Appl. Math. and Comp. Sci., 1996, vol.6, No.1, 77–95

TWO-STAGE IDENTIFICATION OF INTERCONNECTED STEADY-STATE SYSTEMS WITH CASCADE STRUCTURE: A PARAMETRIC APPROACH PART 1: BACKGROUND

ZYGMUNT HASIEWICZ*

The paper deals with selecting, within a stochastic framework, an approximate system model from a parametric candidate model set for a composite steady-state system with cascade structure, assuming mean-squared model output error as a measure of the model accuracy. With applications in mind, particular emphasis is laid on computational simplicity of the model search routine and an easy-to-use but suboptimal two-stage approach is proposed for solving the corresponding identification task. In the first part, a theoretical background for the method is given and the degree of suboptimality of the resultant model is analysed under full probabilistic knowledge of the system. Some illustrative examples are included to inquire into applicability of the method. The empirical counterpart of the algorithm, employing the measured input-output data from the plant, is investigated in Part 2.

1. Introduction

Since the early 1970s, a considerable interest has been drawn to the development of decentralized techniques for the steady-state optimization control of interconnected complex industrial processes, see e.g. (Brdyś and Roberts, 1986; Brdyś *et al.*, 1980; 1986; 1990; Chen *et al.*, 1986; Findeisen *et al.*, 1980; Lin *et al.*, 1988a; 1988b; 1989; 1991; Michalska *et al.*, 1985; Shao and Roberts, 1983; 1985; Tatjewski *et al.*, 1990; Zhang and Roberts, 1991) and the earlier papers cited therein. In general, such studies require respective static mathematical models that preserve the 'structural' properties of the processes, i.e. retain the coupling effects of the interconnected process at hand and are in a sense sufficiently accurate to bring proper analysis or control design tools into efficient use. Since the real processes are typically characterized by substantial complexity, limited prior knowledge and uncertainties in the process equations, some gap between the true description and a model of the process is unavoidable in practice, and only approximate models of reality can be gained in most cases. Then, as a rule, identification of subprocesses consists in selecting, within suitably chosen sets of models, those approximate models which, combined together, ensure the best fit to

* Technical University of Wrocław, Institute of Engineering Cybernetics, ul. Janiszewskiego 11/17, 50–372 Wrocław, Poland

the process as a whole (in the sense of some global performance index), maintaining a particular structure of the process. Unfortunately, methods for the 'exact' solution of the complex identification problems of such kind, developed until now, turn out to be complex and computationally expensive (Hasiewicz and Stankiewicz, 1984a; 1984b; 1986a; 1986b). Accordingly, for systems of even moderate complexity, it can be quite difficult to find an optimum mathematical model in some given set of models. A solution can be a trade-off between the model accuracy and computational complexity, resulting in identification algorithms relevant for practical applications, which will however produce in general suboptimal models only.

In this study, we are concerned with the problem of determining an approximate model for a noisy serially interconnected complex system operating in a steady-state when the global mean-squared model output error (MSE) is assumed as a measure of approximation accuracy, and a two-stage scheme for computing the cascade model is introduced. First, optimum models of suitable aggregated parts of the system are independently established and then the models of component subsystems are derived in such a way that the overall complex model can in effect be obtained in an easy decentralized manner by solving only standard optimization tasks. However, computational efficiency of the method is obtained at the cost of an undesirable side effect of causing the final model to lose optimality (the auxiliary optimum models of system aggregates are anyhow available in the method as a by-product). A thorough analysis of applicability of the approach is the purpose of the work. The considerations are split into two parts. In this part, the background of the method is given and the 'theoretical' version of the two-stage algorithm, based on the use of full probabilistic knowledge of the system, is examined. In Part 2, an empirical counterpart of the scheme, employing only measured data from the system, is presented and analysed.

The present paper is organized as follows. In Section 2, general assumptions about the cascade system and the set of candidate models, referring to the main body of the research, are collected and the system identification problem is formulated. A solution to this problem is discussed in Section 3, and a two-stage identification strategy is proposed in conclusion. The degree of suboptimality of the resulting model is then investigated in Section 4 under some additional 'regularity' restrictions on subsystem models, and respective sufficient conditions for achieving the required relative precision by the model (a prescribed suboptimality index) are provided together with specific conditions which guarantee optimality of the identification results. Next, in Section 5, simple illustrative examples are studied to give an insight into applicability of the method and to explain the statistical meaning of basic factors occuring in the analysis.

2. Problem Statement

Consider an interconnected static system with cascade structure, composed of a collection of n subsystems $S_1, S_2, \ldots, S_n$ (see Fig. 1), whose genuine input-output behaviour is described by the equations

$$y_1 = F_1^*(c_1, z_1), \quad y_i = F_i^*(c_i, u_i, z_i), \quad u_i = y_{i-1}, \quad i = 2, 3, \ldots, n \tag{1}$$

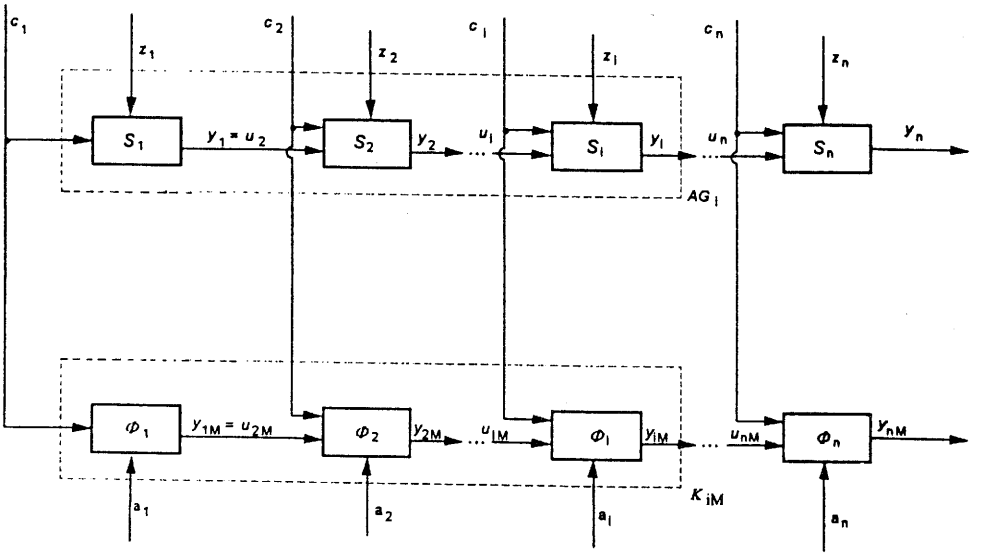

Fig. 1. Series system under consideration and its complex model.

where $y_i \in Y_i \subset \mathbb{R}^{l_i}$, $u_i \in U_i \subset \mathbb{R}^{m_i}$, and $c_i \in C_i \subset \mathbb{R}^{s_i}$ denote respectively outputs, non-manipulable interaction inputs, and external excitations (e.g. control inputs) of the component S_i, while vector $z_i \in Z_i \subset \mathbb{R}^{d_i}$ represents unmeasurable disturbances affecting the subsystem. It is assumed that the true descriptions of subsystems S_i, i.e. the mappings $F_i^* : C_i \times U_i \times Z_i \to \mathbb{R}^{l_i}$, $i = 1, 2, \ldots, n$, are completely unknown and there is a need for a mathematical model of system (1) reflecting the composite structure of the system. The corresponding system identification task consists in determining a complex structural model of system (1), including models of individual system elements and reproducing interactions existing within the complex system. Here, we shall discuss the problem when the search for a suitable model is conducted over a parametric set of candidate models, i.e. the admissible models of individual subsystems S_i are defined in the parametric form—by parameter sets $A_i \subset \mathbb{R}^{p_i}$ and the corresponding parametrized mappings $\Phi_i : \mathbb{R}^{s_i} \times \mathbb{R}^{m_i} \times A_i \to \mathbb{R}^{l_i}$ for $i = 1, 2, \ldots, n$. Then the resultant parametric structural model of a series system as a whole is as follows (see Fig. 1):

$$y_{1M} = \Phi_1(c_1, a_1), \quad y_{iM} = \Phi_i(c_i, u_{iM}, a_i), \quad u_{iM} = y_{i-1,M}, \quad i = 2, 3, \ldots, n \qquad (2)$$

with $a_i \in A_i$, where $y_{iM} \in \mathbb{R}^{l_i}$ and $u_{iM} \in \mathbb{R}^{m_i}$ are respectively the outputs and interaction inputs of the i-th submodel in the complex model, corresponding to external excitations $c_1, c_2, \ldots, c_i$ used in the system. With parameter $a = (a_1, a_2, \ldots, a_n)$ ranging over the set $A = A_1 \times A_2 \times \ldots \times A_n$ this is just a collection of feasible complex models in the case under consideration.

Such a situation arises typically when structural models are intended to be established for real-life industrial processes. Most of them are arranged in some technological lines, with specific order of operations, i.e. serially structured. Let us consider

e.g. the process of ore concentration (being a part of ore dressing line in a typical ore mine). This is a large-scale complex industrial process, poorly recognized due to the complexity of the phenomena and thus unfit to be described in the form (1). Usually $n = 4$ consecutive stages are distinguished: ore milling and next the preliminary, main, and final (purgative) ore flotation, each stage of considerable complexity. The sub-processes in such a cascade have the following external (control) inputs: for the mill block (subsystem S_1 in Fig. 1)—the intensity of water inflow (c_1), for the flotation blocks (subsystems S_i, $i = 2, 3, 4$)—the intensity of water, wetting agents, foaming agents, and compressed air inflow (c_{i1}, c_{i2}, c_{i3}, and c_{i4}, respectively). Interaction inputs u_i and outputs y_i are in turn composed of basic descriptors characterizing the feed and product material on the corresponding block input and output, and include such quantities as e.g. the granulation (classified by the number of grain classes), the content of the useful mineral in the ore-pulp, and the intensity of ore pulp flow. This results in vectors of multiple dimensions. The process operates moreover in a steady-state. This is due to the steadiness of chemical composition of the ore and physical ore body parameters over long time intervals.

It should be emphasized here, and this is essential for our considerations, that—due to the minor prior knowledge of the system (see the above example)—the selected model structure (2) is not assumed to contain the correct description of system (1), i.e. we admit in general that $\Phi_i \neq F_i^*$ for any choice of a. Therefore the system identification task is now that of approximation of unknown system characteristics within the parametric set of models (2), with an acceptable non-unique solution and most likely arbitrary choice of the approximate class of models, and not a parameter estimation problem in the true description of static series system (1). The latter important problem for serially connected systems and other more general structures, as well as the related question of parameter identifiability for complex static systems, were discussed separately in (Hasiewicz, 1987; 1988a; 1988b; 1989).

In the usual stochastic framework, when c_i, z_i and hence also u_i and y_i are random variables in system (1) (e.g. the inflow intensities in the above example due to the inaccuracy of controlling devices (feeders)), the following mean-squared model output error

$$Q_i(a_i \mid a_1, a_2, \ldots, a_{i-1}) = E\|y_i - \Phi_i(c_i, u_{iM}, a_i)\|^2 \tag{3}$$

can be taken as a measure of accuracy of the i-th submodel within the complex model for $i \geq 2$, and respectively

$$Q_1(a_1) = E\|y_1 - \Phi_1(c_1, a_1)\|^2 \tag{4}$$

for $i = 1$ (as for a single-element system), where the parameters $a_1, a_2, \ldots, a_{i-1}$ of the preceding submodels in the cascade (2) are considered as being fixed, $\|\cdot\|$ is the Euclidean vector norm in $\mathbb{R}^{l_i}$, and E denotes the expectation with respect to $(c_1, c_2, \ldots, c_i, y_i)$ provided that such an average is well-defined for the problem at hand. Accordingly, the quality of the overall complex model (2) may be quite

naturally estimated by using the cumulative index

$$\begin{aligned} Q(a_1, a_2, \ldots, a_n) = \; & Q_1(a_1) + \sum_{i=2}^{n} Q_i(a_i \mid a_1, a_2, \ldots, a_{i-1}) \\ = \; & E\|y_1 - \Phi_1(c_1, a_1)\|^2 + \sum_{i=2}^{n} E\|y_i - \Phi_i(c_i, u_{iM}, a_i)\|^2 \end{aligned} \tag{5}$$

which represents the global mean-squared approximation error of composite system outputs (the global mean-squared model output error) for given $c_1, c_2, \ldots, c_n$, applied in parallel to the system and the model.

The evaluation of the series model minimizing the index (5) over the model set (2) is the objective of this work.

Though, if (5) is used, the summands are standard mean-squared loss functions, the entire system identification task (determination of optimum $a^* = (a_1^*, a_2^*, \ldots, a_n^*)$, minimizing the error (5) on the set $A = A_1 \times A_2 \times \ldots \times A_n$ subject to the constraints (2)) is nevertheless not standard due to the interconnections inherent in the complex model. In particular, it cannot be simply solved by separate minimization of the indices Q_i ((3)–(4)) with respect to $a_i \in A_i$ (setting $u_{iM} = u_i$), i.e. as a set of n independent identification tasks of individual system components S_i treated as autonomous objects. (Notice that interactions u_{iM} in the complex model (2) differ in general from those in the real system, $u_{iM} \neq u_i$, and for the i-th submodel they depend on the choice of parameters $a_1, a_2, \ldots, a_{i-1}$ in the preceding components of the composite model (2), cf. Fig. 1.) On the other hand, when the series system is actually large (as in our example), the 'exact' solution—it is the strict minimization of the global index (5) on A with regard to the constraints introduced by the complex model (2)—is not a good choice for practical implementation because, even exploiting the nested modular structure of the problem, it generally leads to severe optimization tasks which require much computation (cf. for instance (Bubnicki, 1975) where the dynamic programming approach was applied). A remedy can be an 'approximate' solution to the problem, where the amount of computation is reduced at the cost of accuracy of the model.

The objective of this paper is to introduce a computationally efficient and convenient in implementation method of finding a suboptimal solution to the considered system identification task, and to examine when the algorithm is applicable to the series system identification in the sense that it provides a reasonable approximation of the optimum model.

Let us denote

$$c = (c_1, c_2, \ldots, c_n), \quad z = (z_1, z_2, \ldots, z_n), \quad y = (y_1, y_2, \ldots, y_n)$$

We shall assume throughout the paper that

a) the external excitation c of the system is a random vector with finite variance;

b) the disturbance z is a c-independent vector random variable of independent components with zero mean and finite variances;

c) the composite parametric model (2) is well-posed, i.e. for $i = 1, 2, \ldots, n$ the following mappings are well-defined:

$$y_{iM} = K_{iM}(\bar{c}_i, b_i) \tag{6}$$

with

$$K_{iM}(\bar{c}_i, b_i) = \Phi_i(c_i, K_{i-1,M}(\bar{c}_{i-1}, b_{i-1}), a_i) \tag{7}$$

and

$$b_i = \varphi_i(b_{i-1}, a_i) \tag{8}$$

for $i \geq 2$, where

$$\bar{c}_i = (\bar{c}_{i-1}, c_i) = (c_1, c_2, \ldots, c_{i-1}, c_i)$$

and $b_1 = a_1$, $\bar{c}_1 = c_1$, $K_{1M}(\bar{c}_1, b_1) = \Phi_1(c_1, a_1)$ for $i = 1$;

d) $\Phi_i(c_i, u_{iM}, a_i)$ are continuous for $i = 1, 2, \ldots, n$;

e) the sets of admissible model parameters $A_i \subset \mathbb{R}^{p_i}, i = 1, 2, \ldots, n$, are compact.

Assumptions (a) and (b) are standard in the system identification framework. As regards assumption (c), it formally means that the models of individual subsystems are compatible, i.e. they can be joined in a cascade giving well-defined models (6) of the aggregates $AG_i = \{S_1, S_2, \ldots, S_i\}$, composed of subsystems $S_1, S_2, \ldots, S_i$ (cf. Fig. 1). These aggregate models are assembled recursively due to (7) and the function (8) describes the way of composing ('glueing') the aggregate AG_i model parameter b_i based on the parameters b_{i-1} and a_i of the models of the aggregate AG_{i-1} and the subsystem S_i. Mappings (6) and (8) are for instance well-defined for the usable class of models (2) being linear in the interactions and parameters (see examples in the following). In turn, under assumption (d) it follows that the functions K_{iM} and φ_i in (6)–(8) are continuous with respect to both arguments and therefore, including assumption (e), the sets of admissible aggregate model parameters $B_i = \varphi_i(B_{i-1} \times A_i)$, $B_1 = A_1$ are compact, which will be essential in what follows.

Note that under the assumed conditions the reference 'theoretical' mean-squared optimum series model does exist.

3. Two-Stage Identification Algorithm

From this point on, full probabilistic knowledge of the system is supposed, i.e. we act as if the expectations in (5) could be theoretically computed.

Denote $B = B_1 \times B_2 \times \ldots \times B_n$ and let B_0 ($\subseteq B$) be defined as follows:

$$B_0 = \Big\{ b = (b_1, b_2, \ldots, b_n) \in B : \exists a = (a_1, a_2, \ldots, a_n) \in A \\ \text{such that } b_i = \varphi_i(b_{i-1}, a_i),\ i \in \overline{2, n} \Big\}$$

By applying (6), the cumulative model quality index (5) can be replaced with the index of the form

$$q(b) = \sum_{i=1}^{n} q_i(b_i) \tag{9}$$

where

$$q_i(b_i) = E\|y_i - K_{iM}(\bar{c}_i, b_i)\|^2 \tag{10}$$

and the identification task from Section 2 may be restated as the minimization problem of $q(b)$ on the set B_0. Let us denote the solution by $b_0^* = (b_{10}^*, b_{20}^*, \ldots, b_{n0}^*)$. The desired optimum model parameters a_i^* can then be immediately obtained as $a_1^* = b_{10}^*$ for $i = 1$, and for $i \geq 2$ by solving the equations

$$\varphi_i(b_{i-1,0}^*, a_i) = b_{i0}^*, \qquad i = 2, 3, \ldots, n \tag{11}$$

in the sets A_i. Noticing in turn that

$$b \in B_0 \quad \text{iff} \quad b \in B \quad \text{and} \quad \mu_i(b_{i-1}, b_i) = 0, \qquad i = 2, 3, \ldots, n$$

where

$$\mu_i(b_{i-1}, b_i) = \min_{a_i \in A_i} \|b_i - \varphi_i(b_{i-1}, a_i)\| \tag{12}$$

we ascertain that the latter optimization problem is, in fact, an optimization task on the set B with additional non-linear equality constraints $\mu_i(b_{i-1}, b_i) = 0$, $i = 2, 3, \ldots, n$ and implicit constraint functions μ_i given by (12). In general, analytical solution is not possible, but the optimal b_0^* can in principle be obtained numerically, by applying e.g. the standard penalty function approach, i.e. by penalizing the quality index $q(b)$ with complementary constraints (12) as follows (q_i and μ_i are continuous and B_0 is a closed set—cf. assumptions (d) and (e) and further remarks in Section 2):

$$q_p(b, \rho) = q(b) + \rho \sum_{i=2}^{n} \mu_i(b_{i-1}, b_i)$$

and implementing a numerical minimum search routine to $q_p(b, \rho)$ on the set B, with b_i varying independently over the sets B_i and gradually increasing the penalty coefficient $\rho > 0$.

Such a strategy for optimum model computation, in the 'pure' form, is however inconvenient to employ, mainly due to the necessity of computing in each step (iteration) of the algorithm the value of penalty function components (12), i.e. determining step by step the minimizers

$$\hat{a}_i(b_{i-1}, b_i) = \arg \min_{a_i \in A_i} \|b_i - \varphi_i(b_{i-1}, a_i)\| \tag{13}$$

for $i = 2, 3, \ldots, n$ (the computations necessary to find the optimal model are then arranged hierarchically in a three-level structure—cf. (Hasiewicz and Stankiewicz, 1986b). A conceptually simple way of overcoming this difficulty and simplifying computations follows from the observation that at the optimum point b_0^* it holds

$$\sum_{i=2}^{n} \mu_i(b_{i-1,0}^*, b_{i,0}^*) = 0$$

Motivated by this fact and owing to (9), the form of $q_p(b,\rho)$, and (12)–(13), one can naturally think of the following approximate two-stage algorithm:

Stage 1. Compute b_i^* minimizing the indices $q_i(b_i)$ in (10) on the sets B_i, independently for $i = 1,2,\ldots,n$.

Stage 2. Using b_i^* from Stage 1, determine the submodel parameters as

$$\hat{a}_i = \arg \min_{a_i \in A_i} \|b_i^* - \varphi_i(b_{i-1}^*, a_i)\| \tag{14}$$

for $i = 2,3,\ldots,n$, taking $\hat{a}_1 = b_1^*$ for $i = 1$ (in what follows, $\hat{a}_i(b_{i-1}^*, b_i^*)$ is briefly denoted by $\hat{a}_i$).

The main idea behind the above two-stage approach consists thus in progressive aggregation of the cascade system and its complex model in the first stage, i.e. in stacking the succeeding subsystems and their models in evolving complexes (aggregates; cf. Fig. 1), for which optimum model parameters are then determined independently, as for the set of independent objects (with inputs $\bar{c}_i$ and outputs y_i). Afterwards, in Stage 2, the respective parameters for the component submodels of the series model, maintaining the structure of the system, are derived in a decentralized manner according to (14), i.e. the models from Stage 1 are disaggregated and a structure preserving the complex model is assembled as a final result.

The advantage consists now in the fact that calculations needed by the algorithm are standard and can be accomplished at both stages in a completely decentralized way. In Stage 1, to compute b_i^* common minimization methods can be adopted, separately for $i = 1,2,\ldots,n$, and in Stage 2 only the solution, in the least-squares sense, of the equations

$$\varphi_i(b_{i-1}^*, a_i) = b_i^* \tag{15}$$

with respect to a_i (in the sets A_i) is additionally required for $i = 2,3,\ldots,n$. They can be solved in turn by applying general-purpose optimization techniques ($\hat{a}_i$ is then treated as the minimum point (14)) or specialized methods of solving non-linear least-squares problems (Dennis and Schnabel, 1983; Forsythe *et al.*, 1977).

The algorithm is however heuristic and therefore only a suboptimal model can be expected. The formal reasons are twofold. Firstly, the vector $b^* = (b_1^*, b_2^*, \ldots, b_n^*)$ composed of b_i^* obtained in Stage 1 minimizes the performance index $q(b)$ in (9) on the set $B \supseteq B_0$. Since most likely the minimum point of $q(b)$ on B differs from that on B_0, $b^* \neq b_0^*$, one has

$$q(b^*) \leq q(b_0^*) \tag{16}$$

and in general $\hat{a}_i \neq a_i^*$ (cf. (15) and (11)). Secondly, for the series model (2) with the parameters $\hat{a}_i$ following from (14) the corresponding aggregate model parameters (8) are $b_i = \hat{b}_i$, where

$$\hat{b}_1 = b_1^*, \quad \hat{b}_i = \varphi_i(\hat{b}_{i-1}, \hat{a}_i), \qquad i = 2,3,\ldots,n \tag{17}$$

and obviously $\hat{b} = (\hat{b}_1, \hat{b}_2, \ldots, \hat{b}_n) \in B_0$. Hence

$$q(b_0^*) \leq q(\hat{b}) \tag{18}$$

Thus, the key problem is evaluation of the loss $q(\hat{b}) - q(b_0^*)$ occuring in the model accuracy when the two-stage approach is used.

4. Accuracy Analysis

For rating the loss $q(\hat{b}) - q(b_0^*)$ in the model quality one may implement as a test quantity the following relative (normalized) error:

$$\Delta_0 \triangleq \frac{q(\hat{b}) - q(b_0^*)}{q(b_0^*)}$$

or, alternatively,

$$\Delta \triangleq \frac{q(\hat{b}) - q(b_0^*)}{q(b^*)}$$

checking the discrepancy $q(\hat{b}) - q(b_0^*)$ against the optimum (minimum in the set B_0) or the least possible (minimum in the set B) value of the quality index q in the candidate model set, respectively, the latter being attainable in the complex model when $b^* \in B_0$. Since by (16) and (18) we obtain

$$\Delta_0 \leq \Delta \leq \Delta^* \tag{19}$$

where

$$\Delta^* = \frac{q(\hat{b}) - q(b^*)}{q(b^*)} \tag{20}$$

we shall further examine the dominant Δ^*, and it will be considered as the sub-optimality index of the model provided by the two-stage approach (in brief: of the two-stage approach). To enhance the clarity and interpretation of the results, apart from $Ez_i = 0$ (assumption (b) in Section 2), we assume here that also $Ec_i = 0$ for $i = 1, 2, \ldots, n$. Then $E\bar{c}_i = 0$ and

$$E\|\bar{c}_i\|^2 \triangleq \bar{\sigma}_{c,i}^2 < \infty, \qquad E\|z_i\|^2 \triangleq \sigma_{z,i}^2 < \infty \tag{21}$$

are simply cumulative (aggregated) variances of the components of $\bar{c}_i$ and z_i for $i = 1, 2, \ldots, n$. More general cases can be analysed in a similar manner with an obvious extension of notions.

Let δ_i^{*2} be the minimum value of the index (10) in Stage 1 (minimum mean-squared model output error), i.e.

$$\delta_i^{*2} \triangleq q_i(b_i^*) = E\|y_i - K_{iM}(\bar{c}_i, b_i^*)\|^2 = \min_{b_i \in B_i} E\|y_i - K_{iM}(\bar{c}_i, b_i)\|^2 \tag{22}$$

and let

$$\hat{\epsilon}_i \triangleq \|b_i^* - \varphi_i(b_{i-1}^*, \hat{a}_i)\| = \min_{a_i \in A_i} \|b_i^* - \varphi_i(b_{i-1}^*, a_i)\| \tag{23}$$

stand for the equation error of (15), i.e. the mismatch arising in Stage 2, for $i = 1, 2, \ldots, n$. Denote next

$$\delta_m^{*2} = \min_{1 \le i \le n} \delta_i^{*2}, \qquad \delta_M^{*2} = \max_{1 \le i \le n} \delta_i^{*2} \tag{24}$$

and

$$\hat{\epsilon}_M = \max_{1 \le i \le n} \hat{\epsilon}_i \tag{25}$$

and assume that $\delta_m^{*2} > 0$. Statistical interpretation of the quantities (21)–(25) is given later, in Example 2 of Section 5.

To gain an insight into the quality of complex models, namely the 'desired' optimum and the 'approximate' one, resulting from the two-stage approach, and to evaluate explicitly suboptimality of the latter, some more specific regularity conditions have to be imposed on parametric system models. To the end of this section we shall assume in addition to (c) and (d) of Section 2 that $K_{iM}(\bar{c}_i, b_i)$ and $\varphi_i(b_{i-1}, a_i)$ in (6) and (8) are Lipschitz with respect to b_i and b_{i-1}, respectively, i.e.

$$\|K_{iM}(\bar{c}_i, b_i^1) - K_{iM}(\bar{c}_i, b_i^2)\| \le k_i(\bar{c}_i)\|b_i^1 - b_i^2\|, \quad \forall b_i^1, b_i^2 \in B_i \tag{26}$$

and

$$\|\varphi_i(b_{i-1}^1, a_i) - \varphi_i(b_{i-1}^2, a_i)\| \le de_i(a_i)\|b_{i-1}^1 - b_{i-1}^2\|, \quad \forall b_{i-1}^1, b_{i-1}^2 \in B_{i-1} \tag{27}$$

for each $i = 1, 2, \ldots, n$ and that

$$\sup_{\bar{c}_i \in \bar{C}_i} \frac{k_i(\bar{c}_i)}{\|\bar{c}_i\|} \triangleq \alpha_i < \infty, \qquad \sup_{a_i \in A_i} de_i(a_i) \triangleq D_i < \infty \tag{28}$$

where $\bar{C}_i = C_1 \times C_2 \times \ldots \times C_i$. Such requirements are satisfied e.g. for a linear set of models, as in (41) of Example 1 in Section 5. In addition to (22)–(23), let us denote

$$\Sigma_{c,i}^2 \triangleq \alpha_i^2 \bar{\sigma}_{c,i}^2 \tag{29}$$

and furthermore (using the same convention as in (24)–(25))

$$\Sigma_{c,M}^2 = \max_{1 \le i \le n} \Sigma_{c,i}^2 \tag{30}$$

$$D_M = \max_{1 \le i \le n} D_i \tag{31}$$

and

$$\frac{1}{\kappa_D(n)} = \begin{cases} \dfrac{1}{D_M - 1}\left[\dfrac{D_M^n - 1}{D_M - 1} - n\right] & \text{if } D_M \neq 1 \\ \dfrac{n(n-1)}{2} & \text{if } D_M = 1 \end{cases} \tag{32}$$

where n is the number of subsystems. The following theorem holds.

Theorem 1. *In the case specified above, if for some* $p \geq 0$ *the maximum equation error (mismatch)* $\hat{\epsilon}_M$ *in Stage 2 (cf. (23) and (25)) satisfies the relation*

$$\hat{\epsilon}_M \leq \epsilon_0(p) \tag{33}$$

where

$$\epsilon_0(p) = \kappa_D(n)\left[\sqrt{\frac{\delta_M^{*2}}{\Sigma_{c,M}^2} + \frac{\delta_m^{*2}}{\Sigma_{c,M}^2} np} - \frac{\delta_M^*}{\Sigma_{c,M}}\right] \tag{34}$$

then the suboptimality index Δ^* *(cf. (20)) does not exceed* p*, i.e.*

$$\Delta^* \leq p \tag{35}$$

Proof. Denote

$$\psi_{i1} \triangleq E\Big\{r_i(\bar{c}_i)\|K_{iM}(\bar{c}_i, \hat{b}_i) - K_{iM}(\bar{c}_i, b_i^*)\|\Big\}$$

and

$$\psi_{i2} \triangleq E\|K_{iM}(\bar{c}_i, \hat{b}_i) - K_{iM}(\bar{c}_i, b_i^*)\|^2$$

where

$$r_i(\bar{c}_i) = E_{y_i|\bar{c}_i}\|y_i - K_{iM}(\bar{c}_i, b_i^*)\| \tag{36}$$

is the conditional expectation with respect to y_i given $\bar{c}_i$. The following bound is valid

$$q(\hat{b}) - q(b^*) \leq 2\sum_{i=1}^{n}\psi_{i1} + \sum_{i=1}^{n}\psi_{i2} \tag{37}$$

Now the upper bound for the index Δ^* is derived as follows. By (26) we obtain

$$\psi_{i1} \leq \|\hat{b}_i - b_i^*\| E\Big\{r_i(\bar{c}_i)k_i(\bar{c}_i)\Big\}$$

and

$$\psi_{i2} \leq \|\hat{b}_i - b_i^*\|^2 E\Big\{k_i^2(\bar{c}_i)\Big\}$$

Next, by the Schwartz inequality and (28)—including moreover (21), (22), (29), and (36)—some simple calculations give

$$E\Big\{r_i(\bar{c}_i)k_i(\bar{c}_i)\Big\} \leq \delta_i^*\Sigma_{c,i}, \qquad E\Big\{k_i^2(\bar{c}_i)\Big\} \leq \Sigma_{c,i}^2$$

Hence, owing to (9), (20), (22), and (37), and using (24) and (30) it is concluded that

$$\Delta^* \leq 2\frac{\delta_M^*\Sigma_{c,M}}{\delta_m^{*2}}\left(\frac{1}{n}\sum_{i=1}^{n}\|\hat{b}_i - b_i^*\|\right) + \frac{\Sigma_{c,M}^2}{\delta_m^{*2}}\left(\frac{1}{n}\sum_{i=1}^{n}\|\hat{b}_i - b_i^*\|^2\right) \tag{38}$$

On account of (17) and (23) one can notice in turn that

$$\|\hat{b}_i - b_i^*\| \leq \hat{\epsilon}_i + \|\varphi_i(\hat{b}_{i-1}, \hat{a}_i) - \varphi_i(b_{i-1}^*, \hat{a}_i)\|$$

which, owing to (27) and (28), yields the following recursive formula

$$\|\hat{b}_i - b_i^*\| \leq \hat{\epsilon}_i + D_i\|\hat{b}_{i-1} - b_{i-1}^*\|$$

This gives immediately

$$\|\hat{b}_i - b_i^*\| \leq \hat{\epsilon}_i + \sum_{r=2}^{i-1}\Big(\prod_{k=r+1}^{i} D_k\Big)\hat{\epsilon}_r$$

Hence, by applying (25) and (31), after some simple algebra we obtain the bound

$$\|\hat{b}_i - b_i^*\| \leq \begin{cases} \hat{\epsilon}_M\left[\dfrac{D_M^{i-1} - 1}{D_M - 1}\right] & \text{if } D_M \neq 1 \\ \hat{\epsilon}_M(i-1) & \text{if } D_M = 1 \end{cases}$$

Thus, making use of (32), it can be readily verified that

$$\sum_{i=1}^{n}\|\hat{b}_i - b_i^*\| \leq \frac{\hat{\epsilon}_M}{\kappa_D(n)}$$

and

$$\sum_{i=1}^{n}\|\hat{b}_i - b_i^*\|^2 \leq \frac{\hat{\epsilon}_M^2}{\kappa_D^2(n)}$$

The above and (38) results in

$$\Delta^* \leq \frac{1}{n}\left[2\frac{\delta_M^*\Sigma_{c,M}}{\delta_m^{*2}}\Big(\frac{\hat{\epsilon}_M}{\kappa_D(n)}\Big) + \frac{\Sigma_{c,M}^2}{\delta_m^{*2}}\Big(\frac{\hat{\epsilon}_M}{\kappa_D(n)}\Big)^2\right] \tag{39}$$

and provides (33)–(34) and (35) as a simple conclusion. ■

Remark 1. From (39) it follows that, for a given approximation accuracy in Stage 1 (evaluated by means of the errors δ_m^{*2} and δ_M^{*2}; cf. (22) and (24)), the degree of sub-optimality depends critically on the maximum equation error (maximum mismatch) $\hat{\epsilon}_M$ ((23) and (25)) occuring in Stage 2 and, as expected, the smaller the value of $\hat{\epsilon}_M$, the better relative accuracy Δ^* is guaranteed. The rate of deterioration of Δ^* in the neighbourhood of $\hat{\epsilon}_M = 0$, for $\hat{\epsilon}_M$ tending to zero, is at least as fast as $O(\hat{\epsilon}_M)$.

Remark 2. If $\hat{\epsilon}_M = 0$, that is to say, the equation errors $\hat{\epsilon}_i$ in Stage 2 vanish for $i = 1, 2, \ldots, n$ (i.e. for a given collection of b_i^* minimizing the indices $q_i(b_i)$ (eqn. (10)) on the parameter sets B_i (a result of Stage 1) there exist exact solutions to the equations (15) in the parameter sets A_i), then $\Delta^* = 0$ (cf. (39)) and the

corresponding complex model resulting from the two-stage approach is optimal (in this case $b^* \in B_0$—see Section 3). Obviously, the conclusion is valid in general, not only under the assumptions of the theorem.

Remark 3. For $\hat{\epsilon}_i \neq 0$ the degree of suboptimality depends on the choice of the acceptable threshold ϵ_{thr} in the requirement $\hat{\epsilon}_i \leq \epsilon_{thr}$, $i = 1, 2, \ldots, n$, i.e. on specifying the range of admissible equation errors $\hat{\epsilon}_i$, $i = 1, 2, \ldots, n$, in Stage 2. From (39), if the equation errors $\hat{\epsilon}_i$ do not exceed the given threshold value ϵ_{thr}, then the suboptimality index Δ^* is

$$\Delta^* \leq p_0(\epsilon_{thr}, n)$$

where

$$p_0(\epsilon_{thr}, n) = \frac{1}{n}\left[2\frac{\delta_M^* \Sigma_{c,M}}{\delta_m^{*2}}\left(\frac{\epsilon_{thr}}{\kappa_D(n)}\right) + \frac{\Sigma_{c,M}^2}{\delta_m^{*2}}\left(\frac{\epsilon_{thr}}{\kappa_D(n)}\right)^2\right] \tag{40}$$

is the maximum relative inaccuracy of the model, guaranteed for given ϵ_{thr}. If in particular $\epsilon_{thr} = \epsilon_0(p)$, where $\epsilon_0(p)$ is as in (34) with $p \geq 0$, then we have $\Delta^* \leq p$ (see Theorem 1).

Notice that for $n = 1$ (a single-element system) $1/\kappa_D(1) = 0$ (cf. (32)) and hence $p_0(\epsilon_{thr}, 1) = 0$ for each choice of ϵ_{thr}, which confirms that the two-stage approach is then optimal (certainly, in such a case $\hat{a}_1 = a_1^* = b_1^*$). For $n = 2$ (a tandem system) we have $1/\kappa_D(2) = 1$ and for the respective ϵ_{thr} the warranted maximum relative error is

$$p_0(\epsilon_{thr}, 2) = \frac{\Sigma_{c,M}}{\delta_m^*}\left[\frac{\delta_M^*}{\delta_m^*} + \frac{1}{2}\frac{\Sigma_{c,M}}{\delta_m^*}\epsilon_{thr}\right]\epsilon_{thr}$$

The corresponding test threshold $\epsilon_0(p)$ which guarantees $\Delta^* \leq p$ is then

$$\epsilon_0(p) = \sqrt{\frac{\delta_M^{*2}}{\Sigma_{c,M}^2} + 2\frac{\delta_m^{*2}}{\Sigma_{c,M}^2}p} - \frac{\delta_M^*}{\Sigma_{c,M}}$$

Summarizing, from the obtained 'theoretical' (derived under the assumption of full probabilistic knowledge of the system) bounds on Δ^* ((33)–(35) and (40)) it follows that for moderate values of equation errors $\hat{\epsilon}_i$ in Stage 2 and properly 'regular' system models (in the sense of the fulfilment of conditions (26)–(28)) the relative error Δ^* may not exceed a reasonable limit (Remark 3) and then the 'approximate' model, obtained by the two-stage approach, and the 'desired' optimum system model are almost equivalent. It should be here emphasized (which can be important for applications) that in realistic conditions, when only measurements from the system are available, the actual value of the suboptimality index Δ^* can be estimated empirically based on the appropriate set of measurements at hand, in each particular case, not only in that specified for the purposes of the theorem (see Section 2 in Part 2 of the paper).

5. Examples

The special cases examined in the following two examples are to gain some insight into the question of the existence of solutions to eqns. (15) (Remark 2) and to provide a clear statistical interpretation of the quantities δ_m^{*2}, δ_M^{*2} and $\Sigma_{c,M}^2$ appearing in the bounds (34) and (40) in Section 4. They will be further developed in Part 2 of the paper.

Example 1. Assume that the components of the complex model (2) are linear, i.e.

$$y_{1M} = B_1 c_1, \quad y_{iM} = A_i u_{iM} + B_i c_i, \qquad i = 2, 3, \ldots, n \tag{41}$$

and the constraints on model parameters (A_i, B_i) are negligible (i.e. the set of admissible model parameters can be arbitrarily large). Then eqn. (15) takes the form

$$[A_i \bar{K}_{i-1,M}^*, B_i] = \bar{K}_{iM}^* \tag{42}$$

where $\bar{K}_{i-1,M}^*$ and $\bar{K}_{iM}^*$ denote optimum, i.e. minimizing (10), parameters of the linear aggregated models (cf. (6) in Section 2)

$$y_{i-1,M} = \bar{K}_{i-1,M} \bar{c}_{i-1}, \qquad y_{iM} = \bar{K}_{iM} \bar{c}_i$$

of the aggregates AG_{i-1} and AG_i of the system (with inputs $\bar{c}_{i-1}$ and $\bar{c}_i = (\bar{c}_{i-1}, c_i)$, and outputs y_{i-1} and y_i, respectively), i.e.

$$\bar{K}_{i-1,M}^* = (E y_{i-1} \bar{c}_{i-1}^T)(E \bar{c}_{i-1} \bar{c}_{i-1}^T)^{-1}$$

$$\bar{K}_{iM}^* = (E y_i \bar{c}_i^T)(E \bar{c}_i \bar{c}_i^T)^{-1}$$

provided that the respective inverses exist. Obviously, a solution to (42) with respect to (A_i, B_i) does exist if

$$\operatorname{rank} \left[\bar{K}_{i-1,M}^{*T}, \bar{K}_{iM}^{(i)*T} \right]^T = \operatorname{rank} \bar{K}_{i-1,M}^* \tag{43}$$

where $\bar{K}_{iM}^{(i)*}$ is the block of $\bar{K}_{iM}^*$ related to the vector $\bar{c}_{i-1}$. If external excitations c_i of consecutive subsystems S_i are stochastically independent, we have

$$\bar{K}_{iM}^* = \left[(E y_i \bar{c}_{i-1}^T)(E \bar{c}_{i-1} \bar{c}_{i-1}^T)^{-1}, (E y_i c_i^T)(E c_i c_i^T)^{-1} \right]$$

that is

$$\bar{K}_{iM}^{(i)*} = (E y_i \bar{c}_{i-1}^T)(E \bar{c}_{i-1} \bar{c}_{i-1}^T)^{-1}$$

and one can easily ascertain that the rank condition (43) is satisfied when the vectors y_{i-1} and y_i are linearly dependent in the cascade system, i.e. the true descriptions F_i^* of the elements S_i are linear in the interaction inputs (cf. (1)). Then the linear series model provided by the two-stage approach is optimal (cf. Remark 2).

Example 2. Let the cascade system (1) be described as follows (a linear-in-the-interactions system):

$$y_1 = F_1^*(c_1, z_1) = f_1^*(c_1) + z_1$$

$$y_i = F_i^*(c_i, u_i, z_i) = f_i^*(c_i) + u_i + z_i, \quad u_i = y_{i-1}, \quad i = 2, 3, \ldots, n$$

which means that the respective aggregates AG_i are now represented by the mappings

$$y_i = \bar{F}_i^*(\bar{c}_i) + \bar{z}_i$$

where

$$\bar{F}_i^*(\bar{c}_i) = \sum_{r=1}^{i} f_r^*(c_r), \quad \bar{z}_i = \sum_{r=1}^{i} z_r$$

and let us assume exceptionally in this example (for an easy interpretation of factors) that in the absence of noise ($\bar{z}_i = 0$) the exact descriptions $\bar{F}_i^*$ of AG_i are accessible within the model sets (cf. (6) and further discussion in Section 2)

$$\left\{ y_{iM} = K_{iM}(\bar{c}_i, b_i), \quad b_i \in B_i \right\}, \qquad i = 1, 2, \ldots, n$$

corresponding to the preassumed collection of models (2). Since $E\bar{z}_i = 0$ (cf. assumption (b) in Section 2), the mean-squared optimum model, minimizing (10) in Stage 1, is as follows:

$$K_{iM}(\bar{c}_i, b_i^*) = \bar{F}_i^*(\bar{c}_i)$$

and in effect (cf. (22) and (21))

$$\delta_i^{*2} = E\|\bar{z}_i\|^2 = \sum_{r=1}^{i} E\|z_r\|^2 = \sum_{r=1}^{i} \sigma_{z,r}^2 \triangleq \bar{\sigma}_{z,i}^2 \tag{44}$$

i.e. the discrepancy δ_i^{*2} becomes actually equal to the aggregated (cumulative) variance of the noise $\bar{z}_i$ disturbing the aggregate AG_i. Therefore we have (cf. (24))

$$\delta_m^{*2} = \delta_1^{*2} = \sigma_{z,1}^2 \triangleq \bar{\sigma}_{z,1}^2, \qquad \delta_M^{*2} = \delta_n^{*2} = \sum_{r=1}^{n} \sigma_{z,r}^2 \triangleq \bar{\sigma}_{z,n}^2$$

i.e. the errors δ_m^{*2} and δ_M^{*2} from Stage 1 are in this case cumulative variances of the corresponding noises $\bar{z}_1$ and $\bar{z}_n$ influencing the two extreme aggregates of the cascade, AG_1 and AG_n. Denoting, analogously to (21), $\sigma_{c,r}^2 \triangleq E\|c_r\|^2$ one can recognize that in (21) we have now

$$\bar{\sigma}_{c,i}^2 \triangleq E\|\bar{c}_i\|^2 = \sum_{r=1}^{i} \sigma_{c,r}^2$$

Hence (cf. (29))

$$\Sigma_{c,i}^2 = \alpha_i^2 \sum_{r=1}^{i} \sigma_{c,r}^2 \triangleq \bar{\sigma}_{c,\alpha,i}^2$$

is, for $E\bar{c}_i = 0$, simply the weighted cumulative variance of the external input vector $\bar{c}_i$ of the aggregate AG_i, with the squared weighting coefficient α_i given by (28). Hence (see (30))

$$\Sigma_{c,M}^2 = \max_{1 \le i \le n} \{\bar{\sigma}_{c,\alpha,i}^2\}$$

represents now the dominant scaled variance in the set of external excitations $\bar{c}_i$ applied to the aggregates AG_i in the complex system. In the particular case when for each subsystem S_i the respective variances $\sigma_{z,i}^2$ and $\sigma_{c,i}^2$ are as follows:

$$\sigma_{z,i}^2 = \sigma_z^2, \quad \sigma_{c,i}^2 = \sigma_c^2, \qquad i = 1, 2, \ldots, n \quad (\sigma_z, \sigma_c > 0)$$

and simultaneously

$$\alpha_m \le \alpha_i \le \alpha_M, \qquad i = 1, 2, \ldots, n$$

for some $\alpha_m, \alpha_M > 0$, we obtain that plainly:

$$\delta_m^{*2} = \sigma_z^2, \qquad \delta_M^{*2} = n\sigma_z^2$$

and

$$n\alpha_m^2 \sigma_c^2 \le \Sigma_{c,M}^2 \le n\alpha_M^2 \sigma_c^2$$

In this case, the threshold value $\epsilon_0(p)$ in (34) and the maximum relative inaccuracy $p_0(\hat{\epsilon}_M, n)$ for a given $\hat{\epsilon}_M$ in Stage 2 ((40) for $\epsilon_{thr} = \hat{\epsilon}_M$) are such that

$$\left(\frac{\kappa_D(n)}{\alpha_M}\right)\left(\frac{\sigma_z}{\sigma_c}\right)\left[\sqrt{1+p} - 1\right] \le \epsilon_0(p) \le \left(\frac{\kappa_D(n)}{\alpha_m}\right)\left(\frac{\sigma_z}{\sigma_c}\right)\left[\sqrt{1+p} - 1\right] \tag{45}$$

and

$$2\alpha_m\left(\frac{\sigma_c}{\sigma_z}\right)\left(\frac{\hat{\epsilon}_M}{\kappa_D(n)}\right) + \alpha_m^2\left(\frac{\sigma_c}{\sigma_z}\right)^2\left(\frac{\hat{\epsilon}_M}{\kappa_D(n)}\right)^2 \le p_0(\hat{\epsilon}_M, n)$$

$$\le 2\alpha_M\left(\frac{\sigma_c}{\sigma_z}\right)\left(\frac{\hat{\epsilon}_M}{\kappa_D(n)}\right) + \alpha_M^2\left(\frac{\sigma_c}{\sigma_z}\right)^2\left(\frac{\hat{\epsilon}_M}{\kappa_D(n)}\right)^2 \tag{46}$$

Moreover, the degree of suboptimality of the two-stage approach depends then on the ratio σ_c/σ_z. The smaller the value of σ_c/σ_z, the less restrictive the requirement (33) on the equation error $\hat{\epsilon}_M$ becomes to attain $\Delta^* \le p$ for given $p > 0$ ((45) and Theorem 1) and the better the relative accuracy of the resulting 'approximate' model can be for the given error (mismatch) $\hat{\epsilon}_M$ in Stage 2 (cf. (46) and Remark 3). Such a

conclusion is not surprising since for large noise (large σ_z) the reference value $q(b^*)$ in the suboptimality index (20), actually of the form (cf. (9)–(10), (22) and (44))

$$q(b^*) = \sigma_z^2 \frac{n(n+1)}{2}$$

is large and thus the loss $q(\hat{b}) - q(b_0^*)$ in the model quality, associated with the two-stage approach, may be less significant. For a tandem system, i.e. for $n = 2$ (then $\hat{\epsilon}_M = \hat{\epsilon}_2, \kappa_D(2) = 1$), the bounds (45) and (46) take the special form

$$\left(\frac{1}{\alpha_M}\right)\left(\frac{\sigma_z}{\sigma_c}\right)\left[\sqrt{1+p}-1\right] \leq \epsilon_0(p) \leq \left(\frac{1}{\alpha_m}\right)\left(\frac{\sigma_z}{\sigma_c}\right)\left[\sqrt{1+p}-1\right]$$

and

$$\left(\frac{\sigma_c}{\sigma_z}\right)(\alpha_m\hat{\epsilon}_2)\left[2+\left(\frac{\sigma_c}{\sigma_z}\right)(\alpha_m\hat{\epsilon}_2)\right] \leq p_0(\hat{\epsilon}_2, 2) \leq \left(\frac{\sigma_c}{\sigma_z}\right)(\alpha_M\hat{\epsilon}_2)\left[2+\left(\frac{\sigma_c}{\sigma_z}\right)(\alpha_M\hat{\epsilon}_2)\right]$$

and the extent of the influence of the ratio σ_c/σ_z on Δ^* depends on the size of weighted equation errors $\alpha_m\hat{\epsilon}_2$ and $\alpha_M\hat{\epsilon}_2$.

6. Final Remarks

A new two-stage 'approximate' algorithm for series system identification has been introduced as an alternative to the 'exact' approach. The motivation was to provide a more handy procedure for complex system identification, taking into account interconnections existing in the system. The proposed scheme exploits in a very natural fashion the particular series structure of the system and is based on progressive aggregation (in Stage 1) and disaggregation (in Stage 2) of the models assumed for individual system components, reducing in this way the usually hard problem of interconnected multicomponent system identification to a set of independent standard identification tasks of single-element systems (aggregates), solved at Stage 1, and a set of complementary local optimization (disaggregation) tasks, solved in a decentralized manner at Stage 2. As a result, a decomposed structural model of a series system is obtained as the main solution and a set of auxiliary (optimum) aggregated models of subsystem complexes (aggregates) as a by-product of the identification procedure. Taking into account the structure of the algorithm, on the one hand it can be included into the class of indirect multistage identification methods (Hsia, 1977; Söderström and Stoica, 1989), since the identification routine is in fact indirect and identification in Stage 1 produces merely the data for the final identification of the system model in Stage 2. On the other hand, the identification task within the approach is as if it was split into two layers, and at the higher layer (Stage 1) the models of large, aggregated parts of the system (subsystem complexes) are established determining in this way the targets for identification at the lower layer (Stage 2); in turn, the models of more detailed system components are computed there so as to achieve the reference (target) model set by the upper layer (a kind of 'follow-up' identification). From this point of view, the approach can thus be traced back to basic concepts of multilayer theory.

In this paper, a theoretical background for the method has been given together with the discussion of accuracy of the resulting 'approximate' complex model—all under the idealistic assumption of full probabilistic knowledge of the system. As regards realistic conditions, in Part 2 an empirical counterpart of the two-stage algorithm is considered, making use only of measurements obtained in the system.

References

Brdyś M., Abdullah N. and Roberts P.D. (1990): *Augmented model-based double iterative loop techniques for hierarchical control of complex industrial processes.* — Int. J. Control, v.52, No.3, pp.549–570.

Brdyś M., Chen S. and Roberts P.D. (1986): *An extension to the modified two-step algorithm for steady-state system optimization and parameter estimation.* — Int. J. Systems Sci., v.17, No.8, pp.1229–1243.

Brdyś M., Findeisen W. and Tatjewski P. (1980): *Hierarchical control for systems operating in steady-state.* — Large Scale Systems, v.1, pp.193–213.

Brdyś M. and Roberts P.D. (1986): *Optimal structures for steady-state adaptive optimizing control of large-scale industrial processes.* —- Int. J. Systems Sci., v.17, No.10, pp.1449–1474.

Bubnicki Z. (1975): *Global and local identification of complex systems with cascade structure.* — Systems Science, v.1, No.1, pp.55–65.

Chen S., Brdyś M. and Roberts P.D. (1986): *An integrated system optimization and parameter estimation technique for hierarchical control of steady-state systems.* — Int. J. Systems Sci., v.17, No.8, pp.1209–1228.

Findeisen W., Bailey F.N., Brdyś M., Malinowski K., Tatjewski P. and Woźniak A. (1980): *Control and Coordination in Hierarchical Systems.* — New York: Wiley.

Findeisen W., Brdyś M., Malinowski K., Tatjewski P. and Woźniak A. (1978): *On-line hierarchical control for steady-state systems.* — IEEE Trans. Automat. Contr., v.23, No.2, pp.189–209.

Forsythe G.E., Malcolm M.A. and Moler C.B. (1977): *Computer Methods for Mathematical Computations.* — Englewood Cliffs, N.J.: Prentice-Hall.

Dennis J.E. and Schnabel R.B. (1983): *Numerical Methods for Unconstrained Optimization and Nonlinear Equations.* — Englewood Cliffs, N.J.: Prentice-Hall.

Hasiewicz Z. (1987): *Identifiability of large-scale interconnected linear zero-memory systems.* — Int. J. Systems Science, v.18, No.4, pp.649–664.

Hasiewicz Z. (1988a): *Two-stage procedure for parameter estimation in large-scale interconnected linear zero-memory systems.* — Int. J. Systems Sci., v.19, No.3, pp.497–516.

Hasiewicz Z. (1988b): *Identifiability analysis of interconnected zero-memory composite systems.* — Int. J. Systems Sci., v.19, No.9, pp.1731-1748.

Hasiewicz Z. (1989): *Applicability of least-squares to the parameter estimation of large-scale no-memory linear composite systems.* — Int. J. Systems Sci., v.20, No.12, pp.2427–2449.

Hasiewicz Z. and Stankiewicz A. (1984a): *On a two-level identification scheme of large-scale systems for open-loop steady-state optimization.* — Int. J. Systems Sci., v.15, No.10, pp.1111–1130.

Hasiewicz Z. and Stankiewicz A. (1984b): *On decentralized identification of large-scale interconnected systems with coordination by price mechanizm.* — Proc. 3rd Int. Conf. *Systems Engineering*, Dayton, pp.67–75.

Hasiewicz Z. and Stankiewicz A. (1986a): *On applicability of interaction balance method to global identification of interconnected steady-state systems.* — IEEE Trans. Automat. Contr., v.31, No.1, pp.77–80.

Hasiewicz Z. and Stankiewicz A. (1986b): *Two-stage scheme for global identification of interconnected steady-state systems.* — Int. J. Systems Sci., v.17, No.11, pp.1635–1643.

Hsia T.C. (1977): *Identification: Least Squares Methods.* — Lexington: Lexington Books.

Lawson C.L. and Hanson R.J. (1974): *Solving Least-Squares Problems.* — Englewood Cliffs, N.J.: Prentice-Hall.

Lin J., Amini Z. and Roberts P.D. (1991): *New coordination strategy for on-line hierarchical optimizing control of large scale industrial processes.* — Int. J. Control, v.54, No.6, pp.1075–1096.

Lin J., Han C., Roberts P.D. and Wan B. (1989): *New approach to stochastic optimizing control of steady-state systems using dynamic information.* — Int. J. Control, v.50, No.6, pp.2205–2235.

Lin J., Hendawy Z.M. and Roberts P.D. (1988a): *Extension of integrated system optimization and parameter estimation to hierarchical control of steady-state systems with output dependent constraints.* — Int. J. Control, v.47, No.2, pp.413–431;

Lin J., Hendawy Z.M. and Roberts P.D. (1988b): *New model-based double-loop iterative strategy for integrated system optimization and parameter estimation of large-scale industrial processes.* — Int. J. Control, v.47, No.3, pp.753–773.

Michalska H., Ellis J.E. and Roberts P.D. (1985): *Joint coordination method for the steady-state control of large-scale systems.* — Int. J. Systems Sci., v.16, No.5, pp.605–618.

Shao F.Q. and Roberts P.D. (1983): *A price correction mechanism with global feedback for hierarchical control of steady-state systems.* — Large Scale Systems, v.4, No.1, pp.67–80.

Shao F.Q. and Roberts P.D. (1985): *Augmented decentralized control with local feedback for steady-state systems.* — Int. J. Systems Sci., v.16, No.5, pp.649–665.

Söderström T. and Stoica P. (1989): *System Identification.* — Englewood Cliffs, N.J.: Prentice-Hall.

Tatjewski P., Abdullah N. and Roberts P.D. (1990): *Comparative study and development of integrated optimization and parameter estimation algorithms for hierarchical steady-state control.* — Int. J. Control, v.51, No.2, pp.421–443.

Zhang H. and Roberts P.D. (1991): *Integrated system optimization and parameter estimation using a general form of steady-state model.* — Int. J. Systems Sci., v.22, No.11, pp.1679–1693.

Received: August 14, 1995
Revised: January 21, 1996