Automatica, Vol. 22. No. 6, pp. 675–694, 1986
Printed in Great Britain.

0005–1098/86 $3.00 + 0.00
Pergamon Journals Ltd.

From Time Series to Linear System—Part II. Exact Modelling*

JAN C. WILLEMS†

The most powerful unfalsified model is defined as that element in a model class which explains a given set of observations and as little else as possible. Algorithms are developed which compute the most powerful unfalsified linear time invariant model starting from an observed time series.

Key Words—Linear systems; time series analysis; modelling; minimal realization; identification; data reduction and analysis; (most powerful unfalsified model).

Abstract—In the second part of this paper the problem of finding an exact model for a q-dimensional infinite time series is considered. First a mathematical vocabulary for discussing exact modelling is developed. It is then shown how the results of Part I guarantee the existence of a most powerful (AR) model for an observed time series. Two algorithms for obtaining such an (AR) model are subsequently derived. One of these algorithms gives a shortest lag input/output model. The problem of obtaining a minimal state space realization of the observed time series is also considered. In order to do that, realization theory based on the truncated behaviour is developed. As an extensive example, the classical situation with impulse response measurements is discussed.

12. INTRODUCTION

ONE OF THE central issues in the modelling of dynamical phenomena may succinctly be formulated as follows:

Given an observed q-dimensional vector time series

$$\tilde{\mathbf{w}}(t_0), \tilde{\mathbf{w}}(t_0+1), \ldots, \tilde{\mathbf{w}}(t_1) \qquad (-\infty \leqslant t_0 \leqslant t \leqslant t_1 \leqslant \infty)$$

with $\tilde{\mathbf{w}}(t) \in \mathbb{R}^q$, find a dynamical model which explains these observations.

This problem is of crucial importance in many diverse areas of application which are concerned with modelling directly on the basis of observations, such as time series analysis, signal processing, econometrics, system theory (identification), automatic control (adaptive control) etc. The usual approach towards providing algorithms for coming up with a model is to postulate a set of equations containing as yet unspecified parameters. These parameters are then fitted by means of the data.

In the process of setting up such a procedure one has to face the conceptual difficulty that many models which one would like to use will be unable to explain the data exactly and it may be objectionable—at least from a Popperian physics oriented point of view—to accept a model which in this sense is refuted by the data. The standard way out of this dilemma is to invoke the philosophy of statistics and to employ models which contain, in addition to unspecified parameters, random elements (for example by assuming that the observed time series is a realization of a stochastic process, or that it is the output of a system driven by a stochastic process). Such an assumption then usually guarantees that every (finite) observed time series can occur with a certain probability and in this sense the data will no longer refute the model. The modelling question is then one of sampling: one has to deduce from the observations the probability measure governing the random system.

Granted, there are many situations in which such a framework is indeed a suitable one. However, as a general philosophy, it has many fundamental drawbacks. The main shortcoming, in the author's opinion is that in most applications the lack of fit between data and model is not in the first place due to randomness or measurement noise but to the fact that one consciously uses a model whose structure is unable to capture the complexity of the phenomenon which one is observing. As such, it is very appealing to formulate this modelling problem as a (deterministic) approximation question instead. This will be pursued in the third part of this paper.

However, before setting up algorithms for obtaining (optimal) approximate models it seems reasonable that one should be able to produce algorithms for obtaining exact models. In other words, the excact modelling issue should precede the approxi-

* Received 3 June 1985; revised 13 January 1986. The original version of this paper was not presented at any IFAC meeting. This paper was recommended for publication in revised form by Associate Editor T. Başar under the direction of Editor H. Kwakernaak.

† Mathematics Institute, University of Groningen, P.O. Box 800, 9700 AV Groningen, The Netherlands.

mate, respectively stochastic one. This is the topic of Part II of this paper: algorithms will be derived for obtaining the best (in the sense of "most powerful") exact linear time invariant model for an observed infinite time series. The underlying existence question of a most powerful model and the specification of what is meant by a dynamical model depends on the theory developed in Part I. The approximate modelling question will be dealt with in Part III.

The motivational premise of this part is: before making headway on realistic but difficult problems such as obtaining non-linear, time-varying, stochastic, approximate models of an observed finite time series, one should have a good understanding of how to obtain an exact linear time invariant deterministic model of an infinite time series.

The reader is referred to some recent papers by Kalman (1982, 1983) in which he discusses (mainly in an econometrics context) the problem of modelling on the basis of data and in which he argues the limited value of the statistician's stochastic approach.

13. MODELLING: A VOCABULARY

First of all, a mathematical language must be set up in which modelling can be discussed in a precise way. In view of the crucial importance which mathematical models play in applications, it is actually quite surprising that such a vocabulary does not seem to be part of the standard arsenal of notions which are taught in introductory mathematics, physics or economics courses. Following the motivation outlined in the introduction, the first issue which should be treated is exact modelling.

Assume that one wishes to model a *phenomenon*. The first step is to "quantify" the variables involved by means of a set S. Elements of the set S are *attributes* of the phenomenon. For example, if the phenomenon is a resistor, then $S = \mathbb{R}^2$, the collection of all possible current/voltage pairs. If the relation between the price, demand and supply of a certain economic resource is to be formalized, then $S = \mathbb{R}^3_+$. When Newton was attempting to describe the relation between the external force and the position of a point mass, he was considering the phenomenon space $S = (\mathbb{R}^3 \times \mathbb{R}^3)^{\mathbb{R}}$, the family of all possible force/position trajectories $t \mapsto (\mathbf{F}(t), \mathbf{q}(t))$.

A *model* for the phenomenon is defined to be a subset $M \subset S$: it is a "law" which says that the phenomenon will only produce outcomes in M. In the case of the resistor, M would be the graph of the current/voltage characteristic of the resistor. In the price/demand/supply example, it would be the graph of the function mapping the price into the corresponding demand and supply. Newton claimed that only those elements of $(\mathbb{R}^3 \times \mathbb{R}^3)^{\mathbb{R}}$, satisfying

$$\mathbf{F} = M\frac{\mathrm{d}^2\mathbf{q}}{\mathrm{d}t^2}$$

are possible force/position trajectories. This yields Newton's second law as a model for the relation between position and external force.

In modelling problems, one usually considers a *class of models* $\mathscr{M}$. In our setting $\mathscr{M}$ will hence be a family of subsets of S, i.e. $\mathscr{M} \subset 2^S$. Ohm's law gives an example of such a situation: Ohm claimed that the voltage across a resistor is a linear function of the current into the resistor. If one postulates that a force/position relation comes from a potential field, then $\mathscr{M} = \{M_V \mid V \in \mathscr{C}^1(\mathbb{R}^3; \mathbb{R})\}$ with

$$M_V := \left\{(\mathbf{q}, \mathbf{F}): \mathbb{R} \to \mathbb{R}^3 \mid \mathbf{F} = \frac{\partial V}{\partial q}(\mathbf{q})\right\}.$$

In other situations, one could assume that $\mathscr{M}$ consists of all linear subspaces of a given vector space, or of all power laws etc. These examples show that $\mathscr{M}$ can express either assumptions of a pragmatic mathematical nature (linearity etc.) or it can express basic physical laws (field equations, conservation laws etc.) which *a priori* limit the possible models.

The model M_1 of a given phenomenon is *more powerful* than the model M_2 if $M_1 \subset M_2$. Indeed, since M_1 allows fewer possibilities than M_2, it has more 'predictive' power and is hence of more value to the user.†

Next, *measurements* will be formalized. Only direct measurements of the attributes of the phenomenon itself will be discussed. Situations in which the outcomes of the measurements are (uncertain) functions of the attributes of the phenomenon will not be treated here, even though many of the ideas generalize to that situation. The measurements will be viewed as a subset $Z \subset S$, and can be thought of either as experimental data or as experimental evidence summarizing, or perhaps extrapolating, observations (e.g. Kepler's laws). In addition it will often be assumed that Z comes from a possible class of sets $\mathscr{Z} \subset 2^S$. The set $\mathscr{Z}$ dictates the measurement sets with which the modeller has to be concerned. For example, $\mathscr{Z}$ could be all finite sets or it could formalize assumptions that the measurements satisfy a boundedness constraint. The set $\mathscr{Z}$ will play an important role

† Cf. Popper (1963, p. 36). "Every good scientific theory is prohibition, it forbids certain things to happen. The more it forbids, the better it is. A theory which is not refutable by any conceivable event is non-scientific. Irrefutability is not a virtue of a theory (as people often think) but a vice."

when approximate modelling is discussed in Part III. A model M is *unfalsified* by the measurements Z if $Z \subset M$.

Assume a phenomenon S is to be modelled on the basis of the measured data, $Z \subset S$ within a given model class $\mathscr{M} \subset 2^S$. Which model should be chosen? Clearly there is a lot to be said in favour of choosing that model in $\mathscr{M}$ which is not contradicted by Z and gives as much predictive power as possible. This is called the *most powerful unfalsified model*. This is summarized in the following definition.

Definition 4. Let S be a set which quantifies a phenomenon. A *model* M for the phenomenon is a subset $M \subset S$. A *model set* $\mathscr{M}$ is a subset $\mathscr{M} \subset 2^S$. A model M_1 is *more powerful* than M_2 if $M_1 \subset M_2$. A *measurement* Z on the phenomenon is a subset $Z \subset S$. The *set of all possible measurements* is a subset $\mathscr{Z} \subset 2^S$. The model M is *unfalsified* by the measurement Z if $Z \subset M$. The unfalsified model M, $Z \subset M \in \mathscr{M}$, is *undominated* in $\mathscr{M}$ if $\{Z \subset M' \in \mathscr{M}$ and $M' \subset M\} \Rightarrow \{M' = M\}$. M_Z^* is the *most powerful unfalsified model in the model class* $\mathscr{M}$ *based on the measurements* Z if $Z \subset M_Z^* \in \mathscr{M}$ and $\{Z \subset M \in \mathscr{M}\} \Rightarrow \{M_Z^* \subset M\}$.

Of course, M_Z^* need not exist, but if it exists, it is obviously unique. In fact, the set $\{M \in \mathscr{M} \mid Z \subset M\}$ may be empty. More to the point, $\mathscr{M}$ may contain many different undominated unfalsified models. There are a number of situations in which it is trivial to see that M_Z^* exists and what it is.

(1) Take $\mathscr{M} = 2^S$. Then $M_Z^* = Z$. This model is often used in daily practice. It is adhered to by those who approach things without (preconceived) ideas ($\mathscr{M} = 2^S$) and consequently only believe what they have observed ($M_Z^* = Z$).

(2) Take $S = \mathbb{R}^n$ and $\mathscr{M} = \{$all linear subspaces of $\mathbb{R}^n\}$. Then $M_Z^* = \text{span}\ \{z \mid z \in Z\}$.

The following, basically trivial, proposition guarantees the existence of M_Z^*.

Proposition 11. Assume that $\mathscr{M} \subset 2^S$ has the intersection property (i.e. $\{\mathscr{M}' \subset \mathscr{M}\} \Rightarrow \left\{\left(\bigcap_{M \in \mathscr{M}'} M\right) \in \mathscr{M}\right\}$ and that for each $Z \in \mathscr{Z}$ there exists a $M \in \mathscr{M}$ such that $Z \subset M$ (e.g. assume $S \in \mathscr{M}$). Then for each $Z \in \mathscr{Z}$ there exists a most powerful unfalsified model M_Z^* in the model class $\mathscr{M}$ based on the measurements Z.

Proof. See Appendix P.

Actually the condition $S \in \mathscr{M}$ together with the intersection property of $\mathscr{M}$ is a familiar condition in topology. Indeed, it is satisfied for the convex sets in what is called a convex structure, and more importantly for the applications which we have in mind, for the closed sets in a topological space etc. The following are examples of this situation:

(i) $\mathscr{M} = 2^S$;

(ii) $S = \mathbb{R}^n$; $\mathscr{M} = \{$all linear subspaces of $\mathbb{R}^n\}$;

(iii) $S =$ a vector space, $\mathscr{M} = \{$all linear subspaces of $S\}$;

(iv) $S =$ a topological vector space; $\mathscr{M} = \{$all closed linear subspaces of $S\}$.

14. MODELLING OF TIME SERIES

If a dynamical system is being modelled, then using the framework of the previous section, the phenomenon space is $S = W^T$, with $T \subset \mathbb{R}$ the time set and W the signal space in which the time signal takes on its values. A model is then simply a subset of $\mathscr{B}$ of W^T. This identifies a model of a dynamical phenomenon with its behaviour and yields the definition of a dynamical system as the triple $\Sigma = \{T, W, \mathscr{B}\}$ of Definition 1. (This provides pleasing circumstantial evidence of the suitability of Definition 1 as the basic definition of a dynamical system.)

Now assume that a q-dimensional real time series $\tilde{\mathbf{w}}: T \to \mathbb{R}^q$ with $T = \mathbb{Z}_+$ or $\mathbb{Z}$ is observed, then in the language of Section 13, $Z = \{\tilde{\mathbf{w}}\}$: the measurement set $\mathscr{Z}$ consists of the singletons. The algorithms which will be obtained are easily extended to the case in which the measurement Z consists of a finite set of time series $\tilde{\mathbf{w}}_1, \tilde{\mathbf{w}}_2, \ldots, \tilde{\mathbf{w}}_k$. However for notational simplicity, the case that Z is a singleton will be used mostly. A model in the class of finite dimensional linear time invariant systems is sought. Part I showed that this is equivalent to looking for an (AR) model or for a model in any of the other equivalent forms introduced in Section 3. It is easy to see that this model class has the intersection property.

Proposition 12. Let $T = \mathbb{Z}_+$ or $\mathbb{Z}$ and $\mathscr{M} = \{\mathscr{B} \subset (\mathbb{R}^q)^T \mid \exists g \in \mathbb{Z}_+$ and $R \in \mathbb{R}^{g \times q}[s]$ such that $\mathscr{B} = \mathscr{B}(R)\}$. Then $(\mathbb{R}^q)^T \in \mathscr{M}$ and $\mathscr{M}$ has the intersection property.

Proof. See Appendix P.

The question of finding the most powerful unfalsified (AR) model for $\tilde{\mathbf{w}}$ thus becomes the following.

Find, for a given $\tilde{\mathbf{w}}: T \to \mathbb{R}^q$, with $T = \mathbb{Z}_+$ or $\mathbb{Z}$, a polynomial matrix $R_{\tilde{\mathbf{w}}}^*$, if one exists such that the corresponding (AR) model explains $\tilde{\mathbf{w}}$ ($: R_{\tilde{\mathbf{w}}}^*(\sigma)\tilde{\mathbf{w}} = \mathbf{0}$), but as little else as possible (i.e. if R is any other polynomial matrix such that $R(\sigma)\tilde{\mathbf{w}} = \mathbf{0}$ then the behaviour induced by R and $R_{\tilde{\mathbf{w}}}^*$ should satisfy $\mathscr{B}(R_{\tilde{\mathbf{w}}}^*) \subset \mathscr{B}(R)$).

The existence of $R_{\tilde{\mathbf{w}}}^*$ follows immediately from Propositions 11 and 12.

Theorem 13. Assume that $\tilde{\mathbf{w}}: T \to \mathbb{R}^q$ with $T = \mathbb{Z}_+$ or $\mathbb{Z}$ is an observed time series. Then there exists a most powerful (AR) model for it:

$$R_{\tilde{\mathbf{w}}}^*(\sigma)\mathbf{w} = \mathbf{0}.$$

Clearly the behaviour of the most powerful unfalsified (AR) model induced by $\tilde{\mathbf{w}}$ is simply the closure in the topology of pointwise convergence in $(\mathbb{R}^q)^T$ of the linear span of the set $\{\sigma^t\tilde{\mathbf{w}} \mid r \in T\}$.

This closed linear shift invariant subspace of $(\mathbb{R}^q)^\perp$ defines the desired behaviour, $\mathscr{B}^*_{\tilde{\mathbf{w}}}$. However, concrete representations are required. In other words, algorithms should be set up which start from a numerical specification of $\tilde{\mathbf{w}}$ and end up with a numerical specification of this most powerful unfalsified behaviour $\mathscr{B}^*_{\tilde{\mathbf{w}}}$. However, as seen in Part 1, numerical specifications of $\mathscr{B}^*_{\tilde{\mathbf{w}}}$ can be obtained in many different forms, for example as an (AR) representation $R^*_{\tilde{\mathbf{w}}}(\sigma)\ \mathbf{w} = \mathbf{0}$, or as a state space representation $\sigma\mathbf{x} = A'\mathbf{x} + B'\mathbf{v}$, $\mathbf{w} = C'\mathbf{x} + D'\mathbf{v}$. Of course, one could also ask for i/o or i/s/o representations or (assuming that $\mathscr{B}^*_{\tilde{\mathbf{w}}}$ is reachable) an (MA) representation, etc.

This paper concentrates on what can be considered to be the most sensible exact modelling questions: first, obtaining an (AR) representation for $\mathscr{B}^*_{\tilde{\mathbf{w}}}$ and second, obtaining a minimal state representation for $\mathscr{B}^*_{\tilde{\mathbf{w}}}$. The first algorithm gives moreover a shortest lag description, i.e. essentially an i/o model and the state space algorithm gives a minimal state/minimal driving input model, i.e. basically an i/s/o model. Schematically, the intention is to develop the diagram shown in Fig. 1.

It is very reasonable to view this problem as a logical 'first' identification problem: given a set of observations, the most powerful unfalsified model is required. It is perhaps an unusual identification problem, because the model is required to be first exact and second, non-stochastic.

15. FROM TIME SERIES TO (AR) MODEL

In this section the polynomial matrix $R^*_{\tilde{\mathbf{w}}}$ of Theorem 13 will be constructed starting from $\tilde{\mathbf{w}}$. In a sense this merely requires an application of Section 8 with $\mathscr{B}_t = \text{span}\ \{\text{col}(\tilde{\mathbf{w}}(\tau), \tilde{\mathbf{w}}(\tau+1), \ldots, \tilde{\mathbf{w}}(\tau+t)), \tau \in T\}$. However, the algorithms will be stated in such a way that they can be applied almost mechanically. As in Section 8, two algorithms for computing $R^*_{\tilde{\mathbf{w}}}$ will be given: one leading to a shortest lag description and one which is more general. Also, recall from Theorem 2 that in $\Sigma = \{T, \mathbb{R}^q, \mathscr{B}(R^*_{\tilde{\mathbf{w}}})\}$, the external variables can be separated componentwise into inputs (unexplained variables) and outputs (explained variables). The first algorithm also recognizes this separation.

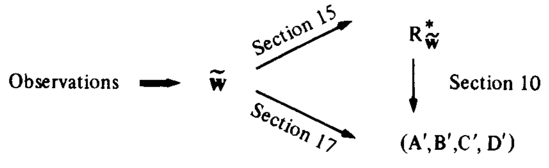


FIG. 1.

The algorithms examine the following infinite (vector) Hankel matrix formed from the data:

$$\mathscr{H}(\tilde{\mathbf{w}}) := \begin{bmatrix} \tilde{\mathbf{w}}(0) & \tilde{\mathbf{w}}(1) & \cdots & \tilde{\mathbf{w}}(t') & \cdots \\ \tilde{\mathbf{w}}(1) & \tilde{\mathbf{w}}(2) & \cdots & \tilde{\mathbf{w}}(t'+1) & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \tilde{\mathbf{w}}(t) & \tilde{\mathbf{w}}(t+1) & \cdots & \tilde{\mathbf{w}}(t+t') & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \end{bmatrix} \quad (\text{case } T = \mathbb{Z}_+)$$

or

$$\mathscr{H}(\tilde{\mathbf{w}}) := \begin{bmatrix} \cdots & \tilde{\mathbf{w}}(-1) & \tilde{\mathbf{w}}(0) & \tilde{\mathbf{w}}(1) & \cdots & \tilde{\mathbf{w}}(t') & \cdots \\ \cdots & \tilde{\mathbf{w}}(0) & \tilde{\mathbf{w}}(1) & \tilde{\mathbf{w}}(2) & \cdots & \tilde{\mathbf{w}}(t'+1) & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \cdots & \tilde{\mathbf{w}}(t-1) & \tilde{\mathbf{w}}(t) & \tilde{\mathbf{w}}(t+1) & \cdots & \tilde{\mathbf{w}}(t+t') & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{bmatrix} \quad (\text{case } T = \mathbb{Z})$$

and its truncations $\mathscr{H}_t(\tilde{\mathbf{w}})$, $t \in \mathbb{Z}_+$ consisting of the first $(t+1)q$ rows (the first $(t+1)$ block rows) of $\mathscr{H}(\tilde{\mathbf{w}})$.

In order to find the desired most powerful (AR) model, a set of generators of $\mathscr{B}^\perp(R^*_{\tilde{\mathbf{w}}})$, viewed as a submodule of $\mathbb{R}^{1\times q}[s]$ (respectively $\mathbb{R}^{1\times q}[s, s^{-1}]$), must be found. Since $\mathscr{B}_t(R^*_{\tilde{\mathbf{w}}})$ is equal to the span of the columns of $\mathscr{H}_t(\tilde{\mathbf{w}})$, the Hankel matrix actually gives us a rather concrete representation of $\mathscr{B}^\perp_t(R^*_{\tilde{\mathbf{w}}})$. Now, if t' is such that $\dim \mathscr{H}_t(\tilde{\mathbf{w}}) - \dim \mathscr{H}_{t-1}(\tilde{\mathbf{w}})$ is constant for $t \geqslant t'$, then, in order to find a set of generators for $\mathscr{B}^\perp_t(R^*_{\tilde{\mathbf{w}}})$, it suffices to find a basis for $\mathscr{B}^\perp_{t'}(R^*_{\tilde{\mathbf{w}}})$. This is a consequence of the fact that $\mathscr{B}^\perp(R^*_{\tilde{\mathbf{w}}})$ then has a set of generators of degree, at most, t'. This observation is the basis of Algorithm 3. However, the construction of a basis of the orthogonal complement of the column span of $\mathscr{H}_{t'}(\tilde{\mathbf{w}})$ can be done recursively by examining $\mathscr{H}_0(\tilde{\mathbf{w}})$, $\mathscr{H}_1(\tilde{\mathbf{w}})$, ... and using the computed orthogonal complement of the column span of $\mathscr{H}_{t'}(\tilde{\mathbf{w}})$ in order to compute that of $\mathscr{H}_{t'+1}(\tilde{\mathbf{w}})$. This is the idea underlying Algorithm 2.

Algorithm 2

Data. $\begin{cases} \tilde{\mathbf{w}}(0), \tilde{\mathbf{w}}(1), \ldots, \tilde{\mathbf{w}}(t), \ldots & (\text{case } T = \mathbb{Z}_+) \\ \ldots, \tilde{\mathbf{w}}(-1), \tilde{\mathbf{w}}(0), \tilde{\mathbf{w}}(1), \ldots, \tilde{\mathbf{w}}(t), \ldots & (\text{case } T = \mathbb{Z}). \end{cases}$

Step 1 (*Determination of the dependency vector*). The first step of the algorithm consists of sorting out the linear dependence of the consecutive rows of $\mathscr{H}(\tilde{\mathbf{w}})$. Check, starting from the top row, which of the (infinite) rows of $\mathscr{H}(\tilde{\mathbf{w}})$ are linearly dependent on the preceding rows. This yields an infinite dependency column vector $\mathbf{d}$ consisting of $*$s and

•s. A $*$ in the ith spot signifies that the ith row is linearly independent of the preceding rows, while a • signifies dependence.

Step 2 (*Determination of the input and output variables*). Observe that, because of the Hankel structure of $\mathscr{H}(\tilde{\mathbf{w}})$, if $\mathbf{d}$ has a • in its ith row, then it will have a • in its $(i + nq)$th row for all $n \in \mathbb{Z}_+$. Now fix $j \in \mathbf{q}$. If there is a $*$ in the $(j + nq)$th row of $\mathbf{d}$ *for all* $n \in \mathbb{Z}_+$ then the jth component of w in $R^*_{\tilde{\mathbf{w}}}(\sigma)\mathbf{w} = \mathbf{0}$ is an input, otherwise it is an output. Let $1 \leqslant j_1 \leqslant j_2 \leqslant \cdots \leqslant j_m \leqslant q$ be the indices of the input variables and $1 \leqslant j'_1 \leqslant j'_2 \leqslant \cdots \leqslant j'_p \leqslant q$ be the indices of the output variables thus obtained.

Step 3 (*Determination of the recursions*). Assume that the ith variable, $i \in \mathbf{q}$, is an output. Let $l_i := \min\{k \in \mathbb{Z}_+ |$ the $(i + kq)$th element of $\mathbf{d}$ is a $\bullet\}$. Then the $(i + l_i q)$th row of $\mathscr{H}(\tilde{\mathbf{w}})$ can be written as a linear combination of the rows preceding it, i.e. there exist $a_{i,0}, a_{i,1}, \ldots, a_{i,l_i} \in \mathbb{R}^{1\times q}$ (with a_{i,l_i} of the form $a_{i,l_i} = [a^1_{i,l_i}, a^2_{i,l_i}, \ldots, a^{i-1}_{i,l_i}, 0, \ldots, 0]$) such that

$$\tilde{\mathbf{w}}_i(t + l_i) = \sum_{k=0}^{l_i} a_{i,k}\tilde{\mathbf{w}}(t + k).$$

(Note that one could also demand that the kth component of $a_{i,j}$ is zero whenever $j \geqslant l_i$. This would actually determine the $a_{i,j}$s uniquely.)

Step 4 (*Specification of* $R^*_{\tilde{\mathbf{w}}}$). Let

$$r_i(s) := [0, \ldots, 0, \underset{\substack{\uparrow \\ i\text{th entry}}}{1}, 0, \ldots, 0]s^{l_i} - \sum_{k=0}^{l_i} a_{i,k}s^k, i \in \{j'_1.j'_2, \ldots, j'_p\}.$$

Then $R^*_{\tilde{\mathbf{w}}} := \mathrm{col}(r_1, r_2, \ldots, r_p)$ is the desired polynomial matrix. This is stated formally as follows.

Theorem 14. Let $\tilde{\mathbf{w}}: T \to \mathbb{R}^q$, $T = \mathbb{Z}_+$ or $\mathbb{Z}$, be an observed time series and let $R^*_{\tilde{\mathbf{w}}}$ be as defined in Algorithm 2. Then

$$R^*_{\tilde{\mathbf{w}}}(\sigma)\mathbf{w} = \mathbf{0}$$

defines the most powerful unfalsified (AR) model for $\tilde{\mathbf{w}}$. In fact, it is a shortest lag i/o description, with the $j_1, j_2, \ldots, j_m$th variables as inputs and the $j'_1, j'_2, \ldots, j'_p$th variables as outputs. The integers $l_1, l_2, \ldots, l_p$ determine, after rearrangement in non-decreasing order, the shortest lag structure of $\mathscr{B}(R^*_{\tilde{\mathbf{w}}})$.

Proof. See Appendix P.

Remarks. (1) As seen in Part 1, which variables are inputs and which outputs is certainly not uniquely specified in a given dynamical system, not even in a (AR) model, even though in that case a separation into inputs and outputs is always possible. Indeed, a lowly (non-zero) ohmic resistor already provides an example of a situation where there is a choice. Algorithm 2 seems to give a unique set of input variables. However, reordering of the components of $\tilde{\mathbf{w}}$ will in general lead to a different set of input and output variables in Step 2, which was clearly prejudiced by the order in which the elements of $w \in \mathbb{R}^q$ were organized. Actually, whether the distinction between inputs and outputs in a model should be emphasized is, to some extent, a matter of opinion. In the remaining algorithms in this paper, the distinction between inputs and outputs will not be made. Recall however Section 6 (iii), where the very strong properties of inputs in an i/o model are spelled out.

(2) Algorithm 2 is closely related to the work of Guidorzi (1981) and similar work which has appeared in the literature, in an input/output setting. The contribution of Algorithm 2 is the following. First, with the search of the most powerful unfalsified model what the algorithm exactly computes is specified in a mathematically unambiguous way and second, the algorithm automatically recognizes what the inputs and the outputs are.

Algorithm 2 is actually an implementation to the case at hand of the procedure treated in Theorem 8. There is also an analogue of Theorem 7 which is now stated in a very concrete way:

Algorithm 3

Data. $\begin{cases} \tilde{\mathbf{w}}(0), \tilde{\mathbf{w}}(1), \ldots, \tilde{\mathbf{w}}(t), \ldots & (\text{case } T = \mathbb{Z}_+) \\ \ldots, \tilde{\mathbf{w}}(-1), \tilde{\mathbf{w}}(0), \tilde{\mathbf{w}}(1), \ldots, \tilde{\mathbf{w}}(t), \ldots & (\text{case } T = \mathbb{Z}). \end{cases}$

Step 1 (*Determination of the lag*). From the Hankel structure it follows immediately that for $t \in \mathbb{Z}_+$

$$\rho_t := \mathrm{rank}\,\mathscr{H}_t(\tilde{\mathbf{w}}) - \mathrm{rank}\,\mathscr{H}_{t-1}(\tilde{\mathbf{w}});$$
$$\rho_0 := \mathrm{rank}\,\mathscr{H}_0(\tilde{\mathbf{w}})$$

is a non-increasing sequence of non-negative integers. Now compute a t' such that $\rho_t = \rho_{t'}$ for $t \geqslant t'$.

Step 2 (*Determination of the linear dependence*). Compute row vectors $r_1, r_2, \ldots, r_g \in \mathbb{R}^{1\times(t'+1)q}$ such that they span the orthogonal complement of the columns of $\mathscr{H}_{t'}(\tilde{\mathbf{w}})$. (There are several ways of doing this. For example, a maximum rank submatrix M of $\mathscr{H}_{t'}(\tilde{\mathbf{w}})$ could be determined. Say that M is formed by the $(l_1, l_2, \ldots, l_r)$th rows and the $(k_1, k_2, \ldots, k_r)$th columns of $\mathscr{H}_{t'}(\tilde{\mathbf{w}})$. Now write the other rows of the matrix formed by the $(k_1, k_2, \ldots, k_r)$th columns of $\mathscr{H}_t(\tilde{\mathbf{w}})$ as a linear combination of the rows of M. The coefficients of these linear combination determine this orthogonal complement. Clearly, this computation may be further refined and simplified by exploiting the Hankel structure of $\mathscr{H}_t(\tilde{\mathbf{w}})$. The coefficients which express this linear dependence determine $r_1, r_2, \ldots, r_g$ in an obvious way.

Step 3 (*Specification of* $R^*_{\tilde{\mathbf{w}}}$). Now write r_i, $i \in \mathbf{g}$, as $r_i := [r_{i,0}, r_{i,1}, \ldots, r_{i,t'}]$ with $r_{i,j} \in \mathbb{R}^{1 \times q}$, define

$$\tilde{r}_i(s) := \sum_{k=0}^{t'} r'_{i,k} s^k, \quad \text{and} \quad R^*_{\tilde{\mathbf{w}}} := \mathrm{col}(\tilde{r}_1, \tilde{r}_2, \ldots, \tilde{r}_g).$$

This yields the result below.

Theorem 15. Let $\tilde{\mathbf{w}}: T \to \mathbb{R}^q$, $T = \mathbb{Z}_+$ or $\mathbb{Z}$, be an observed time series and let $R^*_{\tilde{\mathbf{w}}}$ be as defined in Algorithm 2. Then

$$R^*_{\tilde{\mathbf{w}}}(\sigma)\mathbf{w} = \mathbf{0}$$

defines the most powerful unfalsified (AR) model for $\tilde{\mathbf{w}}$.

Proof. See Appendix P.

Example 5. The purpose of this first example is simply to illustrate, by means of an artificial set of numbers, the procedure described in Algorithms 2 and 3.

Let $q = 2$, $T = \mathbb{Z}_+$, and

$$\tilde{\mathbf{w}}_1 = \{1,0,1,0,0,1,0,0,0,1,0,0,0,0,1,\ldots\}$$
$$\tilde{\mathbf{w}}_2 = \{0,1,1,2,2,2,3,3,3,3,4,4,4,4,4,\ldots\}.$$

The Hankel matrix and the dependency vector become

$$\mathscr{H}(\tilde{\mathbf{w}}) = \begin{bmatrix} 1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & \cdots \\ 0 & 1 & 1 & 2 & 2 & 2 & 3 & 3 & 3 & 3 & 4 & 4 & 4 & 4 & 4 & \cdots \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & \cdots \\ 1 & 1 & 2 & 2 & 2 & 3 & 3 & 3 & 3 & 4 & 4 & 4 & 4 & 4 & 5 & \cdots \\ 1 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & \cdots \\ 1 & 2 & 2 & 2 & 3 & 3 & 3 & 3 & 4 & 4 & 4 & 4 & 4 & 5 & 5 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{bmatrix}$$

$$d = \begin{bmatrix} * \\ * \\ * \\ \bullet \\ * \\ \bullet \\ \vdots \end{bmatrix}.$$

Use Algorithm 2. This shows that w_1 is an input and w_2 is an output. Also l_2 equals 1 and the recursion is obtained by writing the fourth row of $\mathscr{H}(\tilde{\mathbf{w}})$ as a linear combination of the preceding ones:

$$\tilde{\mathbf{w}}_2(t+1) = \tilde{\mathbf{w}}_1(t) + \tilde{\mathbf{w}}_2(t)$$

which yields

$$\sigma\mathbf{w}_2 - \mathbf{w}_1 - \mathbf{w}_2 = \mathbf{0}$$

as the most powerful unfalsified (AR) model for $\tilde{\mathbf{w}}$.

In order to demonstrate the use of Algorithm 3, choose $t' = 2$ and let the submatrix M be

$$M = \underbrace{\begin{bmatrix} 0 & 0 & 0 & 1 \\ 3 & 3 & 3 & 3 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}}_{\substack{\text{columns} \\ 7\ \ 8\ \ 9\ \ 10 \\ \text{of } \mathscr{H}(\tilde{\mathbf{w}})}} \left.\vphantom{\begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}}\right\} \text{rows} \begin{matrix} 1 \\ 2 \\ 3 \\ 5 \end{matrix} \text{ of } \mathscr{H}(\tilde{\mathbf{w}}).$$

Now write the 4th and 6th rows of the matrix formed by the 7, 8, 9 and 10th columns of $\mathscr{H}_2(\tilde{\mathbf{w}})$ as a linear combination of the rows of M. This yields $r_1 = [-1, -1, 0, 1, 0, 0]$ and $r_2 = [-1, -1, -1, 0, 0, 1]$ as a basis for $\mathscr{B}_2^{\perp}(R^*_{\tilde{\mathbf{w}}})$ and leads to the (AR) model

$$\sigma\mathbf{w}_2 - \mathbf{w}_1 - \mathbf{w}_2 = \mathbf{0}$$
$$\sigma^2\mathbf{w}_2 - \sigma\mathbf{w}_1 - \mathbf{w}_1 + \mathbf{w}_2 = \mathbf{0}.$$

Note that the first equation implies the second, and hence the same model is obtained as with the first algorithm. This example will be discussed again in Section 17.

Remark: *Generalization to a finite set of observed time series.* Consider the case that k time series $\tilde{\mathbf{w}}_1, \tilde{\mathbf{w}}_2, \ldots, \tilde{\mathbf{w}}_k$ are observed and that the most powerful unfalsified (AR) model is sought, based on the measurements $Z = \{\tilde{\mathbf{w}}_1, \tilde{\mathbf{w}}_2, \ldots, \tilde{\mathbf{w}}_k\}$. In other words, the algorithm is extended to the case that the measurement set $\mathscr{Z}$ consists of all finite sets. In order to treat this case, it suffices to use Algorithms 2 or 3 verbatim on the (block) Hankel matrix:

$$\mathscr{H}(\tilde{\mathbf{W}}) := \begin{bmatrix} \tilde{\mathbf{W}}(0) & \tilde{\mathbf{W}}(1) & \cdots & \tilde{\mathbf{W}}(t') & \cdots \\ \tilde{\mathbf{W}}(1) & \tilde{\mathbf{W}}(2) & \cdots & \tilde{\mathbf{W}}(t'+1) & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \tilde{\mathbf{W}}(t) & \tilde{\mathbf{W}}(t+1) & \cdots & \tilde{\mathbf{W}}(t+t') & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \end{bmatrix} \quad (\text{case } T = \mathbb{Z}_+)$$

with $\tilde{\mathbf{W}}: T \to \mathbb{R}^{q \times k}$ defined by $\tilde{\mathbf{W}}(t) := [\tilde{\mathbf{w}}_1(t), \tilde{\mathbf{w}}_2(t) \ldots \tilde{\mathbf{w}}_k(t)]$ and with the obvious modification for the case $T = \mathbb{Z}$. This generalization will be used in the next example.

Example 6 (*Impulse response measurements*). The purpose of this second example is to demonstrate that this procedure recovers the results of Kalman's realization theory (Kalman *et al.*, 1969) as a special case.

Let $T = \mathbb{Z}$ and assume that responses to unit impulses are observed in the 1st, 2nd, ..., mth channel of an i/o system with m input and p output channels. Let $\tilde{\mathbf{g}}_1, \tilde{\mathbf{g}}_2, \ldots, \tilde{\mathbf{g}}_m: T \to \mathbb{R}^p$ denote the corresponding impulse responses. Define $q = m + 0$ and $\tilde{\mathbf{w}}_i: T \to \mathbb{R}^q$, $i \in \mathbf{m}$, by $\tilde{\mathbf{w}}_i(t) := \mathrm{col}\,(\tilde{\mathbf{e}}_i, \tilde{\mathbf{g}}_i)$ with $\tilde{\mathbf{e}}_i: T \to \mathbb{R}^m$ given by

$$\tilde{\mathbf{e}}_i(t) = \begin{cases} e_i & \text{for } t = 0 \\ 0 & \text{for } t \neq 0 \end{cases}$$

where $e_i := \mathrm{col}(0, \ldots, 0, 1, 0, \ldots, 0)$. One has the measurement $Z = \{\tilde{w}_1, \tilde{w}_2, \ldots, \tilde{w}_m\}$ and the question is to see if the input/output structure and the dimension of the minimal state space can be predicted from Algorithm 2.

Denote $\tilde{\mathbf{G}}(t) := [\tilde{\mathbf{g}}_1(t), \tilde{\mathbf{g}}_2(t), \ldots, \tilde{\mathbf{g}}_m(t)]$ and $\tilde{\mathbf{W}}(t) := [\tilde{\mathbf{w}}_1(t), \tilde{\mathbf{w}}_2(t), \ldots, \tilde{\mathbf{w}}_m(t)]$. The relevant block Hankel matrix becomes:

$$\mathscr{H}(\tilde{\mathbf{W}}) := \begin{bmatrix} \cdots & 0 & 0 & I & 0 & \cdots \\ \cdots & 0 & 0 & \tilde{\mathbf{G}}(0) & \tilde{\mathbf{G}}(1) & \cdots \\ \cdots & 0 & I & 0 & 0 & \cdots \\ \cdots & 0 & \tilde{\mathbf{G}}(0) & \tilde{\mathbf{G}}(1) & \tilde{\mathbf{G}}(2) & \cdots \\ \cdots & I & 0 & 0 & 0 & \cdots \\ \cdots & \tilde{\mathbf{G}}(0) & \tilde{\mathbf{G}}(1) & \tilde{\mathbf{G}}(2) & \tilde{\mathbf{G}}(3) & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{bmatrix}.$$

Consider now $\mathrm{rank}\,\mathscr{H}_t(\tilde{\mathbf{W}}) = \dim \mathscr{B}_t(R^*_{\tilde{\mathbf{w}}})$. Clearly it equals $(t+1)m + \mathrm{rank}\,\mathscr{H}_t(\tilde{\mathbf{G}})$ with

$$\mathscr{H}(\tilde{\mathbf{G}}) := \begin{bmatrix} \tilde{\mathbf{G}}(1) & \tilde{\mathbf{G}}(2) & \cdots & \tilde{\mathbf{G}}(t') & \cdots \\ \tilde{\mathbf{G}}(2) & \tilde{\mathbf{G}}(3) & \cdots & \tilde{\mathbf{G}}(t'+1) & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \tilde{\mathbf{G}}(t) & \tilde{\mathbf{G}}(t+1) & \cdots & \tilde{\mathbf{G}}(t+t'-1) & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \end{bmatrix}.$$

If $\mathrm{rank}\,\mathscr{H}(\tilde{\mathbf{G}}) = \infty$, then the most powerful (AR) model unfalsified by $Z = \{\tilde{\mathbf{w}}_1, \tilde{\mathbf{w}}_2, \ldots, \tilde{\mathbf{w}}_m\}$ will declare that some of the output channels actually produce unrestricted signals (in the sense that its restrictions, if any, would only have become evident at $t = \infty$). In this context of complete systems, these outputs channels are declared to be input channels.

Assume henceforth that rank $\mathscr{H}(\tilde{\mathbf{G}}) =: n < \infty$. Now use Algorithm 2. The dependency vector $\mathbf{d}$ immediately shows that the 1st, 2nd, ..., mth channels are indeed input channels, while the $(m+1)$th, ..., $(m+p)$th channels are indeed output channels. Furthermore, by Theorem 6, the dimension of the minimal state space, n^*, equals $\lim_{t\to\infty} (\dim \mathscr{H}_t(R_{\mathbf{w}}^*) - (t+1)m) = \lim_{t\to\infty} \operatorname{rank} \mathscr{H}_t(\tilde{\mathbf{G}})$, showing that n^* is indeed equal to what would have been predicted by classical realization theory.

In order to obtain the relevant $R_{\mathbf{w}}^*$, assume that rank $\mathscr{H}(\tilde{\mathbf{G}}) = \operatorname{rank} \mathscr{H}(\tilde{\mathbf{G}}) = n$. Now consider the following truncation:

$\mathscr{H}_{t'}(\tilde{\mathbf{w}}) :=$

$$\begin{bmatrix} \cdots & 0 & \cdots & 0 & I & 0 & \cdots & 0 & \cdots \\ \cdots & 0 & \cdots & 0 & \tilde{\mathbf{G}}(0) & \tilde{\mathbf{G}}(1) & \cdots & \tilde{\mathbf{G}}(t') & \cdots \\ \cdots & 0 & \cdots & I & 0 & 0 & \cdots & 0 & \cdots \\ \cdots & 0 & \cdots & \tilde{\mathbf{G}}(0) & \tilde{\mathbf{G}}(1) & \tilde{\mathbf{G}}(2) & \cdots & \tilde{\mathbf{G}}(t'+1) & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \cdots & I & \cdots & 0 & 0 & 0 & \cdots & 0 & \cdots \\ \cdots & \tilde{\mathbf{G}}(0) & \cdots & \tilde{\mathbf{G}}(t'-1) & \tilde{\mathbf{G}}(t') & \tilde{\mathbf{G}}(t'+1) & \cdots & \tilde{\mathbf{G}}(2t') & \cdots \end{bmatrix}$$

Clearly its rank equals $m(t+1)+n$.

Now apply Algorithm 3. One method for doing this is by first solving the equations

$$P_0\tilde{\mathbf{G}}(t) + P_1\tilde{\mathbf{G}}(t+1) + \cdots + P_{t'}\tilde{\mathbf{G}}(t+t') = 0 \quad t \in \mathbf{t}'$$

for $P_t \in \mathbb{R}^{(p(t'+1)-n)\times p}$, $t = 0, 1, \ldots, t'$, such that the matrix $M = [P_0 \;\; P_1 \ldots P_{t'}]$ has full rank $p(t'+1)-n$. The matrices $Q_0, Q_1, \ldots, Q_{t'}$ can then be determined recursively by

$$Q_t = P_t\tilde{\mathbf{G}}(0) + P_{t+1}\tilde{\mathbf{G}}(1) + \cdots + P_{t'}\tilde{\mathbf{G}}(t'-t) \quad t = 0, 1, \ldots, t'.$$

This yields $R_{\mathbf{w}}^*(s) = [P(s) \vdots -Q(s)]$, with $P(s) = P_{t'}s^{t'} + \cdots + P_0$ and $Q(s) = Q_{t'}s^{t'} + \ldots + Q_0$, as the desired most powerful unfalsified (AR) model.

This example will be reconsidered in Section 17 in order to compute a minimal state model, thus obtaining a complete and convincing generalization of the classical impulse response realization procedure.

Note also that in the impulse response case when $T = \mathbb{Z}_+$, the most powerful unfalsified model will be described by the equation $\sigma\mathbf{u} = \mathbf{0}$ together with the model obtained in the case $T = \mathbb{Z}$. Note that one may now as well set $Q_1 = Q_2 = \cdots = Q_{t'} = 0$. In other words, the most powerful model will also explain the "input". Formally, in this setting, it should therefore also be considered as an output. In fact, in the end the whole system becomes an autonomous one.

A few comments are in order.

First, note that the case in which the impulse response $\tilde{\mathbf{G}}(0)$, $\tilde{\mathbf{G}}(1), \ldots, \tilde{\mathbf{G}}(t), \ldots$, leading up to the model $\mathbf{y}(t) = \sum_{t' \leqslant t} \tilde{\mathbf{G}}(t-t')\mathbf{u}(t')$, is measured cannot logically be viewed as an identification problem, since with the impulse response the model is basically already given. Note, however that this procedure will yield the convolution sum as the most powerful unfalsified model only if rank $\mathscr{H}(\tilde{\mathbf{G}}) < \infty$. If this is the case, then (assume $T = \mathbb{Z}$) the restriction of the most powerful behaviour to sequences $\mathbf{w}$ vanishing on a half line $(-\infty, t]$ will give us exactly the graph of the convolution, whereas in the case rank $\mathscr{H}(\tilde{\mathbf{G}}) = \infty$ there will be additional elements in the restriction of the most powerful behaviour. Thus the closure of the graph of the convolution, restricted to sequences vanishing on a half line, will give the graph of the convolution iff rank $\mathscr{H}(\tilde{\mathbf{G}}) < \infty$.

Second, consider some peculiarities of the case $T = \mathbb{Z}_+$. If rank $\mathscr{H}(\tilde{\mathbf{G}}) < \infty$ then an autonomous system is obtained as a model. The relation between the impulse variables w_1 and the response variables w_2 will be of the type

$$\mathbf{w}_2(t+l) + P_{l-1}\mathbf{w}_2(t+l-1) + \cdots + P_0\mathbf{w}_2(t) = Q_l\mathbf{w}_1(t+l) + Q_{l-1}\mathbf{w}_1(t+l-1) + \cdots + Q_0\mathbf{w}_1(t)$$

but with a certain arbitrariness in the Qs. This can be explained as follows. Even though the most powerful unfalsified model is unique, there are many *transfer functions* with a given set of outputs corresponding to zero input but arbitrary initial conditions. If it is required that the impulse response corresponding to *zero* initial condition yields the model then the Qs which yield the convolution sum are obtained.

Third, note that in the case $T = \mathbb{Z}_+$, rank $\mathscr{H}(\tilde{\mathbf{G}}) = \infty$ will imply that (some of) the impulse variables w_1 will be explained by the variables w_2. This results in the fact that in the most powerful unfalsified model it may occur that the impulse variables are the outputs and the response variables are the inputs (a phenomenon which cannot happen in the case $T = \mathbb{Z}$).

16. CONSTRUCTION OF STATE SPACE MODELS: THE EFFECT OF TRUNCATION

This section is basically a continuation of Section 9 where the problem of passing from the behaviour $\mathscr{B}$ (or $\mathscr{B}^\perp$) to a minimal state space representation of it was discussed. One difficulty with the results in Section 9 is that the spaces $\mathscr{B}^+$, $\mathscr{B}^0$, $\mathscr{B}^1$, $\mathscr{N}^+$, $\mathscr{N}^0$, and $\mathscr{N}^1$ are infinite dimensional and hence that the calculations also appear to be infinite dimensional. However, since in the end everything must reduce to finite dimensional linear algebra it seems plausible that there should exist modified algorithms which avoid infinite dimensional spaces altogether. This is, fortunately, possible by examining truncations of $\mathscr{B}$ to a finite but sufficiently long time interval and by considering only the elements of $\mathscr{N} = \mathscr{B}^\perp$ of a finite but sufficiently high degree.

Of course, the same concern for obtaining algorithms which are based on truncations and/or which are recursive in the data also exist in the classical impulse response realization problem. In fact, many efficient and elegant algorithms based on such ideas have appeared. These will not be reviewed here. The element which is new in this approach: how to identify simultaneously the state space and the driving input space will be emphasized. Efficient implementation of the mathematical constructions and the very important generalization to the finite time case will be pursued elsewhere.

The three basic ideas which yield the realization procedure put forward in this section are: first, truncation; second, isomorphism—when the truncation interval is sufficiently large, no information about the behaviour is lost in the truncation process; and third, that, again when the truncation interval is sufficiently large, the relevant truncated spaces may be computed from the truncated behaviour.

16.1 *Truncation*

Let $\mathscr{B} \in \mathbf{L}$ and consider $\mathscr{B}^+ := \pi^+\mathscr{B}$, $\mathscr{B}^0 := \bigcap_{t\in\mathbb{Z}+} (\sigma^*)^{-t}\mathscr{B}^+$ and $\mathscr{B}^1 := \mathscr{B}^0 \cap \ker \pi^0 = \sigma^*\mathscr{B}^0$ as studied in Section 9. Define their truncations as follows $\mathscr{B}_t := \pi^t\mathscr{B}^+, (\mathscr{B}^0)_t := \pi^t\mathscr{B}^0$ and $(\mathscr{B}^1)_t := \pi^t\mathscr{B}^1$. It follows immediately that there exist maps M_1^t and M_2^t such that the diagram shown in Fig. 2 commutes. Note that S_{t+1}^1 and S_t^0 are automatically surjections.

16.2 *Isomorphism*

If S_{t+q}^1 and S_t^0 are bijections then the construction of (M_1, M_2) required in the primal version of Theorem 9 can be replaced by the construction of (M_1^t, M_2^t). It will be shown that this holds for t sufficiently large.

16.3 *Computability*

In general it will not be possible to compute $(\mathscr{B}^0)_t$ and $(\mathscr{B}^1)_t$ from $\mathscr{B}_t$. However, the following spaces are defined solely in terms of $\mathscr{B}_t$:

$$\bar{\mathscr{B}}_t^0 := \bigcap_{0 \leqslant t' \leqslant t} (\sigma_t^*)^{-t'}\mathscr{B}_t$$

$$\bar{\mathscr{B}}_t^1 := \bar{\mathscr{B}}_t^0 \cap \ker \pi^0$$

Here $\sigma_t^*: \mathscr{L}_t \to \mathscr{L}_t$ is defined by $\sigma_t^* := \pi^t\sigma^*$: it is the forward shift followed by filling in zeros and truncation. In other words,

$$\bar{\mathscr{B}}_t^0 = \{\mathrm{w} \in \mathscr{L}_t \,|\, (\sigma_t^*)^{t'}\mathrm{w} \in \mathscr{B}_t \,\forall\, 0 \leqslant t' \leqslant t\}$$

Note the significance of the spaces $\mathscr{B}_t$, $(\mathscr{B}^0)_t$, $\bar{\mathscr{B}}_t^0$, $(\mathscr{B}^1)_t$, and $\bar{\mathscr{B}}_t^1$. The space $\mathscr{B}_t$ represents all sequences of length $t+1$ compatible with the behaviour; $(\mathscr{B}^0)_t$ represents those which could have been preceded with any number of zeros and hence those which started in the zero state; $(\mathscr{B}^1)_t$ represents those which in addition are zero at time 0; $\bar{\mathscr{B}}_t^0$ represents those sequences of length $t+1$ which, when preceded with any number of zeros, and then truncated to an interval of length $t+1$, will generate elements of $\mathscr{B}_t$; $\bar{\mathscr{B}}_t^1$ represents those which in addition are zero at time 0. This yields the lattice diagram shown in Fig. 3.

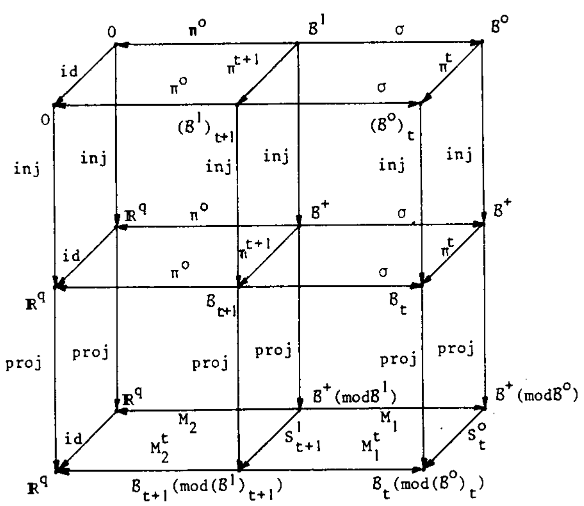


FIG. 2.

The following inclusions hold: $\bar{\mathscr{B}}_t^1 \subset \ker \pi^0$ and $\sigma\bar{\mathscr{B}}_{t+1}^1 \subset \bar{\mathscr{B}}_t^0$. The latter inclusion follows immediately from the definitions, $\sigma\mathscr{B}_{t+1} \subset \mathscr{B}_t$, and the observation that $\sigma\sigma_{t+1}^* = \sigma_t^*\sigma + \pi^0$. It follows from these inclusions that there exist maps $\bar{M}_1^t$ and $\bar{M}_2^t$ such that the important *(primal) truncated structure diagram* (Fig. 4) commutes.

It turns out that $\bar{\mathscr{B}}_t^0$ and $\bar{\mathscr{B}}_t^1$ are equal to $(\mathscr{B}^0)_t$ and $(\mathscr{B}^1)_t$, respectively, for t sufficiently large.

Proposition 16 (*Primal version*). Let $\mathscr{B} \in \mathrm{L}$. Then the following conditions on $t \in \mathbb{Z}_+$ are equivalent:

(i) $\dim \mathscr{B}_{t'+1} - \dim \mathscr{B}_{t'} = \dim \mathscr{B}_{t+1} - \dim \mathscr{B}_t$ for $t' \geqslant t$ (i.e. in terms of the structure indices of Section 7, $t+1 \geqslant t^* := \min_{t \in \mathbb{Z}_+} \{t \,|\, \rho_t = \rho_\infty\}$);

(ii) $\dim \mathscr{B}_t\,(\mathrm{mod}\,(\mathscr{B}^0)_t) = \dim \mathscr{B}(\mathrm{mod}\,\mathscr{B}^0)$;

(iii) $\dim \mathscr{B}_t\,(\mathrm{mod}\,(\mathscr{B}^1)_t) = \dim \mathscr{B}(\mathrm{mod}\,\mathscr{B}^1)$;

(iv) $\bar{\mathscr{B}}_{t+1}^0 = (\mathscr{B}^0)_{t+1}$;

(v) $\bar{\mathscr{B}}_{t+1}^1 = (\mathscr{B}^1)_{t+1}$.

Proof. See Appendix P.

Example 7. This dazzling line-up of subspaces and their differences will be illustrated by means of two simple examples.

Consider first the trivial system consisting of a free input: $\mathscr{B} = (\mathbb{R})^{\mathbb{Z}}$. Then $\mathscr{B}^+ = \mathscr{B}^0 = (\mathbb{R})^{\mathbb{Z}_+}$ and $\mathscr{B}^1 = \{\mathrm{w}: \mathbb{Z}_+ \to \mathbb{R} \,|\, \mathrm{w}(0) = 0\}$. Further, $\bar{\mathscr{B}}_t^0 = \mathscr{B}_t$ for all t and $\bar{\mathscr{B}}_t^1 = \{\mathrm{w} \in \mathscr{L}_t \,|\, \mathrm{w}(0) = 0\}$. It is easy to see that $\min_{t \in \mathbb{Z}_+} \{t \,|\, \rho_t = \rho_\infty\} =: t^* = 0$. Proposition 16 is easily verified.

Next consider the system on $T = \mathbb{Z}_+$ described by the (AR) equation $\mathrm{w}(t+n) = 0$, i.e. $\mathrm{w}(0), \mathrm{w}(1), \ldots, \mathrm{w}(n-1)$ are arbitrary real numbers and $\mathrm{w}(t) = 0$ for $t \geqslant n > 0$. Let $\mathscr{B}^+(n)$ denote the behaviour of this system. Then $\mathscr{B}^+ = \mathscr{B}^+(n)$, $\mathscr{B}^0 = \mathscr{B}^1 = \{0\}$. Hence $(\mathscr{B}^0)_t = (\mathscr{B}^1)_t = \{0\}$. Further $\mathscr{B}_t = \mathscr{L}_t$ for $t \leqslant n-1$ and $\mathscr{B}_t = \{\mathrm{w} \in \mathscr{L}_t \,|\, \mathrm{w}(t') = 0 \text{ for } t' \geqslant n\}$ for $t \geqslant n$. It follows that $\min_{t \in \mathbb{Z}_+} \{t \,|\, \rho_t = \rho_\infty\} =: t^* = n$. A simple calculation shows further that $\bar{\mathscr{B}}_t^0 = \mathscr{B}_t$ for $t \leqslant n-1$ and $\bar{\mathscr{B}}_t^0 = \{0\}$ for $t > n$. Similarly, $\bar{\mathscr{B}}_t^1 = \{\mathrm{w} \in \mathscr{L}_t \,|\, \mathrm{w}(0) = 0\}$ for $t \leqslant n-1$ and $\bar{\mathscr{B}}_t^1 = \{0\}$ for $t \geqslant n$. Proposition 16 is again easily verified.

It is the front level of the truncated structure diagram with $t \geqslant t^*$ which is of particular interest. It is redrawn here for emphasis (Fig. 5). Note that $\bar{M}_1^t$ and $\bar{M}_2^t$ are well defined even when $t < t^*$. This fact is promising for partial realization questions and recursive implementation of these algorithms. This will be briefly discussed at the end of this section.

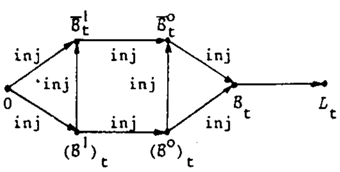


FIG. 3. Lattice diagram.

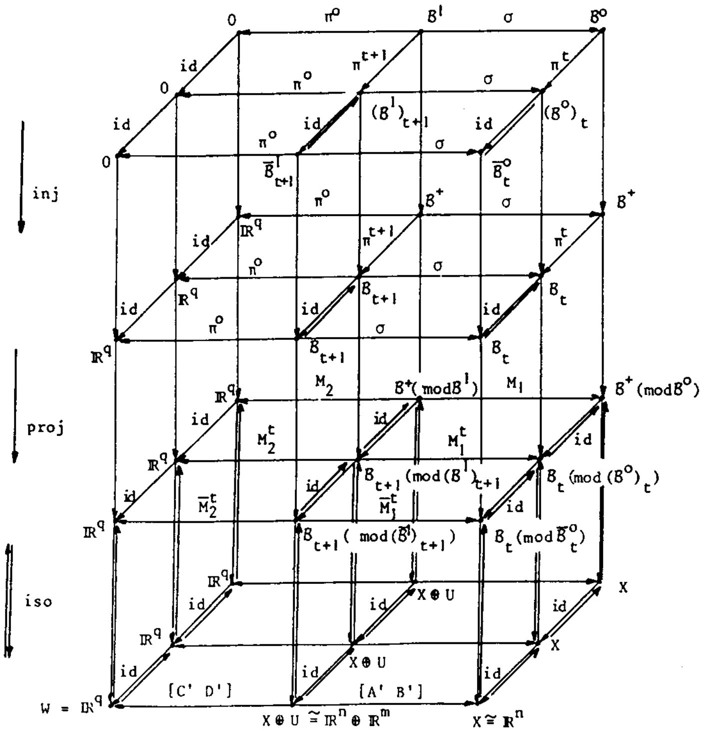


FIG. 4. Primal truncated structure diagram.

Theorem 17 (*Primal version*). Let $\mathscr{B} \in \mathbf{L}$ and let $t^* := \min_{t \in \mathbb{Z}_+} \{t \mid \rho_t = \rho_\infty\}$. Now assume that $t \geqslant t^*$. Then there exist linear maps $\bar{M}_1^t$ and $\bar{M}_2^t$ such that the truncated structure diagram commutes. The map $M = (\bar{M}_1^t, \bar{M}_2^t)$: $\mathscr{B}_{t+1} \pmod{\bar{\mathscr{B}}_{t+1}^1} \to \mathscr{B}_t \pmod{\bar{\mathscr{B}}_t^0} \times \mathbb{R}^q$ yields, identically as in the primal version of Theorem 9, a state space representation $\Sigma_s(A', B', C', D')$ of $\mathscr{B}$ with a minimal number of states and driving inputs.

Proof. See Appendix P.

Remark. There is one more refinement which may be introduced in the truncated diagram and which will be used in the next section. It may be explained as follows. In the truncated structure diagram $\bar{\mathscr{B}}_{t+1}^1$ and $\bar{\mathscr{B}}_t$ have been considered so that the shift $\sigma: \bar{\mathscr{B}}_{t+1}^1 \to \bar{\mathscr{B}}_t$ will be well defined. However, if $t \geqslant t_1^*$, since $\dim\ \mathscr{B}_t \pmod{\bar{\mathscr{B}}_t^1} = \dim \mathscr{B}_{t+1} \pmod{\bar{\mathscr{B}}_{t+1}^1}$, then $\mathscr{B}_t \pmod{\bar{\mathscr{B}}_t^1} \cong \mathscr{B}_{t+1} \pmod{\bar{\mathscr{B}}_{t+1}^1}$. So, provided the action of the shift as a map $\mathscr{B}_t \pmod{\bar{\mathscr{B}}_t^1} \to \mathscr{B}_t \pmod{\bar{\mathscr{B}}_t^0}$ can be followed, no examination of $\mathscr{B}_{t+1}$ and $\bar{\mathscr{B}}_{t+1}^1$ is needed. This remark is summarized in the commutative diagram (Fig. 6), where the map $M = (M_1, M_2)$ can serve as a means of obtaining a realization analogously as in Theorem 17.

The previous discussion will now be dualized, skipping many of the details. Let $\mathscr{B} \in \mathbf{L}$ and consider $\mathscr{N} := \mathscr{B}^\perp$, $\mathscr{N}^+ := (\mathscr{B}^+)^\perp$, $\mathscr{N}^0 = (\mathscr{B}^0)^\perp$ and $\mathscr{N}^1 = (\mathscr{B}^1)^\perp$, as studied in Section 9. Consider also $\mathscr{N}_t := \mathscr{N}^+ \cap \mathscr{L}_t^*$, $(\mathscr{N}^0)_t := \mathscr{N}^0 \cap \mathscr{L}_t^*$ and $(\mathscr{N}^1)_t := \mathscr{N}^1 \cap \mathscr{L}_t^*$, and $\bar{\mathscr{N}}_t^0 := \sum_{0 \leqslant t' \leqslant t} (s^*)^{t'} \mathscr{N}_t$, $\bar{\mathscr{N}}_t^1 := \bar{\mathscr{N}}_t^0 + \mathbb{R}^{1 \times q}[0]$. Note that the lattice diagram holds (Fig. 7).

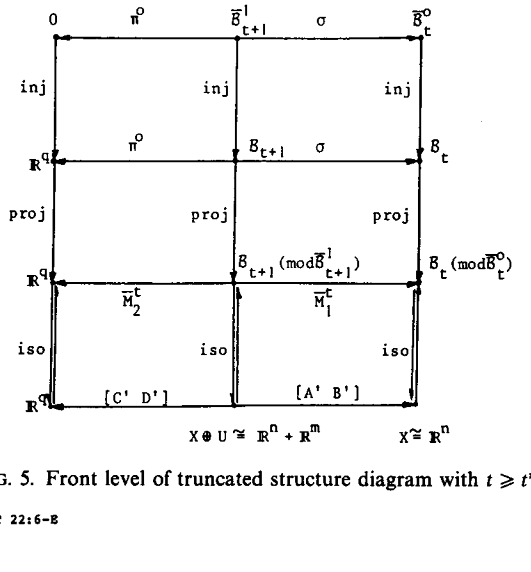


FIG. 5. Front level of truncated structure diagram with $t \geqslant t^*$.

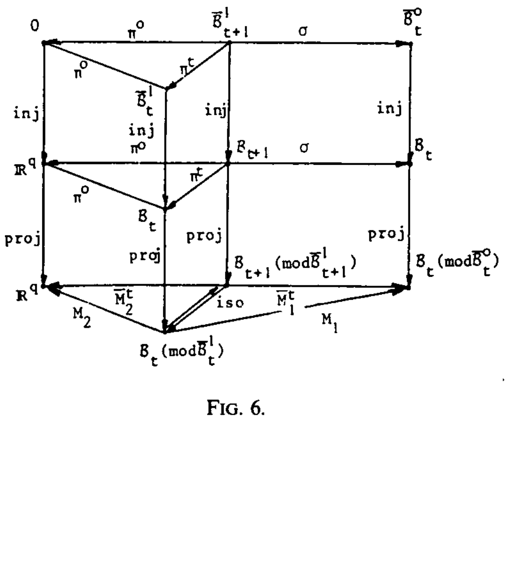


FIG. 6.

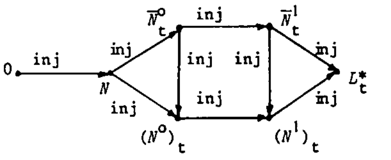


FIG. 7. Lattice diagram.

Further, im $i^0 \subset \bar{\mathcal{N}}_t^{\,1} \subset (\mathcal{N}^1)_t$, $s(\mathcal{N}^0)_t \supset (\mathcal{N}^1)_{t+1}$, and $s\mathcal{N}_t^0 \subset \bar{\mathcal{N}}_{t+1}^{\,1}$.

Note the significance of these spaces: $\mathcal{N}_t$, viewed as a subspace of $\mathbb{R}^{1\times q}[s]$, represents all (AR) relations of lag less than or equal to t satisfied on $\mathcal{B}$; $\bar{\mathcal{N}}_t^{\,0}$ adds all polynomials which can be obtained from $\mathcal{N}_t$ by dropping the lowest order terms and then dividing by the smallest power which was not dropped, while $\bar{\mathcal{N}}_t^{\,1}$ adds in addition all zeroth order polynomials.

Example 7 *revisited.* For the first situation discussed in Example 7, $\mathcal{N}^+ = \mathcal{N}^0 = \{\mathbf{0}\}$ and $\mathcal{N}^1 = \mathbb{R}^{1\times q}[0]$. Hence $\mathcal{N}^+ = (\mathcal{N}^0)_t = \bar{\mathcal{N}}_t^0 = \{\mathbf{0}\}$ and $(\bar{\mathcal{N}}^1)_t = \bar{\mathcal{N}}_t^1 = \mathbb{R}^{1\times q}[0]$.

For the second situation, $\mathcal{N}^+ = s^n\mathbb{R}[s]$, and $\mathcal{N}^0 = \mathcal{N}^1 = \mathbb{R}[s]$. Hence $\mathcal{N}_t = \{\mathbf{0}\}$ for $t < n$ and $\mathcal{N}_t = s^n\mathcal{L}_{t-n}^*$ for $t \geqslant n$. Consequently, for $0 \leqslant t < n$, $(\mathcal{N}^0)_t = (\mathcal{N}^1)_t = \mathcal{L}_t^*$, $\mathcal{N}_t = \bar{\mathcal{N}}_t^0 = \{\mathbf{0}\}$; and $\bar{\mathcal{N}}_t^1 = \mathbb{R}[0]$, while, for $t \geqslant n$, $(\mathcal{N}^0)_t = (\mathcal{N}^1)_t = \bar{\mathcal{N}}_t^0 = \mathcal{N}_t^1 = \mathcal{L}_t^*$.

The dual of Proposition 16, of the truncated structure diagram (Fig. 8), and of Theorem 17 now follows.

Proposition 16 (*Dual version*). Let $\mathcal{B} \in \mathbf{L}$. Then the following conditions on $t \in \mathbb{Z}_+$ are equivalent:

(i) $\exists R \in \mathbb{R}^{g\times q}[s]$ such that $\mathcal{B} = \mathcal{B}(R)$ and $t \geqslant \partial(R)$ (i.e. in terms of the lag structure studied in Section 7, $t \geqslant \partial_{g^*}^*$);

(ii) $\dim(\mathcal{N}^0)_t \pmod{\mathcal{N}_t} = \dim \mathcal{N}^0 \pmod{\mathcal{N}}$;

(iii) $\dim(\mathcal{N}^1)_t \pmod{\mathcal{N}_t} = \dim \mathcal{N}^1 \pmod{\mathcal{N}}$;

(iv) $\bar{\mathcal{N}}_{t+1}^{\,0} = (\mathcal{N}^0)_{t+1}$;

(v) $\bar{\mathcal{N}}_{t+1}^{\,1} = (\mathcal{N}^1)_{t+1}$.

Proof. See Appendix P.

Theorem 17 (*Dual version*). Let $\mathcal{B} \in \mathbf{L}$. Now assume $t \geqslant t^* = \partial_{g^*}^*$. Then there exist linear maps $(\bar{M}_1^t)^*$, $(\bar{M}_2^t)^*$ such that the dual truncated structure diagram (Fig. 8) commutes. The map $M^* = ((\bar{M}_1^t)^*, (\bar{M}_2^t)^*)$: $(\bar{\mathcal{N}}_t^0 \pmod{\mathcal{N}_t}, \mathbb{R}^q) \to \bar{\mathcal{N}}_{t+1}) \pmod{\mathcal{N}_{t+1}}$ yields, identically as in Theorem 9 (dual version), a state space representation $\Sigma_s(A', B', C', D')$ of $\mathcal{B}$ with a minimal number of states and driving inputs.

Proof. See Appendix P.

Example 7 *revisited.* For the second system of Example 7, when $t \geqslant t^* = n$, $\mathcal{B}_t = \{\mathbf{w} \in \mathcal{L}_t \mid \mathbf{w}(t') = 0$ for $t' \geqslant n\}$, $\mathcal{B}_t^0 = \mathcal{B}_t^1 = \{0\}$, $\mathcal{N}_t = s^n\mathcal{L}_{t-n}^*$, $\bar{\mathcal{N}}_t^{\,0} = \bar{\mathcal{N}}_t^* = \mathcal{L}_t^*$. Hence $\mathcal{B}_t \pmod{\mathcal{B}_t^0} \cong X \cong \mathcal{B}_{t+1} \pmod{\mathcal{B}_{t+1}^1} \cong \mathcal{B}_t \pmod{\mathcal{B}_t^1}$, and $X \cong \mathbb{R}^n$, $U \cong 0$. Choose as basis for $X \cong \mathbb{R}^n$ the elements e_i with $e_i \cong \{\mathbf{w} \mid \mathbf{w}(t) = 0$ for $t \neq i-1$ and $\mathbf{w}(i-1) = 1\}$. This yields

$$A' = \begin{bmatrix} 0 & 1 & \cdots & 0 & 0 \\ 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & \cdots & 0 & 0 \end{bmatrix} \text{ and } C' = [1 \quad 0 \quad \cdots \quad 0].$$

The primal structure diagram yields $(A', -, C', -)$ as a natural minimal realization of the system $\sigma^n\mathbf{w} = \mathbf{0}$ with $T = \mathbb{Z}_+$. Now apply the dual structure diagram.

$$\mathcal{N}_t \pmod{\bar{\mathcal{N}}_t^{\,0}} \cong X \cong \mathcal{N}_{t+1} \pmod{\bar{\mathcal{N}}_{t+1}^{\,1}} \cong \mathcal{N}_t \pmod{\bar{\mathcal{N}}_t^{\,1}}$$

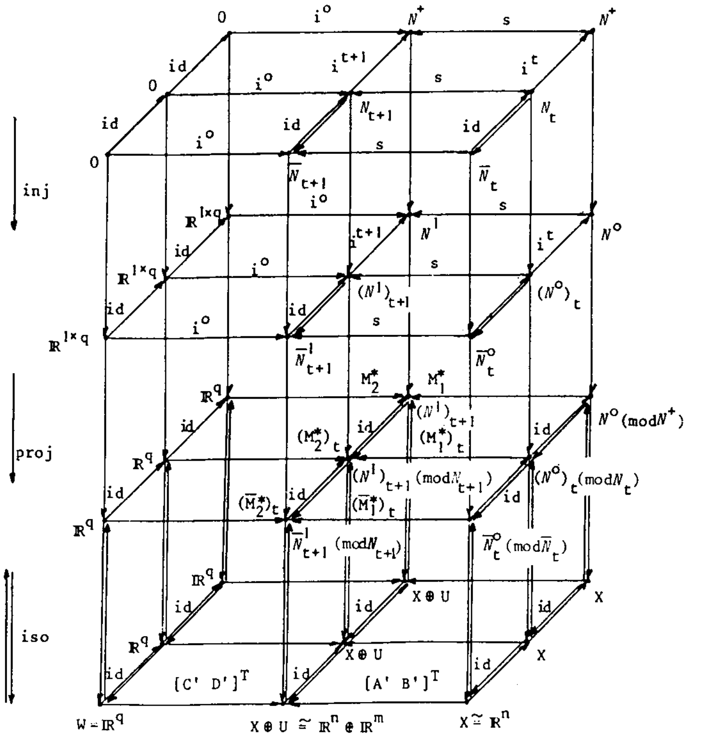


FIG. 8. Dual truncated structure diagram.

and $X \cong \mathbb{R}^n$, $U \cong 0$. Choose as basis on $X \cong \mathbb{R}^n$ the elements e_i with $e_i \cong s^{n-i}$. This yields

$$A'' = \begin{bmatrix} 0 & 0 & \cdots & 0 & 0 \\ 1 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & \cdots & 1 & 0 \end{bmatrix} \text{ and } C'' = [0 \ldots 0 \quad 1]$$

and $(A'', -, C'', -)$ is obtained as another minimal realization of the system.

The dual version of Theorem 17 will now be implemented in an algorithm which is basically a refinement of Theorem 10, where a state space model for a given (AR) system was constructed.

Algorithm 4

Data $R(s) \in \mathbb{R}^{g \times q}[s]$.

Step 1. Determine $t' = \partial(R)$.

Step 2. Determine

$$\mathcal{N}^+_{R,t'} := \{p(s) \in \mathbb{R}^{1 \times q}[s] \mid \partial(p) \leqslant t' \text{ and } p = fR$$

$$\text{for some } f(s) \in \begin{cases} \mathbb{R}^{1 \times q}[s] & \text{case } T = \mathbb{Z}_+ \\ \mathbb{R}^{1 \times q}[s, s^{-1}] & \text{case } T = \mathbb{Z} \end{cases}$$

$$\bar{\mathcal{N}}^0_{R,t'} := \sum_{0 \leqslant t \leqslant t'} (s^*)^t \mathcal{N}^+_{R,t'}$$

$$\bar{\mathcal{N}}^1_{R,t'} := \bar{\mathcal{N}}^0_{R,t'} + \mathbb{R}^{1 \times q}[0].$$

Step 3. Determine complementary bases $f_1, f_2, \ldots, f_n \in \mathbb{R}^{1 \times q}[s]$ for $\mathcal{N}^+_{R,t'}$ (viewed as a subspace of $\mathcal{L}^*_{t'}$) in $\mathcal{N}^0_{R,t'}$ and $u_1, u_2, \ldots, u_m \in \mathbb{R}^{1 \times q}[s]$ for $\mathcal{N}^0_{R,t'}$ in $\mathcal{N}^1_{R,t'}$. Let $F := \text{col}(f_1, f_2, \ldots, f_n)$ and $U := \text{col}(u_1, u_2, \ldots, u_m)$. Note that by Step 1, $\dim \mathcal{N}^+_{R,t} - \dim \mathcal{N}^+_{R,t-1} = \dim \mathcal{N}^+_{R,t'} - \dim \mathcal{N}^+_{R,t'-1}$ for $r \geqslant t'$. Further $\mathcal{N}^+_{R,t'} = \mathcal{N}^+_{R,t'+1} \cap \mathcal{L}^*_{t'}$, $\bar{\mathcal{N}}^0_{R,t'} = \bar{\mathcal{N}}^0_{R,t'+1} \cap \mathcal{L}^*_{t'}$, and $\bar{\mathcal{N}}^1_{R,t'} = \bar{\mathcal{N}}^1_{R,t'+1} \cap \mathcal{L}^*_{t'}$. This shows that $f_1, \ldots, f_n$ and $u_1, \ldots, u_m$ also form complementary bases for $\mathcal{N}^+_{R,t'+1}$ in $\mathcal{N}^0_{R,t'+1}$ and for $\mathcal{N}^0_{R,t'+1}$ in $\mathcal{N}^1_{R,t'+1}$.

Observe that the construction of these bases is very easy to carry out if it is assumed that a shortest lag description $R = \text{col}(r_1, r_2, \ldots, r_g)$ of $\mathcal{B}(R)$ is used. Assume $\partial_i = \partial(r_i)$ and $0 \leqslant \partial_1 \leqslant \cdots \leqslant \partial_g$. Then $t' = \partial_g$ and $s^* r_1, \ldots, (s^*)^{\partial_1} r_1$, $s^* r_2, \ldots, (s^*)^{\partial_2} r_2, \ldots, s^* r_g, \ldots, (s^*)^{\partial_g} r_g$ form a complementary basis for $\mathcal{N}^+_{R,t'}$ in $\mathcal{N}^0_{R,t'}$ and for $u_1, u_2, \ldots, u_m$, a complementary basis for $(s^*)^{\partial_1} r_1, (s^*)^{\partial_2} r_2, \ldots, (s^*)^{\partial_g} r_g$ in $\mathbb{R}^{1 \times q}[0]$ should be constructed. These observations make Algorithm 4 very concrete if a shortest lag (AR) description is used.

Step 4. Find matrices A', B', C', D' and $N = \text{col}(n_1, n_2, \ldots, n_{n+q})$, with $n_i \in \mathcal{N}^+_{R,t'+1}$ for $i \in (\mathbf{n} + \mathbf{q})$, such that

$$\begin{bmatrix} sF(s) \\ \hline I \end{bmatrix} = \begin{bmatrix} A' \\ \hline B' \end{bmatrix} F(s) + \begin{bmatrix} B' \\ \hline D' \end{bmatrix} U(s) + N(s).$$

Corollary 18. Let (A', B', C', D') be computed as in Algorithm 4. Then $\Sigma_s(A', B', C', D')$ defines a minimal state space realization of $\Sigma(R)$ with a minimal number of states and driving inputs.

Proof. This is an immediate consequence of Theorem 10. The only reason for restating it explicitly is to emphasize the finite dimensional nature of the realization algorithm starting from R. □

Example 4 *revisited.* It is clear that $t' = n = \partial(p)$. Hence $\mathcal{N}^+_{R,t'} = \{\alpha[p, -q] \mid \alpha \in \mathbb{R}\}$. The f_is still form a basis for a complement for $\mathcal{N}^+_{R,t'}$ in $\mathcal{N}^0_{R,t'}$ and $[0, 1]$ forms a basis for a complement for $\mathcal{N}^0_{R,t'}$ in $\mathcal{N}^1_{R,t'}$. The calculations for (A', B', C', D') remain unchanged but can now be considered as taking place in the finite dimensional space $\mathcal{L}^*_{t'}$ or, better yet, in span $((f_i, i \in \mathbf{n}), [p, -q], [0, 1])$.

Remark. The development in this section shows the crucial importance of the 'partial' behaviour spaces $\mathcal{B}_0, \mathcal{B}_1, \ldots, \mathcal{B}_t, \ldots$. As seen in Section 7, the dimension of these spaces determine the lag structure of the (AR) equations which describe the system and the observability indices of the i/s/o representations. In Part 3 these dimensions will be used as the definition of the complexity of a system.

These partial behaviour spaces also allow formal study of the finite time structure of a system as follows. Call an element $\mathcal{B} \in \mathbf{L}$ *t-complete* if $\{\mathbf{w} \in \mathcal{B}\} \Leftrightarrow \{\pi^t \sigma^{t'} \mathbf{w} \in \mathcal{B}_t, \forall t' \in T\}$. Let $\mathbf{L}_t \subset \mathbf{L}$ denote the family of all t-complete systems. In fact, $\{\mathcal{B} \in \mathbf{L}_t\} \Leftrightarrow \{t$ satisfies the conditions of Theorem 17$\}$. Actually, contrary to $\mathbf{L}$, $\mathbf{L}_t$ is not closed under addition but it nevertheless has the intersection property and consequently there exists a most powerful t-complete system containing any subset $Z \subset (\mathbb{R}^q)^T$. Consequently, the computation of the most powerful t-complete model for any observed time series $Z = \{\mathbf{w}\}$, or for any behaviour $\mathcal{B} \subset \mathbf{L}$, for that matter can be considered. This yields the *t-completion* of the behaviour $\mathcal{B}$, denoted by $\mathcal{B}^*_t$. Actually these computations are easy to carry out, since in this case relations with unknown long lags need not be considered. When $Z = \{\tilde{\mathbf{w}}\}$ it suffices to use Algorithms 3 or 5 with the t' in Step 1 imposed instead of computed from the data (this has the pleasing consequence that these algorithms then basically operate only on matrices with a finite number of rows. Similarly, when $Z = \mathcal{B}$, a state space realization of $\mathcal{B}^*_t$ can be obtained by using an identical computation as set up in Theorem 17 but with t again imposed.

These refinements and their ramifications will not be pursued further. However, note the relation with the question of persistency of excitation (see Chen, 1986 and references therein). Consider an i/o system with behaviour $\mathcal{B} \in \mathbf{L}$ for which one input/output response pair $\tilde{\mathbf{w}} = (\tilde{\mathbf{u}}, \tilde{\mathbf{y}}) \in (\mathbb{R}^m \times \mathbb{R}^p)^T$ has been observed. Then it is possible to prove that $\mathcal{B}$ itself is the most powerful unfalsified (AR) model for $\{\tilde{\mathbf{w}}\}$ iff $\tilde{\mathbf{u}}, \sigma\tilde{\mathbf{u}}, \ldots, \sigma^t\tilde{\mathbf{u}}, \ldots$ are linearly independent in $(\mathbb{R}^m)^T$, i.e. iff $R^*_{\mathbf{u}} = 0$. It is natural to call such an input sequence *completely exciting*. If *a priori* knowledge is assumed on $\mathcal{B}$ such that $\mathcal{B} \in \mathbf{L}_t$, then $\mathcal{B}$ will be the most powerful unfalsified t-complete (AR) model for $\{\tilde{\mathbf{w}}\}$ iff $\tilde{\mathbf{u}}, \sigma\tilde{\mathbf{u}}, \ldots, \sigma^t\tilde{\mathbf{u}}$ are linearly independent. Such inputs are called *completely exciting for t-complete systems.*

Consideration of t-completeness allows truncation of the number of rows of $\mathcal{H}(\tilde{\mathbf{w}})$. Is it also possible to truncate the number of columns of $\mathcal{H}_t(\tilde{\mathbf{w}})$? For t-complete systems, when trying to construct the behaviour from $\tilde{\mathbf{w}}$, first consider the truncations $\pi^t \sigma^{t'} \tilde{\mathbf{w}}$, $t' \in T$, compute $\mathcal{B}_t = \text{im}\,\mathcal{H}_t(\tilde{\mathbf{w}}) = \text{span}(\pi^t \sigma^{t'} \tilde{\mathbf{w}}, t' \in T)$, and deduce $\mathcal{B}$ from there. if, however, $\text{im}\,\mathcal{H}_t(\tilde{\mathbf{w}}) = \text{span}(\pi^t \sigma^{t'} \tilde{\mathbf{w}}, \pi^t \sigma^{t'+1} \tilde{\mathbf{w}}, \ldots, \pi^t \sigma^{t'+\Delta} \tilde{\mathbf{w}})$ then it is obviously possible to deduce $\mathcal{B}$ from the truncations $\pi^t \sigma^{t'} \tilde{\mathbf{w}}, \pi^t \sigma^{t'+1} \tilde{\mathbf{w}}, \ldots, \pi^t \sigma^{t'+\Delta} \tilde{\mathbf{w}}$. It is possible to prove that this will be the case iff $\pi^t \sigma^{t'} \tilde{\mathbf{u}}, \pi^t \sigma^{t'+1} \tilde{\mathbf{u}}, \ldots, \pi^t \sigma^{t'+\Delta} \tilde{\mathbf{u}}$ are linearly independent as elements of $\mathcal{L}_t \cong \mathbb{R}^{m(t+1)}$. Such an input sequence is *fully exciting on the interval* $[t', t' + \Delta]$ *for t-complete systems.* If this holds for all $t' \in T$ then such an input is called *persistently exciting of order* Δ *for t-complete systems.* These considerations make precise in what sense the combined assumptions of t-completeness and of persistency of excitation of order Δ can be used in order to reduce Algorithms 2 and 3 to algorithms which operate on finite dimensional matrices only. This will be pursued in detail elsewhere. Note also, very importantly, that an input will generically be persistently exciting of order t for t-complete systems.

17. FROM TIME SERIES TO STATE SPACE MODEL

One of the most natural problems which comes up in linear system theory may be formulated as follows:

Given an observed q-dimensional vector time series: $\tilde{\mathbf{w}}(t_0), \tilde{\mathbf{w}}(t_0+1), \ldots, \tilde{\mathbf{w}}(t), \ldots, \tilde{\mathbf{w}}(t_1)$ $(-\infty \leqslant t_0 \leqslant t \leqslant t_1 \leqslant \infty)$, find a minimal state space system: $\sigma\mathbf{x} = A'\mathbf{x} + B'\mathbf{u}$; $\mathbf{w} = C'\mathbf{x} + D'\mathbf{u}$, which explains this time series.

This problem is studied here for the case $t_0 = -\infty$ and $t_1 = \infty$ $(T = \mathbb{Z})$ or $t_0 = 0$ and $t_1 = \infty$ $(T = \mathbb{Z}_+)$. A special case of it has been treated very extensively in the systems theory literature, namely when the observed time series is an impulse response (see Example 6). The resulting theory has become known as *realization theory*. If realization theory is viewed as it should be seen, namely as the problem of *simulating* the i/o map given by the convolution

$$\mathbf{y}(t) = \sum_{t'=0}^{t} G_{t-t'}\mathbf{u}(t')$$

by means of a (minimal) i/s/o system

$$\sigma\mathbf{x} = A\mathbf{x} + B\mathbf{u}; \quad \mathbf{y} = C\mathbf{x} + D\mathbf{u}; \quad \mathbf{x}(0) = 0,$$

then this problem is basically a matter of representation; given a dynamical model of a dynamical system in one form, try to express it in another, more structured, more convenient, more useful one.

The question which is treated in the present section is more akin to (state space) *identification* theory: starting with an arbitrary observed time series, try to model it (exactly) in state space form. As a basic first problem in identification theory, this question is a much more natural one than the impulse response modelling of realization theory (since this corresponds to very special measurements) and it is also more general. In fact, algorithms will be obtained which have the classical realization procedures as a special case.

However, it is not clear *a priori* what is meant by a realization $\sigma\mathbf{x} = A'\mathbf{x} + B'\mathbf{u}$; $\mathbf{w} = C'\mathbf{x} + D'\mathbf{u}$, of an arbitrary observed time series $\tilde{\mathbf{w}}$. In fact, the model $0.\mathbf{w} = \mathbf{0}$, corresponding to the state space model $(-, -, -, I)$, explains every observation and since it is memoryless it has a state space of dimension 0—in a sense it is hence even minimal! Thus it must be made clear how the behaviour of the model is to be related to the data. The precise formulation, which depends heavily on the concepts developed in Part I and the previous sections of Part II, is as follows.

Let $\tilde{\mathbf{w}}$ be a given map from $T = \mathbb{Z}$ or $\mathbb{Z}_+$ to $\mathbb{R}^q$. Consider the model class $\mathcal{M}$ consisting of the external behaviour of the finite dimensional linear time invariant systems with external signal space $\mathbb{R}^q$ (i.e. each element of $\mathcal{M}$ is defined by two integers, m and n, and four matrices: $A' \in \mathbb{R}^{n\times n}$, $B' \in \mathbb{R}^{n\times m}$, $C' \in \mathbb{R}^{q\times n}$, and $D' \in \mathbb{R}^{q\times m}$ which determine the external behaviour of the system $\sigma\mathbf{x} = A'\mathbf{x} + B'\mathbf{u}$; $\mathbf{w} = C'\mathbf{x} + D'\mathbf{u}$ as explained in Section 3.3). The problem is to find a most powerful model in the model class $\mathcal{M}$ which is unfalsified by the measurement $Z = \{\tilde{\mathbf{w}}\}$, and among these, to find one for which n, the dimension of the state space, and m, the dimension of the driving input space, are as small as possible.

It is immediately clear from Theorems 3 and 13 that such a model exists: nothing other than a minimal state space realization of $\mathcal{B}(R^*_{\tilde{\mathbf{w}}})$, the most powerful unfalsified (AR) model, is required. Thus, in a sense, the problem is already solved: Section 14 shows how to obtain $R^*_{\tilde{\mathbf{w}}}$, while Theorem 2 and the proof of Theorem 3 show how to obtain an i/o representation and subsequently a minimal state space realization of $R^*_{\tilde{\mathbf{w}}}$. This route is a very roundabout one however, and in this section algorithms will be obtained which pass directly from $\tilde{\mathbf{w}}$ to (A', B', C', D').

Two algorithms will be presented. The first one uses the Hankel structure of $\mathcal{H}(\tilde{\mathbf{w}})$ in an essential way, while the other does not and is therefore less effective but, in a sense, more general.

Algorithm 5 is an implementation to the problem at hand of the primal version of Theorem 17. On the basis of $\mathcal{H}(\tilde{\mathbf{w}})$, a t must be found which satisfies the conditions of this theorem. Then $\mathcal{B}_t$, $\bar{\mathcal{B}}_t^0$ and $\mathcal{B}_t^1$ will be computed. These yield the state space $X \cong \mathcal{B}_t \,(\mathrm{mod}\,\bar{\mathcal{B}}_t^0)$ and the input space $U = \bar{\mathcal{B}}_t^0$ $(\mathrm{mod}\,\bar{\mathcal{B}}_t^1)$. Finally the representations of the shift and the evaluation map which gives the desired state space realization will be computed. In doing this the primal truncated structure diagram is implemented, making use, however, of the refinement explained in the remark following the primal version of Theorem 17.

Algorithm 5

$$Data. \begin{cases} \tilde{\mathbf{w}}(0), \tilde{\mathbf{w}}(1), \ldots, \tilde{\mathbf{w}}(t), \ldots & (\text{case } T = \mathbb{Z}_+) \\ \ldots, \tilde{\mathbf{w}}(-1), \tilde{\mathbf{w}}(0), \tilde{\mathbf{w}}(1), \ldots & (\text{case } T = \mathbb{Z}). \end{cases}$$

Step 1 (*Determination of the lag*). Determine a $t' \in \mathbb{Z}_+$ as in Step 1 of Algorithm 3.

Step 2 (*Determination of the truncated behaviour*). Determine $\mathcal{B}_{t'} := \mathrm{span}\{\tilde{\mathbf{w}}_{t'}(t);\ t \in T\}$. Here $\tilde{\mathbf{w}}_{t'}(t) := \mathrm{col}(\tilde{\mathbf{w}}(t), \tilde{\mathbf{w}}(t+1), \ldots, \tilde{\mathbf{w}}(t+t'))$.

Step 3 (*Determination of the state space and the input space*). Define $\sigma_{t'}^*: \mathbb{R}^{q(t'+1)} \to \mathbb{R}^{q(t'+1)}$ and

$\pi^0: \mathbb{R}^{q(t'+1)} \to \mathbb{R}^q$ by σ_t^*: col$(a_0, a_1, \ldots, a_{t'}) \mapsto$ col$(0, a_0, \ldots, a_{t'-1})$ and π^0: col$(a_0, a_1, \ldots, a_{t'}) \mapsto a_0$ with $a_t \in \mathbb{R}^q$, $t = 0, 1, \ldots, t'$. Determine $\mathscr{B}_{t'}^0 := \bigcap_{t \leqslant t'} (\sigma_t^*)^{-t} \mathscr{B}_{t'}$ and $\mathscr{B}_{t'}^1 := \mathscr{B}_{t'}^0 \cap \ker \pi^0$.

Step 4 (*Determination of the system parameters*). Determine a $(t'+1)q \times n_2$ submatrix H of $\mathscr{H}_{t'}(\tilde{\mathbf{w}})$, a $(n_2 \times n)$ matrix Q_x, a $(n_2 \times m)$ matrix Q_u, and a $(n \times (t'+1)q)$ matrix P such that the following properties hold:

(4.1) $\mathscr{B}_{t'}^0 \oplus \operatorname{im} HQ_x = \mathscr{B}_{t'}$;

(4.2) $\mathscr{B}_{t'}^1 \oplus \operatorname{im} HQ_u = \mathscr{B}_{t'}^0$;

(4.3) $PHQ_x = I_n, P\mathscr{B}_{t'}^0 = 0$.

It is clear that such matrices indeed exist. Now determine the matrices σH and H^0 defined as follows. Assume that H consists of the $(k_1, k_2, \ldots, k_{n_2})$th columns of $\mathscr{H}_{t'}(\tilde{\mathbf{w}})$. Then σH consists of the $(k_1+1, k_2+1, \ldots, k_{n_2}+1)$th columns of $\mathscr{H}_{t'}(\tilde{\mathbf{w}})$, while H^0 consists of the $(1, 2, \ldots, q)$th rows and the $(k_1, k_2, \ldots, k_{n_2})$th columns of $\mathscr{H}_{t'}(\tilde{\mathbf{w}})$.

Now define the $(n+q) \times (n+m)$ matrix M by

$$M = \begin{bmatrix} P\sigma H \\ \hdashline H^0 \end{bmatrix} [Q_x \quad Q_u].$$

It is worth stating the result formally.

Theorem 19. Let $\tilde{\mathbf{w}}$: $T \to \mathbb{R}^q$, $T = \mathbb{Z}_+$ or $\mathbb{Z}$, be an observed time series, and let $M \in \mathbb{R}^{(n+q)\times(n+m)}$ be as defined in Algorithm 3. Partition M as

$$M =: \begin{bmatrix} A' & B' \\ C' & D' \end{bmatrix}$$

with $A' \in \mathbb{R}^{n\times n}$, $B' \in \mathbb{R}^{n \times m}$, $C' \in \mathbb{R}^{q\times n}$, and $D' \in \mathbb{R}^{q\times m}$. Then $\Sigma_s(A', B', C', D')$ defines a most powerful unfalsified minimal state space model for $\tilde{\mathbf{w}}$ with a minimal number of driving inputs.

Proof. See Appendix P.

The next algorithm is based on the idea of splitting the 'past' and the 'future' of $\tilde{\mathbf{w}}$. As can be seen from Appendix S, the algorithm can in fact also be used in order to split purely static relations as well. Before spelling out the algorithm the notion of the relative row rank $r(M_1; M_2)$ of a partitioned (infinite) matrix $M = \begin{bmatrix} M_1 \\ \hdashline M_2 \end{bmatrix}$ is introduced. Assume first that $M = \text{col}(M_1; M_2)$ is finite: $M_1 \in \mathbb{R}^{k_1 \times k}$, $M_2 \in \mathbb{R}^{k_2 \times k}$. Then $r(M_1; M_2) := \operatorname{rank} M_1 + \operatorname{rank} M_2 - \operatorname{rank} M$. Next, assume that M has an infinite number of columns. Then

$$r(M_1; M_2) := \lim_{t\to\infty} r(M_{1,t}; M_{2,t}) = r(\tilde{M}_1; \tilde{M}_2).$$

Here $M_{i,t}$ denotes the truncation of M_i at its tth column, and $\tilde{M} = \text{col}(\tilde{M}_1, \tilde{M}_2)$ is a maximal rank submatrix of M. This definition is extended in an obvious way to a matrix with a two-sided infinite number of columns. Finally, assume that M_1 and/or M_2 have an infinite number of rows. Then we define

$$r(M_1; M_2) := \lim_{t', t'' \to \infty} r(M_1^{t'}; M_2^{t''}).$$

Here M_1^t and M_2^t denote the truncation of M_1 and M_2 to their t highest and lowest rows, respectively.

It is easy to see that adding rows to M_1^* and/or M_2 never decreases the relative row rank. From there it follows that

$$r(M_1; M_2) = \sup_{t', t''} r(M_1^{t'}; M_2^{t''})$$

and that $r(M_1; M_2)$ is the supremum of the relative rank over all submatrices of M_1 and M_2 obtained by deleting any number of rows in M_1 and M_2.

Now consider $\tilde{\mathbf{w}}$: $\mathbb{Z} \to \mathbb{R}^q$ and define the following partitioned (4 way infinite) Hankel matrix

$$\begin{bmatrix} \mathscr{H}_-(\tilde{\mathbf{w}}) \\ \hdashline \mathscr{H}_+(\tilde{\mathbf{w}}) \end{bmatrix} :=$$

$$\begin{bmatrix} & \vdots & \vdots & & \vdots & \\ \cdots & \tilde{\mathbf{w}}(-t-1) & \tilde{\mathbf{w}}(-t) & \cdots & \tilde{\mathbf{w}}(0) & \cdots \\ & \vdots & \vdots & & \vdots & \\ \cdots & \tilde{\mathbf{w}}(-2) & \tilde{\mathbf{w}}(-1) & \cdots & \tilde{\mathbf{w}}(t'-1) & \cdots \\ \cdots & \tilde{\mathbf{w}}(-1) & \tilde{\mathbf{w}}(0) & \cdots & \tilde{\mathbf{w}}(t') & \cdots \\ \hdashline \cdots & \tilde{\mathbf{w}}(0) & \tilde{\mathbf{w}}(1) & \cdots & \tilde{\mathbf{w}}(t'+1) & \cdots \\ & \vdots & \vdots & & \vdots & \\ \cdots & \tilde{\mathbf{w}}(t-1) & \tilde{\mathbf{w}}(t) & \cdots & \tilde{\mathbf{w}}(t+t') & \cdots \\ & \vdots & \vdots & & \vdots & \end{bmatrix}$$

The relevance of the motion of relative row rank follows from the following result.

Proposition 20. $r(\mathscr{H}_-(\tilde{\mathbf{w}}); \mathscr{H}_+(\tilde{\mathbf{w}})) < \infty$. In fact, it equals the dimension of a minimal state space representation $\Sigma_s(A', B', C', D')$ of $\mathscr{B}(R_{\tilde{\mathbf{w}}}^*)$, the most powerful unfalsified (AR) model for $\tilde{\mathbf{w}}$.

Proof. See Appendix P.

Observe that it follows immediately from Proposition 20 that the relative row rank of any infinite block Hankel or Toeplitz matrix will always be finite.

For the next algorithm, since the case $T = \mathbb{Z}_+$ causes certain complications, only the case $T = \mathbb{Z}$ will be considered. Later on comments on the generalization to the case $T = \mathbb{Z}_+$ will be made.

Algorithm 6

Data. ..., $\tilde{\mathbf{w}}(-1)$, $\tilde{\mathbf{w}}(0)$, $\tilde{\mathbf{w}}(1), \ldots, \tilde{\mathbf{w}}(t), \ldots$.

Step 1 (*Structure determination in* $\mathscr{H}(\tilde{\mathbf{w}})$). Determine matrices H_- and H_+, with H_- consisting of rows of $\mathscr{H}_-(\tilde{\mathbf{w}})$ and H_+ consisting of rows of $\mathscr{H}_+(\tilde{\mathbf{w}})$, such that

$$r(H_-; H_+) = r(\mathscr{H}_-(\tilde{\mathbf{w}}); \mathscr{H}_+(\tilde{\mathbf{w}})) =: n.$$

(Actually, with t' as in Step 1 of Algorithm 3, the $q(t'+1)$ bottom rows of $\mathscr{H}_-(\tilde{\mathbf{w}})$ and the $q(t'+1)$ top rows of $\mathscr{H}_+(\tilde{\mathbf{w}})$ is a suitable selection.) Next, determine a matrix col (H_1, H_2) consisting of a finite set of columns of $H = \text{col}(H_-, H_+)$ such that the columns of $\text{col}(H_1, H_2)$ span those of H. Of course, $r(H_1; H_2) = n$.

Step 2 (*Determination of the state space*). Determine the kernel of H_1 and define $\mathscr{K} := H_2 \ker H_1$. Note that $\dim((\text{im}\, H_2)(\text{mod}\, \mathscr{K})) = n$. Then $X \approx ((\text{im}\, H_2)(\text{mod}\, \mathscr{K}))$. Define $\mathbf{x}(t) := \mathbf{h}_+(t)(\text{mod}\, \mathscr{K})$, with $\mathbf{h}_+(t)$ the tth column of H_+.

Step 3 (*Determination of the input space*). Define $\mathbf{f}(t) := \text{col}(\tilde{\mathbf{w}}(t), \mathbf{x}(t))$, and $S := \text{span}\,\{\mathbf{f}(t), t \in \mathbb{Z}\}$. Obviously the projection $\pi_x: S \to X$ defined by $\pi_x \mathbf{f}(t) := \mathbf{x}(t)$ is surjective (mathematically this specifies S as a vector bundle over X). Identify a vector space U and a surjection $\pi_u := S \to U$ such that $S = X \oplus U$ i.e. such that $\pi := (\pi_x, \pi_u)$ is bijective. Clearly $\dim U = \dim S - \dim X$. Define $\mathbf{u}(t) := P_u \mathbf{f}(t)$.

Step 4 (*Determination of the system parameters*). Determine $t_i, i \in (\mathbf{n} + \mathbf{m})$, such that the $\mathbf{f}(t_i)$s form a basis for S. Then $\mathbf{f}(t_i) = \text{col}(\mathbf{x}(t_i), \mathbf{u}(t_i))$ forms a basis for $X \oplus U$. Now determine the $(n+q) \times (n+m)$ matrix M such that

$$M: \begin{bmatrix} \mathbf{x}(t_i) \\ \cdots \\ \mathbf{u}(t_i) \end{bmatrix} \mapsto \begin{bmatrix} \mathbf{x}(t_i + 1) \\ \cdots \\ \mathbf{w}(t_i) \end{bmatrix}.$$

This yields:

Theorem 21. Let $\tilde{\mathbf{w}}: \mathbb{Z} \to \mathbb{R}^q$ be an observed time series, and let $M \in \mathbb{R}^{(n+q) \times (n+m)}$ be as defined in Algorithm 6. Partition M as

$$M =: \left[\begin{array}{c:c} A' & B' \\ \hdashline C' & D' \end{array}\right]$$

with $A' \in \mathbb{R}^{n \times m}$, $B' \in \mathbb{R}^{n \times m}$, $C' \in \mathbb{R}^{q \times n}$, and $D' \in \mathbb{R}^{q \times m}$. Then $\Sigma_s(A', B', C', D')$ defines a most powerful unfalsified minimal state space model for $\tilde{\mathbf{w}}$ with a minimal number of driving inputs.

Proof. See Appendix P.

Remarks. Algorithm 6 needs some adjustments in the case $T = \mathbb{Z}_+$. These modifications are minor when $\mathscr{B}(R^*_{\#}) = \sigma\mathscr{B}(R^*_{\#})$ but more complicated otherwise. It is easy to decide whether or not this condition is satisfied. Determine (as explained in Step 1 of Algorithm 2) the dependency vectors of $\tilde{\mathbf{w}}$ and $\sigma\tilde{\mathbf{w}}$, denoted respectively by $\mathbf{d}(\tilde{\mathbf{w}})$ and $\mathbf{d}(\sigma\tilde{\mathbf{w}})$. Then

$$\{\mathscr{B}(R^*_{\#}) = \sigma\mathscr{B}(R^*_{\#})\} \Leftrightarrow \{\mathbf{d}(\tilde{\mathbf{w}}) = \mathbf{d}(\sigma\tilde{\mathbf{w}})\}.$$

When this is the case, then Algorithm 6 requires only the following minor modification. Consider the Hankel matrix $\mathscr{H}(\tilde{\mathbf{w}})$. Partition $\mathscr{H}(\tilde{\mathbf{w}})$ into $\mathscr{H}(\tilde{\mathbf{w}}) = \text{col}(\mathscr{H}^t_-(\tilde{\mathbf{w}}), \mathscr{H}^t_+(\tilde{\mathbf{w}}))$ with $\mathscr{H}^t_+(\tilde{\mathbf{w}})$ the first $q(t+1)$ rows of $\mathscr{H}(\tilde{\mathbf{w}})$. Now look for a t such that $\mathbf{r}_t = r(H^t_-(\tilde{\mathbf{w}}); H^t_+(\tilde{\mathbf{w}}))$ equals $\max_{t \in \mathbb{Z}_+} \mathbf{r}_t$. In fact $\max_{t \in \mathbb{Z}_+} \mathbf{r}_t = \lim_{t \to \infty} \mathbf{r}_t$ and any $\mathbf{t}'$ as in Step 1 of Algorithm 2 will yield $\mathbf{r}_{t'} = \max_{t \in \mathbb{Z}_+} \mathbf{r}_t$. Now apply Algorithm 4 with $\mathscr{H}(\tilde{\mathbf{w}})$ replaced by $\mathscr{H}^t_-(\tilde{\mathbf{w}})$ and $\mathscr{H}_+(\tilde{\mathbf{w}})$ replaced by $\mathscr{H}^t_+(\tilde{\mathbf{w}})$.

When $\mathbf{d}(\tilde{\mathbf{w}}) \neq \mathbf{d}(\sigma\tilde{\mathbf{w}})$, the situation is more involved. The difficulty stems from the fact that then there will be nilpotent unreachable models in $R^*_{\#}$ (see comment (i) in Section 6 of Part I). The full details for this case will not be given. The idea is as follows. Compare $\mathbf{d}(\tilde{\mathbf{w}})$ with $\mathbf{d}(\sigma\tilde{\mathbf{w}})$. Determine a t' such that $\mathbf{d}(\tilde{\mathbf{w}})$ and $\mathbf{d}(\sigma\tilde{\mathbf{w}})$ coincide in their entries starting from the $(qt'+1)$th entry. This implies $\mathbf{d}(\sigma^{t'}\tilde{\mathbf{w}}) = \mathbf{d}(\sigma^{t'+1}\tilde{\mathbf{w}})$ and hence a realization for $\sigma^{t'}\tilde{\mathbf{w}}$ can be constructed in the manner explained before. This realization can now be run backwards in time using the driving input obtained directly from $\mathbf{w}$, as explained for example in Algorithms 5 and 6. Now subtract the output obtained this way from $\tilde{\mathbf{w}}$, and denote this difference as $\Delta\tilde{\mathbf{w}}: \mathbb{Z}_+ \to \mathbb{R}^q$. It will have the property $\Delta\tilde{\mathbf{w}}(t) = 0$ for $t \geq t'$. Now realize this $\Delta\tilde{\mathbf{w}}$ minimally: this yields, of course, an autonomous system (no driving inputs are needed) with a nilpotent A' and an observable (A', C'). The parallel connection of this realization with the original one yields the analogue of the realization of Theorem 21 for $T = \mathbb{Z}_+$.

The algorithms obtained in this section at first sight show more than casual similarity with the classical impulse response realization algorithms (see Example 6). Both operate on the Hankel matrix formed by the data. The crucial role played by the Hankel matrix is, of course, no accident: realization algorithms must examine first, the span of $\{\sigma^t\tilde{\mathbf{w}}, t \in T\}$ and second, the action of the shift on this span. Now, both these operations are very effectively displayed by the Hankel matrix of the data. However, the major difference between the classical impulse response case and this one is that here one has to be concerned not just with the rank but with a type of relative rank: a permanent rank increase in the Hankel matrix will be due to an input and, since it does not lead to complex dynamics, will be relatively easy to handle. As such, viewing these algorithms as extensions of the classical impulse response realization algorithms is not a particularly effective way of penetrating what is going on—witness the fact that the proof of Corollary 22 (where it is shown how the impulse response case may be viewed as a special case) is not particularly easy. Algorithm 5 is viewed as an implementation of the abstract theory of Section 15 and Algorithm 6 as an implementation of the abstract theory of Appendix S.

Algorithms 5 and 6 can be refined, streamlined and improved in many different directions: recursivity, truncation, use of efficient numerical procedures etc. These implementations will be pursued at a future stage.

Example 5 revisited. The application of Algorithms 5 and 6 to the time series of Example 5 will be illustrated here.

For Algorithm 5, take $t' = 2$. Then

$$\mathscr{B}'_t = \text{im}\begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 2 \\ 0 & 1 & 0 & 0 \\ 1 & 1 & 2 & 2 \\ 1 & 0 & 0 & 1 \\ 1 & 2 & 2 & 2 \end{bmatrix}, \quad \mathscr{B}^0_{t'} = \text{im}\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix} \quad \text{and} \quad \mathscr{B}^1_{t'} = \text{im}\begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 1 & 0 \\ 0 & 0 \\ 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

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

the 1st and 4th column of $\mathscr{H}_{t'}(\tilde{\mathbf{w}})$, $Q_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$, $Q_u = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $P = [-\frac{1}{4}, \frac{1}{4}, 0, \frac{1}{4}, 0, 0]$. With

$$\sigma H = \begin{bmatrix} 0 & 0 \\ 2 & 1 \\ 1 & 1 \\ 2 & 1 \\ 0 & 0 \\ 3 & 2 \end{bmatrix} \quad \text{and} \quad H^o = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix},$$

this yields $A' = 1$, $B' = \frac{1}{2}$, $C' = \begin{bmatrix} 0 \\ 2 \end{bmatrix}$, $D' = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, i.e. the system $\sigma \mathbf{x} = \mathbf{x} + \frac{1}{2}\mathbf{u}$; $\mathbf{w}_1 = \mathbf{u}$; $\mathbf{w}_2 = 2\mathbf{x}$, which is indeed a minimal i/s/o realization of $\sigma \mathbf{w}_2 = \mathbf{w}_1 + \mathbf{w}_2$.

Now turn to Algorithm 6. Since $\mathbf{d}(\tilde{\mathbf{w}}) = \mathbf{d}(\sigma \tilde{\mathbf{w}})$ there are no nilpotent unreachable modes and the algorithm can be applied in its uncomplicated form even though $T = \mathbb{Z}_+$. In the notation of the above remark, $r_1 = \max_{t \in \mathbb{Z}_+} r_t$ and Algorithm 4 can be applied with

$$H_- = \begin{bmatrix} 1 & 0 & 1 & 0 & 0 & 1 & \cdots \\ 0 & 1 & 1 & 2 & 2 & 2 & \cdots \end{bmatrix}$$

and

$$H_+ = \begin{bmatrix} 0 & 1 & 0 & 0 & 1 & 0 & \cdots \\ 1 & 1 & 2 & 2 & 2 & 3 & \cdots \end{bmatrix}.$$

Take $H_1 = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 2 \end{bmatrix}$ and $H_2 = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 1 & 2 & 2 \end{bmatrix}$. Then

$$\ker H_1 = \operatorname{span} \begin{bmatrix} 1 & 0 \\ 1 & -2 \\ -1 & 0 \\ 0 & 1 \end{bmatrix}$$

and $\mathscr{X} = H_2 \ker H_1 = \operatorname{span} \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, giving

$$\mathbf{x} = (1, 1, 2, 2, 2, 3, \ldots) \quad \text{and} \quad S = \operatorname{im} \begin{bmatrix} 0 & 1 \\ 1 & 1 \\ 1 & 1 \end{bmatrix}$$

Hence $U \cong \mathbb{R}$.

Take $P_u = \begin{bmatrix} s_1 \\ s_2 \\ s_3 \end{bmatrix} \mapsto s_1$. Then $\mathbf{u} = (0, 1, 0, 0, 1, 0, \ldots)$.

Hence M must map $\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ into $\begin{bmatrix} 1 & 2 \\ 0 & 1 \\ 1 & 1 \end{bmatrix}$ which yields $M = \begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 0 \end{bmatrix}$ and the state space model $\sigma \mathbf{x} = \mathbf{x} + \mathbf{u}$; $\mathbf{w}_1 = \mathbf{u}$; $\mathbf{w}_2 = \mathbf{x}$ as a minimal i/s/o realization of $\sigma \mathbf{w}_2 = \mathbf{w}_1 + \mathbf{w}_2$.

Consider now the case that k time series $\tilde{\mathbf{w}}_1, \tilde{\mathbf{w}}_2, \ldots, \tilde{\mathbf{w}}_k$ are observed. As with Algorithms 3 and 4, Algorithms 5 and 6 may be applied unchanged in this case, with, of course, notational adjustments to block Hankel matrices etc. This generalization will be used in the next example.

Example 6 revisited (*Impulse response measurements*). The purpose now is to prove, both as an illustration of Algorithms 5 and 6, and as a result of independent interest, the following algorithm.

Consider the impulse response matrix $\tilde{\mathbf{G}} = \{\tilde{\mathbf{G}}(0), \tilde{\mathbf{G}}(1), \ldots, \tilde{\mathbf{G}}(t), \ldots\}$, and assume that rank $\mathscr{H}(\tilde{\mathbf{G}}) =: n < \infty$. Determine a submatrix $\hat{H}$ of $\mathscr{H}(\tilde{\mathbf{G}})$ such that rank $\hat{H} = n$. Assume that $\hat{H}$ consists of the elements in the $(r_1, r_2, \ldots, r_{n_1})$th rows and the $(k_1, k_2, \ldots, k_{n_2})$th columns of $\mathscr{H}(\tilde{\mathbf{G}})$:

(i) $\sigma\hat{H}$, the matrix formed by the elements in the $(r_1, r_2, \ldots, r_{n_1})$th rows and the $(k_1 + m, k_2 + m, \ldots, k_{n_2} + m)$th columns of $\mathscr{H}(\tilde{\mathbf{G}})$ (equivalently, in the $(r_1 + p, r_2 + p, \ldots, r_{n_1} + p)$th rows and the $(k_1, k_2, \ldots, k_{n_2})$th columns);

(ii) $\hat{H}^o$, the matrix formed by the elements in the first p rows and the $(k_1, k_2, \ldots, k_{n_2})$th columns of $\mathscr{H}(\tilde{\mathbf{G}})$;

(iii) $\hat{H}_o$, the matrix formed by the elements in the $(r_1, r_2, \ldots, r_{n_1})$th rows and the first m columns of $\mathscr{H}(\tilde{\mathbf{G}})$.

This is stated formally below.

Corollary 22. Let $\hat{H}$, $\sigma\hat{H}$, $\hat{H}^o$ and $\hat{H}_o$ be as defined above. Let $\hat{P}$ and $\hat{Q}$ be matrices such that $\hat{P}\hat{H}\hat{Q} = I_n$. Then $\{\hat{P}\sigma\hat{H}\hat{Q}, \hat{P}\hat{H}_o, \hat{H}^o\hat{Q}, \tilde{\mathbf{G}}(0)\}$ defines a minimal i/s/o realization of the matrix impulse response $\tilde{\mathbf{G}}$.

Proof. See Appendix P.

Note the following special cases of Corollary 22.

(1) *The algorithm of B. L. Ho* (*Kalman* et al., 1969)

Let $H_{N,N'}$ be a leading submatrix of $\mathscr{H}(\mathbf{G})$, consisting of its first pN rows and its first nN' columns such that rank $H_{N,N'} = n'$. Determine $R \in \mathbb{R}^{pN \times pN}$ and $S \in \mathbb{R}^{mN' \times mN'}$ such that $RH_{N,N'}S = \begin{bmatrix} I_n & 0 \\ 0 & 0 \end{bmatrix}$. Compute $A = E_{m,mN}R(\sigma H)_{N,N'}SE_{pN',n}$; $B = E_{n,mN}RH_{N,N'}E_{pN',m}$; $C = E_{p,mN}, H_{N,N'}SE_{pN',n}$; $D = \tilde{\mathbf{G}}(0)$. Here $(\sigma H)_{N,N'}$ denotes the $(pN \times mN')$ leading submatrix of $\mathscr{H}(\sigma\tilde{\mathbf{G}})$ and E_{ij} denotes the $(i \times j)$ selection matrix with $(E_{ij})_{k,l} = \delta_{kl}$. This algorithm is a special case of Corollary 22 with $\hat{H} = H_{N,N'}$, $\hat{P} = E_{n,mN}R$, and $\hat{Q} = SE_{pN',n}$.

(2) *Silverman's algorithm* (*Silverman*, 1971)

Determine a non-singular submatrix $F \in \mathbb{R}^{n \times n}$ of $\mathscr{H}(\tilde{\mathbf{G}})$. Say that F consists of the elements in the $(r_1, r_2, \ldots, r_n)$th rows and the $(k_1, k_2, \ldots, k_n)$th columns of $\mathscr{H}(\tilde{\mathbf{G}})$. Let σF denote the matrix formed by the elements in the $(r_1 + p, r_2 + p, \ldots, r_n + p)$th rows and the $(k_1, k_2, \ldots, k_n)$th columns of $\mathscr{H}(\tilde{\mathbf{G}})$, F_1 by those in the $(r_1, r_2, \ldots, r_n)$th rows and the first m columns, and F_2 by those in the first p rows and the $(k_1, k_2, \ldots, k_n)$th columns. Compute $A = (\sigma F)F^{-1}$, $B = F_1$, $C = F_2F^{-1}$, and $D = \tilde{\mathbf{G}}(0)$.

This algorithm is a special case of Corollary 22 with $\hat{H} = F$, $\hat{P} = I$, and $\hat{Q} = F^{-1}$.

Many other realization algorithms for impulse responses, better structured and more efficient numerically, have appeared in the literature. However it would be going too far to demonstrate how they can be viewed as special implementations of Corollary 22. Ho's and Silverman's algorithms have been mentioned primarily because of their historical importance. As a basis for all impulse response realization algorithms, Corollary 22 is both new and elegant: one can start with any maximal rank submatrix of $\mathscr{H}(\tilde{\mathbf{G}})$ and it is completely symmetric in the input and the output.

18. CONCLUSIONS

The appealing modelling language developed in the first sections of this paper allows discussion, on a set theoretic level, of modelling questions in a mathematically precise and consistent way. This yields the notion of an optimal exact model: the most powerful unfalsified model from a given model class. This is the model which explains the observations but as little else as possible. The existence of such an optimal model follows immediately from the intersection property. Linear time invariant finite dimensional systems defined in the way developed in Part I satisfy this condition as a consequence of the fact that their behaviour corresponds precisely to the closed linear shift invariant subspaces of $(\mathbb{R}^q)^T$. Note *en passant* that such a most powerful model may not exist in the traditional input/output setting. The framework set up this way allows an approach to what could be considered as truly the first question in identification theory: to find the most powerful unfalsified linear time invariant system which explains an observed vector time series. Rather concrete algorithms can indeed be set up both for the case in which passage from the time series to an (AR) model or directly to a state space model is required. A pressing refinement is the recursive implementation of these algorithms in terms of the notion of t-completeness.

The two most important contributions of this second part of this paper are the following. First, an elegant framework leading to the notion of the most powerful unfalsified model, a simple concept which, the author believes, will play a basic role in modelling expositions, has been provided. Second, several algorithms for obtaining the most powerful unfalsified (AR) or linear time invariant finite dimensional model for an observed time series have been outlined. This time series was totally arbitrary and not even an *a priori* distinction between inputs and outputs had to be made. These algorithms (or refinements thereof) should eventually be of interest in many identification, adaptive signal processing and adaptive control algorithms.

The 'real' problem in system identification is, of course, approximate modelling. That, in fact, is the subject of Part III of this paper.

REFERENCES

Chen, C. T. (1986). Techniques for identification of linear time-invariant multivariable systems. In Leondes, C. T. (Ed.), *Advances in Control and Dynamic Systems*, Vol. 24. Academic Press, New York (to appear).

Kalman, R. E., P. L. Falb and M. A. Arbib (1969). *Topics in Mathematical System Theory*, McGraw-Hill, New York.

Kalman, R. E. (1982). System identification from noisy data. In Bednarek, A. R. and L. Cesari (Eds), *Dynamical Systems II*, pp. 135–164. Academic Press, New York.

Kalman, R. E. (1983). Identifiability and modelling in econometrics. In Krishnaiah, P. R. (Ed.), *Developments in Statistics*, pp. 97–136. Academic Press, New York.

Popper, K. R. (1963). *Conjectures and Refutations*. Harper and Row, London.

Silverman, L. M. (1971). Realization of linear dynamical systems. *IEEE Trans. Aut. Control*, **AC-16**, 554–567.

APPENDIX N: NOTATION

A few new elements of notation which have been added since Part I are the following.

The collection of all subsets of the set S is denoted 2^S. The continuously differentiable maps from A into B is denoted $\mathscr{C}(A; B)$. The overbar is used for (topological) closure.

A family $\mathscr{M}$ of subsets of a given set will be said to have the intersection property if the intersection of any (possibly uncountably infinite) collection of elements of $\mathscr{M}$ is again an element of $\mathscr{M}$.

If S is a set and $\sim$ is an equivalence relation on it, then the map which takes $s \in S$ into $s(\text{mod}\sim) \in S(\text{mod}\sim)$ is called the natural projection.

A Hankel matrix is a matrix whose (i,j)th element depends on $(i-j)$ only. This nomenclature also applies to block matrices. I_n denotes the $(n \times n)$ identity matrix, while $E_{n,n'}$ denotes the selection matrix whose (i,j)th element equals δ_{ij}, with δ the Kronecker delta. A leading submatrix of a given matrix is simply the matrix formed by a number of its first rows and columns.

Let S be a set. A binary relation, $\leqslant$, on S is called a partial order if (i) $s \leqslant s$ for all $s \in S$, (ii) $\{s_1 \leqslant s_2$ and $s_2 \leqslant s_3\} \Rightarrow \{s_1 \leqslant s_3\}$, and (iii) $\{s_1 \leqslant s_2$ and $s_2 \leqslant s_1\} \Rightarrow \{s_1 = s_2\}$.

APPENDIX S: SPLITTING LINEAR RELATIONS

As can be seen from Proposition 20, the notion of relative row rank is closely related to the problem of state construction. This can be explained by means of a brief discussion of the crux of the problem of state realization: that of *splitting relations*. Only linear relations and finite dimensional vector spaces will be considered. This appendix serves as a preparation for the proof and to provide the intuition needed for Proposition 20 and Theorem 21.

Let $\mathscr{L}_1 = \mathbb{R}^{n_1}$, $\mathscr{L}_2 = \mathbb{R}^{n_2}$ and $\mathscr{L} = \mathbb{R}^{n_1+n_2} \cong \mathscr{L}_1 \times \mathscr{L}_2$. Let $\mathscr{B}$ be a linear subspace of $\mathscr{L}$. Further, let $X = \mathbb{R}^n$ and $\mathscr{B}_1 \subset \mathscr{L}_1 \times \mathscr{X}$, $\mathscr{B}_2 \subset \mathscr{X} \times \mathscr{L}_2$ be linear subspaces. $(\mathscr{X}, \mathscr{B}_1, \mathscr{B}_2)$ is a *splitting* of $\mathscr{B}$ if the composition of the relation $\mathscr{B}_1$ and $\mathscr{B}_2$ equals $\mathscr{B}$, i.e. if $\mathscr{B} = \mathscr{B}_1 \circ \mathscr{B}_2 := \{(l_1, l_2) \in \mathscr{L}_1 \times \mathscr{L}_2 \mid \exists x: (l_1, x) \in \mathscr{B}_1$ and $(x, l_2) \in \mathscr{B}_2\}$. In other words, in the extended relation $\mathscr{B}^{\text{ext}} := \{(l_1, x, l_2) \mid (l_1, x) \in \mathscr{B}_1$ and $(x_1, l_2) \in \mathscr{B}_2\}$, the variable x will 'split' the variables l_1 and l_2 or, said again in an alternative way, l_1 and l_2 will be *independent* (in a set theoretic sense) given x. If $\dim \mathscr{X}$ is as small as possible (under this splitting constraint) then $(\mathscr{X}, \mathscr{B}_1, \mathscr{B}_2)$ is a *minimal splitting* of $\mathscr{B}$. $\mathscr{X}$ in $\mathscr{B}_1 \subset \mathscr{L}_1 \times \mathscr{X}$ is called *accessible* if $P_1^x \mathscr{B}_1 = \mathscr{X}$ and $\mathscr{X}$ in $\mathscr{B}_2 \subset \mathscr{X} \times \mathscr{L}_2$ is called accessible if $P_2^x \mathscr{B}_2 = \mathscr{X}$. Here P_1^x (resp., P_2^x) denotes the natural projection of $\mathscr{L}_1 \times \mathscr{X}$ (resp., $\mathscr{X} \times \mathscr{L}_2$) onto $\mathscr{X}$. x in $\mathscr{B}_1$ is *induced* if there exists a partial map $S_1: \mathscr{L}_1 \to \mathscr{X}$ such that $\{(l_1, x) \in \mathscr{B}_1\} \Leftrightarrow (l_1 \in Do(S_1)$ and $x = S_1(l_1)\}$, i.e. if $\mathscr{B}_1$ is the graph of S_1. Similarly, x in $\mathscr{B}_2$ is induced if there exists a partial map $S_2: \mathscr{L}_2 \to \mathscr{X}$ such that $\{(x, l_2)\} \in \mathscr{L}_2\} \Leftrightarrow (l_2 \in Do(S_2)$ and $x = S_2(l_2)\}$. It is easy to see that because of linearity, $\{x$ is induced in $\mathscr{B}_1\} \Leftrightarrow \{\{(0, x) \in \mathscr{B}_1\} \Rightarrow \{x = 0\}\}$ and $\{x$ is induced in $\mathscr{B}_2\} \Leftrightarrow \{\{0, x) \in \mathscr{B}_2\} \Rightarrow \{x = 0\}\}$. Note also the analogy of accessibility with reachability and controllability, and of inducedness with observability and reconstructibility.

Two splittings $(\mathscr{X}', \mathscr{B}_1', \mathscr{B}_2')$ and $(\mathscr{X}'', \mathscr{B}_1'', \mathscr{B}_2'')$ are *equivalent* if there exists a linear bijection $S: \mathscr{X}' \to \mathscr{X}''$ such that $\{(l_1, x') \in \mathscr{B}_1'\} \Leftrightarrow \{(l_1, Sx') \in \mathscr{B}_1''\}$ and $\{(x', l_2) \in \mathscr{B}_2'\} \Leftrightarrow \{(Sx', l_2) \in \mathscr{B}_2''\}$. Finally, consider the following preorder $\leqq$ on the splitting triples of a given $\mathscr{B}$: $\{(\mathscr{X}', \mathscr{B}_1', \mathscr{B}_2') \leqq (\mathscr{X}'', \mathscr{B}_1'', \mathscr{B}_2'')\} :\Leftrightarrow \{\exists$ a partial surjection $S: \mathscr{X}' \to \mathscr{X}''$ such that $\{(l_1, x') \in \mathscr{B}_1'$ and $x' = f(x'')\} \Rightarrow \{(l_1, x'') \varepsilon \mathscr{B}_1''\}$ and $\{(x', l_2) \in \mathscr{B}_2'$ and $x' = f(x'')\} \Rightarrow \{(x'', l_2) \in \mathscr{B}_2'')\}$.

It is easy to give some canonical constructions for splittings of $\mathscr{B}$. Three important ones are the following.

Define $\mathscr{B}_1^0 := \{l_1 \in \mathscr{L}_1 \mid (l_1, 0) \in \mathscr{B}\}$, $\mathscr{B}_2^0 := \{l_2 \in \mathscr{L}_2 \mid (0, l_2) \in \mathscr{B}\}$ and $\mathscr{B}^0 = \mathscr{B}_1^0 \times \mathscr{B}_2^0$; $P_1: \mathscr{L}_1 \times \mathscr{X} \to \mathscr{L}_1$, the projection onto $\mathscr{L}_1$ and P_2 the projection onto $\mathscr{L}_1$ and $P_2: \mathscr{X} \times \mathscr{L}_2 \to \mathscr{L}_2$ the projection into $\mathscr{L}_2$.

(1) Take $\mathscr{X} = \mathscr{B}(\text{mod } \mathscr{B}^0)$.

$$\mathscr{B}_1 = \{(l_1, x) \mid \exists l_2': (l_1, l_2') \in \mathscr{B} \text{ and } x = (l_1, l_2')(\text{mod } \mathscr{B}^0)\}$$

$$\mathscr{B}_2 = \{(x, l_2) \mid \exists l_1': (l', l_2) \in \mathscr{B} \text{ and } x = (l_1', l_2)(\text{mod } \mathscr{B}^0)\}.$$

(2) Take $X = (P_1 \mathscr{B})(\text{mod } \mathscr{B}_1^0)$.

$$\mathscr{B}_1 = \{(l_1, x) \mid l_1 \in P_1 \mathscr{B} \text{ and } x = l_1(\text{mod } \mathscr{B}_1^0)\}$$

$$\mathscr{B}_2 = \{(x, l_2) \mid \exists l_2: (l_1, l_2) \in \mathscr{B} \text{ and } x = l_1(\text{mod } \mathscr{B}_1^0)\}.$$

(3) Take $\mathscr{X} = (\mathscr{P}_2 \mathscr{B})(\text{mod } \mathscr{B}_2^0)$.

$$\mathscr{B}_1 = \{(l_1, x) \mid \exists l_2: (l_1, l_2) \in \mathscr{B} \text{ and } x = l_2(\text{mod } \mathscr{B}_2^0)$$

$$\mathscr{B}_2 = \{(x, l_2) \mid l_2 \in P_2 \mathscr{B} \text{ and } x = l_2(\text{mod } \mathscr{B}_2^0)\}.$$

It is left to the reader to verify that in all the above cases the defined splittings of $\mathscr{B}$ with x in $\mathscr{B}_1$ and $\mathscr{B}_2$ are both accessible and induced. The next proposition will show that they are therefore also minimal.

An arbitrary splitting can be made into one in which x is accessible and reduced by the following *reduction procedure*. Let $(\mathscr{X}, \mathscr{B}_1, \mathscr{B}_2)$ be a splitting. If x in $\mathscr{B}_1$ is not accessible, define a new splitting $(\mathscr{X}', \mathscr{B}_1', \mathscr{B}_2')$ as follows:

$$\mathscr{X}' = \{x \in \mathscr{X} \mid \exists l_1: (l_1, x) \in \mathscr{B}_1\},$$

$$\mathscr{B}_1' = \{(l_1, x) \mid (l_1, x) \in \mathscr{B}_1\},$$

and

$$\mathscr{B}_2' = \{(x, l_2) \mid x \in \mathscr{X}' \text{ and } (x, l_2) \in \mathscr{B}_2\}.$$

By a similar construction if x in $\mathscr{B}_2$ is not accessible, x-accessibility in both $\mathscr{B}_1$ and $\mathscr{B}_2$ will be obtained. If x in $\mathscr{B}_1$ is not induced, define a new splitting $(\mathscr{X}', \mathscr{B}_1', \mathscr{B}_2')$ as follows: $\mathscr{X}' = \mathscr{X}(\mathrm{mod}\,\mathscr{B}_1(0))$ with

$$\mathscr{B}_1(0) := \{l_1 \in \mathscr{L}_1 \mid (l_1, 0) \in \mathscr{B}_1\},$$
$$\mathscr{B}_1' = \{(l_1, x') \in \mathscr{L}_1 \times \mathscr{X}' \mid \exists (l_1, x) \in \mathscr{B}_1 \text{ such that } x' = x(\mathrm{mod}\,\mathscr{B}_1(0))\},$$
$$\text{and } \mathscr{B}_2 = \{(x', l_2) \in \mathscr{X}' \times \mathscr{L}_2 \mid \exists (x, l_2) \in \mathscr{B}_2 \text{ such that } x' = x(\mathrm{mod}\,\mathscr{B}_2(0)\}.$$

Of course, a similar construction can be applied if x in $\mathscr{B}_2$ is not induced. It is easily verified that by applying these constructions (if necessary all four of them), a splitting in which x is accessible and induced in both $\mathscr{B}_1$ and $\mathscr{B}_2$ is obtained.

Proposition S. Let $\mathscr{B} \subset \mathscr{L}_1 \times \mathscr{L}_2$ be a linear subspace with $\mathscr{L}_1 = \mathbb{R}^{n_1}$ and $\mathscr{L}_2 = \mathbb{R}^{n_2}$. Define P_1, P_2, $\mathscr{B}_1^0$, $\mathscr{B}_2^0$ and $\mathscr{B}^0$ as before. Let $n^*(\mathscr{B})$ denote the minimal splitting dimension of $\mathscr{B}$.

(1) $n^*(\mathscr{B}) = \dim P_1\mathscr{B} - \dim \mathscr{B}_1^0 = \dim P_2\mathscr{B} - \dim \mathscr{B}_2^0$
$= \dim P_1\mathscr{B} + \dim P_2\mathscr{B} - \dim \mathscr{B}$
$= \dim \mathscr{B} - \dim \mathscr{B}_1^0 - \dim \mathscr{B}_2^0$.

(2) Let $(\mathscr{X}_1, \mathscr{B}_1, \mathscr{B}_2)$ be a splitting. Then the following are equivalent:
  (2.1) it is minimal;
  (2.2) x is both accessible and induced in both $\mathscr{B}_1$ and $\mathscr{B}_2$;
  (2.3) any other splitting $(\mathscr{X}', \mathscr{B}_1', \mathscr{B}_2')$ satisfies $(\mathscr{X}, \mathscr{B}_1, \mathscr{B}_2) \leqslant (\mathscr{X}', \mathscr{B}_1', \mathscr{B}_2')$.

(3) (*Isomorphism*): All minimal splittings are equivalent.

Proof. First, if $(\mathscr{X}, \mathscr{B}_1, \mathscr{B}_2)$ satisfies 2.2, then $P_1\mathscr{B}(\mathrm{mod}\,\mathscr{B}_1^0) \cong \mathscr{X} \cong P_2\mathscr{B}(\mathrm{mod}\,\mathscr{B}_2^0)$. In order to see this, observe that accessibility and inducedness of x in $\mathscr{B}_1$ implies that there exists a surjection $S_2: P_1\mathscr{B}_1 \to \mathscr{X}$ such that $\{(l_1, x) \in \mathscr{B}_1\} \Leftrightarrow \bar{S}_1 l_1 = x\}$. Consequently $\{(l_1, 0) \in \mathscr{B}_1\} \Leftrightarrow (l_1, 0) \in \mathscr{B}\}$. Hence $\ker S_1 = \mathscr{B}_1^0$. A similar result holds, of course, for $\mathscr{B}_2$. Therefore the diagrams in Fig. 9 commute and $\bar{S}_1$ and $\bar{S}_2$ are bijections. Now, since $\bar{S}_1$ and $\bar{S}_2$ are bijections, this implies $P_1\mathscr{B}_1(\mathrm{mod}\,\mathscr{B}_1^0)\ \mathscr{X} \cong P_2\mathscr{B}_2(\mathrm{mod}\,\mathscr{B}_2^0)$. Also it follows that all splittings $(\mathscr{X}, \mathscr{B}_1, \mathscr{B}_2)$ for which x is accessible and induced in both $\mathscr{B}_1$ and $\mathscr{B}_2$, are equivalent.

From the reduction procedure, it follows that minimality implies 2.2. Furthermore since all splittings satisfying 2.2 have the same dimension, 2.1 $\Leftrightarrow$ 2.2. This also implies 3. Combining this with the reduction procedure shows that 2.2 $\Leftrightarrow$ 2.3.

Now consider the dimension formulas. It has already been shown that $n^*(\mathscr{B}) = \dim P_1\mathscr{B} - \dim \mathscr{B}_1^0 = \dim P_2\mathscr{B} - \dim \mathscr{B}_2^0$. From the construction of the canonical splitter 1, which satisfies 2.2, $n^*(\mathscr{B}) = \dim P_1\mathscr{B} + \dim P_2\mathscr{B} - \dim \mathscr{B}$. These three expressions for $n^*(\mathscr{B})$ finally yield $n^*(\mathscr{B}) = \dim \mathscr{B} - \dim \mathscr{B}_1^0 - \dim \mathscr{B}_2^0$. □

Note that the proof of Proposition S implies that the canonical splittings 1, 2 and 3 are all minimal. Furthermore, if $\mathscr{B}$ is a linear relation, and $(\mathscr{X}, \mathscr{B}_1, \mathscr{B}_2)$ is a minimal splitting of $\mathscr{B}$, then there exist surjections $S_1: P_1\mathscr{B}_1 \to \mathscr{X}$ and $S_2: P_2\mathscr{B}_2 \to \mathscr{X}$ such that $\{(l_1, l_2) \in \mathscr{B}\} \Leftrightarrow \{S_1 l_1 = S_2 l_2\}$. It is useful to think of $S_1 l_1 = S_2 l_2$ as the (unique up to isomorphism) *common features* in the variables of $\mathscr{L}_1$ and $\mathscr{L}_2$ induced by the relation $\mathscr{B}$. Now consider the problem of *feature extraction*, i.e. that of determining the common features among two sets of variables on the basis of a sequence of observations.

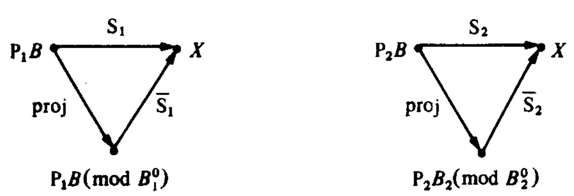


FIG. 9.

Let $a_k \in \mathscr{L}_1 = \mathbb{R}^{n_1}$ and $b_k \in \mathscr{L}_2 = \mathbb{R}^{n_2}$, $k \in \mathbb{N}$, be a sequence of observations. Define $\mathscr{B} = \mathrm{span}\ (c_k \mid k \in \mathbb{N})$ and consider the problem of computing:

(i) $n^*(\mathscr{B})$, the number of common features in the as and the bs;

(ii) a minimal splitting $(\mathscr{X}, \mathscr{B}_1, \mathscr{B}_2)$ of $\mathscr{B}$; and

(iii) the common features in the observations, i.e. $x_k \in \mathscr{X}$, $k \in \mathbb{N}$, such that $(a_k, x_k) \in \mathscr{B}_1$ and $(x_k, b_k) \in \mathscr{B}_2$.

In the context of application to dynamical systems, a can be thought of as the strict past, b as the present and future, and x as the (minimal) state to be constructed.

This problem is a straightforward application of the theory just outlined. Let $A := \mathrm{col}\,(a_1, a_2, \ldots, a_N)$, $B := \mathrm{col}\,(b_1, b_2, \ldots, b_N)$, and $C := \mathrm{col}\ (A, B)$. Then $n^*(\mathscr{B}) = r(A; B)$. This shows the relevance of the notion of relative rank. This formula follows immediately from Proposition S, item 1, since $r(A; B) = \mathrm{rank}\ A + \mathrm{rank}\ B - \mathrm{rank}\ C = \dim P_1\mathscr{B} + \dim P_2\mathscr{B} - \dim \mathscr{B} = n^*(\mathscr{B})$, the dimension of the minimal splitting.

In order to construct the common features, proceed as follows. Define $\mathscr{K} = B \ker A$, $\mathscr{X} = \mathrm{im}\ B(\mathrm{mod}\,\mathscr{K})$, and $x_k = b_k\ (\mathrm{mod}\,\mathscr{K})$, $k \in \mathbb{N}$. That this is indeed a solution follows immediately from the canonical splitter 3 since, with $\mathscr{B} = \mathrm{im\,col}\ (A, B)$, $\mathscr{B}_2^0 = \mathscr{K}$.

APPENDIX P: PROOFS

Proof of Proposition 11

Take $M_Z^* = \bigcap_{Z \supset M \in \mathscr{M}} M$. Clearly M_Z^* is non-empty (since there exists at least one $Z \subset M \in \mathscr{M}$), $M_Z^* \in \mathscr{M}$ (by the intersection property), and $\{Z \subset M \in \mathscr{M}\} \Rightarrow \{M_Z^* \subset M\}$ (since M is one of the intersected sets). □

Proof of Proposition 12

To show that $(\mathbb{R}^q)^T \in \mathscr{M}$, take $R = 0$. Three proofs of the intersection property will be given.

Proof 1. The first proof starts from the polynomial matrices themselves. Let $R_\alpha \in \mathbb{R}^{g_\alpha \times q}[s]$, $\alpha \in A$, with A a parameter set, and consider $\mathscr{B}_A := \{\mathrm{w}: T \to \mathbb{R}^q \mid R_\alpha(\sigma)\mathrm{w} = 0, \forall \alpha \in A\}$. Assume without loss of generally that $g = 1$ for all $\alpha \in A$.

(i) Let $T = \mathbb{Z}_+$. Consider the subset $\mathscr{N} \subset \mathbb{R}^{1 \times q}[s]$ consisting of all elements $r \in \mathbb{R}^{1 \times q}[s]$ such that $r(\sigma)\mathrm{w} = 0$ for all $\mathrm{w} \in \mathscr{B}_A$. Obviously $\{r(s) \in \mathscr{N}\} \Rightarrow \{p(s)r(s) \in \mathscr{N}\}\ \forall p \in \mathbb{R}[s]$. Consequently $\mathscr{N}$ is a submodule of $\mathbb{R}^{1 \times q}[s]$ viewed as a q-dimensional module over $\mathbb{R}[s]$. Hence (6, p. 247) $\mathscr{N}$ is finitely generated, i.e. $\exists r_1, r_2, \ldots, r_g \in \mathbb{R}^{1 \times q}[s]$ such that $\mathscr{N} = r_1\mathbb{R}[s] + r_2\mathbb{R}[s] + \ldots + r_g\mathbb{R}[s]$ implying that $\mathscr{B}_A = \{\mathrm{w}: T \to \mathbb{R}^q \mid r_i(\sigma)\mathrm{w} = 0, i \in \mathbf{g}\} = \mathscr{B}(R)$ with $R := \mathrm{col}(r_1, r_2, \ldots, r_g)$.

(ii) Let $T = \mathbb{Z}$. Repeat the above proof with $\mathbb{R}[s]$ replaced by $\mathbb{R}[s, s^{-1}]$.

Proof 2. The second proof starts from the characterization of $\mathscr{M}$ in terms of linearity, time invariance, and completeness (see Theorem 5). Let $\mathscr{B}_\alpha \in \mathscr{M}$, $\alpha \in A$. Then the $\mathscr{B}_\alpha$s are linear, shift invariant, and complete. Obviously $\bigcap_{\alpha \in A} \mathscr{B}_\alpha$ is also linear. That it is shift invariant follows from $\sigma^t \bigcap_{\alpha \in A} \mathscr{B}_\alpha = \bigcap_{\alpha \in A} \sigma^t\mathscr{B}_\alpha \subset \bigcap_{\alpha \in A} \mathscr{B}_\alpha$. To show completeness, observe that

$$\left\{\mathrm{w} \in \bigcap_{\alpha \in A} \mathscr{B}_\alpha\right\} \Leftrightarrow \{\mathrm{w} \in \mathscr{B}_\alpha, \forall \alpha \in A\}$$
$$\Leftrightarrow \{\mathrm{w}|_{T \cap [t_0, t_1]} \in \mathscr{B}_\alpha|_{T \cap [t_0, t_1]}, \forall \alpha \in A \text{ and } -\infty < t_0 \leqslant t_1 < \infty\}$$
$$\Leftrightarrow \{\mathrm{w}\}_{T \cap [t_0, t_1]} \in \Big(\bigcap_{\alpha \in A} \mathscr{B}_\alpha\Big)\Big|_{T \cap [t_0, t_1]}, \forall -\infty < t_0 \leqslant t_1 < \infty\}$$

Hence $\bigcap_{\alpha \in A} \mathscr{B}_\alpha$ is linear, time invariant, and complete. Now apply Theorem 5.

Proof 3. By Proposition 4 and Theorem 5, $\mathscr{M} = \mathbf{L}$, the family of linear shift invariant, closed subspaces of $(\mathbb{R}^q)^T$, equipped with the topology of pointwise convergence. Clearly $\mathbf{L}$ has the intersection property. Now apply Proposition 11. □

Proof of Theorem 14

The notation of Section 8 is used. Step 3 immediately yields $R^*_{\tilde{w}}(\sigma)\tilde{\mathbf{w}} = \mathbf{0}$, i.e. the model is unfalsified. In order to prove that it is the most powerful unfalsified (AR) model, show that $\mathscr{B}_t(R^*_{\tilde{w}}) = \text{span}\{\sigma^\tau \tilde{\mathbf{w}}|_{T\cap[0,t]}, \tau \in T\}$ for all $t \in T$. From the definition of the r_is it follows that for $j_i' + ql_i \leqslant t \leqslant j_{i+1}' + ql_{i+1}$, $i \in (\mathbf{p}-1)$, or for $t \geqslant j_p' + 9l_p$, there holds that $\{r_1(s), \ldots, s^{t-l_1}r_1(s), r_2(s), \ldots, s^{t-l_2}r_2(s), \ldots, r_i(s), \ldots, s^{t-l_i}r_i(s)\}$ span(span$\{\sigma^\tau \tilde{\mathbf{w}}|_{T\cap[0,t]}, \tau \in T\})^\perp$, while for $t < j_1 + ql_1$, there holds span $\tilde{\mathscr{B}}_t^\perp(R^*_{\tilde{w}}) = 0 = (\text{span}\{\sigma^\tau \tilde{\mathbf{w}}|_{T\cap[0,t]}, \tau \in T\})^\perp$. In order to prove the other claims of the theorem, observe that it is trivial to see that $R^*_{\tilde{w}}$ is a row proper matrix polynomial. Hence by Theorem 6 it defines a shortest lag description in the case $T = \mathbb{Z}_+$. In the case $T = \mathbb{Z}$ it must also be proved that $R^*_{\tilde{w}}(0)$ has full row rank. Assume to the contrary that $\exists i \in \mathbf{p}$ and $\alpha_k \in \mathbb{R}$, $k \in (i-1)$, such that $a_{i,0} = \sum_{k=1}^{i-1} \alpha_k a_{k,0}$. Now consider $r(s) = r_i(s) - \sum_{k=1}^{i-1} \alpha_k r_k(s)$. Then obviously $r(\sigma)\tilde{\mathbf{w}} = 0$ and $r(0) = 0$. Hence $r'(s) := s^{-1}r(s)$ satisfies $r'(\sigma)\tilde{\mathbf{w}} = \mathbf{0}$. Examination of this relation shows that the $(j_i' + q(l_i - 1))$th row of $\mathscr{H}(\tilde{\mathbf{w}})$ is already linearly dependent on the preceding rows. This contradicts the definition of l_i, and proves that $R^*_{\tilde{w}}$ defines a shortest lag description. The statements about the inputs, outputs, and the shortest lag structure are now obvious. □

Proof of Theorem 15

Let $\mathscr{B}_t := \text{span}\{\sigma^\tau \tilde{\mathbf{w}}|_{T\cap[0,t]}, \tau \in T\}$. By assumption $\dim \mathscr{B}_{t+1} - \dim \mathscr{B}_t =$ constant for $t \geqslant t'$. Further $r_1, r_2, \ldots, r_g$ span $\mathscr{B}_{t'}^\perp$. Now apply Theorem 7. □

Proof of Proposition 16

Consider first the following lemma.

Lemma. Let $A \in \mathbb{R}^{n\times n}$, $C \in \mathbb{R}^{p\times n}$, and $t \in \mathbb{Z}^+$, and define

$$\mathscr{K}_t := \bigcap_{0 \leq t' \leq t} A^{-t'} \ker C$$

$$\mathscr{K}_t' := A^{t+1}\mathscr{K}_t$$

$$\mathscr{K}_t'' := \mathscr{K}_{t-1}' \cap \ker C = A^t \mathscr{K}_t.$$

Then, if (A, C) is observable, there holds:

(i) $\dim \mathscr{K}_t = \dim \mathscr{K}_t' = \dim \mathscr{K}_t''$;

(ii) $\{\mathscr{K}_t = 0\} \Leftrightarrow \left\{\bigcap_{t'=0}^{t} (\mathscr{K}_{t'} + \mathscr{K}_{t-t'}') = 0\right\}$.

Proof. (i) It will be proved that $K_t \cap \ker A^{t+1}$ is an A-invariant subspace contained in $\ker C$. Since $\mathscr{K}_t \subset \ker C$ the last part is obvious. Also, $\{a \in \mathscr{K}_t \cap \ker A^{t+1}\} \Leftrightarrow \{Ca = CAa = \cdots = CA^t a = 0;\ A^{t+1}a = 0\} \Rightarrow \{Aa \in \mathscr{K}_t \subset \ker A^{t+1}\}$. Hence, by observability, $\mathscr{K}_t \cap \ker A^{t+1} = 0$ and $\dim \mathscr{K}_t' = \dim A^{t+1}\mathscr{K}_t = \dim \mathscr{K}_t$. The proof that $\dim \mathscr{K}_t'' = \dim \mathscr{K}_t$ is completely analogous.

(ii) ($\Rightarrow$): Assume that $\mathscr{K}_t = 0$. It will be shown that $\mathscr{L} := \bigcap_{t'=0}^{t} (\mathscr{K}_{t'} + \mathscr{K}_{t-t'}')$ is an A-invariant subspace contained in $\ker C$ which, by observability, yields $\mathscr{L} = 0$. The term with $t' = 0$ equals $\mathscr{K}_0 + \mathscr{K}_t = \ker C$, hence $\mathscr{L} \subset \ker C$. Further,

$$A\mathscr{L} = A(\mathscr{L} \cap \ker C) = \bigcap_{t'=0}^{t} ((A\mathscr{K}_{t'} + A\mathscr{K}_{t-t'}') \cap A \ker C)$$

which, since $A \ker C \supset A\mathscr{K}_{t'}$ equals $\bigcap_{t'=0}^{t} (A\mathscr{K}_{t'} + A(\mathscr{K}_{t-t'}' \cap \ker C))$. Now, $\mathscr{K}_{t'+1} = A^{-1}\mathscr{K}_t \cap \ker C$, hence $A\mathscr{K}_{t'+1} = \mathscr{K}_{t'} \cap A \ker C$. Similarly $\mathscr{K}_{t'+1}' = A(\mathscr{K}_{t'}' \cap \ker C)$.

Hence $A\mathscr{L} = \bigcap_{t'=0}^{t} ((\mathscr{K}_{t'-1} \cap A \ker C) + \mathscr{K}_{t-t'+1})$ with $\mathscr{K}_{-1} := \mathbb{R}^n$. Using $\mathscr{K}_{t-t'+1} \subset A \ker C = \mathscr{K}_t + \mathscr{K}_0'$ this yields

$$A\mathscr{L} = \bigcap_{t'=0}^{t} ((\mathscr{K}_{t'-1} + \mathscr{K}_{t-t'+1}) \cap A \ker C) = \bigcap_{t'=0}^{t} (\mathscr{K}_{t'} + \mathscr{K}_{t-t'}) = \mathscr{L},$$

which gives $A\mathscr{L} \subset \mathscr{L}$, as desired.

($\Leftarrow$): This is an immediate consequence of

$$\mathscr{K}_0 \supset \mathscr{K}_1 \supset \cdots \supset \mathscr{K}_t, \text{hence } \mathscr{K}_t \subset \bigcap_{t'=0}^{t} (\mathscr{K}_{t'} + \mathscr{K}_{t-t'}'). \quad □$$

Return now to the proof of Proposition 16. Actually, as will be proved, all the statements of Proposition 16 are equivalent to $\mathscr{K}_t = 0$. Since these claims are basis free, the basis in $W = \mathbb{R}^q$ may as well be chosen to advantage, namely such that $\mathscr{B}^+$ is the external behaviour on $T = \mathbb{Z}_+$ of the observable system $\sigma\mathbf{x} = A\mathbf{x} + B\mathbf{u}$, $\mathbf{y} = C\mathbf{x}$, $\mathbf{w} = \text{col}(\mathbf{u}, \mathbf{y})$. Now, by observability, every element $\mathbf{w} \in \mathscr{B}^+$ is uniquely specified by a $\mathbf{u}$: $\mathbb{Z}_+ \to \mathbb{R}^m$ and an $\mathbf{x}(0) \in \mathbb{R}^n$. Furthermore $\mathbf{w} \in \mathscr{B}^0$ corresponds precisely to $\mathbf{x}(0) = 0$, while $\mathbf{w} \in \mathscr{B}^1$ corresponds to $(\mathbf{x}(0) = 0 \wedge \mathbf{u}(0) = 0)$. This, together with the results of Section 7, yields the following dimension formulas:

$$\dim \mathscr{B}_t = (t+1)m + n - \dim \mathscr{K}_t$$
$$\dim (\mathscr{B}^0)_t = (t+1)m$$
$$\dim (\mathscr{B}^1)_t = tm.$$

Further, $\dim \mathscr{B}^+ (\text{mod}\, \mathscr{B}^0) = n$

and $\dim \mathscr{B}^+ (\text{mod}\, \mathscr{B}^1) = n + m.$

Also, from Section 7, $\{\rho_{t+1} = \rho_\infty\} \Leftrightarrow \{\mathscr{K}_t = 0\}$. From these expressions the equivalence of (i), (ii), and (iii) follows, since they are all equivalent to $\mathscr{K}_t = 0$.

Now consider (iv). Observe the following two facts. First, let $\mathbf{x}(0) \in \mathbb{R}^n$, $\mathbf{u}$: $[0, t] \cap \mathbb{Z}_+ \to \mathbb{R}^m$ be one element which generates a given $\mathbf{w} \in \mathscr{B}_t$. Then it is easy to give all pairs $(\mathbf{x}(0), \mathbf{u})$ which generate this $\mathbf{w}$. In fact, $(\mathbf{x}(0) + \mathscr{K}_t, \mathbf{u})$ are all such elements. Second, if $\mathbf{w} \in \mathscr{B}_t$ is such that $\mathbf{w}(\tau) = \mathbf{0}$ for $0 \leqslant \tau < t' \leqslant t$ then the corresponding underlying $\mathbf{x}(t')$ lies in $\mathscr{K}_{t'}'$.

Now let $\mathbf{w} \in \bar{\mathscr{B}}_{t+1}^0$. Equivalently $(\sigma_{t+1}^*)^{t'}\mathbf{w} \in \mathscr{B}_{t+1}$ for $0 \leqslant t' \leqslant t+1$. Let $a + \mathscr{K}_t$ be the linear variety of initial states explaining that, in particular, $\mathbf{w} \in \mathscr{B}_{t+1}$. Then, by what has just been shown, the underlying state at time t explains that $(\sigma_{t+1}^*)^{t'}\mathbf{w} \in \mathscr{B}_{t+1}$ must on the one hand lie in $a + \mathscr{K}_{t-t'}$, and on the other hand in $\mathscr{K}_{t'-1}$. Hence $(a + \mathscr{K}_{t-t'}) \cap \mathscr{K}_{t'-1}'$ is non-empty. Equivalently $a \in \mathscr{K}_{t-t'} + \mathscr{K}_{t'-1}'$. Hence the initial state corresponding to $\mathbf{w} \in \bar{\mathscr{B}}_{t+1}^0$ is zero if $\bigcap_{t'=0} (\mathscr{K}_{t-t'} + \mathscr{K}_{t'-1}') = 0$. By the lemma, this is the equivalent to $\mathscr{K}_t = 0$. This proves that $\{\mathscr{K}_t = 0\} \Rightarrow \{\bar{\mathscr{B}}_{t+1}^0 = (\mathscr{B}^0)_{t+1}\}$. To show the converse, assume $\mathscr{K}_t \neq 0$. Then by the lemma $\mathscr{K}_t' \neq 0$. Now by what has been shown above, any initial state $\mathbf{x}(0) \in \mathscr{K}_t'$ and any $\mathbf{u}$ will generate a $\mathbf{w} \in \bar{\mathscr{B}}_{t+1}^0$, then there is a response $\mathbf{w} \in \bar{\mathscr{B}}_{t+1}^0$ such that $\mathbf{w}(0) = \cdots = \mathbf{w}(t) = 0$, $\mathbf{u}(t+1) = 0$ but $\mathbf{y}(t+1) \neq 0$. Clearly this yields a $\mathbf{w} \notin (\mathscr{B}^0)_{t+1}$. In order to establish such a response it suffices to show that $C\mathscr{K}_t \neq 0$. Assume that instead $\mathscr{K}_t' \subset \ker C$. Then $\mathscr{K}_t'' \subset \ker C$ satisfies $A\mathscr{K}_t'' = A^{t+1}\mathscr{K}_t = \mathscr{K}_t' \subset \mathscr{K}_{t-1}' \cap \ker C = \mathscr{K}_t''$. Hence $\mathscr{K}_t''$ is then A-invariant and contained in $\ker C$, thus $\mathscr{K}_t'' = 0$, which implies, by the lemma, that $\mathscr{K}_t' = 0$. This gives the desired contradiction.

Finally, $\{\mathscr{K}_t = 0\} \Leftrightarrow \{\bar{\mathscr{B}}_{t+1}^0 = (\mathscr{B}^0)_{t+1}\}$, hence (iv).

Now (iv) $\Rightarrow$ (v) is obvious. The proof that (v) $\Rightarrow \{\mathscr{K}_t = 0\}$ is identical to that of (iv) $\Rightarrow \{\mathscr{K}_t = 0\}$. □

Proof of Theorem 17 (*Primal version*)

By Proposition 16, for $t \geqslant t^*$, $(\mathscr{B}^0)_t = \bar{\mathscr{B}}_t^0$ and $(\mathscr{B}^1)_t = \bar{\mathscr{B}}_t^1$. Hence $\mathscr{B}_t\ (\text{mod}\, \bar{\mathscr{B}}_t^0) \cong \mathscr{B}\, (\text{mod}\, \mathscr{B}^0)$ and $\mathscr{B}_t\ (\text{mod}\, \bar{\mathscr{B}}_t^1) \cong \mathscr{B}\ (\text{mod}\, \mathscr{B}^1)$. This shows that the mappings M_1, M_1^t, $\bar{M}_1^t$ and M_2, M_2^t, $\bar{M}_2^t$ in the truncated structure diagram are isomorphic. The result follows from Theorem 9. □

Proof of the dual versions of Proposition 16 *and Theorem* 17

First observe that as a consequence of Section 7,

$$\min\{t \in \mathbb{Z}_+ \,|\, \rho_t = \rho_\infty\} =: t^* = \partial_{g^*}^* = \min_R\{\partial(R) \,|\, \mathscr{B}(R) = \mathscr{B}\}.$$

The results now follow by straightforward dualization. The details are left to the reader. □

Proof of Theorem 19

It is clear that $\mathscr{B}(R^*_{\tilde{w}}) = \text{span}(\sigma^t \tilde{\mathbf{w}} \,|\, t \in T)$. For simplicity this behaviour is denoted by $\mathscr{B}$. Then $\mathscr{B}_t = \text{span}\{\tilde{\mathbf{w}}_t(\tau); \tau \in T\}$. From this it follows that Steps 1, 2, and 3 of Algorithm 5 determine a t' satisfying the conditions of Theorem 17, such that the realization procedure can be applied with the truncated structure diagram shown in Fig. 10 (incorporating the refinement mentioned following the statement of Theorem 17). To show that the matrices M_1 and M_2 defined in Step 4 of Algorithm 5 indeed correspond to those shown in the above diagram, identify the state space X with $\text{im}\, HQ_x \cong \mathscr{B}_{t'}\ (\text{mod}\, \mathscr{B}^0_{t'})$ and the input space U with $\text{im}\, HQ_u \cong \mathscr{B}^0_{t'}\ (\text{mod}\, \mathscr{B}^1_{t'})$. Now the shift, σ, maps HQ_x into σHQ_x and HQ_u into σHQ_u, while the evaluation map π^0 maps HQ_x into H^0Q_x and HQ_u into H^0Q_u. Take the columns of HQ_x as a basis for X and those of HQ_u as a basis for U. Then $H^0[Q_x \vdots Q_u] = [C' \vdots D']$. In order to determine A' and B', $\sigma H[Q_x \vdots Q_u]$ should now be expressed as a linear combination of the columns of HQ_x and elements of $\mathscr{B}^0_{t'}$, i.e. $\sigma H[Q_x \vdots Q_u] = \sigma H[A' \vdots B'] + F_0$ should be solved for the unknowns A', B', and F_0, with $\text{im}\, F_0 \subset \mathscr{B}^0_{t'}$. However, in the end, only A' and B' are needed. Now, premultiply both sides of this equation with P in order to obtain $PH[Q_x \vdots Q_u] = [A' \vdots B']$, as required. □

Proof of Proposition 20 *and Theorem* 21

Preamble. The reader is referred to Appendix S where the basic ideas underlying the construction of Algorithm 6 have been explained in a simple 'static' setting.

Define

$$\mathscr{L}^+_t = \{\mathbf{w}: \mathbb{Z} \cap [0, t] \to \mathbb{R}^q\} \cong \mathbb{R}^{q(t+1)}$$

and

$$\mathscr{L}^-_t = \{\mathbf{w}: \mathbb{Z} \cap [-t, 0) \to \mathbb{R}^q\} \cong \mathbb{R}^{qt}.$$

Let $T = \mathbb{Z}$ and consider $\mathscr{B} \in \mathbf{L}$. Define for t', $t'' \in \mathbb{Z}_+$, $\mathscr{B}(t', t'') := \mathscr{B}|_{\mathbb{Z} \cap [-t', t'']}$. Clearly $\mathscr{B}(t', t'')$ can be viewed in a natural way as a linear subspace of $\mathscr{L}^-_{t'} \times \mathscr{L}^+_{t''}$. As such it admits a minimal splitting, say $(\mathscr{X}', \mathscr{B}'_1, \mathscr{B}'_2)$, which by Proposition S is unique up to isomorphism. Now define $\mathscr{B}'_i \subset (\mathbb{R}^q \times \mathscr{X})^{\mathbb{Z}}$ as follows: $\{(\mathbf{w}, \mathbf{x}) \in \mathscr{B}'_i\} \Leftrightarrow \{(\sigma^t \mathbf{w}|_{\mathbb{Z} \cap [-t', 0)}, \mathbf{x}(t)) \in \mathscr{B}'_1$ and $(\mathbf{x}(t), \sigma^t \mathbf{w}|_{\mathbb{Z} \cap [0, t'']}) \in \mathscr{B}'_2$, for each $t \in \mathbb{Z}\}$.

Now if $t'' + 1 \geqq t^*$ and $t' \geqq t^* := \min_{t \in \mathbb{Z}_+} \{t \mid \rho_t = \rho_\infty\}$, then $\mathscr{B}'_i$ defines a minimal linear time invariant state space representation of $\mathscr{B}$. Linearity and time invariance are obvious. In order to see that it is a minimal state space representation, assume that $\mathscr{B}_i \subset (\mathbb{R}^q \times \mathbb{R}^n)^{\mathbb{Z}}$ is a minimal linear time invariant state space representation of $\mathscr{B}$. Define $\mathscr{B}_1 \subset \mathscr{L}^-_{t'} \times \mathbb{R}^n$ by $\mathscr{B}_1 := \{(\mathbf{w}, \mathbf{x}) \mid \exists (\mathbf{w}', \mathbf{x}) \in \mathscr{B}_i$ such that $\mathbf{w} = \mathbf{w}'|_{\mathbb{Z} \cap [t', 0)}$ and $\mathbf{x}(0) = x\}$ and $\mathscr{B}_2 := \{(x, \mathbf{w}) \mid \exists (\mathbf{x}, \mathbf{w}')$ such that $\mathbf{w} = \mathbf{w}'|_{\mathbb{Z} \cap [0, t'']}$ and $\mathbf{x}(0) = x\}$. It is an easy consequence of the axiom of state that $(\mathbb{R}^n, \mathscr{B}_1, \mathscr{B}_2)$ is a splitting of $\mathscr{B}(t', t'')$. From the fact that $\mathscr{B}_i$ is minimal, it follows immediately that x in both $\mathscr{B}_1$ and $\mathscr{B}_2$ is accessible. Exploiting the representation for $\mathscr{B}_i$ and the lemma used in the proof of Proposition 16 it follows, using minimality, which implies observability, that $\{(0, x) \in \mathscr{B}_1\} \Leftrightarrow \{\mathscr{X}_{t'-1} = 0\}$ and $\{(x, 0) \in \mathscr{B}_2\} \Leftrightarrow \{\mathscr{X}_{t''} = 0\}$. Further, $\{\mathscr{X}_t = 0\} \Leftrightarrow \{\mathscr{X}'_t = 0\} \Leftrightarrow \{t + 1 \geqq t^*\}$. Consequently, if $t' \geqq t^*$ and $t'' + 1 \geqq t^*$, then x is induced in both $\mathscr{B}_1$ and $\mathscr{B}_2$. Hence if $t' \geqslant t^*$ and $t'' + 1 \geqq t^*$, then $(\mathbb{R}^n, \mathscr{B}_1, \mathscr{B}_2)$ is a minimal splitting of $\mathscr{B}(t', t'')$, consequently, by Proposition S, item 3, isomorphic to $(\mathscr{X}', \mathscr{B}'_1, \mathrm{B}'_2)$, and this $\mathscr{B}'_i$ is a minimal state space representation of $\mathscr{B}$.

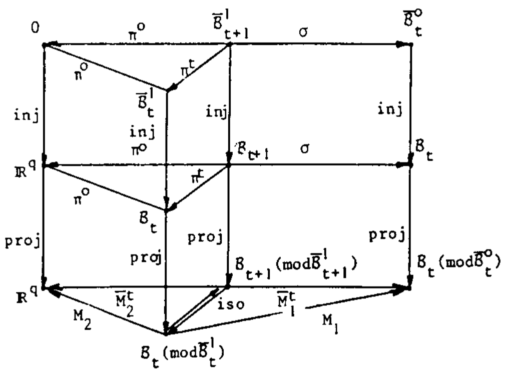


FIG. 10.

Proof of Proposition 20.

Let $\mathscr{H}^{t'}_-(\tilde{\mathbf{w}})$ and $\mathscr{H}^{t''}_+(\tilde{\mathbf{w}})$ denote the bottom qt' rows of $\mathscr{H}_-(\tilde{\mathbf{w}})$ and the top $q(t'' + 1)$ rows of $\mathscr{H}_+(\tilde{\mathbf{w}})$ respectively. Then im col $(\mathscr{H}^{t'}_-(\tilde{\mathbf{w}}), \mathscr{H}^{t'}_+(\tilde{\mathbf{w}})) = \mathscr{B}(R^*_\#)\ (t', t'')$. By the preamble and Appendix S this yields that for $t' \geqq t^*$ and $t'' + 1 \geqq t^*$, $r(\mathscr{H}^{t'}_-(\tilde{\mathbf{w}}); \mathscr{H}^{t''}_+(\tilde{\mathbf{w}}))$ equals the minimal state space dimension of $\mathscr{B}(R^*_\#)$, as desired. □

Proof of Theorem 21.

First observe the following refinement of the construction of Appendix S. Let $\mathscr{B} \subset \mathscr{L}_1 \times \mathscr{L}_2$, $T_1: \mathscr{L}_1 \to \mathscr{L}'_1$, $T_2: \mathscr{L}_2 \to \mathscr{L}'_2$ all be linear. Denote $T_1 \mathscr{B} T_2 := \{(l'_1, l'_2) \mid \exists (l_1, l_2) \in \mathscr{B}$ such that $l' = T_1 l_1\ l'_2 = T_2 l_2\}$. Now if $(\mathscr{X}, \mathscr{B}_1, \mathscr{B}_2)$ is a splitting for $\mathscr{B}$, then it follows that $(X, T_1 \mathscr{B}_1, \mathscr{B}_2 T_2)$ is a splitting for $T_1 \mathscr{B} T_2$. It may not be minimal, but if it is then, by Proposition S item 3, $(x, l'_2) \in \mathscr{B}_2 T_2$ and $l'_2 = T_2 l_2$ will imply $(x, l_2) \in \mathscr{B}_2$.

Now assume that t' and t'' are sufficiently large such that $\mathscr{H}^{t'}_-(\tilde{\mathbf{w}})$ (resp., $\mathscr{H}^{t''}_+(\tilde{\mathbf{w}})$) contains the rows of H_- (resp. H_+). Then applying the above in the situation where T_1 and T_2 constitute the obvious selection of components (yielding H_- from $\mathscr{H}^{t'}(\tilde{\mathbf{w}})$ and H_+ from $\mathscr{H}^{t''}_+(\tilde{\mathbf{w}})$), it is clear that a minimal splitting for the columns of col(H_-, H_+) will induce a minimal splitting for the columns of col $(\mathscr{H}^{t'}_-(\tilde{\mathbf{w}}), \mathscr{H}^{t''}_+(\tilde{\mathbf{w}}))$ and hence by the preamble a minimal state space representation of $\mathscr{B}(R^*_\#)$.

The path to be followed in making this into an algorithm is now laid out. Consider $\mathscr{B}$ defined as the span of the columns of col (H_-, H_+), construct a minimal splitting $(\mathscr{X}, \mathscr{B}_1, \mathscr{B}_2)$ for it, and obtain $\mathbf{x}(t)$, $t \in \mathbb{Z}$, as the splitting variable for the tth column of col(H_-, H_+). In Step 2 of Algorithm 6 the procedure of Appendix S is applied to the case at hand. In this way, the state/external trajectory $(\mathbf{w}, \mathbf{x})$ is obtained and matrix representations of this trajectory must be determined. This is in fact exactly what is done in Steps 3 and 4. Further details are left to the reader. □

Proof of Corollary 22

Corollary 22 is an application of Algorithm 5. Schematically,

$H_{t'}(\tilde{\mathbf{w}}) =$

$$\begin{bmatrix}
\cdots & 0 & \cdots & 0 & \mathrm{I} & 0 & \cdots & 0 & \cdots \\
\cdots & 0 & \cdots & 0 & \tilde{\mathbf{G}}(0) & \tilde{\mathbf{G}}(1) & \cdots & \tilde{\mathbf{G}}(t'') & \cdots \\
\cdots & 0 & \cdots & I & 0 & 0 & \cdots & 0 & \cdots \\
\cdots & 0 & \cdots & \tilde{\mathbf{G}}(0) & \tilde{\mathbf{G}}(1) & \tilde{\mathbf{G}}(2) & \cdots & \tilde{\mathbf{G}}(t''+1) & \cdots \\
\cdots & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & \cdots \\
\cdots & 0 & \cdots & \tilde{\mathbf{G}}(1) & \tilde{\mathbf{G}}(2) & \tilde{\mathbf{G}}(2) & \cdots & \tilde{\mathbf{G}}(t''+2) & \cdots \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
\cdots & I & \cdots & 0 & 0 & 0 & \cdots & 0 & \cdots \\
\cdots & \tilde{\mathbf{G}}(0) & \cdots & \tilde{\mathbf{G}}(t'-2) & \tilde{\mathbf{G}}(t'-1) & \tilde{\mathbf{G}}(t') & \cdots & \tilde{\mathbf{G}}(t'+t'') & \cdots
\end{bmatrix}$$

spans $\mathscr{B}^1_{t'}$ — spans U — spans X

where t' is chosen such that $\hat{H}$ is contained in $\mathscr{H}_{t'}(\tilde{\mathbf{w}})$ as a submatrix.

Choose H in Algorithm 5 to be the columns of $\mathscr{H}_{t'}(\tilde{\mathbf{w}})$ making up the m columns marked 'spans U' in the above scheme and the columns which also appear in $\hat{H}$. Now define $Q_u = \text{col}\, (I_m, 0)$ $Q_x = \text{col}\, (0, \hat{Q})$, and P as follows. For the columns of P corresponding to the rows of $\mathscr{H}_{t'}(\tilde{\mathbf{w}})$ which also appear in $\hat{H}$, choose the corresponding column of $\hat{P}$. In the columns of P corresponding to the other rows of $\mathscr{H}_{t'}(\tilde{\mathbf{w}})$ containing the matrices $\mathbf{G}(t)$, put zeros. Finally choose the columns of P corresponding to the rows of $\mathscr{H}_{t'}(\tilde{\mathbf{w}})$ containing the Is, such that $PHQ_u = 0$ and $P\mathscr{B}^1_{t'} = 0$. It is obvious that this is possible, and, in fact, since the numerical values of these entries are not needed, it suffices to observe that it can be done. Note that Equations 4.1–4.3 of Algorithm 5 will be satisfied by this P. Now it is easy to verify the following equalities:

$$P\sigma HQ_x = \hat{P}\sigma \hat{H}\hat{Q}$$

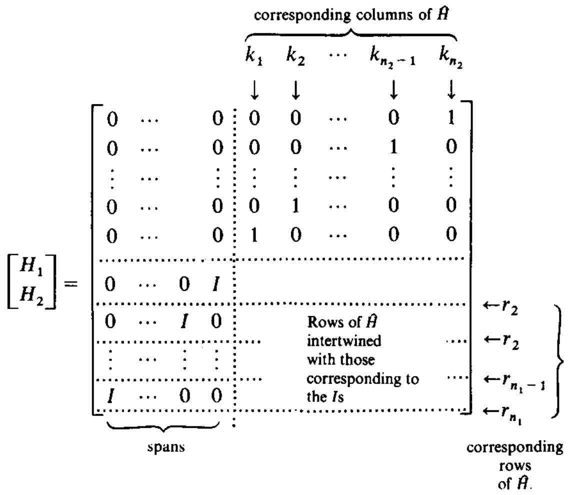
$$P\sigma HQ_u = \hat{P}\hat{H}_0$$

$$H^0 Q_x = \begin{bmatrix} 0 \\ \hat{H}_0 \hat{Q} \end{bmatrix}$$

$$H^0 Q_u = \begin{bmatrix} I \\ \mathrm{G}(0) \end{bmatrix}.$$

This yields Corollary 22.

In order to illustrate the application of Algorithm 6, how Algorithm 6 yields Corollary 21 is demonstrated. It is easily verified that the choices of H_1 and H_2 schematically shown below satisfy the relative rank and other requirements of Algorithm 6;



The state at time t can now identified with the tth column of $H_+(\text{mod}\,\mathscr{X})$ with $\mathscr{X}$ as identified in the above diagram. Now examine $\mathbf{f}(t) = \text{col}(\mathbf{w}(t), \mathbf{x}(t))$. Denote the columns of $H_+(\text{mod}\,\mathscr{X})$ corresponding to those of $\hat{H}$ and $\sigma\hat{H}$ by $\hat{H}(\text{mod}\,\mathscr{X})$ and $\sigma\hat{H}(\text{mod}\,\mathscr{X})$, respectively. The following matrix has as its image span $\{\mathbf{f}(t), t \in \mathbb{Z}\}$

$$S = \begin{bmatrix} I_m & 0 \\ \mathbf{G}(0) & \hat{H}^0 \\ 0 & \hat{H}(\text{mod}\,\mathscr{X}) \end{bmatrix}.$$

Clearly the first m entries in the columns of S can be identified with the input. Then $D: I_m \to \begin{bmatrix} I_m \\ \tilde{\mathbf{G}}(0) \end{bmatrix}$, i.e. $D = \begin{bmatrix} I_m \\ \tilde{\mathbf{G}}(0) \end{bmatrix}$. In order to compute matrices (A, B, C) a suitable coordinate representation of the columns of $\hat{H}\,(\text{mod}\,\mathscr{X})$ should be chosen with:

A as the matrix representing the map which takes the columns of $\hat{H}(\text{mod}\,\mathscr{X})$ into those of $\sigma\hat{H}(\text{mod}\,\mathscr{X})$;

B as the matrix representing the columns of $\hat{H}_0\,(\text{mod}\,\mathscr{X})$; and

C as the matrix representation of the map $\hat{H}$ which takes the columns of $\hat{H}(\text{mod}\,\mathscr{X})$ into those of $\begin{bmatrix} 0 \\ \hat{H}^0 \end{bmatrix}$

Now choose the state vector x corresponding to the column h of $\hat{H}$ as $x = \hat{P}h$. Then $A: \hat{P}\hat{H} \to \hat{P}\sigma\hat{H}$, $B = \hat{P}\hat{H}^0$, and $C: \hat{P}\hat{H} \to \begin{bmatrix} 0 \\ \hat{H}^0 \end{bmatrix}$. Consequently $A = \hat{P}\sigma\hat{H}\hat{Q}$, $B = \hat{P}\hat{H}^0$, $C = \begin{bmatrix} 0 \\ \hat{H}^0\hat{Q} \end{bmatrix}$ and $D = \begin{bmatrix} I_m \\ \tilde{\mathbf{G}}(0) \end{bmatrix}$ is a minimal realization, which yields Corollary 22.

□