

OPTIMAL ACTIVATION STRATEGY OF DISCRETE SCANNING SENSORS FOR FAULT DETECTION IN DISTRIBUTED-PARAMETER SYSTEMS¹

Maciej Patan*, Dariusz Uciński*

* *Institute of Control and Computation Engineering*
University of Zielona Góra
ul. Podgórna 50, 65-246, Zielona Góra
{M.Patan,D.Ucinski}@issi.uz.zgora.pl

Abstract: The problem under consideration is to determine an activation policy of discrete scanning sensors in such a way as to maximize the power of a simple parametric hypothesis test, which verifies the nominal state of the considered distributed system specified over a given multi-dimensional spatial domain. The optimal locations of sensors are determined based on the D_s -optimality criterion defined on the respective Fisher Information Matrix. The proposed approach exploits the notion of directly constrained design measures recently introduced in modern optimum experimental design theory, which leads to an extremely fast iterative procedure of exchange type. In this work, a general scheme of such an approach leading to maximization of the fault detection efficiency in distributed-parameter systems is delineated and tested via computer simulations regarding an advection-diffusion problem. *Copyright*© 2005 *IFAC*

Keywords: distributed-parameter systems; fault detection; hypothesis testing, scanning observations.

1. INTRODUCTION

The systems with spatio-temporal dynamics, commonly known as Distributed-Parameter Systems (DPS's) constitute one of the most general and important classes of systems which are widely used in modelling a wide variety of real-world engineering problems. This is a direct consequence of the fact that recent developments in technical systems, especially regarding control and fault detection, force engineers to search for more precise mathematical models of the considered phenomena. Consequently, lumped descriptions often become unsatisfactory as they may not provide a sufficient approximation of the investigated sys-

tem. This, in turn, leads directly to the description of the system at hand using partial differential equations (PDE's) subject to appropriate boundary and initial conditions. Despite the sophisticated formulation, such models provide a high quality and efficiency of simulations and control techniques (Curtain and Zwart, 1995; Lucquin and Pironneau, 1998; Lasiecka and Triggiani, 2000; Demetriou and Moghani, 2004).

Although we witness an extremely fast development of methods of Fault Detection and Isolation (FDI) for dynamical systems in the last decade (Basseville and Nikiforov, 1993; Isermann, 1997; Frank and Köppen-Seliger, 1997; Chen and Patton, 1999; Patton *et al.*, 2000; Chiang *et al.*, 2001; Korbicz *et al.*, 2004), there is a grave lack of such effective techniques dedicated for DPS's.

¹ The work was supported by the State Committee for Scientific Research in Poland under the grants 4 T11A 014 25 and 3 T11C 035 27

Furthermore, within the framework of FDI systems, the optimization of data acquisition process which increases the reliability of the diagnosis is most often neglected and the contributions are very scarce. The main aim of this work is to develop a novel approach originated in the parameter identification and related to a wide class of processes. One of the most important problems in the context of fault detection in DPS's is a proper determination of observation schedules. It consists in allocation of a limited number of measurement transducers in a given spatial domain in such a way as to maximally reduce the uncertainty of the detection. The strong necessity of developing some systematic approaches stems directly from the fact, that in an overwhelming majority of practical problems the allocation of the sensors does not follow directly from the physical considerations or experimenter intuition. Moreover, additional requirements appear regarding reduction of the instrumentation cost and increasing the effectiveness of diagnostic methods.

The sensor location problem for DPS's has already attained a lot of attention, but mainly in the context of state and/or estimation accuracy (for surveys, see Kubrusly and Malebranche, 1985; Uciński, 2000; van de Wal and de Jager, 2001; Demetriou and Borggaard, 2004; Patan, 2004; Uciński, 2004). Within the framework of diagnostics, the crucial difficulty is the definition of a suitable criterion describing the relations between the quality of the system diagnosis and the observation strategy. In this paper, as the appropriate performance measure, we propose the D_s -optimality criterion, operating on the Fisher Information Matrix (FIM) related to the parameters. Furthermore, we shall indicate how an extremely efficient approach based on directly-constrained design measures proposed in (Uciński and Patan, 2002) can be applied in order to derive the most informative observations for fault detection.

2. PARAMETER ESTIMATION APPROACH TO FAULT DETECTION IN DPS'S

Parameter estimation is one of the fundamental methods from among all analytical techniques of fault detection (Korbicz *et al.*, 2004). Its role is of great significance in situations, when the abnormal system state appears not only in the form of output changes but also as fluctuations in model parameters. This is a very common situation in practice if only the parameters have a physical interpretation built upon a proper analysis of various quantities which are crucial for the considered process. Unfortunately, they are usually directly non-measurable and application of effective parameter estimation algorithms is required in order

to obtain their estimates. The basic idea is a comparison of such estimates with some known nominal values of parameters treating possible differences as residuals which contain information about potential faults. Then, based on some thresholding techniques, the appropriate decision making system can be constructed.

Let $y = y(x, t)$ denote the scalar system state at point x of the spatial domain $\Omega \subset \mathbb{R}^d$ and let time t belong to the observational interval $T = [0, t_f]$. Mathematically, we handle a dynamic system described by the following PDE:

$$\frac{\partial y}{\partial t} = \mathcal{F}(x, t, y, \nabla y, \nabla^2 y; \theta^*), \quad (1)$$

$$(x, t) \in \Omega \times T \subset \mathbb{R}^{d+1},$$

subject to the boundary and initial conditions

$$\mathcal{E}(x, t, y, \nabla y; \theta^*) = 0, \quad (x, t) \in \partial\Omega \times T, \quad (2)$$

$$y(x, 0) = y_0(x), \quad x \in \Omega, \quad (3)$$

where \mathcal{E}, \mathcal{F} and y_0 are some known functions, ∇ and ∇^2 denote the gradient and Hessian, respectively, and θ^* stands for the vector of constant system parameters.

The state y is observed (possibly indirectly) by N moving sensors, which can switch their locations at time instants $0 < t_0 < t_1 < \dots < t_K = t_f$ and will be remaining stationary for the duration of each subinterval $T_k = [t_{k-1}, t_k]$, $k = 1, \dots, K$. Forming such an arbitrary partition on the time interval T , the considered 'scanning' observation strategy can be formally represented as

$$z^j(t) = h(y(x_k^j, t; \theta^*), t) + \varepsilon^j(t), \quad t \in T_k \quad (4)$$

for $j = 1, \dots, N$ and $k = 1, \dots, K$, where $h(\cdot, \cdot)$ is a given function, $z^j(t)$ is a scalar output, $x_k^j \in X$ stands for the location of the j -th sensor on the subinterval T_k , X signifies the part of Ω where the measurements can be made, and $\varepsilon^j(\cdot)$ denotes the zero-mean, Gaussian and uncorrelated measurement noise, i.e. $E[\varepsilon^j(t)] = 0$ and $\text{var}(\varepsilon^j(t)) = \sigma^2$.

A further assumption is that the estimation of the unknown parameter vector θ^* is performed via minimization of the least-squares criterion

$$J(\theta) = \sum_{j=1}^N \sum_{k=1}^K \int_{T_k} |z^j(t) - h(\hat{y}(x_k^j, t; \theta), t)|^2 dt, \quad (5)$$

with $\theta \in \Theta_{\text{ad}}$, where Θ_{ad} is the set of admissible parameters and $\hat{y}(\cdot, \cdot; \theta)$ denotes the solution of (1)–(3) corresponding to a given θ . A vector $\hat{\theta}$ minimizing $J(\theta)$ stands for the estimate of the true value of θ^* . Generally, in practical situations, only a subset of all parameters can be useful for detection. If such a situation takes place, then without loss of generality the parameters of interest can be distinguished by partitioning the parameter vector into

$$\theta^T = [\theta_1 \dots \theta_s \vdots \theta_{s+1} \dots \theta_m] = [\alpha^T \vdots \beta^T] \quad (6)$$

where α is a vector of s parameters which are essential for a proper fault detection and β is the vector of some parameters which are a part of the model but are not significant for detection (although they can be used for fault isolation or are some representation of the model uncertainty). Based on the observations, it is possible to test the simple hypothesis $H^0 : \alpha = \alpha^*$, where α^* is the nominal value of the vector α corresponding to the normal system performance.

The ‘continuous’ generalization of the likelihood function for the considered experiment takes the following form (Goodwin and Payne, 1977):

$$L(z; \theta) = \left(\frac{1}{2\pi\sigma^2} \right)^{N/2} \exp \left\{ -\frac{1}{2\sigma^2} \sum_{j=1}^N \sum_{k=1}^K \int_{T_k} |z^j(t) - h(\hat{y}(x^j, t; \theta), t)|^2 dt \right\}. \quad (7)$$

Setting $\Theta_0 = \{\theta \in \Theta : \alpha = \alpha^*\}$, we can define the following generalized likelihood ratio:

$$\lambda(z) = \frac{\sup_{\theta \in \Theta_{\text{ad}}} L(z; \theta)}{\sup_{\theta \in \Theta_0} L(z; \theta)} = \exp \left\{ -\frac{1}{2\sigma^2} (J(\hat{\theta}) - J(\tilde{\theta})) \right\} \quad (8)$$

where

$$\hat{\theta} = \arg \min_{\theta \in \Theta_{\text{ad}}} J(\theta), \quad \tilde{\theta} = \arg \min_{\theta \in \Theta_0} J(\theta). \quad (9)$$

The generalized log-likelihood ratio test is widely used in statistics, because it can be shown that assuming the validity of the null hypothesis H^0 the sequence $\{2 \ln \lambda(y)\}$ for $N, K \rightarrow \infty$ is weakly convergent to a χ^2 random variable on s degrees of freedom (Goodwin and Payne, 1977, Thm. 3.6.1, p. 55). The meaning of this fact is that we can compare the observed value of $2 \ln \lambda(z)$ with some threshold k_γ obtained from the cumulative χ^2 distribution on s degrees of freedom where k_γ is such that $100(1-\gamma)\%$ of the distribution lies to the left of k_γ . The decision rule for a given significance level γ , which represents a fixed range of model uncertainty, takes the following form:

$$S = \begin{cases} S^1 & \text{if } 2 \ln \lambda(y) \geq k_\gamma \quad (\text{reject } H^0) \\ S^0 & \text{if } 2 \ln \lambda(y) < k_\gamma \quad (\text{accept } H^0) \end{cases} \quad (10)$$

The potential rejection of H^0 indicates an essential deviation of the vector α from the nominal value of this parameter and is a base for detection of abnormal states in the system.

3. D_s -OPTIMUM SENSOR LOCATION PROBLEM

At this point, it is at least clear that in order to achieve low probabilities of false alarms, and missed detection, we have to keep the low values for the probabilities of rejecting H^0 when H^0

is true and accepting H^0 when the alternative hypothesis $H^1 : \alpha = \alpha^1 \neq \alpha^*$ is true, respectively.

If only a subset of s parameters is of interest, with the parameter vector partitioning given by (6) and for a fixed significance level γ (the probability of a false alarm), the power of the proposed test for alternative hypothesis H^1 (the probability of proper detection) can be increased by minimization of the D_s -optimality criterion (Goodwin and Payne, 1977):

$$\Psi_s(M) = -\ln \det[M_{\alpha\alpha} - M_{\alpha\beta}M_{\beta\beta}^{-1}M_{\alpha\beta}^T], \quad (11)$$

where $M \in \mathbb{R}^{m \times m}$ is the FIM corresponding to the vector θ , which can be partitioned into blocks

$$M = \begin{bmatrix} M_{\alpha\alpha} & \cdots & M_{\alpha\beta} \\ \vdots & \ddots & \vdots \\ M_{\alpha\beta}^T & \cdots & M_{\beta\beta} \end{bmatrix} \quad (12)$$

and

$$M_{\alpha\alpha} \in \mathbb{R}^{s \times s}, \quad M_{\alpha\beta} \in \mathbb{R}^{s \times (m-s)}, \\ M_{\beta\beta}^{-1} \in \mathbb{R}^{(m-s) \times (m-s)}.$$

In the considered case, the average information matrix is defined as (Uciński and Patan, 2002)

$$M = \frac{1}{N} \sum_{j=1}^N \sum_{k=1}^K \Upsilon_k(x_k^j), \quad (13)$$

where

$$\Upsilon_k(x) = \frac{1}{t_f} \int_{T_k} G^T(x, t)G(x, t) dt, \quad (14)$$

$$G(x, t) = \frac{\partial h(y(x, t; \theta), t)}{\partial y} \frac{\partial y(x, t; \theta)}{\partial \theta} \Big|_{\theta=\theta^*}. \quad (15)$$

Minimization of the criterion (11) can be achieved by a suitable sensor allocation. For the D_s -optimality case, the process is equivalent to minimizing the determinant of the estimate of the covariance matrix for vector α . Furthermore, we have the relation

$$\Psi_s[M] = \ln \det M_{\beta\beta} - \ln \det M. \quad (16)$$

An introduction of the optimality criterion (11) operating on the FIM elements make it possible to formulate our sensor location problem for fault detection as an optimization one:

$$\Psi_s[M] \longrightarrow \min \quad (17)$$

with respect to sensor positions x_k^j , $j = 1, \dots, N$, $k = 1, \dots, K$ chosen from the set of admissible locations X .

4. PROPOSED STRATEGY OF SCANNING OBSERVATIONS

In order to avoid the so-called *sensor clusterization* phenomenon (i.e. the tendency to taking measurements at the same spatial locations by different sensors), being a direct consequence of

the assumption of independent observations commonly employed in classical methods of experimental design, an extremely efficient approach developed by Uciński and Patan (2002) can be applied. The main idea is to operate on the density of sensors (i.e. the number of sensors per unit area)

$$\xi(dx) = \lim_{\Delta X \rightarrow 0} \frac{N(\Delta X)}{N}, \quad (18)$$

where $N(\Delta X)$ denotes the number of sensors in the area X , than on the sensors' locations directly, which is justified for a sufficiently large total number of sensors N .

In contrast to the classical designs encountered in experimental design theory, we impose the crucial restriction that the density of sensor allocation must not exceed some prescribed level, i.e. (Uciński and Patan, 2002; Patan, 2004)

$$\xi_k(dx) \leq \omega(dx), \quad k = 1, \dots, K, \quad (19)$$

where $\omega(dx)$ signifies the maximal possible 'number' of sensors per dx such that $\int_X \omega(dx) \geq 1$. Furthermore, we assume that the considered domain X where the measurements can be taken is a finite P -element set. Then, any K -element sequence of N -element subsets of X is called a *design of experiment* and will be denoted by ξ . In the sequel, the set containing all designs of such form will be denoted by Ξ . In such a manner, we are faced with the following optimization problem:

$$\xi^* = \arg \min_{\xi \in \Xi} \Psi[M(\xi)] \quad (20)$$

subject to

$$\xi_k(dx) \leq \omega(dx), \quad k = 1, \dots, K. \quad (21)$$

The problem of determining the optimal design ξ^* can be readily solved with the use of the approach based on the so-called directly constrained design measures (Fedorov, 1989; Fedorov and Hackl, 1997). Its extensions related to the optimal planning of the activation schedules for discrete scanning sensors were addressed in the works (Uciński and Patan, 2002; Patan, 2004; Uciński, 2004). In order to apply it to the considered case, let us notice that introducing the decomposition

$$M^{-1} = \begin{bmatrix} D_{\alpha\alpha} & \vdots & D_{\alpha\beta} \\ \dots & \dots & \dots \\ D_{\alpha\beta}^T & \vdots & D_{\beta\beta} \end{bmatrix} \quad (22)$$

where $D_{\alpha\alpha} \in \mathbb{R}^{s \times s}$, $D_{\alpha\beta} \in \mathbb{R}^{s \times (m-s)}$, $D_{\beta\beta}^{-1} \in \mathbb{R}^{(m-s) \times (m-s)}$, from the Equivalence Theorem for D_s -optimum designs (Fedorov, 1972) an analogue of the variance of the prediction for each time interval T_k can be established from the system response. In particular, we have the following form of this function and its components, which are crucial for the potential numerical algorithms, which can be used for determining the optimal solution:

$$\psi_k(x, \xi) = \phi_k(x, \xi) - \varsigma_k(\xi)/K, \quad (23)$$

$$\varsigma_k(\xi) = -\text{trace} [\overset{\circ}{\Psi}[M(\xi)]M(\xi)] = s, \quad (24)$$

$$\phi_k(x, \xi) = -\text{trace} [\overset{\circ}{\Psi}[M(\xi)]\Upsilon_k(x)], \quad (25)$$

where

$$\overset{\circ}{\Psi}[M(\xi)] = \frac{\partial \Psi(M)}{\partial M} = \begin{bmatrix} D_{\alpha\alpha} & \vdots & D_{\alpha\beta} \\ \dots & \dots & \dots \\ D_{\alpha\beta}^T & \vdots & D_{\beta\beta} - M_{\beta\beta}^{-1} \end{bmatrix}. \quad (26)$$

Analytical determination of the optimal designs is possible only in simple situations and for general systems it is usually the case that some iterative design procedures will be required. The extremely efficient computational procedure for that purpose was developed in (Uciński and Patan, 2002), where the interested reader can find the detailed scheme of simple exchange-type routine. The main idea of the algorithm mentioned above consist in replacing in the actual design the worst measurement points (in the sense of the function $\psi_k(x, \xi)$) with the best candidates from among the rest of all admissible measurement locations which are not the elements of design. This operation is repeated iteratively until the design will consist of the best spatial locations for measurements. The maximal numbers of sensors allocated to the spatial element dx can be usually expressed as $\omega(dx) = \varrho(x)dx$, $k = 1, \dots, K$, where the ϱ plays the role of a density function. But then it is always possible to propose an appropriate change of coordinates which allows us to restrict attention to a constant ϱ . Additionally, a computer implementation forces the replacement of all integration operators (w.r.t. the spatial elements dx) by summing over some suitable regular grid elements. As a result, the rather abstract form of the iterative procedure presented in (Uciński and Patan, 2002; Patan, 2004) is reduced to an efficient exchange-type algorithm with the additional constraint that every grid element must not contain more than one support point in each time subinterval.

5. ILLUSTRATIVE EXAMPLE

As an illustrative example consider an atmospheric pollutant transport-chemistry process over an urban area being normalized to a unit square. For the sake of clarity, assume that the velocity of the transport medium is constant over the considered domain and defined by a vector $v = (-1, -1)$, (i.e. the air moves with constant speed rate along the axis $x_1 = x_2$). In addition, we take into account an active source of pollution and reaction, which leads to changes in the pollutant concentration $y(x, t)$. The entire process over the observation interval $T = [0, 1]$ is described by the following advection-diffusion-reaction equation:

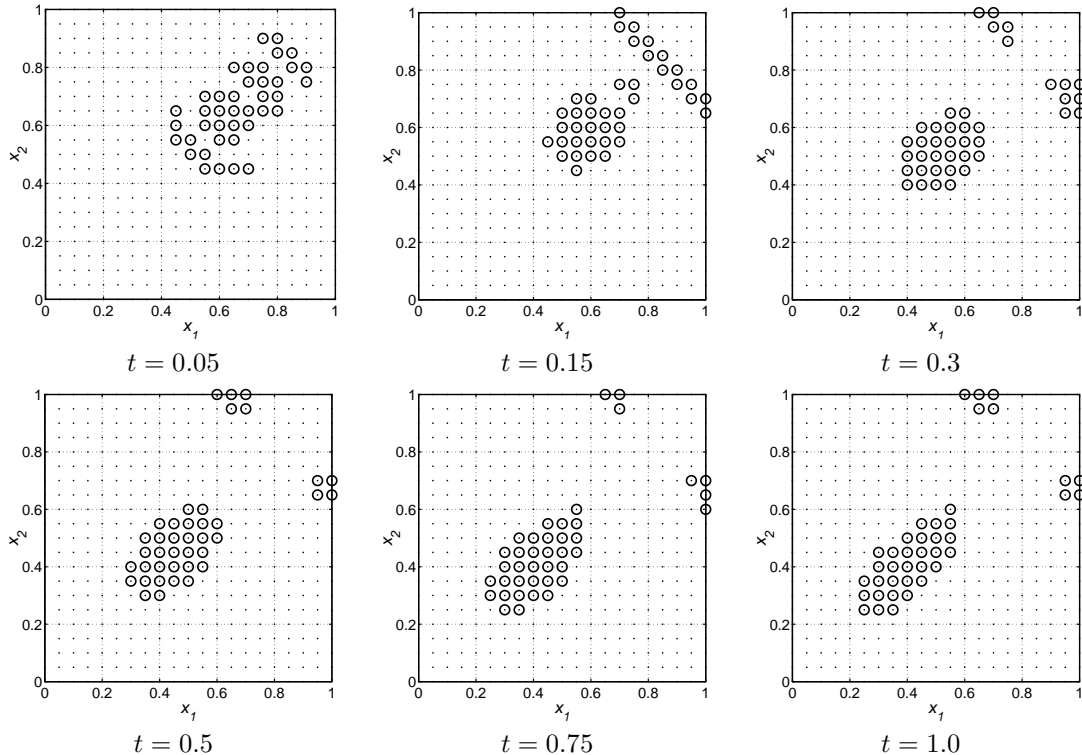


Fig. 1. D_s -optimal sensor switchings

$$\begin{aligned} \frac{\partial y(x, t)}{\partial t} + \nabla \cdot (v(x)y(x, t)) \\ = \nabla \cdot (d(x)\nabla u(x, t)) + f(x), \quad x \in \Omega \end{aligned} \quad (27)$$

subject to the boundary and initial conditions:

$$\begin{cases} \frac{\partial y(x, t)}{\partial n} = 0, & \Gamma \times T, \\ y(x, 0) = 0, & \Omega, \end{cases} \quad (28)$$

where term $f(x) = e^{-50[(x_1-0.7)^2+(x_2-0.7)^2]}$ represents an active source of pollutant located at point $(0.7, 0.7)$, and $\partial y/\partial n$ stands for the partial derivative of y with respect to the outward normal to the boundary Γ . In simulations, the following form of the distributed diffusion coefficient was applied

$$d(x) = \theta_1 + \theta_2 x_1^2 + \theta_3 x_2^2, \quad (29)$$

where $\theta_1 = 0.1$, $\theta_2 = \theta_3 = -0.02$.

The fault scenario was that at time instant $t = 0.5$ the intensity of the pollutant emission from the source dramatically increases and a symptom of this abrupt fault is an excessive deviation of the parameter θ_1 from its nominal value. Our task was to determine the D_s -optimal activation schedule for $N = 40$ sensors to maximize the reliability of the detection. Measurements can be taken on the grid of size 21×21 (cf. Fig. 1). The algorithm of Section 4 generated the solution after 57 iterations ($\eta = 10^{-3}$), practically below 1 second on a low-cost PC (Pentium IV, 1.7 GHz, using the Lahey/Fujitsu Fortran 95 compiler v.5.7), and the final sensor allocation is presented in Fig. 1, where open circles indicate the activated sensors. Observe that the symmetry of the problem along

the $x_1 = x_2$ axis is retained in the optimal solution. The sensors tend to take measurements in the areas where the changes in pollutant concentration are the greatest, clearly following the wind direction. Surprisingly, the measurements in the close vicinity of the pollution source are not very attractive for the fault detection, which is contradictory with our intuition and is very difficult to predict when armed only with the experimenter experience. The fault scenario was performed with additional measurement noise whose standard deviation was assumed on the level of 5% of the simulated system output. The detection with the use of fixed thresholds is illustrated in Fig. 2. It is clear that the appropriate level of the threshold should be suitably chosen in such a way as to obtain the reasonably high sensitivity of detection system and the small false alarms rate.

6. CONCLUSIONS

The proposed approach can be considered as a promising attempt to establish an interconnection between parameter-estimation-based fault detection for DPS's and the sensor location problem. As a result, an extremely simple and efficient algorithm is adopted for the purpose of determining an optimal measurement schedule for fault detection. However, there still remain some open problems which need close attention. One of the essential issues is a necessity of developing some adaptive thresholding techniques which allow to effectively establish a compromise between a low number

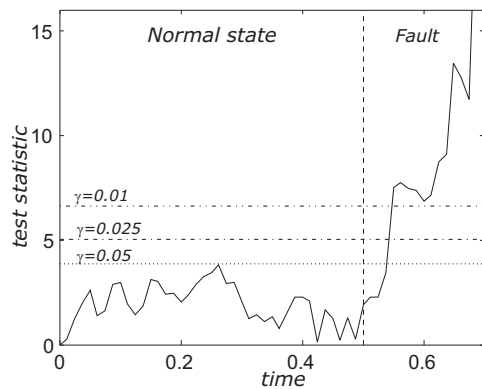


Fig. 2. Comparison of different significance levels for fault detection

of false alarms and a decrease of the detection system sensitivity. Another not trivial problem is the dependence of the optimal solution on the system parameters and therefore some robust approaches to parameter uncertainty of the model are required. This leads directly to the notion of so-called *robust designs of experiment* (Walter and Pronzato, 1997) which try to make the optimal solutions independent on the parameters to be identified.

REFERENCES

- Basseville, M. and I. Nikiforov (1993). *Detection of Abrupt Changes*. Prentice-Hall. New York.
- Chen, J. and R. J. Patton (1999). *Robust Model-Based Fault Diagnosis for Dynamic Systems*. Kluwer Academic Publishers. Berlin.
- Chiang, L. H., E. L. Russel and R. D. Braatz (2001). *Fault Detection and Diagnosis in Industrial Systems*. Springer. London.
- Curtain, R. F. and H. Zwart (1995). *An Introduction to Infinite-Dimensional Linear Systems Theory*. Texts in Applied Mathematics. Springer-Verlag. New York.
- Demetriou, M. A. and J. Borggaard (2004). Optimization of a joint sensor placement and robust estimation scheme for distributed parameter processes subject to worst case spatial disturbance distributions. In: *Proc. of the 2004 American Control Conference*. Boston. (published on CD-ROM).
- Demetriou, M. A. and T. Moghani (2004). Robust actuator placement in flexible plates subject to worst-case spatial distribution of disturbances. In: *Proc. of the 2004 American Control Conference*. Boston. (published on CD-ROM).
- Fedorov, V. V. (1972). *Theory of Optimal Experiments*. Academic Press. New York.
- Fedorov, V. V. (1989). Optimal design with bounded density: Optimization algorithms of the exchange type. *Journal of Statistical Planning and Inference* **22**, 1–13.
- Fedorov, V. V. and P. Hackl (1997). *Model-Oriented Design of Experiments*. Lecture Notes in Statistics. Springer-Verlag. New York.
- Frank, P. M. and B. Köppen-Seliger (1997). New developments using AI in fault diagnosis. *Artificial Intelligence* **10**(1), 3–14.
- Goodwin, G. C. and R. L. Payne (1977). *Dynamic System Identification. Experiment Design and Data Analysis*. Mathematics in Science and Engineering. Academic Press. New York.
- Isermann, R. (1997). *Supervision, Fault Detection and Diagnosis of Technical Systems*. Special Section Control Engineering Practice. Vol. 5, No. 5.
- Korbicz, J., J.M. Kocielny, Z. Kowalczyk and W. Cholewa (2004). *Fault Diagnosis. Models, Artificial Intelligence, Applications*. Springer-Verlag. Berlin Heidelberg.
- Kubrusly, C. S. and H. Malebranche (1985). Sensors and controllers location in distributed systems—A survey. *Automatica* **21**(2), 117–128.
- Lasiecka, I. and R. Triggiani (2000). *Control Theory for Partial Differential Equations: Continuous and Approximation Theories*. Vol. I and II of *Encyclopedia of Mathematics and Its Applications*. Cambridge University Press. Cambridge.
- Lucquin, B. and O. Pironneau (1998). *Introduction to Scientific Computing*. John Wiley & Sons. Chichester.
- Patan, M. (2004). *Optimal Observation Strategies for Parameter Estimation of Distributed Systems*. Technical University Press. Zielona Góra.
- Patton, R. J., P. M. Frank and R.N. Clark (2000). *Issues of Fault Diagnosis for Dynamic Systems*. Springer-Verlag. Berlin.
- Uciński, D. (2000). Optimal sensor location for parameter estimation of distributed processes. *International Journal of Control* **73**(13), 1235–1248.
- Uciński, D. (2004). *Optimal Measurement Methods for Distributed Parameter System Identification*. CRC Press. Boca Raton.
- Uciński, D. and M. Patan (2002). Optimal location of discrete scanning sensors for parameter estimation of distributed systems. In: *Proc. 15th Triennial World Congress of the IFAC, Barcelona, Spain*. Pergamon Press. Oxford. (published on CD-ROM).
- van de Wal, M. and B. de Jager (2001). A review of methods for input/output selection. *Automatica* **37**, 487–510.
- Walter, É. and L. Pronzato (1997). *Identification of Parametric Models from Experimental Data*. Communications and Control Engineering. Springer-Verlag. Berlin.