Ewaryst Rafajłowicz

Sequential identification algorithm and controller choice for a certain class of distributed systems

Kybernetika, Vol. 22 (1986), No. 6, 471--486

Persistent URL: http://dml.cz/dmlcz/124449

SEQUENTIAL IDENTIFICATION ALGORITHM AND CONTROLLER CHOICE FOR A CERTAIN CLASS OF DISTRIBUTED SYSTEMS

EWARYST RAFAJŁOWICZ

In the paper a computationally simple and reliable algorithm for parameter estimation in stochastic distributed-parameter systems (DPS) is proposed. In opposite to most of earlier proposed algorithms it is of noniterative type, but it is applicable to a class of linear DPS with eigenfunctions not depending on unknown parameters. Application of the sequential approach in the proposed algorithm resulted in nontypical for parameter estimators in DPS properties. Namely, the estimators are shown to be unbiased and efficient for a finite but random observation time. Extensive simulation studies confirm computational and theoretical advantages of the algorithm and indicate that modifications of the sequential approach, proposed in the paper, essentially reduce mean value of the observation time.

1. INTRODUCTION

In recent years several identification algorithm for distributed parameter systems (DPS) have been proposed (see [5], [8], [9] for survey papers). They are designed in such a way that they try to cover as large class of systems as possible. Such attempts resulted in computationally complicated algorithms of a searching type. For this reason they are not always reliable in finding estimates and their properties are difficult to investigate. Even if statistical properties of estimators such as obtained by output least squares minimization or maximum likelihood estimators are investigated (see [2]) it occurs that they are only asymptotically unbiased and efficient. However nothing can be said about their behaviour for a finite observation time. Also experiment design considerations for such estimators are based on asymptotic reasoning, which utilize the inverse of the information matrix instead of covariance matrix (see e.g. [4], [7], [10], [11]).

These circumstance have inspired the author to try to design an easily implementable and noniterative identification algorithm for a more narrow class of DPS (see Section 2 for details). This resulted in a prototype algorithm based on modal decomposition, presented in Section 4. However, finite sample accuracy of this algo-

rithm is also difficult to investigate. For this reason, the prototype algorithm is modified in Section 5 by applying an idea of the sequential estimation. The sequential approach to hypothesis testing has been introduced by Wald an then it has been developed in a spirit of the decision theory (see [13] and the bibliography therein).

More recently, Shiryayev [6] for the first time applied the sequential approach to the parameter estimation problem in lumped parameter system (LPS), described by the first order stochastic differential equation (see Section 6). His theoretically appealing results could not find way to practical applications since a random observation time is usually too long for stable systems (see Section 7, where this opinion is confirmed by a simulation study). For this reason two modifications of his approach are proposed in Section 5. Firstly, nonzero set points of modal controllers allow to obtain shorter random observation times for negative feedback gains, and secondly identifieries of each mode measurements are allowed to have their own random processing times. Combining these modifications with the prototype algorithm of Section 4 we get the sequential estimators which are the best in the mean square sense in the class of all unbiased estimators with finite variances (see Section 6) and these properties hold for a finite but random observation time. Extensive simulation studies, partly reported in Section 7, confirm these theoretical advantages.

2. SYSTEM DESCRIPTION

Let H be a separable Hilbert space. In the sequel we consider a system described by the following abstract stochastic differential equation:

$$\mathrm{d}q + A(a)\, q(t) \,.\, \mathrm{d}t - u(t)\, \mathrm{d}t = \mathrm{d}w\,, \quad t \in (0, T) \tag{2.1}$$

with the initial condition $q(0) = q_0$. Above, $q(t) \in H$, $u(t) \in H$, $w(t) \in H$; $t \in (0, T)$ denotes a system state, an input signal and a disturbance, respectively.

We shall use the same framework as in [3] in understanding the meaning of a solution of (2.1). In (2.1), $A(a)$ is a closed, densely defined linear operator $A(a)$: $D_A \to H$. It is assumed that $A(a)$ is of the form:

$$A(a)\, h = Q_0 h + \sum_{i=1}^{r} a_i \,.\, Q_i h\,, \quad h \in D_A \tag{2.2}$$

where Q_i: $D_A \to H$, $i = 0, 1, 2, \ldots, r$, are given linearly independent operators, while $a = [a_1, a_2, \ldots, a_r]'$ is a vector of unknown constant parameters to be estimated. Furthermore, we assume that $A(a)$ is the infinitesimal generator of a strongly continuous semigroup $T_t(a)$ and that $A(a)$ is selfadjoint with compact resolvent (see e.g. [1] for definitions), which imply (see e.g. [1]) that $A(a)$ possesses, orthonormal and complete in H, set of eigenfunctions $v_k \in H$; $k = 1, 2, \ldots$. These assumptions also imply that $T_t(a)$ has the representation

$$T_t(a)\, h = \sum_{k=1}^{\infty} \mathrm{e}^{[-\lambda_k(a) t]} \,.\, \langle v_k \,.\, h \rangle \,.\, v_k \tag{2.3}$$

where $\lambda_k(a)$, $k = 1, 2, \ldots$, are the eigenvalues of $A(a)$ associated with v_k, $k = 1, 2, \ldots$; i.e. $A(a)\, v_k = \lambda_k(a)\, v_k$, while $\langle \cdot, \cdot \rangle$ denotes the inner product in H.

In general, the eigenfunctions of $A(a)$ may depend on a, but we restrict a class of considered systems to those, for which v_k, $k = 1, 2, \ldots$, do not depend on the vector a in the sense that

$$\frac{\partial v_k}{\partial a_i} = 0\,, \quad i = 1, 2, \ldots, r\,; \quad k = 1, 2, \ldots. \tag{2.4}$$

The disturbance $w(\cdot)$ is assumed to be an H-valued Wiener process (see [3] for definitions) which has the representation:

$$w(t) = \sum_{k=1}^{\infty} \beta_k(t)\,.\, v_k \quad \text{with probability 1}\,, \tag{2.5}$$

where $\beta_k(t)$, $k = 1, 2, \ldots$, are mutually independent real Wiener processes with incremental covariances γ_k, $k = 1, 2, \ldots$, and

$$\sum_{k=1}^{\infty} \gamma_k\,.\,\lambda_k(a) < \infty\,. \tag{2.6}$$

The initial condition q_0 and $w(\cdot)$ are required to be independent and the coefficients q_{0k}, $k = 1, 2, \ldots$, in the following representation

$$q_0 = \sum_{k=1}^{\infty} q_{0k} v_k$$

are assumed to be mutually independent random variables. Concerning the input signal $u(\cdot)$ it is assumed that for every $s, t > 0$, $(T_t(a)\, u(s)) \in D_A$ and

$$\sum_{k=1}^{\infty} \lambda_k(a) \int_0^t |u_k(s)|\, \mathrm{d}s < \infty \tag{2.7}$$

where $u_k(t)$, $k = 1, 2, \ldots$, are the coefficients in the representation

$$u(t) = \sum_{k=1}^{\infty} u_k(t)\,.\, v_k\,. \tag{2.8}$$

Using the above assumptions it is easy to verify that (2.6) and (2.7) imply conditions (5.41), (5.42) of Theorem 5.35 in [3] to hold. This means that the unique strong solution of (2.1) exists and it is of the form

$$q(t) = T_t(a)\, q_0 + \int_0^t T_{t-s}(a)\, u(s)\, \mathrm{d}s + \int_0^t T_{t-s}(a)\, \mathrm{d}w(s)\,. \tag{2.9}$$

Remark 1. Taking into account that a is unknown we must require the above conditions to hold for every vector a from a certain set of admissible parameters.

3. EXPERIMENT CONDITIONS

We assume that the vector a is estimated from measurements of the form

$$y_k(t) = \langle q(t), v_k \rangle , \quad k = 1, 2, \ldots, K , \quad t \in [0, T] \tag{3.1}$$

where $K < \infty$ is a given number. This form of measurements implicitly assumes that all measurement and modelling errors are included in the disturbance $w(\cdot)$ in (2.1).

In practice, measurements are possible in a finite number of spatial points and thus integrals in (3.1) have to be approximated by a quadrature type formulas. Taking into account that, for carefully selected measurement points and weights, quadrature formulas are very accurate, we take measurements (3.1) for further considerations.

Concerning the input signal $u(\cdot)$ we assume that if it is an open loop control then at least the functions $u_k(t) = \langle u(t), v_k \rangle$, $k = 1, 2, \ldots, K$, $t \in (0, T)$ are precisely known (possible errors being again included into $w(\cdot)$). In Sections 5, 6, 7 we also consider a feedback type control. In this case we assume that the feedback law is exactly known and it is nonanticipating, i.e. $u(t)$ is measurable with respect to sigma field generated by $q(s)$, $0 \leqq s \leqq t$.

4. PROTOTYPE OF THE IDENTIFICATION ALGORITHM

From (2.2), (2.9) and (3.1) one can see that application of the output least squares as an identification criterion would lead to a complicated, iterative type optimization algorithm. Also application of the maximum likelihood approach, although possible, yields an estimator, which is difficult to investigate. For this reason a simple, non-iterative type algorithm is proposed.

The suggested algorithm uses minimization of the least squares twice. At the first step the estimates $\hat{\lambda}_k$, $k = 1, 2, \ldots, K$, of the values $\lambda_k(a)$, $k = 1, 2, \ldots, K$, are obtained. At the second step, the least square criterion is used to derive the estimate $\hat{a}$ by minimizing

$$I_2(a) = \sum_{k=1}^{K} (\hat{\lambda}_k - \lambda_k(a))^2 \tag{4.1}$$

with respect to $a \in \mathbb{R}^r$.

A crucial point in performing the second step lies in the fact that (2.4) implies the affine dependence of $\lambda_k(a)$ on the vector a. More precisely, we have:

Lemma 1. If condition (2.4) holds and the eigenfunctions of $A(a)$ are such that

$$\frac{\partial}{\partial a_j} Q_i v_k = Q_i \frac{\partial v_k}{\partial a_j} ; \quad i, j = 1, 2, \ldots, r , \quad k = 1, 2, \ldots \tag{4.2}$$

then there exists a sequence of vectors $\alpha_k = [\alpha_k^{(1)}, \alpha_k^{(2)}, \ldots, \alpha_k^{(r)}]'$, and numbers $\alpha_k^{(0)}$,

$k = 1, 2, \ldots$, for which

$$\lambda_k(a) = \alpha_k^{(0)} + \alpha_k' \,.\, a\,, \quad k = 1, 2, \ldots, \tag{4.3}$$

where prime "′" denotes transposition.

Furthermore, every v_k is simultaneously the eigenfunction of all the operators Q_i, $i = 0, 1, \ldots, r$, and

$$Q_i v_k = \alpha_k^{(i)} v_k\,, \quad i = 0, 1, \ldots, r\,; \quad k = 1, 2, \ldots. \tag{4.4}$$

Proof. Assumption (4.2) allows us to differentiate both sides of the equality $A(a)\, v_k = \lambda_k(a)\, v_k$ with respect to a_j, $j = 1, 2, \ldots, r$. These yield

$$A(a) \frac{\partial v_k}{\partial a_j} = \lambda_k(a) \frac{\partial v_k}{\partial a_j} + \frac{\partial \lambda_k}{\partial a_j} v_k - Q_j v_k\,, \quad j = 1, 2, \ldots, r \tag{4.5}$$

and together with (2.4) we obtain

$$\frac{\partial \lambda_k}{\partial a_j} v_k = Q_j v_k\,, \quad j = 1, 2, \ldots, r\,. \tag{4.6}$$

Taking into account that neither Q_j, nor v_k depend on a, we conclude from (4.6) that $\partial_k/\partial a_j$ is a constant not depending on a.

Denoting $\alpha_k^{(j)} = \partial_k/\partial a_j$ one can see from (4.6) that (4.4) holds for $j = 1, 2, \ldots, r$. Using this fact we get from the definition of v_k and

$$Q_0 v_k = \left(\lambda_k(a) - \sum_{i=1}^{r} \alpha_k^{(i)} \,.\, a_i\right) v_k\,. \tag{4.7}$$

Again Q_0 and v_k do not depend on a and thus $\alpha_k^{(0)} \triangleq \lambda_k(a) - \alpha_k' a$ must be a constant, which does not depend on a. This statement proves (4.4) for $i = 0$ and simultaneously justifies (4.3). □

In order to present the identification algorithm let us take inner products of (2.1) and v_k, $k = 1, 2, \ldots, K$ (existence of and uniqueness of the solution of (2.1) justifies these operations). Using (2.5), (2.8) and the data (3.1) these yield:

$$\mathrm{d}y_k + \lambda_k(a) \,.\, y_k \,.\, \mathrm{d}t = u_k \,.\, \mathrm{d}t + \mathrm{d}\beta_k\,, \quad k = 1, 2, \ldots, K \tag{4.8}$$

with given initial conditions $y_k(0) = q_{k0}$, $k = 1, 2, \ldots, K$.

Remark 2. The assumptions made in Section 2 imply that $y_k(t)$, $t \in [0, T]$ are mutually independent stochastic processes.

Formulas (4.3) and (4.8) are crucial for the following prototype of the proposed identification algorithm.

Step 1: Basing on (4.8) find estimates $\hat{\lambda}_k^T$ of $\lambda_k(a)$ according to the formulas:

$$\hat{\lambda}_k^T = -\left[\int_0^T y_k \,\mathrm{d}y_k + \int_0^T y_k u_k \,\mathrm{d}t\right] \Big/ \int_0^T y_k^2 \,\mathrm{d}t\,, \quad k = 1, 2, \ldots, K. \tag{4.9}$$

Step 2: Using $\hat{\lambda}_k^T$, $k = 1, 2, \ldots, K$, and (4.3) find the estimate $\hat{a}_T$ by solving the system of normal equations

$$Ba = \sum_{k=1}^{K} (\hat{\lambda}_k^T - \alpha_k^{(0)})\, \alpha_k \tag{4.10}$$

with respect to $a \in \mathbb{R}^r$, where $B \triangleq \sum_{k=1}^{K} \alpha_k \alpha_k'$.

One can notice that (4.9) are the maximum likelihood estimates of $\lambda_k(a)$ (see [6] for a derivation of the likelihood ratio), which coincide with the least squares estimates due to the Gaussian distributions of the Wiener processes $\beta_k(\cdot)$, $k = 1, 2, \ldots, K$. We also remark that the system (4.10) always has a solution, since $\hat{a}_T$ minimizes $I_2(a)$. Uniqueness of $\hat{a}_T$ can be assured if the matrix $[\alpha_1, \alpha_2, \ldots, \alpha_K]$ is of the rank r (see [12] for the proof), what can be achieved by suitable choice of the modes $v_1, v_2, \ldots, v_K$ and it is further assumed.

5. ESTIMATION ACCURACY AND SEQUENTIAL IDENTIFICATION ALGORITHM

From Remark 2 and (4.9) it follows that the estimators $\hat{\lambda}_k^T$, $k = 1, 2, \ldots, K$, are mutually independent random variables. For this reason the covariance matrix of $\hat{a}_T$ is given by

$$\operatorname{cov}(\hat{a}_T) = C \sum_{k,m=1}^{K} \alpha_k \alpha_m' \, \mathsf{E}[(\hat{\lambda}_k^T - \mathsf{E}\hat{\lambda}_k^T)(\hat{\lambda}_m^T - \mathsf{E}\hat{\lambda}_m^T)]\, C = \tag{5.1}$$

$$= C[\sum_{k=1}^{K} \alpha_k \alpha_k' \operatorname{var}(\hat{\lambda}_k^T)]\, C\,,$$

where $C = B^{-1}$, while var $(\cdot)$ means the variance. Taking the arithmetic mean of parameter estimate variances as a global measure of the estimation accuracy we obtain (5.1):

$$J_T(u) = \operatorname{tr}[\operatorname{cov}(\hat{a}_T)] = \sum_{k=1}^{K} c_k \operatorname{var}(\hat{\lambda}_k^T)\,, \tag{5.2}$$

where the symbol $J_T(u)$ is used in order to indicate the dependence of the estimation accuracy on the input signal u and the observation time T. In (5.2) $c_k = \alpha_k' C^2 \alpha_k$, $k = 1, 2, \ldots, K$, are positive constants depending only on the numbers of modes that are measured.

It is a common practice in experiment design for identification of LPS (see e.g. [4], [7]) as of DPS (see e.g. [10], [11]) to use the Cramér-Rao lower bound (CRLB) for evaluation of an estimation accuracy instead of exact but complicated expressions. We shall adopt the same approach here. Assuming that for a large observation time T the estimation bias is negligible we have the following CRLB (see [6] for its

derivation)

$$\text{var}\,(\hat{\lambda}_k^T) \geqq 1/\mathsf{E}\int_0^T y_k^2(t)\,\mathrm{d}t\,, \quad k = 1, 2, \ldots, K\,. \tag{5.3}$$

Hence, (5.2) yield

$$J_T(u) \geqq \sum_{k=1}^{K} c_k/\mathsf{E}\left[\int_0^T y_k^2(t)\,\mathrm{d}t\right] \tag{5.4}$$

and one can see that input signals which produce high energies of the modes 1, 2, ..., K are desirable from the estimation accuracy point of view. On the other hand, behaviour of higher modes does not influence the estimation accuracy and thus we shall exclude them from further consideration assuming that they are controlled in a certain way, e.g. by the minimum variance control law or uncontrolled but stable. However, high energy outputs are usually not admissible in practice and for this reason the following expression is proposed as the experiment design criterion:

$$\tilde{J}_T(u) \triangleq \sum_{k=1}^{K} c_k/\mathsf{E}\int_0^T [y_k^2(t)]\,\mathrm{d}t + \nu\sum_{k=1}^{K}\mathsf{E}\int_0^T [y_k^2(t)]\,\mathrm{d}t\,. \tag{5.5}$$

In (5.5), $\nu > 0$ is a given constant, which can be interpreted as a cost per unit of the average output energy (given by the last term in (5.5)).

Minimization of $\tilde{J}_T(u)$ with respect to $u(\cdot)$, under constraints (2.8) and (4.8), can be performed in two steps. Namely, by finding the values $\mathsf{E}\int_0^T [y_k^2(t)]\,\mathrm{d}t$, $k = 1, 2, \ldots$, minimizing (5.5) and then synthetizing an input signal, which assure these values. Realization of the first step yields

$$\mathsf{E}\int_0^T [y_k^*(t)]^2\,\mathrm{d}t = \sqrt{(c_k/\nu)}\,, \quad k = 1, 2, \ldots, K\,, \tag{5.6}$$

where $y_k^*(\cdot)$, $k = 1, 2, \ldots, K$, denote stochastic processes minimizing (5.5). For these processes the minimum value of (5.5) equals

$$\tilde{J}_T^* = 2\sum_{k=1}^{K}\sqrt{(\nu c_k)}\,, \tag{5.6'}$$

while

$$J_T(u) \geqq \tfrac{1}{2}\tilde{J}_T^*\,. \tag{5.6''}$$

It remains to show that such processes can be generated by (4.8) using nonanticipating control laws. For this purpose let us try to find the constants δ_k, $k = 1, 2, \ldots, K$, in feedback laws of the form:

$$u_k(t) = \delta_k(w_k(t) - y_k(t))\,, \quad k = 1.\,2, \ldots, K \tag{5.7}$$

where $w_k(\cdot)$, $k = 1, 2, \ldots, K$, are set point functions of the modal controllers that can be chosen by an experimenter.

A traditional approach to this problem would be as follows. Substituting (5.7) into (4.8), solving the obtained equations and calculating $\mathsf{E}\int_0^T y_k^2(t)\,dt$ in terms of δ_k one can convert (5.6) into nonlinear equations with respect to δ_k. Sufficient conditions for existence of solutions of these equations can be given, at least for a large T and $w_k(\cdot) \equiv 0$. This approach, however, has an important drawback. Namely, it is difficult, if ever possible, to show that the equality sign in (5.4) can be attained for finite T, and thus, only simulation experiments can confirm advantages of such a choice of δ_k. For this reason another approach is proposed here. It is based on certain properties of sequential estimators established in [6] (see also Section 6).

We need, however, to extend slightly a notion of the stopping time (see [6] for definitions). Let $E_k(t)$ be the sigma-field of events generated by the process $y_k(s)$, $s \in [0, t]$, $k = 1, 2, \ldots, K$. Any nonnegative random variable τ_k, which is such that $\forall_{t\in[0,\infty)}\{\omega\colon \tau_k(\omega) \leqq t\} \in E_k(t)$ will be further called the kth processing time.

The above condition means that a decision to stop processing of kth measurement $y_k(\cdot)$ or to continue its observation and processing should be based on observations prior to time t only. In the above terms the stopping time τ, defined as $\tau = \max\limits_{1\leqq k\leqq K} \tau_k$, is a random variable, which defines end of all the observations processing. For our purposes we define the following processing times $\tilde{\tau}_k$, $k = 1, 2, \ldots, K$

$$\tilde{\tau}_k = \inf\left\{s\colon \int_0^s y_k^2(t)\,dt = \sqrt{(c_k/v)}\right\}. \tag{5.8}$$

Remark 3. From (5.8) it is clear that if processing of the kth measurement is stopped when for the first time $\int_0^s y_k^2(t)\,dt = \sqrt{(c_k/v)}$ then condition (5.6) is fulfilled. This is true for every value of gain constants δ_k, $k = 1, 2, \ldots, K$, provided that $\tilde{\tau}_k$, $k = 1, 2, \ldots, K$, are finite almost surely (see Section 6). These definitions and substitution of (5.7) into (4.8) and (4.9) yield the following sequential identification algorithm with random processing times:

Step 1. Basing on measurements of the system

$$dy_k = -(\lambda_k(a) + \delta_k)\, y_k\,dt + \delta_k w_k\,dt + d\beta_k, \quad k = 1, 2, \ldots, K, \tag{5.9}$$

find the estimates $\tilde{\lambda}_k$ of $\lambda_k(a)$ according to the formulas

$$\tilde{\lambda}_k = -\sqrt{(v/c_k)}\left[\int_0^{\tilde{\tau}_k} y_k\,dy_k - \delta_k \int_0^{\tilde{\tau}_k} w_k y_k\,dt\right] - \delta_k, \quad k = 1, 2, \ldots, K, \tag{5.10}$$

where $\tilde{\tau}_k$ are defined by (5.8).

Step 2. Find the estimate $\tilde{a}$ by solving the system of equations:

$$Ba = \sum_{k=1}^{K} (\tilde{\lambda}_k - \alpha_k^{(0)})\,\alpha_k \tag{5.11}$$

Performance of the above algorithm together with the control law is shown in Fig. 1, where $I_1, \ldots, I_K$ denote the eigenvalue identifiers (5.10).

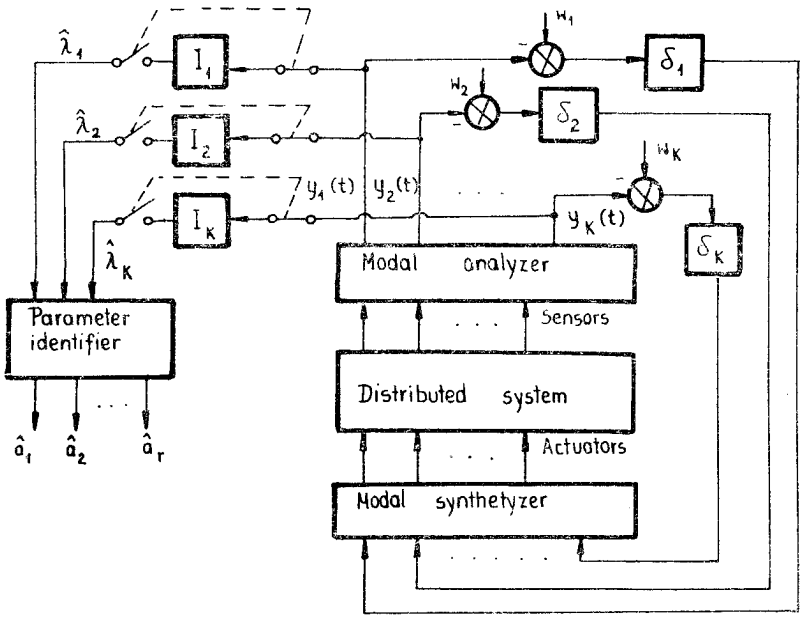

Fig. 1. Scheme of the sequential identification algorithm and the control law.

Computer implementation of this algorithm will be discussed in Section 7, while its theoretical advantages are shown in the next section.

6. PROPERTIES OF THE SEQUENTIAL IDENTIFICATION ALGORITHMS

Properties of the above algorithm can easily be deduced from the following result of Shiryayev [6]. Shiryayev has considered the system

$$\mathrm{d}y = \alpha y \,\mathrm{d}t + \mathrm{d}\beta(t), \quad t > 0, \tag{6.1}$$

where α is a parameter to be estimated from the process $y(t)$, $t \in (0, \hat{T})$, while $\beta(\cdot)$ is a Wiener process with unknown incremental variance. As the estimator $\hat{\alpha}$ of α he has proposed

$$\hat{\alpha} = \frac{1}{H} \int_0^{\hat{T}} y \,\mathrm{d}y, \tag{6.2}$$

where the stopping time $\hat{T}$ is defined by

$$\hat{T} = \inf \left\{ s : \int_0^s y^2(t) \,\mathrm{d}t = H \right\}, \tag{6.3}$$

for a given $H > 0$. The following properties of $\hat{\alpha}$ have been proved in [6]:

P 1) Unbiasedness; $\mathsf{E}\hat{\alpha} = \alpha$.

P 2) $\mathsf{E}(\hat{\alpha} - \alpha)^2 = 1/H$.

P 3) The probability distribution of $\hat{\alpha}$ is Gaussian.

P 4) The stopping time $\hat{T}$ is finite, i.e.

$$\mathsf{P}\{\hat{T} < \infty\} = 1 \quad \text{for every } \alpha .$$

P 5) Let $\mathscr{A}$ be a class of all unbiased estimators of α, which are such that its stopping times $\overline{T}$ are finite and

$$\mathsf{E}\int_0^{\overline{T}} y^2(t)\,\mathrm{d}t \leqq H . \tag{6.4}$$

The estimator $\hat{\alpha} \in \mathscr{A}$ is the best in this class in the sense that

$$\mathsf{E}(\hat{\alpha} - \alpha)^2 \leqq \mathsf{E}(\bar{\alpha} - \alpha)^2 \quad \text{for any} \quad \bar{\alpha} \in \mathscr{A} . \tag{6.5}$$

P 6) The following bounds hold for the expected value of the stopping time:

$$\mathsf{E}\hat{T} \leqq 2 . (|\alpha| . H + 2\sqrt{H}) + (34H + 8 . \alpha^2 . H^2)^{1/2} \tag{6.6}$$

and

$$\mathsf{E}\hat{T} \geqq -2 . \alpha . H \quad \text{for} \quad \alpha < 0 . \tag{6.7}$$

Proofs of P 1)–P 6) are based on the equality

$$\hat{\alpha} = \alpha + \frac{1}{H}\int_0^{\overline{T}} y\,\mathrm{d}\beta(t) \tag{6.8}$$

and on the fact that the integral in (6.8) is a zero mean Gaussial random variable with the variance H (see [6]). Let us note that for theoretical purposes (5.10) can be rewritten as follows:

$$\hat{\lambda}_k = \lambda_k(a) - \sqrt{(v/c_k)}\int_0^{\tilde{\tau}_k} y_k\,\mathrm{d}\beta_k , \quad k = 1, 2, \ldots, K . \tag{6.9}$$

Comparison of (6.8) and (6.9) shows that statistical properties of $\hat{\lambda}_k$ are analogous to P 1)–P 6) although the systems (5.9) and (6.1) are different. In particular, we have

R 1) $\mathsf{E}\hat{\lambda}_k = \lambda_k(a)$, $k = 1, 2, \ldots, K$.

R 2) $\mathsf{E}(\hat{\lambda}_k - \lambda_k(a))^2 = \sqrt{(v/c_k)}$, $k = 1, 2, \ldots, K$.

R 3) $\hat{\lambda}_k$, $k = 1, 2, \ldots, K$, are mutually independent Gaussian random variables.

R 4) $\mathsf{P}\{\tilde{\tau}_k < \infty\} = 1$, $k = 1, 2, \ldots, K$.

R 5) $\hat{\lambda}_k$ is the best unbiased estimate of $\lambda_k(a)$ (in the same sense as in P 5).

In order to investigate properties of $\tilde{a}$ let us note that from R 1), R 2), R 3) and (4.3) it follows that $\hat{\lambda}_k$ can be expressed as follows

$$\hat{\lambda}_k - \alpha_k^{(0)} = \alpha_k' . a + \varepsilon_k , \quad k = 1, 2, \ldots, K , \tag{6.10}$$

where ε_k, $k = 1, 2, \ldots, K$, are mutually independent zero mean Gaussian random variables with the variances

$$\text{var}(\varepsilon_k) = \sqrt{(c_k/\nu)}, \quad k = 1, 2, \ldots, K. \tag{6.11}$$

These facts and (5.11) imply that

$$\tilde{a} = C \cdot \sum_{k=1}^{K} (\hat{\lambda}_k - \alpha_k^{(0)}) \cdot \alpha_k \tag{6.12}$$

is the least squares estimate of a and all assumptions of the Gauss-Markov linear regression scheme are fulfilled (see e.g. [12]) if $(\hat{\lambda}_k - \alpha_k^{(0)})$, $k = 1, 2, \ldots, K$, are treated as observations. Thus, all the following properties of $\tilde{a}$ follow at once from the theory of a regression function parameters estimation (see e.g. [12]):

T 1) Unbiasedness: $\mathsf{E}\tilde{a} = a$.

T 2) Estimator $\tilde{a}$ possesses Gaussian distribution with the covariance matrix

$$\text{cov}(\tilde{a}) = C \cdot \left[\sum_{k=1}^{K} \alpha_k \cdot \alpha_k' \cdot \sqrt{(\nu/c_k)}\right] \cdot C.$$

T 3) Estimator $\tilde{a}$ is efficient in the sense that the Cramér-Rao lower bound in (5.6″) is attained.

T 4) The sequential identification algorithm of Section 5 ensures the minimum value of the $\tilde{J}_T(u)$ criterion given by (5.6′).

Remark 4. Concerning T 3) it is to be noted that the Cramér-Rao lower bound (5.4) is valid also for sequential estimates (see [6] for the proof).

It should be stressed that T 1) ÷ T 4) hold for finite observation time, since from R 4) it follows that $\max_{1 \leq k \leq K} \tilde{\tau}_k$ is almost surely finite. This is in contrast with estimators proposed so far for parameter estimation in DPS, which can be only asymptotically unbiased and efficient.

It is also worthy to notice that the estimation accuracy, given by (5.6″), does not depend on system parameters such as unknown parameters a, initial conditions, variance of disturbances etc., In particular, the estimation accuracy does not depend on feedback gains δ_k and set points w_k, $k = 1, 2, \ldots, K$. These robust type properties are also in contrast with properties of estimators used so far and they put an experiment design problem in a new light. Namely, the stopping time τ can be large if the above mentioned variables are not properly chosen as it is confirmed by simulation studies in the next section. Thus, the experiment design problem is to minimize $\mathsf{E}\tau$ rather then the estimation accuracy, since the last one is determined by ν and the measured modes.

Using the same technique as in [6] bounds on $\mathsf{E}\tau_k$, similar to P 6), can also be derived in our case. This will not be done here, since bounds P 6) applied to our problem with $w_k = 0$, $k = 1, 2, \ldots, K$, occurred to be very rough in comparison with simulation results.

7. SIMULATION EXPERIMENTS

In this section results of extensive simulation studies are reported. Their aims are twofold. Namely, to verify experimentally robust properties of the sequential estimator with respect to system parameters and, what is more important, to investigate their influence on expected value of the stopping time.

For simulation purposes the following system has been modelled

$$\frac{\partial q(x,t)}{\partial t} = a_1 \frac{\partial^2 q(x,t)}{\partial x^2} - a_2 \cdot q(x,t) + u(x,t) + \varepsilon(x,t), \tag{7.1}$$

for $x \in (0, \pi)$ with the boundary conditions

$$q(0,t) = q(\cdot,t) = 0. \tag{7.2}$$

In (7.1) $\varepsilon(x,t)$ denotes zero mean Gaussian noise uncorrelated in space and time.

The operators $A(a)\,h = -a_1(\partial^2 h/\partial x^2) + a_2 h$, defined on the Sobolev space H_0^1 (see e.g. [3] for definition), possesses all the properties required in Section 2. In particular, its eigenfunctions and eigenvalues are as follows:

$$v_k(x) = \sqrt{(2/\pi)} \sin kx\,, \quad \lambda_k(a) = a_1 k^2 + a^2\,, \quad k = 1, 2, \ldots. \tag{7.3}$$

The system (7.1), (7.2) is assumed to be controlled by the memoryless, time-invariant feedback:

$$u(x,t) = \sum_{k=1}^{K} v_k(x) \cdot \delta_k \cdot (w_k - y_k(t))\,, \tag{7.4}$$

where the constant w_k and δ_k, $k = 1, 2, \ldots, K$, are the set points and the amplification factors of each mode and they can be chosen by an experimenter. In (7.4), $y_k(t)$, $k = 1, 2, \ldots, K$, denote available output signals defined as follows

$$y_k(t) = \int_0^{\pi} q(x,t)\, v_k(x)\, \mathrm{d}x\,, \quad k = 1, 2, \ldots, K\,. \tag{7.5}$$

Using (7.1) and (7.4) we infer that they are governed by:

$$\dot{y}_k(t) = -\lambda_k(a)\, y_k(t) + \delta_k(w_k - y_k(t)) + \varepsilon_k(t)\,, \tag{7.6}$$

where $\varepsilon_k(t) = \int_0^{\pi} \varepsilon(x,t)\, v_k(x)\, \mathrm{d}x$, $k = 1, 2, \ldots, K$. The equations (7.6) have been solved by the simplest scheme:

$$y_k(t_{n+1}) - y_k(t_n) = h[-\lambda_k(a) \cdot y_k(t_n) + \delta_k(w_k - y_k(t_n) + \varepsilon_k(t_n))] \tag{7.7}$$

where $t_{n+1} - t_n \triangleq h$, $t_0 = 0$, $n = 1, 2, \ldots$. Estimators of eigenvalues defined by (5.10) have been replaced by their discrete analogs

$$\hat{\lambda}_k = -\delta_k - \Big(\sqrt{(y/c_k)} \sum_{n=0}^{N_K} y_k(t_n)\,[y_k(t_{n+1}) - y_k(t_n)] - \delta_k \cdot w_k \cdot \sum_{n=0}^{N_K} y_k(t_n) \cdot h\Big)\,. \tag{7.8}$$

Remark 5. According to (5.8) the number of steps N_k in (7.8) should be chosen such that $\tilde{\tau}_k = N_k . h$. This, however, requires a large computational burden and therefore the following simplified condition were used instead of (5.8). The number of processing steps of the kth mode – N_k is the smallest integer, for which

$$0{\cdot}9\sqrt{(c_k/\nu)} \leqq \sum_{n=0}^{N_k} y_k^2(t_n) \,.\, h \leqq \sqrt{(c_k/\nu)} \,. \tag{7.9}$$

During all the simulations reported below the step size $h = 0{\cdot}05$, while other system parameters were changes around the following – nominal – values: $a_1 = 1$, $a_2 = 0{\cdot}1$, $\nu = 10^{-6}$, $K = 5$, $y_k(0) = 10 + h \,.\, z_k$, $k = 1, 2, \ldots, K$, where z_k are zero mean Gaussian random variables with the variances equal to 100. Also variances of the system disturbances $\mathrm{var}\,(\varepsilon_k(t_n)) = 100$, $k = 1, 2, \ldots, K$, for all n. Each run reported below has been obtained from averaging results of 20 simulations performed for the same set of system parameters. Empirical mean values and variances obtained by averaging 20 realizations are denoted by $\mathsf{E}^{\sim}\tilde{a}_1$, $\mathsf{E}^{\sim}\tilde{a}_2$, $\mathrm{var}^{\sim}(\tilde{a}_1)$ etc.

Simulation runs have been divided into the following groups:

Group I. In this group all the parameters have their nominal values, while the set points $w_k = 0$, $k = 1, 2, \ldots, 5$. This means that each mode, has been controlled

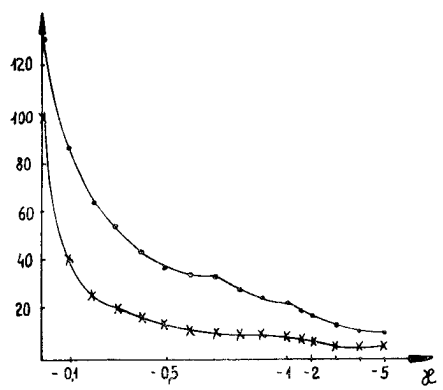

Fig. 2. Empirical means of the number of processing steps: $\mathsf{E}^{\sim}N_1$ – ·—·—·, $\mathsf{E}^{\sim}N_2$ – ×—×—× (Group I).

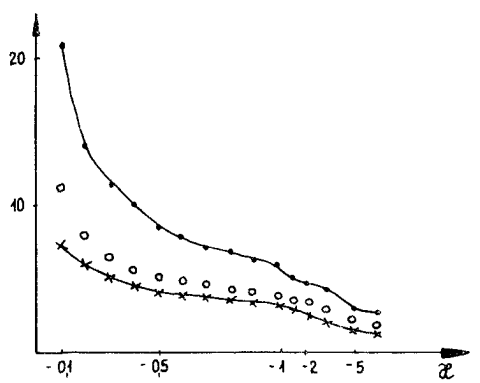

Fig. 3. Empirical means of the number of processing steps: $\mathsf{E}^{\sim}N_3$ – ·—·—·, $\mathsf{E}^{\sim}N_3$ – ×—×—×, $\mathsf{E}N_5$ – o—o—o (Group I).

exactly in a way proposed in [6]. Simulations have been performed for $\varkappa = -1, -2, \ldots, -5$ where $\varkappa \triangleq \delta_k/\lambda_k$, $k = 1, 2, \ldots, 5$, i.e. for positive feedback gains of each mode. Results for negative feedback gains were not obtained since for $\varkappa = 0{\cdot}1$ we had $\mathsf{E}^{\sim}N = 410$, where $N = \max_{1 \leqq k \leqq K} N_k$.

Group II. The aim of these experiments was to investigate the influence of nonzero set points on mean stopping time and the estimation accuracy, keeping other para-

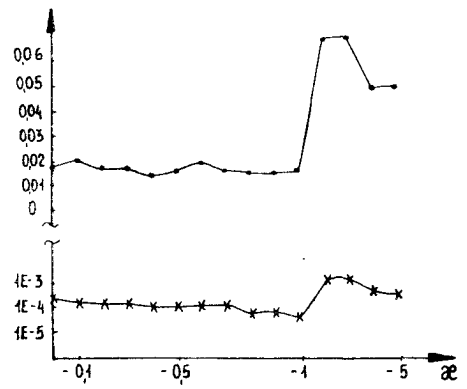

Fig. 4. Empirical variances of the estimates $\widetilde{\text{var}}(\hat{a}_1)$ – ×—×—×, $\widetilde{\text{var}}(\hat{a}_2)$ – ·—·—·
(Group I).

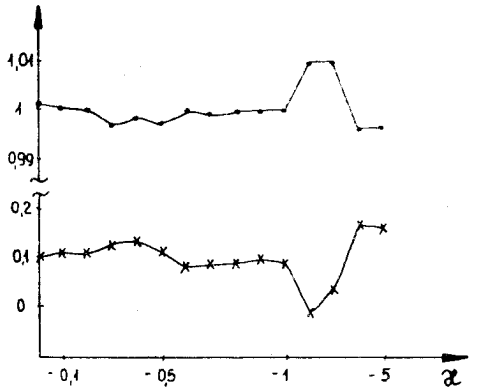

Fig. 5. Empirical means of the estimates $\widetilde{E}\hat{a}_1$ – ·—·—·, $\widetilde{E}\hat{a}_2$ – ×—×—×
(Group I).

meters at their nominal values. Results are reported for $\varkappa = 1, 2, \ldots, 5$ (i.e. for negative feedback gains) with the same value of each mode set point $\hat{w} = w_1 = \ldots = w_5$ equal to: a) $\hat{w} = 10$, b) $\hat{w} = 15$, c) $\hat{w} = 20$, d) $\hat{w} = 30$.

Group III. Simulations have been performed for different values of the control

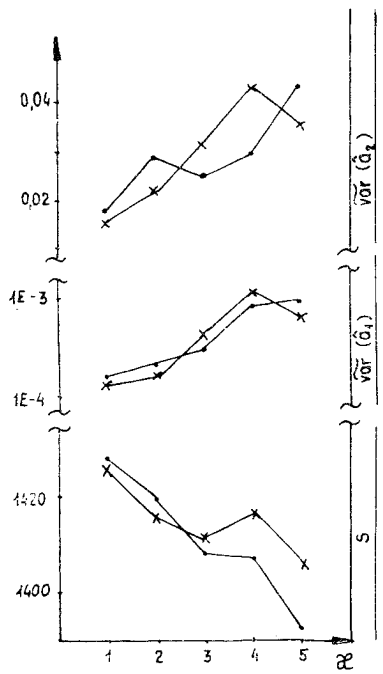

Fig. 6. Empirical variances of the estimates
·—·—· – set points $w = 10$,
×—×—× – set points $w = 15$ (Group II).

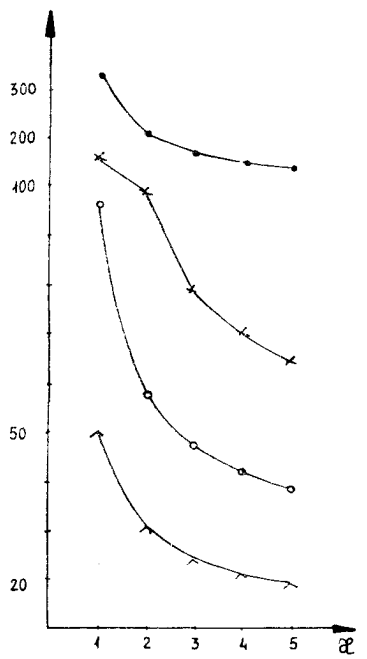

Fig. 7. Empirical means of the stopping times for different set points: ·—·—· – $w = 10$,
×—×—× – $w = 15$, ○—○—○ – $w = 20$,
∧—∧—∧ – $w = 30$ (Group II).

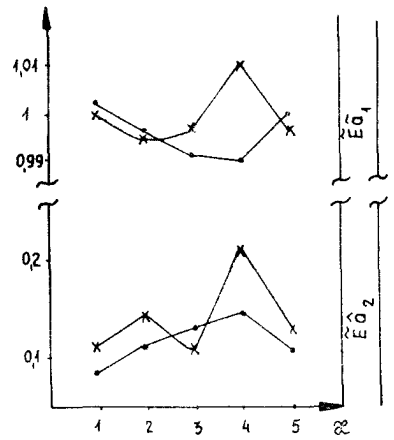

Fig. 8. Empirical means of the estimates: ·—·—· – w = 10, ×—×—× – w = 15 (Group II).

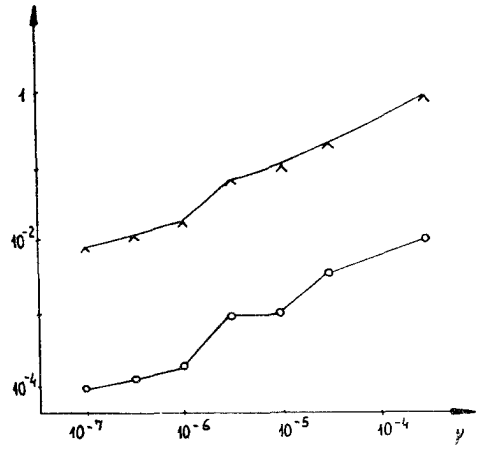

Fig. 9. Empirical variances of the estimates versus energy price ν: o—o—o – $\widetilde{\text{var}}\,(\hat{a}_1)$, ^—^—^ – $\widetilde{\text{var}}\,(\hat{a}_2)$.

cost ν, while the system parameters retain their nominal values. Results are reported for $\varkappa = 1$ and $w_1 = \ldots = w_5 = 20$.

Group IV. The aim of these experiments was to illustrate that the estimation accuracy does not depend on "true" values of the unknown parameters. The rest of the system parameters was as in Group III with $\nu = 10^{-6}$. These results are not drawn in figures, since the estimates variances were almost perfectly constant.

Remark. In Fig. 6 the empirical mean S of the following values

$$h \sum_{k=1}^{K} \sum_{n=0}^{N_k} y_k^2(t_n)$$

is drawn.

The obtained results would suggest the following conclusions:

C 1) From Fig. 2 and Fig. 3 it follows that sequential estimators of eigenvalues in the form proposed in [6] (i.e. with zero set points) can be recommended in practice only in cases when positive feedback is admissible. For negative feedback expected processing times can be too long.

C 2) The proposed here, sequential estimator with nonzero modal set points has sufficiently short expected stopping time also for negative feedback. This time is smaller for larger set points (see Fig. 7). Also greater feedback gains are desirable in order to achieve set points faster.

C 3) From Fig. 6 and Fig. 8 it follows that empirical mean and variance of the estimates are almost independent of δ_k, $k = 1, 2, \ldots, K$. Observed fluctuations are rather caused by difficulties in keeping S constant (see Remark 5).

C 4) Fig. 9 shows that the estimation accuracy depends strongly on the energy cost ν. Small values of ν are desirable from this viewpoint but one should take into account related increase of $E\tau$.

8. CONCLUDING REMARKS

In the paper a sequential parameter estimation algorithm has been proposed and investigated. The estimators possess all good statistical properties usually required and they are proved for observations of finite expected duration. It was shown that this duration can be decreased by suitable choice of modal controllers set points and gains without influencing the estimation accuracy. The algorithm is of non-searching (noniterative) type what results in its high reliability in finding exact estimates. In the form presented here it is applicable to the class of linear DPS with eigenfunctions independent on unknown parameters. This limitation can be partly overcome by using linearization of eigenvalues with respect to a and repeated application of the algorithm. Further efforts seems to be desirable in order to obtain bounds on the expected stopping time of the algorithms and to investigate an influence of PI and PD control laws on it.

(Received June 6, 1985.)

REFERENCES

[1] A. V. Balakrishnan: Applied Functional Analysis. Springer-Verlag, Berlin—New York 1976.

[2] A. V. Balakrishnan: Identification-inverse problems for partial differential equations: A stochastic formulation. In: Proceedings of IFIP Technical Conference on Optimization Techniques, Novosibirsk 1974 (G. I. Marchuk, ed.), Springer-Verlag, Berlin—New York 1975.

[3] R. F. Curtain and A. J. Pritchard: Infinite Dimensional Linear Systems Theory. Springer-Verlag, Berlin—New York 1978.

[4] R. E. Goodwin and R. L. Payne: Dynamic System Identification: Experiment Design and Data Analysis. Academic Press, New York 1977.

[5] C. S. Kubrusly: Distributed parameter system identification. A Survey. Internat. J. Control *26* (1977), 4, 509—535.

[6] R. S. Liptzer and A. N. Shiryayev: Statistics of Stochastic Processes (in Russian). Nauka, Moscow 1974.

[7] R. K. Mehra: Optimal input signals for parameter estimation in dynamic systems. Survey and new results. IEEE Trans. Automat. Control *AC-21*, (1976), 55—64.

[8] M. P. Polis and R. E. Goodson: Parameter identification in distributed systems. A synthesizing overview. Proc. IEEE *64* (1976), 45—61.

[9] M. P. Polis: The distributed system parameter identification problem: A survey of recent results. In: Proceedings of IFAC 3rd Symposium on Control of Distributed Parameter Systems, Toulouse, France 1982 (J. P. Babary and L. LeLetty, eds.), Pergamon Press, Oxford—New York 1983.

[10] E. Rafajłowics: Optimal experiment design for identification of linear distributed parameter systems. Frequency domain approach. IEEE Trans. Automat. Control *AC-28* (1983), 806—808.

[11] E. Rafajłowicz: Optimum experiment design for parameter identification in distributed systems: Brief survey and new results. In: Proceedings of 9th World Congress IFAC 84, Budapest 1984. Pergamon Press, Oxford—New York 1985.

[12] R. C. Rao: Linear Statistical Inference and its Applications. Wiley, New York–London 1966.

[13] A. Shiryayev: Sequential Statistical Analysis (in Russian). Nauka, Moscow 1976.

Dr. Ewaryst Rafajłowicz, Institute of Engineering Cybernetics, Technical University of Wrocław, Wybrzeże St. Wyspiańskiego 27, 50 370 Wrocław. Poland.