Ch. 2

Techniques for the Selection of Identifiable Parametrizations for Multivariable Linear Systems

MICHEL GEVERS
VINCENT WERTZ

Laboratoire d'Automatique,
de Dynamique et d'Analyse des Systèmes
Louvain University
B-1348 Louvain la Neuve, Belgium

I. INTRODUCTION

This article is a survey. The objective is to give an up-to-date account of the use of identifiable representations for linear multivariable systems as well as to briefly survey methods for the estimation of the structure of these representations from measured data. Linear multivariable systems can be represented in a number of ways using finitely parametrized models; in this article we will consider state-space models (SSs), matrix fraction descriptions (MFDs), and autoregressive moving average models with exogeneous inputs (ARMAXs), also called vector difference equation models (VDEs). This last class of models is of particular importance in the context of system identification from observed data because most available parameter estimation methods are well suited to ARMAX models. Transformations between the various classes of finite-dimensional

models are easy to establish, and within each class there exists an infinite number of equivalent models, all producing the same sequence of Markov parameters (or impulse response matrices $\{K_0, K_1, K_2, \ldots\}$ or, equivalently, the same matrix transfer function $K(z) \triangleq \Sigma_{i=0}^{\infty} K_i z^{-i}$.

A system is uniquely defined by its sequence of Markov parameters. If the system is finite dimensional (see Section II), it can be represented in an infinite number of ways by a finitely parametrized model (SS, MFD, or ARMAX). Obviously, in an identification context, one would like to use a parametrized model set $M^* = \{M(\theta) | \theta \in D\}$ (either an SS, MFD, or ARMAX model set) that is able to represent the system for a unique value of the parameter vector θ, so that the parameter estimation algorithm would converge. This is the problem of constructing identifiable model sets to which this article is devoted. The structure of these model sets is determined by a set of integer-valued indices, called structure indices. They determine the locations of 0 and 1 elements and of free parameters in the matrices of an SS model set or the degrees of the polynomials in an MFD or ARMAX model set.

A fundamental property of linear multivariable systems is that no unique model set is able to represent all systems of given order, say n. The set S(n) of all systems of order n can only be represented as a finite union of model sets, each characterized by its structure indices. This is what makes the determination of identifiable model sets nontrivial in the multivariable case. Now there are basically two ways to let each system in S(n) be represented by an identifiable model. One way is to let S(n) be described by a disjoint union of identifiable model sets: that is, each system $\sigma \in S(n)$ is

represented by a unique model specified by a unique set of integer-valued structure indices and a unique set of real-valued parameters. These identifiable model sets are then called canonical. Given a particular choice of canonical form (e.g., observer canonical form), there exists one such form for each system. The alternative is to let $S(n)$ be described by a union of identifiable but overlapping model sets, each set being characterized by its set of integer-valued structure indices and each model within that set by a unique set of real-valued parameters. In this approach the structure indices are not defined by the system. As a matter of fact, each of the model sets is dense in $S(n)$ (see [1]-[4] for a description of the structure of $S(n)$); this means that any such identifiable model set M^* will be able to represent almost any given system $\sigma \in S(n)$ for a particular value of θ. If $M_1(\theta_1)$ and $M_2(\theta_2)$ are the representations of a same system σ in two different identifiable model sets, then the parameter vectors θ_1 and θ_2 are related by a transformation that corresponds to a coordinate transformation in Euclidean space.

Canonical forms for multivariable systems were first introduced as a tool to simplify some observer or controller design problems (see, e.g., [5]). Because a canonical form is also a uniquely defined representative of an equivalence class of (SS, MFD, or ARMAX) models, its importance in identification was soon recognized (see [1] and [6]-[12]). When using canonical forms for system identification, the most critical part of the problem is the estimation of the structure indices (or Kronecker indices) of the system. If they have been wrongly estimated, then the parameter vector θ cannot converge to the true θ_0 [13].

The estimation of the structure indices becomes very critical if the system happens to lie close to the boundary between two of the disjoint subsets of S(n) mentioned earlier.

Around 1974, the structure of S(n) came to be much better understood thanks to the important work of Glover and Willems [8], Hazewinkel and Kalman [14], and Clark [2], who showed that S(n) has the structure of an analytic manifold that can be covered by a union of overlapping subsets and that an identifiable parametrization can be defined for each subset. These parametrizations were called overlapping forms or pseudocanonical forms and their use for identification of multivariable linear systems, as an alternative to canonical forms, was studied by a number of authors: overlapping state-space models were examined in [15] through [18], ARMAX models in [19] through [20], while both state-space and ARMAX models, and the relationships between them, were studied in [21] through [26]. The advantage of overlapping forms over canonical forms is that, instead of having to estimate p structure indices $n_1, \ldots, n_p$ (p being the number of observed outputs of the system), one has to estimate only the order of the system. Given the order n, one can choose any set of structure indices $n_1, \ldots, n_p$ adding up to n, and, almost surely, one can identify the system in the corresponding pseudocanonical (or overlapping) form. Numerical considerations may lead one to choose a particular set of indices (see [16]-[17]). The disadvantage is that this form may contain a few parameters more than the canonical form.

In any case, whether canonical or pseudocanonical forms are used, one has to estimate one or more integer-valued indices from the data: the order n or the structure indices $n_1, \ldots, n_p$. This is called structure estimation; it is the most critical

step in the identification of multivariable systems. (In econometrics, structure estimation is actually called identification.) Recent work of Hannan and Kavalieris [26]-[28] has shown that, under mild conditions, the order n (for pseudocanonical forms) and the Kronecker indices $n_1, \ldots, n_p$ (for canonical forms) can be consistently estimated. See also [25] for an excellent presentation of these recent results and for a discussion of the consequences of misspecifying the structure.

Here we will survey most of the results mentioned. Because almost all results are available in the literature, we shall give them without proof, but refer to the appropriate references. In Section II we shall present the different model sets discussed here and establish their interconnection. The concept of identifiability will be presented in Section III, while the structure of S(n) will be analyzed in Section IV. It will be shown that a rational system can be represented by a point (i.e., a coordinate vector) in an appropriate coordinate system. In Section V we will present a class of canonical and pseudocanonical model structures, in SS, MFD, and ARMAX form; these structures are all derived from the coordinates of the system, defined in a more abstract way in Section IV. In Section VI we will briefly present some other identifiable model structures; they are not directly derived from the coordinate vector defined in Section IV and tend to have a larger number of parameters. Finally, in Section VII we will survey the most recent results on structure estimation; this means estimation of the Kronecker indices when canonical forms are used, of the order of the system when pseudocanonical forms are used.

II. MODELS

As a starting point, we consider a p-vector stationary stochastic process $y(t)$ generated as follows:

$$y(t) = G(z)u(t) + H(z)e(t), \tag{1}$$

where $y(t) \in R^p$, $u(t) \in R^m$, $e(t) \in R^p$, and $G(z)$ and $H(z)$ are causal rational transfer function matrices. The following assumptions are made about the system:

$$\text{(i)} \quad E\{e(t)\} = 0, \quad E\{e(t)e^T(s)\} = \Sigma\delta_{ts}, \quad \Sigma > 0, \tag{2a}$$

$$\text{(ii)} \quad G(z) = G_1 z^{-1} + G_2 z^{-2} + \cdots, \tag{2b}$$

$$H(z) = I + H_1 z^{-1} + H_2 z^{-2} + \cdots, \tag{2c}$$

$$\text{(iii)} \quad G(z) \text{ is analytic in } |z| \geq 1, \tag{2d}$$

$$\text{(iv)} \quad H(z) \text{ has full rank so that } H^{-1}(z) \text{ exists and } H(z) \text{ and } H^{-1}(z) \text{ are analytic in } |z| \geq 1, \text{ and} \tag{2e}$$

(v) $u(t)$ is an observed input signal, which can be either deterministic or stochastic, but we assume that $|u(t)|$ is bounded and that the following limits exist:

$$\lim_{N\to\infty} \frac{1}{N} \sum_{1}^{N} E[u(t)] = 0,$$

$$\lim_{N\to\infty} \frac{1}{N} \sum_{1}^{N} E[u(t)u^T(t-\tau)] = R_u(\tau), \tag{2f}$$

where the expectation is discarded if $u(t)$ is deterministic. We then define the spectrum of $u(\cdot)$ as

$$\phi_u(\omega) \triangleq \sum_{\tau=-\infty}^{\infty} R_u(\tau)e^{i\omega\tau}. \tag{3}$$

The model (1)-(2) is very general and can be justified as follows. If

$$\tag{4}$$

where $s(t)$ is some useful signal and $v(t)$ is noise, and if the spectrum $\phi_v(\omega)$ of $v(t)$ is rational, Hermitian, and positive definite for all ω in $[-\pi, \pi]$, then $\phi_v(\omega)$ can be decomposed uniquely as

$$\phi_v(\omega) = H(e^{i\omega})\Sigma H^T(e^{-i\omega}), \tag{5}$$

where $H(z)$ and Σ satisfy the conditions (2a, c, e).

With these conditions, the $e(t)$ are the linear innovations, that is, the prediction errors of the best linear one-step-ahead prediction of $v(t)$ from its infinite past. The model (1) will be called a transfer function (TF) model.

Comment 1. The condition (2d) has been introduced to make $y(t)$ a stationary process, so that covariances and spectra can be defined. It is necessary for most of the consistency results on structure estimation, which will be presented in Section VII. However, it is not required for the proper definition of canonical and pseudocanonical forms; obviously one should be able to identify unstable dynamical systems. The important feature for this is that the predictor $\hat{y}(t|t-1)$ be stable, rather than the data-generating model; see Comment 3 later.

With the model (1)-(2) we shall associate a Hankel matrix $H_{1,\infty}[G \vdots H]$ defined as follows. Let $\{K_1, K_2, \ldots\}$ be a sequence of matrices; then

$$H_{1,N}[K] \triangleq \begin{bmatrix} K_1 & K_2 & \cdots & K_N \\ K_2 & K_3 & \cdots & K_{N+1} \\ \vdots & & & \\ K_N & K_{N+1} & \cdots & K_{2N-1} \end{bmatrix} \tag{6}$$

In particular $H_{1,\infty}[G \vdots H]$ is obtained by replacing K_i in (6) by $[G_i \vdots H_i]$ and by letting N go to ∞. The rank of $H_{1,\infty}[G \vdots H]$ is then called the order of the system (1).

Given the spectra $\phi_y(\omega)$, $\phi_u(\omega)$, and $\phi_{yu}(\omega)$ (all of which can theoretically be estimated from arbitrarily long input and output records), in principle one can determine $G(z)$ and $H(z)$ uniquely under weak conditions on $\phi_u(\omega)$ from

$$\phi_{yu}(\omega) = G(e^{i\omega})\phi_u(\omega), \tag{7}$$

$$\phi_y(\omega) = G(e^{i\omega})\phi_u(\omega)G^T(e^{-i\omega}) + H(e^{i\omega})\Sigma H^T(e^{-i\omega}). \tag{8}$$

However, to actually construct an estimate of $G(z)$ and $H(z)$ from second-order statistics, one needs to represent the process $y(t)$ in a finitely parametrized form and then construct an algorithm for the estimation of these parameters. In other words, one needs to define a coordinate space and to let the system be represented by a point in that coordinate space. This is the problem of parametrization, that is, of defining finite-dimensional identifiable model sets for the process $y(t)$. From now on we shall consider (1)-(2) as the given system, and we shall study three classes of finitely parametrized model sets that are input-output equivalent with the system (1)-(2). Notice that (1)-(2) is an infinite-dimensional representation; that is, the system is specified by the infinite sequence of matrices $[G_1, H_1, G_2, H_2, \ldots]$.

A. *STATE-SPACE MODELS*

One way of representing the system (1)-(2) is through a state-space model:

$$\begin{cases} x(t+1) = Ax(t) + Bu(t) + Ke(t), & (9a) \\ y(t) = Cx(t) + e(t), & (9b) \end{cases}$$

where $y(t)$, $u(t)$, and $e(t)$ are the same as in (1), and where $\dim x(t) \triangleq n$ is minimal, with A, B, C, K such that

$$[I - G(z)]^{-1}[G(z) \quad H(z)] = C(zI - A)^{-1}[B \quad K] + [0 \quad I] \tag{10}$$

By conditions (2d, e) this implies that all the eigenvalues of A have modulus less than 1. It is a standard result of linear system theory that

$$n \triangleq \dim x(t) = \operatorname{rank} H_{1,\infty}[G \vdots H], \tag{11}$$

where n is called the order of the system. Now there exists an infinity of SS models {A, B, C, K} that are input-output equivalent with the system (1)-(2), that is, for which the relation (10) holds. They are all obtained from an arbitrary model by the following similarity transformations

$$A^* = T^{-1}AT, \quad B^* = T^{-1}B, \quad K^* = T^{-1}K, \quad C^* = CT, \tag{12}$$

where T is any nonsingular $n \times n$ matrix. Our task is to define identifiable model sets, that is, to parametrize $A(\theta)$, $B(\theta)$, $C(\theta)$, $K(\theta)$ in such a way that any arbitrary system $\sigma \in S(n)$ (defined by the sequences G_i, H_i; $i = 1, 2, \ldots$) can be represented in this parametrized model set by a unique value of the parameter vector θ.

B. MATRIX FRACTION DESCRIPTIONS

Another representation of the system (1)-(2) is

$$P(z)y(t) = Q(z)u(t) + R(z)e(t), \tag{13}$$

where P(z), Q(z), and R(z) are left-coprime polynomial matrices (see, e.g., [29]) satisfying the relationship

$$P^{-1}(z)[Q(z) \vdots R(z)] = [G(z) \vdots H(z)]. \tag{14}$$

In (13) z is the forward shift operator: $zy(t) \triangleq y(t + 1)$. If n is the order of the system (see above), then

$$\deg \det P(z) = n. \tag{15}$$

Also, by (2d, e), det P(z) and det R(z) have all their roots in $|z| < 1$.

Again the MFD (13) is nonunique. If $\{P(z), Q(z), R(z)\}$ is an arbitrary left-coprime MFD for (1)-(2), then all input-output-equivalent left-coprime MFD models are obtained from this arbitrary model by the following unimodular transformations:

$$[P^*(z) \vdots Q^*(z) \vdots R^*(z)] = U(z)[P(z) \vdots Q(z) \vdots R(z)], \tag{16}$$

where $U(z)$ ranges over the set of all unimodular matrices of dimension $p \times p$. (A unimodular matrix is a square polynomial matrix whose determinant is a nonzero constant.) Again, to obtain identifiable MFD model sets, our task will be to parametrize $P(z, \theta)$, $Q(z, \theta)$, $R(z, \theta)$ (i.e., to define their structure) in such a way that (14) holds for a unique value of θ.

C. *VECTOR DIFFERENCE EQUATION OR ARMAX MODELS*

One of the most widely used model sets in econometrics, but also in engineering, is the ARMAX[1] model set

$$\bar{P}(D)y(t) = \bar{Q}(D)u(t) + \bar{R}(D)e(t), \tag{17}$$

where $\bar{P}(D)$, $\bar{Q}(D)$, and $\bar{R}(D)$ are left-coprime polynomial matrices with

$$\bar{P}(D) = \bar{P}_0 + \bar{P}_1 D + \cdots + \bar{P}_p D^p, \tag{18a}$$

$$\bar{Q}(D) = \bar{Q}_1 D + \cdots + \bar{Q}_q D^q, \tag{18b}$$

$$\bar{R}(D) = \bar{R}_0 + \bar{R}_1 D + \cdots + \bar{R}_r D^r, \tag{18c}$$

where D is the delay operator: $Dy(t) \triangleq y(t-1)$.

In order for (17)-(18) to be a representation of (1)-(2), the following conditions must hold:

$$\bar{P}^{-1}(D)[\bar{Q}(D) \vdots \bar{R}(D)] = [\bar{G}(D) \vdots \bar{H}(D)], \tag{19}$$

[1] *When there is no exogenous input u(t), Eq. (17) becomes*

where

$$\bar{G}(D) \triangleq G(D^{-1}) = \sum_{1}^{\infty} G_i D^i, \tag{20}$$

and similarly for $\bar{H}(D)$. In particular (19) implies that

$$\text{(i)} \quad \bar{P}_0 \text{ is nonsingular and } \bar{P}_0 = \bar{R}_0, \tag{21}$$

$$\text{(ii)} \quad \det \bar{P}(z) \text{ and } \det \bar{R}(z) \text{ have all their roots in } |z| > 1. \tag{22}$$

In addition to (21), we will often want ARMAX models for which

$$\bar{P}_0 = \bar{R}_0 = I_p. \tag{23}$$

This allows one to write

$$y(t) = -\sum_{1}^{p} \bar{P}_i y(t - i) + \sum_{1}^{q} \bar{Q}_i u(t - i) + \sum_{1}^{r} \bar{R}_i e(t - i) + e(t). \tag{24}$$

We shall see later that condition (23) introduces some additional complications.

Comment 2. We could have used z^{-1} for the delay operator in lieu of D. We have not done so because we want to keep powers of z^{-1} for power series and to stress that $\bar{P}(D)$, $\bar{Q}(D)$, and $\bar{R}(D)$ are polynomials.

There is an obvious relationship between MFD models and ARMAX models. Let (14) be a left-coprime MFD of $[G(z) \vdots H(z)]$ and let $n_1, \ldots, n_p$ be the row degrees of the matrix $[P(z) \vdots Q(z) \vdots R(z)]$.

Define $M(z) = \text{diag}\{z^{n_1}, \ldots, z^{n_p}\}$, and let $z^k D^k = D^k z^k = 1$ so that $M(z)M(D) = M(D)M(z) = I_p$. Then

$$M(D)[P(z) \vdots Q(z) \vdots R(z)] = [\bar{P}(D) \vdots \bar{Q}(D) \vdots \bar{R}(D)]. \tag{25a}$$

Similarly

$$M(z)[\bar{P}(D) \vdots \bar{Q}(D) \vdots \bar{R}(D)] = [P(z) \vdots Q(z) \vdots R(z)]. \tag{25b}$$

Note that the row degrees of $[P(z) \vdots Q(z) \vdots R(z)]$ are identical to the row degrees of $[\bar{P}(D) \vdots \bar{Q}(D) \vdots \bar{R}(D)]$; however, the row degrees of $\bar{P}(D)$ are not equal to those of $P(z)$, and deg det $\bar{P}(D)$ is not equal to the order of the system (see [30] and [31] for details).

Comment 3. An alternative to using the "data-generating model" (1)-(2) as our starting point is to use a "predictor model," that is, the model that generates the one-step-ahead predictions of $y(t)$ given the infinite past. It can be written

$$\hat{y}(t|t-1) = W_u(z)u(t) + W_y(z)y(t), \tag{26}$$

where $W_u(z)$ and $W_y(z)$ are stable rational transfer function matrices

$$W_u(z) = \sum_1^\infty W_u(k)z^{-k}, \quad W_y(z) = \sum_1^\infty W_y(k)z^{-k}. \tag{27}$$

The model (26) is easily derived from the data-generating model (1)-(2):

$$W_u(z) = H^{-1}(z)G(z), \quad W_y(z) = I - H^{-1}(z). \tag{28}$$

The use of the prediction model (26) as a starting point for the analysis of identifiable model sets is justified by the fact that most identification methods are based on minimizing prediction errors. This viewpoint is taken in [32]. All our subsequent analysis for the representation of (1) by finitely parametrized identifiable model sets applies equally well to the representation (26), which can also be modeled by SS, MFD, or ARMAX models. In SS form, (26) leads to the Kalman filter. Notice finally that the condition "$\{W_u(z), W_y(z)\}$ stable" is

really what is required to apply prediction error methods; it enables one to identify some models of the form (1) with unstable $G(z)$. For example, an MFD model (13) with $\det P(z)$ having roots in $|z| \geq 1$ and $\det R(z)$ having all its roots in $|z| < 1$ leads to an unstable $G(z)$, but yields a stable $W_u(z)$ and $W_y(z)$.

III. IDENTIFIABILITY

There are several ways of defining identifiability. Here we propose a setup and a definition inspired by, but not identical to, [32]. As a starting point we assume that we have an input process $u(t)$ and an output process $y(t)$ that can be described by (1):

$$y(t) = G(z)u(t) + v(t) = G(z)u(t) + H(z)e(t) \tag{29}$$

with the properties (2). This is our basic model. By the constraints (2) $H(z)$ is a unique factorization of $\phi_v(\omega)$. Therefore $G(z)$, $H(z)$, and Σ are uniquely defined by the spectra ϕ_y, ϕ_{uy}, and ϕ_u (see Section II), and therefore (29) can actually be seen as an input/output (I/O) model from the input process

$$\begin{bmatrix} u \\ e \end{bmatrix}$$

to the output process y with transfer function matrix $[G(z) \vdots H(z)]$. Note that we could also have chosen (26) as our basic model, which would similarly be uniquely defined by the spectra.

Since we also want to consider other classes of models than the TF model, such as SS, MFD, or ARMAX models, we introduce the following definition.

Definition 1. A model M is a stable algebraic operator that transforms a given input process

$$\begin{bmatrix} u(t) \\ e(t) \end{bmatrix}$$

into a unique output process y(t).

Therefore (1), (9), (13), and (17) are all models by our definition, provided that specific values (rational functions or real-valued parameters or polynomials) are put into the operators G(z), H(z), A, B, C, K, ... Whatever the description chosen for our model (SS, MFD, ARMAX), we can always compute the corresponding transfer functions G(z) and H(z) from it [see, e.g., (10)]. We can therefore introduce the following definition.

Definition 2. Let $M^{(1)}$ and $M^{(2)}$ be two models relating an input vector

$$\begin{bmatrix} u \\ e \end{bmatrix}$$

to an output vector y, and let $\{G^{(1)}(z), H^{(1)}(z)\}$ and $\{G^{(2)}(z), H^{(2)}(z)\}$ be the corresponding transfer functions. Then the two models are equivalent [we write $M^{(1)} = M^{(2)}$] if and only if $G^{(1)}(z) = G^{(2)}(z)$ and $H^{(1)}(z) = H^{(2)}(z)$.

In identification a search will typically be conducted over a set of models. Most often this model set is noncountable; it is obtained as the range of a smoothly parametrized algebraic operator where the parameter vector θ is allowed to cover a subset D of R^d ($d = \dim \theta$). The search for the "best" model is then performed over all $\theta \in D$. We formalize this as follows.

Definition 3. A model set M^* is a set of models: $M^* = \{M_\alpha | \alpha \in I_\alpha\}$, where I_α is an index set.

Most often the index set is noncountable. The model set is then smoothly parametrized by a parameter vector θ of dimension d, where $\theta \in D_M \subset R^d$.

Definition 4. A model structure $\tilde{M}$ is a differentiable mapping from a subset D_M of R^d to a model set: $\tilde{M} : \theta \in D_M \to M(\theta) \in M^*$, where $M(\theta)$ is defined by any one of the stable parametrized algebraic operators described earlier.

Example 1. We illustrate by a scalar example.

(1) $M : y(t) = 0.85\, y(t-1) + 1.5\, u(t-1)$.

(2) $M^* = \left\{ y(t) = \alpha y(t-1) + \beta u(t-1) \,|\, (\alpha, \beta) \in D_M \subset R^2 \right\}$, where $D_M = \{\alpha, \beta \,|\, |\alpha| < 1, |\beta| \leq 100\}$, for example.

(3) $\tilde{M} : (\alpha, \beta) \in D_M \subset R^2 \to M^* = \{y(t) = \alpha y(t-1) + \beta u(t-1) \,|\, (\alpha, \beta) \in D_M\}$.

The range of a model structure defines a model set: Range $\tilde{M} = \{M(\theta) \,|\, \theta \in D_M\} = M^*$. Note that a given model set can typically be described as the range of different model structures. An important problem, which will be discussed at length in this paper, is to find a model structure (i.e., a parametrization) whose range equals a given model set. It turns out that the set S(n) of all linear systems of order n with dim $y_t = p > 1$ and dim $u_T = m > 1$ cannot be described as the range of a single model structure. The remedy will be to describe S(n) as a union of ranges of different model structures:

$$M^* = \bigcup_{i=1}^{l} R\left(\tilde{M}_i\right). \tag{30}$$

We can now define identifiability. The concept of identifiability involves several aspects and has therefore given rise to several definitions. In its broad sense, identifiability is concerned with whether the identification procedure yields

a unique value of θ, and possibly whether the resulting model $M(\theta)$ is equivalent to the true system σ. This involves aspects of whether $\sigma \in M^*$, whether the data set is informative enough, and whether the "model structure is identifiable." This last problem, which we shall call structural identifiability, concerns only the invertibility of the map of Definition 4.

Definition 5. A model structure $\tilde{M}$ is called *globally identifiable at* θ^* if

$$M(\theta) = M(\theta^*), \quad \theta \in D_m \Rightarrow \theta = \theta^*. \tag{31}$$

It is called *strictly globally identifiable* if it is globally identifiable $\forall \theta^* \in D_M$.

This last condition is almost never met, as we shall see in a moment. Therefore we introduce a more realistic property.

Definition 6. A model structure $\tilde{M}$ is *globally identifiable* if it is globally identifiable at almost all $\theta^* \in D_M$.

Similar definitions can be introduced for local identifiability, that is, where (31) holds for θ in a neighborhood of θ^*. We stress that these definitions cover only one aspect of identifiability; in particular they are a property of the model structure only and are totally independent of a possible "true system." To stress the relevance of these definitions to our problem of constructing identifiable model structures, we point out the following facts:

(1) The set of all single-input/single-output models of order up to n cannot be described as the range of a strictly globally identifiable model structure without restrictions on θ. This is because of pole-zero cancellations on certain hypersurfaces. It can, however, be described as the range of a globally identifiable model structure.

(2) The set of all multivariable systems of order up to n (or even exactly n) cannot be described as the range of a unique globally identifiable model structure. We shall see in the next section that it can be described as a union of ranges of globally identifiable model structures. The concept of structural identifiability is very old in econometrics. It was first introducted in the engineering literature in [33]. It is important for numerical reasons. It guarantees that if the data are informative enough, the parameter estimation problem will be well posed, since different θ yield different I/O properties. The definition we have adopted is by no means unique. Instead of using the uniqueness of $G(z, \theta)$ and $H(z, \theta)$, some definitions use the uniqueness of the parametrized joint probability function of the data $p(Y, U; \theta)$ as a starting point (see, e.g., [18], [23], [35]). Other definitions are based on the Kullback-Leibler information or on the theory of estimable functions (see [35,36]). Notice finally that the term "structural identifiability" has a different meaning in econometrics [25].

IV. THE STRUCTURE OF S(n)

Consider $K(z) \triangleq [G(z) \vdots H(z) - I] = \Sigma_1^\infty K_i z^{-i}$ of dimension $p \times s$, with $s \triangleq m + p$. Then $K(z)$ is strictly proper, that is,

$$\lim_{z \to \infty} K(z) = 0. \tag{32}$$

Definition 1. We call $S(n)$ the set of all strictly proper stable rational transfer function matrices $K(z)$ of order n (i.e., such that rank $H_{1,\infty}[K] = n$).

Notice that the problem of representing $[G(z) \vdots H(z)]$ by an identifiable model is a trivial modification of the problem of representing $[G(z) \vdots H(z) - I]$, and this last problem is that of

representing an arbitrary strictly proper transfer function $K(z)$. This is why we now study the structure of $S(n)$. If we further denote

$$m(t) \triangleq \begin{bmatrix} u(t) \\ e(t) \end{bmatrix},$$

then the basic model (1) can be written

$$y(t) = K(z)m(t) + e(t), \tag{33}$$

with $m(t) \in R^s$, and the problem becomes one of finding an identifiable representation of $K(z)$. The cases where $u(t) = 0$ (ARMA model) or $e(t) = 0$ (I/O model) are special cases of our setup.

The topology on $S(n)$ is the relative pointwise topology T_{pt} in which a sequence $K^{(n)}(z)$ converges to $K(z)$ if and only if the coefficient matrices $K_i^{(n)}$ converge to K_i (in the relative Euclidean topology) for all $i = 1, 2, \ldots$ (see [3]). If S is a set in this topological space, its closure will be denoted by $\bar{S}$. In particular $\bar{S}(n)$ denotes the set of all strictly proper rational transfer functions of order less than or equal to n.

Now let $K(z)$ be a point in $S(n)$, with $K(z)$ of dimensions $p \times s$, and consider the Hankel matrix $H_{1,\infty}[K]$. This Hankel matrix is made up of blocks of p rows. Call r_{ij} the ith row of the jth block of $H_{1,\infty}[K]$. We shall also make the following standing technical assumption.

Technical Assumption. The rows of the first block of $H_{1,\infty}[K]$ are linearly independent.

This really only eliminates degenerate transfer functions $K(z)$ whose rows would be linearly dependent. We now state without proof a few important results on $S(n)$.

Result 1. If $K(z) \in S(n)$, then there exists at least one partition $n = n_1 + n_2 + \cdots + n_p$ of n such that the set of rows

$$\{r_{ij} : i = 1, \ldots, p;\ j = 1, \ldots, n_i\} \triangleq R(n_1, \ldots, n_p) \tag{34}$$

form a basis of $H_{1,\infty}[K]$.

This follows immediately from the Hankel structure. Note that the rows in (34) are chosen in such a way that if $r_{ij} \in R(n_1, \ldots, n_p)$ for $j > 1$, then $r_{i,j-1} \in R(n_1, \ldots, n_p)$. The set of the row indices corresponding to the rows of $R(n_1, \ldots, n_p)$ is completely determined by the partition $\mu \triangleq (n_1, \ldots, n_p)$. The indices $n_1, \ldots, n_p$ are called structure indices. We also denote $|\mu| = \sum_1^p n_i = n$. It is easy to see that there are

$$\begin{bmatrix} n + p - 1 \\ p - 1 \end{bmatrix}$$

such partitions.

Definition 2. We call U_μ the set of all points in $S(n)$ for which the corresponding set of rows (34) specified by μ forms a basis for the rows of $H_{1,\infty}[K]$.

We now show that U_μ can be completely coordinatized by $n(p + s)$ coordinates. Row r_{i,n_i+1} is a unique linear combination of the basis rows

$$r_{i,n_i+1} = \sum_{j=1}^{p} \sum_{l=1}^{n_j} \alpha_{ijl} r_{jl}, \quad i = 1, \ldots, p. \tag{35}$$

It follows again from the Hankel structure that knowing the first block of elements (i.e., the first s elements) of the basis rows (34) and the coefficients $\{\alpha_{ijl};\ i,j = 1, \ldots, p;\ l = 1, \ldots, n_j\}$ allows one to compute any other row of $H_{1,\infty}[K]$

and therefore to specify $K(z)$ completely. Therefore any element $K(z) \in U_\mu$ can be mapped into a vector τ_μ in R^d by the following $d \triangleq n(p + s)$ coordinates:

$$\tau_\mu = \left\{ \begin{matrix} \alpha_{ijl}; & i, j = 1, \ldots, p; \quad l = 1, \ldots, n_j \\ k_l(i, j); & i = 1, \ldots, p; \quad j = 1, \ldots, s; \quad l = 1, \ldots, n_i \end{matrix} \right\} \tag{36}$$

where $k_l(i, j)$ is the (i, j)th element of the matrix K_l. We shall call ϕ_μ the mapping from $K(z)$ into τ_μ:

$$\Phi_\mu: U_\mu \to \Theta_\mu \subset R^d: K(z) \in U_\mu \to \tau_\mu = \Phi_\mu(K(z)) \in \Theta_\mu. \tag{37}$$

We now have the following important result, originally proved by Clark [2], and further extended by several authors (see, e.g., [14] and [19]).

Result 2.

(1) $S(n)$ is a real analytic manifold of dimension $n(p + s)$.

(2) $S(n)$ is the union of the U_μ such that $|\mu| = n$. Each U_μ is open and dense in $S(n)$; Θ_μ is open and dense in R^d.

(3) Φ_μ described in (37) is a homeomorphism between U_μ and an open and dense subset Θ_μ of R^d, with $d = n(p + s)$: $\Theta_\mu = \Phi_\mu(U_\mu)$.

(4) $\bar{S}(n) \triangleq \bigcup_{i \leq n} S(i) = \bar{U}_\mu$ if $|\mu| = n$.

Comment 1. Since U_μ, $|\mu| = n$, is open and dense in $S(n)$, it follows that almost all points of $S(n)$ are in U_μ for any such μ. The choice of a partition μ specifies a local coordinate system. Therefore, once n is chosen, a system in $S(n)$ can be described in almost any coordinate system such that $|\mu| = n$. We shall see later that the partition μ and the corresponding coordinate vector τ_μ specify a pseudocanonical (SS, MFD, or ARMAX) form. The message therefore is that a given system of order n can be almost surely represented by any of the

$$\begin{bmatrix} n + p - 1 \\ p - 1 \end{bmatrix}$$

pseudocanonical forms corresponding to the

$$\begin{bmatrix} n + p - 1 \\ p - 1 \end{bmatrix}$$

partitions μ of n.

Comment 2. Let μ and ν be two partitions such that $|\mu| = |\nu| = n$. The intersection $U_\mu \cap U_\nu$ is also dense in S(n). A point in that intersection can be represented by either $\tau_\mu \in \Theta_\mu$ or $\tau_\nu \in \Theta_\nu$. Since S(n) is an analytic manifold, it follows that these two sets of coordinates are analytically related; that is the mapping

$$\Phi_\mu \cdot \Phi_\nu^{-1}: \Phi_\nu(U_\mu \cap U_\nu) \to \Phi_\mu(U_\mu \cap U_\nu) \tag{38}$$

is analytic.

Comment 3. S(n) can be covered by the

$$\begin{bmatrix} n + p - 1 \\ p - 1 \end{bmatrix}$$

open sets U_μ, $|\mu| = n$. Whether it can be covered by fewer sets is still an open question. It is known that if $p > 1$, no unique set U_μ can cover all of S(n). This has an important consequence: it means that the set of all systems in S(n) cannot be described by a unique identifiable representation.

Having described the structure of S(n), one can now think of the identification problem in the following terms. Estimate the order n, and then take any partition μ of n such that $|\mu| = n$ and compute the maximum likelihood estimate $\hat{\tau}_\mu$ of the corresponding vector τ_μ that completely specifies the system. However, this requires an algorithm that necessitates an I/O description of the point τ_μ. It turns out that the pseudo-

canonical SS, MFD, or ARMAX representations obtained from the τ_μ are not the simplest ones, as we shall see in the next sections. This is due to the overlap between the subsets U_μ for different μ. An alternative is to cover $S(n)$ by disjoint subsets V_μ, which can again be coordinatized by nonoverlapping coordinate systems. These will give rise to the somewhat simpler canonical SS, MFD, and ARMAX forms.

Definition 3. We call V_μ the subset of U_μ for which the rows (34) specified by μ are the first n linearly independent rows of $H_{1,\infty}[K]$.

Since the row r_{i,n_i+1} is now a linear combination of the basis rows above it, (35) is replaced by

$$r_{i,n_i+1} = \sum_{j=1}^{p} \sum_{l=1}^{n_{ij}} \alpha_{ijl} r_{jl}, \quad i = 1, \ldots, p, \tag{39}$$

where

$$\begin{aligned} n_{ij} &\triangleq \min(n_i, n_j) && \text{if } i \leq j, \\ &\triangleq \min(n_i + 1, n_j) && \text{if } i > j. \end{aligned} \tag{40}$$

It follows, by the same argument as before, that any element $K(z) \in V_\mu$ can be mapped into a vector ρ_μ in $R^{d(\mu)}$ defined by the following coordinates:

$$\rho_\mu = \begin{cases} \alpha_{ijl}; & i, j = 1, \ldots, p; \quad l = 1, \ldots, n_{ij}, \\ k_l(i, j); & i = 1, \ldots, p; \quad j = 1, \ldots, s; \quad l = 1, \ldots, n_i, \end{cases} \tag{41}$$

where

$$d(\mu) = n(s + 1) + \sum_{i<j} \{\min(n_i, n_j) + \min(n_i, n_j + 1)\}. \tag{42}$$

We call Ψ_μ the mapping from $K(z)$ into ρ_μ:

$$\Psi_\mu: V_\mu \to X_\mu \subset R^{d(\mu)}: K(z) \in V_\mu \to \rho_\mu = \Psi_\mu[K(z)] \in X_\mu. \quad (43)$$

The following result now holds about the V_μ (see, e.g., [4,14, 19]).

Result 3.

(1) The V_μ are disjoint, $V_\mu \subseteq U_\mu$ and $\bigcup_{|\mu|=n} V_\mu = S(n)$.

(2) Ψ_μ described in (43) is a homeomorphism between V_μ and an open and dense subset X_μ of $R^{d(\mu)}$, with $d(\mu)$ given by (42): $X_\mu = \Psi_\mu(V_\mu)$.

(3) X_μ is an open and dense subset of $R^{d(\mu)}$.

Comment 4. For $p > 1$, $S(n)$ is partitioned into the

$$\begin{bmatrix} n + p - 1 \\ p - 1 \end{bmatrix}$$

disjoint sets V_μ, $|\mu| = n$, which are of different dimensions $d(\mu)$. Hence every system $\sigma \in S(n)$ belongs to one of the V_μ and therefore has a set of structure indices $\mu = (n_1, \ldots, n_p)$ attached to it. Those structure indices are usually called observability indices or left Kronecker indices; they determine the local coordinate system in which that system is described by the vector ρ_μ. These structure indices and these coordinates will in turn define canonical SS, MFD, or ARMAX forms, as we shall see in the next section. They form a complete system of independent invariants for $K(z)$ (see [9,11,37]). Note that in all cases $d(\mu) \leq d = n(p + s)$. The canonical forms therefore will generally have fewer parameters than the corresponding pseudocanonical ones. This is one advantage of canonical forms.

Comment 5. If the first n rows of $H_{1,\infty}[K]$ are linearly independent, then μ has $n_1 = n_2 = \cdots = n_q = n_{q+1} + 1 = \cdots = n_p + 1$

for some q. Then $V_\mu = U_\mu$. Hence this subset is open and dense in S(n), and $\overline{V}_\mu = \overline{U}_\mu = \overline{S}(n)$. This particular V_μ is called the generic neighborhood, because generically a system will have Kronecker indices $n_1 = \cdots = n_q = n_{q+1} + 1 = \cdots = n_p + 1$ for some q. Hence, in practice, the generic neighborhood (and its corresponding canonical parametrization) is sufficient to represent almost any system. However, other nongeneric μ's (and their corresponding parametrizations) might be preferred for numerical reasons. For the generic μ, $d(\mu) = n(p + s)$; the other V_μ are mapped into spaces of lower dimension.

Recall now that a model structure was defined (see Section III, Definition 4) as a mapping from a parameter vector θ to a particular model $M(\theta)$ and that a structure was called globally identifiable if that mapping was injective for almost all θ in a subset D_M. Now we have just shown that (1) given a system $K(z) \in S(n)$, there exists a uniquely defined set of structure indices $\mu = (n_1, \ldots, n_p)$ and a uniquely defined mapping Ψ from K(z) to a parameter vector ρ_μ, and that (z) given a system $K(z) \in S(n)$, for almost any arbitrary set of structure indices $\mu = (n_1, \ldots, n_p)$ such that $\Sigma_1^p n_i = n$, there exists a uniquely defined mapping Φ_μ from K(z) to a parameter vector τ_μ. Therefore if we can now define model structures (in SS, MFD, or ARMAX form) that are entirely specified by the integer valued structure indices and the real-valued parameter vectors (ρ_μ or τ_μ), then these model structures will be identifiable since the sought-after inverse mappins are precisely Ψ_μ and Φ_μ. This is what we will set out to do in the next sections.

V. CANONICAL AND PSEUDOCANONICAL FORMS

We now describe SS, MFD, and ARMA canonical forms for a $p \times s$ transfer function $K(z)$ of order n, assuming that $K(z) \in V_\mu$, where $\mu = (n_1, \ldots, n_p)$. Recall that these Kronecker indices are determined by the linear dependence relations between the rows of $H_{1,\infty}[K]$. For simplicity of notation we assume that

$$y(t) = K(z)m(t) \tag{44}$$

rather than the original model (33) or (1). It is trivial, of course, to split $K(z)$ into $[G(z) \vdots H(z) - I]$, to replace $m(t)$ by

$$\begin{bmatrix} u(t) \\ e(t) \end{bmatrix},$$

and to add $e(t)$ to the right-hand side [see (33)]. An ARMA model would then be converted back into an ARMAX model. However, at this stage the only issue is one of parametrizing a SS, MFD, or ARMA model of a strictly causal rational transfer function.

A. CANONICAL STATE-SPACE FORM

A canonical state-space model A, B, C such that $K(z) = C(zI - A)^{-1}B$ is obtained from the complete set of invariants defining the vector ρ_μ [see (41)] as follows:

$$\underset{(p\times n)}{C} = \left[\begin{array}{ccc|ccc|c|ccc} 1 & 0 \cdots & 0 & 0 & \cdots & 0 & & 0 & \cdots & 0 \\ 0 & & & 1 & 0 \cdots & 0 & & & & \\ & & & 0 & & & \cdots & \vdots & & \vdots \\ \vdots & & \vdots & \vdots & & \vdots & & & & \\ & & & & & & & 0 & & 0 \\ 0 & \cdots & 0 & 0 & \cdots & 0 & & 1 & 0 \cdots & 0 \end{array}\right], \tag{45a}$$

$$\underbrace{\qquad\qquad}_{n_1} \quad \underbrace{\qquad\qquad}_{n_2} \quad \underbrace{\qquad\qquad}_{n_p}$$

$$A_{(n\times n)} =$$

$$\left[\begin{array}{c|c|c|c} \begin{matrix}0\\ \vdots\\ 0\end{matrix} \; I_{n_1-1} & \begin{matrix}0 & \cdots & 0\\ \vdots & & \vdots\\ 0 & \cdots & 0\end{matrix} & \cdots & \begin{matrix}0 & \cdots & 0\\ & & \vdots\\ 0 & & 0\end{matrix} \\ \alpha_{111} \cdots \alpha_{11n_1} & \alpha_{121} \cdots \alpha_{12n_{12}}\; 0 \cdots 0 & & \alpha_{1p1} \cdots \alpha_{1pn_{1p}}\; 0 \cdots 0 \\ \hline \vdots & \vdots & \cdots & \vdots \\ \hline \begin{matrix}0 & \cdots & 0\\ \vdots & & \vdots\\ 0 & \cdots & 0\end{matrix} & \begin{matrix}0 & \cdots & 0\\ \vdots & & \vdots\\ 0 & \cdots & 0\end{matrix} & & \begin{matrix}0\\ \vdots\\ 0\end{matrix} \; I_{n_p-1} \\ \alpha_{p11} \cdots \alpha_{p1n_{p1}}\; 0 \cdots 0 & \alpha_{p21} \cdots \alpha_{p2n_{p2}}\; 0 \cdots 0 & & \alpha_{pp1} \cdots \alpha_{ppn_p} \end{array}\right], \tag{45b}$$

$$B_{(n\times s)} = \left[\begin{array}{ccc} k_1(1,1) & \cdots & k_1(1,s) \\ \vdots & & \vdots \\ k_{n_1}(1,1) & \cdots & k_{n_1}(1,s) \\ \hline & \vdots & \\ \hline k_1(p,1) & \cdots & k_1(p,s) \\ \vdots & & \vdots \\ k_{n_p}(p,1) & \cdots & k_{n_p}(p,s) \end{array}\right]. \tag{45c}$$

Note that A and C have a very specific structure while B is fully parametrized. In an identification context, once the structure indices n_i have been estimated, the structure of A and C is completely specified by the 0 and 1 elements, while the α_{ijk} and $k_l(i, j)$ are free parameters. Any arbitrary state-space representation of an nth order system with Kronecker indices $(n_1, \ldots, n_p)$ can be transformed to this canonical form by a similarity transformation (see, e.g., [11] or [38] for details). Finally, note that this form has $d(\mu)$ parameters.

B. CANONICAL MFD FORM

A canonical MFD form $P(z)$, $Q(z)$ such that $K(z) = P^{-1}(z)Q(z)$ is obtained as follows from ρ_μ.

Let

$$\underset{p\times p}{P(z)} \triangleq [p_{ij}(z)] \quad \text{and} \quad \underset{p\times s}{Q(z)} \triangleq [q_{ij}(z)].$$

Then

$$p_{ii}(z) = z^{n_i} - \alpha_{iin_i} z^{n_i-1} - \cdots - \alpha_{iil}, \tag{46a}$$

$$p_{ij}(z) = -\alpha_{ijn_{ij}} z^{n_{ij}-1} - \cdots - \alpha_{ijl} \quad \text{for} \quad i \neq j, \tag{46b}$$

$$q_{ij}(z) = \beta_{ijn_i} z^{n_i-1} + \cdots + \beta_{ijl}, \tag{46c}$$

where the coefficients β_{ijl} are bilinear functions of the coefficients α_{ijl} and $k_l(i, j)$ obtained as follows. Let

$$\underset{(n\times s)}{G} \triangleq \begin{bmatrix} G_1 \\ \vdots \\ G_p \end{bmatrix}, \qquad \underset{(n_i\times s)}{G_i} \triangleq \begin{bmatrix} \beta_{i11} & & \beta_{isl} \\ \vdots & & \vdots \\ \beta_{iln_i} & \cdots & \beta_{isn_i} \end{bmatrix},$$

$$\underset{(n\times n)}{M} \triangleq [M_{ij}], \quad i, j = 1, \ldots, p, \tag{47}$$

with

$$\underset{(n_i\times n_i)}{M_{ii}} \triangleq \begin{bmatrix} -\alpha_{ii2} & -\alpha_{ii3} & \cdots & -\alpha_{iin_i} & 1 \\ -\alpha_{ii3} & & \cdots & 1 & 0 \\ \vdots & & & & \vdots \\ -\alpha_{iin_i} & 1 & & & \\ 1 & 0 & \cdots & & 0 \end{bmatrix},$$

$$M_{\substack{ij \\ (n_i \times n_j)}} \triangleq \begin{bmatrix} -\alpha_{ij2} & -\alpha_{ij3} & \cdots & -\alpha_{ijn_{ij}} & 0 & \cdots & 0 \\ -\alpha_{ij3} & \cdots & & 0 & \cdots & & 0 \\ \vdots & & & & & & \\ -\alpha_{ijn_{ij}} & & \ddots & & & & \vdots \\ 0 & & & & & & \\ \vdots & & & & & & \\ 0 & & \cdots & & & & 0 \end{bmatrix}. \tag{48}$$

Then

$$G = MB. \tag{49}$$

See [11] for a proof. This canonical form also has $d(\mu)$ parameters. It has the following properties, which actually define its structure:

(i) The polynomials on the main diagonal of $P(z)$ are monic with

$$\deg(p_{ii}) = n_i, \tag{50a}$$

(ii) $$\deg(p_{ij}) \leq \deg(p_{ii}) \quad \text{for} \quad j \leq i, \tag{50b}$$

$$\deg(p_{ij}) < \deg(p_{ii}) \quad \text{for} \quad j > i, \tag{50c}$$

$$\deg(p_{ji}) < \deg(p_{ii}) \quad \text{for} \quad j \neq i, \tag{50d}$$

(iii) $\deg(q_{ij}) < \deg(p_{ii})$ and $P(z)$, $Q(z)$ are left coprime. (50e)

The form $P(z)$, $Q(z)$ with the properties (50) was first proposed by Guidorzi [10] and is called the canonical echelon form in econometrics (see [25]). In the control engineering literature a closely related canonical MFD is called the canonical echelon form (see [29]). It is obtained from the Guidorzi form by permuting the rows of $P(z)$ [and correspondingly of $Q(z)$] such that in the transformed $\bar{P}(z)$ (1) the row degrees are arranged in increasing order and (2) if in $P(z)$ $n_i = n_j$ with $i < j$, then

If we denote by P_{hc} the highest column degree coefficient matrix of $P(z)$ [i.e., the matrix whose columns are the coefficients of the highest power of z in each column of $P(z)$] and by P_{hr} the highest row degree coefficient matrix, then it follows easily from the properties (50) that

$$\text{(i)} \quad P_{hc} = I_p, \tag{51a}$$

$$\text{(ii)} \quad P_{hr} \text{ is lower triangular with unit diagonal elements.} \tag{51b}$$

It follows that $P(z)$ is both column reduced and row reduced (see [29]) with row degrees and column degrees equal to $n_1, \ldots, n_p$. It also follows that

$$\deg \det P(z) = \sum_{1}^{p} n_i = n = \text{order of} \quad K(z). \tag{52}$$

In an identification context, once the Kronecker indices n_i have been estimated, the structure of $P(z)$ and $Q(z)$ is completely specified by (46) or, equivalently, by the degree relations (50), where the α_{ijk} and β_{ijk} are free parameters to be estimated from the data; note that the number of free parameters β_{ijk} is identical to the number of parameters $k_\ell(i, j)$ in ρ_μ, that is, ns.

C. *CANONICAL ARMA FORM*

Using (25) it is easy to obtain a canonical ARMA model from the echelon MFD model:

$$[\overline{P}(D) \vdots \overline{Q}(d)] = M(D)[P(z) \vdots Q(z)], \tag{53}$$

where $M(D) \triangleq \text{diag}\{D^{n_1}, \ldots, D^{n_p}\}$. It is easy to see, using (50), that $\overline{P}(D)$, $\overline{Q}(D)$ have the following properties:

$$\text{(i)} \quad \overline{P}(D) \text{ and } \overline{Q}(D) \text{ are left coprime.} \tag{54a}$$

(ii) $\bar{P}(0) = P_{hr}$ is lower triangular and nonsingular. (54b)

(iii) The row degrees of $[\bar{P}(D) \vdots \bar{Q}(D)]$ are $n_1, \ldots, n_p$. (54c)

The parameters in this canonical form are identical to those appearing in the echelon MFD form, and their positions are again determined by the Kronecker indices. It is therefore identifiable. Notice that the row degrees of $\bar{P}(D)$ are not necessarily $n_1, \ldots, n_p$ (this depends on the particular system, i.e., on the values of the coefficients α_{ijk}), while the column degrees of $\bar{P}(D)$ are generically equal to $r \triangleq \max_i \{n_i\}$. Recall also Section II, Comment 2. One major disadvantage of this ARMA canonical form is that $\bar{P}(0) \neq I$; that is, we cannot write $y(t)$ in the form (24). One way to obtain an ARMA (or ARMAX) form such as (24) is to multiply $[\bar{P}(D) \vdots \bar{Q}(D)]$ to the left by P_{hr}^{-1}. However, this increases some of the row degrees of $\bar{P}(D)$ and $\bar{Q}(D)$ (and hence the lag structure of the model) and therefore increases the number of parameters in the model. In fact, it can be shown [31] that a system in $S(n)$ with Kronecker indices $n_1, \ldots, n_p$ can in general not be represented by an ARMA model

$$\bar{P}(D)y(t) = \bar{Q}(D)m(t), \quad \bar{P}(0) = I, \tag{55}$$

such that the row degrees of $[\bar{P}(D) \vdots \bar{Q}(D)]$ are $n_1, \ldots, n_p$ and that an ARMA model of the form (55) will have more than $d(\mu)$ parameters. Moreover, ARMA models of the form (55) will generically represent systems whose order is a multiple of p, the dimension of $y(t)$ (see [31] for details).

We return now to the mapping (43). If we denote by $X_\mu(SS)$ [resp. $X_\mu(MFD)$, $X_\mu(ARMA)$] the set of all free parameters in the canonical SS (resp. MFD, ARMA) form, then it follows from

Section IV, Result 3 and the fact that the parameters of these three canonical forms are bijectively related to the components of $\rho_\mu \in X_\mu$, that the sets X_μ, $X_\mu(SS)$, $X_\mu(MFD)$, and $X_\mu(ARMA)$ are homeomorphic. Finally, note that the canonical forms we have described here are just one possible set of canonical forms. Using different (but uniquely defined) selection rules for the basis vectors of $H_{1,\infty}[K]$, one can obtain a number of other SS, MFD, and ARMA forms (see, e.g., [7,29,38,39]).

The main disadvantage of canonical forms for identification is that one has to estimate the Kronecker indices. In Section VII we shall briefly describe different methods for doing this, but in any case it is a time-consuming and numerically sensitive procedure. An alternative is to use pseudocanonical forms; this requires the estimation of only one integer-valued parameter, the order n. For almost every system $\sigma \in S(n)$, any set of structure indices $\mu = (n_1, \ldots, n_p)$ such that $\Sigma_1^p h_i = n$ can then be used to define the structure of a pseudocanonical form. We now describe SS, MFD, and ARMA pseudocanonical forms, which are very similar to the canonical forms just described.

D. PSEUDOCANONICAL STATE-SPACE FORM

Let $K(z) \in U_\mu$. Then a pseudocanonical SS form for $K(z)$ is obtained from the parameter vector τ_μ by taking C as in (45a), B as in (45c), and

$$\underset{(n\times n)}{A} \triangleq [A_{ij}], \quad \text{with} \quad \dim A_{ij} = n_i \times n_j, \tag{56}$$

where the $[A_{ii}]$ are as in (45b) and where

$$\underset{(n_i\times n_j)}{A_{ij}} = \begin{bmatrix} 0 & \cdots & 0 \\ \vdots & & \vdots \\ 0 & \cdots & 0 \\ \alpha_{ij1} & \cdots & \alpha_{ijn_j} \end{bmatrix}. \tag{57}$$

(See, e.g., [23] for a derivation of this form.) Note that the free parameters are exactly the coordinates of τ_μ, which uniquely describe the system in the coordinate space defined by μ. Hence this form is identifiable.

E. PSEUDOCANONICAL MFD FORM

A pseudocanonical MFD form for $P(z)$, $Q(z)$ is obtained from the coordinates of τ_μ. [Compare with (46)-(50).]

$$p_{ii}(z) = z^{n_i} - \alpha_{iin_i} z^{n_i-1} - \cdots - \alpha_{iil}, \tag{58a}$$

$$p_{ij}(z) = -\alpha_{ijn_j} z^{n_j-1} - \cdots - \alpha_{ijl} \quad \text{for} \quad i \neq j, \tag{58b}$$

$$q_{ij}(z) = \beta_{ij\rho_i} z^{\rho_i-1} + \cdots + \beta_{ijl}, \tag{58c}$$

where $\rho_i \triangleq$ ith row degree of $P(z) = \max(n_i, \max_j\{n_j\} - 1)$. The bilinear relations (49) between the β_{ijl} and $k_l(i, j)$ still hold, with B as before but with

$$\underset{(\rho_i \times s)}{G_i} \triangleq \begin{bmatrix} \beta_{ill} & & \beta_{isl} \\ \vdots & & \vdots \\ \beta_{il\rho_i} & \cdots & \beta_{is\rho_i} \end{bmatrix}, \tag{59}$$

$$\underset{(\rho_i \times n_i)}{M_{ii}} \triangleq \begin{bmatrix} -\alpha_{ii2} & \cdots & \alpha_{iin_i} & 1 \\ \vdots & & & 0 \\ -\alpha_{iin_i} & & \ddots & \\ 1 & & & \vdots \\ 0 & & & \\ \vdots & & & \\ 0 & \cdots\cdots & & 0 \end{bmatrix} \begin{matrix} \\ \\ \\ \\ \\ \\ \rho_i - n_i \end{matrix},$$

$$M_{ij} \triangleq \underset{(\rho_i \times n_j)}{\begin{bmatrix} -\alpha_{ij2} & \cdots & -\alpha_{ijn_j} & 0 \\ \vdots & & & \\ -\alpha_{ijn_j} & \cdots & & \vdots \\ 0 & & & \\ \vdots & & & \\ 0 & \cdots & & 0 \end{bmatrix}}_{\rho_i - n_j + 1}. \tag{60}$$

The relationship (49) for this pseudocanonical form, with G and M as just described, was derived independently in [21,40,41]. The pseudocanonical form has the following properties:

$$\text{(i)} \quad \text{The } p_{ii} \text{ are monic with } \deg(p_{ii}) = n_i. \tag{61a}$$

$$\text{(ii)} \quad \deg(p_{ji}) < \deg(p_{ii}) = n_i \quad \text{for} \quad j \neq i. \tag{61b}$$

$$\text{(iii)} \quad P(z) \text{ and } Q(z) \text{ are left coprime, and deg det } P(z) = n. \tag{61c}$$

It follows from (61b) that $P(z)$ is column reduced with column degrees $(n_1, \ldots, n_p)$. However, it is not row reduced and this could make some of the I/O relations apparently nonstrictly causal if $\rho_i - 1 \geq n_i$. This will be the case if $\max_i\{n_i\} - \min_i\{n_i\} \geq 2$. The problem arises because the parameters β_{ijl} are not all free; the relationship (49) is not a bijection between the $k_l(i, j)$ and the β_{ijl} here (as it was in the canonical echelon form), since the number of elements in G is larger than that in B. It was shown in [19] that the following set of $n(p + s)$ parameters may be chosen as free:

$$\begin{cases} \alpha_{ijl}; & i, j = 1, \ldots, p; \quad l = 1, \ldots, n_j, \\ \beta_{ijl}; & i = 1, \ldots, p; \quad j = 1, \ldots, s; \quad l = 1, \ldots, n_i, \end{cases} \tag{62}$$

Compare with (36). The other β_{ijl} are then nonlinear combinations of the parameters in (62). When these nonlinear con-

causalities disappear. In an identification context, if the α_{ijl} and β_{ijl} are estimated independently, the constraints will not be exactly satisfied, and noncausal relations may appear. However, Correa and Glover have explicitly computed these constraints and they have shown that by reordering the output variables such that $n_1 \geq \cdots \geq n_p$ (this is always possible), one can treat the I/O relations one by one and successively eliminate the dependent β_{ijl} [24].

F. PSEUDOCANONICAL ARMA FORM

Because $P(z)$ is not row proper, P_{hr} can be singular. Therefore, if we multiply $[P(z) \vdots Q(z)]$ to the left by $M(D)$ as in (53), there is no guarantee that $\bar{P}(0)$ is nonsingular, making the ARMA pseudocanonical form difficult to use. In Section VI we shall present an alternative ARMA form with $\bar{P}(0) = I$. While this form has more than $n(p + s)$ parameters [recall that $n(p + s)$ is the dimension of the space of the overlapping subsets U_μ] and requires a larger number of integer-valued structure indices for its definition, it is identifiable. Other identifiable ARMA forms that have more than the minimum number of parameters have been proposed in [23].

The choice between using either the canonical forms or the pseudocanonical forms described in this section is a fairly subjective one. Canonical forms may have slightly fewer parameters leading to more efficient estimates; on the other hand, the structure estimation step of the identification requires the estimation of p structure indices while, with pseudocanonical forms, only the order must be estimated. If pseudocanonical forms are used and if the (arbitrarily chosen) set of structure indices leads to a numerically ill-conditioned parametrization,

a coordinate transformation can be used [see (38)] to move to a better-conditioned parametrization [16]. If pseudocanonical forms are used, it is always a good idea to start with the generic μ; see Comment 5.

VI. OTHER IDENTIFIABLE PARAMETRIZATIONS

The canonical and pseudocanonical parametrizations of Section V were directly derived from the coordinates τ_μ (resp. ρ_μ) of $K(z)$ in the coordinate spaces spanning U_μ (resp. V_μ). The number of free parameters in these parametrizations is entirely determined by the order of the system (resp. the Kronecker indices) and that number is minimal; it equals the dimension of the space U_μ (resp. V_μ). However, these parametrizations are by no means the only identifiable ones. In this section we describe some other identifiable parametrizations; they will most often have more free parameters than the ones described earlier, but they have some other useful properties.

A. FULLY PARAMETRIZED ARMA MODELS

We consider the class of stable rational strictly proper $K(z)$ of dimension $p \times s$ that can be modeled as

$$\overline{K}(D) \triangleq K(D^{-1}) = \overline{P}^{-1}(D)\overline{Q}(D), \tag{63}$$

where

$$\text{(i)} \quad \overline{P}(D) = I_p + \overline{P}_1 D + \cdots + \overline{P}_u D^u, \tag{64a}$$

$$\overline{Q}(D) = \overline{Q}_1 D + \cdots + \overline{Q}_v D^v, \tag{64b}$$

$$\text{(ii)} \quad \overline{P}(D) \text{ and } \overline{Q}(D) \text{ are left coprime,} \tag{64c}$$

$$\text{(iii)} \quad \operatorname{rank}[\overline{P}_u \vdots \overline{Q}_v] = p. \tag{64d}$$

We call $S(u, v)$ the set of all $K(z)$ that can be modeled by (63)-(64) with prescribed degrees u, v, and we denote by $\Theta(u, v)$ the set of all parameters in $\bar{P}(D)$ and $\bar{Q}(D)$ that are not identically 0 or 1 and for which (64) holds. The set $S(u, v)$ has the following properties (see [20,25,26,42,43]).

Result 1.

(1) $S(u, v)$ is mapped homeomorphically into an open set $\Theta(u, v) \subset R^d$, where $d = p(p \times u + s \times v)$ by the mapping

$$\phi_{u,v}: S(u, v) \to \Theta(u, v): K(z) \in S(u, v) \to \tau_{u,v}[K(z)] \in \Theta(u, v), \tag{65}$$

where $\tau_{u,v}$ is the vector of the coefficients appearing in (64a) and (64b), and hence $\Theta(u, v)$ is identifiable.

(2) $\{S(u, v), u, v \in Z_+\}$ is not a cover of $S(n)$: that is, there exist $K(z) \in S(n)$ for which no u, v exist such that $K(z) \in S(u, v)$.

(3) The $S(u, v)$ are not disjoint: that is, a given $K(z)$ can be in $S(u_1, v_1)$ and $S(u_2, v_2)$ for $(u_1, v_1) \neq (u_2, v_2)$.

Note that d here will always be larger than the dimension $n(p + s)$ of the overlapping submanifolds of Section IV and V. In addition, the order of such models will generically be a multiple of p [31]. The problem raised in the preceding paragraph (2) can be eliminated if, instead of prescribing the highest column degrees (u, v) of $\bar{P}(D)$ and $\bar{Q}(D)$, we prescribe the column degrees $(u_1, \ldots, u_p; v_1, \ldots, v_s)$ of each column of $[\bar{P}(D) \vdots \bar{Q}(D)]$. We denote by a_i (resp. b_i) the vector of coefficients of D^{u_i} (resp. D^{v_i}) in the ith column of $\bar{P}(D)$ [resp. $\bar{Q}(D)$]. Now we denote by $S(u_1, \ldots, u_p; v_1, \ldots, v_s)$ the set of all $K(z)$ that can be modeled by (63), where $\bar{P}(D)$ and $\bar{Q}(D)$ obey

(64a)-(64c) with $u = \max_i u_i$, $v = \max_i v_i$, and where

$$\text{rank}[a_1, \ldots, a_p; b_1, \ldots, b_s] = p \tag{66}$$

for prescribed column degrees $(u_1, \ldots, u_p; v_1, \ldots, v_s)$, and we denote by $\Theta(u_1, \ldots, u_p; v_1, \ldots, v_s)$ the set of all parameters in these $\bar{P}(D)$, $\bar{Q}(D)$ that are not identically 0 and 1. We then have the following result.

Result 2 [20,26,44]

(1) $S(u_1, \ldots, u_p; v_1, \ldots, v_s)$ is mapped homeomorphically onto an open and dense subset $\Theta(u_1, \ldots, u_p; v_1, \ldots, v_s)$ of R^d, where $d = p\left(\Sigma_1^p u_i + \Sigma_1^s v_i\right)$, and hence $\Theta(u_1, \ldots, u_p; v_1, \ldots, v_s)$ is identifiable.

(2) For every $K(z) \in S(n)$, there exists $(u_1, \ldots, u_p; v_1, \ldots, v_s)$ such that $K(z) \in S(u_1, \ldots, u_p; v_1, \ldots, v_s)$.

(3) The $S(u_1, \ldots, u_p; v_1, \ldots, v_s)$ are not disjoint: that is, a $K(z)$ can be modeled uniquely by ARMA models having different u_i, v_i, each obeying the constraints (64a), (64c), and (66).

The most detailed discussion of this identifiable model structure is given in [20]. A major disadvantage of this form is that p + s integer-valued parameters must be prescribed.

B. A "SCALAR" ARMA MODEL

A commonly used representation for $K(z)$ is

$$\bar{p}(D)y(t) = \bar{Q}(D)m(t), \tag{67}$$

where

$$\text{(i)} \quad \bar{p}(D) = 1 + \bar{p}_1 D + \cdots + \bar{p}_u D^u \text{ is a scalar polynomial with } \bar{p}_u \neq 0. \tag{68a}$$

(ii) $\bar{Q}(D) = \bar{Q}_1 D + \cdots + \bar{Q}_v D^v$ has dimension $p \times s$ with

$$\bar{Q}_v \neq 0. \tag{68b}$$

(iii) $\bar{p}(D)I$ and $\bar{Q}(D)$ are left coprime. (68c)

We call $S_{sc}(u, v)$ the set of all stable rational strictly proper $K(z)$ that can be modeled by (67) under the constraints (68). We then have the following result.

Result 3 [42].

(1) $S_{sc}(u, v)$ is mapped homeomorphically onto an open and dense subset $\Theta_{sc}(u, v)$ of R^d with $d = u + p \times s \times v$, and hence $\Theta_{sc}(u, v)$ is identifiable.

(2) $S_{sc}(u, v)$ covers $S(n)$.

This last result follows immediately from the fact that the form (67) is obtained by taking $P(z) = z^r \bar{p}(D)$, with $r \triangleq \max(u, v)$, as the least common denominator of the elements of $K(z)$. Notice finally that the form (67) contains in general more parameters than the canonical or pseudocanonical forms.

C. *ELEMENTARY SUBSYSTEM REPRESENTATIONS*

The elementary subsystem (ESS) representation for multivariable linear systems, introduced in [45]-[46], is based on a decomposition of the monic least common denominator $p(z)$ of the elements of $K(z)$ into irreducible first- and second-degree polynomials:

$$p(z) = \prod_{i=1}^{n_r} p_{ri}(z) \prod_{j=1}^{n_c} p_{cj}(z), \tag{69}$$

where

$$p_{ri}(z) = z + a_i, \quad p_{cj}(z) = z^2 + b_{1j} z + b_{2j}. \tag{70}$$

It is assumed that the poles of $K(z)$ have multiplicity one. Collecting all terms associated with $P_{ri}(z)$ into a matrix $K_{ri}(z)$ and all terms associated with $p_{cj}(z)$ into a matrix $K_{cj}(z)$, one can then write $K(z)$ as the sum of these partial fraction matrices (PFM):

$$K(z) = \sum_{i=1}^{n_r} K_{ri}(z) + \sum_{j=1}^{n_c} K_{cj}(z). \tag{71}$$

The PFM $K_{ri}(z)$ [resp. $K_{cj}(z)$] can then be realized as a direct sum of first-order (resp. second-order) elementary subsystems using dyadic decompositions of the numerator matrices. Now $K(z)$ is finally realized as a direct sum of these direct sums. The procedure is explained in great detail in [46], where a structure estimation scheme is also proposed. The authors claim that this elementary subsystem structure is identifiable; however, it is not clear that this will be the case if several elements of $K(z)$ have the same pole, unless some additional rules are imposed to make the dyadic decompositions unique. On the other hand, the advantage of this structure is that it is often close to the physical model of the system or its subsystems. Since the poles of $K(z)$ are directly estimated, stability can easily be checked, or stability constraints can be introduced.

VII. THE ESTIMATION OF THE STRUCTURE

In this section we shall briefly survey some important theoretical results on structure estimation and point to a number of practical methods that have been proposed for the estimation of the order n or the Kronecker indices $(n_1, \ldots, n_p)$ or the

lag lengths (u, v). In the case where pseudocanonical forms are used, any set of structure indices $\mu = (n_1, \ldots, n_p)$ adding up to n can in principle be used; however, also in this case methods have been suggested for selecting a partition μ that leads to a numerically better-conditioned estimation algorithm. There is a very abundant literature on structure estimation, mainly originating from statisticians. For reasons of space, we shall be able to present only the main thrust of the results without going into technical details. Our starting point will be the model (1) again. We shall sometimes specialize to methods that do not allow deterministic inputs or others that consider only deterministic inputs. For brevity of notation, we shall also sometimes use $K(z) \triangleq [G(z) \vdots H(z) - I]$ as before.

A. *RESULTS USING THE MAXIMUM LIKELIHOOD*

We first discuss some important consistency results for maximum likelihood estimation of parameters. We will assume that the model (1) is subjected to the conditions (2) and that the u(t) are observable. We shall denote by F_t the σ-algebra of events determined by $\{y(s), s \leq t\}$; equivalently, since u(t) is observable and (2e) is assumed, F_t is determined by $\{e(s), s \leq t\}$. It is now assumed that there exists a true system $K_0(z) \triangleq [G_0(z) \vdots H_0(z) - I]$ with innovations e(t) obeying the following assumptions:

$$\text{(i)} \quad E\{e(t) \mid F_{t-1}\} = 0 \text{ a.s.}, \quad E\left\{e(t)e^T(t) \mid F_{-\infty}\right\} = \Sigma_0. \tag{72a}$$

$$\text{(ii)} \quad e(t) \text{ is ergodic.} \tag{72b}$$

$$\text{(iii)} \quad E\left\{[e_j(t)]^4\right\} < \infty \quad \text{for} \quad j = 1, \ldots, p. \tag{72c}$$

Conditions (72a) ensure that $e(t)e^T(t)$ is purely nondeterministic and that the best predictor of $y(t)$ given F_{t-1} is the best linear predictor. We now denote by U any of the sets of transfer functions $K(z)$ described in Sections IV through VI by $\bar{U}$ its closure w.r.t. T_{pt}, by Θ the parameter set of a corresponding identifiable model structure, by Φ the mapping from U to Θ such that $\Phi(K(z)) = \tau \in \Theta$, and by Π the inverse mapping such that $\Pi(\tau) = K(z) \in U$ for $\tau \in \Theta$. For example, if U is taken as U_μ, then $K(z)$ is the set of transfer functions such that $K(z) \in U_\mu$, $\bar{U} = \bar{S}(n)$, where $n = |\mu|$ by Result 2 in Section IV (4), Θ could be taken as $\Theta_\mu(SS)$ (see Section V), any element $K(z) \in U_\mu$ is then mapped into $\tau_\mu = \Phi_\mu(K(z))$ [see (37)] and $\Pi(\tau_\mu) = K(z)$. Setting the initial values of $u(\cdot)$ and $y(\cdot)$ to zero, and denoting $V_N = [v^T(1), \ldots, v^T(N)]^T$, where $v(t) \triangleq y(t) - \sum_1^t G_i u(t-i)$ and

$$\Gamma_N(\tau, \Sigma) \triangleq E\{V_N V_N^T\}, \tag{73}$$

then the likelihood function is given by

$$L_N(\tau, \Sigma) = \frac{1}{N} \log \det \Gamma_N(\tau, \Sigma) + \frac{1}{N} V_N^T \Gamma_N^{-1}(\tau, \Sigma) V_N. \tag{74}$$

Now the important point is that $L_N(\tau, \Sigma)$ depends on the parameter vector τ only through $\Pi(\tau) = K(z)$. Therefore the likelihood function $L_N(\tau, \Sigma)$ can be considered as "coordinate free;" the particular parametrization is unimportant. We then have the following important consistency result [3].

Result 1. Assume that $y(t)$ is generated by an ARMAX process (10) with the assumptions (2) and (72) and assume that $K_0(z) \in \bar{U}$. If $\hat{\tau}_N$, $\hat{\Sigma}_N$ are the MLEs obtained by optimizing $L_N(\tau, \Sigma)$ over

$\bar{U} \times \{\Sigma | \Sigma > 0\}$, and if $\hat{K}_N(z) = \Pi(\hat{\tau}_N)$, then

$$\hat{K}_N(z) \to K_0(z) \quad \text{in } T_{pt} \text{ a.s.}, \quad \text{and} \quad \hat{\Sigma}_N \to \Sigma_0 \quad \text{a.s.} \tag{75}$$

This result has the following consequences.

(1) If, say, U is taken as U_μ, $|\mu| = n$, and if $K_0(z) \in S(j)$, $j \leq n$, then $\bar{U} = \bar{S}(n) = \bigcup_{j \leq n} S(j)$, so that the MLE is consistent even when the true order is exceeded.

(2) Result 1 is about consistency of transfer function estimates, not about consistency of the parameter vector $\hat{\tau}_N$. If U is taken as U_μ and $K_0(z) \in U_\mu$, then $\hat{\tau}_N \to \tau_0$ a.s., but other situations can arise when $K_0(z) \notin U_\mu$ (see [25]).

Result 1 makes the crucial assumption that $K_0(z) \in \bar{U}$. This almost amounts to saying that the structure of $K_0(z)$ is known. We now turn to results on structure estimation. A maximum likelihood criterion cannot be used for, say, order estimation for the following reason. If n_0 is the true order and if $n_1 > n_0$, then $\bar{S}(n_0) \subset \bar{S}(n_1)$ and $S(n_1)$ is dense in $\bar{S}(n_1)$; therefore the MLE over $\bar{S}(n_1)$ will almost surely be attained in $S(n_1)$; MLE will almost always overestimate the order.

Order estimation criteria therefore add a penalty term on the dimension d of the parameter space. They are generally obtained by minimizing a criterion of the form

$$A_N(n) = \log \det \hat{\Sigma}_N(n) + d(n)[C(N)/N], \qquad n = 0, 1, \ldots, n_{max}, \tag{76}$$

where $\hat{\Sigma}_N(n)$ is the MLE of Σ_0 over $\bar{S}(n) \times \{\Sigma | \Sigma > 0\}$, n_{max} is the maximum order to be considered, d(n) is the dimension of the parameter space, N is the number of observations, and C(N) can take different values. Such criteria were first proposed by

Akaike. If $C(N) = 2$, $A_N(n)$ is called AIC [47]; if $C(N) = C \log N$, $A_N(n)$ is called BIC [48]. Other criteria, based on minimum description length of data, are due to Rissanen [49,50]; a third term is added to the expression (76).

The criterion (76) has been expressed as a function of the order n. It can just as well be expressed as a function of the set of Kronecker indices $\mu = (n_1, \ldots, n_p)$ when the search is performed over the disjoint canonical forms instead of the overlapping pseudocanonical forms:

$$A_N(\mu) = \log \det \hat{\Sigma}_N(\mu) + d(\mu)[C(N)/N], \quad |\mu| \leq n_{max}. \tag{77}$$

Under a reasonable set of assumptions, the following results have been obtained in [26] through [28].

Result 2.

(1) For $U = V_\mu$ (i.e., disjoint neighborhoods), BIC gives strongly consistent estimates $\hat{\mu}_N$ of the Kronecker indices.

(2) For $U = U_\mu$ (i.e., overlapping neighborhoods), BIC gives a strongly consistent estimate $\hat{n}_N$ of the true order n_0.

(3) AIC is not consistent in cases (1) and (2) (it overestimates the order).

However, in practice there will be no true order or Kronecker indices: that is, the true system will not have a rational transfer function. AIC seems to be directed at this situation. Using a particular criterion of optimality for this situation, Shibata has shown that AIC has some optimal properties for spectral estimation using autoregressive models [51,52].

For the practical applicability of these structure estimation results, it is important to observe that the criteria (76)-(77) use the MLE $\hat{\Sigma}_N(n)$ or $\hat{\Sigma}_N(\mu)$. Since the search for the Kronecker

indices is to be performed over a large number of candidate models, this represents a formidable computational task. One way of reducing the effort is to search first for the largest Kronecker index by performing the search over $\mu_i = (n_1, \ldots, n_p)$, where $n_1 = \cdots = n_p = i$ for increasing i. If $\hat{\mu}_l = \min_{i=1,2,\ldots} A_N(\mu_i)$, then l is an estimate of the largest Kronecker index. Subsequently a much smaller number of V_μ need be examined to estimate the remaining Kronecker indices. However, this method still requires the minimization of a number of likelihood functions. Therefore other simpler methods have been sought, which we will briefly describe next.

B. *OTHER METHODS FOR THE ESTIMATION OF THE KRONECKER INDICES*

Hannan and Kavalieris [26] (see also [53]) have proposed a method that minimizes the criterion (77), but where the MLE $\hat{\Sigma}_N(\mu)$ is computed cheaply using only linear equations. The method assumes that K(z) belongs to a generic neighborhood, that is, that $n_1 = n_2 = \cdots n_q = n_{q+1} + 1 = \cdots = n_p + 1$ for some q. This is not a severe limitation, since generic neighborhoods are dense in $\bar{S}(n)$ (see Comment 5). On the other hand, it greatly reduces the number of candidate models. It is also assumed that y(t) is generated by an ARMAX model. The method uses a three-stage procedure inspired by Durbin [54]. In stage I an autoregressive model, whose order increases with the number of data, is fitted to the data; it is used to compute estimates of the innovations. Using these estimates, canonical and generic ARMAX models are fitted in stage II; the residual variances from these models are used in (77) to estimate μ. Once the model structure is chosen, stage III computes MLEs of the

parameters. It is shown that this procedure gives consistent estimates of the Kronecker indices under assumptions that are only slightly stronger than those required for Result 2.

A number of other simple methods have been proposed for the estimation of the Kronecker indices; however, no consistency results are available for these methods. They are all based on the fact that the Kronecker indices can be inferred from the linear dependence relations on the rows of $H_{1,\infty}[K]$. The problem is that K(z) is not known. One procedure is to first estimate the K_i by a long autoregression. If there are no deterministic inputs, then the covariance matrix between the vectors of "future" outputs and "past" outputs can be used to establish the linear dependence relationships; the rows of that block-Hankel covariance matrix have the same linear dependences as the rows of $H_{1,\infty}[K]$. This observation has led to the canonical correlation analysis proposed by Akaike [55]. A similar method based on rank tests of covariance matrices has been proposed in [12]. For deterministic I/O models, on the other hand, structure estimation methods have been proposed based on rank tests of the product-moment matrix of the input and output data [10, 11]. These methods will also work when the inputs and outputs are measured with noise, provided that the noise is white or its statistics are known. In [56] another method is proposed that applies to least squares models, that is, H(z) = I in (1) (see also [57]).

C. *STRUCTURE ESTIMATION METHODS USING PSEUDOCANONICAL FORMS*

When pseudocanonical forms are used, it is in principle only necessary to estimate the order n of the system. Result 2

Once the order has been estimated, say $\hat{n}$, any partition μ such that $|\mu| = \hat{n}$ can normally be used to obtain an identifiable model structure. It might well be, however, that the true system is close to the boundary of the selected neighborhood: that is, the coordinates could well be ill conditioned. For numerical reasons, therefore, it may be worth selecting, if not the best, at least a well-conditioned parametrization among the finite set of admissible ones (considering $\hat{n}$ as fixed). A number of methods have been proposed. Ljung and Rissanen have proposed a method based on the complexity of various submatrices of $R_{\hat{y}}$, where $R_{\hat{y}}$ is the covariance matrix of a vector $\hat{Y}_t^N$ made up of a finite set of predictors $\hat{y}(t + 1|t), \ldots, \hat{y}(t + N|t)$ [15]. Wertz, Gevers, and Hannan have proposed a Q - R factorization of Akaike's covariance matrix between future and past outputs, where at each step of the factorization the most independent remaining vector is added to the basis [17]. Van Overbeek and Ljung have proposed a method that is based on the conditioning of the information matrix; their procedure is not to search a priori for the "best" coordinate system, but to perform a coordinate transformation if the parameters in the present coordinate system become ill conditioned [16]. All these methods apply only to state-space models; in addition, they are covariance methods, which will only work when no deterministic inputs are present.

A lot more should be said about structure estimation. In particular, we have hardly touched on the consistency results for $\hat{K}_N(z)$ (or for $\hat{\tau}_N$) when the model order is either larger or smaller than that of the true system. We refer to the work of Hannan, Deistler, and Kavalieris for a discussion of these issues. A very readable discussion can be found in [58].

VIII. CONCLUSIONS

We have given a broad overview of the issues involved in selecting identifiable parametrizations and have presented most of the commonly used parametrizations and the techniques for selecting them. In the past decade researchers in this field have gained a much better understanding of the structure of multivariable systems, and yet there is still no consensus on a universal technique for the representation of such systems. This is due to several reasons. First there is still no agreement on a universal order or structure selection criterion, and there might never be one, because it is recognized more and more that an order or structure selection criterion will always have to incorporate a degree of subjectivity. The research of the past few years has focused on finding structure selection criteria that converge to the true structure when a true system is assumed to exist. Those "optimal" criteria require an enormous amount of computations, as they require the maximization of a large number of likelihood functions. This has led people to search for computationally cheaper suboptimal methods, most of which rely on very heuristic arguments. This is a major reason for the wide range of existing methods.

A second reason is that most researchers now recognize the fact that in most practical applications the true system being identified is infinite dimensional. It is not certain that an "optimal" criterion that converges to the true structure when such exists will also be the best one in the more realistic situation when no true system exists. In this situation the structure selection criterion should most certainly incorporate the intended use of the model. Very little if no effort has

Because of this lack of universal agreement, the methods people use will be influenced very much by their familiarity with a particular method, by the intended use of the model, and by the availability of a particular software package. All of this probably accounts for the relatively small number of successful applications reported in the literature and the difficulty of comparing results obtained with different methods.

Finally, we wish to conclude on two practical notes. First, there is no denying that tremendous new insights have been gained from the theoretical research of the past decade on the estimation of structure of multivariable systems. However, since the generic parametrization is able to represent almost all systems (see Comment 5), in practice it is most often sufficient to estimate the order of the system, rather than its entire structure. Second, companion (or bloc-companion) forms are notoriously sensitive to numerical errors in the parameters. From a practical point of view, more research should be spent on finding numerically insensitive multivariable parametrizations

REFERENCES

1. R. E. KALMAN, "Algebraic Geometric Description of the Class of Linear Systems of Constant Dimension," *Annu. Princeton Conf. Inf. Sci. Syst., 8th, Princeton, New Jersey* (1974).

2. J. M. C. CLARK, "The Consistent Selection of Parametrizations in Systems Identification," *Joint Automatic Control Conf.*, Purdue University (1976).

3. W. DUNSMUIR and E. J. HANNAN, "Vector Linear Time-Series Models," *Adv. Appl. Probab. 8*, 339-364 (1976).

4. M. HAZEWINKEL, "Moduli and Canonical Forms for Linear Dynamical Systems II: The Topological Case," *Math. Syst. Theory 10*, 363-385 (1977).

5. D. G. LUENBERGER, "Canonical Forms for Linear Multivariable Systems," *IEEE Trans. Autom. Control AC-12*, 290-293 (1976).

6. D. Q. MAYNE, "A Canonical Model for Identification of Multivariable Linear Systems," *IEEE Trans. Autom. Control AC-17*, 728-729 (1972).

7. M. J. DENHAM, "Canonical Forms for the Identification of Multivariable Linear Systems," *IEEE Trans. Autom. Control AC-19*, 646-656 (1974).

8. K. GLOVER and J. C. WILLEMS, "Parametrization of Linear Dynamical Control Systems, Canonical Forms and Identifiability," *IEEE Trans. Autom. Control AC-19*, 640-646 (1974).

9. J. RISSANEN, "Basis of Invariants and Canonical Forms for Linear Dynamical Systems," *Automatica 10*, 175-182 (1974).

10. R. P. GUIDORZI, "Canonical Structures in the Identification of Multivariable Systems," *Automatica 11*, 361-374 (1975).

11. R. P. GUIDORZI, "Invariants and Canonical Forms for Systems Structural and Parametric Identification," *Atuomatica 17*, 117-133 (1981).

12. E. TSE and H. L. WEINERT, "Structure Determination and Parameter Identification for Multivariable Stochastic Linear Systems," *IEEE Trans. Autom. Control AC-20*, 603-613 (1975).

13. P. E. CAINES and J. RISSANEN, "Maximum Likelihood Estimation of Parameters in Multivariate Gaussian Stochastic Processes," *IEEE Trans. Inf. Theory IT-20*, 102-104 (1974).

14. M. HAZEWINKEL and R. E. KALMAN, "On Invariants, Canonical Forms and Moduli for Linear Constant Finite-Dimensional Dynamical Systems," *in* "Lecture Notes Economical-Mathematical System Theory, Vol. 131, pp. 48-60, Springer-Verlag, Berlin and New York, 1976.

15. L. LJUNG and J. RISSANEN, "On Canonical Forms, Parameter Identifiability and the Concept of Complexity," *Proc. IFAC Symp. Identif. Syst. Parameter Estim., 4th, Tbilisi*, 58-69 (1976).

16. A. J. M. VAN OVERBEEK and L. LJUNG, "On-Line Structure Selection for Multivariable State-Space Models," *Automatica 18*, 529-544 (1982).

17. V. WERTZ, M. GEVERS, and E. J. HANNAN, "The Determination of Optimum Structures for the State-Space Representation of Multivariate Stochastic Processes," *IEEE Trans. Autom. Control AC-27*, 1200-1211 (1982).

18. G. PICCI, "Some Numerical Aspects of Multivariable Systems Identification," *Math. Program. Stud. 18*, 76-101 (1982).

19. M. DEISTLER and E. J. HANNAN, "Some Properties of the Parametrization of ARMA Systems with Unknown Order," *J. Multivariate Anal. 11*, 474-484 (1981).

20. M. DEISTLER, "The Properties of the Parametrization of ARMAX Systems and Their Relevance for Structural Estimation," *Econometrica 51*, 1187-1207 (1983).

21. R. P. GUIDORZI and S. BEGHELLI, "Input-Output Multistructural Models in Multivariable Systems Identifaction," *Proc. IFAC Symp. Identif. Syst. Parameter Estim., 6th, Washington D.C.*, 461-466 (1982).

22. M. GEVERS and V. WERTZ, "Overlapping Parametrizations for the Representation of Multivariate Stationary Time Series," *in* "Geometry and Identification," pp. 73-99 (P. E. Caines and R. Hermann, eds.), Math. Sci. Press, Massachusetts, 1983.

23. M. GEVERS and V. WERTZ, "Uniquely Identifiable State-Space and ARMA Parametrizations for Multivariable Linear Systems," *Automatica 20*, 333-348 (1984).

24. G. O. CORREA and K. GLOVER, "Pseudocanonical Forms, Identifiable Parametrizations and Simple Parameter Estimation for Linear Multivariable Systems; Part I: Input-Output Models; Part II: Parameter Estimation," *Automatica 20*, 443-452 (1984).

25. M. DEISTLER, "General Structure and Parametrization of ARMA and State Space Systems and Its Relation to Statistical Problems," *in* "Handbook of Statistics," Vol. 5 (E. J. Hannan, P. R. Krishnaiah, and M. M. Rao, eds.), North Holland Publ., Amsterdam, 1983.

26. E. J. HANNAN and L. KAVALIERIS, "Multivariate Linear Time Series Models," *Adv. Appl. Probab. 16*, 492-561 (1984).

27. E. J. HANNAN, "The Estimation of the Order of an ARMA Process," *Ann. Stat. 8*, 1071-1081 (1980).

28. E. J. HANNAN, "Estimating the Dimension of a Linear System," *J. Multivariate Anal. 11*, 459-473 (1981).

29. T. KAILATH, "Linear Systems," Prentice-Hall, Englewood Cliffs, New Hersey, 1980.

30. W. A. WOLOWICH and H. ELLIOTT, "Discrete Models for Linear Multivariable Systems," *Int. J. Control 38*, 337-357 (1983).

31. M. GEVERS, "On ARMA Models, Their Kronecker Indices and Their McMillan Degree," *Int. J. Control 43*, 1745-1761 (1986).

32. L. LJUNG, "System Identification: Theory for the User," Prentice-Hall, Englewood Cliffs, New Jersey, 1986. 1986.

33. R. BELLMAN and K. J. ÅSTRÖM, "On Structural Identifiability," *Math. Biosci. 7*, 329-339 (1970).

34. L. LJUNG, "On Consistency and Identifiability," *Math. Program. Stud. 5*, 169-190 (1976).

35. V. SOLO, "Topics in Advanced Time Series Analysis," *in* "Lecture Notes in Mathematics," Springer-Verlag, Berlin and New York, 1983.

36. M. DEISTLER and H. G. SEIFERT, "Identifiability and Consistent Estimability in Econometric Models," *Econometrica 46*, 969-980 (1978).

37. V. M. POPOV, "Invariant Description of Linear, Time Invariant Controllable Systems," *SIAM J. Control 10*, 252-264 (1972).

38. A. C. ANTOULAS, "On Canonical Forms for Linear Constant Systems," *Int. J. Control 33*, 95-122 (1981).

39. W. A. WOLOWICH, "Linear Multivariable Systems," *Appl. Math. Sci. 11* (1974).

40. G. O. CORREA and K. GLOVER, "Multivariable Identification Using Pseudo-Canonical Forms," *Proc. IFAC Symp. Identif. Syst. Parameter Estim., 6th, Washington D.C.*, 1110-1115 (1982).

41. M. GEVERS and V. WERTZ, "On the Problem of Structure Selection for the Identification of Stationary Stochastic Processes," *Proc. IFAC Symp. Identif. Syst. Parameter Estim., 6th, Washington D.C.*, 287-292 (1982).

42. E. J. HANNAN, "The Identification and Parametrization of ARMAX and State Space Forms," *Econometrica 44*, 713-723 (1976).

43. P. STOICA, "On Full Matrix Fraction Descriptions," *Bull. Inst. Politech. Buc. Ser. Electro.*, 99-104 (1982).

44. M. DEISTLER, W. DUNSMUIR, and E. J. HANNAN, "Vector Linear Time Series Models: Corrections and Extensions," *Adv. Appl. Probab. 10*, 360-372 (1978).

45. L. KEVICZKY, J. BOKOR, and CS. BANYASZ, "A New Identification Method with Special Parametrization for Model Structure Determination," *Proc. IFAC Symp. Identif. Syst. Paramater Estim., 5th, Darmstadt, Federal Republic of Germany*, 561-568 (1979).

46. J. BOKOR and L. KEVICZKY, "Structure and Parameter Estimation of MIMO Systems Using Elementary Sub-System Representations," *Int. J. Control 39*, 965-986 (1984).

47. H. AKAIKE, "Information Theory and an Extension of the Maximum Likelihood Principle," *Proc. Int. Symp. Info. Theory, 2nd*; Supplement to "Problems of Control and Information Theory," 267-281 (1972).

48. H. AKAIKE, "On Entropy Maximisation Principle," *in* "Applications of Statistics," (P. R. Krishnaiah, ed.), North-Holland Publ., Amsterdam, 1977.

49. J. RISSANEN, "Modeling by Shortest Data Description," *Automatica 14*, 465-471 (1978).

50. J. RISSANEN, "Estimation of Structure by Minimum Description Length," *Circuits Syst. Signal Process. 1*, 395-406 (1982).

51. R. SHIBATA, "Asymptotically Efficient Selection of the Order of the Model for Estimating Parameters of a Linear Process," *Ann. Stat. 8*, 147-164 (1980).

52. R. SHIBATA, "An Optimal Autoregression Spectral Estimate," *Ann. Stat. 9*, 300-306 (1981).

53. L. KAVALIERIS, "Statistical Problems in Linear Systems," Ph.D. Dissertation, Australian National University, 1984.

54. J. DURBIN, "Fitting of Time Series Models," *Rev. Inst. Stat. 28*, 233-244 (1961).

55. H. AKAIKE, "Canonical Correlation Analysis of Time Series and the Use of an Information Criterion," *in* "System Identification: Advances and Case Studies," pp. 27-96 (R. K. Mehra and D. G. Lainiotis, eds.), Academic Press, New York, 1976.

56. R. P. GUIDORZI, M. LOSITO, and T. MURATORI, "The Range Error Test in the Structural Identification of Linear Multivariable Systems," *IEEE Trans. Autom. Control Ac-27*, 1044-1054 (1982).

57. P. STOICA, "Comments on 'The Range Error Test in the Structural Identification of Linear Multivariable Systems'," *IEEE Trans. Autom. Control AC-29*, 379-381 (1984).

58. M. DEISTLER, "Multivariate Time Series and Linear Dynamic Systems," *Adv. Stat. Methods Comput. 1*, in press (1985).