ESTIMATION AND RECOGNITION IN STOCHASTIC SYSTEMS WITH TIME-DELAY (MEMORY) CONTINUOUS-DISCRETE OBSERVATIONS

N. Dyomin*, **S. Rozhkova****

**Department of Applied Mathematics and Cybernetics,*
Tomsk State University, 36 Lenin ave., 634050 Tomsk, Russia
svrhm@rambler.ru
***Department of Natural Sciences and Mathematics,,*
Tomsk Polytechnic University, 30 Lenin ave., 634034 Tomsk, Russia
oni@cam.tpu.ru

Abstract: The paper considers: 1) the determination problem of optimal filtering and interpolation estimations in the mean square sense for stochastic processes with continuous time on continuous-discrete time-delay (memory) observations; 2) the determination problem of likelihood ratio for recognition of the stochastic processes on observations of the same type; 3) the anomalous noises detection in discrete channel with memory observations.

Keywords: stochastic systems, filtering, interpolation, recognition, time-delay, memory.

1. INTRODUCTION

The modern stage in the synthesis theory of the algorithms for stochastic processes treatment began from paper (Kalman, 1960). The pair $\{x_t; y_t\}$, $t \geq 0$, where x_t is an unobservable process and y_t is an observable process, is the basic mathematical object in the Kalman systems. This theory received further development in (Kallianpur, 1980; Liptser and Shiryayev, I. 1977, II. 1978; Sage and Melse, 1972; Van Trees, 1971). A new class of problems is the situation when: 1) the observable process is a set of process with continuous time z_t and discrete time $\eta(t_m)$, i.e. $y_t = y(t, t_m) = \{z_t; \eta(t_m);\ m = 0, 1, \cdots\}$; 2) the processes z_t and $\eta(t_m)$ are time-delay processes and possess the memory relatively unobservable process, i.e. depend not only on the current values but also on the past values of process x_t. For continuous-discrete observations with memory the filtering problem was considered in (Abakumova *et. al.*, 1995a, 1995b), the extrapolation problem was considered in (Dyomin *et. al.*, 1997; 2000; 2003), the recognition problem with fixed memory was investigated in (Dyomin *et. al.*, 2001). The present paper considers problems of adaptive in sense (Lainiotis, 1971) estimation and recognition with moving memory (see Remark 2).

Notations: $P\{\cdot\}$ is event probability, $E\{\cdot\}$ is expectation; $\mathcal{N}\{y; a, B\}$ denotes Gaussian probability density function with given parameters a and B; $|D|$ is a determinant of matrix; D^{-1} is the inversion matrix of D; D^T denotes transpose of a matrix or a vector; I and O are unit and zero matrixes of appropriate size; $B > 0$ is positively defined matrix; a vector is column-vector.

2. STATEMENT OF THE PROBLEM

On the probability space (Liptser and Shiryayev, I. 1977, II. 1978) $(\Omega, \mathcal{F}, F = (\mathcal{F}_t)_{t\geq 0}, \mathcal{P})$ the unobservable n-dimensional process x_t and the observable l-dimensional process z_t with continuous time are described by the stochastic differential equations (in the Ito sense)

$$dx_t = f(t, x_t, z, \theta)dt + \Phi_1(t, x_t, z, \theta)dw_t,\ \ t \geq 0, \quad (1)$$

$$dz_t = h(t, x_t, x_{t-t_1^*}, \cdots, x_{t-t_N^*}, z, \theta)dt + \Phi_2(t, z)dv_t, \quad (2)$$

and the observable q – dimensional process with discrete time $\eta(t_m)$ has the form

$$\eta(t_m) = g(t_m, x_{t_m}, x_{t_m - t_1^*}, \cdots, x_{t_m - t_N^*}, z, \theta) + \Phi_3(t_m, z, \theta)\xi(t_m), \quad m = 0, 1, \cdots, \tag{3}$$

where $0 < t - t_N^* < \cdots < t - t_1^* < t_m \le t$, $t_k^* = const$, $k = \overline{1;N}$. It is assumed: 1) the parameter θ can accept values from the set $\Theta = \{\theta_0, \theta_1, \cdots, \theta_r\}$ with the a priori probabilities $p_0(\theta_j) = P\{\theta = \theta_j\}$, $j = \overline{0;r}$; 2) w_t and v_t are standard Wiener processes; 3) for all $\theta \in \Theta$ coefficients of the equations (1), (2) satisfy conditions (Kallianpur, 1980; Liptser and Shiryayev, I. 1977, II. 1978) and $g(\cdot)$ is continuous for all arguments; 4) $\xi(t_m)$ is the standard white Gaussian sequence; 5) x_0, w_t, v_t, $\xi(t_m)$, θ are assumed to be statistically independent; 6) $Q(\cdot) = \Phi_1(\cdot)\Phi_1^T(\cdot) > 0$, $R(\cdot) = \Phi_2(\cdot)\Phi_2^T(\cdot) > 0$, $V(\cdot) = \Phi_3(\cdot)\Phi_3^T(\cdot) > 0$ for all $\theta \in \Theta$; 7) the initial density functions $p_0(x|\theta_j) = \partial P\{x_0 \le x|\theta = \theta_j\}/\partial x$, $j = \overline{0;r}$, are given.

Problem 1 (estimation). To determine the optimal estimations $\mu(t)$ and $\mu(t - t_k^*, t)$ in mean-square sense for x_t and $x_{t-t_k^*}$, $k = \overline{1;N}$, respectively, on the realizations set $\mathcal{F}_t^{z,\eta} = \{z_0^t; \eta_0^m\}$ where $z_0^t = \{z_s; 0 \le s \le t\}$, $\eta_0^m = \{\eta(t_0), \eta(t_1), \cdots, \eta(t_m); 0 \le t_0 < \cdots < t_m \le t\}$.

Problem 2 (recognition). To determine the likelihood ratio $\Lambda_t = (\theta_j : \theta_\alpha)$ for the hypotheses recognition problem $\mathcal{H}_j\{\theta = \theta_j\}$ and $\mathcal{H}_\alpha\{\theta = \theta_\alpha\}$ on the realizations set $\mathcal{F}_t^{z,\eta}$.

Remark 1. The term "delay" means time delay in mathematical models of both observed and nonobserved processes (Dion *et. al.*, 1999; Basin and Martinez-Zuniga, 2003;Wang and Ho, 2003). Since in this paper delay is present only in mathematical models of observed processes, the given paper uses the term "memory" meaning that the current values of observed processes have a memory relative to the past values of nonobserved process according to (Abakumova *et. al.*, 1995a, 1995b; Dyomin et. al., 1997, 2000, 2001, 2003).

Remark 2. The models of the processes z_t, $\eta(t_m)$ of (2), (3) are adequate to the observations with moving memory (Abakumova *et. al.*, 1995a, 1995b; Dyomin *et. al.*, 1997; 2000). If $t - t_k^* = \tau_k = const$ in (2) and $t_m - t_k^* = \tau_k = const$ in (3), $k = \overline{1;N}$, observations have fixed memory (Dyomin *et. al.*, 2001; 2003).

3. MAIN RESULTS

For $k = \overline{1;N}$, we shall enter extended processes

$$\widetilde{x}_{t-t^*}^{N-k+1} = \begin{bmatrix} x_{t-t_k^*} \\ x_{t-t_{k+1}^*} \\ \cdots \\ x_{t-t_N^*} \end{bmatrix}, \ \widetilde{x}_{t,t-t^*}^{N+1} = \begin{bmatrix} x_t \\ \widetilde{x}_{t-t^*}^N \end{bmatrix} \tag{4}$$

and operators

$$L_{s,y}[\varphi(s,y,\cdot)] = -\sum_{i=1}^n \frac{\partial[f_i(s,y,\cdot)\varphi(s,y,\cdot)]}{\partial y_i} + \frac{1}{2}\sum_{i,j=1}^n \frac{\partial^2[Q_{ij}(s,y,\cdot)\varphi(s,y,\cdot)]}{\partial y_i \partial y_j},$$

$$L_{s,y}^*[\varphi(s,y,\cdot)] = \sum_{i=1}^n f_i(s,y,\cdot)\frac{\partial\varphi(s,y,\cdot)}{\partial y_i} + \frac{1}{2}\sum_{i,j=1}^n Q_{ij}(s,y,\cdot)\frac{\partial^2\varphi(s,y,\cdot)}{\partial y_i \partial y_j}, \tag{5}$$

$$\mathcal{L}_{s,y}[\varphi_1(s,y,\cdot);\varphi_2(s,y,\cdot)] = \frac{\varphi_1(s,y,\cdot)}{\varphi_2(s,y,\cdot)} \times L_{s,y}[\varphi_2(s,y,\cdot)] - \varphi_2(s,y,\cdot)L_{s,y}^*\left[\frac{\varphi_1(s,y,\cdot)}{\varphi_2(s,y,\cdot)}\right].$$

Let for $j = \overline{0;r}$, $k = \overline{1;N}$,

$$p_t(\theta_j) = P\{\theta = \theta_j | z_0^t, \eta_0^m\}, \tag{6}$$

$$p_t(x;\widetilde{x}_N|\theta_j) = \partial^{N+1}P\{x_t \le x; \widetilde{x}_{t-t^*}^N \le \widetilde{x}_N | \theta = \theta_j, z_0^t, \eta_0^m\}/\partial x \partial \widetilde{x}_N, \tag{7}$$

$$p_{t-t_k^*}(\widetilde{x}_{N-k+1}|\theta_j) = \partial^{N-k+1}P\{\widetilde{x}_{t-t^*}^{N-k+1} \le \widetilde{x}_{N-k+1} | \theta = \theta_j, z_0^{t-t_k^*}, \eta_0^{m_k}\}/\partial \widetilde{x}_{N-k+1}. \tag{8}$$

Theorem 1. The posterior probabilities (6) and the posterior conditional probability densities (7) on the time intervals $t_m \le t < t_{m+1}$ are determined by the equations

$$d_t p_t(\theta_j) = p_t(\theta_j)[\overline{h(t,z|\theta_j)} - \overline{h(t,z)}]^T \times R^{-1}(t,z)[dz_t - \overline{h(t,z)}dt], \tag{9}$$

$$d_t p_t(x;\widetilde{x}_N|\theta_j) = \{L_{t,x}[p_t(x;\widetilde{x}_N|\theta_j)) + \sum_{k=1}^N \mathcal{L}_{t-t_k^*,x_k}[p_t(x;\widetilde{x}_N|\theta_j);p_{t-t_k^*}(\widetilde{x}_{N-k+1}|\theta_j)]\}dt + p_t(x;\widetilde{x}_N|\theta_j)[h(t,x,\widetilde{x}_N,z,\theta_j) - \overline{h(t,z|\theta_j)}]^T \times R^{-1}(t,z)[dz_t - \overline{h(t,z|\theta_j)}dt] \tag{10}$$

with the initial conditions

$$p_{t_m}(\theta_j) = \frac{C(\eta(t_m),z|\theta_j)}{C(\eta(t_m),z)}p_{t_m-0}(\theta_j), \tag{11}$$

$$p_{t_m}(x;\widetilde{x}_N|\theta_j) = \frac{C(\eta(t_m),z,x,\widetilde{x}_N,\theta_j)}{C(\eta(t_m),z|\theta_j)}p_{t_m-0}(x;\widetilde{x}_N|\theta_j), \tag{12}$$

where

$$\overline{h(t,z)} = E\{h(t,x_t,\widetilde{x}_{t-t^*}^N,z,\theta)|z_0^t,\eta_0^m\}, \tag{13}$$

$$\overline{h(t,z_t|\theta_j)} = E\{h(t,x_t,\widetilde{x}_{t-t^*}^N,z,\theta)|\theta = \theta_j, z_0^t,\eta_0^m\}, \tag{14}$$

$$C(\eta(t_m),z) = E\{C(\eta(t_m),z,x_{t_m}, \widetilde{x}_{t_m-t^*}^N,\theta)|z_0^{t_m},\eta_0^{m-1}\}, \tag{15}$$

$$C(\eta(t_m),z|\theta_j) = E\{C(\eta(t_m),z,x_{t_m}, \widetilde{x}_{t_m-t^*}^N,\theta)|\theta = \theta_j, z_0^{t_m},\eta_0^{m-1}\}, \tag{16}$$

$$C(\eta(t_m),z,x,\widetilde{x}_N,\theta_j) = |V(t_m,z,\theta_j)|^{-1/2} \times \exp\{-\tfrac{1}{2}[\eta(t_m) - g(t_m,x,\widetilde{x}_N,z,\theta_j)]^T \times V^{-1}(t_m,z,\theta_j)[\eta(t_m) - g(t_m,x,\widetilde{x}_N,z,\theta_j)]\}. \tag{17}$$

Proof. For $(j = \overline{0;r})$

$$p_t(x;\widetilde{x}_N;\theta_j) = \partial^{N+1}\mathcal{P}\{x_t \le x; \widetilde{x}^N_{t-t^*} \le \widetilde{x}_N; \theta = \theta_j | z_0^t, \eta_0^m\} / \partial x \partial \widetilde{x}_N \quad (18)$$

on the time intervals $t_m \le t < t_{m+1}$ we have the equations

$$d_t p_t(x;\widetilde{x}_N;\theta_j) = \{L_{t,x}[p_t(x;\widetilde{x}_N;\theta_j)] + \sum_{k=1}^{N} \mathcal{L}_{t-t_k^*,x_k}[p_t(x;\widetilde{x}_N;\theta_j); p_{t-t_k^*}(\widetilde{x}_{N-k+1};\theta_j)]\}dt + p_t(x;\widetilde{x}_N;\theta_j)[h(t,x,\widetilde{x}_N,z,\theta_j) - \overline{h(t,z)}]^T \times R^{-1}(t,z)[dz_t - \overline{h(t,z)}dt], \quad (19)$$

$$p_{t-t_k^*}(\widetilde{x}_{N-k+1};\theta_j) = \partial^{N-k+1}\mathcal{P}\{\widetilde{x}^{N-k+1}_{t-t^*} \le \widetilde{x}_{N-k+1}; \theta = \theta_j | z_0^{t-t_k^*}, \eta_0^{m_k}\} / \partial \widetilde{x}_{N-k+1} \quad (20)$$

with the initial conditions

$$p_{t_m}(x;\widetilde{x}_N;\theta_j) = \frac{C(\eta(t_m), z, x, \widetilde{x}_N, \theta_j)}{C(\eta(t_m), z)} p_{t_m-0}(x;\widetilde{x}_N;\theta_j), \quad (21)$$

which follows from (Abakumova *et. al.*, 1995a; Theorems 1, 3). Since

$$p_t(x;\widetilde{x}_N;\theta_j) = p_t(x;\widetilde{x}_N|\theta_j) p_t(\theta_j), \quad (22)$$

integrating (19), (21) with respect to $\{x;\widetilde{x}_N\}$ taking into account (14), (16) yields (9), (11), and (12) follows from (11), (21). As innovation process

$$\widetilde{z}_t = \int_0^t [dz_s - \overline{h(s,z)ds}] \quad (23)$$

is such that $\widetilde{Z}_t = (\widetilde{z}_t, \mathcal{F}_t^z)$ is Wiener process (Kallianpur, 1980; Liptser and Shiryayev, I. 1977, II. 1978) with

$$E\{\widetilde{z}_t \widetilde{z}_t^T | \mathcal{F}_t^z\} = \int_0^t R(s,z) ds, \quad (24)$$

then differentiating according to Ito formula (Kallianpur, 1980; Liptser and Shiryayev, I. 1977, II. 1978)

$$p_t(x;\widetilde{x}_N|\theta_j) = p_t(x;\widetilde{x}_N;\theta_j) / p_t(\theta_j) \quad (25)$$

taking into account (9), (19), (23), (24) yields (10).

Theorem 2. The mean-square optimal filtering estimation $\mu(t)$ and interpolation estimations $\mu(t-t_k^*, t)$, $k = \overline{1;N}$ are determined by

$$\mu(t) = \sum_{j=0}^{r} \mu(t|\theta_j) p_t(\theta_j),$$
$$\mu(t-t_k^*, t) = \sum_{j=0}^{r} \mu(t-t_k^*, t|\theta_j) p_t(\theta_j), \quad (26)$$

where (see (4), (7))

$$\mu(t|\theta_j) = \int \cdots \int x p_t(x;\widetilde{x}_N|\theta_j) dx d\widetilde{x}_N,$$
$$\mu(t-t_k^*, t|\theta_j) = \int \cdots \int x_k p_t(x;\widetilde{x}_N|\theta_j) dx d\widetilde{x}_N, \quad (27)$$

and $p_t(\theta_j)$, $p_t(x;\widetilde{x}_N|\theta_j)$ are determined by Theorem 1.

Proof. As the meant-square optimal estimations are posterior expectations (Kallianpur, 1980; Liptser and Shiryayev, I. 1977, II. 1978) then

$$\mu(t) = E\{x_t | z_0^t, \eta_0^m\},$$
$$\mu(t-t_k^*, t) = E\{x_{t-t_k^*} | z_0^t, \eta_0^m\}. \quad (28)$$

As $\mathcal{F}_t^{z,\eta} \subseteq \mathcal{F}_t^{z,\eta,\theta}$ then according to (27), (28)

$$\mu(t) = E\{E\{x_t | \mathcal{F}_t^{z,\eta,\theta}\} | \mathcal{F}_t^{z,\eta}\} = E\{\mu(t|\theta) | \mathcal{F}_t^{z,\eta}\}, \quad (29)$$
$$\mu(t-t_k^*, t) = E\{E\{x_{t-t_k^*} | \mathcal{F}_t^{z,\eta,\theta}\} | \mathcal{F}_t^{z,\eta}\} = E\{\mu(t-t_k^*|\theta) | \mathcal{F}_t^{z,\eta}\}.$$

Then formulas (26) follow from (6), (7), (27)–(29).

Remark 3. In Lainiotis sense (Lainiotis, 1971) estimations (26) are adaptive estimations of filtering and interpolation with respect to the set of continuous and discrete observations with moving memory.

Theorem 3. For the likelihood ratio $\Lambda_t(\theta_j : \theta_\alpha)$ in the problem of hypotheses recognition $\mathcal{H}_j\{\theta = \theta_j\}$ and $\mathcal{H}_\alpha\{\theta = \theta_\alpha\}$, $j = \overline{0;r}$, $\alpha = \overline{0;r}$, $j \ne \alpha$, we have

$$\Lambda_t(\theta_j : \theta_\alpha) = \Lambda_{t-t_1^*-0}(\theta_j : \theta_\alpha) \times \exp\left\{ \sum_{t-t_1^* \le t_i \le t} \ln\left[\frac{C(\eta(t_i), z|\theta_j)}{C(\eta(t_i), z|\theta_\alpha)}\right] + \int_{t-t_1^*}^{t} [\overline{h(s,z|\theta_j)} - \overline{h(s,z|\theta_\alpha)}]^T R^{-1}(s,z) \times [dz_s - \tfrac{1}{2}\overline{h(s,z|\theta_j)}ds - \tfrac{1}{2}\overline{h(s,z|\theta_\alpha)}ds] \right\}. \quad (30)$$

Proof. Let us denote $P_t(\theta_j : \theta_\alpha) = p_t(\theta_j)/p_t(\theta_\alpha)$. As $\Lambda_t(\theta_j : \theta_\alpha) = [p_0(\theta_\alpha)/p_0(\theta_j)] P_t(\theta_j : \theta_\alpha)$ (Sage and Melse, 1972; Van Trees, 1971) then differentiating according to Ito formula taking into account (9), (23), (24) yields for $t_m \le t < t_{m+1}$ the equation

$$d_t \Lambda_t(\theta_j : \theta_\alpha) = \Lambda_t(\theta_j : \theta_\alpha)[\overline{h(t,z|\theta_j)} - \overline{h(t,z|\theta_\alpha)}]^T \times R^{-1}(t,z)[\overline{h(t,z)} - \overline{h(t,z|\theta_\alpha)}]dt + \Lambda_t(\theta_j : \theta_\alpha)[\overline{h(t,z|\theta_j)} - \overline{h(t,z|\theta_\alpha)}]^T \times R^{-1}(t,z)[dz_t - \overline{h(t,z)}dt] \quad (31)$$

with the initial condition

$$\Lambda_{t_m}(\theta_j : \theta_\alpha) = \frac{C(\eta(t_m), z|\theta_j)}{C(\eta(t_m), z|\theta_\alpha)} \Lambda_{t_m-0}(\theta_j : \theta_\alpha), \quad (32)$$

which follows from (11). Let us denote $\widetilde{\Lambda}_t(\theta_j : \theta_\alpha) = \ln\{\Lambda_t(\theta_j : \theta_\alpha)\}$. Differentiating according to Ito formula taking into account (23), (24) yields for $t_m \le t < t_{m+1}$ the equation

$$d_t\widetilde{\Lambda}_t(\theta_j : \theta_\alpha) = [\overline{h(t,z|\theta_j)} - \overline{h(t,z|\theta_\alpha)}]^T R^{-1}(t,z) \times[dz_t - (1/2)(\overline{h(t,z|\theta_j)} + \overline{h(t,z|\theta_\alpha)})dt] \quad (33)$$

with the initial condition

$$\widetilde{\Lambda}_{t_m}(\theta_j : \theta_\alpha) = \widetilde{\Lambda}_{t_m-0}(\theta_j : \theta_\alpha) + \ln\left[\frac{C(\eta(t_m)|\theta_j)}{C(\eta(t_m)|\theta_\alpha)}\right]. \quad (34)$$

Then (30) follows from (33), (34).

From Theorem 1–3 follows that the effective calculation $\mu(t)$, $\mu(t-t_k^*, t)$, $\Lambda_t(\theta_j : \theta_\alpha)$ are realized provided that there is a possibility of effective calculation $\mu(t|\theta_j)$, $\mu(t-t_k^*, t|\theta_j)$, $\overline{h(t,z|\theta_j)}$, $C(\eta(t_m), z|\theta_j)$. In following item the particular case of the processes x_t, z_t, $\eta(t_m)$ supposing such possibility is considered.

4. CONDITIONALLY-GAUSSIAN CASE

Proposition 1. Assume (see (1)–(4))

$$\begin{aligned} f(\cdot) &= f(t,z,\theta) + F(t,z,\theta)x_t, \\ \Phi_1(\cdot) &= \Phi_1(t,z,\theta), \\ h(\cdot) &= h(t,z,\theta) + H_{0,N}(t,z,\theta)\widetilde{x}^{N+1}_{t,t-t^*}, \quad (35) \\ g(\cdot) &= g(t_m,z,\theta) + G_{0,N}(t_m,z,\theta)\widetilde{x}^{N+1}_{t_m,t_m-t^*}, \\ p_0(x|\theta_j) &= \mathcal{N}\{x;\mu_0^j,\Gamma_0^j\},\ j=\overline{0;r}, \end{aligned}$$

where

$$H_{0,N}(\cdot) = [H_0(\cdot)\vdots H_1(\cdot)\vdots\cdots\vdots H_N(\cdot)],$$

$$G_{0,N}(\cdot) = [G_0(\cdot)\vdots G_1(\cdot)\vdots\cdots\vdots G_N(\cdot)]. \quad (36)$$

Then $(t-\widetilde{t}_N^* = [t-t_1^*, \cdots, t-t_N^*])$

$$p_t(x;\widetilde{x}_N|\theta_j) = p_t(\widetilde{x}_{N+1}|\theta_j) = \mathcal{N}\{\widetilde{x}_{N+1}; \widetilde{\mu}_{N+1}(t-\widetilde{t}_N^*, t|\theta_j), \widetilde{\Gamma}_{N+1}(t-\widetilde{t}_N^*, t|\theta_j)\}, \quad (37)$$

$$\widetilde{\mu}_{N+1}(\cdot|\theta_j) = \begin{bmatrix} \mu(t|\theta_j) \\ \widetilde{\mu}_N(t-\widetilde{t}_N^*, t|\theta_j) \end{bmatrix} = \begin{bmatrix} \mu(t|\theta_j) \\ \mu(t-t_k^*, t|\theta_j) \end{bmatrix},\quad k=\overline{1;N}, \quad (38)$$

$$\widetilde{\Gamma}_{N+1}(\cdot|\theta_j) = \begin{bmatrix} \Gamma(t|\theta_j) & \widetilde{\Gamma}_{0N}(t-\widetilde{t}_N^*, t|\theta_j) \\ \widetilde{\Gamma}_{0N}^T(\cdot) & \widetilde{\Gamma}_N(t-\widetilde{t}_N^*, t|\theta_j) \end{bmatrix} = \begin{bmatrix} \Gamma(t|\theta_j) & \Gamma_{01}(t-t_1^*, t|\theta_j) & \Gamma_{0k}(t-t_k^*, t|\theta_j) \\ \Gamma_{01}^T(\cdot) & \Gamma_{11}(t-t_1^*, t|\theta_j) & \Gamma_{lk}(t-t_l^*, t-t_k^*, t|\theta_j) \\ \Gamma_{0k}^T(\cdot) & \Gamma_{lk}^T(\cdot) & \Gamma_{kk}(t-t_k^*, t|\theta_j) \end{bmatrix}, (39)$$

$$l=\overline{1;N-1},\quad k=\overline{2;N},\quad k>l,$$

and block components of distribution (37) parameters $\widetilde{\mu}_{N+1}(\cdot)$, $\widetilde{\Gamma}_{N+1}(\cdot)$ for all $j=\overline{0;r}$ are defined by system of differential-recurrent equations (11)–(20), (24)–(29) in (Abakumova *et. al.*, 1995b).

The formulated proposition follows immediately from the results of (Abakumova *et. al.*, 1995b).

Theorem 4. Let (35) be satisfied. Then we have Theorems 1–3 where $\mu(t|\theta_j)$, $\mu(t-t_k^*, t|\theta_j)$ are determined by Proposition 1 and

$$\overline{h(t,z|\theta_j)} = h(t,z,\theta_j) + H_{0,N}(t,z,\theta_j)\widetilde{\mu}_{N+1}(t-\widetilde{t}_N^*, t|\theta_j), \quad (40)$$

$$\begin{aligned} C(\eta(t_m),z|\theta_j) &= |V(t_m,z,\theta_j)|^{-1/2} \\ &\times\exp\{-(1/2)[\eta(t_m)-g(t_m,z,\theta_j)]^T \\ &\times V^{-1}(t_m,z,\theta_j)[\eta(t_m)-g(t_m,z,\theta_j)]\} \\ &\times\frac{|\widetilde{\Gamma}_{N+1}(t_m-\widetilde{t}_N^*, t_m|\theta_j)|^{1/2}}{|\widetilde{\Gamma}_{N+1}(t_m-\widetilde{t}_N^*, t_m-0|\theta_j)|^{1/2}} \\ &\times\frac{\exp\{(1/2)a(t_m|\theta_j)\}}{\exp\{(1/2)a(t_m-0,\theta_j)\}}, \quad (41) \end{aligned}$$

$$\begin{aligned} a(t_m|\theta_j) &= \widetilde{\mu}^T_{N+1}(t_m-\widetilde{t}_N^*, t_m|\theta_j) \\ &\times\widetilde{\Gamma}^{-1}_{N+1}(t_m-\widetilde{t}_N^*, t_m|\theta_j)\widetilde{\mu}_{N+1}(t_m-\widetilde{t}_N^*, t_m|\theta_j), \\ a(t_m-0|\theta_j) &= \widetilde{\mu}^T_{N+1}(t_m-\widetilde{t}_N^*, t_m-0|\theta_j) \\ &\times\widetilde{\Gamma}^{-1}_{N+1}(t_m-\widetilde{t}_N^*, t_m-0|\theta_j) \\ &\times\widetilde{\mu}_{N+1}(t_m-\widetilde{t}_N^*, t_m-0|\theta_j). \quad (42) \end{aligned}$$

Proof. The formula (40) follows from (14), (35)–(38), and (41) are proved by analogy with (4.26) in (Dyomin *et. al.*, 2001).

Let us consider the case when coefficients of (35), (36) are independent of z and solutions are made in moments t_m of discrete observations recept, i.e. only by values of $\Lambda_{t_m}(\theta_j : \theta_\alpha)$.

Theorem 5. Let coefficients of (35), (36) are independent of z. Then

$$\Lambda_{t_m}(\theta_j : \theta_\alpha) = \frac{|W(t_m|\theta_j)|^{-1/2}}{|W(t_m|\theta_\alpha)|^{-1/2}} \times\frac{\exp\{-\frac{1}{2}A^T(t_m,\theta_j)W^{-1}(t_m|\theta_j)A(t_m,\theta_j)\}}{\exp\{-\frac{1}{2}A^T(t_m,\theta_\alpha)W^{-1}(t_m|\theta_\alpha)A(t_m,\theta_\alpha)\}}, (43)$$

$$A(t_m,\theta_i) = [\widetilde{\eta}(t_m|\theta_i) - g(t_m,\theta_i)],$$

where for all $i=\overline{0;r}$

$$\widetilde{\eta}(t_m|\theta_i) = \eta(t_m) - G_{0,N}(t_m,\theta_i)\widetilde{\mu}_{N+1}(t_m-\widetilde{t}_N^*, t_m-0|\theta_i), \quad (44)$$

$$W(t_m|\theta_i) = V(t_m,\theta_i) + G_{0,N}(t_m,\theta_i) \times\widetilde{\Gamma}_{N+1}(t_m-\widetilde{t}_N^*, t_m-0|\theta_i)G^T_{0,N}(t_m,\theta_i). \quad (45)$$

Proof. From (3), (35), (36), (44) taking into account independence coefficients from z follow that process $\widetilde{\eta}(t_m|\theta_i)$ by hypothesis $\mathcal{H}_i(\theta=\theta_i)$ is Gaussian with parameters $g(t_m,\theta_i)$ and $W(t_m|\theta_i)$, i.e.

$$p_{t_m}(\widetilde{\eta}|\mathcal{H}_i) = \mathcal{N}\{\widetilde{\eta}; g(t_m,\theta_i), W(t_m|\theta_i)\}. \quad (46)$$

Then for

$$\Lambda_{t_m}(\theta_j:\theta_\alpha)=p_{t_m}(\widetilde{\eta}(t_m|\theta_j))/p_{t_m}(\widetilde{\eta}(t_m|\theta_\alpha)) \quad (47)$$

from (46), (47) follows formula (43).

Theorem 6. Let conditions of Theorem 5 are satisfied and $f(\cdot)$, $F(\cdot)$, $\Phi_1(\cdot)$, $h(\cdot)$, $H_{0,N}(\cdot)$, $G_{0,N}(\cdot)$ are independent of θ. Then for Kullback divergences (Kullback, 1960) for all $j=\overline{0;r}$, $\alpha=\overline{0;r}$, $j\neq\alpha$,

$$I_{t_m}(j:\alpha)=E\{\ln[\Lambda_{t_m}(\theta_j:\theta_\alpha)]|\mathcal{H}_j\} \quad (48)$$

take place formulas

$$\begin{aligned} I_{t_m}(j:\alpha) &= \frac{1}{2}\ln\frac{|W(t_m|\theta_\alpha)|}{|W(t_m|\theta_j)|} \\ &+\frac{1}{2}tr[W^{-1}(t_m|\theta_\alpha)W(t_m|\theta_j)] \\ &+\frac{1}{2}[g(t_m,\theta_j)-g(t_m,\theta_\alpha)]^TW^{-1}(t_m|\theta_\alpha) \\ &\times[g(t_m,\theta_j)-g(t_m,\theta_\alpha)]-\frac{q}{2}. \quad (49)\end{aligned}$$

Formula (49) arises from (43), (46), (48).

5. EXAMPLE

Let x_t, z_t, $\eta(t_m)$ are scalar stationary processes and are represented by

$$\begin{aligned} dx_t &= -ax_tdt+\Phi_1dw_t, \quad a>0, \\ dz_t &= H_0x_tdt+\Phi_2dv_t, \\ \eta(t_m) &= G_0x_{t_m}+G_1x_{t_m-t^*}+\xi_0(t_m)+\theta\xi_1(t_m), \quad (50)\end{aligned}$$

i.e. x_t is Ornstein-Uhlenbeck process (Kallianpur, 1980), z_t is process without memory, $\xi_0\sim\mathcal{N}\{0,V_0\}$, $\xi_1\sim\mathcal{N}\{b,V_1\}$, $\Theta=\{\theta_0;\theta_1\}=\{0;1\}$, $\xi_0(t_m)$ is a regular noise, $\xi_1(t_m)$ is an anomalous noise. Therefore the problem of hypotheses recognition $\mathcal{H}_0\{\theta=\theta_0\}$ and $\mathcal{H}_1\{\theta=\theta_1\}$ is problem of anomalous noise detection. The case of rare observations is considered when solution is made only with respect to the current discrete observation $\eta(t_m)$.

Proposition 2. For $\Lambda_{t_m}(\theta_1:\theta_0)$ and Kullback divergences $I_{t_m}(1:0)$, $I_{t_m}(0:1)$ we have

$$\begin{aligned} \Lambda_{t_m}(\theta_1:\theta_0) &= (W_0/W_1)^{1/2} \\ &\times\exp\left\{-\frac{[\widetilde{\eta}(t_m)-q\sqrt{V_0}]^2}{2W_1}\right\}\exp\left\{\frac{[\widetilde{\eta}^2(t_m)]}{2W_0}\right\}, \quad (51)\\ I_{t_m}(1:0) &= \frac{1}{2}\left[\ln\left(\frac{V_0+\gamma g(t^*)}{(l+1)V_0+\gamma g(t^*)}\right)\right. \\ &\left.+\frac{(q^2+l)V_0}{V_0+\gamma g(t^*)}\right], \quad (52)\\ I_{t_m}(0:1) &= \frac{1}{2}\left[\ln\left(\frac{(l+1)V_0+\gamma g(t^*)}{V_0+\gamma g(t^*)}\right)\right. \\ &\left.+\frac{(q^2-l)V_0}{(l+1)V_0+\gamma g(t^*)}\right], \quad (53)\end{aligned}$$

where l and q are such that $V_1=lV_0$, $b=q\sqrt{V_0}$,

$$\begin{aligned} \widetilde{\eta}(t_m) &= \eta(t_m)-G_0\mu(t_m-0) \\ &-G_1\mu(t_m-t^*,t_m-0), \\ \gamma &= (\lambda-a)/\delta, \quad \lambda=\sqrt{a^2+\delta Q}, \\ \delta &= H_0^2/R, \quad Q=\Phi_1^2, \quad R=\Phi_2^2, \\ g(t^*) &= G_0^2+G_1^2[æ+(1-æ)\exp\{-2\lambda t^*\}] \\ &+2G_0G_1\exp\{-\lambda t^*\}, \quad æ=(\lambda+a)/2\lambda. \quad (54)\end{aligned}$$

This Proposition follows from Theorems 5, 6 taking into account (50) and formula (3.19) in (Dyomin *et. al.*, 2000).

Let $\widetilde{I}_{t_m}(1:0)$ and $\widetilde{I}_{t_m}(0:1)$ denote the corresponding values in the case of observations $\eta(t_m)$ without memory, when $G_1=0$. Then $\Delta I_{t_m}(1:0)=I_{t_m}(1:0)-\widetilde{I}_{t_m}(1:0)$, $\Delta I_{t_m}(0:1)=I_{t_m}(0:1)-\widetilde{I}_{t_m}(0:1)$ will characterize the observation effectiveness with memory in regard to the observations without memory with respect to probable errors $\alpha=P\{\theta=\theta_1|\theta=\theta_0\}$ and $\beta=P\{\theta=\theta_0|\theta=\theta_1\}$. Lower boundaries $\alpha^*=\inf\alpha$ and $\beta^*=\inf\beta$ can be find from Kullback inequalities (Kullback, 1960)

$$\begin{aligned} I_{t_m}(1:0) &\geq \beta\ln[\beta/(1-\alpha)] \\ &+(1-\beta)\ln[(1-\beta)/\alpha], \quad (55)\\ I_{t_m}(0:1) &\geq \alpha\ln[\alpha/(1-\beta)] \\ &+(1-\alpha)\ln[(1-\alpha)/\beta]. \quad (56)\end{aligned}$$

Since inequalities $\Delta I_{t_m}(1:0)>0$, $\Delta I_{t_m}(0:1)>0$ correspond to inequalities $\Delta\alpha^*<0$, $\Delta\beta^*<0$ according to (55) and (56), then a channel with memory is more effective then a channel without memory with respect to a probability of false detection α and probability of anomalous noise omission β. From (52), (53) for $\Delta I_{t_m}(1:0)$, $\Delta I_{t_m}(0:1)$ follow formulas

$$\begin{aligned} \Delta I_{t_m}(1:0) &= \Delta I_1(1:0)+\Delta I_2(1:0), \\ \Delta I_{t_m}(0:1) &= \Delta I_1(0:1)+\Delta I_2(0:1), \quad (57)\\ \Delta I_1(1:0) &= (1/2)\ln[d_2/d_1], \\ \Delta I_1(0:1) &= (1/2)\ln[d_1/d_2], \quad (58)\\ d_1 &= [(l+1)V+g(t^*)\gamma][V+G_0^2\gamma], \\ d_2 &= [V+g(t^*)\gamma][(l+1)V+G_0^2\gamma], \quad (59)\\ \Delta I_2(1:0) &= (1/2)(q^2+l)V([V+g(t^*)\gamma]^{-1} \\ &-[V+G_0^2\gamma]^{-1}), \quad (60)\\ \Delta I_2(0:1) &= (1/2)(q^2-l)V([(l+1)V \\ &+g(t^*)\gamma]^{-1}-[(l+1)V+G_0^2\gamma]^{-1}). \quad (61)\end{aligned}$$

The research of the dependencies $\Delta I_{t_m}(1:0)$, $\Delta I_{t_m}(0:1)$ as functions of the memory depth t^* brings us to the following result.

Proposition 3. Let

$$(G_0,G_1)\in\mathbf{G}=\{(G_0,G_1):|G_0+G_1|<|G_0|\}. \quad (62)$$

Then $\Delta I_{t_m}(1:0)$ and $\Delta I_{t_m}(0:1)$ by $t^* \uparrow_0^\infty$ are monotonically decreasing from the values $\Delta I^0_{t_m}(1:0) > 0$ and $\Delta I^0_{t_m}(0:1) > 0$ up to the values $\Delta I^\infty_{t_m}(1:0) < 0$ and $\Delta I^\infty_{t_m}(0:1) < 0$. The value t^* subject to $I_{t_m}(1:0) = \Delta I_{t_m}(0:1) = 0$ can be defined as the effective depth of memory determined by the formula

$$t^*_{eff} = \frac{1}{\lambda} \ln \frac{|G_0| + \sqrt{|G_0|^2 - æ(1-æ)|G_1|^2}}{æ|G_1|}. \quad (63)$$

The values $\Delta I^0_{t_m}(1:0)$, $\Delta I^0_{t_m}(0:1)$ and $\Delta I^\infty_{t_m}(1:0)$, $\Delta I^\infty_{t_m}(0:1)$ are determined by the formulas (57)–(61), where $g(t^*) = g_0 = (G_0 + G_1)^2$ and $g(t^*) = g_\infty = G_0^2 + æG_1^2$, respectively.

The influence of continuous observations z_t on quality detection is carried out by means of parameter æ, and when z_t is absent ($\delta = 0$) then æ $= 1$. In this case from (63) it follows that

$$t^*_{eff} = (1/a))\ln(2|G_0|/|G_1|), \quad (64)$$

i.e. formula (5.22) in (Dyomin *et. al.*, 2001) is referred to as a particular case (63).

6. CONCLUSION

Under conditions (35) the obtained solution includes solutions of particular problems: 1) there is no discrete observations ($G_{0,N}(\cdot) \equiv O$) and continuous observations with memory, without memory ($H_k(\cdot) \equiv O, k = \overline{1;N}$), with lag ($H_0(\cdot) \equiv O$); 2) there is no continuous observations ($H_{0,N}(\cdot) \equiv O$) and discrete observations with memory, without memory ($G_k(\cdot) \equiv O, k = \overline{1;N}$), with lag ($G_0(\cdot) \equiv O$).

As it follows from the example considered the presence of memory can either improve or worsen the quality of recognition procedure and estimation procedure (Abakumova *et. al.*, 1995b; Dyomin *et. al.*, 1997; 2000).

REFERENCES

Abakumova, O.L. , N.S. Dyomin and T.V. Sushko (1995a). Filtering of stochastic processes for continuous and discrete observations with memory I. Main equation of non-linear filtering. *Automat. and Telemekh.* **9**, 49-59 (in Russian).

Abakumova, O.L., N.S. Dyomin and T.V. Sushko (1995b). Filtering of stochastic processes for continuous and discrete observations with memory II. Synthesis of filters. *Automat. and Remote Control.* **56**(10), 1383–1393.

Basin, M. and R. Martinez-Zuniga (2003). Optimal filtering for linear systems with multiple delays in observations. *13th IFAC Symp. on Syst. Ident., Preprints*, 1042–1047.

Dion, J.M., J.L. Dugard and M. Fliess (1999). *Linear Time-Delay Systems.* Pergamon, London.

Dyomin, N.S., S.V. Rozhkova and O.V. Rozhkova (2000). Generalized moving extrapolation of stochastic processes based on the totality of continuous and discrete observations with memory. *J. Comput. Syst. Sci. Intern.,* **39**(4), 523-535.

Dyomin, N.S., S.V. Rozhkova and O.V. Rozhkova (2001). Likelihood ratio determination for stochastic processes recognition problem with respect to the set of continuous and discrete memory observations. *Informatica,* **12**(2), 263–284.

Dyomin, N.S., I.E. Safronova and S.V. Rozhkova (2003). Information amount determination for joint problem of filtering and generalized extrapolation of stochastic processes with respect to the set of continuous and discrete observations. *Informatica,***14**(3), 295–322.

Dyomin N.S., T.V. Sushko and A.V. Yakovleva (1997). Generalized inverse extrapolation of stochastic processes by an aggregate of continuous-discrete observations with memory. *J. Comput. Syst. Sci. Intern.* **36**(4), 543–554.

Kallianpur G. (1980). *Stochastic Filtering Theory.* Springer-Verlag, New York.

Kalman, R.E. (1960). A new approach to linear filtering and prediction problems. *Trans. ASME. J.Basic Eng., Ser.D.,* **82**(March), 35-45.

Kullback, S (1960). *Information Theory and Statistics.* John Wiley, New York.

Lainiotis D.G. (1971). Optimal adaptive estimation: structure and parameter adaptation. *IEEE Trans. Aut. Contr.* **AC-16**, N.2,160-170.

Liptser, R. Sh. and A.N. Shiryayev (I.1977, II.1978). *Statistics of Random Processes.* Springer-Verlag, New York.

Sage, A.P. and J.L. Melse (1972). *Estimation Theory with Application to Communication and Control.* Mc Graw-Hill, New York.

Van Trees H. (1971). *Detection, Estimation and Modulation Theory.* Wiley, New-York.

Wang, Z. and D.W.C. Ho (2003). Filtering on nonlinear time-delay stochastic systems. *Automatica,* **39**(1), 101–109.