The Determination of Optimum Structures for the State Space Representation of Multivariate Stochastic Processes

VINCENT WERTZ, MICHEL GEVERS, MEMBER, IEEE, AND EDWARD J. HANNAN

Abstract—**When identifying a model for a multivariate stationary stochastic process, an important problem is that of determining the structure of the state-variable model. Several "overlapping" parameterizations can usually be fitted to a given process, and the question arises as to which structure leads to the most accurate parameter estimates. The accuracy of parameter estimates is often measured by the determinant of the Fisher information matrix. We show that all admissible structures will give asymptotically the same value to this criterion. For finite data some structures may still be better than others, and two heuristic structure estimation methods are analyzed. Some simulation results are also presented.**

I. INTRODUCTION

AN important and widely studied problem in the theory of identification of linear systems is the problem of determining the structure of a state space representation of the type

$$x_t = Fx_{t-1} + Ke_t$$
$$y_t = Hx_t. \tag{1.1}$$

Here y_t is a stationary purely random stochastic vector process, e_t is a vector white noise of the same dimension as y_t, and x_t is the state. In multivariable systems, the structure estimation problem does not consist only of the determination of the dimension of the state vector x_t (as is the case for single-input–single-output systems); it also involves the determination of a special structure (or parameterization) for the F, K, and H matrices such that they can be uniquely determined in a parameter estimation procedure. These structures are problem-dependent, i.e., there is no universal structure that could be used for all multivariable linear systems of the same order. Another difficulty is that a given system can often be uniquely parameterized with more than one structure. Hence the structure estimation problem is to first determine a set of admissible structures for a given y_t process and then to choose the "best" structure within that set, i.e., the structure within which the parameter estimation problem will be best conditioned.

This problem has been given partial answers by various authors (see [1]–[8] and [18]), but most of the proposed methods are based on heuristic arguments. They are mainly based on the idea that the components of the state vector should be selected in such a way that they are as orthogonal as possible to one another. The motivation behind this idea is that this should make the parameter estimation problem well-conditioned. In fact, we show in this paper that if the determinant of the information matrix is used as the criterion to be maximized, then all admissible structures are asymptotically equivalent. This will be our main result. The paper is organized as follows. In Section II we discuss what is meant by admissible structures, and we show that some constraints on the choice of these structures can lead to a minimum number of parameters to be estimated.

In Section III we give the main result of this paper. Since this is an asymptotic result, the problem remains as to which structure to choose when only a finite number of observations is available. In addition, other criteria could be used, such as the trace (rather than the determinant) of the information matrix, which might enable one to discriminate between various admissible structures. Sections IV and V provide heuristic answers to this question. Finally, we give some simulation results in Section VI.

The main body of this paper is the work of V. Wertz and M. Gevers. The results of Section III triggered the interest of E. J. Hannan, who provided the Appendix. The theorem proved in the Appendix is important because it specifically establishes that the main result of this paper, namely Lemma 3.1, is valid also in the case where certain elements in the system matrices are constrained.

Finally, we want to stress that the problem discussed in this paper is whether or not and how an optimal structure can be selected among a finite number of equivalent state space representations. The more general question of how to define uniquely identifiable parameterizations for multivariable state space and ARMA models is treated in [21], using an algebraic-geometric approach and the concept of intrinsic invariants.

Manuscript received April 24, 1981; revised December 15, 1981. Paper recommended by B. W. Dickinson, Past Chairman of the Identification Committee. This work was supported by IRSIA (Belgium) and the Australian Research Grants Committee.

V. Wertz and M. Gevers were on leave at the University of Newcastle, N.S.W., 2308, Australia. They are with the Department of Electrical Engineering, Louvain University, Louvain-la-Neuve, Belgium.

E. J. Hannan is with the Department of Statistics, Australian National University, Canberra, ACT 2600, Australia.

0018-9286/82/1200-1200$00.75

II. Statement of the Problem

We will be dealing hereafter only with linear stationary full-rank stochastic processes y_t that can be modeled by finite-dimensional state space representations

$$\begin{aligned} x_t &= Fx_{t-1} + Ke_t \\ y_t &= Hx_t. \end{aligned} \tag{2.1}$$

y_t is the observation process of dimension p, x_t is the state vector of dimension n, e_t is a Gaussian white noise, also of dimension p, and F, K, H are constant matrices of appropriate dimensions. We consider that the components of the y_t's, $t \in Z$, span a Hilbert space $\boldsymbol{H}$. The stochastic input e_t can then be viewed as the innovation of the y_t process, defined as follows:

$$e_t \triangleq y_t - \hat{y}_{t|t-1} \tag{2.2}$$

where $\hat{y}_{t|t-1}$, the one-step-ahead predictor of y_t, is the projection of y_t onto the space $(Y_{-\infty}^{t-1})$, the closed linear subspace of $\boldsymbol{H}$ spanned by the components of y_τ for $\tau \leqslant t-1$. By the Gaussian assumption

$$\hat{y}_{t|t-1} \triangleq E\{y_t \leqslant Y_{-\infty}^{t-1}\}.$$

The identification of the structure of a state space representation such as (2.1) consists of 1) the determination of the dimension n of the state vector and 2) the location, in the matrix F, of the parameters to be estimated.

In fact, as we shall show later, the determination of the structure of the model for the process y_t leads to some entries of the F matrix being set to 0 or 1, and thus not having to be estimated. The location of these "0" and "1" entries is an important part of the structure determination because it reduces the number of unknown parameters and ensures that the estimation problem will be "well posed." This aspect also constitutes the main difference between the identification of a multivariate process and that of a univariate process.

We now show that the structure determination problem amounts to the selection of a particular basis in the prediction space spanned by the components of $\hat{y}_{t+k|t} \triangleq E\{y_{t+k}|Y_{-\infty}^{t}\}$ for $k = 0, 1, \cdots$.

Consider the following impulse response representation for y_t:

$$y_t = \sum_{i=0}^{\infty} H_i e_{t-i}. \tag{2.3}$$

Then $H_i = HF^iK$, and, because of the fact that the innovations are a zero-mean white-noise process, we may write $\hat{y}_{t+k|t}$ as

$$\hat{y}_{t+k|t} = \sum_{i=k}^{\infty} H_i e_{t+k-i}. \tag{2.4}$$

Note that the filtered estimate $\hat{y}_{t|t}$ is equal to y_t. We now define the Np-dimensional prediction vector $\hat{Y}_t^N$ as follows:

$$\hat{Y}_t^N = \begin{bmatrix} \hat{y}_{t|t} \\ \hat{y}_{t+1|t} \\ \vdots \\ \hat{y}_{t+N-1|t} \end{bmatrix}.$$

Using (2.4) we can write

$$\hat{Y}_t^N = H_{N,\infty} \begin{bmatrix} e_t \\ e_{t-1} \\ \vdots \end{bmatrix} \tag{2.5}$$

where $H_{N,\infty}$ is a semi-infinite Hankel matrix defined as follows:

$$H_{N,M} \triangleq \begin{bmatrix} H_0 & H_1 & \cdots & H_{M-1} \\ H_1 & H_2 & \cdots & \\ \vdots & & & \vdots \\ H_{N-1} & & \cdots & H_{N+M-2} \end{bmatrix}.$$

Because of the assumption that the process y_t is of finite order n, the rank of the matrix $H_{N,M}$ is less than or equal to n whatever the values of N and M. Hence, for N sufficiently large, we can choose n independent rows in $H_{N,\infty}$ and correspondingly n independent components of the prediction vector $\hat{Y}_t^N$ that will constitute a basis for the space spanned by all the components of $\hat{Y}_t^N$. Let x_t be the vector formed with these independent components of $\hat{Y}_t^N$. Then it is easy to show that x_t can be related to x_{t-1} by the equation

$$x_t = Fx_{t-1} + Ke_t \tag{2.6a}$$

and that y_t is a linear combination of x_t as follows:

$$y_t = Hx_t. \tag{2.6b}$$

Equations (2.6a) and (2.6b) together constitute a state space representation for the process y_t.

The following definition will be very useful in the sequel. We define h, the structural vector of the state space representation, as the vector containing the indexes of the rows of $H_{N,\infty}$ (or equivalently, the components of the prediction vector) that have been chosen in the basis (to form x_t).

We shall now illustrate by a simple example the fact that several entries of the F matrix can be made 0 or 1. Suppose we have a process y_t of dimension 2 (i.e., $p = 2$) and of order 3 (i.e., $n = 3$). If we are looking for an arbitrary basis of the prediction space, it is clear that many selections of three independent components of the prediction vector can be made. We therefore introduce some restrictive conditions for the selection of a basis. First, note that because of the form of the Hankel matrix $H_{N,\infty}$, if the jth component of $\hat{y}_{t+i|t}$, $\hat{y}_{t+i|t}^j$, is in the span of the preceding components of the prediction vector $\hat{Y}_t^N$, then so will be $\hat{y}_{t+k|t}^j$ for $k > i$.

Secondly, since y_t is a full-rank process, it is clear that the first p components of the prediction vector $\hat{Y}_t^N$ are

independent. Hence, we impose two conditions on the selection of the basis: *condition 1)* if $i \in h$, then $i - p \in h$, and *condition 2)* $i \in h$ for $i = 1, 2, \cdots p$.

Definition 2.1: If the selection procedure obeys conditions 1) and 2), then the corresponding structural vector h will be called "nice." If the state x_t is formed by selecting n components of $\hat{Y}_t^N$ obtained from a nice structural vector h, then the corresponding F, K, H structure will be called "admissible."

Returning to our simple example, and subject to conditions 1) and 2), there will be only two possible choices for a basis of the prediction space, corresponding to the nice structural vectors $h_1 = (1,2,3)$ and $h_2 = (1,2,4)$. The F matrix corresponding to h_1 is

$$F_1 = \begin{pmatrix} 0 & 0 & 1 \\ X & X & X \\ X & X & X \end{pmatrix}$$

where X denotes parameters to be estimated. With h_2 the matrix F can take two forms as follows:

$$F_{21} = \begin{pmatrix} X & X & 0 \\ 0 & 0 & 1 \\ X & X & X \end{pmatrix} \quad \text{or} \quad F_{22} = \begin{pmatrix} X & X & X \\ 0 & 0 & 1 \\ X & X & X \end{pmatrix}.$$

The matrix F_{21} corresponds to the situation where $\hat{y}_{t+1|t}^1$ lies in the span of $\hat{y}_{t|t}^1$ and $\hat{y}_{t|t}^2$. The matrix F_{21} can also be seen as the result of a procedure selecting the *first* independent components of the predictor space, while the matrix F_{22} could be the result of a procedure selecting the "best" independent components of the prediction space, subject to conditions 1) and 2). In this paper, we shall not insist on selecting the first n independent components of $\hat{Y}_t^N$, but rather the "most linearly independent set" in a sense that will be made precise later. The objective will be to obtain a well-conditioned parameter estimation problem. In the above example this will correspond to choosing between the admissible structures F_1 or F_{22}, depending on whether or not the components corresponding to h_1 or h_2 constitute the "most linearly independent set."

An advantage of this approach is that it allows overlapping between different parameterizations, as was stated by Glover and Willems [9], Rissanen and Ljung [4], and Van Overbeek and Ljung [5], and hence it is possible to change parameterization via a similarity transform during the parameter estimation process without losing any information.

Now it is not easy to define the "best" independent predictors. One selection procedure has been proposed by Ljung and Rissanen [3], [4] using some complexity ideas. However, their method was based on heuristic arguments and did not relate the choice of the "best" basis to the errors of the parameter estimates in that given basis. Clearly, the optimal choice would be the parameterization that leads to the most accurate parameter estimates. Assuming a maximum likelihood estimation method is used to estimate the parameters, then an optimal structure would be the parameterization for which the determinant of the Fisher information matrix M_θ is maximized, with

$$M_\theta = E_{Y|\theta}\left\{\left(\frac{\partial \log p(Y|\theta)}{\partial \theta}\right)^T\left(\frac{\partial \log p(Y|\theta)}{\partial \theta}\right)\right\} \quad (2.7)$$

where

$$\frac{\partial \log p(Y|\theta)}{\partial \theta}$$

is a row vector. The main result of this paper is that all structures of the same order will asymptotically give the same value to the criterion $J = \log\det M$ (provided, of course, that each state vector is made up of independent components of the prediction vector or independent linear combinations of these).

Since our result is an asymptotic result, the question remains as to whether or not any one structure would be better conditioned than others when only a finite observation record is available for the estimation of the parameters of H, F, and K. Since finite data results are very scarce, the answer to this question is far from trivial. This justifies the use of heuristic approaches to select the "best" independent set of predictor components. In Section IV we briefly present the method of Rissanen and Ljung, and we propose a modification to make their search method iterative. In Section V we present an alternative procedure, also based on heuristic arguments, whose major advantage is that it requires less computation.

We conclude this section by justifying the use of the model (2.6) and conditions 1) and 2) for the selection of the components of the basis. We shall show that this leads to a minimum number of unknown parameters to be estimated.

The example illustrates a more general result, namely, that the rows of the matrix F are of two different types: type 1) $(0 \quad 0 \quad 1)$ or more generally $(0 \cdots 0 \quad 1 \quad 0 \cdots 0)$, and type 2) $(X \quad X \quad X)$, i.e., fully parameterized. In fact, from the way F expresses the linear dependences in the prediction vector (see, e.g., [10], [11]), one can see that there are $n - n_1$ type 1) rows in F, where n_1 is the number of rows of the first block row of $H_{N,\infty}$ $(n_1 \leqslant p)$ that are selected in the basis. The rows of H are also of type 1) or 2), and there are n_1 type 1) rows in H. Hence, there are nn_1 parameters to be estimated in F and $(p - n_1)n$ in H. The total number of parameters in F and H is np, independently of the value of n_1. However, n_1 does have an effect on the number of unknown parameters in K.

Instead of the Markovian model (2.6) one can also choose the following model:

$$\begin{aligned} x_{t+1} &= Fx_t + K_p e_t \\ y_t &= Hx_t + e_t. \end{aligned} \quad (2.8)$$

This model is obtained if the prediction vector $\hat{Y}_t^N$ is started with $\hat{y}_{t+1|t}$ rather than $\hat{y}_{t|t}$. We now examine the number of parameters to be estimated in various structures.

Case 1: Consider the state space representation (2.6).

i) If x_t is chosen so as to satisfy conditions 1) and 2), one can see that $H = [I \quad 0]$, $K = \begin{bmatrix} I \\ K^* \end{bmatrix}$, and that the

number of parameters to be estimated is $2np - p^2$, where n is the dimension of the state of model (2.6).

ii) If x_t is chosen so as to satisfy condition 1) only, then the number of parameters is $2np - n_1 p$, where $n_1 < p$.

iii) If x_t is chosen without satisfying conditions 1) and 2), the number of parameters to be estimated in F may be greater than $2np - n_1 p$.

Case 2: Consider now the state space representation (2.8).

i) With condition 1) being satisfied, the number of parameters to be estimated is $2\bar{n}p$, whether or not condition 2) is satisfied. Here $\bar{n}$ is the dimension of the state of model (2.8).

ii) With neither condition 1) nor 2) being satisfied, the number of parameters may be greater than $2\bar{n}p$.

Note that the number of states n of model (2.6) is, in general, larger than the number of states $\bar{n}$ of model (2.8). For a comparison of models (2.6) and (2.8), see the discussion in, e.g., [19], [20].

Following this analysis, it is reasonable to restrict the selection of basis vectors to "nice" structural vectors, i.e., to impose conditions 1) and 2).

III. Asymptotic Equivalence of All Canonical Forms

Two important methods for parameter estimation are the prediction error method and the maximum likelihood method. In the case of Gaussian innovations, it is well known that these methods yield asymptotically efficient parameters, i.e., the covariance matrix of the estimation errors is asymptotically equal to the inverse of the Fisher information matrix M_θ. Therefore, it seems logical to try to discriminate between various structures by maximizing some scalar measure of this information matrix over all admissible structures. The most obvious (and widely used) criterion is the determinant of the information matrix; this corresponds to minimizing the determinant of the asymptotic error covariance matrix.

We shall show in this section that all structures F, K, and H containing the same number of parameters will asymptotically yield the same value for the determinant of the information matrix. This result is important because it shows that, at least asymptotically, all such structures are equivalent as far as the accuracy of the parameter estimates is concerned if this criterion is used. As a consequence, this criterion is unable to discriminate between two admissible structures for the same process.

The main argument is contained in the following lemma.

Lemma 3.1: Let F, K, H and F^*, K^*, H^* be the matrices of the state space representation (2.1) in two different structures and let T be the $n \times n$ transformation matrix from the triple (F, K, H) to the triple (F^*, K^*, H^*), i.e.,

$$\begin{aligned} F^* &= TFT^{-1} \\ K^* &= TK \\ H^* &= HT^{-1}. \end{aligned} \tag{3.1}$$

Then the absolute value of the Jacobian of the transformation from the parameters in (F, K, H) to the parameters in (F^*, K^*, H^*) is equal to 1.

Proof: Let θ be the k-vector of all parameters in F, K, H and θ^* the corresponding k-vector for F^*, K^*, H^*. We recall that in a general one-to-one functional transformation from the variables θ to the variables $\theta^* = f(\theta)$, the Jacobian of the transformation is the determinant of the matrix C defined by

$$C = \frac{\partial \theta^*}{\partial \theta} = \left(\frac{\partial \theta_i^*}{\partial \theta_j} \right).$$

Writing $\theta_i^* = f_i(\theta_1, \theta_2, \cdots, \theta_k)$ and taking the differential of this expression, we have $d\theta_i^* = \sum \partial \theta_i^* / \partial \theta_j (d\theta_j)$ and hence $d\theta^* = C(d\theta)$.

Thus we see that the transformation of the differential is linear and that the Jacobian is the determinant of this transformation. We shall now look at the transformation from the variables in F, K, H, and T to the variables in F^*, K^*, H^*, and T. This is indeed a one-to-one transformation. Given F, K, H, T, then F^*, K^*, H^*, and T are uniquely determined and the converse is also true. In particular, we shall study the transformation between the differential of these matrices.

Following (3.1) and because $(dT^{-1}) = -T^{-1}(dT)T^{-1}$, we have

$$\begin{aligned} dF^* &= (dT)FT^{-1} + T(dF)T^{-1} - TFT^{-1}(dT)T^{-1} \\ dK^* &= (dT)K + T(dK) \\ dH^* &= (dH)T^{-1} - HT^{-1}(dT)T^{-1} \\ dT &= dT. \end{aligned} \tag{3.2}$$

Hence

$$\begin{aligned} T^{-1}(dF^*)T &= T^{-1}(dT)F + dF - FT^{-1}(dT) \\ T^{-1}(dK^*) &= T^{-1}(dT)K + dK \\ (dH^*)T &= dH - HT^{-1}(dT) \\ T^{-1}(dT) &= T^{-1}(dT). \end{aligned} \tag{3.3}$$

Let

$$\begin{aligned} T^{-1}(dF^*)T &= d\bar{F}^* \\ T^{-1}(dK^*) &= d\bar{K}^* \\ (dH^*)T &= d\bar{H}^* \\ dW &= T^{-1}(dT). \end{aligned} \tag{3.4}$$

Then (3.3) becomes

$$\begin{aligned} d\bar{F}^* &= (dW)F + dF - F(dW) \\ d\bar{K}^* &= (dW)K + dK \\ d\bar{H}^* &= dH - H(dW) \\ dW &= dW. \end{aligned} \tag{3.5}$$

The transformation from (F, K, H, T) to (F^*, K^*, H^*, T) can be viewed as the composition of three transformations as shown in the following diagram:

$$\begin{array}{ccc} (F,K,H,T) & \xrightarrow{J} & (F^*,K^*,H^*,T) \\ J_1 \downarrow & & \uparrow J_3 \\ (F,K,H,W) & \xrightarrow[J_2]{} & (\bar{F}^*,\bar{K}^*,\bar{H}^*,W). \end{array} \tag{3.6}$$

To each of these transformations there corresponds a Jacobian J, J_1, J_2, J_3, which is the determinant of the transformation of the differentials, and

$$J = J_1 \cdot J_2 \cdot J_3. \tag{3.7}$$

Let us now look at the following transformation:

$$dY = A(dX) \tag{3.8}$$

where Y, A, and X are matrices of compatible dimensions.

By defining dy (respectively, dx) as a vector obtained by putting the columns of the matrix dY (respectively, dX) under one another, (3.8) is transformed into

$$dy = (I \otimes A)\, dx \tag{3.9}$$

where $(A \otimes B)$ is the direct product of the matrices A and B, defined by

$$A \otimes B = \begin{bmatrix} a_{11}B & a_{12}B & \cdots & a_{1n}B \\ \vdots & & & \\ a_{n1}B & \cdots & & a_{nn}B \end{bmatrix}.$$

In a similar way, a transformation of the type

$$dY = (dX)B \tag{3.10}$$

leads to

$$dy = (B^T \otimes I)\, dx. \tag{3.11}$$

If A and B are square matrices of orders r and s, then the following is a well-known property of the direct product

$$\det(A \otimes B) = (\det A)^r (\det B)^s. \tag{3.12}$$

Using this, it is now easy to compute the value of the Jacobians in (3.6) as follows: by (3.4d), (3.9), and (3.12), $J_1 = \det(T)^{-n}$; by (3.5) and the fact that dW does not depend on dF, dK, dH, $J_2 = 1$; by (3.4), (3.9), (3.11), and (3.12), $J_3 = \det(T)^n$ and hence $J = J_1 J_2 J_3 = 1$, which completes the proof.

The proof of Lemma 3.1 assumes implicitly that the parameters of F, K, H are free, and that the elements of T are differentiable functions of F, K, and H. Now in order to apply Lemma 3.1 to our problem, we must make sure that the lemma is valid also for the case where the matrices F, K, H and F^*, K^*, H^* are restricted to the set of admissible structures, i.e., where certain elements of these matrices are constrained to be 0 or 1. It is shown in the Appendix that Lemma 3.1 does apply to this case.

We can now prove the following theorem.

Theorem 3.1: Given two admissible parameterizations with the same number of parameters for the same process of order n, then the determinants of the information matrices corresponding to these two parameterizations are identical.

Proof: Let θ and θ^* be the vectors of parameters corresponding to the two parameterizations (F, K, H) and (F^*, K^*, H^*). The information matrices M_θ and M_{θ^*} are related by[1]

$$\begin{aligned} M_{\theta^*} &= E_{Y|\theta^*}\left\{\left(\frac{\partial \log p(Y|\theta^*)}{\partial \theta^*}\right)^T \left(\frac{\partial \log p(Y|\theta^*)}{\partial \theta^*}\right)\right\} \\ &= E_{Y|\theta^*}\left\{\left(\frac{\partial \log p(Y|\theta)}{\partial \theta}\frac{\partial \theta}{\partial \theta^*}\right)^T \left(\frac{\partial \log p(Y|\theta)}{\partial \theta}\frac{\partial \theta}{\partial \theta^*}\right)\right\} \\ &= \left(\frac{\partial \theta}{\partial \theta^*}\right)^T E_{Y|\theta}\left\{\left(\frac{\partial \log p(Y|\theta)}{\partial \theta}\right)^T \cdot \left(\frac{\partial \log p(Y|\theta)}{\partial \theta}\right)\right\}\left(\frac{\partial \theta}{\partial \theta^*}\right) \\ &= \left(\frac{\partial \theta}{\partial \theta^*}\right)^T M_\theta \left(\frac{\partial \theta}{\partial \theta^*}\right). \end{aligned}$$

Hence $\det M_{\theta^*} = (\det(\partial\theta/\partial\theta^*))^2 \det M_\theta$. By Lemma 3.1 $\det(\partial\theta/\partial\theta^*) = 1$, and hence $\det M_{\theta^*} = \det M_\theta$.

It follows immediately from Theorem 3.1 and the discussion above that, in the Gaussian case and with a maximum likelihood estimation scheme, any two parameterizations will asymptotically yield the same value for the determinant of the parameter error covariance matrix.

Of course other criteria could be used that might be able to discriminate, even asymptotically, between different admissible structures.

An important consequence of our results of this section is that attention should be directed toward finite data results: some parameterizations might be better conditioned than others if only a finite data record is available. Since theoretical results are difficult to obtain for finite data, we shall now turn to two suboptimal structure estimation methods that are based on the idea of selecting the "most independent components" in the state vector.

IV. A Method Based on the Concept of Complexity

In this section, we sketch a method proposed by Ljung and Rissanen [3], [4] and based on the concept of complexity of a random vector defined by Van Emden [12]. We shall also propose a new iterative procedure that is closely related to that of Ljung and Rissanen.

Complexity is, in fact, a measure of the interaction between the components of a random vector. The more interaction there is, the larger the complexity. Van Emden shows that the complexity can be expressed using the covariance matrix and derives the following expression:

[1] If x is a scalar and θ is a k-vector, then $\partial x/\partial\theta$ denotes the row vector $[\partial x/\partial\theta_1 \cdots \partial x/\partial\theta_k]$.

$$C = -\frac{1}{2}\sum_{i=1}^{n}\log(n\lambda_i) \tag{4.1}$$

where λ_i are the eigenvalues of the covariance matrix of the random vector (provided this covariance matrix has been normalized so that its trace equals 1).

Suppose now that we know the covariance matrix $R_{\hat{Y}}$ of $\hat{Y}_t^N$ and that the order of the process, n, is also known. Then one can compute the complexity of various subvectors of order n of $\hat{Y}_t^N$ because the corresponding $n \times n$ covariance matrices are submatrices of $R_{\hat{Y}}$. The idea proposed in [3] is then to select as the state that subvector of $\hat{Y}_t^N$ that has the smallest complexity among all subvectors of dimension n that obey the conditions 1) and 2) of Section II. The components obtained in this way are called by Ljung and Rissanen the "most independent components" of Y_t^N. The procedure they suggest is as follows.

i) Compute estimates of the predictors $\hat{y}_{t+k|t}$ by first fitting a high-order autoregressive model to the data.

ii) Compute the sample covariance matrix $R_{\hat{Y}}$ from the estimated predictions.

iii) For a given value of n, compute the complexity of various submatrices of $R_{\hat{Y}}$, subject to the constraints that the $p \times p$ upper left submatrix of $R_{\hat{Y}}$ is always included and that the jth row of the matrix $R_{\hat{Y}}$ is chosen only if the $(j-p)$th row is also chosen. (These constraints amount to meeting conditions 1) and 2) of Section II.)

iv) Select the basis for the predictor space that corresponds to the submatrix with smallest complexity.

v) Repeat the procedure for higher order models and take the order that minimizes a criterion such as Akaike's AIC criterion.

A major disadvantage of this method is that one needs to first estimate the covariance matrix of the prediction vector by fitting a high-order autoregressive model to the data and then computing sample predictions. However, the procedure has the following interesting feature: if the parameters of the F matrix are estimated by least squares, one can show (see [13]) that the covariance matrix of the error of the parameter estimates is related to the inverse of the submatrix of $R_{\hat{Y}}$ selected by the procedure of Ljung and Rissanen. This seems to justify using the matrix $R_{\hat{Y}}$ as a starting point for the selection of the basis components, even though this matrix is not directly available. It also suggests minimizing some scalar measure of the inverse of the various submatrices of $R_{\hat{Y}}$ in order to discriminate between the corresponding subvectors of $\hat{Y}_t^N$. With this idea in mind, we suggest the following procedure.

The first two steps are identical to those in Ljung and Rissanen's method.

iii) compute the inverse of the upper left $p \times p$ submatrix of $R_{\hat{Y}}$.

iv) for an order n equal to $p+1$, select all the $(p+1)\times(p+1)$ submatrices of $\hat{R}_Y$, which contain the $p \times p$ upper left submatrix, and such that condition 1) of Section II is also satisfied. Compute the inverses of these submatrices, using the fact that the inverse of a matrix $A^1 = \begin{bmatrix} A & b \\ b^T & \alpha \end{bmatrix}$, where b is a vector and α is a scalar, is given by the following inversion formulas:

$$(A^1)^{-1} = \begin{bmatrix} E & g \\ g^T & \epsilon \end{bmatrix}$$

where

$$E = A^{-1} + A^{-1}b\epsilon b^T A^{-1}$$
$$g = -A^{-1}b\epsilon$$
$$\epsilon = [\alpha - b^T A^{-1} b]^{-1}.$$

(Hence, no other matrix inversion is needed once the $p \times p$ submatrix A^{-1} has been computed.)

v) Select the submatrix for which the trace of the inverse is minimized.

vi) Repeat the last two steps with all $(n+1)\times(n+1)$ submatrices that contain the selected $n \times n$ submatrix and whose additional row and column is chosen so as to satisfy condition 1) of Section II.

vii) A stopping criterion is needed. One criterion that seems to work well in practice is to stop when $J = 1/n\,\mathrm{tr}[(A^1)^{-1}]$ does not decrease anymore. This is justified by the fact that when the order of the submatrices is greater than the actual order of the process, these submatrices will be ill-conditioned. Their inverses will be large and so will be J.

V. A $Q-R$ Factorization Method

In this section, we present a new method of structure identification, based on the assumption that a "good" structure will be one in which the predictors selected to form the basis of the prediction space are most independent. We will first assume, as has been done by Akaike [7], that we have chosen $M \in \boldsymbol{N}$ large enough so that $\boldsymbol{H}(Y_{t-M}^t)$ (i.e., the Hilbert space spanned by the components of $y(t-k)$ for $0 \leqslant k \leqslant M$) is close enough to the space $\boldsymbol{H}(Y_{-\infty}^t)$. This allows us to replace the analysis of dependence of the $\hat{y}_{t+k|t}^j$ by the analysis of dependence of the $\hat{y}_{t+k|t,t-M}^j$, which are defined as the projections of y_{t+k}^j onto the space $\boldsymbol{H}(Y_{t-M}^t)$.

Denote by Y_{t-M}^t and Y_{t+N}^t the vectors

$$Y_{t-M}^t = \begin{bmatrix} y_t \\ y_{t-1} \\ \vdots \\ y_{t-M} \end{bmatrix} \qquad Y_{t+N}^t = \begin{bmatrix} y_t \\ y_{t+1} \\ \vdots \\ y_{t+N} \end{bmatrix} \tag{5.1}$$

and by Σ_{11}, Σ_{12}, and Σ_{22} the covariance and cross-covariance matrices

$$\Sigma_{11} = E\left\{Y_{t+N}^t\left(Y_{t+N}^t\right)^T\right\},$$
$$\Sigma_{12} = E\left\{Y_{t+N}^t\left(Y_{t-M}^t\right)^T\right\}$$
$$\Sigma_{22} = E\left\{Y_{t-M}^t\left(Y_{t-M}^t\right)^T\right\}. \tag{5.2}$$

By the full-rank assumption Σ_{22} is positive definite so that we can have

$$\Sigma_{22}^{-T/2}\Sigma_{22}\Sigma_{22}^{-1/2} = I \tag{5.3}$$

for some nonsingular matrix $\Sigma_{22}^{1/2}$.

Let us perform the following transformation:

$$Z_{t-M}^{t} = \Sigma_{22}^{-T/2} Y_{t-M}^{t} \tag{5.4}$$

so that

$$E\left\{\left(Z_{t-M}^{t}\right)\left(Z_{t-M}^{t}\right)^{T}\right\} = I. \tag{5.5}$$

Now $\hat{y}_{t+k|t,t-M}^{j}$ admits a representation

$$\hat{y}_{t+k|t,t-M}^{j} = \sum_{m=0}^{M}\sum_{l=1}^{p} a(j,k,l,m) z_{t-m}^{l} \tag{5.6}$$

where

$$E\left\{\hat{y}_{t+k|t,t-M}^{j} z_{t-m}^{l}\right\} = a(j,k,l,m) = E\left\{y_{t+k}^{j} z_{t-m}^{l}\right\}. \tag{5.7}$$

Hence, the elements of the $(kp+j)$th row of the matrix $\Sigma_{12}\Sigma_{22}^{-1/2}$ are the coordinates of $\hat{y}_{t+k|t,t-M}^{j}$ in the basis Z_{t-M}^{t}. Therefore, because of (5.5), we can replace the analysis of independence of the components $\hat{y}_{t+k|t,t-M}^{j}$ by the analysis of independence of the rows of $\Sigma_{12}\Sigma_{22}^{-1/2}$. Now, a well-known (and numerically well-conditioned) method to search for the most independent rows of a given matrix is to perform a $Q-R$ factorization by means of Householder transformations (see [14] and [15]) with row interchange.

The method we propose is thus as follows.

1) Compute an upper-triangular square root of Σ_{22}, i.e., an upper-triangular matrix $\Sigma_{22}^{1/2}$ such that $\Sigma_{22}^{T/2}\Sigma_{22}^{1/2} = \Sigma_{22}$ (Cholesky factorization).

2) Compute the product $\Sigma_{12}\Sigma_{22}^{-1/2}$. Note that the first p rows of this product are lower triangular. This saves p steps in the subsequent $Q-R$ factorization procedure.

3) Compute a recursive $Q-R$ factorization of $\Sigma_{12}\Sigma_{22}^{-1/2}$ using Householder transformations, where the triangularization is always performed on the row leading to the largest pivot.

4) Stop when the pivots do not significantly decrease any more.

We briefly illustrate one step of the $Q-R$ factorization.

After k steps of the triangularization procedure, we have the following factorization:

$$S_k^T \Sigma_{12}\Sigma_{22}^{-1/2} P_k = T_k$$

where S_k is the product of k permutation matrices, P_k is the product of k Householder transformation matrices, and T_k is a lower-triangular matrix of the following form:

$$T_k = \left[\begin{array}{cccccc|ccc} t_1 & 0 & \cdots & & 0 & & 0 & & 0 \\ X & t_2 & 0 & \cdots & 0 & & & \cdots & \\ \vdots & & & & \vdots & & \vdots & & \vdots \\ & & & & 0 & & & \cdots & \\ X & \cdots & & X & t_k & & 0 & & 0 \\ \hline & & & T_{21} & & & & T_{22} & \end{array}\right].$$

It is clear from the structure of T_k that the Euclidean norm of the rows in T_{22} are the distances of the last $(N+1)p-k$ rows of T_k to the space spanned by the first k rows. From the properties of Householder transformations, the pivot of the next triangularization step is the Euclidean norm of the selected row in T_{22}. So choosing the row leading to the largest pivot amounts to choosing the row with the largest projection onto the space orthogonal to the span of the first k rows. This is what we call the "most independent" row.

Upon completion of the triangularization procedure, the following factorization is obtained:

$$\Sigma_{12}\Sigma_{22}^{-1/2} = STP^T \tag{5.8}$$

where S is a permutation matrix, T is a lower-triangular matrix (at least for its first n rows), and P is a product of Householder transformation matrices (which implies that $P^TP = PP^T = I$). The permutation matrix S indicates which components of $\hat{Y}_t^N$ are to be chosen in the basis. In order to be consistent with the analysis at the end of Section II, we should also introduce some constraints in the procedure of row selection for the triangularization: the first p rows are to be triangularized, and one can choose the jth row only if the $(j-p)$th row has already been chosen. These constraints are taken in order to meet conditions 1) and 2) of Section II which, we recall, lead to a smaller number of parameters to be estimated.

We should also add that with a few more computations (i.e., the solution of a linear system of equations with triangular matrix) one can easily obtain estimates of the parameters of the F matrix, which can be used as initial values for a maximum likelihood algorithm (see Canuto and Menga [8]).

We now show that our triangularization method can also be related very nicely to Akaike's method [7] which is based on a canonical correlation analysis on the vectors of future and past observations (see Anderson [16]). We shall not go back in detail to this method, but we recall that it amounts to a singular value decomposition of the matrix $\Sigma_{11}^{-T/2}\Sigma_{12}\Sigma_{22}^{-1/2}$ and that the idea of the canonical correlation analysis in this case is to search for independent linear combinations of the vector Y_{t+N}^{t} which are most correlated with independent linear combinations of Y_{t-M}^{t}. Actually, in Akaike's method, this only gives the number of independent components of a subvector of the prediction vector, and hence it leads to the choice of the *first* independent components of the prediction vector in the basis. It is

not possible with this method to discriminate between various independent components and to take the "most independent" ones, and the reason for this is that the canonical correlation analysis uses linear combinations of Y_{t+N}^t rather than single components of this vector. By searching for the *single* components of Y_{t+N}^t that are most correlated with independent linear combinations of the past, one can establish a close relationship between our procedure and the method of Akaike.

In the canonical correlation analysis, one searches in a first step for two vectors α_1 and γ_1 such that $\alpha_1^T \Sigma_{12} \gamma_1$ is maximized subject to $\alpha_1^T \Sigma_{11} \alpha_1 = \gamma_1^T \Sigma_{22} \gamma_1 = 1$. In a second step, two linear combinations $\alpha_2 Y_{t+N}^t$ and $\gamma_2 Y_{t-M}^t$ are sought, which are orthogonal to the first ones, and which have maximum correlation with one another. Now consider the following variations.

In the first step we maximize $a_1(\Sigma_{12})_{j_1}\gamma_1$, where a_1 is a scalar, $(\Sigma_{12})_{j_1}$ is the j_1th row of Σ_{12}, γ_1 is a vector of dimension $(M+1)p$, under the constraints $a_1^2 = 1$, $\gamma_1^T \Sigma_{22} \gamma_1 = 1$. The maximum is taken over the values of a_1, γ_1, and the index j_1. For the second step, we maximize $a_2(\Sigma_{12})_{j_2}\gamma_2$, $j_2 \neq j_1$, under the constraints $a_2^2 = 1$, $\gamma_2^T \Sigma_{22} \gamma_2 = 1$, $\gamma_2^T \Sigma_{22} \gamma_1 = 0$, and so on.

With computations similar to those of the canonical correlation analysis, this amounts to finding two matrices A and Γ such that $A^T A = I$ and A is a permutation matrix (with possibly some changes of sign), $\Gamma^T \Sigma_{22} \Gamma = I$, and

$$A^T \Sigma_{12} \Gamma = T. \tag{5.9}$$

T is a lower-triangular matrix with decreasing pivots. If we compare (5.8) and (5.9), we can identify S with A and $\Sigma_{22}^{-1/2} P$ with Γ, and hence establish the equivalence between the two approaches.

In this section we have proposed a new structure estimation method as an alternative to the method of Rissanen and Ljung. It is based on another heuristic definition of "most independent rows" of a matrix. We do not claim that this new method is superior from a theoretical point of view. However, from a computational point of view, our method has the major advantage that it works with the covariance function of the observation process, R_Y, which can be readily estimated from the data, while the method of Rissanen and Ljung works with the covariance of the predictors, $R_{\hat{Y}}$, which requires that the data be filtered first, using an AR model that has to be identified.

VI. Simulation Results

In this section, we present some simulation results in order to compare the various methods that we have introduced. Data sequences have been generated from the Markovian model as follows:

$$x_k = Ax_{k-1} + Be_k$$
$$y_k = Cx_k.$$

TABLE I
PARAMETER VALUES FOR MODEL 1

$$A_1 = \begin{pmatrix} -0.50 & 4.83 & -0.63 & 0.72 \\ 0.20 & -3.67 & 0.50 & -0.58 \\ -0.22 & -2.42 & -0.19 & -0.36 \\ -1.55 & 22.50 & -2.87 & 3.69 \end{pmatrix} \qquad B_1 = \begin{pmatrix} 3.38 & -3.13 \\ -2.5 & 2.5 \\ -0.69 & 1.56 \\ 15.13 & -14.88 \end{pmatrix}$$

$$C_1 = \begin{pmatrix} -7.0 & 38.0 & 16.0 & 9.0 \\ -4.39 & 85.33 & 37.5 & 17.92 \end{pmatrix}$$

TABLE II
PARAMETER VALUES FOR MODEL 2

$$A_2 = \begin{pmatrix} 0 & 1 & 0 & 0 \\ -0.1 & 0.65 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ -2/3 & -5/3 & -0.25 & 1 \end{pmatrix} \qquad B_2 = \begin{pmatrix} 0 & 2 \\ 0.25 & 0.8 \\ 0 & 0 \\ 1 & 1 \end{pmatrix}$$

$$C_2 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

A number of different models has been simulated, but for reasons of brevity, we restrict ourselves here to two models which will hopefully give a sufficient illustration for the previous section. In both models, y_k is a two-dimensional vector process, x_k is of dimension 4, and e_k is a two-dimensional Gaussian white noise with mean zero and unit covariance matrix. Tables I and II give the values of the matrices A, B, and C for each of the two models.

In the first model, the entries of the matrices have been chosen such that the third row of the Hankel matrix $H_{N,M}$ is nearly in the linear span of the first two rows. Hence, the best structural vector of the process is $h_1 = (1,2,4,6)$.

The second model has been taken from a paper by H. El Sherief and N. K. Sinha [17]; using their own method, they find the structural vector $h_2 = (1,2,3,4)$ for this model.

In the sequel, the method of Ljung and Rissanen will be referred to as the LR method, the variant that we have introduced in Section IV will be called WG1, while the QR factorization procedure will be called method WG2.

Table III shows the results of the three methods applied to the first model. Note that in the LR method, the program computes the complexity of all admissible structures for each different order; method WG1, on the other hand, is recursive: for an order n, only those structural vectors are considered which contain all the rows selected in the optimal $(n-1)$th order structural vector. The stopping criterion J used with WG1 has been defined at the end of Section IV.

All three methods reject the odd rows in the structural vector. (Recall that the first p rows are always chosen following condition 2) of Section II.) The LR method gives no estimate of the order. In the WG1 method, the criterion $J = 1/n \operatorname{tr}(A^1)^{-1}$ is minimum for $n = 4$, while in method WG2, after the triangularization of row 4 (third step of the

TABLE III
METHODS LR, WG1, AND WG2 APPLIED TO MODEL 1

Order	Structural vector	LR method Complexity	WG1 method trace	J	WG2 method selected row	|pivot|
3	(1,2,3)	4.37	4.27			
3	(1,2,4)	3.72 ←	4.03 ←	1.33	1	0.408
					2	0.054
4	(1,2,3,4)	7.49	21.52		4	0.110
4	(1,2,4,6)	5.00 ←	4.14 ←	1.04	6	0.104
4	(1,2,3,5)	6.72			8	0.100
					10	0.105
5	(1,2,3,4,6)	8.71	21.68			
5	(1,2,4,6,8)	7.17 ←	7.01 ←	1.40		
5	(1,2,3,4,5)	9.86				
5	(1,2,3,5,7)	9.08				

TABLE IV
METHODS LR, WG1, AND WG2 APPLIED TO MODEL 2

Order	Structural vector	method LR complexity	method WG1 trace	J	method WG2 selected row	|pivot|
3	(1,2,3)	0.92 ←	2.72			
3	(1,2,4)	1.14	2.28 ←	0.76	1	0.735
					2	0.964
4	(1,2,3,4)	2.06 ←	4.89 ←	1.22	4	0.230
4	(1,2,4,6)	2.68	9.22		3	0.063
4	(1,2,3,5)	4.08			5	0.062
					7	0.063
5	(1,2,3,4,5)	7.17	7266.8			
5	(1,2,3,4,6)	3.76 ←	15.11←	5.04		
5	(1,2,4,6,8)	7.91				
5	(1,2,3,5,7)	6.24				

procedure) the decrease of the pivots is not significant anymore, which suggests an order 3.

Finally, we add that the use of Akaike's canonical correlation method [7] on this model leads to the structural vector $h = (1,2,4)$.

In Table IV we see that the three methods lead to a structural vector $h = (1,2,3,4)$ if the supposed order is 4, but if $n = 3$, the LR method leads to $h = (1,2,3)$ while the other two methods indicate $h = (1,2,4)$. The estimation of the order in WG1 leads to $n = 3$, while visual inspection of the pivots in WG2 leads to $n = 4$. Again, the simulation of Akaike's method leads to $h = (1,2,4)$.

The conclusion we have drawn from our simulations is that in most cases (with results of Table IV being the sole exception) method WG1 seems to give the best estimate of the order of the model. We recall that Akaike's method gives an order estimate, but does not select a "best basis," while the LR method does not estimate the order, but selects a "best basis" within a prescribed order. As for method WG2, a better criterion than the visual inspection of the decrease of the pivots could probably be found.

We believe that all three methods give fairly good results for the determination of the structure, the advantage of the methods WG1 and WG2 being that they provide some estimate of the order as well, which avoids the fitting of too many parameterizations.

VII. CONCLUSION

There are several ways of defining a "best" parameterization for the representation of a stationary finite-dimensional multivariate stochastic process. But obviously a logical method would be to select the structure that minimizes some scalar measure of the information matrix that corresponds to each parameterization. We have shown here that the determinants of all these information matrices are asymptotically equivalent, and that this criterion is therefore unable to discriminate between different structures, at least asymptotically. This does not mean that other criteria might not be able to discriminate, or that some structures might not be better than others when only a finite data record is available. Two heuristic schemes have therefore been analyzed in some detail, and simulation results have been presented.

APPENDIX

The interesting argument in Section III establishes that if F, H, K and F^*, H^*, K^* are related by (3.1), where T is a differentiable function of F, H, K, the Jacobian determinant of the transformation from F, H, K, to F^*, H^*, K^* is unity. However, for the application of this result to the systems studied in Section II we must restrict F, H, K to satisfy the reachability and observability conditions

$$\operatorname{rank}\left[K, FK, F^2K, \cdots\right] = \operatorname{rank}\left[H', F'H', (F')^2H', \cdots\right] = n. \quad \text{(A.1)}$$

In addition, the F, H, K considered in Section II are restricted to the set of "nice" structures (i.e., they contain "zeros" and "ones" in fixed places and only $2np$ free parameters).

This is a very "thin" set in the set of all F, H, K. Consequently, it is not apparent that Theorem 3.1 holds, or in what sense it holds.

Consider the Hankel matrix $\boldsymbol{H}$:

$$\boldsymbol{H} = \begin{bmatrix} H(1) & H(2) & H(3) & \cdots \\ H(2) & H(3) & H(4) & \cdots \\ H(3) & H(4) & H(5) & \cdots \\ \vdots & \vdots & \vdots & \end{bmatrix}$$

or, alternatively, $H_{\infty,\infty}$ [see below (2.5)]. Since the theories for $\boldsymbol{H}$ and $H_{\infty,\infty}$ are essentially the same, we shall, for brevity, henceforth speak only of $\boldsymbol{H}$. Call $M(n)$ the set of all systems F, H, K, satisfying (A.1). Then $\boldsymbol{H}$ is of rank n for a system in $M(n)$ and, as is explained in Section II, we may choose as a basis for the rows of $\boldsymbol{H}$ a set of rows of the form

$$r(u,j); \qquad u = 1,\cdots,n_j;\ j = 1,\cdots,p;\ \Sigma n_j = n \quad \text{(A.2)}$$

where $r(u,j)$ denotes the jth row of the uth block of p rows.

Let $U(\{n_j\})$ be the subset of $M(n)$ for which the rows in (A.2) are linearly independent. Then, for suitable $\alpha_{ij}(u)$,

$$r(n_i+1,i)+\sum_{j=1}^{p}\sum_{u=1}^{n_j}\alpha_{ij}(u)r(u,j)=0, \qquad i=1,\cdots,p. \quad \text{(A.3)}$$

Let $h_{ij}(u)$ be the element in row i, column j of $H(u)$. Then the $2np$ numbers

$$\alpha_{ij}(u), \qquad u=1,\cdots,n_j;$$
$$h_{ij}(u), \qquad u=1,\cdots,n_i; \quad i,j=1,\cdots,p$$

coordinatize $U(\{n_j\})$, i.e., they map that set, in a one-to-one manner, on Euclidean space of dimension $2np$. The union of the $U(\{n_j\})$ is, of course, $M(n)$. If $\alpha^*_{ij}(u)$, $h^*_{ij}(u)$ are the coordinates of the same point in $M(n)$ in the coordinate system corresponding to $U(\{n^*_j\})$, $\Sigma n^*_j = n$, then of course the $\alpha^*_{ij}(n)$, $h^*_{ij}(n)$ are functions of the $\alpha_{ij}(u)$, $h_{ij}(u)$ and indeed, as is well known, they are analytic functions on $U(\{n_j\})\cap U(\{n^*_j\})$. Indeed, the proof given below shows that they are finite rational functions. This shows that $M(n)$ may be topologized as an analytic manifold and each $U(\{n_j\}$ is dense in $M(n)$.

We wish to prove the following theorem.

Theorem A.1: The Jacobian of the transformation from the $\alpha_{ij}(u)$, $h_{ij}(u)$ to the $\alpha^*_{ij}(u)$, $h^*_{ij}(u)$ is unity if each of these sets of functions on $M(n)$ is arranged in dictionary order according to i, then j, then u, with i, j, u increasing from 1 to the maximum possible value.

Of course, the same theorem is true for $H_{\infty,\infty}$. The result is, in a sense, special to these parameterizations. For example, it clearly is not true if one such parameterization is given by $\alpha_{ij}(u)$, $h_{ij}(u)$ and the other by $2\alpha^*_{ij}$, $h^*_{ij}(u)$. It is possible that the theorem may be proved more simply via Lemma 3.1, but that is not obvious.

Proof: We may think of the transformation from the unstarred to the starred quantities as the product of a sequence of transformations in each of which one of the indexes, n_j, is increased by unity, and one is decreased by unity, the others being unchanged. Thus, if $n_1=2$, $n_2=3$, $n_3=4$, $n^*_1=5$, $n^*_2=1$, $n^*_3=3$, we may consider $(2,3,4)\to(3,2,4)\to(4,1,4)\to(5,1,3)$. It is therefore necessary and sufficient to prove the result for the case where $n_j\to n_j+1$, $n_k\to n_k-1$, $n_a=n^*_a$, $a\neq j,k$. It is then clearly necessary and sufficient to prove the result when $n^*_1=n_1+1$, $n^*_2=n_2-1$, $n_j=n^*_j$, $j\neq 1,2$. We must assume that $\alpha_{12}(n_2)\neq 0$, for otherwise from (A.3), $r(n_1+1,1)$ is a linear combination of the $r(u,j)$, other than $r(n_2,2)$, and so the set of rows of $\boldsymbol{H}$ corresponding to the n^*_j is not linearly independent. Conversely, if $\alpha_{12}(n_2)\neq 0$ these rows are linearly independent and the point in $M(n)$ lies in $U(\{n_j\})\cap U(\{n^*_j\})$. Thus

$$r(n_2,2)=-\frac{1}{\alpha_{12}(n_2)}\sum_{j=1}^{p}\sum_{u=1}^{n^*_j}\alpha_{1j}(u)r(u,j),$$

putting $\alpha_{11}(n^*_1)=1$. Also using (A.3) for $i=1$ and (A.4) we may express the $\alpha^*_{ij}(u)$ as follows.

$$\alpha^*_{1j}(u)=\alpha_{1j}(u-1)-\frac{\alpha_{12}(n_2-1)}{\alpha_{12}(n_2)}\alpha_{1j}(u)$$
$$-\sum_{k=2}^{p}\alpha_{1k}(n_k)\left\{\alpha_{kj}(u)-\frac{\alpha_{k2}(n_2)}{\alpha_{12}(n_2)}\alpha_{1j}(u)\right\}$$
$$j=1,\cdots,p, \qquad u=1,\cdots,n^*_j.$$

Here, in addition to $\alpha_{11}(n^*_1)=1$, we have put $\alpha_{i1}(n^*_1)=0$ for $i=2,\cdots,p$ and $\alpha_{1j}(0)=0$ for $j=1,\cdots,p$.

$$\alpha^*_{2j}(u)=\alpha_{1j}(u)/\alpha_{12}(n_2), \qquad j=1,\cdots,p;\ u=1,\cdots,n^*_j.$$

$$\alpha^*_{ij}(u)=\alpha_{ij}(u)-\frac{\alpha_{i2}(n_2)}{\alpha_{12}(n_2)}\alpha_{1j}(u),$$
$$i=3,\cdots,p;\ j=1,\cdots,p;\ u=1,\cdots,n^*_j,$$

with $\alpha_{i1}(n^*_1)=0$, $i=3,\cdots,p$.

To evaluate $|\partial\alpha^*_{ij}(u)/\partial\alpha_{ab}(v)|$ we take i in the order $i=2,3,\cdots,p,1$ and a in the order $1,3,\cdots,p,2$. For i fixed (a fixed) we order the row first according to j increasing (b increasing) and then u increasing (v increasing), except that we take $\alpha^*_{i1}(n^*_1)$ last ($\alpha_{a2}(n_2)$ last) in the set of n rows for each i (in the set of n columns for each a). We shall now show that the Jacobian matrix, $\partial\alpha^*_{ij}(u)/\partial\alpha_{ab}(v)$, is upper triangular and that its determinant is $(-1)^{p+n-1}\alpha_{12}(n_2)^{-p}$.

First,

$$\frac{\partial\alpha^*_{ij}(u)}{\partial\alpha_{ab}(v)}=0, \qquad i>2,\ a\neq i,\ a\neq 1;$$
$$\frac{\partial\alpha^*_{2j}(u)}{\partial\alpha_{ab}(v)}=0, \qquad a\geqslant 2.$$

Thus all blocks below the main diagonal are null and we need only evaluate the determinants of the diagonal blocks. Now

$$\frac{\partial\alpha^*_{2j}(u)}{\partial\alpha_{1b}(v)}=0, \qquad \text{unless } b=j,\ v=u \text{ or } b=2,\ v=n_2.$$

Thus the top left-hand block is itself upper triangular. Since

$$\frac{\partial\alpha^*_{2j}(u)}{\partial\alpha_{1j}(u)}=\alpha_{12}(n_2)^{-1}, \qquad \frac{\partial\alpha^*_{21}(n^*_1)}{\partial\alpha_{12}(n_2)}=-\{\alpha_{12}(n_2)\}^{-2}$$

this upper left-hand block has determinant $-\{\alpha_{12}(n_2)\}^{-n-1}$. The blocks $\partial\alpha^*_{ij}(u)/\partial\alpha_{ib}(v)$, $i=3,4,\cdots,p$, are again upper triangular with diagonal elements unity save for $\partial\alpha^*_{i1}(n^*_1)/\partial\alpha_{i2}(n_2)=-\{\alpha_{12}(n_2)\}^{-1}$. Thus the product of the determinants of these $p-2$ diagonal blocks is $\{-\alpha_{12}(n_2)\}^{-p+2}$. Finally, $\partial\alpha^*_{1j}(u)/\partial\alpha_{2b}(v)$ is also upper triangular. Indeed for $j=1$, $u=n^*_1$, $b=2$, $v=n_2$ the derivative is unity while otherwise

$$\frac{\partial\alpha^*_{1j}(u)}{\partial\alpha_{2b}(v)}=\delta_{jb}\delta_{uv}+\delta_{2b}\delta_{n_2v}\alpha_{1j}(u)$$

so that the determinant is $\{\alpha_{12}(n_2)\}^{n-1}$. Thus $|\partial\alpha^*_{ij}(u)/$

$\partial\alpha_{ab}(v)|$ is $(-1)^{p+n-1}\alpha_{12}(n_2)^{-p}$, as stated above.

Since $\partial\alpha^*_{ij}(u)/\partial h_{ab}(v) \equiv 0$, we need only evaluate $\partial h^*_{ij}(u)/\partial h_{ab}(v)$. This is $\delta_{ia}\delta_{jb}\delta_{uv}$ unless $i=1, u=n^*_1, a=2, v=n_2$ when we obtain $-\delta_{jb}\alpha_{12}(n_2)$. Thus the determinant is $\{-\alpha_{12}(n_2)\}^p$ and the Jacobian determinant required for the theorem is, in the ordering of rows and columns used in this proof, $(-1)^{n-1}$. Reordering rows and columns in the natural order stated in the theorem we obtain the required result.

In order that (2.8) should represent a stationary process with e_t the innovation sequence, then it is necessary and sufficient that $k(z) = H(zI_n - F)^{-1}K + I_p$ should be analytic for $|z| \geq 1$ and $\det k(z) \neq 0$, $|z| > 1$. If $\det k(z) \neq 0$, $|z| \geq 1$, then the resulting set of F, H, K is an open submanifold of $M(n)$ so that precisely the same result holds.

Although the Jacobian determinant of the transformation is unity, it is evident that the smallest and largest eigenvalues of the Jacobian matrix approach 0 and ∞, respectively, as the edge of a coordinate neighborhood is approached. If a stationary random process is parameterized by points in $M(n)$, then one might seek to check the appropriateness of a coordinate system through the information matrix.

Although its determinant is independent of the partition, $n = \Sigma n_j$, used to define a coordinate system, the eigenvalues will not be. It might be computationally difficult to determine these eigenvalues, but the trace is easy to compute. The determinant will be useful in checking whether n is too large as the information matrix will be singular when this is so.

REFERENCES

[1] M. Hazewinkel and R. E. Kalman, "On invariants, canonical forms and moduli for linear constant finite dimensional dynamical systems," in *Proc. CNR CISM Symp. Algebraic Syst. Theory*, Udine, 1975.

[2] M. Hazewinkel, "Moduli and canonical forms for linear dynamical systems, II: The topological case," *Math. Syst. Theory*, vol. 10, no. 4, 1976.

[3] L. Ljung and J. Rissanen, "On canonical forms, parameter identifiability and the concept of complexity," in *Proc. IFAC Symp. Identification*, Tbilisi, U.S.S.R., 1976, vol. 3.

[4] J. Rissanen and L. Ljung, "Estimation of optimum structures and parameters for linear systems," in *Proc. CNR CISM Symp. Algebraic Syst. Theory*, Udine, 1975.

[5] A. J. M. Van Overbeek and L. Ljung, "On line structure selection for multivariable state space models," in *Proc. IFAC Symp. Identification Parameter Est.*, Darmstadt, Germany, 1979, vol. 1.

[6] J. Rissanen, "Minimax entropy estimation of models for vector processes," in *System Identification: Advances and Case Studies*, R. Mehra and D. Lainiotis, Eds. New York: Academic, 1976.

[7] H. Akaike, "Canonical correlation analysis of time series and the use of an information criterion," in *System Identification: Advances and Case Studies*, R. Mehra and D. Lainiotis, Eds. New York: Academic, 1976.

[8] E. Canuto and G. Menga, "On line identification of multivariable linear stochastic systems," in *Proc. IFAC Symp. Discontinuous Comput. Contr. Syst.*, Prague, Czechoslovakia, 1977.

[9] K. Glover and J. C. Willems, "Parameterizations of linear dynamical systems: Canonical forms and identifiability," *IEEE Trans. Automat. Contr.*, vol. AC-19, Dec. 1974.

[10] J. Rissanen, "Basis of invariants and canonical forms for linear dynamic systems," *Automatica*, vol. 10, 1974.

[11] L. M. Silverman, "Realization of linear dynamical systems," *IEEE Trans. Automat. Contr.*, vol. AC-16, Dec. 1971.

[12] M. Van Emden, "Analysis of complexity," *Math. Cent. Tracts*, vol. 35, Amsterdam, The Netherlands, 1971.

[13] V. Wertz, "Structure selection for the identification of multivariate processes," Ph. D. dissertation, Louvain Univ., Louvain-la-Neuve, Belgium, May 1982.

[14] G. H. Golub, V. Klema, and G. W. Stewart, "Rank degeneracy and least squares problems," Stanford Univ., Stanford, CA, Tech. Rep. Stan. CS76.559, Aug. 1976.

[15] G. H. Golub and G. P. H. Styan, "Numerical computations for univariate linear models," *J. Stat. Comput. Simulation*, vol. 2, pp. 253–274, 1973.

[16] T. W. Anderson, *An Introduction to Multivariate Statistical Analysis*. New York: Wiley, 1958.

[17] M. El Sherief and N. K. Sinha, "Determination of the structure of a canonical model for the identification of linear multivariable systems," in *Proc. IFAC Symp. Identification Parameter Est.*, Darmstadt, Germany, Sept. 1979, vol. 1.

[18] G. Picci, "Some numerical aspects of multivariable systems identification," in *Proc. Workshop Numer. Methods Syst. Eng. Problems*, Lexington, KY, June 1980.

[19] M. Pavon, "Stochastic realization and invariant directions of the matrix Riccati equation," *SIAM J. Contr. Optimiz.*, vol. 18, pp. 155–180, Mar. 1980.

[20] E. J. Hannan, "The statistical theory of linear systems," in *Developments in Statistics, vol. 2*, P. Krishnaiah, Ed. New York: Academic, 1979, pp. 83–121.

[21] M. Gevers and V. Wertz, "On the problem of structure selection for the identification of stationary stochastic processes," in *Proc. IFAC Symp. Identification Parameter Est.*, Washington, DC, June 1982.

Vincent Wertz was born in Liège, Belgium, in 1955. He received the Engineering degree in applied mathematics and the Ph. D. degree from Louvain University, Louvain-la-Neuve, Belgium, in 1978 and 1982, respectively.

Since September 1978, he has been a Research Assistant at the Laboratoire d'Automatique et d'Analyse des Systèmes, Louvain University. From July 1980 to January 1981, he was on leave at the Department of Electrical Engineering, University of Newcastle, N.S.W., Australia, where he was sponsored by a fellowship of IRSIA (Institut pour la Recherche Scientifique dans l'Industrie et l'Agriculture). From February 1981 to July 1981, he was an Assistant at the Centre Universitaire de Stif, Algeria. His main interests are in multivariable systems and in system identification. His Ph.D. dissertation will be on the subject of parameterization and identifiability of multivariable systems.

Michel Gevers (S'66–S'70–M'72) was born in Antwerp, Belgium, in 1945. He received the electrical engineering degree from Louvain University, Louvain-la-Neuve, Belgium, in 1968, and the Ph.D. degree from Stanford University, Stanford, CA, in 1972.

He went to Stanford University, where he was supported by a Harkness fellowship and then an ESRO/NASA fellowship, in 1969, following a one-year research assistantship in the Solid State Laboratory, Louvain University. Since 1972 he has been Assistant Professor and, subsequently, Professor at the Laboratoire d'Automatique et d'Analyse des Systèmes at Louvain University. He was head of the laboratory from 1976 to 1980. In 1980 he was on sabbatical leave at the University of Newcastle, N.S.W., Australia. His main research interests are in estimation, identification, stochastic processes, and multivariable system theory. He has done applied work on hydrological, biomedical, and industrial problems. He is the author or coauthor of about forty papers and conference papers.

Edward J. Hannan was born in Melbourne, Australia, in 1921. He received the Bachelor's degree in commerce (business administration) in 1949.

He received his early education in Melbourne. He commenced a five-year period of Army service in an infantry battalion in 1940. In early 1946 he commenced his tertiary education under an assistance scheme for ex-servicemen. After some introduction to elementary mathematics and statistics during his degree, he continued to study mathematics while he worked as an economist with what is now the Reserve Bank of Australia. In 1953 he was sent by this institution to study

for one year at the newly formed Australian National University, Canberra. He has remained at ANU since then and has worked almost entirely in the field of time series analysis, in which he has published two books. His interest in problems of a geophysical nature has arisen from consulting with geophysicists and oceanographers. His most recent work has predominantly been in the statistical theory of linear systems.

Regulator Synthesis Using (C, A, B)-Pairs

J. M. HANS SCHUMACHER, MEMBER, IEEE

Abstract—**The concept of a (C, A, B)-pair of subspaces was introduced in an earlier paper by the author to study problems of compensator design. Here, it is shown that the same concept plays a basic role in the theory of tracking and regulation. From a few fundamental properties, the main results of regulator theory are derived in a relatively simple and straightforward manner. Another new feature is that we are able to allow the presence of an unknown input, thereby unifying the fields of "tracking and regulation" and "disturbance decoupling." All results are fully constructive, and an example is given to illustrate the theory.**

I. Introduction

IT IS often true that insight in the practical handling of a computational problem depends crucially on which theoretical concepts are used to describe the basic question of solvability. Thus, the invertibility of a matrix can be described in terms of its determinant or in terms of its singular values, but from a numerical point of view the first standpoint is much less significant than the second. It is always desirable to have at least a clear conceptual framework available.

In this paper, we consider servo and regulator problems for linear multivariable systems. The basic solvability question has been studied extensively during the past decade, mostly within the state-space framework that we shall also use; see, for instance, [1]–[9]. ([9] will be our main reference.) The purpose of the present paper is to put forward a new concept which can be used to obtain the solution of the regulator problem in a direct and transparent way. Our method is to be compared with the procedure of [9], which proceeds stepwise from "restricted" problems (using state feedback with certain restrictions) to "extended" problems (using dynamic output feedback). Of course, the "extended" version is the form in which the problem has to be solved ultimately, and the main benefit of our approach is that it considers output feedback directly, without any intermediate steps. The key concept is that of a "(C, A, B)-pair of subspaces," which relates to dynamic output feedback in roughly the same way as the concept of "(A, B)-invariant subspace" relates to state feedback. The notion of a (C, A, B)-pair has been used before by the author [14] to study stabilization by dynamic output feedback.

An interesting feature of our method, which acts as a further corroboration of its naturalness, is that it is able to bring the disturbance decoupling aspect into the regulator problem. Up to now, the topics of "tracking and regulation" and "disturbance decoupling" have been treated separately; the solution of the latter problem involving dynamic output feedback is quite recent [10]–[13] and it has, in fact, been the origin of the notion "(C, A, B)-pair" (see [14]). We give a unified formulation from which both the tracking and regulation problems and the disturbance decoupling problems can be recovered as special cases. This general formulation has an appealing form: we consider a system which has two kinds of inputs, one that is at the controller's disposal and one that is controlled by "nature," and which also has two kinds of outputs, of which one describes the goal of the control and the other gives the available information.

For the general problems we shall pose, an algorithm will be given which checks their solvability, and we shall also give a method to construct a solution if one exists. While our treatment of these problems is thus completely constructive, it is not claimed that the presented algorithms are optimal with respect to numerical behavior. However, coming back to the point made at the beginning of this paper, it is believed that the conceptual clarification will be helpful to improve practical solution methods.

We lay down some notation and terminology in Section II where we also give the definition of a "(C, A, B)-pair." The next two sections give some basic results involving (C, A, B)-pairs; these are Theorem 3.1, which will be needed for the "necessity" part of our later results, and Theorem 4.1 which will be used for the constructive side. We shall pose the generalized regulator problem in two different versions, depending on the stability requirements. In Section V we consider "output stability," and "maximal stability" is treated in Section VI. Complete solutions of both problems are given, together with specializations to

Manuscript received December 17, 1980; revised September 2, 1981. Paper recommended by E. W. Kamen, Chairman of the Linear Systems Committee.

The author is with the Department of Mathematics, Erasmus Universiteit, Rotterdam, The Netherlands.

0018-9286/82/1200-1211$00.75