INT. J. CONTROL, 1985, VOL. 41, NO. 3, 589–613

On multivariable partial realization†

A. A. H. DAMEN‡, R. P. GUIDORZI§, A. K. HAJDASINSKI|| and P. M. J. VAN DEN HOF‡

In this paper some aspects of the partial realization problem for multivariable formal power series are considered. The dynamic structure of the systems underlying such series is analysed on the basis of the results of Kalman, and a canonical realization algorithm is described based on this concept of structure. Special attention is paid to the aspect of uniqueness of the realization; necessary and sufficient conditions are given for the uniqueness of the extension sequence. A specific class of power series is defined for which a very simple uniqueness criterion is derived. Generally, noise-disturbed data sequences belong to this specific class; it is shown that the unique partial realization for such a noise-disturbed power series may serve as an intermediate step to arrive at an approximate realization of limited order.

1. Introduction

The comparatively young field of system theory has already seen many fashions that sometimes have emphasized specific problems and sometimes specific methodologies. Moreover the large advances, undoubtedly achieved, have sometimes given the impression of the attainment of every possible degree of knowledge in some sectors. This feeling, which periodically appears in every field, has almost invariably been proved vacuous by subsequent results.

The fundamental importance of the realization problem has, in fact, never been questioned in system theory. However, it seems to be a common opinion regarding the realization of linear time-invariant systems that almost everything has been said about this problem after the work of Kalman (1960) and subsequent developments (Kalman 1963, 1968, 1971, Ho and Kalman 1966, Tether 1970, Ackermann and Bucy 1971, Silverman 1971, Roman and Bullock 1975, Ledwich and Moore 1976, Brockett 1978).

Yet new insights into the partial realization of formal power series can be found in a recent paper by Kalman (1979) where many implications of this problem are investigated in depth. That paper treats the scalar case; the multivariable case requires, as usual, more complex technicalities and also exhibits some peculiar aspects essentially related to the concept of multivariable structure which is not relevant for the scalar case (see, for example, Zazworsky *et al.* (1979), Bosgra (1983) and Bistritz (1983)).

Received 23 August 1983; received in revised form 11 June 1984.

† This research has been supported by Ministero della Pubblica Istruzione, Roma, Italia, by Samenwerkingsorgaan KHT-THE and by Eindhoven University of Technology, The Netherlands.

‡ Department of Electrical Engineering, Eindhoven University of Technology, Eindhoven, The Netherlands.

§ Dipartimento di Elettronica Informatica e Sistemistica, University of Bologna, Bologna, Italy.

|| Samenwerkingsorgaan KHT-THE, Catholic University Tilburg, Tilburg, The Netherlands.

Nevertheless Kalman's approach can easily be generalized for the multivariable case. After a short restatement of the problem this will be done in this paper.

A multivariable formal power series is given by an infinite sequence of $q \times p$ matrices of generally real entries

$$\{M(i)\} = M(1), M(2), \ldots \tag{1}$$

In fact there are no restrictions on the field to which these coefficient matrices belong. In control engineering it is usual to think of such sequences as a matrix power series in the delay operator z^{-1}:

$$M(1)z^{-1} + M(2)z^{-2} + \ldots$$

This explains the names 'power series' and 'coefficient matrices'. In the practice of system engineering the coefficients may then be, for example, the Markov parameters of a process or the covariances of a set of stochastic signals.

The dimension n of such a sequence is said to be finite if and only if it can be represented by

$$M(i) = HF^{i-1}G, \quad i \in N^+ \tag{2}$$

where F, G, H are matrices of sizes $n \times n$, $n \times p$, $q \times n$ and n is the finite dimension of this triple. N^+ is the set of positive integers. In the case that we are dealing with, Markov parameters, this means that F, G, H, are coefficient matrices of a state-space representation for the process concerned of dimension n with p inputs and q outputs.

A triple $\{F, G, H\}$ is called a *minimal realization* if and only if the dimension n is minimal among all possible $\{F', G', H'\}$ fulfilling (2). A necessary and sufficient condition for the existence of a realization of dimension n is given by the *realizability criterion*† (see, for example, Tether (1970)):

There exists an integer N_0 such that

$$\rho H[N_0, N_0] = \rho H[N_0 + i, N_0 + j] = n, \quad \forall i, j \in N^+ \tag{3}$$

where $\rho H[a, b]$ is the rank of the block Hankel matrix of block size $a \times b$, constructed from the sequence of matrices $M(i)$.

If we are dealing with a sequence of finite length L:

$$\{M(i)\}_L = M(1), M(2), \ldots, M(L), \quad L \in N^+ \tag{4}$$

we can always find a triple $\{F, G, H\}$ of finite size such that

$$M(i) = HF^{i-1}G, \quad i = 1, 2, \ldots, L \tag{5}$$

and consequently the dimension n of this sequence will always be finite. Then such a triple $\{F, G, H\}$ is a partial realization of the sequence $\{M(i)\}_L$. The attribute 'partial' is being used since just a finite sequence has been fitted by a realization. Such a finite sequence may be part of an infinite sequence which is not fully available.

An important aspect in minimal partial realization is the question of whether the extension is unambiguously determined by the finite sequence if a minimal complexity (i.e. dimension n) of the realization is required. This characteristic is confined in the property of uniqueness (of the extension). If the extension is not unique, this implies

† In Silverman (1971) an *equivalent* criterion is proposed where i is fixed to 1. This could be done by taking advantage of the particular structure of a block Hankel matrix.

that the partially available sequence does not provide enough information about the underlying process. On the other hand, if the extension is unique this does not guarantee that the partial realization contains all information about the underlying process. In both cases, the partial realization will incorporate all information contained in the sequence of finite length $\{M(i)\}_L$.

So the crucial questions are what the minimal value of the dimension n is and whether the continuation $M(i)=HF^{i-1}G$ for $i>L$ is unique for all triples with minimal dimension satisfying (5).

Following the definition of Kalman (1979) we can state the problem more formally.

Definition 1. *Minimal partial realization problem* (*MPR*)

(*a*) Find all exact realizations of sequence $\{M(i)\}_L$ whose dimension n is minimal over the family of all possible realizations for this sequence.

(*b*) Give the necessary and sufficient conditions under which there is only one minimal partial realization ('uniqueness' in the sense of the extension of $\{M(i)\}_L$, which forms an equivalence class modulo the choice of basis in the state space).

(*c*) In the case of non-uniqueness, parametrize all minimal partial realizations.

In § 2 we will elaborate a partial realization for multivariable sequences. The multi-companion forms appear to be closely related to this analysis of the minimal partial realization problem (MPR). Subsequently the MPR in a multi-companion output form will be formally stated and elucidated with an example in § 3. In § 4 the conditions for the uniqueness will be analysed and it will be shown that in the generic case such conditions will take a very simple form. Finally, in § 5 some reflections are made on the approximate partial realization problem. In connection with the MPR a straightforward and proper definition for this problem will be proposed.

2. Minimal partial realization

It is known (Tether 1970, Kalman 1971) that the minimal dimension of a finite matrix sequence can be found by means of the rank of the partial behaviour matrix, which we will now define.

The *partial behaviour matrix* associated to the finite sequence $\{M(i)\}_L$ is the following (block) Hankel matrix:

$$B_L = B(M(1), \ldots, M(L)) = \begin{bmatrix} M(1) & M(2) & \cdots & M(L) \\ M(2) & M(3) & \cdots & ? \\ \vdots & & ? & \vdots \\ M(L-1) & & ? & \vdots \\ M(L) & ? & \cdots & ? \end{bmatrix} \tag{6}$$

where the elements denoted by '?' correspond to values of the sequence which are not part of the data given for the partial realization problem. The rank of this partial behaviour matrix is defined as the minimal possible rank obtained by proper choice of the elements indicated by the question marks.

In detail we may develop the actual determination of the rank by the following procedure:

2.1. *Rank of a partial behaviour matrix*

Examine the first p columns of B_L (which belong to the vector space R^{qL}). If they are null the rank of B_L is zero and no other operations are required. Otherwise, if n_1 independent columns have been found, consider the second group of p columns of B_L and check if it contains vectors linearly independent of those of the first group (we are now in $R^{q(L-1)}$ since the last q entries in the vectors of the second group are not given). If only dependent vectors are found *the procedure ends*, otherwise, if n_2 independent vectors have been found, the third and following groups are considered until no independent vectors are found or until the whole matrix has been examined. Now the rank of the partial matrix B_L is defined as $n = \sum_i n_i$.

The rank of a partial matrix as defined contains the definition of the rank of a complete (ordinary) matrix. Moreover, rank B_L cannot be decreased by adding any element to B_L and is therefore *a lower bound on rank* B_∞ for any extension of the given partial sequence.

Remark 1

If rank $B_L = 0$ the partial realization problem is trivial since the dimension is zero (all entries are zero). There exists no *minimal* realization. When $n = \text{rank } B_L > 0$ it is well known that n is the minimal dimension for the possible partial realizations of the given sequence.

Remark 2

The rank of a partial behaviour matrix can be indifferently computed (just as the rank of a complete matrix) checking the independent columns (as in the given definition) or checking the independent rows.

Remark 3

Only in exceptional cases will the partial behaviour matrix have full rank. For example, it may happen for a single delay line of L samples, as can easily be checked:

$$B_L = \overbrace{\begin{bmatrix} 0 & 0 & . & . & 0 & 1 \\ 0 & 0 & . & 0 & 1 & ? \\ . & . & 0 & 1 & ? & . \\ . & 0 & 1 & ? & . & . \\ 0 & 1 & ? & . & . & ? \\ 1 & ? & . & . & . & ? \end{bmatrix}}^{L} \Bigg\updownarrow L \qquad \text{rank } B_L = L$$

In the generic case however, where the elements of the formal power series show no interdependence, all submatrices of B_L with known elements will have full rank, contrary to the example of the delay line. Consequently in the generic case the rank of the partial behaviour matrix will be equal to the dimension of the biggest square submatrix of known elements. This is substantially smaller than the maximal rank L. We will elaborate on this in § 4.

As we are dealing with the multivariable case, the aspect of model structure becomes important when analysing a data sequence, in contrast to the scalar case. The input structure (respectively, output structure) is reflected in the status of columns (respectively, rows) in the partial behaviour matrix. We will now analyse this topic in more detail.

2.2. *Structure of a partial behaviour matrix*

Consider the partial behaviour matrix (6) and apply the rank determination procedure to its columns. Denote with μ_1 the number of previously found independent vectors (if any) appearing in the first position of the first, second, ..., and Lth block column considered. Similarly denote with $\mu_2, \ldots, \mu_p$ the numbers of independent vectors which have been found in the second, ..., pth position of the same block columns. The integers $(\mu_1, \ldots, \mu_p)$ will be called 'control invariants' of B_L. Now apply the rank determination procedure to the rows of B_L and define, similarly, the integers $(\nu_1, \ldots, \nu_q)$ which will be called 'observation invariants' of B_L. The two sets of integers $(\mu_1, \ldots, \mu_p)$ and $(\nu_1, \ldots, \nu_q)$ define the structure of B_L. Obviously

$$n = \operatorname{rank} B_L = \sum_{i=1}^{p} \mu_i = \sum_{i=1}^{q} \nu_i$$

Remark 1

It is well known that the integers $(\mu_1, \ldots, \mu_p)$ and $(\nu_1, \ldots, \nu_q)$ are the control and observation invariants of every minimal state-space realization (F, G, H) of the given formal power series (Kronecker invariants).

Remark 2

It is well known that all μ_i independent vectors selected from the ith position within the different block columns in B_L belong to adjacent block columns 1 to μ_i. Thus, if, for instance, $\mu_2 = 2$, the independent vectors selected from the second position in the various block columns come from the first and second block column.

Example 1

Consider the matrices

$$M(1) = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad M(2) = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$$

then

$$B_2 = \begin{array}{cc} \left[\begin{array}{cccc} 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 0 & ? & ? \\ 1 & 1 & ? & ? \end{array}\right] & \begin{array}{c} \ast \\ \ast \\ \circ \\ \ast \end{array} \\ \begin{array}{cccc} \ast & \ast & \ast & \circ \end{array} & \end{array} \tag{7}$$

The application of the rank determination procedure to the columns of B_2 shows that in the first block column two independent vectors (in R^4) are present. When the second block column is considered only its first vector is linearly independent of the preceding ones (now in R^2). In (7) the independent and dependent vectors are denoted with the symbols '$*$' and '$\circ$', respectively. The determined rank is thus $n=3$. The same procedure applied to the rows of B_2 shows both rows of the first block row and the second row of the second block row to be independent. Again, the determined rank is $n=3$. The control invariants of B_2 are given by $\mu_1=2$ and $\mu_2=1$ while its observation invariants are given by $\nu_1=1$ and $\nu_2=2$.

2.3. *Additional tools for structure analysis*

Next the influence of the addition of new samples $M(L+1)$, $M(L+2)$, ... to available ones $M(1)$, $M(2)$, ..., $M(L)$ of a sequence can be analysed as a means to solve the MPR problem. For this purpose the following lemma and corollaries will play a fundamental role. The first lemma is an extended version of the lemmas given by Kalman (1979) and Tether (1970) and presented here for the general multivariable situation. The proof is given in Appendix 1.

Main lemma

Consider the partial matrix

$$\Omega(?)=\begin{bmatrix} A & \vdots & B \\ \cdots & \vdots & \cdots \\ C & \vdots & ? \end{bmatrix} \quad \text{where } \rho[A \;\vdots\; B]=\rho\begin{bmatrix} A \\ \cdots \\ C \end{bmatrix}=\rho[A] \tag{8}$$

with A, B and C of proper dimensions. There exists one and only one matrix D, which satisfies the condition

$$\rho\begin{bmatrix} A & \vdots & B \\ \cdots & \vdots & \cdots \\ C & \vdots & D \end{bmatrix}=\rho[A] \tag{9}$$

This holds for any size of matrix D, so in particular it is true when D is a scalar d. Then B and C are a column vector $\mathbf{b}$ and a row vector $\mathbf{c}^{\mathrm{T}}$, respectively.

Corollary 1

Consider a partial matrix $\Omega(?)$ with

$$\Omega(?)=\begin{bmatrix} A & \vdots & \mathbf{b} \\ \cdots & \vdots & \cdots \\ \mathbf{c}^{\mathrm{T}} & \vdots & ? \end{bmatrix}$$

(1) $\rho[\Omega(d)]=\rho[\Omega(?)]$ *if and only if*

(*a*) $\rho[A \;\vdots\; \mathbf{b}]>\rho[A]$,

or (*b*) $\rho\begin{bmatrix} A \\ \cdots \\ \mathbf{c}^{\mathrm{T}} \end{bmatrix}>\rho[A]$,

or (c) $\rho[A \;\vdots\; \mathbf{b}]=\rho\begin{bmatrix} A \\ \cdots \\ \mathbf{c}^{\mathrm{T}} \end{bmatrix}=\rho[A]$ and d is given the unique singular value according to the main lemma.

(2) $\rho[\Omega(d)] = \rho[\Omega(?)] + 1$ *if and only if*

$$\rho[A \;\vdots\; \mathbf{b}] = \rho\begin{bmatrix} A \\ \cdots \\ \mathbf{c}^{\mathrm{T}} \end{bmatrix} = \rho[A]$$ and d is not given the unique singular value according to the main lemma.

The proof is by inspection.

In other words, substitution of d in the partial matrix $\Omega(?)$ can only increase the rank of Ω if d appears in both a dependent row and a dependent column, and d is not given the unique singular value defined by the main lemma.

When analysing the influence of individual entries $M_{ij}(L+1)$ on the rank of B_{L+1}, one is confronted with a situation similar to Corollary 1. By inspection and use of Corollary 1 the following corollary can be verified.

Corollary 2

Given a partial behaviour matrix B_{L+1} where at least the elements $M(1), M(2), \ldots, M(L), M_{kl}(L+1)$ for $k=1, \ldots, i$; $l=1, \ldots, j$ and $(k, l) \neq (i, j)$ have been specified. Insertion of entry $M_{ij}(L+1)$ into B_{L+1} can affect the rank of B_{L+1} *if and only if* $M_{ij}(L+1)$ occurs in B_{L+1} at a crossing of a dependent row and a dependent column.

The main lemma and both corollaries will be used in analysing the partial behaviour matrix in detail, and in studying the particular influence of each entry in a sample $M(L+1)$ on the structural invariants of the system underlying $\{M(i)\}_L$. These entries may cause jumps on the structural invariants ν_i and μ_j, as will be shown.

2.4. *Insertion of a single new entry of $M(L+1)$ in B_{L+1}*

Consider the insertion of the (i,j)th entry of $M(L+1)$ that was unspecified before. To refer to a completely specified scenario it will be assumed that in each column and each row where this unknown entry appears, all entries up to the entry (i,j) under study have been specified before, as stated in Corollary 2. The control and observation invariants of B_{L+1} before the insertion of the considered entry are denoted by $(\mu_1, \ldots, \mu_p)$ and $(\nu_1, \ldots, \nu_q)$.

Let us first study the effect with the help of Example 1. For entry (1, 1) of Markov parameter $M(3)$ we deal with the situation:

$$B_3 = \begin{array}{c} \\ * \\ * \\ \circ \\ * \\ \circ \\ \circ \end{array}\!\!\begin{array}{c} \begin{array}{cccccc} * & * & * & \circ & \circ & \circ \end{array} \\ \begin{bmatrix} 1 & 0 & 1 & 0 & * & ? \\ 0 & 0 & 1 & 1 & ? & ? \\ 1 & 0 & * & ? & \cdot & \\ 1 & 1 & ? & ? & \cdot & \\ * & ? & \cdot & \cdot & ? & ? \\ ? & ? & & & ? & ? \end{bmatrix} \end{array} \qquad \begin{array}{ll} \nu_1 = 1, & \mu_1 = 2 \\ \nu_2 = 2, & \mu_2 = 1 \end{array}$$

Applying Corollary 2 to this partial behaviour matrix, noting that for all three positions of $M_{11}(3)$ in B_3 the entry is placed in an independent row or an independent column, it follows that no value of $M_{11}(3)$ exists that can affect the structure of B_3. Therefore this entry can freely be chosen and is not fixed by $M(1)$, $M(2)$ and the requirement of an extension with minimal dimension. $M_{11}(3)$ can be considered as a free parameter. Let us indicate this with '$\times$' and continue to study $M_{12}(3)$:

$$
B_3 = \begin{array}{c} \\ * \\ * \\ \circ \\ * \\ \circ \\ \circ \end{array}
\begin{array}{c} \begin{array}{cccccc} * & * & * & \circ & \circ & \circ \end{array} \\
\left[\begin{array}{cccccc}
1 & 0 & 1 & 0 & \times & * \\
0 & 0 & 1 & 1 & ? & ? \\
1 & 0 & \times & * & & \cdot \\
1 & 1 & ? & ? & & \cdot \\
\times & * & \cdot & \cdot & ? & ? \\
? & ? & & & ? & ?
\end{array}\right] \end{array}
\qquad
\begin{array}{ll} \nu_1 = 1, & \mu_1 = 2 \\ \nu_2 = 2, & \mu_2 = 1 \end{array}
$$

In this partial matrix at position (3, 4) a dependent row and column meet at the unknown entry; therefore by Corollary 1 this entry is uniquely determined if one is to avoid affecting the structure. If $M_{12}(3)$ is inserted and is not equal to this unique singular value, the third row and the fourth column of B_3 will become independent and the structure indices ν_1 and μ_2, and rank B_3 will increase by 1. Similarly it can be derived that both entries $M_{21}(3)$ and $M_{22}(3)$ cannot affect the structure.

It has to be noted that an entry of Markov parameter $M(L+1)$ appears $L+1$ times in the partial behaviour matrix B_{L+1}. Several times this entry may appear on intersections of dependent rows and columns. On each crossing then, a unique value is specified by the main lemma that will not affect the structure of B_{L+1}†. These values are the same for all different positions in B_{L+1}. This can be proved by contradiction, inspecting dependence and independence of rows and columns in B_{L+1}.

The results pointed out in the example above can be extended to the general case. The entry $M_{ij}(L+1)$ appears $L+1$ times in the partial behaviour matrix. Suppose that among these positions there are l positions at crossings of dependent rows and dependent columns. If $M_{ij}(L+1)$ is given a value conflicting the main lemma, then with Corollary 2 the l mentioned positions in B_{L+1} will change dependent rows/columns into independent rows/columns. This means that after insertion of this value of $M_{ij}(L+1)$, ν_i and μ_j will jump to $\nu_i + l$ and $\mu_j + l$, respectively. As a result the minimal dimension n of the partial realization, i.e. the rank of B_{L+1}, will also increase by l.

After this analysis of the influence of new entries, we may formally state the result as follows. Consider two sets of integers

$$S_c^j = \{\mu_j p + j, (\mu_j + 1)p + j, \ldots, Lp + j\} \tag{10}$$

$$S_r^i = \{\nu_i q + i, (\nu_i + 1)q + i, \ldots, Lq + i\} \tag{11}$$

(S_c^j indicating the dependent columns in block position j and S_r^i indicating the dependent rows in block position i) and the set of ordered pairs of integers given by

$$S_{ij}^{(L+1)} = \{(i, Lp + j), (q + i, (L-1)p + j), \ldots, (Lq + i, j)\} \tag{12}$$

i.e. the positions of the considered entry $M_{ij}(L+1)$ in the partial behaviour matrix.

† Formally the main lemma cannot be applied in the case this situation occurs in the first column or the first row of B_{L+1}. In that case the unique value is given by zero, as can easily be verified.

Result

The insertion of $M_{ij}(L+1)$ in B_{L+1} does not modify the integers ν_k for $k \neq i$ and μ_h for $h \neq j$. The jump of ν_i and μ_j is given by zero or by the number of elements of the set $Q_{ij}^{(L+1)} = S_{ij}^{(L+1)} \cap (S_r^i \times S_c^j)$. When $Q_{ij}^{(L+1)}$ is not empty there exists one and only one value for the considered entry which leads to a zero jump. Every other value increases the rank of B_{L+1} by the number of elements of the set $Q_{ij}^{(L+1)}$. According to the given definition of structure for a partial behaviour matrix, the considered increase in rank B_{L+1} due to $M_{ij}(L+1)$ will be entirely observed in the invariants ν_i and μ_j and not in any other invariant.

Example 2

Consider the sequence

$$M(1) = \begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}, \quad M(2) = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 1 \end{bmatrix}, \quad M(3) = \begin{bmatrix} 0 & 0 \\ * & ? \\ ? & ? \end{bmatrix}$$

The partial behaviour matrix B_3 becomes

$$\begin{array}{cc} & \begin{array}{cccccc} * & * & \circ & * & \circ & \circ \\ \downarrow & & \downarrow & & \downarrow & \end{array} \\ \begin{array}{c} * \\ \circ \rightarrow \\ * \\ \circ \\ \circ \rightarrow \\ * \\ \circ \\ \circ \rightarrow \\ \circ \end{array} & \left[\begin{array}{cccccc} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & * & ? \\ 0 & 0 & 0 & 1 & ? & ? \\ 0 & 0 & 0 & 0 & ? & ? \\ 0 & 0 & * & ? & ? & ? \\ 0 & 1 & ? & ? & ? & ? \\ 0 & 0 & ? & ? & ? & ? \\ * & ? & ? & ? & ? & ? \\ ? & ? & ? & ? & ? & ? \end{array}\right] \end{array} = B_3 \tag{13}$$

and the insertion of $M_{21}(3)$ is considered. Leaving this entry undetermined, the rank of the partial behaviour matrix B_3 is 3 while its structure is given by $\nu_1 = 1$, $\nu_2 = 0$, $\nu_3 = 2$, $\mu_1 = 1$, $\mu_2 = 2$. Thus it follows that

$$S_r^2 = \{2, 5, 8\}, \quad S_c^1 = \{3, 5\} \tag{14}$$

$$S_{21}^{(3)} = \{(2, 5), (5, 3), (8, 1)\} \tag{15}$$

$$\begin{aligned} Q_{21}^{(3)} &= S_{21}^{(3)} \cap \{(2, 3), (2, 5), (5, 3), (5, 5), (8, 3), (8, 5)\} \\ &= \{(2, 5), (5, 3)\} \end{aligned} \tag{16}$$

Thus the jump of rank B_3, μ_2 and ν_1 is 2 when the (2, 1)th entry of $M(3)$ is generic and zero when it assumes the singular value (0 in this case) in correspondence with the main lemma.

Now it has been shown that the control and observation invariants of a partial behaviour matrix play a fundamental role in the description of the dynamical behaviour of formal power series. These invariants play a fundamental role also in the description of input–output and state-space models belonging to canonical classes with respect to given equivalence relations (Beghelli and Guidorzi 1976, Guidorzi 1979, 1981). In the next section a partial realization algorithm based on the use of multi-companion canonical forms is presented.

3. A partial realization algorithm in multi-companion canonical form

A complete description regarding this approach and also considerations regarding the selection of stable models in the family of possible ones can be found in Roman and Bullock (1975). The considered realization algorithm can be based either on the information carried by the control invariants or on the information carried by the observation invariants; dual models are obtained following this method. The control invariants of the partial behaviour matrix will be considered.

Consider a partial behaviour matrix B_L whose control invariants $\mu_1, \ldots, \mu_p$ are assumed to be non-zero. A minimal partial realization (F, G, H) of B_L in a multi-companion canonical form can be obtained by means of the following steps.

Step 1

Construct the input distribution matrix G by means of the first, $(\mu_1 + 1)$th, ..., and $(\mu_1 + \ldots + \mu_{p-1} + 1)$th column of the identity matrix I_n where $n = \mu_1 + \ldots + \mu_p$.

$$G = \begin{array}{c} \begin{bmatrix} 1 & 0 & \ldots & 0 \\ 0 & 0 & & 0 \\ \vdots & 0 & & \vdots \\ \vdots & 1 & & \vdots \\ \vdots & \vdots & & 0 \\ \vdots & \vdots & & 1 \\ 0 & 0 & \ldots & 0 \end{bmatrix} \end{array} \begin{array}{l} \leftarrow 1 \\ \\ \\ \leftarrow (\mu_1 + 1) \\ \\ \leftarrow (\mu_1 + \ldots + \mu_{p-1} + 1) \\ \\ \end{array} \tag{17}$$

Step 2

Denote with $M^k(i)$ the kth column of the coefficient matrix $M(i)$. Then construct the output distribution matrix H as follows

$$H = [M^1(1)\, M^1(2) \ldots M^1(\mu_1) \ldots M^p(1)\, M^p(2) \ldots M^p(\mu_p)] \tag{18}$$

Step 3

Denote with B_L^k the kth column of B_L. Construct then the dynamical matrix F as follows

$$F = [F_{ij}] \quad (i, j = 1, \ldots, p) \tag{19 a}$$

$$F_{jj} = \begin{bmatrix} 0 & \cdots & \cdots & 0 & \alpha_{jj1} \\ 1 & \ddots & & \varnothing & \alpha_{jj2} \\ & \ddots & \ddots & & \vdots \\ & \varnothing & \ddots & & \vdots \\ & & & 1 & \alpha_{jj\mu_j} \end{bmatrix}_{\mu_j \times \mu_j} \tag{19 b}$$

$$F_{ij} = \begin{bmatrix} 0 & \cdots & 0 & \alpha_{ij1} \\ \vdots & & \vdots & \vdots \\ \vdots & \varnothing & \vdots & \alpha_{ij\mu_{ij}} \\ \vdots & & \vdots & 0 \\ 0 & \cdots & 0 & 0 \end{bmatrix}_{\mu_i \times \mu_j}, \quad i \neq j \tag{19 c}$$

where the scalars

$$\alpha_{1j1} \ldots \alpha_{1j\mu_{1j}} \ldots \alpha_{jj1} \ldots \alpha_{jj\mu_j} \ldots \alpha_{pj1} \ldots \alpha_{pj\mu_{pj}} \tag{20}$$

are selected in order to express the linear dependence of the vector $B_L^{p\mu_j+j}$ (if this vector is present in B_L) from the previous linearly independent columns of B_L, i.e.

$$\begin{aligned} B_L^{p\mu_j+j} = \alpha_{1j1} B_L^1 + \ldots + \alpha_{1j\mu_{1j}} B_L^{p(\mu_{1j}-1)+1} + \ldots + \alpha_{jj1} B_L^j + \ldots \\ + \alpha_{jj\mu_j} B_L^{p(\mu_j-1)+j} + \ldots + \alpha_{pj\mu_{pj}} B_L^{p(\mu_{pj}-1)+p} \end{aligned} \tag{21}$$

Note that since only independent columns at the left of $B_L^{p\mu_j+j}$ are considered, the integers μ_{ij} assume the following values

$$\left. \begin{aligned} \mu_{ij} &= \min(\mu_i, \mu_j + 1) \quad \text{for } i < j \\ \mu_{ij} &= \min(\mu_i, \mu_j) \qquad \text{for } i > j \end{aligned} \right\} \tag{22}$$

If the vector $B_L^{p\mu_j+j}$ does not belong to B_L then all the scalars (20) can take arbitrary values. Note, however, that also when relation (21) must be satisfied, some degrees of freedom can remain if, for example, the number of defined entries in $B_L^{p\mu_j+j}$ is lower than the number of independent vectors of B_L at its left.

Proof

The property of the obtained triple (F, G, H) to realize B_L can be easily verified considering that, by construction, $[HF^{k-1}G]^j = M^j(k)$ for $k = 1, \ldots, \mu_j$ and that condition (21), because of the block Hankel structure of B_L, assures the fulfilment of the previous condition for $k = \mu_j + 1, \ldots, L$.

Example 3

The canonical partial realization of the sequence

$$M(1) = \begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}, \quad M(2) = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 1 \end{bmatrix}, \quad M(3) = \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 2 & 3 \end{bmatrix}$$

is considered. The control invariants of B_3 are, in this case, $\mu_1 = 3$ and $\mu_2 = 2$.

Thus the order of the associated minimal partial realization is $n = 5$.

$$B_4 = \begin{array}{cc} & \begin{array}{cccccccc} * & * & * & * & * & \circ & \circ & \circ \end{array} \\ \begin{array}{c} * \\ * \\ * \\ \circ \\ * \\ * \\ \circ \\ \circ \\ \circ \\ \circ \\ \circ \\ \circ \end{array} & \left[\begin{array}{cccccccc} 1 & 0 & 0 & 0 & 0 & 0 & ? & ? \\ 0 & 0 & 0 & 0 & 1 & 0 & ? & ? \\ 0 & 0 & 0 & 1 & 2 & 3 & ? & ? \\ 0 & 0 & 0 & 0 & ? & ? & \multicolumn{2}{c}{.} \\ 0 & 0 & 1 & 0 & ? & ? & \multicolumn{2}{c}{.} \\ 0 & 1 & 2 & 3 & ? & ? & \multicolumn{2}{c}{.} \\ 0 & 0 & ? & ? & ? & ? & \multicolumn{2}{c}{.} \\ 1 & 0 & ? & ? & ? & ? & \multicolumn{2}{c}{.} \\ 2 & 3 & ? & ? & ? & ? & \multicolumn{2}{c}{.} \\ ? & ? & & & & & ? & ? \\ ? & ? & . & . & . & . & ? & ? \\ ? & ? & & & & & ? & ? \end{array}\right] \end{array} \qquad \begin{array}{l} \mu_1 = 3 \\ \mu_2 = 2 \\ \\ \nu_1 = 1 \\ \nu_2 = 2 \\ \nu_3 = 2 \end{array}$$

Step 1

The input distribution matrix is given by

$$G = \begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}$$

Step 2

The output distribution matrix is given by

$$H = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 & 1 \end{bmatrix}$$

Step 3

The dynamical matrix F is of the type

$$F = \left[\begin{array}{ccc|cc} 0 & 0 & \alpha_{111} & 0 & \alpha_{121} \\ 1 & 0 & \alpha_{112} & 0 & \alpha_{122} \\ 0 & 1 & \alpha_{113} & 0 & \alpha_{123} \\ \hdashline 0 & 0 & \alpha_{211} & 0 & \alpha_{221} \\ 0 & 0 & \alpha_{212} & 1 & \alpha_{222} \end{array}\right]$$

The scalars $[\alpha_{111}\ \alpha_{112}\ \alpha_{113}\ \alpha_{211}\ \alpha_{212}]$ and $[\alpha_{121}\ \alpha_{122}\ \alpha_{123}\ \alpha_{221}\ \alpha_{222}]$ must be selected in order to express the linear dependence of the vectors B_4^7 and B_4^6 from previous independent vectors in the matrix B_4. Since the vector B_4^7 does not belong to B_3 the scalars $[\alpha_{111}\ \alpha_{112}\ \alpha_{113}\ \alpha_{211}\ \alpha_{212}]$ can be arbitrary. The vector B_4^6, on the contrary, is the sixth vector in B_3.

Since only the first three entries of B_4^6 are defined, the selection of the scalars $[\alpha_{121}\ \alpha_{122}\ \alpha_{123}\ \alpha_{221}\ \alpha_{222}]$ must be such to satisfy the relation

$$\begin{bmatrix} 0 \\ 0 \\ 3 \end{bmatrix} = \alpha_{121} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + \alpha_{122} \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} + \alpha_{123} \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix} + \alpha_{221} \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} + \alpha_{222} \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

The following values are thus obtained

$$\begin{aligned} &\alpha_{121} = 0, \quad \alpha_{221} = \text{arbitrary} \\ &\alpha_{122} = \text{arbitrary}, \quad \alpha_{222} = 3 \\ &\alpha_{123} = 0 \end{aligned}$$

The dynamical matrix F is thus given by

$$F = \left[\begin{array}{ccc|cc} 0 & 0 & \times & 0 & 0 \\ 1 & 0 & \times & 0 & \times \\ 0 & 1 & \times & 0 & 0 \\ \hdashline 0 & 0 & \times & 0 & \times \\ 0 & 0 & \times & 1 & 3 \end{array}\right]$$

where the elements denoted by '$\times$' are arbitrary entries.

If we had taken the observation invariants, we would have obtained:

$$F = \left[\begin{array}{c|cc|cc} 0 & \times & 0 & 0 & 0 \\ \hdashline 0 & 0 & 1 & 0 & 0 \\ 1 & \times & \times & \times & 0 \\ \hdashline 0 & 0 & 0 & 0 & 1 \\ 2 & \times & \times & \times & 3 \end{array}\right]$$

where the entries denoted by '$\times$' are completely arbitrary. The number of arbitrary parameters (7) is the same as in the realization performed by means of control invariants.

Remark 1

In the considered partial realization algorithm all the invariants have been assumed different from zero. The case of zero invariants has modest interest and can be easily treated. In fact, the presence of null invariants indicates that the associated inputs and/or outputs (control and/or observation invariants) of the model (F, G, H) must be assumed linearly dependent from the remaining ones. Thus it is sufficient to avoid the introduction of any dynamics (blocks of F) associated with the considered inputs or outputs. The algebraical links with remaining inputs or outputs can be easily introduced in the matrices G or H.

Remark 2

The conditions leading to *non-unique* minimal partial realizations are very clear in the proposed algorithm (and in the numerical examples). From an algebraic point of view these conditions correspond to the possibilities of expressing the linear dependence of a vector from a set of non-independent vectors. The canonical forms used, because of their *minimal* parametrization, directly exhibit the various sets of scalars which can be used to solve the previous dependence problem. This responds to aspect (*c*) of the minimal partial realization problem (Definition 1).

Remark 3

The obtained state-space canonical models completely describe the whole class of minimal partial realizations for a given power series since every other minimal realization can be obtained from these models via the equivalence relation induced by a change of coordinates in the state space.

4. Uniqueness of minimal partial realizations; finite generic sequences

In the previous sections we have shown that by detailed and thus meticulous inspection it is possible to decide whether a unique minimal partial realization exists or not and how to parametrize in the latter case. However, a more general and less detailed, necessary and sufficient condition for uniqueness is not given. For the partial realization, we can find this in the partial realizability criterion, which is a reflection of the realizability criterion (3) for sequences of infinite length. For finite length sequences a realization of finite dimension always exists; the following criterion gives the conditions for the uniqueness of the minimal partial realization of the finite sequence $\{M(i)\}_L$.

Definition 2. *Partial realizability criterion* (*PRC*)

There exist positive integers N and N' such that:

(*a*) $N + N' = L$

(*b*) $\rho H[N', N] = \rho H[N' + 1, N] = \rho H[N', N + 1](= n)$

Note that $\rho H[N', N]$ is the rank of the Hankel matrix of size $qN' \times pN$ whose elements belong to $\{M(i)\}_{L-1}$.

If the partial realizability criterion is fulfilled the Ho–Kalman algorithm can be applied (Ho and Kalman 1966) leading to a realization. It is easy to prove that this realization is minimal and moreover that there exists only one such realization (see Tether (1970) and/or Appendix 2). Moreover it is proved in Appendix 2 that if the PRC is not satisfied, a unique MPR does not exist, where this necessity has not been proved before. Consequently the following theorem holds.

Theorem 1

A finite sequence $\{M(i)\}_L$ has a unique minimal partial realization if and only if the partial realizability criterion is fulfilled.

We will now work out an example, illustrating the effect of this criterion.

Example 4

Let $\{M(i)\}$ be the Markov parameters of a single-input single-output system (SISO), which consists of two delay lines of 1 and 5 samples:

$$\{M(i)\} = 1, 0, 0, 0, 1, 0, 0 \ldots$$

The infinite Hankel matrix is then given by

$$\begin{matrix}
1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & \ldots \\
0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & \\
0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \\
0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \\
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & . & \\
\vdots & & & & & & & & & \ddots &
\end{matrix}$$

and it is easy to show that:

For $2 \leqslant L \leqslant 4$,

(i) the rank will be 1;

(ii) the PRC is satisfied for all N, N' where $N + N' = L$;

(iii) there is a unique MPR with transfer function z^{-1} and, for example, $F = 0$, $G = H = 1$.

For $5 \leqslant L \leqslant 7$,

(i) the rank n will be 4;

(ii) the PRC is not satisfied for any N, N' where $N + N' = L$, for $N, N' \geqslant 4$ will require $L \geqslant 8$;

(iii) there is not a unique MPR.

For $L = 8$,

(i) the rank n equals 4;

(ii) the PRC is fulfilled for $N = N' = 4$;

(iii) there is a unique MPR with transfer function

$$z^{-1}+z^{-5}+z^{-9}+\dots = z^3/(z^4-1)$$

for example, $H=[1 \quad 0 \quad 0 \quad 0]$

$$F=\begin{bmatrix}0&1&0&0\\0&0&1&0\\0&0&0&1\\1&0&0&0\end{bmatrix},\quad G=\begin{bmatrix}1\\0\\0\\0\end{bmatrix}$$

For $L=9$,

(i) the rank n equals 5;

(ii) the PRC is not satisfied for any N, N' where $N+N'=L$, for $N, N' \geqslant 5$ will require $L \geqslant 10$;

(iii) there is not a unique MPR.

For $L \geqslant 10$,

(i) the rank n equals 5;

(ii) the PRC is satisfied for $N \geqslant 5$, $N' \geqslant 5$ and $N+N'=L$;

(iii) there exists a unique MPR with transfer function $z^{-1}+z^{-5}=(z^4+1)/z^5$, for example

$$F=\begin{bmatrix}0&1&0&0&0\\0&0&1&0&0\\0&0&0&1&0\\0&0&0&0&1\\0&0&0&0&0\end{bmatrix},\quad G=\begin{bmatrix}1\\0\\0\\0\\1\end{bmatrix}$$

$$H=[1 \quad 0 \quad 0 \quad 0 \quad 0]$$

Formally the necessary and sufficient condition for a unique MPR is stated by the PRC. In practice this criterion shows two disadvantages:

(1) It is a rank test on at least three different matrices, which implies quite a lot of work. We say 'at least' as we do not know N and N' beforehand, which could fulfil PRC. Therefore we have to check all possible pairs (N, N') as long as no pair is found that satisfies the PRC.

(2) A rank test becomes rather indefinite when it concerns inaccurate data.

In cases where one is dealing with data which is not disturbed by noise, the second disadvantage may be overcome, but numerical (truncational) errors may still be troublesome. If we are dealing with noisy data, the exact minimal partial realization will be of a high dimension and will also incorporate the noise contribution. We will now show that the partial realizability criterion can be substantially simplified for such 'noisy sequences'. A 'noisy sequence' will be defined to belong to the following class of sequences:

Definition 3. *Finite generic sequence* (*FGS*)

A finite generic sequence $\{\tilde{M}(i)\}_L$ is a sequence of L $(q \times p)$ matrices for which *every* Hankel matrix constructed from this sequence has full rank.

Comment

In practice it concerns a sequence of stochastic data. Only in the exceptional case of certain deterministic (particular, singular) sequences, dependences may appear among the rows (or columns) of the Hankel matrices concerned. Consequently the probability is one that a random sequence will be an FGS.

Because of the special properties of a FGS, a concrete statement can be made with respect to the minimal dimension of the corresponding partial realization, reflected in the next theorem.

Theorem 2

The minimal dimension $\tilde{n}$ of a finite generic sequence is the rank of the partial behaviour matrix and equals the dimension of the biggest square submatrix of known elements.

The verification of this theorem is easy and will be left to the reader.

The biggest square submatrix of B_L has to contain either an integer number of block rows or an integer number of block columns or both. Based on this statement and on Theorem 2 an expression for the minimal dimension $\tilde{n}$ can be derived

$$\tilde{n} = \max\left\{q \text{ entier}\left[(L+1)\frac{p}{p+q}\right], p \text{ entier}\left[(L+1)\frac{q}{p+q}\right]\right\} \tag{23}$$

(for derivation see Appendix 3).

For a finite generic sequence the partial realizability criterion cannot be satisfied by linear dependence between the elements in the rows/columns in the Hankel matrix; the stochastic nature of the sequence will prevent this. Therefore the only situation in which the PRC can be satisfied, is when the rank of an enlarged Hankel matrix cannot increase because it is restricted by the smallest dimension of H. This result is stated in the next lemma.

Lemma 2

For a finite generic sequence the partial realizability criterion is satisfied *if and only if* there exist integers N', N such that $N' + N = L$ and $H[N', N]$ is a square matrix.

The proof of this lemma is given in Appendix 4. If the condition mentioned in this lemma is fulfilled, the rank of the square Hankel matrix will equal the minimal dimension $\tilde{n}$ of the finite generic sequence.

The results above give us instruments for developing a very concrete and manageable criterion for the uniqueness of a minimal partial realization of an FGS.

Theorem 3

A finite generic sequence has a unique minimal partial realization *if and only if*

$$L = \frac{a(p+q)}{k} \tag{24}$$

where $a \in N^+$ and k is the greatest common divisor (GCD) of (p, q).

Proof

In Theorem 1 it has been proved that the PRC is a necessary and sufficient condition for the existence of a unique MPR which, of course, also holds for an FGS.

According to Lemma 2, the fulfilment of the PRC is equivalent to the existence of a square Hankel matrix of size $qN' \times pN$. Consequently this is true if and only if

$$\begin{cases} L = N' + N \\ N'q = Np = \tilde{n} \\ N', N \in N^+ \end{cases}$$

$$\begin{Bmatrix} N', N \text{ integer} \\ N'q = Np = \tilde{n} \end{Bmatrix} \Leftrightarrow \begin{cases} \tilde{n} = \dfrac{apq}{k}, \quad a \in N^+ \\ k = \text{GCD of } (p, q) \\ N' = \dfrac{ap}{k}, \quad N = \dfrac{aq}{k} \end{cases}$$

furthermore $L = N' + N = a(p+q)/k$, and as a result

$$L = \frac{a(p+q)}{k}, \quad \tilde{n} = L\frac{pq}{p+q} \qquad \square$$

Comment

(i) This criterion for uniqueness is a very suitable, simply decountable tool.

(ii) These specific expressions for L and $\tilde{n}$ correspond with the general formulation for $\tilde{n}$, as expressed in (23).

(iii) The rank of the partial behaviour matrix was $\tilde{n}$. Consequently the number of degrees of freedom for a state-space description equals $\tilde{n}(p+q)$. By putting this equal to the number of elements Lpq in the FGS $\{\tilde{M}(i)\}_L$ the same condition (24) can be derived.

Finally, if, for a finite generic sequence the partial realizability criterion is fulfilled or equivalently L satisfies Theorem 3 the realization can easily be found, for example, by means of the Ho–Kalman algorithm, or the algorithm presented in the previous section. If the PRC is not fulfilled application of a multi-companion realization algorithm will lead to a minimal partial realization with a number of free parameters. These parameters can be given arbitrary values, reflecting the addition of extra matrices $\tilde{M}(i)$ to the given sequence, until L satisfies the uniqueness criterion. Another approach is, of course, to neglect the last part of the data sequence, such that L satisfies the given condition. However, in this case these neglected matrices will not be realized exactly and therefore the solution to the MPR problem will not be correct in the sense of Definition 1.

5. Some considerations of the approximate partial realization problem

In this section we would like to ponder on a realization problem that is strongly connected to the minimal partial realization problem as analysed before. In fact the MPR problem can be considered both as a problem of exact modelling of a finite data set, and as an approximation method. The latter interpretation will be actual if a restricted part of an available matrix sequence is used to construct a—low order—approximation of the complete sequence.

In the field of system identification, a related problem arises, where one wants to fit the impulse response of a linear system of low dimension to a finite sequence of matrices. This approximate realization problem is met when one has to develop a state-space description of a process, of which a sequence of Markov parameters has been measured or estimated, and where this obtained data consequently is noise-corrupted (see, for example, Damen and Hajdasinski (1982)). Also in the spectral factorization related to the stochastic realization problem one has to deal with this approximate realization (see, for example, van Zee and Bosgra (1979)). It also plays a fundamental role when applying dimension reduction in optimal filter design (see, for example, Silvermann and Bettayeb (1980)).

In view of this, and building on Definition 1, that was initiated by Kalman (1979), the approximate partial realization problem can be defined as follows.

Definition 4. *Approximate partial realization problem* (*APR*)

Given a sequence $\{M(i)\}_L$ and the dimension n of a corresponding MPR, where $L \in N^+$ (possibly infinite)

(*a*) find all realizations of a given dimension $n_r < n$, where a prescribed norm is minimized;

(*b*) give the necessary and sufficient conditions under which this approximate realization is unique;

(*c*) in the case of non-uniqueness, parametrize all APRs.

This problem has turned out to be a difficult one to solve.

The characteristic of uniqueness is still important, though one might expect that this is trivial as $n_r < n$. The following example shows the non-triviality of the uniqueness.

Example 5

Let $\{M(i)\}_L = 0, 0, 1$ so $L = 3$, $p = q = 1$

$$n = \text{rank } B_L = \text{rank} \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & ? \\ 1 & ? & ? \end{bmatrix} = 3$$

and take $n_r = 2 < n$. Such a system ($n_r = 2$) may be parametrized as

$$\bar{M}(1) = \beta_1, \quad \bar{M}(2) = \beta_2, \quad \bar{M}(k) = \alpha_1 \bar{M}(k-1) + \alpha_2 \bar{M}(k-2), \quad k > 2$$

The prescribed norm can be taken a Frobenius norm on the finite matrix sequence

$$V = \beta_1^2 + \beta_2^2 + (\alpha_1 \beta_2 + \alpha_2 \beta_1 - 1)^2$$

The minimum of this norm is zero, but can not be obtained by finite values for the parameters. Nevertheless the minimum may be approximated in an arbitrarily accurate way by choosing β_1 and β_2 very small and α_1 and α_2 very big, such that $\alpha_1 \beta_2 + \alpha_2 \beta_1 = 1$. So there are infinitely many solutions for $V = \delta$, where $\delta \ll 1$.

There was a tendency to attack the APR problem in a way similar to that of the minimal realization problem itself (Zeiger and McEwen 1974, Kung 1978, van Zee and Bosgra 1979, Staar *et al.* 1981, Silverman and Bettayeb 1980, Damen and Hajdasinski 1982, Damen *et al.* 1982). However all methods presented by these authors fail to solve the problem in the case of a finite data sequence and for a specified norm as mentioned in part (*a*) of Definition 4. In many algorithms this failure is reflected in the problem of approximating a finite block Hankel matrix with a block Hankel matrix of lower rank. This problem is still unsolved.

However, if an infinite number of matrices is available and the corresponding n is finite, the APR problem may be solved using, for example, the optimal model-reduction results of Adamjan, Arov and Krein (see Kung and Lin (1981)). This algorithm is capable of finding a realization of prescribed dimension n_r according to the minimization of the spectral norm of the infinite Hankel matrix, once an exact model of the original sequence $\{M(i)\}_L$ has been found. Such an exact model of the original finite sequence can be constructed with the use of the theory on minimal partial realizations as discussed in this paper. In fact all this means that given a finite data sequence $\{M(i)\}_L$ first a minimal partial realization is constructed, based on these L matrices. With a model-reduction algorithm this realization is reduced in dimension, according to some specified criterion.

At this moment we restrict ourselves to indicating a few problems that may occur when following this procedure. The MPR corresponding to the finite $\{M(i)\}_L$ may not be unique. This means that different models of minimal dimension n exist that reconstruct the given sequence exactly. There will be some freedom to choose a model from a class of 'good' models, in which all models have different extension sequences $\{M(i)\}_{L+1,\ldots,\infty}$. By choosing a specific model in this class and applying an optimal model-reduction algorithm as described above, the resulting reduced-order model will depend heavily on the chosen extension sequence. This is caused by the fact that the model-reduction algorithm takes into account an infinite sequence of matrices to approximate the model. It has to be stressed that the extension sequence was not part of the data given, and therefore to some extent will be indeterminate. Consequently the condition of uniqueness of the MPR remains an important one.

Of course the free parameters in the MPR could be chosen according to a criterion measuring the errors between the original *finite* matrix sequence and the corresponding part of the finally obtained APR. However in that case the result will be a two-step approximation method, applying two different approximation criteria: one based on a finite sequence; the other based on an infinite sequence. The overall result will not necessarily be well-defined.

A more pragmatic solution can probably be found when considering the given matrix sequence $\{M(i)\}_L$ to be a measured or estimated sequence of Markov parameters (or covariance matrices) of a stable system. In most cases the length of the sequence to be measured or estimated (L) will be chosen in such a way that the values of the elements in the extension can be neglected. In addition to this, the MPR of the finite sequence $\{M(i)\}_L$ could be chosen in such a way that the elements in the extension sequence are as small as possible.

Apart from the possible non-uniqueness of the MPR, a second problem may arise that is related to the preceding remark. As analysed in Byrnes and Lindquist (1982), no statements can be made, in general, on the stability of a MPR, even if it is known that the finite data sequence is a part of an infinite stable sequence of matrices. As we can only apply the model-reduction algorithm to stable models, in fact we have to find minimal partial *stable* realizations. This problem has still to be studied.

6. Conclusions

Some considerations on the multivariable partial realization problem have been formulated. The lines followed have been suggested by the work of Kalman (1979) who has introduced new and important concepts regarding the dynamic structure of formal power series.

The algorithm introduced shows, once again, the fundamental importance of the concept of system structure in multivariable realization problems. Nevertheless, this greatly complicates the test of the uniqueness of the extension. Tether (1970) introduced a sufficient condition for uniqueness in the form of the 'partial realizability criterion' (PRC). In the present paper it has been proved that this criterion is also necessary. Although this is quite an acceptable general criterion for realizability, we felt the need to find a simpler test for a generic sequence. First, this generic sequence was defined in the class of 'finite generic sequences' (FGSs). Next we were able to show that a simple algebraic test on the length L of such an FGS can be applied as a necessary and sufficient one for the existence of a unique extension. This powerful condition then indicates the number of terms in the formal power series (Markov parameters or covariances) by estimation or measurement, which are necessary to obtain a unique extension.

The uniqueness of the extension turns out to be non-trivial for the approximate partial realization problem (APR) too, where one wants to find a realization of prescribed low dimension, which fits the formal power series as well as possible. This APR has been elaborated upon and it is clear that this is still an open-ended problem.

ACKNOWLEDGMENT

The authors would like to thank Professor P. Eykhoff and the reviewer for their substantial remarks while preparing the final version of this paper.

Appendix 1

Proof of the Main Lemma†

As

$$\rho[A \,\vdots\, B] = \rho \begin{bmatrix} A \\ \cdots \\ C \end{bmatrix} = \rho[A] = n$$

it follows that $B = AV$ and $C = UA$, where U and V are matrices of proper dimensions and not necessarily unique. Consequently, as the matrix D has to follow both the row and the column dependences in order to keep the rank equal to n, we necessarily

† While preparing the final version of this paper, the authors came across a paper of Godbole (1972), giving similar results.

obtain

$$\Omega(D)=\left[\begin{array}{c:c} A & B \\ \hdashline C & D \end{array}\right]=\left[\begin{array}{c:c} A & AV \\ \hdashline UA & D \end{array}\right]=\left[\begin{array}{c:c} A & AV \\ \hdashline UA & UAV \end{array}\right]$$

By using the generalized inverse A^+ of matrix A one can write

$$D = UAV = UAA^+AV = CA^+B \tag{A 1.1}$$

So there exists a solution.

Now suppose that there exists another solution D^* which necessarily has to fulfil

$$D^* = U^*AV^* \quad \text{while } B = AV^* \text{ and } C = U^*A$$

Because $AV = AV^*$ and $UA = U^*A$ it follows that

$$D^* = U^*AV^* = U^*AV = UAV = D$$

As a result the solution (A 1.1) for D is unique. □

Appendix 2

Proof of Theorem 1†

Sufficiency ⟨PRC⟩⇒⟨unique MPR⟩

If the PRC is satisfied for integers N' and N, by applying the main lemma, a unique extension $M(L+1)$ can be found such that

$$\rho H[N'+1, N+1] = \rho H[N'+1, N] = n$$

Because of the structure of the Hankel matrix it follows that

$$\rho H[N'+2, N] = \rho H[N'+1, N] = n \tag{A 2.1}$$

Then the application of the main lemma once again supplies a unique extension $M(L+2)$. Proceeding in this way by induction, it is proved that there exists a unique minimal partial realization. As the Ho–Kalman algorithm provides a partial realization, it necessarily has to be this unique one. (See also Tether (1970).)

Necessity ⟨unique MPR⟩⇒⟨PRC⟩

The necessary condition is going to be proved by proving the implication

$$\langle \neg \text{PRC} \rangle \Rightarrow \langle \text{non-unique MPR} \rangle \tag{A 2.2}$$

Suppose B_L is the partial behaviour matrix corresponding to the sequence $\{M(i)\}_L$, having structure indices $\nu_1, \ldots, \nu_q$ and $\mu_1, \ldots, \mu_p$. Let us define

$$\nu = \max_i \nu_i, \quad \mu = \max_j \mu_j$$

Kalman (1971) has proved that for any finite sequence $\{M(i)\}_L$ the integer L' defined by $L' = \nu + \mu$ indicates the *shortest* possible length of the corresponding sequence $\{M'(i)\}_{L'}$, where $M'(i) = M(i)$, $i = 1, 2, \ldots, \min(L, L')$, such that:

† An alternative proof of this theorem, along the lines given by Bosgra (1983), can be found in Van den Hof *et al.* (1984).

(i) the PRC is fulfilled for $\{M'(i)\}_{L'}$ reflected in

$$\rho H'[\nu, \mu] = \rho H[\nu+1, \mu] = \rho H'[\nu, \mu+1] = n$$

and consequently this sequence $\{M'(i)\}_{L'}$ has a unique extension;

(ii) rank $B_L = \text{rank } B'_{L'} = n$, which means that the minimal dimensions of the partial realizations of both sequences are equal.

If the PRC is fulfilled for $\{M(i)\}_L$ it is easy to see that $L' \leqslant L$. In this case the uniqueness of the realization is guaranteed. If the PRC is not fulfilled for $\{M(i)\}_L$ then $L' > L$ and the question is whether the extension $\{M'(L+1), \ldots, M'(L')\}$ can be unique so that PRC is fulfilled for $\{M'(i)\}_{L'}$. If this $\{M'(L+1), \ldots, M'(L')\}$ is unique, then the MPR of $\{M(i)\}_L$ is unique although PRC was *not* fulfilled for $\{M(i)\}_L$. We will now prove that this cannot be true.

Suppose an extension of the sequence $\{M(i)\}_L$ has been chosen

$$\{M'(L+1), M'(L+2), \ldots, M'(L'-1)\}$$

in such a way that rank $B'_{L'-1} = \text{rank } B'_{L'}$. Such a choice is always possible, but not necessarily unique. In the situation $L' = L+1$, this extension has zero length but this does not violate the following reasoning concerning the last extension $M'(L')$.

It will be proved that $M'(L')$ cannot be determined uniquely in order to have rank $B'_{L'} = \text{rank } B'_{L'-1}$, which implies that the complete extension $\{M'(L+1), \ldots, M'(L')\}$ cannot be unique.

The PRC for $\{M'(i)\}_{L'}$ can be split up:

(*a*) rank $H'[\nu, \mu] = \text{rank } H'[\nu+1, \mu]$ (A 2.3)

$$H'[\nu+1, \mu] = \begin{bmatrix} M'(1) & M'(2) & \cdots & M'(\mu-1) & M'(\mu) \\ \cdot & & & \vdots & \vdots \\ M'(2) & & & \vdots & \vdots \\ \vdots & & & \vdots & \vdots \\ M'(\nu) & \cdots & \cdots & M'(L'-2) & M'(L'-1) \\ M'(\nu+1) & \cdots & \cdots & M'(L'-1) & ? \end{bmatrix}$$

Markov parameter $M'(L')$ has to be substituted in position '?'. Because of the definition of μ the last block column in $H'[\nu, \mu]$ will contain at least one independent column vector. Then with Corollary 1, all entries of $M'(L')$ appearing in this independent column can be arbitrary, concerning this rank condition, as they cannot affect rank $H[\nu+1, \mu]$.

(*b*) The dual proof for at least one particular row in $M'(L')$ can be given by means of the condition

$$\rho H'[\nu, \mu] = \rho H'[\nu, \mu+1]$$

Finally we may conclude that at least the element on the cross-section of the independent row and the independent column is completely arbitrary. This implies that $M'(L')$ is non-unique, as is the extension $\{M'(L+1), \ldots, M'(L')\}$.

Result

If $\{M(L+1), \ldots, M(L'-1)\}$ is chosen in such a way that rank $B_{L'-1} = \text{rank } B_L$, it is not possible to find a unique $M(L')$ such that rank $B_{L'}$ = rank B_L. Consequently the extension $\{M(L+1), \ldots, M(L')\}$ will always be non-unique, which means that (A 2.2) has been proved.

Appendix 3

Derivation of the expression for the minimal dimension of a finite generic sequence

According to Theorem 2, the minimal dimension $\tilde{n}$ of a FGS equals the dimension of the biggest square submatrix in the partial behaviour matrix B_L. Because of the structure of B_L the biggest square submatrix has to contain an integer number of block rows *or* an integer number of block columns (or both).

(*a*) Consider the situation where this biggest square submatrix contains an integer number of block rows l. The upper limit for the number of rows $n' = lq$ is given by the minimal length of these rows being $(L+1-l)p$ in order to keep the rank equal to $\tilde{n}$. This restriction $lq \leqslant (L+1-l)p$ becomes

$$l \leqslant (L+1)\frac{p}{p+q}$$

Because lq is maximized and l is an integer it follows that

$$lq = \tilde{n} = q \text{ entier}\left[(L+1)\frac{p}{p+q}\right]$$

(*b*) Considering the dual situation of an integer number of block columns m in the biggest square submatrix will lead to a dual result:

$$mp = \tilde{n} = p \text{ entier}\left[(L+1)\frac{q}{p+q}\right]$$

Situation (*a*) or (*b*) will appear depending on which expression gives a maximal value of the dimensions of the square submatrix. As a result

$$\tilde{n} = \max\left\{q \text{ entier}\left[(L+1)\frac{p}{p+q}\right],\ p \text{ entier}\left[(L+1)\frac{q}{p+q}\right]\right\}$$

It may happen that the two terms in this expression are equal so that both situation (*a*) and situation (*b*) occur. This is precisely the case for $l + m = L$ and for which the PRC is fulfilled.

Appendix 4

Proof of Lemma 2

Sufficiency $\langle H[N', N]$ is square and of full rank because it concerns an FGS$\rangle \Rightarrow \langle$PRC satisfied$\rangle$

$\rho H[N', N] = \tilde{n}$ while $H[N', N]$ is a $\tilde{n} \times \tilde{n}$ matrix. Consequently $H[N'+1, N]$ is a matrix of dimensions $(\tilde{n}+q) \times \tilde{n}$ and has rank $\tilde{n}$. Also $H[N', N+1]$ is a matrix of dimensions $\tilde{n} \times (\tilde{n}+p)$ and has rank $\tilde{n}$ which means that the PRC has been fulfilled.

Necessity

PRC is satisfied, thus

$$\rho H[N', N] = \rho H[N'+1, N] = \rho H[N', N+1] = \tilde{n}$$

It concerns finite generic sequences, so that these matrices have full rank, which implies

$$\tilde{n} = \min [N'q, Np] = \min [(N'+1)q, Np] = \min [N'q, (N+1)p]$$

As $(N'+1)q > N'q$ and $(N+1)p > Np$ it follows

$$\tilde{n} = Np = N'q \text{ so that } H[N', N] \text{ is a square matrix.} \qquad \square$$

References

Ackermann, J. E., and Bucy, R. S., 1971, *Inform. Control*, **19**, 224.

Beghelli, S., and Guidorzi, R. P., 1976, *I.E.E.E. Trans. autom. Control*, **21**, 692.

Bistritz, Y., 1983, *SIAM J. Control Optim.*, **21**, 804.

Bosgra, O. H., 1983, *Syst. Control Lett.*, **3**, 181.

Bosgra, O. H., and van der Weiden, A. J. J., 1980, *I.E.E.E. Trans. autom. Control*, **25**, 20.

Brockett, R. W., 1978, *Proc. I.E.E.E. Conf. on Decision and Control*, p. 1048.

Byrnes, C. I., and Lindquist, A., 1982, *Syst. Control Lett.*, **2**, 99.

Damen, A. A. H., and Hajdasinski, A. K., 1982, *Proc. 6th IFAC Symp. on Identification and System Parameter Estimation*, Washington DC, pp. 903–908.

Damen, A. A. H., Van den Hof, P. M. J., and Hajdasinski, A. K., 1982, *Syst. Control Lett.*, **2**, 202.

Godbole, S. S., 1972, *I.E.E.E. Trans. autom. Control*, **17**, 173.

Guidorzi, R. P., 1979, Equivalence, invariance and dynamical systems canonical modelling. Research Report, Istituto di Automatica, University of Bologna, Italy; 1981, *Automatica*, **17**, 117.

Ho, B. L., and Kalman, R. E., 1966, *Regelungstechnik*, **14**, 545.

Kalman, R. E., 1960, *Proc. 1st IFAC Congress*, Moscow, Vol. 1, edited by J. F. Coales (London: Butterworth), pp. 481–492; 1963, *SIAM J. Control*, **1**, 152; 1968, *Proc. C.I.M.E. Summer School*, Pontecchio Marconi 1968, Edizioni Cremonese, Roma, 1969; 1971, *Aspects of Network and Systems Theory*, edited by R. E. Kalman and N. DeClaris (Holt, Rinehart and Winston), p. 385; 1979, *Acta polytech. scand.*, **31**, 9.

Kung, S., 1978, *12th Asilomar Conf. on Circuits, Systems and Computers*, Pacific Grove, California.

Kung, S., and Lin, D. W., 1981, *I.E.E.E. Trans. autom. Control*, **26**, 832.

Ledwich, G., and Moore, J. B., 1976, *Automatica*, **12**, 497.

Roman, J. R., and Bullock, T. E., 1975, *I.E.E.E. Trans. autom. Control*, **20**, 529.

Silverman, L. M., 1971, *I.E.E.E. Trans. autom. Control*, **16**, 554.

Silverman, L. M., and Bettayeb, M., 1980, *Joint Automatic Control Conf.*, San Francisco, paper FA8-A.

Staar, J., Vandewalle, J., and Wemans, M., 1981, *8th IFAC World Congress*, Kyoto, Vol. 1, pp. 7–12.

Tether, A. J., 1970, *I.E.E.E. Trans. autom. Control*, **15**, 427.

Van den Hof, P. M. J., Damen, A. A. H., and Hajdasinski, A. K., 1984, *Proc. 9th IFAC World Congress*, Budapest.

Zazworsky, R. M., Jensen, D., and Baker, W. P., 1979, *I.E.E.E. Trans. autom. Control*, **24**, 966.

Zee, G. A. van, and Bosgra, O. H., 1979, *Proc. 5th IFAC Symp. on Identification and System Parameter Estimation*, Darmstadt, pp. 477–484.

Zeiger, H. P., and McEwen, J., 1974, *I.E.E.E. Trans. autom. Control*, **19**, 153.