Pergamon

PII:S0005-1098(96)00021-0

Automatica, Vol. 32, No. 6, pp. 875–883, 1996

0005-1098/96 $15.00+0.00

Test for Local Structural Identifiability of High-order Non-linearly Parametrized State Space Models

HANS G. M. DÖTSCH† and PAUL M. J. VAN DEN HOF†

Structural identifiability of high-order, physically parametrized, and possibly unstable state space models can be assessed by rank evaluation of an information matrix that can be computed based on simple analytical expressions.

Key Words—System identification; physical models; identifiability; state space models; gradients.

Abstract—In this paper the issue of local structural identifiability of high-order state space models with non-linear parametrizations is addressed. Two methods are presented that provide analytical expressions for information matrices of which the rank determines identifiability. The first method requires solving a Lyapunov equation of high dimension and is applicable only to stable models. The second method applies also to unstable models and is based on a gradient computation algorithm by dynamic programming, that normally is used in an identification framework. Copyright ©1996 Elsevier Science Ltd.

1. INTRODUCTION

The issue addressed in this paper concerns identifiability of linear, time invariant model structures where the parametrization is based on, for example, physical a priori knowledge of the process. If the physically interpretable parameters are to be determined using system identification, identifiability of the parameters is imperative for two reasons.

Firstly, if different parameter values correspond to the same input–output behavior, the parameters do not provide a unique physical description of the process. Secondly, a non-unique model parametrization might result in an ill-posed identification problem. For these reasons investigation of the existence of a one-to-one relation between the input–output behavior of the model and the physical parameters is a necessary exercise before the actual parameter estimation procedure is carried out. The problem of investigating this relation is referred to as structural identifiability of a model parametrization and is the subject of this paper.

In literature the topic of structural identifiability has received an extensive amount of attention. Bellman and Åström (1970) stated the concept; the subject has been thoroughly studied in the field of biological and biochemical modeling (Rubinow and Winzer, 1971; Cobelli *et al.*, 1979; Norton, 1980; Godfrey, 1983).

State space model parametrizations are analyzed by Glover and Willems (1974), Grewal and Glover (1976) and Walter (1981). The problem has been treated in a stochastic framework by Rothenberg (1971) and Tse (1973). Structural identifiability of models in terms of differential-algebraic expressions has been investigated by Ljung and Glad (1991). Dasgupta and Anderson (1987) consider the situation of processes with transfer functions that are multilinear in the unknown parameters.

We focus on the case that available a priori knowledge leads to formulation of a linear, time invariant high-order state space model where the model parameters are non-linear expressions of the physically interpretable parameters. Investigation of local structural identifiability using existing techniques in this case has its limitations. On one hand methods proposed in the field of biological modeling are restricted to a specific type of parametrization, i.e. compartmental model parametrizations (Godfrey, 1983; Norton, 1980). On the other hand the calculation effort required by methods based on non-linearly parametrized state space models increases drastically or may even become impossible in case of increasing model order (Grewal and

Received 7 November 1994; revised 24 July 1995; received in final form 8 January 1996. The original version of this paper was presented at the 10th IFAC Symposium on System Identification, which was held in Copenhagen, Denmark, during 4–6 July 1994. The Published Proceedings of this IFAC meeting may be ordered from Elsevier Science Limited, The Boulevard, Langford Lane, Kidlington, Oxford OX5 1GB, U.K. This paper was recommended for publication in revised form by Associate Editor J. Bokor under the direction of Editor Torsten Söderström. Corresponding author Dr P. M. J. Van den Hof. Tel. +31 15 278 4509; Fax +31 15 278 4717; E-mail p.m.j.vandenhof@wbmt.tudelft.nl.

† Mechanical Engineering Systems and Control Group, Delft University of Technology, Mekelweg 2, 2628 CD Delft, The Netherlands.

Glover, 1976; Glover and Willems, 1974).

In this paper two methods are proposed for analysis of local structural identifiability of non-linearly parametrized, high-order state space model representations. Both methods are based on rank evaluation of a Jacobian matrix, being similar to a rank test on the information matrix as formulated by Rothenberg (1971) and Tse (1973). A problem in their work still is the computation of the information matrix in the related identification problem.

We will present two different methods where the required rank test is formulated on information matrices related to an infinite-time and a finite-time identification criterion. The first method provides a compact solution to the problem but requires solving a Lyapunov equation of (very) high dimension. The method only holds for stable models which may pose problems in case the physics of the process dictate, for instance, an integrator. An alternative method provides an extension to unstable systems. Here, calculation of an analytical expression for the (finite-time) information matrix is enabled through use of an algorithm for gradient computation, based on a dynamic programming formulation and normally used in a prediction-error identification framework. Both methods require analytical expressions for the partial derivatives of the state space matrices with respect to the argument parameters and show limited complexity.

In our notation we will consider the discrete-time case, however, it will be explained a posteriori that both methods are also applicable to continuous-time state space realizations. Firstly we will briefly review the local structural identifiability problem. The approach based on calculation of the infinite-time information matrix is presented in Section 3. Next, the computational aspects of local structural analysis based on gradient computation are elaborated in Section 4. A simulation example is treated in Section 5.

2. LOCAL STRUCTURAL IDENTIFIABILITY

Consider a linear, time invariant single-input, single-output discrete-time state space model structure, parametrized in θ:

$$x(k+1) = A(\theta)x(k) + b(\theta)u(k), \qquad (1)$$

$$y(k) = c(\theta)x(k), \qquad (2)$$

where $x(k) \in \mathbb{R}^n$, $u(k) \in \mathbb{R}$, $y(k) \in \mathbb{R}$ and $\theta \in \Theta \subset \mathbb{R}^q$. Without loss of generality the direct feedthrough $d(\theta)$ is assumed to be 0. The elements of $A(\theta)$, $b(\theta)$, $c(\theta)$ are supposed to be algebraic functions of the elements of θ.

For a definition of local structural identifiability we adopt the formulation by Glover and Willems (1974):

Definition 1. A model structure (1), (2) is said to be locally identifiable from the transfer function at the point $\hat{\theta} \in \Theta$ if there exists $\epsilon > 0$ such that the following conditions:

(1) $\|\theta_1 - \hat{\theta}\| < \epsilon$, $\|\theta_2 - \hat{\theta}\| < \epsilon$; and

(2) $c(\theta_1)(zI - A(\theta_1))^{-1}b(\theta_1) = c(\theta_2)(zI - A(\theta_2))^{-1}b(\theta_2)$
for all $z \in \mathbb{C}\backslash\{\lambda(A(\theta_1)), \lambda(A(\theta_2))\}$ where $\lambda(\cdot)$ indicates the set of eigenvalues of a matrix;

imply that $\theta_1 = \theta_2$. □

In words, in the neighborhood of $\hat{\theta}$ there are no two models with distinct parameters which have the same transfer function. In Definition 1 equality of the transfer functions is related to equality of the parameters. To mould this in a mathematical framework the following lemma on injective maps is presented (Glover and Willems, 1974):

Lemma 1. Let Ω be an open set in $\mathbb{R}^n$ and $f : \Omega \to \mathbb{R}^m$ be a k-times continuously differentiable map with $k \geq 1$. If $\partial f(x)/\partial x$ has constant rank r in a neighborhood of $\hat{x}$, then f is locally injective at $\hat{x}$ if and only if $r = n$.

Now Definition 1 together with Lemma 1 lead to the following proposition (see also Grewal and Glover, 1976; Norton, 1980).

Proposition 1. Consider the map $S_m : \Theta \subset \mathbb{R}^q \to \mathbb{R}^m$ defined by:

$$S_m(\theta) := [h(1,\theta)\; h(2,\theta)\; \ldots\; h(m,\theta)]^{\mathrm{T}}, \qquad (3)$$

where $h(k,\theta) = c(\theta)A^{k-1}(\theta)b(\theta)$ $(k = 1, 2, \ldots, m)$ are the first m Markov parameters of the model. Then, the model structure (1), (2) is locally structural identifiable in $\theta = \theta_0$ if for any $m \geq 2n$, $\mathrm{rank}(\partial S_m/\partial\theta) = q$ in $\theta = \theta_0$. □

So, analysis of local structural identifiability of (1), (2) amounts to evaluation of the rank of the Jacobian matrix

$$\left.\frac{\partial S_m(\theta)}{\partial \theta}\right|_{\theta=\theta_0} = \begin{bmatrix} \frac{\partial h(1)}{\partial \theta_1} & \frac{\partial h(1)}{\partial \theta_2} & \cdots & \frac{\partial h(1)}{\partial \theta_q} \\ \frac{\partial h(2)}{\partial \theta_1} & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ \vdots & & & \vdots \\ \frac{\partial h(m)}{\partial \theta_1} & \cdot & \cdots & \frac{\partial h(m)}{\partial \theta_q} \end{bmatrix}_{\theta=\theta_0} \qquad (4)$$

with dimensions $m \times q$, while $m \geq 2n$. A test on the rank of (4) is equivalent to a rank test on the matrix

$$\left(\left(\frac{\partial S_m(\theta)}{\partial \theta}\right)^{\mathrm{T}} \frac{\partial S_m(\theta)}{\partial \theta}\right)\Bigg|_{\theta=\theta_0}, \qquad (5)$$

and so local structural identifiability can be determined by a rank test on (5). With (slight) abuse of terminology we will refer to (5) as an information matrix, as it can be shown to be equal to the Fisher information matrix for specifically formulated identification problems. In the next two sections analytical expressions for two information matrices are presented for the situations of $m \to \infty$ and for $m = 2n$.

3. IDENTIFIABILITY ANALYSIS BASED ON INFINITE-TIME INFORMATION MATRIX

In this section we will present a result that is based on the evaluation of the separate elements in the infinite-time information matrix, of which element (i, j) is determined by

$$\left[\left(\frac{\partial S_\infty(\theta)}{\partial \theta}\right)^{\mathrm{T}} \frac{\partial S_\infty(\theta)}{\partial \theta}\right]_{ij} = \sum_{k=1}^{\infty} \frac{\partial h(k)}{\partial \theta_i} \frac{\partial h(k)}{\partial \theta_j}. \qquad (6)$$

Following an approach that is taken by Peeters (1994), for any dynamical system in state space form (1), (2) the derivative of the output with respect to θ_i can be written as:

$$\begin{bmatrix} x(k+1) \\ x^{(i)}(k+1) \end{bmatrix} = \begin{bmatrix} A & 0 \\ A^{(i)} & A \end{bmatrix} \begin{bmatrix} x(k) \\ x^{(i)}(k) \end{bmatrix} + \begin{bmatrix} b \\ b^{(i)} \end{bmatrix} u(k) \qquad (7)$$

$$y^{(i)}(k) = \begin{bmatrix} c^{(i)} & c \end{bmatrix} \begin{bmatrix} x(k) \\ x^{(i)}(k) \end{bmatrix}, \qquad (8)$$

where the notation $(\cdot)^{(i)}$ refers to the derivative $\partial/\partial\theta_i$. This expression was employed by Peeters (1994) for finding explicit expressions for an information matrix in a closely related situation. We can now employ the following proposition.

Proposition 2. Consider a state space model structure (1), (2) with u a stationary stochastic white noise process. If $A(\theta_0)$ has all eigenvalues within the open unit disc, then

$$E \frac{\partial y(t)}{\partial \theta_i} \frac{\partial y(t)}{\partial \theta_j}\Bigg|_{\theta=\theta_0} = \sum_{k=1}^{\infty} \frac{\partial h(k)}{\partial \theta_i} \frac{\partial h(k)}{\partial \theta_j}\Bigg|_{\theta=\theta_0}. \qquad (9)$$

Proof. The result follows straightforwardly by writing $\frac{\partial y(t)}{\partial \theta} = \frac{\partial h(1)}{\partial \theta} u(t-1) + \frac{\partial h(2)}{\partial \theta} u(t-2) + \cdots$ and employing the fact that u is a white noise process. Stability of (7), (8) is required to ensure boundedness of (9).

■

For convergence of the right-hand side of (9), this approach is confined to stable state space models. The expression in the proposition is equal to the (i, j)-element of the infinite-time information matrix, as shown in (6). Building on the state space expressions (7), (8) we can now write

$$\begin{bmatrix} x(k+1) \\ x^{(1)}(k+1) \\ x^{(2)}(k+1) \\ \vdots \\ x^{(q)}(k+1) \end{bmatrix} = \begin{pmatrix} A & 0 & 0 & \cdots & 0 \\ A^{(1)} & A & 0 & \cdots & 0 \\ A^{(2)} & 0 & A & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ A^{(q)} & 0 & 0 & \cdots & A \end{pmatrix} \begin{bmatrix} x(k) \\ x^{(1)}(k) \\ x^{(2)}(k) \\ \vdots \\ x^{(q)}(k) \end{bmatrix} + \begin{bmatrix} b \\ b^{(1)} \\ b^{(2)} \\ \vdots \\ b^{(q)} \end{bmatrix} u(k) \qquad (10)$$

$$\begin{pmatrix} y^{(1)}(k) \\ y^{(2)}(k) \\ \vdots \\ y^{(q)}(k) \end{pmatrix} = \begin{pmatrix} c^{(1)} & c & 0 & \cdots & 0 \\ c^{(2)} & 0 & c & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ c^{(q)} & 0 & 0 & \cdots & c \end{pmatrix} \times \begin{bmatrix} x(k) \\ x^{(1)}(k) \\ x^{(2)}(k) \\ \vdots \\ x^{(q)}(k) \end{bmatrix}. \qquad (11)$$

This leads to the following theorem.

Theorem 1. Let $A(\theta_0)$ have all eigenvalues in the open unit disc. Then

$$\left(\frac{\partial S_\infty(\theta)}{\partial \theta}\right)^{\mathrm{T}} \frac{\partial S_\infty(\theta)}{\partial \theta}\Bigg|_{\theta=\theta_0} = C_x P C_x^{\mathrm{T}} \qquad (12)$$

with P the solution to the Lyapunov equation

$$P = A_x P A_x^{\mathrm{T}} + B_x B_x^{\mathrm{T}}, \qquad (13)$$

where A_x, B_x and C_x are the respective matrices in the state space representation (10), (11) for $\theta = \theta_0$.
□

Proof. It follows from Proposition 2 that the considered information matrix is given by the covariance matrix of the output process (11). This covariance can directly be calculated by using the state covariance P.

■

Local identifiability of the model structure can now be tested in $\theta = \theta_0$ by testing the rank of (12).

Note that for solving (13) a Lyapunov equation has to be solved having dimension $(q+1)n$. This comes down to the calculation of the controllability Gramian of the state space model (A_x, B_x, C_x). Especially for high-order state space models (large n), the dimension of this problem can become very large.

To solve the equation (13) accurately, the specific structural properties of the matrix A_x should be efficiently used. A further analysis of this problem is outside the scope of this paper.

With respect to the high-dimensionality of the concerned problem, it has to be stressed that in the approach of Glover and Willems (1974), a rank test has to be performed on a square matrix of dimension n^2+q, which in general will be essentially larger than $(q+1)n$.

Remark 1. For the method presented in this section there exists an equivalent formulation for a continuous-time model structure. In that case the discrete-time Lyapunov equation (13) is naturally replaced by the continuous-time Lyapunov equation $A_xP + PA_x^{\mathrm{T}} = B_xB_x^{\mathrm{T}}$.

4. IDENTIFIABILITY ANALYSIS BASED ON FINITE-TIME INFORMATION MATRIX

As stated in the previous section, calculation of an infinite-time information matrix only holds for stable state space models. As will be shown in the example in Section 5, there are many processes that are marginally stable, e.g. when their dynamic behavior contains an integrator.

In this section it is shown that the problem of evaluating the rank of (4) for finite values of m enables identifiability analysis of unstable models and can be solved by adequately rephrasing the identifiability problem and applying an algorithm for gradient computation.

4.1. *Rephrasing the identifiability problem*

Consider the following "thought" identification problem, i.e. the identification is not actually carried out. Let there be given a linear, time invariant system that generates input–output data $\{u(k), y(k)\}_{k=0,\cdots,N}$, determined by

$$z(k+1) = A(\theta_0)z(k) + b(\theta_0)u(k), \tag{14}$$

$$y(k) = c(\theta_0)z(k) + e(k), \tag{15}$$

where state $z(k) \in \mathbb{R}^n$, input $u(k) \in \mathbb{R}$, output $y(k) \in \mathbb{R}$ and $\{e(k)\}$ a sequence of independent, identically distributed random variables (white noise) with unit variance. When identifying a model on the basis of measurement data obtained from this data generating system, we consider a state space model structure

$$\begin{aligned} x(k+1) &= A(\theta)x(k) + b(\theta)u(k), \\ \varepsilon(k,\theta) &= y(k) - c(\theta)x(k), \end{aligned} \tag{16}$$

where $A(\theta), b(\theta), c(\theta)$ is a parametrized state space model with $\theta \in \Theta \subset \mathbb{R}^q$ and $\varepsilon(k,\theta)$ the one-step-ahead prediction error (Ljung, 1987). The corresponding least squares identification criterion is given by *

$$V_N(\theta) = \frac{1}{2}\sum_{k=1}^{N} \varepsilon^2(k,\theta). \tag{17}$$

To this end the following proposition will be fruitful.

Proposition 3. Consider a parametrized model structure (16), and any $\theta_0 \in \Theta$. Let a data generating system $(A(\theta_0), b(\theta_0), c(\theta_0))$ generate input–output data according to (14), (15) with $u(k)$ a unit pulse signal, i.e. $u(k) = 1, k = 0$; $u(k) = 0, k > 0$. Then

$$E\left(\frac{\partial V_N(\theta)}{\partial\theta}\right)\left(\frac{\partial V_N(\theta)}{\partial\theta}\right)^{\mathrm{T}}\Bigg|_{\theta=\theta_0} = \left(\frac{\partial S_N(\theta)}{\partial\theta}\right)^{\mathrm{T}}\frac{\partial S_N(\theta)}{\partial\theta}\Bigg|_{\theta=\theta_0} \tag{18}$$

with $V_N(\theta)$ and $\frac{\partial S_N(\theta)}{\partial\theta}$ as expressed in (17), (4), and E the expectation operator. □

Proof. Under pulse input conditions, the prediction error satisfies

$$\varepsilon(k,\theta) = y(k) - h(k,\theta), \tag{19}$$

where $h(k,\theta)$ represents the k-th element of the pulse response of the model. The gradient $\partial V_N(\theta)/\partial\theta$ can now be expressed as

$$\frac{\partial V_N(\theta)}{\partial\theta} = -\sum_{k=1}^{N}\varepsilon(k,\theta)\frac{\partial h(k,\theta)}{\partial\theta}. \tag{20}$$

Consequently

$$\begin{aligned} &E\,\frac{\partial V_N}{\partial\theta}\left(\frac{\partial V_N}{\partial\theta}\right)^{\mathrm{T}}\Bigg|_{\theta=\theta_0} \\ &= \sum_{k=1}^{N}\sum_{l=1}^{N}\frac{\partial h(k,\theta_0)}{\partial\theta}E\varepsilon(k,\theta_0)\varepsilon(l,\theta_0)\left(\frac{\partial h(l,\theta_0)}{\partial\theta}\right)^{\mathrm{T}}. \\ &= \sum_{k=1}^{N}\frac{\partial h(k,\theta_0)}{\partial\theta}\left(\frac{\partial h(k,\theta_0)}{\partial\theta}\right)^{\mathrm{T}}, \end{aligned} \tag{21}$$

where the last equality follows from the unit-variance white noise property of $\varepsilon(k,\theta_0)$.

* For reasons that will become clear later on, we will not consider a factor $\frac{1}{N}$ in $V_N(\theta)$.

So, for any $N > 0$

$$E\, \frac{\partial V_N}{\partial \theta}\left(\frac{\partial V_N}{\partial \theta}\right)^{\mathrm{T}}\Bigg|_{\theta=\theta_0} = \sum_{k=1}^{N} \begin{bmatrix} \frac{\partial h(k)}{\partial \theta_1}\frac{\partial h(k)}{\partial \theta_1} & \frac{\partial h(k)}{\partial \theta_1}\frac{\partial h(k)}{\partial \theta_2} & \cdots & \frac{\partial h(k)}{\partial \theta_1}\frac{\partial h(k)}{\partial \theta_q} \\ \frac{\partial h(k)}{\partial \theta_2}\frac{\partial h(k)}{\partial \theta_1} & & & \cdot \\ \cdot & & & \cdot \\ \vdots & & & \vdots \\ \cdot & & & \cdot \\ \frac{\partial h(k)}{\partial \theta_q}\frac{\partial h(k)}{\partial \theta_1} & \cdot & \cdots & \frac{\partial h(k)}{\partial \theta_q}\frac{\partial h(k)}{\partial \theta_q} \end{bmatrix} = \left(\frac{\partial S_N(\theta)}{\partial \theta}\right)^{\mathrm{T}} \frac{\partial S_N(\theta)}{\partial \theta}\Bigg|_{\theta=\theta_0}. \tag{22}$$

■

Using the result of Proposition 3 local structural identifiability analysis amounts to rank evaluation of a finite-time information matrix under the conditions as specified in the "thought" experiment of Proposition 3, and by chosing a data interval $N \geq 2n$ according to Proposition 1.

4.2. *An algorithm for gradient computation*

In Van Zee and Bosgra (1982) an algorithm has been proposed that, based on measurements $\{u(k), y(k)\}_{k=0,\dots,N}$ and analytic expressions for the partial derivatives $\frac{\partial}{\partial\theta_r}A(\theta)$, $\frac{\partial}{\partial\theta_r}b(\theta)$ and $\frac{\partial}{\partial\theta_r}c(\theta)$ with respect to θ_r ($r = 1, \dots, q$), generates the gradient $\partial V_N(\theta)/\partial\theta_r$ for a prechosen value of θ. This algorithm that is based on a dynamic programming formulation, is sketched next.

Consider the "thought" identification experiment as stated in Section 4.1. Dynamic programming formulation is based on the definition of a partial criterion function:

$$W_i(x(i), \theta) := \frac{1}{2}\sum_{k=i}^{N} \varepsilon^2(k, \theta) \tag{23}$$

and derivation of the partial derivatives of $W_i(x(i), \theta)$ with respect to $x(i)$ and θ on the given trajectory $x(i)$, where $W_i(x(i), \theta)$ is expressed as

$$W_i(x(i), \theta) = \frac{1}{2}\varepsilon^2(i, \theta) + W_{i+1}(x(i+1), \theta). \tag{24}$$

This leads to

$$\begin{aligned}\frac{\partial}{\partial x(i)}W_i(x(i), \theta) &= \frac{1}{2}\frac{\partial}{\partial x(i)}\varepsilon^2(i, \theta) \\ &\quad + \frac{\partial}{\partial x(i+1)}W_{i+1}(x(i+1), \theta)\frac{\partial x(i+1)}{\partial x(i)} \\ &= -\varepsilon(i, \theta)c(\theta) \\ &\quad + \frac{\partial}{\partial x(i+1)}W_{i+1}(x(i+1), \theta)A(\theta);\end{aligned} \tag{25}$$

$$\begin{aligned}\frac{\partial}{\partial \theta}W_i(x(i), \theta) &= \frac{1}{2}\frac{\partial}{\partial \theta}\varepsilon^2(i, \theta) \\ &\quad + \frac{\partial}{\partial x(i+1)}W_{i+1}(x(i+1), \theta)\frac{\partial x(i+1)}{\partial \theta} \\ &\quad + \frac{\partial}{\partial \theta}W_{i+1}(x(i+1), \theta).\end{aligned} \tag{26}$$

Adopting the following notation:

$$\lambda_i^{\mathrm{T}} := \frac{\partial}{\partial x(i)}W_i(x(i), \theta) \tag{27}$$

$$H_i(x(i), \lambda_{i+1}, \theta) := \frac{1}{2}\varepsilon^2(i, \theta) + \lambda_{i+1}^{\mathrm{T}}[Ax(i) + bu(i)] \tag{28}$$

calculation of the gradient $\frac{\partial}{\partial\theta}V_N(\theta)$ is carried out in the following two steps:

(1) λ_i is computed for $i = 1, \dots, N$ from

$$\lambda_i^{\mathrm{T}} = \lambda_{i+1}^{\mathrm{T}}A(\theta) - \varepsilon(i, \theta)c(\theta), \quad \lambda_{N+1}^{\mathrm{T}} = 0. \tag{29}$$

This implies a backward run on the error-sequence $\varepsilon(i, \theta)$;

(2) $\frac{\partial}{\partial\theta}H_k$ is calculated for $k = 0, \dots, N$ as

$$\frac{\partial}{\partial \theta}H_k = -\varepsilon(k, \theta)\frac{\partial}{\partial \theta}c(\theta)x(k) + \lambda_{k+1}^{\mathrm{T}}\left[\frac{\partial}{\partial \theta}A(\theta)x(k) + \frac{\partial}{\partial \theta}b(\theta)u(k)\right], \tag{30}$$

where $\frac{\partial}{\partial\theta}A(\theta)$, $\frac{\partial}{\partial\theta}b(\theta)$ and $\frac{\partial}{\partial\theta}c(\theta)$ are determined analytically.

The gradient $\frac{\partial V_N(\theta)}{\partial\theta}$ is finally found as:

$$\frac{\partial V_N(\theta)}{\partial \theta} = \sum_{k=0}^{N}\frac{\partial}{\partial \theta}H_k(x(k), \lambda_{k+1}, \theta). \tag{31}$$

4.3. *An analytical expression for the information matrix*

The results of the previous two subsections can now be employed to formulate an expression for the derivative $\frac{\partial V_{2n}}{\partial\theta_r}$.

Lemma 2. Consider the gradient expressions (31), (28) and (27) and the experimental conditions as specified in Proposition 3. Then

$$\begin{aligned}\frac{\partial V_{2n}(\theta)}{\partial \theta_r} = &-\sum_{k=1}^{2n}\varepsilon(k, \theta)\frac{\partial c}{\partial \theta_r}A^{k-1}b \\ &-\sum_{k=1}^{2n}\varepsilon(k, \theta)cA^{k-1}\frac{\partial b}{\partial \theta_r} \\ &-\sum_{j=2}^{2n}\sum_{\ell=1}^{j-1}\varepsilon(j, \theta)cA^{j-\ell-1}\frac{\partial A}{\partial \theta_r}A^{\ell-1}b.\end{aligned} \tag{32}$$

Proof. An explicit expression for λ_k^{T} is derived which follows from (27) in a straightforward way:

$$\lambda_k^{\mathrm{T}} = -\sum_{j=k}^{2n} \varepsilon(j,\theta) c A^{j-k}, \tag{33}$$

where $k = 1, \dots, 2n$ and $\lambda_{2n+1}^{\mathrm{T}} = 0$. Substitution of (33) in (31) and taking into account the "thought" experimental conditions as stated in Proposition 3 result in expression (32) for the gradient. The third term of the expressions follows by rewriting the summation $\sum_{\ell=1}^{2n}\sum_{j=k+1}^{2n}$ into $\sum_{j=2}^{2n}\sum_{\ell=1}^{j-1}$. ■

The information matrix can now be calculated element-wise by multiplying (32) with a similar expression for $\frac{\partial V_{2n}}{\partial \theta_s}$ and taking the expectation. Since $\varepsilon(k,\theta_0)$ is assumed to be a standard white noise, the terms of this product containing $E\,\varepsilon(i,\theta_0)\varepsilon(j,\theta_0)$, where $i \neq j$ will disappear. The exact elaboration is quite technical and is presented in the Appendix. Here, we only state the result.
Before we do so, some notational conventions are introduced:

(1) the extended controllability and observability matrices W_c are W_o are defined as

$$W_c := \begin{bmatrix} b & Ab & A^2 b & \dots & A^{2n-1} b \end{bmatrix}; \tag{34}$$

$$W_o := \begin{bmatrix} c^{\mathrm{T}} & A^{\mathrm{T}} c^{\mathrm{T}} & \dots & (A^{\mathrm{T}})^{2n-1} c^{\mathrm{T}} \end{bmatrix}^{\mathrm{T}}; \tag{35}$$

(2) for $i \le j$, $E_c(i,j) \in \mathbb{R}^{2n \times (j-i+1)}$ is defined as the matrix that selects the i-th up to the j-th column of a matrix with $2n$ columns by post-multiplication:

$$E_c(i,j) := \begin{bmatrix} 0_{(i-1)\times(j-i+1)} \\ I_{(j-i+1)\times(j-i+1)} \\ 0_{(2n-j)\times(j-i+1)} \end{bmatrix}, \tag{36}$$

where $i, j = 1, \dots, 2n$. $0_{p\times r}$ denotes the $(p \times r)$ 0-matrix; $I_{p\times p}$ denotes the $(p \times p)$ identity matrix;

(3) for $i \le j$, $E_o(i,j) \in \mathbb{R}^{(j-i+1)\times 2n}$ is defined as the matrix that selects the i-th up to the j-th row of a matrix with $2n$ rows by pre-multiplication:

$$E_o(i,j) := \begin{bmatrix} 0_{(j-i+1)\times(i-1)} & I_{(j-i+1)\times(j-i+1)} & 0_{(j-i+1)\times(2n-j)} \end{bmatrix}, \tag{37}$$

where $i, j = 1, \dots, 2n$.
With these notational conventions we state the following result:

Proposition 4. Under the conditions stated in Proposition 3 and using the notational conventions (34), (35), (36) and (37) the (r, s)-element of the information matrix is expressed as:

$$\begin{aligned} E&\left(\frac{\partial V_{2n}}{\partial \theta_r}\right)\left(\frac{\partial V_{2n}}{\partial \theta_s}\right) \\ &= \frac{\partial b^{\mathrm{T}}}{\partial \theta_r} W_o^{\mathrm{T}} W_o \frac{\partial b}{\partial \theta_s} + \frac{\partial c}{\partial \theta_r} W_c W_c^{\mathrm{T}} \frac{\partial c^{\mathrm{T}}}{\partial \theta_s} \\ &\quad + f_1(\theta_r,\theta_s) + f_1(\theta_s,\theta_r) + f_2(\theta_r,\theta_s) + f_2(\theta_s,\theta_r) \\ &\quad + f_3(\theta_r,\theta_s) + f_3(\theta_s,\theta_r) + f_4(\theta_r,\theta_s) \end{aligned} \tag{38}$$

with

$$f_1(\theta_1,\theta_2) = \frac{\partial c}{\partial \theta_1} W_c W_o \frac{\partial b}{\partial \theta_2} \tag{39}$$

$$\begin{aligned} f_2(\theta_1,\theta_2) &= \sum_{k=1}^{2n-1} \frac{\partial c}{\partial \theta_1} W_c E_c(k+1, 2n) \\ &\quad \times E_o(1, 2n-k)\left(W_o \frac{\partial A}{\partial \theta_2} W_c\right) E_c(k,k). \end{aligned} \tag{40}$$

$$\begin{aligned} f_3(\theta_1,\theta_2) &= \sum_{k=1}^{2n-1} \frac{\partial b^{\mathrm{T}}}{\partial \theta_1} W_o^{\mathrm{T}} E_o^{\mathrm{T}}(k+1, 2n) \\ &\quad \times E_o(1, 2n-k)\left(W_o \frac{\partial A}{\partial \theta_2} W_c\right) E_c(k,k). \end{aligned} \tag{41}$$

$$\begin{aligned} f_4(\theta_r,\theta_s) &= \sum_{k=1}^{2n-1}\sum_{\ell=1}^{2n-1} E_c^{\mathrm{T}}(k,k)\left(W_o \frac{\partial A}{\partial \theta_r} W_c\right)^{\mathrm{T}} \\ &\quad \times E_o^{\mathrm{T}}(\max(k,\ell) - k + 1, 2n - k) \\ &\quad \times E_o(\max(k,\ell) - \ell + 1, 2n - \ell) \\ &\quad \times \left(W_o \frac{\partial A}{\partial \theta_s} W_c\right). \end{aligned} \tag{42}$$

Proof. See the Appendix.

Although the analytical expression (38) seems quite elaborate, the calculation is straightforward.

Computation of the partial derivatives of the state space matrices $(A(\theta), b(\theta), c(\theta))$ with respect to θ_i $(i = 1, \dots, q)$ for $\theta = \theta_0$ is done by hand and the extended controllability and observability matrices are calculated straightforwardly.

Note that using the notation based on row- and column-selection the (r, s)-element of the information matrix is expressed in terms of:

$$W_o\left(\frac{\partial A}{\partial \theta_i}\right) W_c, \quad W_o\left(\frac{\partial b}{\partial \theta_i}\right), \quad \left(\frac{\partial c}{\partial \theta_i}\right) W_c, \qquad i = r, s. \tag{43}$$

From a computational point of view expression (38) constitutes an attractive basis for formulation of an algorithm for local structural analysis since the main computation effort consists of calculation of summations of matrix products of reasonable dimensions. Since W_o, W_c are derived from $A(\theta)$, $b(\theta)$, $c(\theta)$ this algorithm has to be supplied with analytical expressions for

$$A(\theta_0),\ b(\theta_0),\ c(\theta_0),$$
$$\left.\frac{\partial A}{\partial \theta_i}\right|_{\theta=\theta_0},\quad \left.\frac{\partial b}{\partial \theta_i}\right|_{\theta=\theta_0},\quad \left.\frac{\partial c}{\partial \theta_i}\right|_{\theta=\theta_0},$$

for $i = r, s$ to compute the (r, s)-element of the information matrix.

Expression (38) also states a very nice intuitive result, since parameter identifiability is expressed in terms of controllability and observability properties of the model and the sensitivity of the state space matrices with respect to the elements of the argument parameter vector θ. In this respect these results closely relate in spirit to the infinite-time results of the previous section.

Remark 2. As the presented test for identifiability is based on the injective map from parameters to the sequence of Markov parameters $\{h(k, \theta)\}_{k=1,\cdots 2n}$, the results also directly apply to continuous-time state space parametrizations, where this same injectivity reflects identifiability. As a consequence, continuous-time model structures can also be directly substituted in the expressions of Proposition 4 and consecutively in the rank test of Proposition 3.

5. EXAMPLE

In this section both tests for local structural identifiability are applied to a physically structured model of an industrial process for mixing solid materials (Dötsch, 1991). The process under investigation concerns the mixing behavior in case two components are stirred in a vessel. Here we consider two different situations. The first situation concerns mixing while the components remain in the vessel. In this case the process has a structural integrator, structural meaning independent of the parameter values. In the second situation mixing takes place while there is a continuous mass flow in and out of the vessel; in this case the process is structurally stable.

The structure of the model is based on discrete-time mass conservation laws that describe mass exchange between an arbitrarily chosen number of compartments. The states represent the concentration of one component in each compartment. The model order is determined by the number of compartments that constitute the vessel, denoted as $n_1 + n_2$.

The state space matrices are partitioned accordingly and are expressed as follows:

$$A(\theta) = \begin{bmatrix} A_{11}(\theta) & A_{12}(\theta) \\ A_{21}(\theta) & A_{22}(\theta) \end{bmatrix},$$

where

$$A_{11}(\theta) = \begin{bmatrix} 1-\theta_1-\theta_2 & 0 & \dots & & 0 & \theta_4^{-1}\theta_6 \\ \theta_6^{-1}\theta_1\theta_4 & 0 & & & & 0 \\ 0 & 1 & 0 & & & . \\ \vdots & & \ddots & \ddots & & \vdots \\ & & & 1 & 0 & \\ . & & & 0 & 1 \ 0 & 0 \\ 0 & & \dots & & 0 \ 1 & 0 \end{bmatrix}$$

with dimension $n_1 \times n_1$ and

$$A_{22}(\theta) = \begin{bmatrix} 1-\theta_3 & 0 & \dots & & . & 0 \\ \theta_3 & 1-\theta_3 & 0 & & & 0 \\ 0 & \theta_3 & \ddots & & & \\ \vdots & & \ddots & \ddots & & \vdots \\ & & 0 & \theta_3 & 1-\theta_3 & 0 \\ 0 & \dots & & 0 & \theta_3 & f_o-\theta_3 \end{bmatrix}$$

with dimension $n_2 \times n_2$. In the case where the process has a structural integrator, the parameter f_o is equal to 1; in case of a continuous mass flow (structurally stable process) f_o must satisfy $f_o < 1$ and is chosen equal to 0.55.

The elements of $A_{12}(\theta)$ and $A_{21}(\theta)$ are all equal to zero except the $(1, 1)$ element of $A_{21}(\theta) = \theta_5^{-1}\theta_2\theta_4$ and the $(1, n_2)$ element of $A_{12}(\theta) = \theta_4^{-1}\theta_3\theta_5$.

The input and output vector are:

$$b(\theta) = \left[0 \dots 0 \ \vdots \ (\theta_4 + n_2\theta_5 + n_1\theta_6)\ 0 \dots 0\right]^{\mathrm{T}};$$
$$c(\theta) = \left[0 \dots 0\ 1\ 0 \ \vdots \ 0 \dots 0\right].$$

The model has six parameters with a physical interpretation: $\theta_{1,2,3}$ represent flow-rates and $\theta_{4,5,6}$ compartment volumes. The choice of parameter for which structural analysis is carried out is based on physical insight and model simulations:

$$\theta_0 = [0.0792\ \ 0.1759\ \ 0.9500\ \ 9.4406\ \ 1.6608\ \ 0.7867]^{\mathrm{T}}.$$

For $f_o = 0.55$ the infinite and finite-time information matrix is calculated; for $f_o = 1$ only the finite-time information matrix is calculated. The results are shown in terms of the singular values of the calculated information matrices in Table 1.

From the zero singular value it may be concluded that for $f_o = 0.55$ and $f_o = 1$ the model is not locally structural identifiable at $\theta = \theta_0$.

A natural step to take now is to eliminate those parameters from the identification problem that show very small contribution to the input–output

Table 1. Singular values of the information matrices ($\times 10^4$)

$f_0 = 0.55$		$f_0 = 1$
infinite IM	finite IM	
2.1375542	1.7522639	3.8864028
0.9885777	0.8815305	1.9976371
0.0540982	0.0409189	0.0967318
0.0014324	0.0010992	0.0030977
0.0000210	0.0000026	0.0000099
0.0000000	0.0000000	0.0000000

Table 2. Diagonal elements of the information matrices ($\times 10^4$)

$f_0 = 0.55$		$f_0 = 1$
infinite IM	finite IM	
1.7962620	1.4037856	3.3047459
0.2337035	0.1445630	0.3544174
1.1234486	1.1042209	2.2674753
0.0001111	0.0000202	0.0000551
0.0136405	0.0112508	0.0277019
0.0145179	0.0119745	0.0294838

behavior of the model. A nice feature of the information matrix is that the diagonal elements express parameter sensitivity of the Markov parameters (see (22)). The choice which parameters to eliminate can thus be based on the diagonal elements of the information matrix. In Table 2 the diagonal elements of the information matrices are shown. In this case the model dynamics seem to be the least sensitive with respect to θ_4.

6. CONCLUSIONS

In the case of model parametrizations based on physical modeling it is imperative to test local structural identifiability before parameter estimation is carried out. In this paper two methods are presented to test local structural identifiability of SISO, non-linearly parametrized state space models of high order. They are based on calculation and rank evaluation of an infinite-time and finite-time information matrix. In the infinite-time case a compact solution is presented that requires solving a high-dimensional Lyapunov equation and only holds for stable models. A finite-time information matrix enables identifiability analysis for stable as well as unstable models. For calculation of a finite-time information matrix a specific "thought" identification experiment is formulated and an algorithm for gradient computation, as formulated by Van Zee and Bosgra (1982), is applied. Both methods apply to discrete-time and continuous-time model structures. Their applicability is verified by showing the results of tests with a stable and marginally stable model of an industrial mixing process of order 40.

Acknowledgements—The work of Hans Dötsch is supported by Philips Research Laboratories, Eindhoven, The Netherlands. The authors gratefully acknowledge Ralf Peeters, for bringing his work to their attention.

REFERENCES

Bellman, R. and K. J. Åström (1970). On structural identifiability. *Mathematical Biosciences*, **7**, 329–339.

Cobelli, C., A. Lepschy and G. Romanin Jacur (1979). Identifiability of compartmental systems and related structural properties. *Mathematical Biosciences*, **44**, 1–18.

Dasgupta, S. and B. D. O. Anderson (1987). Physically based parameterizations for designing adaptive algorithms. *Automatica*, **23**, 469–477.

Dötsch, J. G. M. (1991). Identification of physically structured models of a Nauta mixer. Report A-513. Mech. Eng. Systems and Control Group, Delft Univ. of Technology, (in Dutch).

Glover, K. and J. C. Willems (1974). Parametrizations of linear dynamical systems: canonical forms and identifiability. *IEEE Trans. Automat. Control*, **AC-19**, 640–645.

Godfrey, K. (1983). *Compartmental Models and their Application.* Academic Press, London.

Grewal, M. S. and K. Glover (1976). Identifiability of linear and nonlinear dynamical systems. *IEEE Trans. Automat. Control, Short Papers*, 833–837.

Ho, B. L. and R. E. Kalman (1966). Effective construction of linear state-variable models from input–output functions. *Regelungstechnik, Heft* **12**, 545–548.

Ljung, L. (1987). *System Identification: Theory for the User.* Prentice Hall, Englewood Cliffs, NJ.

Ljung, L. and T. Glad (1991). On global identifiability for arbitrary model parametrizations. Report LiTH-ISY-I-1308, Dept. of Electrical Engineering, Linköping University, Sweden.

Norton, J. P. (1980). Normal-mode identifiability analysis of linear compartmental systems in linear stages. *Mathematical Biosciences*, **50**, 95–115.

Peeters, R. L. M. (1994). System identification based on Riemannian geometry theory and algorithms. PhD thesis, Tinbergen Institute Research Series, Vol. 64.

Rothenberg, T. J. (1971). Identification in parametric models. *Econometrica*, **39**, 577–591.

Rubinow, S. I. and A. Winzer (1971). Compartment analysis: an inverse problem. *Mathematical Biosciences*, **11**, 203–247.

Tse, E. (1973). Information matrix and local identifiability of parameters. In *Proc. Joint Automat. Control Conf.*, Columbus, OH, pp. 611–619.

Van Zee, G. A. and O. H. Bosgra (1982). Gradient computation in prediction error identification of linear, discrete-time systems. *IEEE Trans. Automat. Control*, **AC-27**, 738–739.

Walter, E. (1981). Identifiability of State Space Models. In *Lecture Notes in Biomathematics, vol.* 46, Springer, Berlin.

APPENDIX A. PROOF OF PROPOSITION 4

Calculation of the product $E\frac{\partial V_{2n}}{\partial \theta_r}\frac{\partial V_{2n}}{\partial \theta_s}$ is carried out by employing the expression (32) in Lemma 2 together with $E\varepsilon(i, \theta_0)\varepsilon(j, \theta_0) = 0$ for $i \neq j$.

The first two terms in the expression (38) follow directly by considering the quadratic terms of the first and second terms in (32).

The f_1-terms in (38) follow from

$$\sum_{k=1}^{2n} \frac{\partial c}{\partial \theta_r} A^{k-1} b \left(c A^{k-1} \frac{\partial b}{\partial \theta_s} \right) = \frac{\partial c}{\partial \theta_r} W_c W_o \frac{\partial b}{\partial \theta_s}.$$

The f_2-terms reflect the product of terms dependent on $\frac{\partial c}{\partial \theta}$ and $\frac{\partial A}{\partial \theta}$, and so

$$f_2(\theta_r, \theta_s) = \sum_{j=2}^{2n} \frac{\partial c}{\partial \theta_r} A^{j-1} b \sum_{\ell=1}^{j-1} c A^{j-\ell-1} \frac{\partial A}{\theta_s} A^{\ell-1} b.$$

Rewriting the summations shows that this is equivalent to the expression

$$\sum_{\ell=1}^{2n-1} \sum_{j=\ell+1}^{2n} \frac{\partial c}{\partial \theta_r} A^{j-1} b c A^{j-\ell-1} \frac{\partial A}{\theta_s} A^{\ell-1} b$$

from which the expression for f_2 in the proposition follows directly.

The f_3-terms reflect the product of terms dependent on $\frac{\partial b}{\partial \theta}$ and $\frac{\partial A}{\partial \theta}$. They are calculated in a way that is similar to the calculation of f_2.

The f_4-term reflects the quadratic term of $\frac{\partial A}{\partial \theta}$. It follows from (32) that this term is given by

$$\begin{aligned}
f_4(\theta_r, \theta_s) &= \\
&= \sum_{j=2}^{2n} \sum_{k=1}^{j-1} \sum_{\ell=1}^{j-1} \left(c A^{j-k-1} \frac{\partial A}{\partial \theta_r} A^{k-1} b \right) \\
&\quad \times \left(c A^{j-\ell-1} \frac{\partial A}{\partial \theta_s} A^{\ell-1} b \right) \\
&= \sum_{k=1}^{2n-1} \sum_{\ell=1}^{2n-1} \sum_{j=\max(k,\ell)+1}^{2n} \left(\frac{\partial A}{\partial \theta_r} A^{k-1} b \right)^{\mathrm{T}} \\
&\quad \times (c A^{j-k-1})^{\mathrm{T}} (c A^{j-\ell-1}) \frac{\partial A}{\partial \theta_s} A^{\ell-1} b \\
&= \sum_{k=1}^{2n-1} \sum_{\ell=1}^{2n-1} \left(\frac{\partial A}{\partial \theta_r} A^{k-1} b \right)^{\mathrm{T}} W_o^{\mathrm{T}} \\
&\quad \times E_o^{\mathrm{T}}(\max(k, \ell) - k + 1, 2n - k) \\
&\quad \times E_o(\max(k, \ell) - \ell + 1, 2n - \ell) W_o \frac{\partial A}{\partial \theta_s} A^{\ell-1} b,
\end{aligned}$$

which proves the result. □