OPTIMAL OPERATIONAL STRATEGIES FOR AN INSPECTED COMPONENT – SOLUTION TECHNIQUES

U. Pulkkinen and S. Uryas'ev

CP-91-13
July 1991

Mr. Urho Pulkkinen works at the Technical Research Centre of Finland, Laboratory of Electrical and Automation Engineering
Dr. Stanislav Uryas'ev works at the International Institute for Applied Systems Analysis

IIASA International Institute for Applied Systems Analysis □ A-2361 Laxenburg □ Austria
Telephone: +43 2236 715210 □ Telex: 079 137 iiasa a □ Telefax: +43 2236 71313

Foreword

This is the third report on work done on time dependent probabilities at IIASA. The International Atomic Energy Agency (IAEA) and the Technical Research Centre of Finland (VTT) have cooperated in this work which was initiated in 1990. The underlying mathematical model was described for two different cases in the earlier papers. This paper is directed towards solution techniques by which the optimal solution for the problem can be found. A special consideration in the paper is devoted to the calculation of the gradient, because the nonsmooth character of the model makes this especially cumbersome. The assumptions of the model are relatively simple, but can be refined accordingly when necessary. The model is in its first phase and is intended to be used to obtain qualitative insights on relationships between main variables. The model has been tested using a computer code and the results obtained show agreement with practical results.

Comments or proposals for applications of this modeling approach are invited.

Björn Wahlström
Leader
Social & Environmental Dimensions
of Technology Project

Contents

OPTIMAL OPERATIONAL STRATEGIES FOR AN INSPECTED COMPONENT – SOLUTION TECHNIQUES

U. Pulkkinen and S. Uryas'ev

1 Introduction

This paper is a continuation of an earlier IIASA Working Paper entitled "Optimal Operational Strategies for an Inspected Component - Statement of the Problem" [3]. This earlier working paper presented a mathematical formulation of the optimization problem concerning the operational strategies of an inspected component. More specifically, a probabilistic failure model of a component which has some initial defects due to an imperfect manufacturing process was presented, where the sizes of the defects may increase with random amounts whenever a random shock hit the system. The sizes of the defects may be measured with an imperfect device, and the measurements were described with another probability model. The inspection measurements, repairs and failures of the component were included in the cost model. The aim of the model was to find such an operational strategy to minimize the expected costs. The model is in the form of a stochastic optimization problem, and the aim of this follow-up paper is to construct a solution technique for the problem.

The solution of the problem requires the differentiation of the objective function which is, in this case, a mathematical expectation of a discontinuous function. The objective function cannot be differentiated with standard techniques, since interchange of the gradient and the expectation operator is not allowed (see, for example [5]). To cope with the problem, we have to use formulae for differentiation of integrals over set depending on the parameters. To date, the theory for the differentiation of such integrals is not fully developed. Formulae for such derivatives, in general case, are described in the papers of E. Raik [4], N. Roenko [6], J. Simon [8]. The gradient expressions in these papers have the form of surface integrals, and they are computationally inconvenient since the probability measure of such surfaces is equal to zero.

Uryas'ev [13] expressed the gradient in the form of an integral over a volume. This kind of formula is more convenient because stochastic quasi-gradient algorithms (see, for example, [1],[2],[10] and [14]) can be used for the minimization of integral functions. The paper by

Uryas'ev [13] provides a formula for the differentiation of an integral over a set given by only one nonlinear inequality. In our model the objective function can be expressed as a sum of multidimensional integrals over sets defined by several nonlinear inequalities. The general expressions for such gradients are known, but these results are not yet published. In this paper we apply these expressions to evaluate the gradient of our objective function. As a special case of the general differentiation formula, we prove two lemmas, which are used in calculating the gradient of the objective function.

The solution of the optimization problem by applying the stochastic quasigradient approach requires the sampling of the trajectories of the stochastic system. In the case of our model this is rather simple due to the simple probability distributions. We have assumed discrete distributions for the numbers of the defects; exponential distributions for the sizes of the defects, the size increments of the defects and for the time between successive shocks; and normal distributions for the inspection models. The conditional distribution of the failure time is with the piecewise linear failure intensity. The methods to generate (pseudo) random numbers are well known. In our case we apply the package C-RAND [9] to generate all random distributions.

The notation of this paper is based on the earlier Working Paper [3].

2 The Calculation of the Conditional Expectations for the Cost Function

The cost function $F(x) = E\,[\,\tilde{f}(x,\omega)\,]$ can be calculated with the function

$$\tilde{f}(x,\omega) \;=\; G_f\int_0^{T_{stop}} \lambda(t)\,dt \;-\; G_p\,T_{stop} \;+\; \sum_{j=1}^{j_{stop}} \left[G_1\chi_1(t_j) + G_2\chi_2(t_j) + G_r\chi_r(t_j)\right],$$

(see p.18 in [3]). The first term in this formula is a mathematical expectation of costs of the component failure; the second is a profit from the component exploitation; the third is a cost of the large inspections, small inspections and repairs, respectively. We consider that all random values are specified on the probability space $(P, \mathcal{F}, \Omega)$ [1]. The function $\tilde{f}(x,\omega)$ is a discontinuous one with respect to x , since indicator functions $\chi_1(t_j), \chi_2(t_j), \chi_r(t_j)$ and time moment T_{stop} are discontinuous functions of x. For this reason, it is not possible to calculate the gradient of the function $F(x)$ as a mathematical expectation of the gradient of the function $\tilde{f}(x,\omega)$. To calculate the gradient we have to rewrite the objective function $F(x)$. Objective function is presented as a mathematical expectation of some random function continuously and piecewice smoothly depending upon the parameter x. To smooth over the function $\tilde{f}(x,\omega)$ we integrate

[1] We denote the random variables without the index $\omega \in \Omega$, i.e. the random variable $\theta(\omega)$ is denoted simply by θ, whenever it is possible without confusion.

it with respect to random variables $\theta^{\nu}(t_j)$, $0 \le t_j \le T_{stop}$. More formally, it could be presented as a mathematical expectation of the function $\tilde{f}(x,\omega)$ with respect to a proper σ–algebra.

Further, to simplify investigation we clearly specify random values which generate all stochastic behaviour of the model:

1. Initial number of the defects $M(0)$;
2. Initial defect sizes $C_1(0), \ldots, C_{M(0)}(0)$;
3. Number of the defects $M(t_j^r)$ after repair, $j = 1, 2, \ldots$;
4. Defect sizes $C_1(t_j^r), \ldots, C_{M(t_j^r)}(t_j^r)$ after repair, $j = 1, 2, \ldots$;
5. Shock points τ_l, $l = 1, 2, \ldots$;
6. Exponentially distributed independent variables $\delta_1, \delta_2, \ldots$, i.e. these variables have density function

   $$g(z) = \zeta \exp\{-\zeta z\} \ . \tag{1}$$

   Defect increases $\Delta C_i(t_j)$ (here $i = 1, \ldots, M_{\max}$, $j = 1, 2, \ldots$) are generated by these identically exponentially distributed independent variables. Defect increase $\Delta C_i(t_j)$ is equal to the product of the proper value δ_n on $C_i(t_j)$;
7. Signals $\theta_i(t_j)$ caused by the defects ($i = 1, \ldots, M_{\max}$, $j = 1, 2, \ldots$) .

We introduce a new σ–algebra $\mathcal{F}_1$ belonging to the σ–algebra $\mathcal{F}$. The σ–algebra $\mathcal{F}_1$ is generated by the random values specified in items 1-6. It means that if we calculate the mathematical expectation of the function $\tilde{f}(x,\omega)$ with respect to the σ–algebra $\mathcal{F}_1$, the integration will be made only with respect to the random values $\theta_i(t_j)$ specified by item 7.

Let us fix the variables specified in items 1–6. For given τ_l, $l = 1, 2, \ldots$ denote by V_τ union of all possible inspection time points. This set includes the points τ_l, $l = 1, 2, \ldots$, the inspection points $t_1^I, t_2^I, \ldots$ planed in advance (before operation) and also the inspection points which could appear in case of the shift of the schedule for the regular inspections.

Let us define decision functions $\varphi_1, \varphi_2, \varphi_3$ analogously to [3]. We slightly change the definition in comparison with [3]. The sense of the functions φ_1, φ_2, φ_3 can be explained as follows:

- if $\varphi_1\big(\hat{\lambda}(t), v_1(t)\big) > 0$, then terminate the operation of the component after a large inspection;

- if $\varphi_2(\hat{\lambda}(t), v_2(t)) > 0$, then make a repair of the component after a large inspection (in case of $\varphi_1(\hat{\lambda}(t), v_1(t)) \leq 0$) ;
- if $\varphi_2(\hat{\lambda}(t), v_2(t)) \leq 0$, then continue the operation after a large inspection (in this case $\varphi_1(\hat{\lambda}(t), v_1(t)) \leq 0$) ;
- if $\varphi_3(\hat{\lambda}(t), v_3(t)) \leq 0$ then continue the operation after a small inspection without a large inspection.

The decision function $\varphi_3(\hat{\lambda}(t), v_3(t))$ is used after a small inspection. The decision functions $\varphi_2(\hat{\lambda}(t), v_2(t))$ and $\varphi_1(\hat{\lambda}(t), v_1(t))$ are used after a large inspection. Since repairs can be made only if the operation of the component is not stopped, then

$$\varphi_2(\hat{\lambda}(t), v_2(t)) \; < \; \varphi_1(\hat{\lambda}(t), v_1(t)) \, .$$

We denote $\theta^\nu = (\theta_1, \ldots, \theta_\nu)$. The set of inequalities $\theta_1 \geq \theta^{tr}, \ldots, \theta_\nu \geq \theta^{tr}$, we present as follows $\theta^\nu \geq (\theta^{tr})^\nu$. For each $t \in V_\tau$, we denote by $\Theta_{all}(t)$ a family of sets in the space of signals with the structure

$$\left\{\theta^{\nu(t)}(t) \mid \theta^{\nu(t)}(t) \geq (\theta^{tr})^{\nu(t)} , \ \varphi_3(\hat{\lambda}(t), v_3(t)) \leq 0\right\} , \tag{2}$$

$$\left\{\theta^{\nu(t)}(t) \mid \theta^{\nu(t)}(t) \geq (\theta^{tr})^{\nu(t)} , \ \varphi_3(\hat{\lambda}(t), v_3(t)) > 0\right\} , \tag{3}$$

$$\left\{\theta^{\nu(t)}(t) \mid \theta^{\nu(t)}(t) \geq (\theta^{tr})^{\nu(t)} , \ \varphi_2(\hat{\lambda}(t), v_2(t)) \leq 0\right\} , \tag{4}$$

$$\left\{\theta^{\nu(t)}(t) \mid \theta^{\nu(t)}(t) \geq (\theta^{tr})^{\nu(t)} , \ \varphi_2(\hat{\lambda}(t), v_2(t)) > 0, \ \varphi_1(\hat{\lambda}(t), v_1(t)) \leq 0\right\} , \tag{5}$$

$$\left\{\theta^{\nu(t)}(t) \mid \theta^{\nu(t)}(t) \geq (\theta^{tr})^{\nu(t)} , \ \varphi_1(\hat{\lambda}(t), v_1(t)) > 0\right\} . \tag{6}$$

Let $2^{\{0,\ldots,t\}}$ be a set of all subsets of the time points $\{0, \ldots, t\}$. We denote all possible Cartesian products of sets with structure (2) – (6) by $\Theta_{all}^{prod}(t)$, where the time moments are varied from 0 to t. For example, the set $\Theta_{all}(0) \times \ldots \times \Theta_{all}(t)$ belongs to the set $\Theta_{all}^{prod}(t)$. More formally, it can be written as follows

$$\Theta_{all}^{prod}(t) \; = \; \bigcup_{\mathcal{A} \in 2^{\{0,\ldots,t\}}} \; \prod_{z \in \mathcal{A}} \; \Theta_{all}(z) \, .$$

We denote all possible elements from the $\Theta_{all}^{prod}(t)$ which can lead to a repair at the time t by $\Xi_r(t) \subset \Theta_{all}^{prod}(t)$. Analogously,

$$\Xi_{stop}(t) \subset \Theta_{all}^{prod}(t) \, ,$$

$$\Xi_1(t) \subset \Theta_{all}^{prod}(t) \, ,$$

$$\Xi_2(t) \subset \Theta_{all}^{prod}(t)\,,$$

are sets, which lead to a termination of the operation, a big inspection and a small inspection respectively.

The mathematical expectation of the function $\tilde{f}(x,\omega)$ with respect to the σ–algebra $\mathcal{F}_1$ can be written as follows

$$\begin{aligned}
E\,[\tilde{f}(x,\omega) \mid \mathcal{F}_1\,] \stackrel{\text{def}}{=} \hat{f}(x,\omega) \;&=\; G_f E\,[\int_0^{T_{stop}} \lambda(t)\,\mathrm{d}t \mid \mathcal{F}_1\,] \;-\; G_p\,E\,[T_{stop} \mid \mathcal{F}_1\,] \;+\\
&+\; E\,\Big[\sum_{j=1}^{j_{stop}} [G_1\chi_1(t_j) + G_2\chi_2(t_j) + G_r\chi_r(t_j)] \mid \mathcal{F}_1\,\Big] \;=\\
&=\; \sum_{t\in V_r}\;\sum_{\Theta\in\Xi_{stop}(t)} \big(G_f\int_0^t \lambda(t) \;-\; G_p\,t\big)P(\Theta \mid \mathcal{F}_1\,) \;+\\
&+\; \sum_{t\in V_r}\Big[\,G_1\sum_{\Theta\in\Xi_1(t)} P(\Theta \mid \mathcal{F}_1\,) \;+\; G_2\sum_{\Theta\in\Xi_2(t)} P(\Theta \mid \mathcal{F}_1\,) \;+\; G_r\sum_{\Theta\in\Xi_r(t)} P(\Theta \mid \mathcal{F}_1\,)\Big]\,. \qquad (7)
\end{aligned}$$

The last expression can also be presented as

$$\begin{aligned}
\hat{f}(x,\omega) \;&=\; \sum_{t\in V_r}\;\sum_{\Theta\in\Xi_{stop}(t)} \Big[\,G_f\int_0^t \lambda(t) \;-\; G_p\,t \;+\\
&+\; G_1N_1(\Theta) \;+\; G_2N_2(\Theta) \;+\; G_rN_r(\Theta)\,\Big]\,P(\Theta \mid \mathcal{F}_1\,)\,, \qquad (8)
\end{aligned}$$

where $N_1(\Theta)\,, N_2(\Theta)\,$, and $N_r(\Theta)$ denotes number of the large inspections, the small inspections, and the repairs for the sequence of sets Θ . Denote

$$\Upsilon(\Theta,t) \;=\; G_f\int_0^t \lambda(t) \;-\; G_p\,t \;+\; G_1N_1(\Theta) \;+\; G_2N_2(\Theta) \;+\; G_rN_r(\Theta)\,,$$

then

$$\hat{f}(x,\omega) \;=\; \sum_{t\in V_r}\;\sum_{\Theta\in\Xi_{stop}(t)} \Upsilon(\Theta,t)\,P(\Theta \mid \mathcal{F}_1\,)\,. \qquad (9)$$

Since all random variables specified in items 1–6 (see page 3) are fixed, then for

$$\Theta = (\Theta_0 \times \ldots \times \Theta_t) \in \Xi_{stop}(t)\,,$$

the probability measure is presented with the product

$$P(\Theta \mid \mathcal{F}_1\,) = P(\Theta_0 \mid \mathcal{F}_1\,)\,P(\Theta_1 \mid \Theta_0, \mathcal{F}_1\,)\cdot\ldots\cdot P(\Theta_t \mid \Theta_{t-1},\ldots,\Theta_0,\mathcal{F}_1\,)\,. \qquad (10)$$

Let us calculate a conditional probability measure for some concrete set $\Theta_t \in \Theta_{all}(t)$. If, for example,

$$\Theta_t = \big\{\theta^{\nu(t)}(t) \mid \theta^{\nu(t)}(t) \geq (\theta^{tr})^{\nu(t)}\,,\; \varphi_1\big(\hat{\lambda}(t), v_1(t)\big) \leq 0\,,\; \varphi_2\big(\hat{\lambda}(t), v_2(t)\big) > 0\big\}\,, \qquad (11)$$

then (see [3], page 9)

$$P(\Theta_t \mid \Theta_{t-1},\dots,\Theta_0,\mathcal{F}_1\,) = \int\limits_{\substack{\varphi_1(\hat\lambda(t),v_1(t)) \le 0\,, \\ \varphi_2(\hat\lambda(t),v_2(t)) > 0\,, \\ \theta^{\nu(t)} \ge (\theta^{tr})^{\nu(t)}}} \prod_{i=1}^{\nu(t)} g^{tr}\big(\theta_i \mid C(t),\theta_i \ge \theta^{tr}\big)\,\mathrm{d}\theta^{\nu(t)}\,. \tag{12}$$

Analogously, $P(\Theta_z \mid \Theta_{z-1},\dots,\Theta_0,\mathcal{F}_1\,)$ can be calculated for all Θ_z from the sets $\Theta_{all}(z)$, $1 \le z \le T_{\max}$. Since the function $P(\Theta_z \mid \Theta_{z-1},\dots,\Theta_0,\mathcal{F}_1\,)$ is continuous and piecewise smooth with respect to variable x, formula (7) can be used for the calculation of the gradient estimate of the cost function $F(x) = E\,[\,\hat f(x,\omega)\,]$.

3 The Calculation of the Gradients for the Cost Function

The function $\hat f(x,\omega)$ is continuous and piecewise smooth with respect to variable x (see formula (7)). Analogous to [5], we can interchange the gradient sign

$$\nabla_x E\,[\,\hat f(x,\omega)\,] \;=\; E\,[\,\nabla_x \hat f(x,\omega)\,]\,.$$

Let us calculate $\nabla_x \hat f(x,\omega)$ (see (7))

$$\nabla_x \hat f(x,\omega) = \sum_{t\in V_\tau}\;\sum_{\Theta\in\Xi_{stop}(t)} \Upsilon(\Theta,t)\,\nabla_x P(\Theta\;\mid \mathcal{F}_1\,)\,. \tag{13}$$

Thus, to calculate $\nabla_x \hat f(x,\omega)$ we have to calculate $\nabla_x P(\Theta \mid \mathcal{F}_1\,)$ for $\Theta \in \Xi_{stop}(t)$ for $t \in V_\tau$, $0 \le t \le T_{\max}$. If $\Theta = (\Theta_0 \times \ldots \times \Theta_t)$, then (see (10))

$$\nabla_x P(\Theta \mid \mathcal{F}_1\,) \;=\; \nabla_x\big[\,P(\Theta_0 \mid \mathcal{F}_1\,)\,P(\Theta_1 \mid \Theta_0,\mathcal{F}_1\,)\cdot\ldots\cdot P(\Theta_t \mid \Theta_{t-1},\dots,\Theta_0,\mathcal{F}_1\,)\big] \;=$$

$$=\; P(\Theta \mid \mathcal{F}_1\,)\sum_{z=0}^{t}\frac{\nabla_x P(\Theta_z \mid \Theta_{z-1},\dots,\Theta_0,\mathcal{F}_1\,)}{P(\Theta_z \mid \Theta_{z-1},\dots,\Theta_0,\mathcal{F}_1\,)}\,. \tag{14}$$

Now with (13) and (14) we have

$$\nabla_x \hat f(x,\omega) \;=\; \sum_{t\in V_\tau}\;\sum_{\Theta\in\Xi_{stop}(t)}\big[\,\Upsilon(\Theta,t)\sum_{z=0}^{t}\frac{\nabla_x P(\Theta_z \mid \Theta_{z-1},\dots,\Theta_0,\mathcal{F}_1\,)}{P(\Theta_z \mid \Theta_{z-1},\dots,\Theta_0,\mathcal{F}_1\,)}\,\big]\,P(\Theta \mid \mathcal{F}_1\,)\,. \tag{15}$$

Further we show that for any $\Theta_t \in \Theta_{all}(t)$ the calculation of $\nabla_x P(\Theta_t \mid \Theta_{t-1},\dots,\Theta_0,\mathcal{F}_1\,)$ can be reduced to the calculation of the gradient of the functions

$$Z(x) \;=\; \int\limits_{\substack{h_\nu(\theta^\nu) \le x\,, \\ \theta^\nu \ge (\theta^{tr})^\nu}} \tilde g_\nu(\theta^\nu,C)\;\mathrm{d}\theta^\nu\,,\quad Z_1(x) \;=\; \int\limits_{\substack{h_\nu(\theta^\nu) \ge x\,, \\ \theta^\nu \ge (\theta^{tr})^\nu}} \tilde g_\nu(\theta^\nu,C)\;\mathrm{d}\theta^\nu\,, \tag{16}$$

where

$$h_\nu(\theta^\nu) \;=\; \sum_{i=1}^{\nu}\theta_i^{1/\beta}\,, \tag{17}$$

$$q^{tr} = \frac{Q^{tr}}{\sigma_\xi \sqrt{2\pi}}\,, \quad \text{and} \tag{18}$$

$$\tilde{g}_\nu(\theta^\nu, C) = \prod_{i=1}^{\nu} \frac{q^{tr}}{\theta_i} \exp\left\{ -\frac{[\ln\theta_i - (\beta_0 + \beta_1 \ln C_i)]^2}{2\sigma_\xi^2} \right\}. \tag{19}$$

The calculation of a gradient for the conditional probability of sets (2) and (4) can be reduced to the calculation of the gradient of the function $Z(x)$. The calculation of a gradient for the conditional probability of sets (3) and (6) can be reduced to the calculation of a gradient of the function $Z_1(x)$. The calculation of a gradient for conditional probability of set (5) can be reduced to the calculation of a gradient of the function $Z(x)$ or the function $Z_1(x)$. To illustrate this let us calculate, for example, the partial derivative of function (12) with respect to the variable x_1^2

$$\frac{\partial}{\partial x_1^2} P(\Theta_t \mid \Theta_{t-1}, \ldots, \Theta_0, \mathcal{F}_1) = \frac{\partial}{\partial x_1^2} \int\limits_{\substack{\varphi_1(\hat\lambda(t), v_1(t)) \le 0, \\ \varphi_2(\hat\lambda(t), v_2(t)) \ge 0, \\ \theta^{\nu(t)} \ge (\theta^{tr})^{\nu(t)}}} \tilde{g}_{\nu(t)}(\theta^{\nu(t)}, C(t))\, d\theta^{\nu(t)} =$$

$$\frac{\partial}{\partial x_1^2} \left(\int\limits_{\substack{\varphi_1(\hat\lambda(t), v_1(t)) \le 0, \\ \theta^{\nu(t)} \ge (\theta^{tr})^{\nu(t)}}} \tilde{g}_{\nu(t)}(\theta^{\nu(t)}, C(t))\, d\theta^{\nu(t)} - \int\limits_{\substack{\varphi_2(\hat\lambda(t), v_2(t)) \le 0, \\ \theta^{\nu(t)} \ge (\theta^{tr})^{\nu(t)}}} \tilde{g}_{\nu(t)}(\theta^{\nu(t)}, C(t))\, d\theta^{\nu(t)} \right) =$$

$$- \frac{\partial}{\partial x_1^2} \int\limits_{\substack{\varphi_2(\hat\lambda(t), v_2(t)) \le 0, \\ \theta^{\nu(t)} \ge (\theta^{tr})^{\nu(t)}}} \tilde{g}_{\nu(t)}(\theta^{\nu(t)}, C(t))\, d\theta^{\nu(t)}\,. \tag{20}$$

Further, we have (see (31),(32) and (26) in [3])

$$\varphi_2\big(\hat\lambda(t), v_2(t)\big) = \hat\lambda(t) - v_2(t_j) =$$

$$\lambda^1(t) + \mu_1 \sum_{i=1}^{\nu(t)} \exp\left\{ \frac{\ln\theta_i(t) - \beta_0}{\beta_1} \right\} - (x_1^2 + x_2^2 t) =$$

$$\lambda^1(t) + \mu_1 \sum_{i=1}^{\nu(t)} \Big(\exp\{\ln\theta_i(t) - \beta_0\} \Big)^{1/\beta_1} - (x_1^2 + x_2^2 t) =$$

$$\lambda^1(t) + \mu_1 \exp\{-\beta_0/\beta_1\} \sum_{i=1}^{\nu(t)} \theta_i^{1/\beta_1}(t) - (x_1^2 + x_2^2 t) =$$

$$\lambda^1(t) + \mu_1 \exp\{-\beta_0/\beta_1\} h_{\nu(t)}(\theta^{\nu(t)}(t)) - (x_1^2 + x_2^2 t)\,.$$

Thus,

$$\frac{\partial}{\partial x_1^2} P(\Theta_t \mid \Theta_{t-1}, \ldots, \Theta_0, \mathcal{F}_1) =$$

$$= -\frac{\partial}{\partial x_1^2} \int_{\substack{(\lambda^1(t)-x_2^2 t) + \mu_1 \exp\{-\beta_0/\beta_1\} h_{\nu(t)}(\theta^{\nu(t)}) \le x_1^2 , \\ \theta^{\nu(t)} \ge (\theta^{tr})^{\nu(t)}}} \tilde{g}_{\nu(t)}(\theta^{\nu(t)}, C(t)) \ d\theta^{\nu(t)} .$$

and the calculation of the derivative

$$\frac{\partial}{\partial x_1^2} P(\Theta_t \mid \Theta_{t-1}, \ldots, \Theta_0, \mathcal{F}_1)$$

is reduced to the calculation of the derivative of the function $Z(x)$ (see (16)).

For some vectors $y^\nu \in R^\nu$, $\theta^\nu \in R^\nu$, we denote

$$(\theta_i \mid y^\nu) = (y_1, \ldots, y_{i-1}, \theta_i, y_{i+1}, \ldots, y_\nu) ,$$

$$y^{\nu,-i} = (y_1, \ldots, y_{i-1}, y_{i+1}, \ldots, y_\nu) ,$$

$$\rho^\nu(\theta^\nu) = \big(\rho_1(\theta_1), \ldots, \rho_\nu(\theta_\nu)\big) = \frac{\beta}{\nu}\big(\theta_1^{1-\frac{1}{\beta}}, \ldots, \theta_\nu^{1-\frac{1}{\beta}}\big) ,$$

$$\operatorname{div} \rho^\nu(\theta^\nu) = \sum_{i=1}^{\nu} \frac{\partial \rho_i(\theta_i)}{\partial \theta_i} .$$

Now we present a lemma about the derivative of the function $Z(x)$.

Lemma 3.1 *Let the set* $\big\{\theta^\nu \in R^\nu \mid h^\nu(\theta^\nu) \le x , \ \theta^\nu \ge (\theta^{tr})^\nu \big\}$ *be non-empty for some open neighborhood of the point* x. *The function* $Z(x)$ *(see formula* (16)*) is differentiable, and the derivative is given by the formula*

$$\frac{\partial Z(x)}{\partial x} = \int_{\substack{h_\nu(\theta^\nu) \le x , \\ \theta^\nu \ge (\theta^{tr})^\nu}} \operatorname{div}[\, \tilde{g}_\nu(\theta^\nu, C) \rho^\nu(\theta^\nu)] \ d\theta^\nu + \tag{21}$$

$$+ \sum_{i=1}^{\nu} \Big[\rho_i(\theta_i^{tr}) \int_{\substack{h_\nu(\theta_i^{tr} \mid \theta^\nu) \le x , \\ \theta^{\nu,-i} \ge (\theta^{tr})^{\nu,-i}}} \tilde{g}_\nu\big((\theta_i^{tr} \mid \theta^\nu), C\big) \ d\theta^{\nu,-i} \Big] . \tag{22}$$

Proof. Here we give only an idea of the proof. A more detailed proof can be written analogously to [13]. In this case, it is not possible directly to apply the results of the paper [13], since the integration (see (16)) is made over the set with the simple constraints $\theta^\nu \ge (\theta^{tr})^\nu$.

Let us calculate $Z(x + \Delta x)$. Using the Taylor's series it is easy to show that

$$h_\nu\big(y^\nu + \Delta x \rho^\nu(y^\nu)\big) - (x + \Delta x) = h_\nu(y^\nu) - x + o(\,|\, \Delta x \,|\,) . \tag{23}$$

We will make an increment Δx in the argument of the function $Z(x)$ and change the variables $\theta^\nu = y^\nu + \Delta x \, \rho^\nu(y^\nu)$ in the integral

$$Z(x + \Delta x) = \int_{\substack{h_\nu(\theta^\nu) \le x + \Delta x , \\ \theta^\nu \ge (\theta^{tr})^\nu}} \tilde{g}_\nu(\theta^\nu, C) \ d\theta^\nu =$$

$$= \int\limits_{\substack{h_\nu(y^\nu+\Delta x\rho^\nu(y^\nu)) \le x+\Delta x\,, \\ y^\nu+\Delta x\rho^\nu(y^\nu) \ge (\theta^{tr})^\nu}} \tilde g_\nu\big(y^\nu + \Delta x \rho^\nu(y^\nu), C\big) \prod_{i=1}^{\nu}\Big(1 + \Delta x \frac{\partial}{\partial y_i}\rho_i(y_i)\Big)\ \mathrm{d}y^\nu =$$

$$= \int\limits_{\substack{h_\nu(y^\nu) \le x - o(|\Delta x|)\,, \\ y^\nu+\Delta x\rho^\nu(y^\nu) \ge (\theta^{tr})^\nu}} [\tilde g_\nu(y^\nu, C) + \langle \nabla_{y^\nu}\tilde g_\nu(y^\nu, C)\,,\, \Delta x\,\rho^\nu(y^\nu)\rangle + o(\,|\,\Delta x\,|\,)] \times$$

$$\times \left[1 + \Delta x \sum_{i=1}^{\nu} \frac{\partial}{\partial y_i}\rho_i(y_i) + o(\,|\,\Delta x\,|\,)\right]\ \mathrm{d}y^\nu =$$

$$= \int\limits_{\substack{h_\nu(y^\nu) \le x\,, \\ y^\nu+\Delta x\rho^\nu(y^\nu) \ge (\theta^{tr})^\nu}} \left[\tilde g_\nu(y^\nu, C) + \Delta x\, \tilde g_\nu(y^\nu, C) \sum_{i=1}^{\nu} \frac{\partial}{\partial y_i}\rho_i(y_i) + \right.$$

$$\left. + \langle \nabla_{y^\nu}\tilde g_\nu(y^\nu, C)\,,\, \Delta x\,\rho^\nu(y^\nu)\rangle\right]\ \mathrm{d}y^\nu + o(\,|\,\Delta x\,|\,) =$$

$$= \int\limits_{\substack{h_\nu(y^\nu) \le x\,, \\ y^\nu+\Delta x\rho^\nu(y^\nu) \ge (\theta^{tr})^\nu}} \big\{\tilde g_\nu(y^\nu, C) + \Delta x\,\mathrm{div}\,[\,\tilde g_\nu(y^\nu, C)\,\rho^\nu(y^\nu)\,]\big\}\ \mathrm{d}y^\nu + o(\,|\,\Delta x\,|\,) =$$

$$= \int\limits_{\substack{h_\nu(y^\nu) \le x\,, \\ y^\nu+\Delta x\rho^\nu(y^\nu) \ge (\theta^{tr})^\nu}} \tilde g_\nu(y^\nu, C)\ \mathrm{d}y^\nu + \tag{24}$$

$$+ \Delta x \int\limits_{\substack{h_\nu(y^\nu) \le x\,, \\ y^\nu+\Delta x\rho^\nu(y^\nu) \ge (\theta^{tr})^\nu}} \mathrm{div}\,[\,\tilde g_\nu(y^\nu, C)\,\rho^\nu(y^\nu)\,]\ \mathrm{d}y^\nu + o(\,|\,\Delta x\,|\,)\,. \tag{25}$$

Integral (25) can be presented as

$$\Delta x \int\limits_{\substack{h_\nu(y^\nu) \le x\,, \\ y^\nu+\Delta x\rho^\nu(y^\nu) \ge (\theta^{tr})^\nu}} \mathrm{div}\,[\,\tilde g_\nu(y^\nu, C)\,\rho^\nu(y^\nu)\,]\ \mathrm{d}y^\nu =$$

$$= \Delta x \int\limits_{\substack{h_\nu(y^\nu) \le x\,, \\ y^\nu \ge (\theta^{tr})^\nu}} \mathrm{div}\,[\,\tilde g_\nu(y^\nu, C)\,\rho^\nu(y^\nu)\,]\ \mathrm{d}y^\nu + o(\,|\,\Delta x\,|\,)\,. \tag{26}$$

We can write integral (24) as follows

$$\int\limits_{\substack{h_\nu(y^\nu) \le x\,, \\ y^\nu+\Delta x\rho^\nu(y^\nu) \ge (\theta^{tr})^\nu}} \tilde g_\nu(y^\nu, C)\ \mathrm{d}y^\nu =$$

$$= Z(x) + \int\limits_{\substack{h_\nu(y^\nu) \le x\,, \\ y^\nu+\Delta x\rho^\nu(y^\nu) \ge (\theta^{tr})^\nu}} \tilde g_\nu(y^\nu, C)\ \mathrm{d}y^\nu - \int\limits_{\substack{h_\nu(y^\nu) \le x\,, \\ y^\nu \ge (\theta^{tr})^\nu}} \tilde g_\nu(y^\nu, C)\ \mathrm{d}y^\nu = Z(x) +$$

$$\sum_{i=1}^{\nu}\left[\ \int\limits_{\substack{h_\nu(y^\nu) \le x\,, \\ \theta_i^{tr} - \Delta x\rho_i(y_i) \le y_i \le \theta_i^{tr}\,, \\ y^{\nu,-i} \ge (\theta^{tr})^{\nu,-i}}} \tilde g_\nu(y^\nu, C)\ \mathrm{d}y^\nu - \int\limits_{\substack{h_\nu(y^\nu) \le x\,, \\ \theta_i^{tr} \le y_i \le \theta_i^{tr} - \Delta x\rho_i(y_i)\,, \\ y^{\nu,-i} \ge (\theta^{tr})^{\nu,-i}}} \tilde g_\nu(y^\nu, C)\ \mathrm{d}y^\nu\right] + o(\,|\,\Delta x\,|\,) =$$

$$= Z(x) + \Delta x \sum_{i=1}^{\nu} \Bigg[\rho_i(\theta_i^{tr}) \int\limits_{\substack{h_\nu(\theta_i^{tr} | y^\nu) \le x, \\ y^{\nu,-i} \ge (\theta^{tr})^{\nu,-i}}} \tilde{g}_\nu\big((\theta_i^{tr} \mid y^\nu), C\big) \ \mathrm{d}y^{\nu,-i} \Bigg] + o(|\Delta x|). \tag{27}$$

Formulae (24) - (27) imply

$$Z(x + \Delta x) = Z(x) + \Delta x \Bigg(\int\limits_{\substack{h_\nu(y^\nu) \le x, \\ y^\nu \ge (\theta^{tr})^\nu}} \operatorname{div}[\tilde{g}_\nu(y^\nu, C)\, \rho^\nu(y^\nu)] \ \mathrm{d}y^\nu +$$

$$+ \sum_{i=1}^{\nu} \Bigg[\rho_i(\theta_i^{tr}) \int\limits_{\substack{h_\nu(\theta_i^{tr} | y^\nu) \le x, \\ y^{\nu,-i} \ge (\theta^{tr})^{\nu,-i}}} \tilde{g}_\nu\big((\theta_i^{tr} \mid y^\nu), C\big) \ \mathrm{d}y^{\nu,-i} \Bigg] \Bigg) + o(|\Delta x|).$$

The last equation proves the lemma. □

Analogously to lemma 3.1 we can prove a lemma about the derivative of the function $Z_1(x)$.

Lemma 3.2 *Let the set* $\{\theta^\nu \in R^\nu \mid h^\nu(\theta^\nu) \le x,\ \theta^\nu \ge (\theta^{tr})^\nu\}$ *be non-empty for some open neighborhood of the point* x. *The function* $Z_1(x)$ *(see formula* (16)*) is differentiable, and the derivative is given by the formula*

$$\frac{\partial Z_1(x)}{\partial x} = \int\limits_{\substack{h_\nu(\theta^\nu) \ge x, \\ \theta^\nu \ge (\theta^{tr})^\nu}} \operatorname{div}[\tilde{g}_\nu(\theta^\nu, C)\, \rho^\nu(\theta^\nu)] \ \mathrm{d}\theta^\nu + \tag{28}$$

$$+ \sum_{i=1}^{\nu} \Bigg[\rho_i(\theta_i^{tr}) \int\limits_{\substack{h_\nu(\theta_i^{tr} | \theta^\nu) \ge x, \\ \theta^{\nu,-i} \ge (\theta^{tr})^{\nu,-i}}} \tilde{g}_\nu\big((\theta_i^{tr} \mid \theta^\nu), C\big) \ \mathrm{d}\theta^{\nu,-i} \Bigg]. \tag{29}$$

Remark. Lemmas 3.1 and 3.2 are also true for an arbitrary function $\tilde{g}_\nu(\theta^\nu, C)$ and any functions $h_\nu(\theta^\nu)$ and $\rho^\nu(\theta^\nu)$ satisfying equation (23). We suppose that for these functions integrals (21),(22), (28) and (29) exist.

Finally, we can calculate the gradient of the objective function $F(x) = E[\hat{f}(x,\omega)]$ using formula (15) and lemmas 3.1 and 3.2.

4 A Solution Algorithm

4.1 A Stochastic Quasi–Gradient Algorithm

Let us consider the following optimization problem

$$F(x) = E[\hat{f}(x,\omega)] \quad \rightarrow \quad \min_{x \in X}, \tag{30}$$

subject to the dynamics of the process $\lambda(t), \hat{\lambda}(t)$ (see Sections 2 and Subsection 3.2. in [3]). By X we denote a feasible set for the decision vector x , i.e.

$$X = \{x \in R^6 \colon \underline{x}_l \le x_l \le \overline{x}_l \text{ , for } l = 1, \ldots, 6\},$$

where $\underline{x}_l$, $\overline{x}_l$, $l = 1,\ldots,6$ are low and upper bounds for variables x_l, $l = 1,\ldots,6$.

To solve this problem one can use a gradient-based method. Note that for the computation of the gradient $\nabla_x F(x)$, formulae (15) and

$$\nabla_x F(x) \;=\; E\,[\,\nabla_x \hat{f}(x,\omega)\,] \tag{31}$$

can be used. However, these formulae would be difficult to implement, since it would require computation of number multi–dimensional integrals. In order to avoid this, stochastic quasi-gradient algorithms can be used (see, for example, [1] and [2]). One of the most simple stochastic quasi-gradient algorithms has the form

$$x^{s+1} = \Pi_X(x^s - \rho_s \xi^s)\,, \tag{32}$$

where s is a number of the algorithm iteration; x^s is the approximation of the extremum on the s^{th} iteration; $\Pi_X(\cdot)$ is the orthoprojection operation on the convex set X; $\rho_s > 0$ is a step size; and ξ^s is a stochastic quasi-gradient satisfying the following property

$$E\,[\,\xi^s \mid x^0, \xi^0, x^1, \xi^1, \ldots, x^s\,] = \nabla_x F(x^s)$$

i.e. the conditional expectation of the vector ξ^s is equal to the gradient of the function $F(x)$ at the point x^s. Results of computation experiments show that the algorithm (32) rapidly leads to the point of the extremum if the objective function is not ill-conditioned, i.e., for non-"ravine" functions. In cases where the function $f(x)$ is "ravine", the algorithm gets stuck "at the bottom of the ravine". This difficulty may be overcome by using more complicated stochastic quasigradient algorithms with averaging (see, for example, [10]), or variable metrics algorithm [12] with metric transformation.

For improving the convergence rate of the algorithm (32), we use here stochastic quasi-gradient algorithm with adaptively controlled step sizes (see [12] and [14]) and the scaling procedure, suggested by Saridis [7]

$$x^{s+1} = \Pi_X(x^s - \rho_s H^s \xi^s)\,, \tag{33}$$

where H^s is a scaling matrix. The scaling procedure and adaptive control of step size considerably improve the practical convergence rate of the algorithm (32). The scaling matrix is calculated as follows

$$H^{s+1} = \begin{pmatrix} h_{1,s+1} & & 0 \\ & \ddots & \\ 0 & & h_{n,s+1} \end{pmatrix}, \quad H^0 = \begin{pmatrix} 1/n & & 0 \\ & \ddots & \\ 0 & & 1/n \end{pmatrix},$$

$$h_{i,s+1} = \alpha h_{i,s} + (1-\alpha)\beta_{i,s+1} \, ,$$

$$\beta_{i,s+1} = \begin{cases} 0\, , & \text{if } \ \xi_i^{s+1}(x_i^s - x_i^{s+1}) \leq 0, \ k_{s+1} \neq 0 \, , \\ 1/n\, , & \text{if } \ k_{s+1} = 0 \, , \\ 1/k_{s+1}\, , & \text{if } \ \xi_i^{s+1}(x_i^s - x_i^{s+1}) > 0, \ k_{s+1} \neq 0 \, , \end{cases}$$

where

$n = 6$ is the dimension of the control vector x ;

k_{s+1}, $0 \leq k_{s+1} \leq n$ is the quantity of numbers i for which $\xi_i^{s+1}(x_i^s - x_i^{s+1}) > 0$.

Step size ρ_s is given by the following recursive relations

$$T_s = \langle \xi^{s+1}, \Delta x^{s+1} \rangle \, ,$$

$$\Delta x^{s+1} = x^{s+1} - x^s \, ,$$

$$z_s \doteq z_{s-1} + (|T_s| - z_{s-1})\, D \, , \quad z_{-1} = 0 \, ,$$

$$\tilde{\rho}_{s+1} = \rho_s a^{T_s/z_s} \begin{cases} 1\, , & \text{if } \ T_s > 0 \, , \\ U\, , & \text{if } \ T_s \leq 0 \, , \end{cases} \tag{34}$$

$$\rho_{s+1} = \begin{cases} \rho_s 3\, , & \text{if } \quad \tilde{\rho}_{s+1}/\rho_s > 3, \\ \rho_s/4\, , & \text{if } \quad \tilde{\rho}_{s+1}/\rho_s < 1/4, \\ \tilde{\rho}_{s+1}\, , & \text{otherwise} \quad . \end{cases} \tag{35}$$

In formulae (34) and (35), an additional reduction of the step size occurs only if the value T_s is negative. The recommended values of parameters are

$$a = 2, \ U = 0.8, \ D = 0.2, \ \alpha = 0.5 \, .$$

Note that the considered algorithm has a natural termination criterion for the iteration process. In the neighborhood of extremum the value $\|\Delta x^{s+1}\|$ becomes small and tends to zero. Therefore, for the procedure to terminate, we may use the following averaged value Q_s

$$Q_s = Q_{s-1} + (\|\Delta x^{s+1}\| - Q_{s-1})\, D \, , \quad Q_{-1} = 0 \, .$$

If $Q_s \leq \epsilon$, the process is terminated. Here ϵ is some positive constant which characterizes the required precision of the solution.

4.2 A Stochastic Quasi-Gradient Calculation

Equation (31) implies that in the case considered a stochastic quasi-gradient can be computed by the formula $\nabla_x \hat{f}(x^s, \omega^s)$. Here ω^s denotes a sample at the s^{th} iteration of random values specified in items 1-6 (see, page 3). It is much simpler to calculate $\nabla_x \hat{f}(x^s, \omega^s)$ than

$E\,[\nabla_x \hat{f}(x^s,\omega)]$, because it does not require integration with respect to random variables specified in items 1-6. Nevertheless, it is necessary to calculate $\nabla_x P(\Theta \mid \mathcal{F}_1)$ for some $\Theta \in \Theta_{all}^{prod}(t)$ for $t \in V_\tau$, $0 \le t \le T_{\max}$ (see 13) . We do not actually need to calculate $\nabla_x P(\Theta \mid \mathcal{F}_1)$ exactly. As it was mentioned before, it is sufficient to use Monte-Carlo estimate of this vector.

The gradient $\nabla_x \hat{f}(x^s,\omega^s)$ can be estimated during one run of the model. Suppose, some sample of the model is known. It means that all random values specified in items 1-7 (see page 3) are sampled. In this case, some sequence of sets $\Theta_0^s \times \ldots \times \Theta_{T_{stop}}^s$ is calculated together with the value $\Upsilon(\Theta^s, T_{stop})$. It can be seeing from formulae (9) and (15) that the difference between them is in the term

$$\sum_{z=0}^{T_{stop}} \frac{\nabla_x P(\Theta_z \mid \Theta_{z-1},\ldots,\Theta_0,\mathcal{F}_1)}{P(\Theta_z \mid \Theta_{z-1},\ldots,\Theta_0,\mathcal{F}_1)} \,.$$

Therefore, to calculate an estimate of the gradient $\nabla_x \hat{f}(x^s,\omega^s)$, we can multiply the value $\Upsilon(\Theta^s, T_{stop})$ by

$$\sum_{z=0}^{T_{stop}} \frac{\nabla_x P(\Theta_z^s \mid \Theta_{z-1}^s,\ldots,\Theta_0^s,\mathcal{F}_1)}{P(\Theta_z^s \mid \Theta_{z-1}^s,\ldots,\Theta_0^s,\mathcal{F}_1)} \,. \tag{36}$$

As it was mentioned before, the calculation of $\nabla_x P(\Theta_z^s \mid \Theta_{z-1}^s,\ldots,\Theta_0^s,\mathcal{F}_1)$ can be reduced to the calculation of the derivative of the functions $Z(x)$ and $Z_1(x)$ (see (16)). According to lemmas 3.1 and 3.2 the derivatives $\frac{\partial Z(x)}{\partial x}$ and $\frac{\partial Z_1(x)}{\partial x}$ are presented as a sum of integrals. The calculation of a gradient for the conditional probability of sets (2), (4) and (5) can be reduced to the calculation of a gradient of the function $Z(x)$. According to lemma 3.1,

$$\frac{\partial Z(x)}{\partial x} = \int\limits_{\substack{h_\nu(\theta^\nu) \le x, \\ \theta^\nu \ge (\theta^{tr})^\nu}} \frac{\operatorname{div}[\tilde{g}_\nu(\theta^\nu,C)\,\rho^\nu(\theta^\nu)]}{\tilde{g}_\nu(\theta^\nu,C)} \tilde{g}_\nu(\theta^\nu,C)\;d\theta^\nu \; + \tag{37}$$

$$+ \sum_{i=1}^{\nu} \Big[\rho_i(\theta_i^{tr}) \int\limits_{\substack{h_\nu(\theta_i^{tr}|\theta^\nu) \le x, \\ \theta^{\nu,-i} \ge (\theta^{tr})^{\nu,-i}}} \tilde{g}_\nu\big((\theta_i^{tr} \mid \theta^\nu),C\big)\;d\theta^{\nu,-i} \Big] \,. \tag{38}$$

Integrals (37) and (38) are calculated with the density function $\tilde{g}_\nu(\theta^\nu,C)$. Thus, we can use for the estimation of the $\frac{\partial Z(x)}{\partial x}$ the same random generators as in the generation of the trajectory of the model. Analogously, the calculation of a gradient for the conditional probability of sets (3) and (6) can be reduced to the calculation of the gradient of the function $Z_1(x)$

$$\frac{\partial Z_1(x)}{\partial x} = \int\limits_{\substack{h_\nu(\theta^\nu) \ge x, \\ \theta^\nu \ge (\theta^{tr})^\nu}} \frac{\operatorname{div}[\tilde{g}_\nu(\theta^\nu,C)\,\rho^\nu(\theta^\nu)]}{\tilde{g}_\nu(\theta^\nu,C)} \tilde{g}_\nu(\theta^\nu,C)\;d\theta^\nu \; + \tag{39}$$

$$+ \sum_{i=1}^{\nu} \Big[\rho_i(\theta_i^{tr}) \int\limits_{\substack{h_\nu(\theta_i^{tr}|\theta^\nu) \ge x, \\ \theta^{\nu,-i} \ge (\theta^{tr})^{\nu,-i}}} \tilde{g}_\nu\big((\theta_i^{tr} \mid \theta^\nu),C\big)\;d\theta^{\nu,-i} \Big] \,. \tag{40}$$

Instead of the exact calculation of the value $P(\Theta_z^s \mid \Theta_{z-1}^s, \ldots, \Theta_0^s, \mathcal{F}_1)$ in the formula (36), we can also use some approximation of this integral (for example, Monte-Carlo approximation).

The trajectories and stochastic quasi-gradients, ξ^s , should be sampled for each value of the argument of the objective function, x^s .

5 Conclusions

In the paper we have proposed a stochastic quasi–gradient algorithm for the optimization of operational strategies. We have derived an expression for calculating the gradient of the objective function. The gradient is presented as a sum of rather simple integrals. Furthermore, the gradient is estimated with Monte-Carlo techniques. This expression is used for the calculation of stochastic quasi–gradient (stochastic estimate of the gradient).

The model discussed in the earlier working paper [3] and in this paper is intended for finding of the optimal operational strategies for an inspected component. Similar problems are encountered in many contexts in the field of reliability and risk analysis of technical systems. The models of failure phenomena in these analyses are case dependent, but the structure of the optimization problems is the same as here and the same solution techniques can be applied. One of the most fruitful application areas is the optimization of the operational and maintenance strategies of components, which are subject to aging (see, for example, [15]).

References

[1] Ermoliev, Yu. (1983): Stochastic Quasi-Gradient Methods and Their Applications to System Optimization. *Stochastics*, **4**. pp. 1-36.

[2] Ljung, L. and T. Söderström (1983): *Theory and Practice of Recursive Identification.* MIT Press. 529 p.

[3] Pulkkinen, A. and S. Urya'sev (1990): *Optimal Operational Strategies for an Inspected Component - Statement of the Problem.* Working Paper, International Institute for Applied Systems Analysis, Laxenburg, Austria, WP-90-62, 22 p.

[4] Raik, E. (1975): The Differentiability in the Parameter of the Probability Function and Optimization of the Probability Function via the Stochastic Pseudogradient Method, *Izvestiya Akad. Nayk Est. SSR,* Phis. Math., **24, 1** pp. 3-6 (in Russian).

[5] Rockafellar R.T., and J.-B. Wets (1982): On the Interchange of Subdifferentiation and Conditional Expectation for Convex Functionals. *Stochastics*, **7**, pp. 173-182.

[6] Roenko, N. (1983): *Stochastic Programming Problems with Integral Functionals over Multivalued Mappings.* Synopsis of Thesis, USSR, Kiev, (in Russian).

[7] Saridis, G.M. (1970): Learning applied to successive approximation algorithms, *IEEE Trans. Syst. Sci. Cybern.*, 1970, SSC-6, Apr., pp. 97–103.

[8] Simon, J. (1989): *Second Variation in Domain Optimization Problems,* In "International Series of Numerical Mathematics, **91**, Eds. F.Kappel, K.Kunish and W.Schappacher, Birkhauser Verlag, pp. 361–378

[9] Stadlober, E. and R.Kremer (1991): *Sampling from Discrete and Continuous Distributions with C-RAND,* Lecture Notes in Economical Math. Systems, Springer, New York (to appear).

[10] Syski W. (1988): A Method of Stochastic Subgradients with Complete Feedback Stepsize Rule for Convex Stochastic Approximation Problems. *J. of Optim. Theory and Applic.* Vol. **39**, No. **2**, pp. 487–505.

[11] Uryas'ev, S.P. (1986): *Stochastic Quasi-Gradient Algorithms with Adaptively Controlled Parameters.* Working Paper, International Institute for Applied Systems Analysis, Laxenburg, Austria, WP-86-32, 27 p.

[12] Uryas'ev, S.P. (1989): *A Stochastic Quasi-Gradient Algorithms with Variable Metric*, Working Paper, International Institute for Applied Systems Analysis, Laxenburg, Austria, WP-89-98, 13 p.

[13] Uryas'ev, S.P. (1989): A differentiation Formula for Integrals over Sets Given by Inclusion. *Numerical Functional Analysis and Optimization.* **10**(**7** & **8**), (1989), 827–841.

[14] Uryas'ev, S.P. (1990): *Adaptive Algorithms of Stochastic Optimization and Game Theory*, Nauka, Moscow, 182 p. (in Russian).

[15] Vesely, W.E. (1991): Incorporating Aging Effects into Probabilistic Risk Analysis Using a Taylor Expansion Approach. *Reliability Engineering & System Safety.* **32**, 315–337.