Proceedings of 25th Conference
on Decision and Control
Athens, Greece • December 1986

SOME ASYMPTOTIC PROPERTIES OF MULTIVARIABLE MODELS IDENTIFIED BY EQUATION ERROR TECHNIQUES

Paul Van den Hof *)
Peter Janssen

Eindhoven University of Technology, Department of Electrical Engineering,
P.O. Box 513, 5600 MB Eindhoven, The Netherlands.
*) Now with: Delft University of Technology, Dep. of Mech. Engin.,
Mekelweg 2, 2628 CD Delft, The Netherlands.

Summary

In this paper some interesting properties are derived for simple equation error identification techniques, least squares and basic instrumental variable methods, applied to a class of linear time-invariant time-discrete multivariable models. The system at hand is not supposed to be contained in the chosen model set. Assuming that the input is unit variance white noise, it is shown that the Markov parameters of the system are estimated asymptotically unbiased over a certain interval around t = 0.

1. Introduction

In system identification literature there is a growing interest in identification methods that give reliable results in situations where the process at hand is not necessarily contained in the chosen model set. This aspect is considered to be a valuable robustness property[1]. Its importance is indicated by realizing that in many practical situations of system identification, a model will be required that is of restricted complexity, approximating the essential characteristics of the - possibly very complex - process, rather than a very sophisticated model that exactly models the process behaviour.

If the problem of system identification, or rather approximate modelling, is considered in this context, an important item now becomes: which criterion has to be used to approximate the original process, and which model set has to be chosen.
It has been recognized that these choices highly determine the performance of the model, when used for specific purposes such as simulation, prediction, (minimum variance) control etc.[2,3]. In many situations the performance of an identified model is judged upon its ability to simulate the process under consideration. However, output error methods, being most appropriate if the simulation behaviour of the model is concerned, are much more complex than equation error methods. It is therefore important to analyse the simulation behaviour of an equation error model. A frequency domain analysis of this aspect of approximate models, identified by prediction error methods, is given in [4,5]. By considering the Markov parameters of the identified model, we will focus on properties of the approximate model in the time domain.

Earlier work on this subject has been published in 1976 by Mullis and Roberts [6], who established a connection between results in model reduction and asymptotic results in least squares system identification. An extension to the multivariable case has been worked out by Inouye [7], but restricted to a full polynomial parametrization. The previous work will be extended to a general class of multivariable models, while also the basic instrumental variable method will be considered.

We will consider a discrete-time, linear, time-invariant process:

$$y(t) = H(z)\,u(t) + \xi(t) \qquad (1)$$

with $y(t)$, $\xi(t)\in\mathbb{R}^q$, $u(t)\in\mathbb{R}^p$, $t\in\mathbb{Z}$, $H(z)$ a $(q\times p)$-transfer matrix, z the forward shift operator, and $\xi(t)$ a random signal that is uncorrelated with the input signal $u(t)$; $u(t)$ and $\xi(t)$ are supposed to be jointly wide-sense stationary and ergodic.
The class of models that we use is a parametrized set of linear time-invariant discrete-time multivariable I/O models, given by the following general MFD (Matrix Fraction Description)-form:

$$P(z;\theta)y(t) = Q(z;\theta)u(t) + \varepsilon(t;\theta) \qquad (2)$$

where $P(z;\theta):=[p_{ij}(z)]_{q\times q}$ and $Q(z;\theta):=[q_{ij}(z)]_{q\times p}$ are $(q\times q)$, resp. $(q\times p)$-polynomial matrices; for ease of notation the explicit dependency of the polynomial entries on the parameter θ has been omitted.
The polynomials are specified by:

$$p_{ij}(z) = \delta_{ij}\, z^{\nu_j} - \alpha_{ij\nu_{ij}} z^{\nu_{ij}-1} - \ldots - \alpha_{ijr_{ij}} z^{r_{ij}-1} \qquad 1\leq i,j\leq q \qquad (3)$$

where $\delta_{ij} = 1$ for $i=j$ and $\delta_{ij} = 0$ for $i\neq j$.

$$q_{ij}(z) = \beta_{ij\mu_{ij}} z^{\mu_{ij}-1} + \ldots + \beta_{ijs_j} z^{s_j-1} \qquad 1\leq i\leq q,\ 1\leq j\leq p. \qquad (4)$$

The integer indices ν_i, ν_{ij}, μ_{ij}, r_{ij} and s_j determine the structure of the model set (2); Without loss of generality they are restricted to: $\nu_i \geq 0$, $\nu_{ij} \geq 0$, $\mu_{ij} \geq 0$, $r_{ij} \geq 1$, $s_j \geq 1$.
If $\nu_{ij} < r_{ij}$ then $p_{ij}(z) = \delta_{ij} z^{\nu_j}$, and if $\mu_{ij} < s_j$ then $q_{ij}(z) = 0$.

As a restriction on the model set we will require that

$$\nu_j \geq \nu_{ij} \quad \text{for} \quad 1 \leq i,j \leq q. \qquad (5)$$

This means that the leading column coefficient matrix of $P(z)$ is equal to the identity matrix.

In this paper we will consider the situation where the vector θ of unknown parameters consists of all the coefficients α and β occurring in the polynomial matrices $P(z;\theta)$ and $Q(z;\theta)$. The equation error

$$\varepsilon(t;\theta) = P(z;\theta)y(t) - Q(z;\theta)u(t) \qquad (6)$$

is dependent on the parameter vector θ, but is not parametrized itself.
The model set (2)-(5) is very general and encompasses most uniquely identifiable MFD-forms, currently used in the identification of multivariable systems. For

most forms it will follow that $r_{ij} = 1$, $s_j = 1$. Note that the model set is not necessarily restricted to causal models, and that it admits both one step ahead and more general k-step ahead prediction models.

Asymptotic least squares results are presented in section 2, and applied to a number of different parametrizations. The results are illustrated with a practical experiment in section 3. In section 4 the analysis is extended to the basic instrumental variable method. An overall discussion on the results follows in section 5.

2. Asymptotic results for least squares equation error minimization

Equation error methods are very popular in system identification. The main reason for this is the simplicity of the corresponding identification algorithm, due to the linearity-in-the-parameters of the model.

A common equation error method for obtaining an estimate of θ in (2) is the simple least-squares estimator, minimizing

$$V_N(\theta) = \frac{1}{N} \sum_{t=1}^{N} \varepsilon(t;\theta)^T \varepsilon(t;\theta) \tag{7}$$

with respect to θ. N denotes the number of data samples. Being interested only in asymptotic results, the asymptotic analogon of this problem will be considered, minimizing

$$V(\theta) = E\, \varepsilon(t;\theta)^T \varepsilon(t;\theta) \tag{8}$$

with respect to θ, under assumption of stationarity and ergodicity of the input and output signals. (E denotes the expectation-operator).

Remark 1
In our theoretical analysis we do not impose the condition that the parameter vector θ minimizing $V(\theta)$ is unique. With respect to the identification algorithm, however, one would like to use model sets which guarantee uniqueness. Examples of these will be given later on. For the analysis, there is no objection to the non-uniqueness of $\hat{\theta}$. ■

As a direct result of the linearity-in-the-parameters of $\varepsilon(t,\theta)$ the asymptotic least squares estimator $\hat{\theta}$ satisfies:

$$E\, \varepsilon_i(t;\hat{\theta})\, z^{\ell-1} y_j(t) = 0 \qquad r_{ij} \leq \ell \leq \nu_{ij},\quad 1 \leq i,j \leq q \quad \text{and} \tag{9}$$

$$E\, \varepsilon_i(t;\hat{\theta}) z^{\ell-1} u_j(t) = 0 \qquad s_j \leq \ell \leq \mu_{ij},\quad 1 \leq i \leq q,\ 1 \leq j \leq p \tag{10}$$

Now we define $\hat{P}(z) := P(z;\hat{\theta})$ (11)

and $\hat{Q}(z) := Q(z;\hat{\theta})$ (12)

The output signal $\hat{y}(t)$ of the estimated model, when excitated by the original input signal, is given by:

$$\hat{P}(z)\, \hat{y}(t) = \hat{Q}(z)\, u(t) \qquad -\infty < t < \infty \tag{13}$$

If we assume the input signal u(.) to be a zero mean stationary white noise sequence with unit variance (i.e. $E\, u(k)u^T(\ell) = \delta(k-\ell).I$), and we denote:

$$M(k) = Ey(t)u^T(t-k) \qquad -\infty < k < \infty \tag{14a}$$

$$\hat{M}(k) = E\hat{y}(t)u^T(t-k) \qquad -\infty < k < \infty \tag{14b}$$

then $M(k)$ is the k-th Markov parameter of the process and $\hat{M}(k)$ the k-th Markov parameter associated with the estimated model.

Proposition 1
Let $\hat{P}(z)$ and $\hat{Q}(z)$ be as defined above and let $\hat{\theta}$ be an asymptotic estimator fulfilling (10). Then the Markov parameters of the estimated model and the process satisfy:

$$\hat{P}_{i*}(z) M_{*j}(k) = \hat{Q}_{ij}(z)\delta(k) \qquad 1-\mu_{ij} \leq k \leq 1-s_j,\quad 1 \leq i \leq q,\ 1 \leq j \leq p \tag{15}$$

and

$$\hat{P}(z)\hat{M}(k) = \hat{Q}(z)\delta(k) I \qquad -\infty < k < \infty \tag{16}$$

where $\hat{P}_{i*}(z)$, resp. $M_{*j}(k)$ denote the i^{th} row vector of $\hat{P}(z)$, resp. the j^{th} column vector of $M(k)$.

Proof The result follows directly from (6), (10) and (13).

Note that the Markov parameters of the process satisfy the same relationship as the Markov parameters of the identified model, however on a restricted interval. It will be shown that, as a result of this, under some conditions, the two sequences of Markov parameters are equal on a restricted interval.

Theorem 1
Consider a multivariable I/O model as defined in (2)-(5). If this model is used for identifying a linear time invariant system by an equation error technique, fulfilling (10), if the input signal is zero mean, stationary unit variance white noise and if the number of data samples tends to infinity, then the Markov parameters $\hat{M}(t)$ of the identified model satisfy:

$$\hat{M}_{ij}(t) = M_{ij}(t) \quad \text{for } -\tilde{\mu}_j + \tilde{r}_i \leq t \leq 1-s_j + \nu_i \tag{17}$$

under the condition that the Markov parameters of the original process satisfy,

$$M_{ij}(t) = 0 \qquad \text{for } -\tilde{\mu}_j + \tilde{r}_i \leq t \leq \gamma_{ij} \tag{18}$$

where

$$\tilde{r}_j := \min_i r_{ij} \qquad 1 \leq j \leq q \tag{19}$$

$$\tilde{\mu}_j := \max_i \mu_{ij} \qquad 1 \leq j \leq p \tag{20}$$

$$\gamma_{ij} := \max\left[\nu_i - \mu_{ij},\ \max_\ell(\nu_{\ell i} - 1 - \mu_{\ell j})\right] \tag{21}$$

Proof The proof of this theorem is given in the appendix.

Theorem 1 has been stated in a general setting, due to the generality of the chosen model set (2)-(5). Its full implications will become clear if it is applied to specific parametrizations, with specific restrictions on the structure indices. If condition (18) is fulfilled, the result shows that apparently a Padé type of approximation is involved, where the length of the matching interval is determined by the chosen parametrization and the chosen structure indices of the model.

Note that by (21), γ_{ij} can be interpreted as the maximal difference between the degrees of $p_{\ell i}(z)$ and $q_{\ell j}(z)$ minus 1:

$$\gamma_{ij} = \max_\ell \left[\text{degr}\left(p_{\ell i}(z)\right) - \text{degr}\left(q_{\ell j}(z)\right)\right] - 1 \tag{22}$$

When modelling causal systems by MFD model sets the degrees of the polynomials $q_{ij}(z)$ are usually chosen equal to the i-th row degree of P(z):

$$\mu_{ij} - 1 = \max_\ell \left[\text{degr } p_{i\ell}(z)\right] \qquad 1 \leq j \leq p,\ 1 \leq i \leq q \tag{23}$$

This choice is motivated by the fact that for causal MFD's the row degrees of P(z) have to be greater than

	CANONICAL OBSERVABILITY FORM	PSEUDO-CANONICAL FORM	HERMITE FORM	DIAGONAL FORM	GENERALIZED PSEUDO-CANONICAL FORM
MFD Structure specification	$\nu_{ij}=\min(\nu_i+1,\nu_j), i>j$ $=\min(\nu_i,\nu_j), i<j$ $\mu_{ij} = \nu_i + 1$ $r_{ij} = 1;\ s_j = 1$	$\nu_{ij} = \nu_j$ $\mu_{ij} = \nu_m + 1$ if $\nu_i=\nu_m$ $\mu_{ij} = \nu_m$ if $\nu_i<\nu_m$ $r_{ij} = 1;\ s_j = 1$	$\nu_{ij} = 0,\ i<j$ $\nu_{ij} = \nu_j,\ i>j$ $\mu_{ij} = \max_{\ell<i}(\nu_i+1,\nu_\ell)$ $r_{ij} = 1;\ s_j = 1$	$\nu_{ij} = 0,\ i\neq j$ $\nu_{ii} = \nu_i$ $\mu_{ij} = \nu_i + 1$ $r_{ij} = 1;\ s_j = 1$	$\nu_{ij} = m_j + n_j - 1$ $\nu_i = \max_j(m_j+n_j-1)=:s$ $\mu_{ij} = s + 1$ $r_{ij} = m_j;\ s_j = 1$
Auxiliary variables	$\tilde{r}_j = 1;\ \tilde{\mu}_j = \nu_m+1$	$\tilde{r}_j = 1;\ \tilde{\mu}_j = \nu_m+1$	$\tilde{r}_j = 1;\ \tilde{\mu}_j = \nu_m+1$	$\tilde{r}_j = 1;\ \tilde{\mu}_j = \nu_m+1$	$\tilde{r}_j = m_j;\ \tilde{\mu}_j = s+1$
Theorem 1: $\hat{M}_{ij}(t)=M_{ij}(t)$	$-\nu_m \leq t \leq \nu_i$	$-\nu_m \leq t \leq \nu_i$	$-\nu_m \leq t \leq \nu_i$	$-\nu_m \leq t \leq \nu_i$	$-\nu_m + m_i - 1 \leq t \leq \nu_i$

$\nu_m := \max_i(\nu_i)$

Table -1- Structure specifications for a number of multivariable parametrizations in MFD-form; results for the asymptotic fit of Markov parameters.

or equal to the corresponding row degrees of $Q(z)$. As a result, the corresponding model sets will fulfil $\gamma_{ij} \leq -1$.
We can now present the following result as a direct consequence of theorem 1:

<u>Proposition 2</u>
Consider the situation as described in theorem 1. If the original system is causal and if $\gamma_{ij} \leq -1$ then

$$\hat{M}_{ij}(t) = M_{ij}(t) \quad \text{for} \quad t \leq 1 - s_j + \nu_i \qquad (24)$$
$$1 \leq i \leq q,\ 1 \leq j \leq p.$$

<u>Proof:</u> The result follows directly from theorem 1 and the observation that $\hat{M}_{ij}(t) = 0$ for $t \leq \nu_i - \tilde{\mu}_j$. ■

A remarkable result is that the estimated model will necessarily be causal too, notwithstanding the fact that the applied model set possibly admits non-causal models.
For most common MFD-models, $s_j = 1$ and the above result (24) can be formulated as

$$\hat{M}_{ij}(t) = M_{ij}(t)\ ,\quad t \leq \nu_i \qquad (25)$$
$$1 \leq i \leq q,\ 1 \leq j \leq p.$$

The asymptotic matching of Markov parameters, as given in theorem 1, finds its source in the correlation result (10). Bearing this in mind, the number of Markov parameter entries that is forced to be matched by the equation error method, will be equal to the number of β-parameters in the model. If the transfer matrix $[\hat{P}(z)]^{-1}\hat{Q}(z)$ of the model (2) is proper, the sequence of matched Markov parameter entries lies completely in the causal range ($t \geq 0$); if not, also a non-causal part will be involved.

It is illustrated in table 1 how the results of this section work out for a number of - commonly used - parametrizations of the original model set. The parametrizations will not be described in detail; for a thorough description the reader is referred to the literature. A general and up-to-date account on the use of identifiable parametrizations for multivariable linear systems is given in [8].

Table 1 shows the structure specification, and results of Theorem 1, for a number of MFD forms: the canonical observability form [9]; the pseudo-canonical (overlapping) form [10]; the Hermite form [11]; the diagonal form [12] and the generalized pseudo-canonical form [13].

Although the analysis and discussion has, until now, been based on MFD-model sets, defined in the forward shift-operator z, the obtained results can also be used for ARMAX model sets, which are defined in the backward shift operator z^{-1}. By transforming the ARMAX-model to a corresponding MFD-form, the results of this section can still be applied.
In Table 2, the results are listed for two commonly used ARMAX-model forms: the "prescribed maximum lag"-form $S(n_1,\dots,n_q,m_1,\dots,m_p)$ [14], and the full polynomial form $S(n,m)$ [15].
The asymptotic fit of Markov parameters for the full polynomial ARMAX form has been presented before by Mullis and Roberts [6] for the single input-single output (SISO) case, and by Inouye [7], for the multivariable (MIMO) case.

	PRESCRIBED MAXIMAL LAG FORM	FULL POLYNOMIAL FORM
MFD Structure specification	$S(n_1,..n_q,m_1..,m_p)$ $\nu_i = s := \max_{k,\ell}(n_k,m_\ell)$ $\nu_{ij} = \nu_j = s$ $\mu_{ij} = s + 1$ $r_{ij} = s-n_j+1;\ s_j = s-m_j+1$	$S(n,m)$ $\nu_i = s := \max(n,m)$ $\nu_{ij} = \nu_j = s$ $\mu_{ij} = s + 1$ $r_{ij} = s-n+1;\ s_j = s-m+1$
Auxiliary variables	$\tilde{r}_j = s-n_j+1;\ \hat{\mu}_j = s+1$	$\hat{r}_j = s-n+1;\ \tilde{\mu}_j = s+1$
Theorem 1: $\hat{M}_{ij}(t)=M_{ij}(t)$	$-n_i \leq t \leq m_j$	$-n \leq t \leq m$

Table -2- MFD structure specifications for two multivariable parametrizations in ARMAX-form; results for the asymptotic fit of Markov parameters.

3. A practical example

In order to illustrate the theoretical results of the previous section, some results will be shown of an equation error identification algorithm applied to a multivariable industrial process.
The process concerns the shaping part of a glass tube production process, where the output variables wall-thickness and diameter can be adjusted by manipulating two input variables:

- the pressure of the air through the tube and
- the pulling speed at the end of the tube.

A two input, two output model has been constructed by applying a least squares equation error technique to a model set with dimension 6, parametrised in pseudo-canonical overlapping form.
The process has been excitated with a sufficiently rich input signal, and 500 samples of input and output signals have been used for the identification.

Figures 1 and 2 show the Markov parameters of the estimated model for two transmittancies within the model, and for three sets of structure indices. The results for the estimated models are compared with Markov parameters that are directly estimated from input/output data.

Figure 1 shows the result of the transfer from input 1 (pressure) to output 2 (diameter). The fit of the start sequence of Markov parameters can be recognized, as well as an increasing length of the fit if the structure index corresponding to output 2 increases. A similar effect for output 1 is shown in Fig. 2, showing the results of the transfer from input 2 (pulling speed) to output 1 (wall-thickness).

The results of this experiment show how the choice of a set of structure indices can influence the performance of the approximate model.

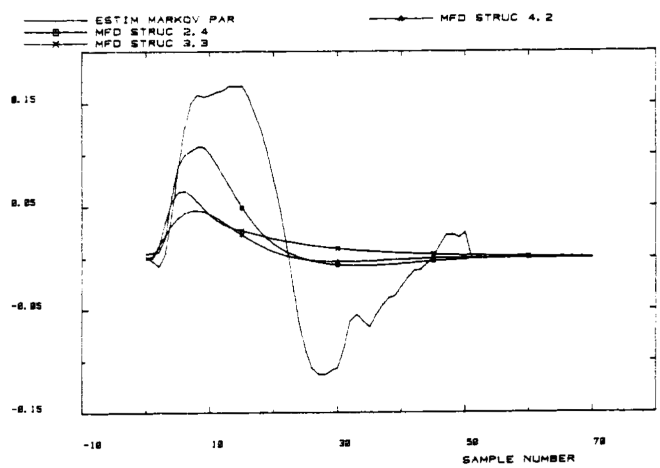

Fig. 1 Markov parameters $\hat{M}_{21}(k)$ of estimated MFD model, parametrized in overlapping form, for structures (2,4), (3,3) and (4,2); directly estimated Markov parameters.

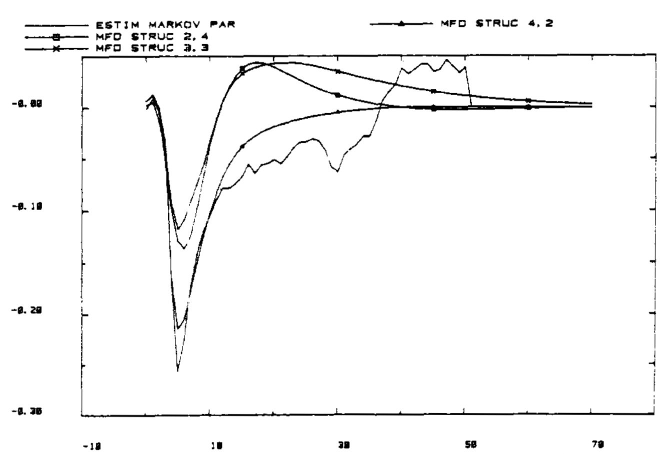

Fig. 2 Markov parameters $\hat{M}_{12}(k)$ of estimated MFD model, parametrized in overlapping form, for structures (2,4), (3,3) and (4,2); directly estimated Markov parameters.

4. Extension of the results for basic instrumental variable methods

In section 2 we presented some asymptotic results for the least-squares equation error method. Another useful technique for estimating the parameters in equation error models is formed by the instrumental variable methods. Many variants of this technique have been proposed (for an overview see [16]). In this discussion only the most elementary one, the basic IV-method, will be considered.
For this method we will establish similar results such as those presented in section 2.

Asymptotically, the basic IV-method for estimating θ amounts to the solution of the set of equations

$$E\, Z(t)\varepsilon(t;\theta) = 0 \qquad (26)$$

where $Z(t)$ is a $n_\theta xq$ matrix consisting of properly chosen instrumental variables, and n_θ is the number of parameters in the model. In our analysis it will be assumed that (26) indeed has a solution. Typical choices for $Z(t)$ are [16]:

$$Z(t) = \mathrm{diag}\left[z_1(t),\ldots,z_q(t)\right] \qquad (27)$$

where the length of vector $z_i(t)$ corresponds to the number of parameters in the i-th equation of (2), and $z_i(t)$ contains:

$$z^{\ell-1}\phi_{ij}(t),\ r_{ij} \leq \ell \leq \nu_{ij};\ 1 \leq j \leq q \qquad (28a)$$

and

$$z^{\ell-1}u_j(t),\ \ s_j \leq \ell \leq \mu_{ij};\ 1 \leq j \leq p \qquad (28b)$$

with $\phi_{ij}(t)$ a filtered or delayed input signal.

Because of (27), equation (26) now becomes:

$$E\, z_i(t)\ \varepsilon_i(t;\theta) = 0 \qquad 1 \leq i \leq q \qquad (29)$$

Substituting (28b) gives:

$$E\, z^{\ell-1}u_j(t)\ \varepsilon_i(t;\theta) = 0 \quad 1 \leq j \leq p,\ 1 \leq i \leq q,$$
$$s_j \leq \ell \leq \mu_{ij} \qquad (30)$$

which is exactly the same expression as (10).

Using (30) it is obvious that Proposition 1 and Theorem 1 will also hold for the basic IV-estimator (27) - (28).

These results can even be extended for a special choice of the basic IV-estimator as proposed in [17]. Using namely delayed input components $z^{x-1}u_j(t)$ with $x < s_j$ for all the IV-variables in (28a), we will obtain an IV-vector $z_i(t)$ consisting of the components:

$$z^{\ell-1}u_j(t) \qquad \lambda_{ij} \leq \ell \leq \mu_{ij};\ 1 \leq j \leq p \qquad (31)$$

Here λ_{ij} are integers such that $\lambda_{ij} \leq s_j$ for $1 \leq j \leq p$ and such that $\sum_{j=1}^{p} (\mu_{ij}-\lambda_{ij}+1)$ is equal to the number of parameters in the i-th equation of (2).

Using (31) as IV-vector, it is obvious that equation (30) will be valid on the extended interval $\lambda_{ij} \leq \mu_{ij}$. Therefore the relationship (15) in Proposition 1 will hold for $1-\mu_{ij} \leq k \leq 1-\lambda_{ij}$, while relationship (16) remains valid.

By using similar arguments as in the proof of Theorem 1, it can now be shown that Theorem 1 will hold true for this specific basic IV-estimator if we replace s_j in (17) by f_{ij}, where f_{ij} are integers which fulfill the following conditions:

$$f_{ij} \geq \lambda_{ij} \qquad 1 \leq i \leq q,\ 1 \leq j \leq p \qquad (32)$$

and

$$\min_{1 \le i,\, k \le q} \quad (\nu_k - \nu_{ik} + 1 + f_{ij} - f_{kj}) \ge 0 \quad 1 \le j \le p \tag{33}$$

It can easily be seen that the special choice

$$f_{ij} := \tilde{\lambda}_j := \max_{1 \le \ell \le q} \lambda_{\ell j} \quad 1 \le i \le q,\ 1 \le j \le p$$

satisfies (32)-(33).
This choice illustrates that the asymptotic matching of the Markov parameters of process and model will hold on a (possibly) larger interval than in the original results of Theorem 1 (namely $f_{ij} := \tilde{\lambda}_j \le s_j$ since $\lambda_{\ell j} \le s_j$ for $1 \le \ell \le q$).

The consequences of this general result will be illustrated by a simple example:

Example 4.1:
Consider a causal SISO-system. Suppose that we use the model set

$$y(t) = -\sum_{i=1}^{n} a_i y(t-i) + \sum_{j=0}^{m} \beta_j\, u(t-j) + \varepsilon(t;\theta) \tag{35}$$

The instrumental variable vector is chosen to consist of delayed input signals $u(t-\ell)$ $(0 \le \ell \le m+n)$.
Application of the foregoing results gives that:

$$\hat{M}(t) = M(t) \qquad \text{for } 0 \le t \le m+n \tag{36}$$

All the parameters of the model are used to match the Markov parameters. Therefore the model is completely determined by the first m+n+1 Markov parameters of the system, a result which is, if fact, equivalent to a Padé approximation. As a consequence of this, use of this IV-method can result in unstable models, a phenomenon which is inherent to Padé approximations. ■

5. Discussion

The analysis in this paper has been motivated by the consideration that in many practical situations the performance of an identified model is assessed by its ability to simulate the given process.
In the asymptotic equation error results, all β-parameters are, in fact, determined in such a way that the Markov parameters of the process are matched over a certain range. The α-parameters determine how the start-sequence of Markov parameters of the model is extended to infinity. As a result, the choice of a parametrization can have a severe effect on the emphasis that is adjusted to the start-sequence c.q. the extension sequence of Markov parameters; the specific set of structure indices determines the length of the matching interval for the various input-output transmittances within the model.

It has been illustrated in [3] for some SISO examples that the asymptotic unbiasedness of the start-sequence of Markov parameters causes the equation error model to generate a bad sequence of Markov parameters in the extension to infinity. This effect especially occurs when the system impulse response is small in the start-sequence, and increases outside this interval[18], a very common situation in practice (e.g. because of time delays). A similar result will probably hold for MIMO systems.

We have to note that the results of section 3 cannot simply be generalized to situations of non-white input signals. Although the quantitites

$$\Phi_{yu}(t) \quad \text{and} \quad \Phi_{\hat{y}u}(t)$$

satisfy the same relationship on a restricted interval (see Proposition 1), they will not match in general, since the associated initial conditions for the recursive relations are different in the case of non-white input signals.

The results in this paper, presented in terms of Markov parameters can of course also be formulated in terms of step responses.

6. Conclusions

For a general class of linear multivariable models, asymptotic properties are derived for the Markov parameters of approximate models, when identified by equation error identification methods.
The results of this paper are valid for general linear time-invariant systems corrupted by output noise that is not correlated with the input signal. Under the condition of white input noise, it has been shown that the Markov parameters of the system are estimated asymptotically unbiased over a certain interval around t=0. The position of the interval is dictated by the chosen structure indices of the model. Moreover it has been shown that for causal systems the identified model is asymptotically causal, notwithstanding the fact that the applied model set might contain non-causal models.
The results have been obtained for a very general class of linear models, having the property of linearity in the parameters. This class of models covers all commonly used MFD (Matrix Fraction Description)-forms and ARMAX forms.
The theoretical analysis has been illustrated by results of a practical identification experiment.
For basic IV techniques, it has been shown that the method of Wouters [17] is asymptotically equivalent to Padé approximation in model reduction.

Appendix

Proof of Theorem 1.
From (16) we obtain:

$$z^{\nu_i}\hat{M}_{ij}(t) = -\hat{P}^*_{i*}(z)\hat{M}_{*j}(t) + \hat{Q}_{ij}(z)\delta(t) \qquad -\infty < t < \infty \tag{A-1}$$

where $P^*(z) = P(z) - \mathrm{diag}\left[z^{\nu_1}, z^{\nu_2}, \ldots, z^{\nu_q}\right]$.
Since $\hat{Q}_{ij}(z)\ \delta(t) = 0$ for $t \le -\tilde{\mu}_j$ it follows directly that $z^{\nu_i}\hat{M}_{ij}(t) = 0$ for $t \le -\tilde{\mu}_j$. Combining this with condition (18) results in:

$$z^{\nu_i}\hat{M}_{ij}(t) = z^{\nu_i}M_{ij}(t) \quad \text{for } -\tilde{\mu}_j + \tilde{r}_i - \nu_i \le t \le -\tilde{\mu}_j \tag{A-2}$$

In order to prove the remaining part of (17) we observe with (15) that

$$z^{\nu_i}M_{ij}(t) = -\hat{P}^*_{i*}(z)M_{*j}(t) + \hat{Q}_{ij}(z)\delta(t) \qquad 1-\mu_{ij} \le t \le 1-s_j \tag{A-3}$$

Under condition (18) this relation will also hold for $1-\tilde{\mu}_j \le t \le -\mu_{ij}$.
Using (A-1), (A-2) and (A-3) on the extended interval $1-\tilde{\mu}_j \le t \le 1-s_j$ it follows by induction that

$$z^{\nu_i}\hat{M}_{ij}(t) = z^{\nu_i}M_{ij}(t), \qquad 1-\tilde{\mu}_j \le t \le 1-s_j \tag{A-4}$$

In order to proof this, consider that (A-4) holds for $t = 1 - \tilde{\mu}_j$.
For the inductive continuation we assume that
$z^{\nu_i}\hat{M}_{ij}(t) = z^{\nu_i}M_{ij}(t)$ for $1 - \tilde{\mu}_j \le t \le s^*$,

where $1 - \tilde{\mu}_j \leq s^* \leq - s_j$. As a result of (A-3):

$$z^{\nu_i} M_{ij}(s^*+1) = -P^*_{i*}(z)M_{*j}(s^*+1) + Q_{ij}(z)\delta(s^*+1) \tag{A-5}$$

Using the inductive assumption and (A-2) it follows that

$$P^*_{i*}(z)M_{*j}(s^*+1)=P^*_{i*}(z)\hat{M}_{*j}(s^*+1) \tag{A-6}$$

Combination of (A-1), (A-5) and (A-6) finally gives

$z^{\nu_i} \hat{M}_{ij}(s^{\bullet}+1) = z^{\nu_i} M_{ij}(s^*+1)$, providing the inductive argument leading to (A-4).

References

[1] L. Ljung and T. Söderström, "Theory and practice of recursive identification", Cambridge, Mass.: MIT Press. MIT Press series in signal processing, optimization, and control, Vol. 4, 1983.

[2] M. Gevers and G. Bastin, "What does system identification have to offer?" In: Identification and System Parameter Estimation. Proc. 6th IFAC Symp., Washington, D.C., 7-11 June 1982. Ed. by G.A. Bekey and G.N. Saridis. Oxford: Pergamon, 1983. P. 77-84.

[3] A.A.H. Damen, Y. Tomita and P.M.J. Van den Hof, "Equation error versus output error methods in system identification", Report ER 85/06 Eindhoven Univ. Technology, Dept. Electr. Eng., 1985.

[4] B. Wahlberg and L. Ljung, "Design variables for bias distribution in transfer function estimation," IEEE Trans. Autom. Control, vol. AC-31, pp. 134-144, 1986.

[5] L. Ljung, "Asymptotic variance expression for identified black-box transfer function models", IEEE Trans. Autom. Control, Vol. AC-30, pp. 834-844, 1985.

[6] C.T. Mullis and R.A. Roberts, "The use of second order information in the approximation of discrete time linear systems," IEEE Trans. Acoust., Speech and Signal Process., vol. ASSP-24, pp. 226-238, 1976.

[7] Y. Inouye, "Approximation of multivariable linear systems with impulse response and autocorrelation sequences," Automatica, vol. 19, pp. 265-277, 1983.

[8] M. Gevers and V. Wertz, "Techniques for the selection of identifiable parametrizations for multivariable linear systems", To appear in: Control and Dynamic Systems. Advances in theory and applications. Vol. 24. Ed. by C.T. Leondes, New York: Academic Press, 1985.

[9] R.P. Guidorzi, "Invariants and canonical forms for systems; structural and parametric identification", Automatica, Vol. 17, pp. 117-133, 1981.

[10] G.O. Correa and K. Glover, "Pseudo-canonical forms, identifiable parametrizations and simple parameter estimation for linear multivariable system: input-output models", Automatica, Vol. 20, pp. 429-442, 1984.

[11] B.W. Dickinson, T. Kailath and M. Morf, "Canonical matrix fraction and state space descriptions for deterministic and stochastic linear systems", IEEE Trans. Autom. Control, Vol. AC-19, pp. 656-666, 1974.

[12] A. Gauthier and I.D. Landau, "On the recursive identification of multi-input, multi-ouput systems", Automatica, Vol. 14, pp. 609-614, 1978.

[13] M. Gevers and A.C. Tsoi, "Structural identification of linear multivariable systems using overlapping forms: a new parametrization", Int. J. Control, Vol. 40, pp. 971-987, 1984.

[14] M. Deistler, "The properties of the parametrization of ARMAX systems and their relevance for structural estimation", Econometrica, Vol. 51, pp. 1187-1207, 1983.

[15] E.J. Hannan, "The identification of vector mixed autoregressive moving-average systems", Biometrika, Vol. 56, pp. 223-225, 1969.

[16] T. Söderström and P.G. Stoica, "Instrumental variable methods for system identification," Berlin: Springer, Lecture notes in control and information sciences, vol. 57, 1983.

[17] W.R. Wouters, "On-line identification in an unknown stochastic environment", IEEE Trans. Syst., Man and Cybern., Vol. SMC-2, pp. 666-668, 1972.

[18] A.A.H. Damen, R.L. Moses, P.M.J. Van den Hof and Y. Tomita, "Bounds on the simulation power of equation error estimates", To be presented at this Conference, 1986.