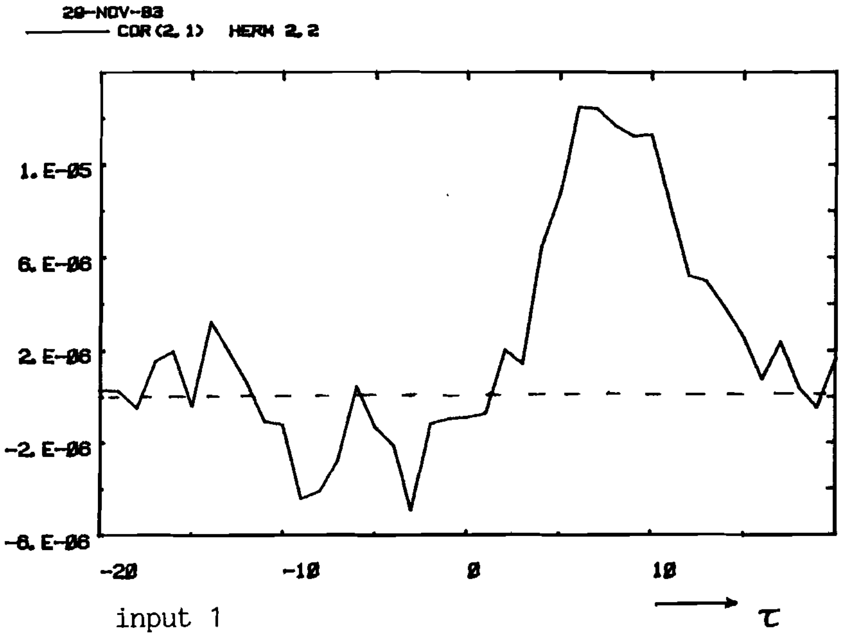
input 2

Fig 2.5

Correlation between estimated output error and inputs.
Hermite parametrisation, structure (2,2); output 1.
S/N = 55 dB.
The vertical axis has to be multiplied by 256.



input 3



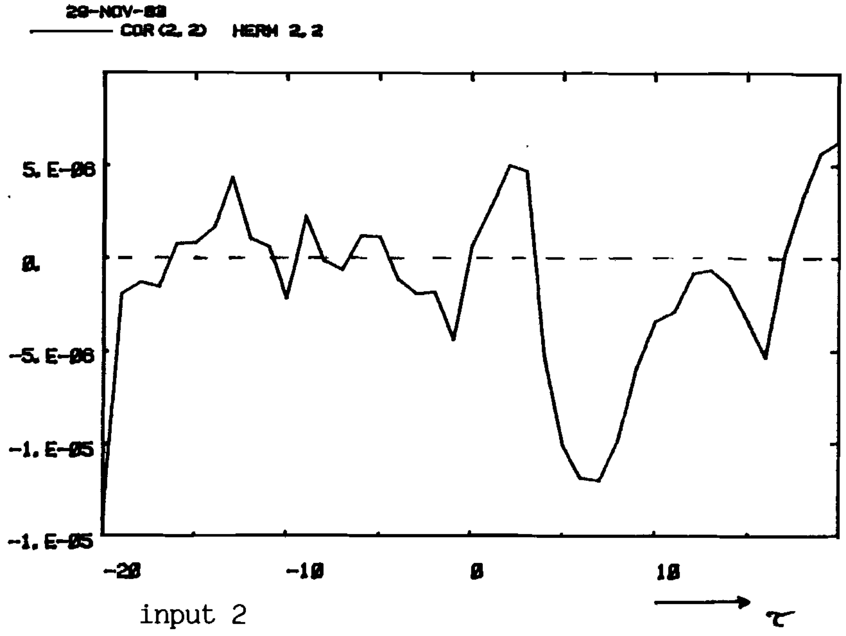
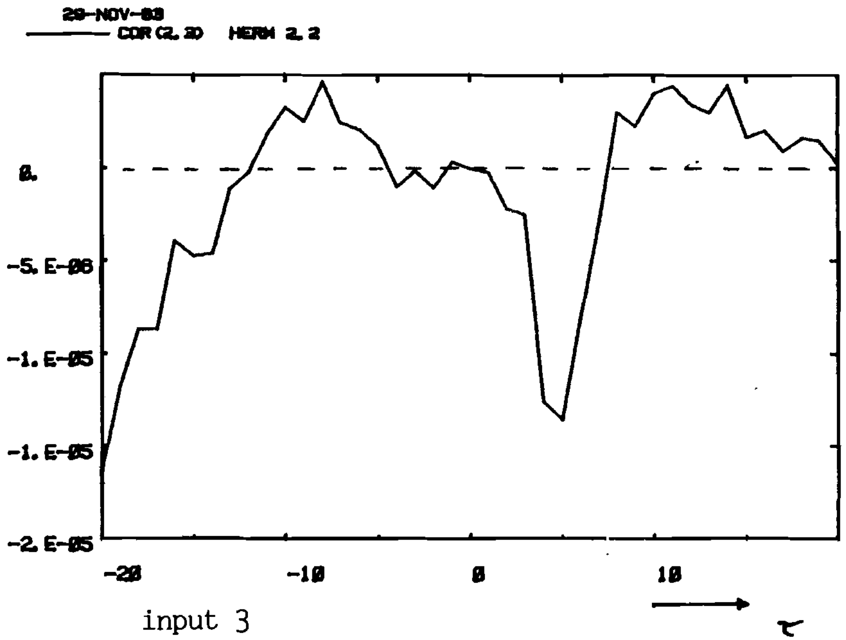
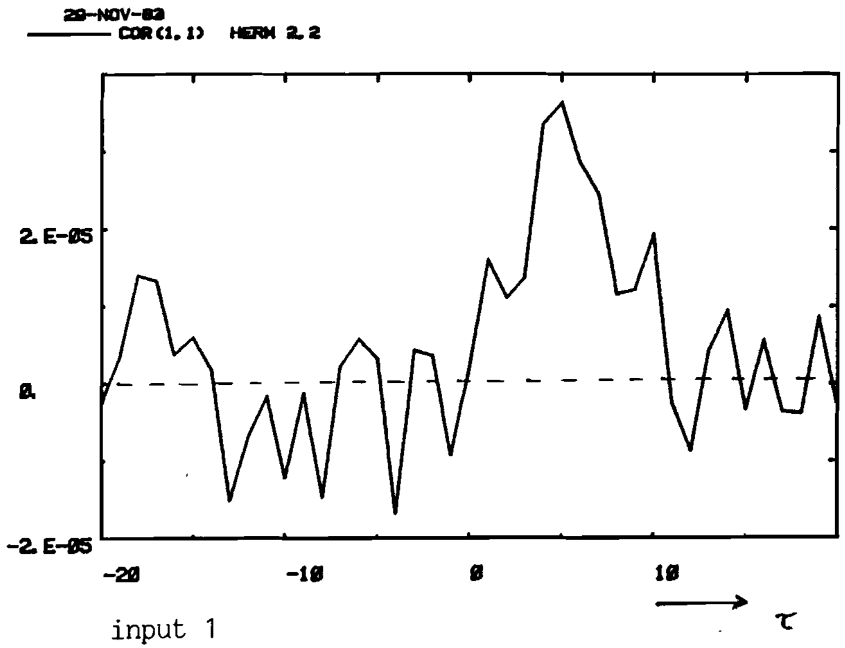


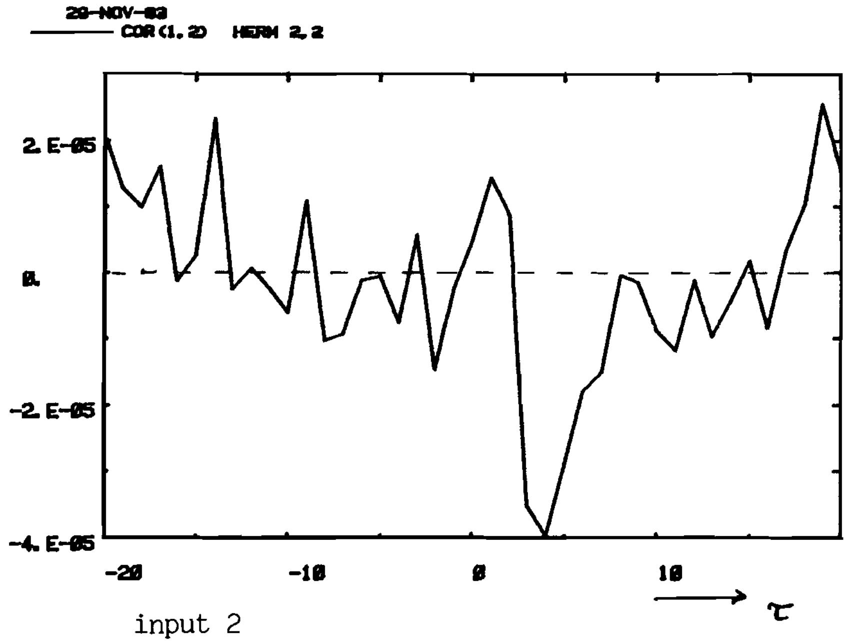
Fig 2.6

Correlation between estimated output error and inputs. Hermite parametrisation, structure (2,2); output 2. S/N = 55 dB. The vertical axis has to be multiplied by 256.





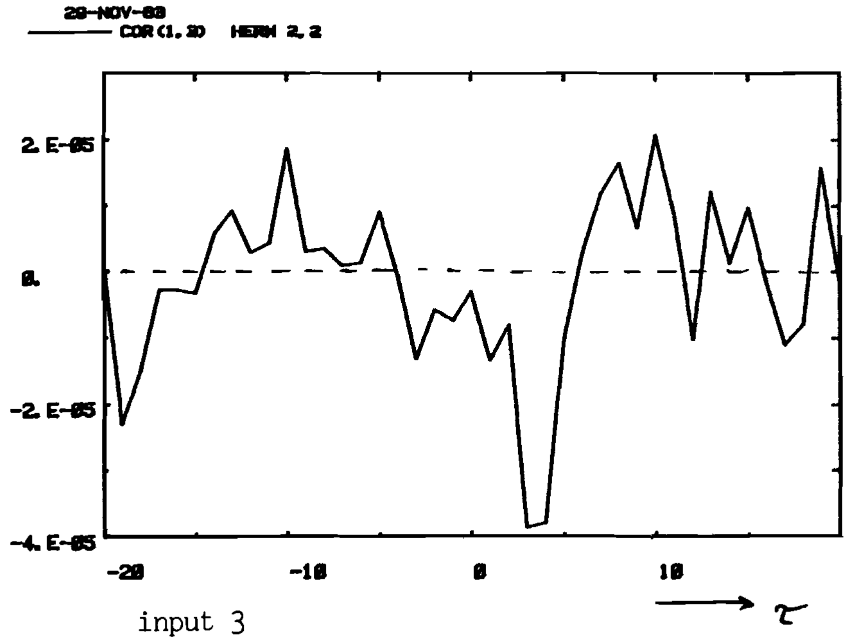
input 1



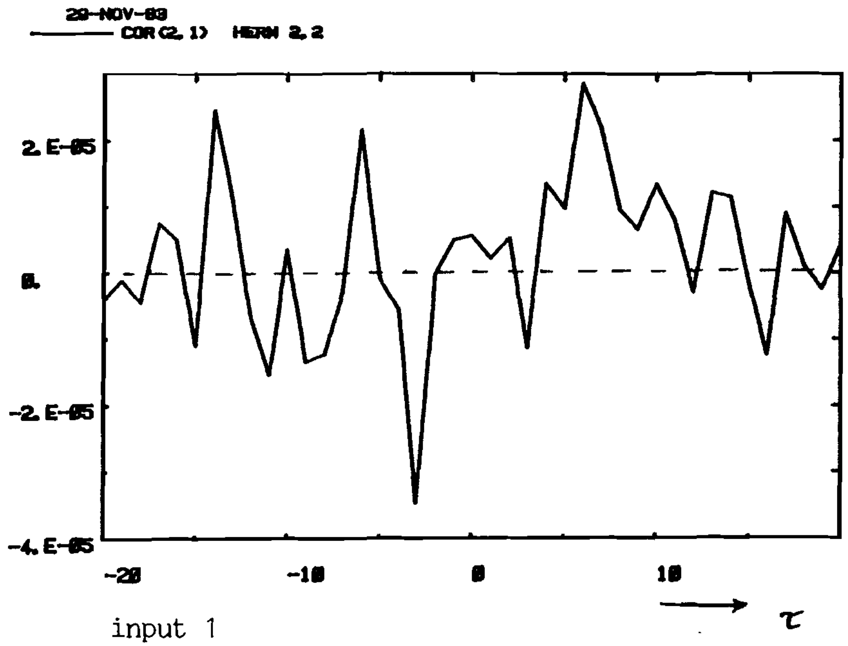
input 2

Fig 2.7

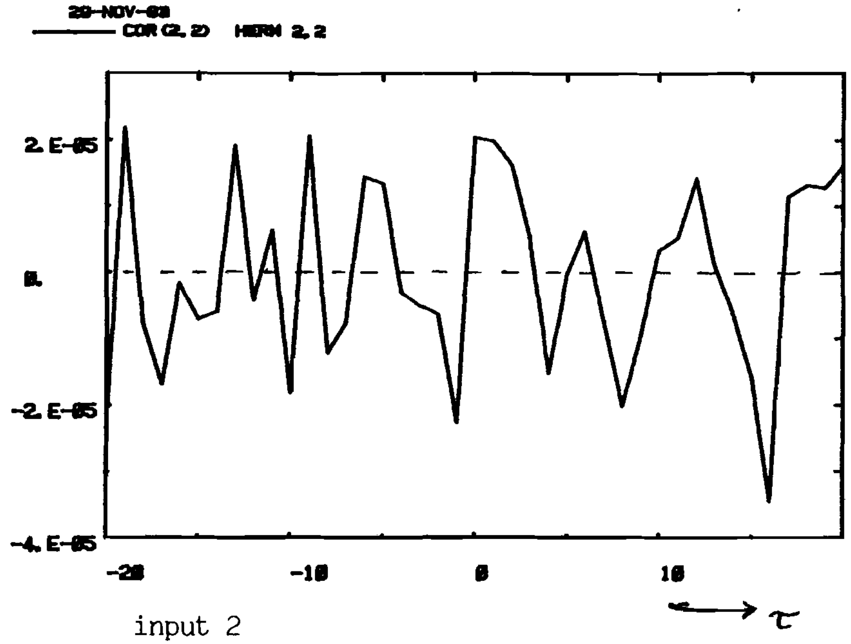
Correlation between estimated output error and inputs. Hermite parametrisation, structure (2,2); output 1. S/N = 35 dB. The vertical axis has to be multiplied by 256.



input 3



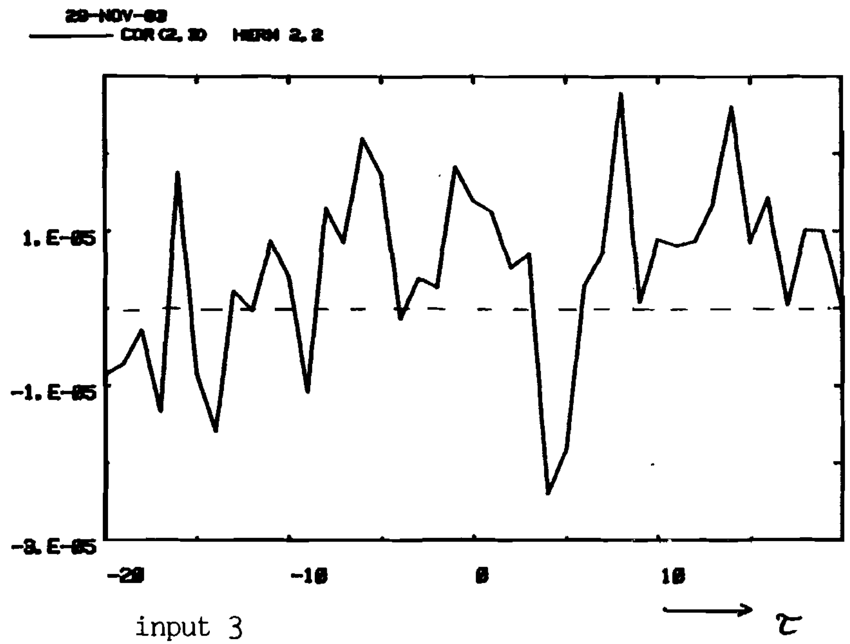
input 1



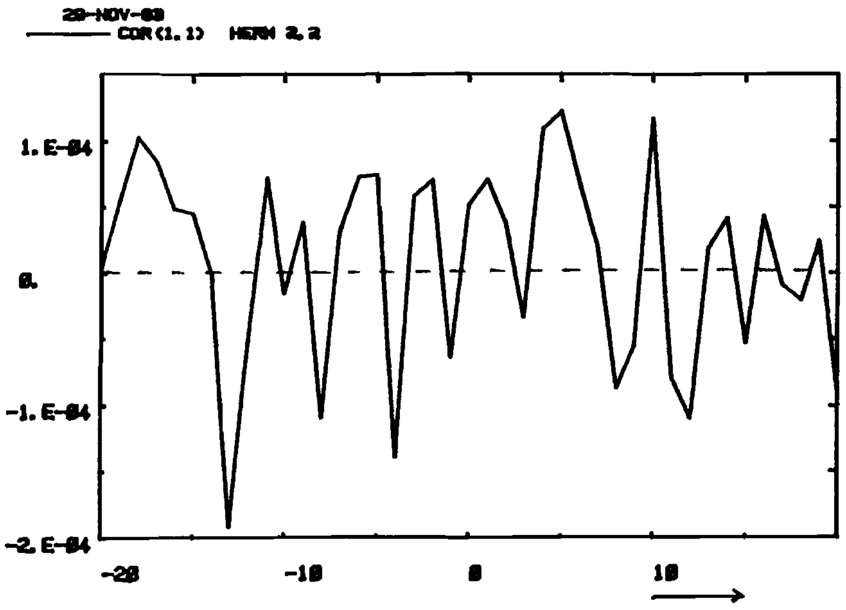
input 2

Fig 2.8

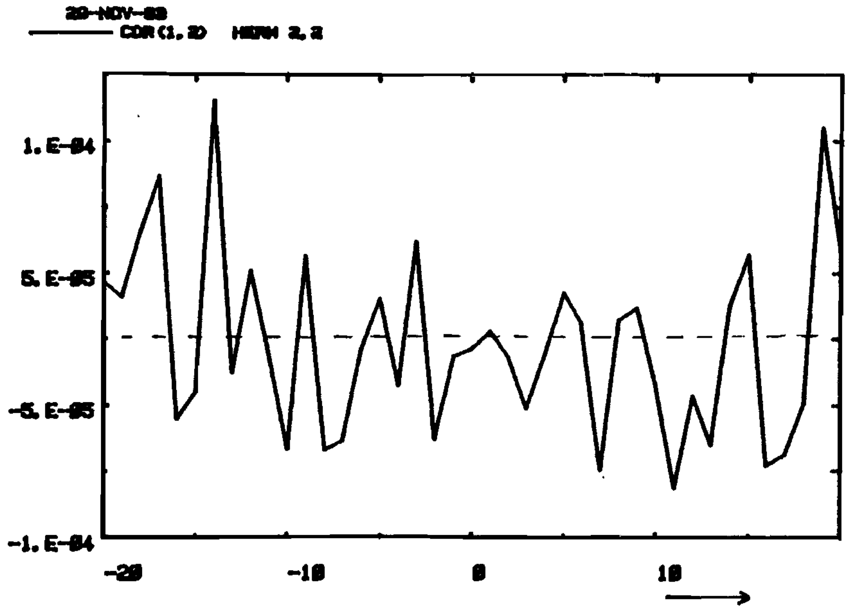
Correlation between estimated output error and inputs. Hermite parametrisation, structure (2,2); output 2. S/N = 35 dB. The vertical axis has to be multiplied by 256.



input 3



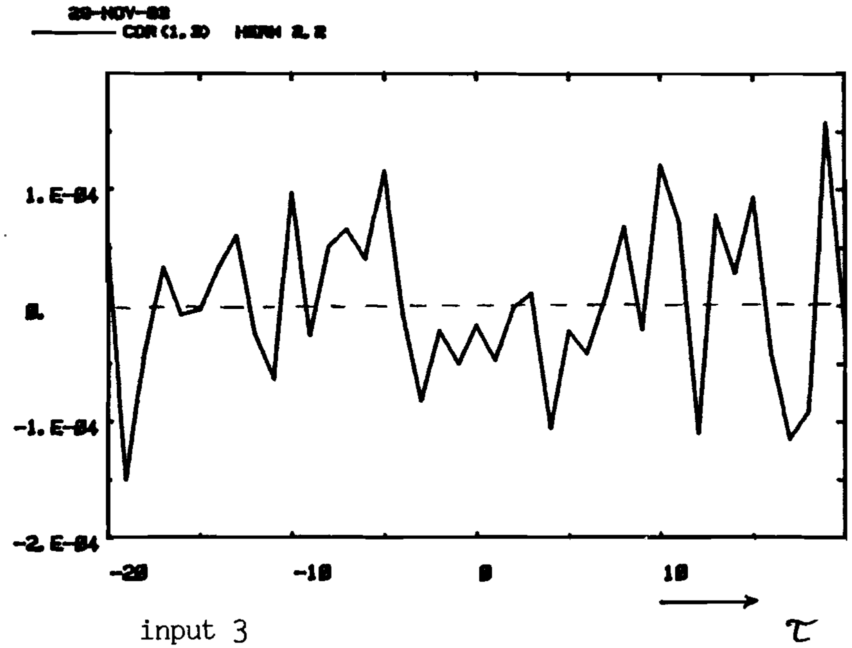
input 1



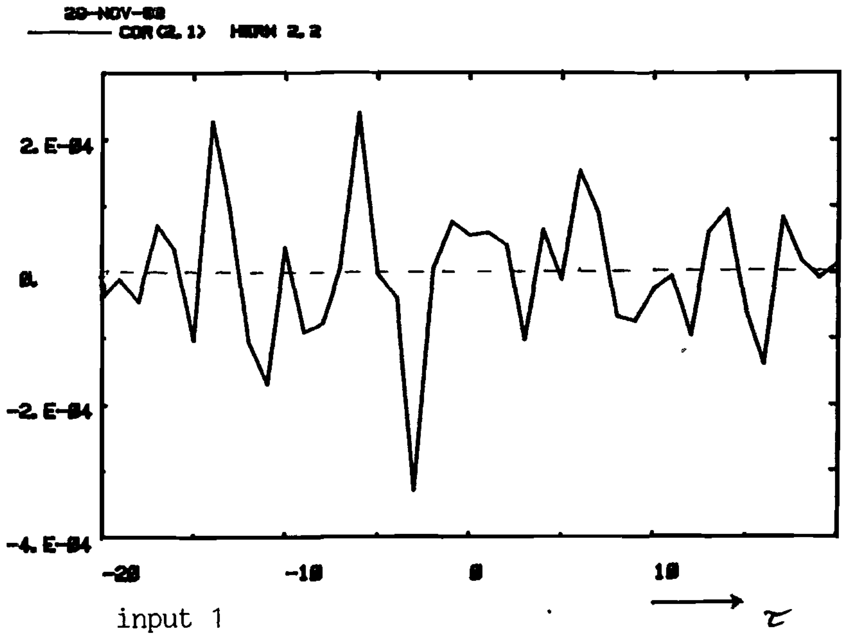
input 2

Fig 2.9

Correlation between estimated output error and inputs. Hermite parametrisation, structure (2,2); output 1. S/N = 15 dB. The vertical axis has to be multiplied by 256.



input 3



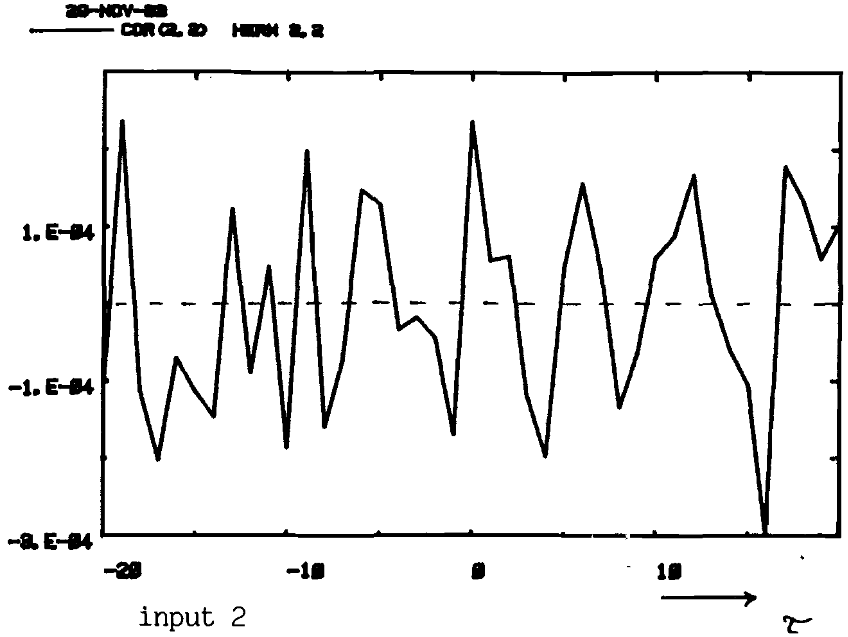
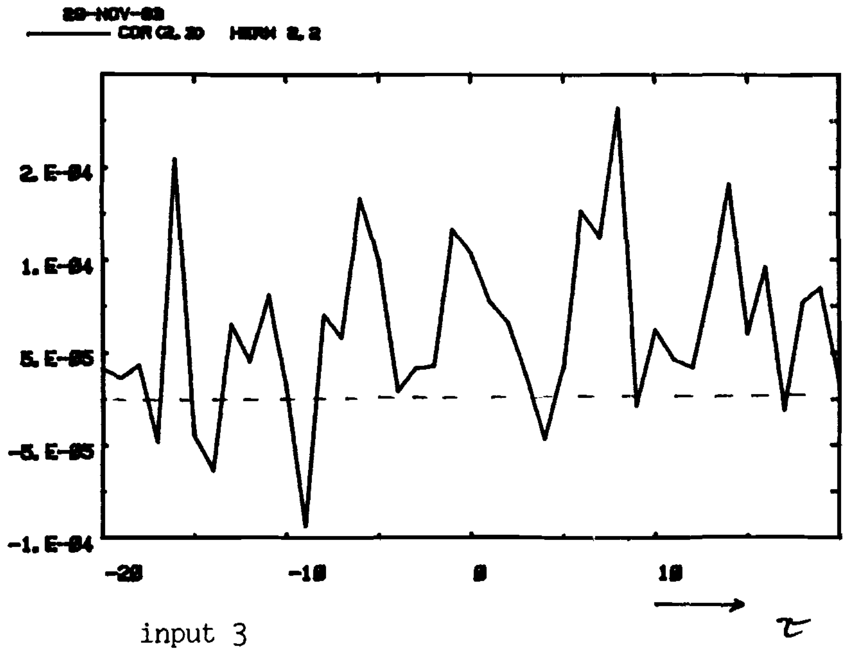
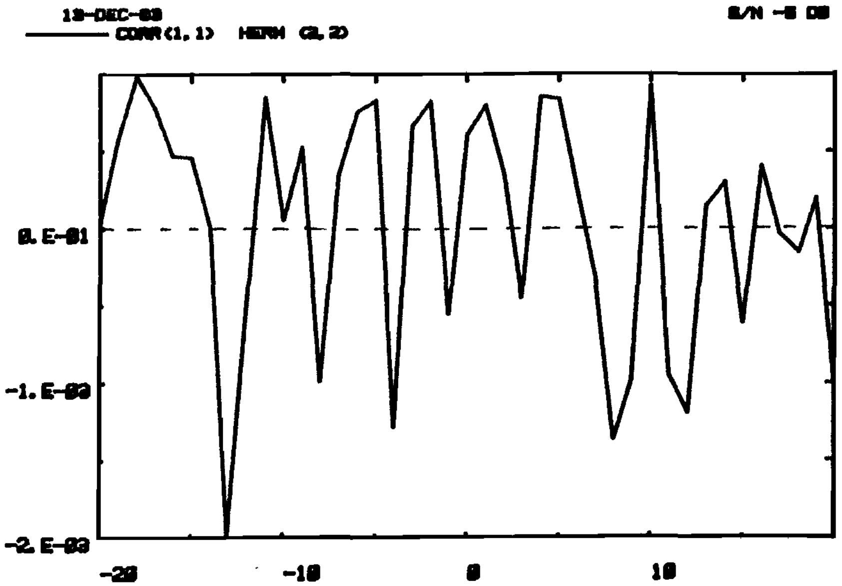


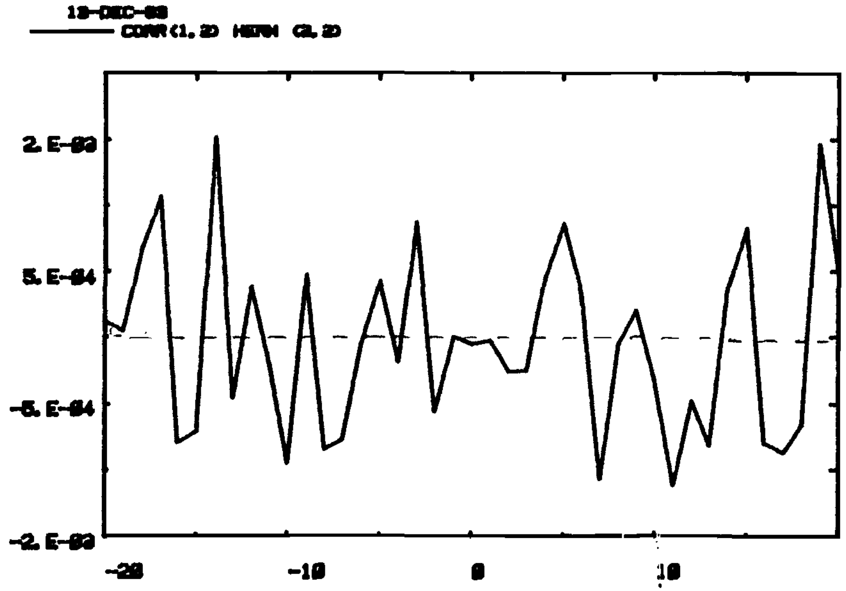
Fig 2.10

Correlation between estimated output error and inputs. Hermite parametrisation, structure (2,2); output 2. S/N = 15 dB. The vertical axis has to multiplied by 256.





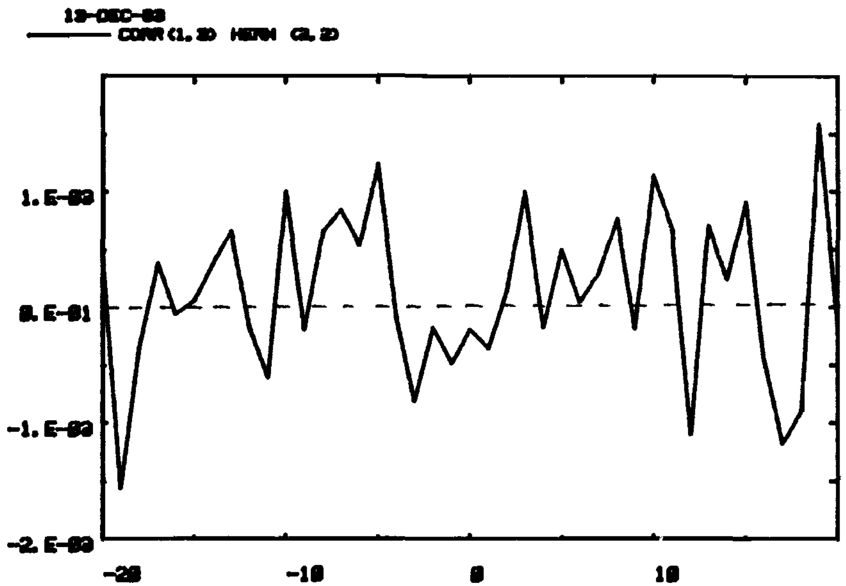
input 1 $\longrightarrow \tau$



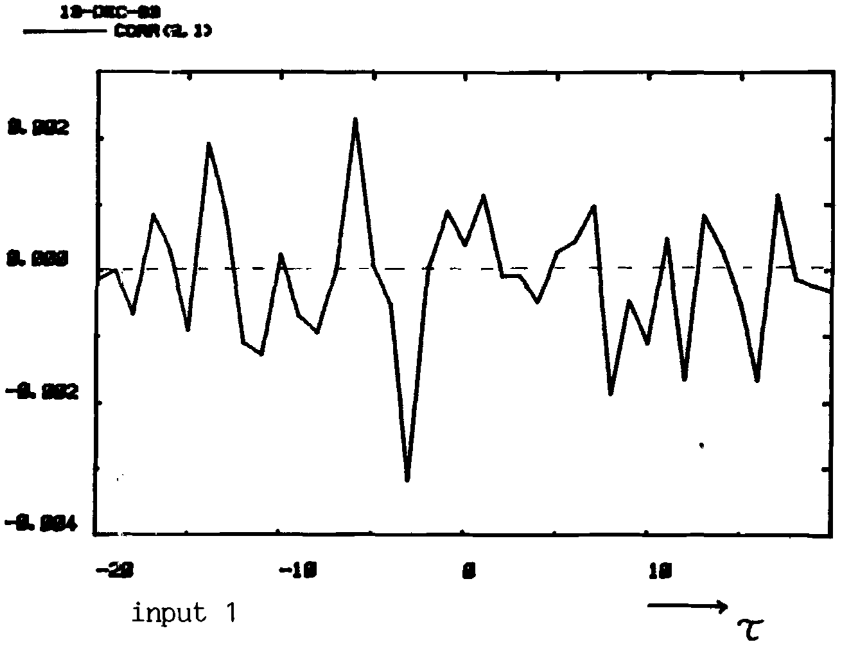
input 2 $\longrightarrow \tau$

Fig 2.11

Correlation between estimated output error and inputs. Hermite parametrisation, structure (2,2); output 1. S/N = -5 dB. The vertical axis has to be multiplied by 256.



input 3 $\longrightarrow \tau$



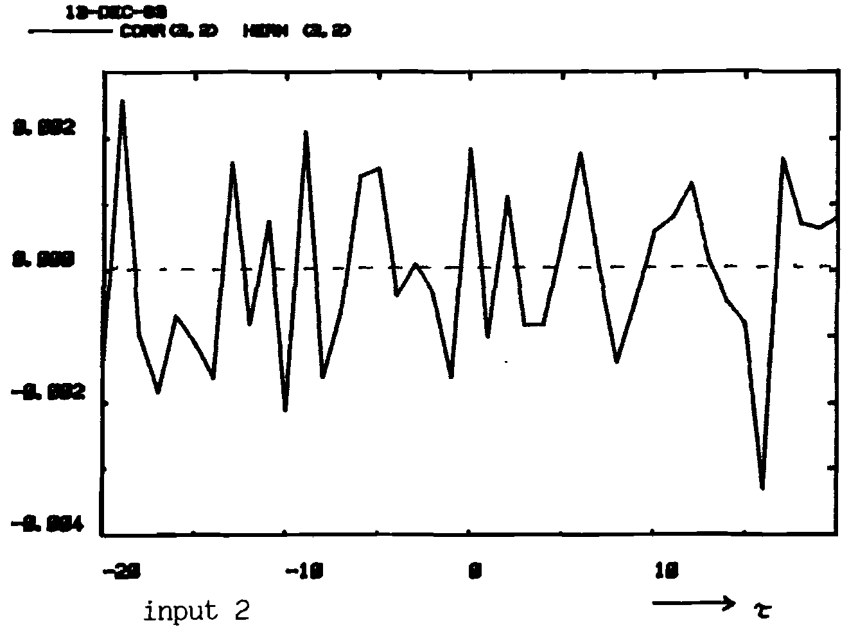
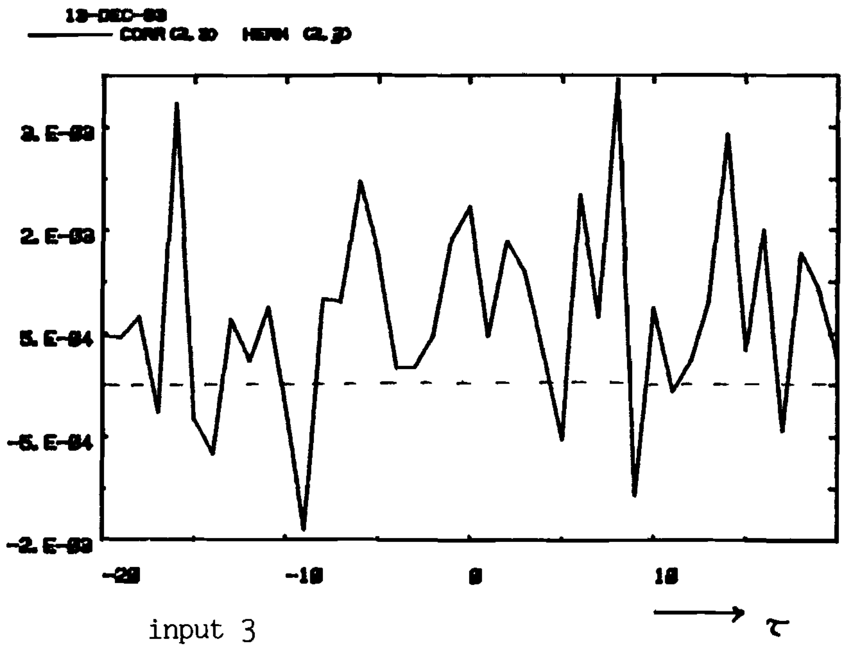


Fig 2.12

Correlation between estimated output error and inputs. Hermite parametrisation, structure (2,2); output 2. S/N = -5 dB. The vertical axis has to be multiplied by 256.



-5dB.

Extended list of references

-Antoulas, A.C. (1981)
ON CANONICAL FORMS FOR LINEAR CONSTANT SYSTEMS.
Intern. J. of Control, vol. 33 (1981), nr. 1,p. 95-122.

-Anderson, B.D.O , Bitmead, R.R. (1978)
A DYNAMICAL INTERPRETATION OF COLUMN AND ROW PROPERNESS OF MFDS.
IEEE Trans. Autom. Contr., vol. AC-23 (1978), nr. 6, p. 1101-1102.

-Anderson, B.D.O., Moore, J.B. (1979)
OPTIMAL FILTERING
Prentice Hall, Englewood Cliffs ,1979.

-Barnett , S. (1973)
MATRICES, POLYNOMIALS, LINEAR TIME-INVARIANT SYSTEMS.
IEEE Trans. Autom. Contr., vol. AC-18 (1973),nr. 1, pp 3-10.

-Barrett-Lenard, I. , Blair, J.R. (1981)
ESTIMATIONS OF COEFFICIENTS FOR MULTIPLE INPUT SYSTEMS MODELS WITHOUT EMPLOYING COMMON DENOMINATOR STRUCTURE;
Intern. J. of Control, vol. 33 (1981), nr. 1, pp. 123-136

-Beghelli, S. , Guidorzi, R.P. (1976)
A NEW INPUT-OUTPUT CANONICAL FORM FOR MULTIVARIABLE SYSTEMS.
IEEE Trans. Autom. Contr., vol. AC-21 (1976), nr. 5, pp. 692-696.

-Beghelli, S. , Guidorzi, R.P. (1982)
TRANSFORMATION BETWEEN INPUT-OUTPUT MULTISTRUCTURAL MODELS: PROPERTIES AND APPLICATIONS.
(To appear)

-Birkhoff, G. , MacLane, S. (1977)
A survey of modern algebra.
MacMillan Publ. Co. Inc., New York, 1977.

-Denham, M.J. (1974)
CANONICAL FORMS FOR THE IDENTIFICATION OF MULTIVARIABLE SYSTEMS
IEEE Trans. Autom. Contr., vol. AC-19 (1974), nr.6, pp. 646-656

-Dickinson, B.W. (1974)
PROPERTIES AND APPLICATIONS OF MATRIX FRACTION DESCRIPTIONS OF LINEAR SYSTEMS.
PhD-thesis, Stanford University , 1974

-Dickinson, B.W. , Kailath, T., Morf, M. (1974)
CANONICAL MATRIX FRACTION AND STATE SPACE DESCRIPTIONS FOR DETERMINISTIC AND STOCHASTIC LINEAR SYSTEMS.
IEEE Trans. Autom. Contr., vol. AC-19 (1974), nr. 6, pp. 656-667.

-Dickinson, B.W., Morf, M. , Kailath, T. (1974)
A MINIMAL REALIZATION ALGORITHM FOR MATRIX SEQUENCES.
IEEE Trans. Autom. Contr., vol. AC-19 (1974), nr. 1, pp. 31-37.

-El-Sherief, H., Sinha, N.K. (1979)
CHOICE OF MODELS FOR THE IDENTIFICATION OF LINEAR MULTIVARIABLE DISCRETE-TIME SYSTEMS.
Proc. IEE, vol. 126 (1979), nr. 12, pp 1326-1330.

-Emre, E. (1980)
ON A NATURAL REALIZATION OF MATRIX FRACTION DESCRIPTIONS.
IEEE Trans. Autom. Contr., vol. AC-25 (1980), nr. 2, pp. 288-289.

-Gantmacher, F.R. (1959)
The theory of matrices, vol. 1 and 2
Chelsea, New York, 1959.

-Gauthier, A. , Landau, J.D. (1978)
ON THE RECURSIVE IDENTIFICATION OF MULTI-INPUT MULTI-OUTPUT SYSTEMS.
Automatica, vol. 14 (1978), pp. 609-614.

-Guidorzi,R.P. ,Losito, M.P., Muratori, T. (1982)
THE RANGE ERROR TEST IN THE STRUCTURAL IDENTIFICATION OF LINEAR MULTIVARIABLE SYSTEMS.
IEEE Trans. Autom. Contr., vol. AC-27 (1982), nr. 5, pp 1044-1053.

-Guidorzi, R.P. (1975)
CANONICAL STRUCTURES IN THE IDENTIFICATION OF MULTIVARIABLE SYSTEMS.
Automatica , vol. 11 (1975), nr. 11, pp 361-374.

-Guidorzi, R.P. (1981)
INVARIANTS AND CANONICAL FORMS FOR SYSTEMS STRUCTURAL AND PARAMETRIC IDENTIFICATION.
Automatica, vol. 17 (1981), nr. 1, pp. 117-133.

-Hajdasinski, A.K., Eykhoff, P., Damen, A.A.H. , v.d. Boom, A.J.W.
THE CHOICE AND USE OF DIFFERENT MODEL SETS FOR SYSTEMS IDENTIFICATION.
Prepr./Proc. 6-th IFAC Symp. Ident. and Syst. Param. Estim., Arlington, Va, 1982, pp. 2-10.

-Jakeman, A.J. (1979)
MULTIVARIABLE IV ESTIMATORS: THE CHOICE BETWEEN ALTERNATIVES.
Prepr./Proc. 5-th IFAC Symp. Ident. and Syst. Param. Estim., Darmstadt, 1979,pp. 459-467.

-Jakeman, A.J., Young, P. (1979)
REFINED INSTRUMENTAL VARIABLE METHODS OF RECURSIVE TIME-SERIES ANALYSIS.
Intern. J. of Control, vol. 29 (1979), nr. 4, pp. 621-644.

-Joseph, P., Lewis, J., Tou, J. (1961)
Trans. AIEE, vol. 80 (1961), nr. 18, 1961.

-Kailath, T. (1980)
LINEAR SYSTEMS.
Prentice Hall, Englewood Cliffs, 1980.

-Kashyap, R.L., Nasburg,R.E.(1974)
PARAMETER ESTIMATION IN MULTIVARIABLE STOCHASTIC DIFFERENCE EQUATIONS.
IEEE Trans. Autom. Contr., vol. AC-19 (1974), pp. 784-797.

-Khargonekar,P.P. , Sontag, E.D. (1982)
ON THE RELATION BETWEEN STABLE MATRIX FRACTION FACTORIZATIONS AND REGULABLE REALIZATIONS OF LINEAR SYSTEMS OVER RINGS.
IEEE Trans. Autom. Contr., vol. AC-27 (1982), nr. 3, pp. 627-638.

-De Keyser, R. (1979)
ON THE RELATIONSHIP BETWEEN SOME WELLKNOWN RECURSIVE IDENTIFICATION METHODS- THE RECURSIVE TALLY PRINCIPLE.
Prepr./Proc. 5-th IFAC Symp. Ident. and Syst. Param. Estim., Darmstadt, 1979, pp. 251-256.

-Kotta, U. (1979)
STRUCTURE AND PARAMETER ESTIMATION OF MULTIVARIABLE SYSTEMS USING EIGENVECTOR METHOD.
Prepr./Proc. 5-th IFAC Symp. Ident. and Syst. Param. Estim., Darmstadt, 1979, pp. 453-458.

-Lancaster,P.,Gohberg, I., Rodman, L. (1982)
MATRIX POLYNOMIALS.
Academic Press, New York, 1982.

-Ljung, L. Rissanen, J. (1976)
Prepr./Proc. 4-th IFAC Symp. Ident. and Syst. Param. Estim., Tbilisi, 1976, pp. 58-69.

-Luenberger, D.G. (1967)
CANONICAL FORMS FOR LINEAR MULTIVARIABLE SYSTEMS.
IEEE Trans. Autom. Contr., vol. AC-12 (1967), nr. 3, pp. 290-293.

-MacDuffee, C.C. (1956)
THEORY OF MATRICES.
Chelsea, New York, 1956.

-Maroulas,J., Barnett, S. (1978)
THE GREATEST COMMON DIVISOR OF GENERALIZED POLYNOMIALS AND POLYNOMIAL MATRICES.
Lin. Algebra and Applic., vol. 22 (1978), pp. 195-210.

-Mayne, D.Q. (1967)
ADVANCES IN COMPUTER CONTROL.
IEE Publications, London, 1967.

-Mayne, D.O. , Firoozan, F. (1982)
LINEAR IDENTIFICATION OF ARMA PROCESSES.
Automatica, vol. 18 (1982), nr. 4, pp. 461-466.

-Meertens, J.B. (1983)
Structural and parametric identification of MIMO systems according to Guidorzi. Further developments and applications.
Masters thesis, Dep. of El. Eng., Eindhoven Univ. of Techn.,1983.

-Patel, R.V. (1981)
COMPUTATION OF MATRIX FRACTION DESCRIPTIONS OF LINEAR TIME-INVARIANT SYSTEMS.
IEEE Trans. Autom. Contr., vol. AC-26 (1981), nr. 1, pp. 148-161.

-Peterka, V. , Smuk, K. (1969)
IFAC congress, Warsaw, 1969.

-Rosenbrock, H.H. (1970)
STATE-SPACE AND MULTIVARIABLE THEORY.
John Wiley, New York, 1970.

-Rowe, I.H. (1970)
A BOOTSTRAP METHOD FOR THE STATISTICAL ESTIMATION OF MODEL PARAMETERS.
Int. J. of Control, vol. 12 (1970), pp. 721-738.

-Salehi, S.V. (1978)
ON THE STATE-SPACE REALIZATIONS OF MFDS FOR MULTIVARIABLE SYSTEMS.
IEEE Trans. Autom. Contr., vol. AC-23 (1981), nr. 6, pp. 1054-1058.

-Sinha, S., Caines, P.E. (1977)
ON THE USE OF SHIFT REGISTER SEQUENCES AS INSTRUMENTAL VARIABLES FOR THE RESURSIVE IDENTIFICATION OF MULTIVARIABLE LINEAR SYSTEMS;
Int. J. Systems Sci., vol. 8, pp. 1041-1055.

-Stoica, P. (1981)
ON A PROCEDURE FOR TESTING THE ORDER OF TIME SERIES.
IEEE Trans. Autom. Contr., vol. AC-26 (1981), nr. 2, pp. 572-573.

-Stoica, P. Söderström, T. (1982)
IDENTIFICATION OF MULTIVARIABLE SYSTEMS USING INSTRUMENTAL VARIABLE METHODS.
Proc. 6-th IFAC Symp. Ident. and Syst. Param. Estim.,
Washington DC., 1982.

-Söderström, T. , Stoica, P. (1983)
INSTRUMENTAL VARIABLE METHODS FOR SYSTEMS IDENTIFICATION.
Lecture notes in Control and Information Sciences, nr. 57.
Springer Verlag, Berlin, 1983.

-Söderström, T. , Stoica, P. (1981)
COMPARISON OF SOME INSTRUMENTAL VARIABLE METHODS - CONSISTENCY AND ACCURACY ASPECTS.
Automatica, vol. 17 (1981), nr. 1, pp. 101-115.

-Tse, E., Weinert, H.L. (1975)
STRUCTURE DETERMINATION AND PARAMETER IDENTIFICATION FOR MULTIVARIABLE STOCHASTIC LINEAR SYSTEMS.
IEEE Trans. Autom. Contr., vol. AC-20 (1975), nr. 5, pp 603-613.

-Wellstead, P.E. (1978)
AN INSTRUMENTAL PRODUCT MOMENT TEST FOR MODEL ORDER ESTIMATION.
Automatica, vol. 14 (1978), nr. 1, pp. 89-91.

-Wolovich, W.A. (1974)
LINEAR MULTIVARIABLE SYSTEMS.
Springer Verlag, New York, 1974.

-Wong , K.Y. , Polak, E. (1967)
IDENTIFICATION OF LINEAR DISCRETE TIME SYSTEMS USING THE INSTRUMENTAL VARIABLE METHOD.
IEEE Trans. Autom. Contr., vol. AC-12 (1967), nr. 6, pp. 707-718.

-Young,P. (1976)
SOME OBSERVATIONS ON INSTRUMENTAL VARIABLE METHODS OF TIME SERIES ANALYSIS.
Intern. J. of Control, vol. 23 (1976), nr. 5, pp. 593-612.

-Young, P., Jakeman, A., McMurtrie, R. (1980)
AN INSTRUMENTAL VARIABLE METHOD FOR MODEL ORDER IDENTIFICATION.
Automatica, vol. 16 (1980), nr. 3 , pp. 281-294.

-Young, P., Shellswell, S.H., Neethling, C.G. (1971)
Report nr. CUED/ B- Control/ TR 16.
Dep. of Eng., Univ. of Cambridge, 1971.