INFORMATICA, 2003, Vol. 14, No. 3, 295–322

Information Amount Determination for Joint Problem of Filtering and Generalized Extrapolation of Stochastic Processes with Respect to the Set of Continuous and Discrete Memory Observations

Nikolas DYOMIN, Irina SAFRONOVA

Department of Applied Mathematics and Cybernetics, Tomsk State University
36 Lenin ave., Tomsk, Russia
e-mail: svrhm@mail2000.ru

Svetlana ROZHKOVA

Department of Natural Sciences and Mathematics, Tomsk Politechnic University
30 Lenin ave., Tomsk, Russia
e-mail: svrhm@mail2000.ru

Received: December 2002

Abstract. This paper considers an information aspect of the problem of the joint filtering and generalized extrapolation, when the output of observation channels (data transmission) is the realizations set of the processes with continuous and discrete time, which depend on both the current and the past values of unobservable process (useful signal). The relations defining time evolution of Shannon information are obtained. The particular problems of the memory channels information efficiency and optimal transmission of stochastic processes, with applying the general results are considered.

Key words: filtering, extrapolation, information amount, memory, feedback.

1. Introduction

In the Kalman systems (Kalman, 1960; Kalman and Bucy, 1961) the pair of processes $\{x_t; y_t\}$ with continuous or discrete time, where x_t is an unobservable process, and y_t is an observable process, is the basic mathematical object. The situation is generalized, when x_t is the process with continuous time, and $y_t = y(t, t_m) = \{z_t, \eta(t_m)\}$, $m = 0, 1, \ldots$, i.e., one can observe set of the processes with continuous z_t and discrete $\eta(t_m)$ time, which possess the memory relatively unobservable process and depend on the current and the past values of process x_t. For similar class of processes the filtering problem was considered in (Abakumova *et al.*, 1995a; 1995b), the generalized extrapolation problem was considered in (Dyomin *et al.*, 1997; 2000) and the recognition problem was considered in (Dyomin *et al.*, 2001).

Any statistic problem has an informative aspect (Stratonovitch, 1975), the essence of which is to find corresponding information amounts about unobservable process values (useful signal), which are contained in the realizations of the observable processes (an output signals of a transmission channels). Furthermore, awareness of information amount makes possible to investigate the questions those are specific in information theory, such as minimization of the error of signal reproduction (Shannon and Weaver, 1949; Gallager, 1968), maximization of the capacity of transmission channels (Ihara, 1990), optimal transmission of signals (Liptser, 1974), as well as the questions of information substantiation of estimation problems (Arimoto, 1971; Tomita, *et al.*, 1976). Basing on the results (Abakumova *et al.*, 1995a; 1995b; Dyomin *et al.*, 1997; 2000), with the use of the methods (Liptser, 1974; Dyomin and Korotkevich, 1983; 1987) this paper considers the questions of finding of Shannon measures of the information amount about the values of the unobservable process in the current x_t and the arbitrary number $x_{s_1}, \dots, x_{s_L}$ of future instants, which are contained in the realizations of the observable processes z_t, $\eta(t_m)$, depending on the current x_t and on the arbitrary number $x_{\tau_1}, \dots, x_{\tau_N}$ of the past values of unobservable process. The research of informative efficiency of memory channels relative to memoryless channels and the optimal transmission of stochastic processes under feedback are carried out on the basis of general results in particular cases.

Used notations: $\mathcal{P}\{\cdot\}$ is event probability; $M\{\cdot\}$ denotes the expectation operator; $\mathcal{N}\{y; a, B\}$ denotes Gaussian probability density function with given parameters a and B; $|\cdot|$ is a determinant of the matrix; $\mathrm{tr}[\cdot]$ is a trace of the matrix; I_k is the $(k \times k)$ identity matrix; O is the zero matrix of the corresponding dimension; B^{-1} is the inversion matrix of B; $B > 0$ and $B \geqslant 0$ are the properties of positive and nonnegative definiteness of the matrix B, respectively; vector x is a column-vector; if $\varphi(x)$ is scalar function of n -dimensional argument x, then $\partial\varphi/\partial x$ is a column-vector with the components $\partial\varphi/\partial x_k$, $k = \overline{1;n}$, and $\partial^2\varphi/\partial x^2$ is a matrix with the components $\partial^2\varphi/\partial x_k \partial x_l$, $k = \overline{1;n}$, $l = \overline{1;n}$; $\partial\varphi(x_t)/\partial x_t$ and $\partial^2\varphi(x_t)/\partial x_t^2$ denote $\partial\varphi(x)/\partial x|_{x=x_t}$ and $\partial^2\varphi(x)/\partial x^2|_{x=x_t}$.

2. Statement of the Problem

On the probability space $(\Omega, \mathcal{F}, F = (\mathcal{F}_t)_{t\geqslant 0}, \mathcal{P})$ the unobsevable n-dimensional process x_t (useful signal) and observable l-dimensional process z_t (an output signal of a continuous transmission channel) are defined by the stochastic differential equations (Kallianpur, 1980; Liptser and Shiryayev, 1977; 1978)

$$\mathrm{d}\,x_t = f(t, x_t)\mathrm{d}\,t + \Phi_1(t)\mathrm{d}\,w_t, \quad t \geqslant 0, \tag{2.1}$$

$$\mathrm{d}\,z_t = h(t, x_t, x_{\tau_1}, \dots, x_{\tau_N}, z)\mathrm{d}\,t + \Phi_2(t, z)\mathrm{d}\,v_t, \tag{2.2}$$

and observable q-dimansional process with discrete time $\eta(t_m)$ (an output signal of the discrete transmission channel) has the form

$$\eta(t_m) = g(t_m, x_{t_m}, x_{\tau_1}, \dots, x_{\tau_N}, z) + \Phi_3(t_m, z)\xi(t_m), \quad m = 0, 1, \dots, \tag{2.3}$$

where $0 \leqslant t_0 < \tau_N < \ldots < \tau_1 < t_m \leqslant t$. It is assumed: 1) w_t and v_t are r_1- and r_2-dimensional standard Wiener processes, respectively, $\xi(t_m)$ is the r_3-dimensional standard white Gaussian sequence; 2) x_0, w_t, v_t, $\xi(t_m)$ are assumed to be statistically independent; 3) $h(\cdot)$, $\Phi_2(\cdot)$ and $g(\cdot)$, $\Phi_3(\cdot)$ are nonanticipating functionals of the realizations $z = z_0^t = \{z_\sigma;\, 0 \leqslant \sigma \leqslant t\}$ and $z = z_0^{t_m}$, of observable process z_t, respectively; 4) coefficients of equations (2.1) and (2.2) are satisfied conditions (Kallianpur, 1980; Liptser and Shiryayev, 1977; 1978), providing an existence of solutions, and $g(\cdot)$ is continuous for all arguments; 5) $Q(\cdot) = \Phi_1(\cdot)\Phi_1^T(\cdot) > 0$, $R(\cdot) = \Phi_2(\cdot)\Phi_2^T(\cdot) > 0$, $V(\cdot) = \Phi_3(\cdot)\Phi_3^T(\cdot) > 0$; 6) the initial density function $p_0(x_0) = = \partial\mathcal{P}\{x_0 \leqslant x\}/\partial x$ is given.

The following problem is stated: for a sequence of moments $t < s_1 < \ldots < s_L$ is to be found relations defining time evolution of joint information amount $I_s^t[x_t, x_{s_1}, \ldots, x_{s_L}; z_0^t, \eta_0^m]$ about the current values x_t and the future values $x_{s_1}, \ldots, x_{s_L}$ of the unobservable process which is contained in the realizations set $z_0^t = \{z_\sigma : 0 \leqslant \sigma \leqslant t\}$ and $\eta_0^m = \{\eta(t_0), \eta(t_1), \ldots, \eta(t_m); t_m \leqslant t\}$ of the observable processes. In this case $s_l = \text{const}$, $l = \overline{1; L}$, i.e., the extrapolation is inverse (Dyomin *et al.*, 1997; Dyomin *et al.*, 2000).

The abstract variant of the formula for Shannon joint information which is contained in the realizations $x = x_0^t$ and $y = y_0^t$, (Dobrushin, 1963; Kolmogorov, 1963) where μ_x, μ_y, $\mu_{x,y}$ are measures agreeable to the processes x_t, y_t, $\{x_t; y_t\}$ (see (Duncan, 1971) and (Liptser and Shiryayev, 1978, chap. 16])

$$I_t[x; y] = M\left\{\ln \frac{\mathrm{d}\,\mu_{x,y}}{\mathrm{d}\,[\mu_x\mu_y]}(x, y)\right\}, \tag{2.4}$$

can't be used in the stated problem. Thus, the solution of the stated problem can be realized by the presentation of the information amount through probabilities distribution densites with the use of Ito formula (Kallianpur, 1980; Liptser and Shiryayev, 1977; 1978) and Ito-Ventzel formula (Rozovsky, 1973; Ocone and Pardoux, 1989), analogously (Liptser, 1974; Dyomin and Korotkevich, 1983; Dyomin and Korotkevich, 1987).

If similar to (Dyomin *et al.*, 1997; Dyomin *et al.*, 2001), we introduce extended processes and variables

$$\tilde{x}_\tau^N = \begin{bmatrix} x_{\tau_1} \\ \vdots \\ x_{\tau_N} \end{bmatrix}, \quad \tilde{x}_s^L = \begin{bmatrix} x_{s_1} \\ \vdots \\ x_{s_L} \end{bmatrix}, \quad \tilde{x}_{t,\tau,s}^{N+L+1} = \begin{bmatrix} x_t \\ \tilde{x}_\tau^N \\ \tilde{x}_s^L \end{bmatrix} = \begin{bmatrix} \tilde{x}_{t,\tau}^{N+1} \\ \tilde{x}_s^L \end{bmatrix} = \begin{bmatrix} \tilde{x}_\tau^N \\ \tilde{x}_{t,s}^{L+1} \end{bmatrix},$$

$$\tilde{x}_N = \begin{bmatrix} x_1 \\ \vdots \\ x_N \end{bmatrix}, \quad \tilde{x}^L = \begin{bmatrix} x^1 \\ \vdots \\ x^L \end{bmatrix}, \quad \tilde{x}_{N+L+1} = \begin{bmatrix} x \\ \tilde{x}_N \\ \tilde{x}^L \end{bmatrix} = \begin{bmatrix} \tilde{x}_{N+1} \\ \tilde{x}^L \end{bmatrix} = \begin{bmatrix} \tilde{x}_N \\ \tilde{x}^{L+1} \end{bmatrix}, \tag{2.5}$$

then in the assumption of the probability densities existence ($\tilde{\tau}_N = [\tau_1, \ldots, \tau_N]$, $\tilde{s}_L = [s_1, \ldots, s_L]$)

$$p_s^t(x; \tilde{x}^L) = \partial^{L+1}\mathcal{P}\left\{x_t \leqslant x; \tilde{x}_s^L \leqslant \tilde{x}^L | z_0^t, \eta_0^m\right\}/\partial x \partial \tilde{x}^L, \tag{2.6}$$

$$p(t,x;\tilde{s}_L,\tilde{x}^L)=\partial^{L+1}\mathcal{P}\left\{x_t\leqslant x;\tilde{x}_s^L\leqslant\tilde{x}^L\right\}/\partial x\partial\tilde{x}^L \tag{2.7}$$

the formula takes place

$$I_s^t[x_t,\tilde{x}_s^L;z_0^t,\eta_0^m]=M\left\{\ln\left[p_s^t(x;\tilde{x}_s^L)/p(t,x;\tilde{s}_L,\tilde{x}_s^L)\right]\right\}. \tag{2.8}$$

REMARK 1. Similar to (Liptser, 1974; Dyomin and Korotkevich, 1983; Dyomin and Korotkevich, 1987) it is assumed: 1^0) application conditions of Ito formula and Ito–Ventzel formula are satisfied; 2^0) for stochastic integrals $J_t=\int_0^t\Psi(\tau,\omega)d\chi_\tau$ with respect to Wiener processes χ_τ the condition $M\{\int_0^t\Psi^2(\tau,\omega)d\tau\}<\infty$, providing the property $M\left\{\int_0^t\Psi(\tau,\omega)d\chi_\tau\right\}=0$ (Kallianpur, 1980; Liptser and Shiryayev, 1977; 1978) is satisfied; 3^0) scalar finite functions $\varphi(\tau,y,\cdot)$, $\varphi_1(\tau,y,\cdot)$, $\varphi_2(\tau,y,\cdot)$ and their derivatives up to second-order, and vector-function $f(\tau,y)$, are assumed so that operators

$$L_{\tau,y}[\varphi(\tau,y,\cdot)]=-\sum_{i=1}^n\frac{\partial[f_i(\tau,y)\varphi(\tau,y,\cdot)]}{\partial y_i}+\frac{1}{2}\sum_{i,j=1}^n\frac{\partial^2[Q_{ij}(\tau)\varphi(\tau,y,\cdot)]}{\partial y_i\partial y_j}, \tag{2.9}$$

$$L^*_{\tau,y}[\varphi(\tau,y,\cdot)]=\sum_{i=1}^n f_i(\tau,y)\frac{\partial\varphi(\tau,y,\cdot)}{\partial y_i}+\frac{1}{2}\sum_{i,j=1}^n Q_{ij}(\tau)\frac{\partial^2\varphi(\tau,y,\cdot)}{\partial y_i\partial y_j}, \tag{2.10}$$

$$\mathcal{L}_{\tau,y}\left[\varphi_1(\tau,y,\cdot);\ \varphi_2(\tau,y,\cdot)\right]=\frac{\varphi_1(\tau,y,\cdot)}{\varphi_2(\tau,y,\cdot)}L_{\tau,y}[\varphi_2(\tau,y,\cdot)]-\varphi_2(\tau,y,\cdot)L^*_{\tau,y}\left[\frac{\varphi_1(\tau,y,\cdot)}{\varphi_2(\tau,y,\cdot)}\right] \tag{2.11}$$

are nonsingular. In accordance with (2.1), $L[\varphi]$ and $L^*[\varphi]$ are the direct and inverse Kolmogorov operators, corresponding to n-dimensional Markovian diffusion process (Kallianpur, 1980; Liptser and Shiryayev, 1977; 1978), and $\mathcal{L}[\varphi_1;\varphi_2]$ as superposition of $L[\cdot]$ and $L^*[\cdot]$ takes part in presentation of the solution of generalized extrapolation problem (Dyomin *et al.*, 1997; Dyomin *et al.*, 2000).

REMARK 2. The models of the processes z_t and $\eta(t_m)$ of form (2.2), (2.3) are adequate to the observations with fixed memory if $\tau_k=\text{const}$, and observations with sliding memory if $\tau_k=t-t_k^*$ in (2.2) and $\tau_k=t_m-t_k^*$ in (2.3), where $t_k^*=\text{const}$, $k=\overline{1;N}$ (Dyomin *et al.*, 1997; Dyomin *et al.*, 2000). The present paper consideres the case of the fixed memory. The dependence $h(\cdot)$ and $g(\cdot)$ of z means that observation channels possess silent feedback relatively the process z_t (Ihara, 1990; Liptser, 1974; Liptser and Shiryayev, 1977; 1978). The absence of feedback, when $h(\cdot)$ and $g(\cdot)$ do not depend on z, is a particular case.

3. The General Relations

The solution of the stated problem is realized by the use of the posterior density

$$p_s^t(x;\tilde{x}_N;\tilde{x}^L)=\partial^{N+L+1}\mathcal{P}\left\{x_t\leqslant x,\ \tilde{x}_\tau^N\leqslant\tilde{x}_N,\ \tilde{x}_s^L\leqslant\tilde{x}^L|z_0^t,\eta_0^m\right\}/\partial x\partial\tilde{x}_N\partial\tilde{x}^L. \tag{3.1}$$

PROPOSITION 1. The density (3.1) on the time intervals $t_m \leqslant t < t_{m+1}$ is defined by the equation

$$\mathrm{d}_t p_s^t(x;\tilde{x}_N;\tilde{x}^L) = \mathcal{L}_{t,x}\left[p_s^t(x;\tilde{x}_N;\tilde{x}^L);p_t(x;\tilde{x}_N)\right]\mathrm{d}\,t$$
$$+p_s^t(x;\tilde{x}_N;\tilde{x}^L)\left[h(t,x,\tilde{x}_N,z)-\overline{h(t,z)}\right]^T R^{-1}(t,z)\left[\mathrm{d}\,z_t-\overline{h(t,z)}\mathrm{d}\,t\right], \quad (3.2)$$

subject to the initial condition

$$p_s^{t_m}(x;\tilde{x}_N;\tilde{x}^L) = \left[C(x;\tilde{x}_N;\eta(t_m),z)/C(\eta(t_m),z)\right]p_s^{t_m-0}(x;\tilde{x}_N;\tilde{x}^L), \quad (3.3)$$

where

$$p_t(x;\tilde{x}_N) = \partial^{N+1}\mathcal{P}\left\{x_t \leqslant x,\ \tilde{x}_\tau^N \leqslant \tilde{x}_N|z_0^t,\eta_0^m\right\}/\partial x\partial\tilde{x}_N, \quad (3.4)$$

$$\overline{h(t,z)} = M\left\{h(t,x_t,\tilde{x}_\tau^N,z)|z_0^t,\eta_0^m\right\}, \quad (3.5)$$

$$C\left(\eta(t_m),z\right) = M\left\{C\left(x_{t_m},\tilde{x}_\tau^N,\eta(t_m),z\right)|z_0^{t_m},\eta_0^{m-1}\right\}, \quad (3.6)$$

$$C\left(x,\tilde{x}_N,\eta(t_m),z\right) = \exp\Big\{-\frac{1}{2}\left[\eta(t_m)-g(t_m,x,\tilde{x}_N,z)\right]^T V^{-1}(t_m,z)$$
$$\times\left[\eta(t_m)-g(t_m,x,\tilde{x}_N,z)\right]\Big\}, \quad (3.7)$$

and $p_s^{t_m-0}(x;\tilde{x}_N;\tilde{x}^L) = \lim p_s^t(\cdot)$ subject to $t \uparrow t_m$.

This proposition is valid, taking into account (2.9)–(2.11), from Corollary 1 in (Dyomin *et al.*, 1997).

Theorem 1. *The information amount* (2.8) *on the time intervals* $t_m \leqslant t < t_{m+1}$ *is determined by equation*

$$\mathrm{d}\,I_s^t[x_t,\tilde{x}_s^L;z_0^t,\eta_0^m]/\mathrm{d}\,t = (1/2)tr\Big[M\Big\{R^{-1}(t,z)\Big[\overline{h(\tilde{\tau}_N,z|x_t,\tilde{x}_s^L)}$$
$$-\overline{h(t,z)}\Big]\Big[\overline{h(\tilde{\tau}_N,z|x_t,\tilde{x}_s^L)}-\overline{h(t,z)}\Big]^T\Big\}\Big]$$
$$-\frac{1}{2}tr\Big[Q(t)M\Big\{\frac{\partial\ln p_s^t(x_t;\tilde{x}_s^L)}{\partial x_t}\Big(\frac{\partial\ln p_s^t(x_t;\tilde{x}_s^L)}{\partial x_t}\Big)^T$$
$$-\frac{\partial\ln p(t,x_t;\tilde{s}_L,\tilde{x}_s^L)}{\partial x_t}\Big(\frac{\partial\ln p(t,x_t;\tilde{s}_L,\tilde{x}_s^L)}{\partial x_t}\Big)^T\Big\}\Big]$$
$$+tr\Big[Q(t)M\Big\{\Big[\frac{\partial\ln p_s^t(x_t;\tilde{x}_s^L)}{\partial x_t}-\frac{\partial\ln p_t(x_t)}{\partial x_t}\Big]\Big(\frac{\partial\ln p_t(x_t)}{\partial x_t}\Big)^T$$
$$-\Big[\frac{\partial\ln p(t,x_t;\tilde{s}_L,\tilde{x}_s^L)}{\partial x_t}-\frac{\partial\ln p(t,x_t)}{\partial x_t}\Big]\Big(\frac{\partial\ln p(t,x_t)}{\partial x_t}\Big)^T\Big\}\Big], \quad (3.8)$$

subject to the initial condition

$$I_s^{t_m}[x_{t_m},\tilde{x}_s^L;z_0^{t_m},\eta_0^m] = I_s^{t_m-0}[x_{t_m},\tilde{x}_s^L;z_0^{t_m},\eta_0^{m-1}] \\ +\Delta I_s^{t_m}\left[x_{t_m},\tilde{x}_s^L;z_0^{t_m},\eta(t_m)\right], \tag{3.9}$$

where

$$p_t(x) = \partial\mathcal{P}\left\{x_t \leqslant x|z_0^t,\ \eta_0^m\right\}/\partial x, \quad p(t,x) = \partial\mathcal{P}\{x_t \leqslant x\}/\partial x, \tag{3.10}$$

$$\overline{h(\tilde{\tau}_N,z|x,\tilde{x}^L)} = M\left\{h(t,x_t,\tilde{x}_\tau^N,z)|x_t = x,\tilde{x}_s^L = \tilde{x}^L,z_0^t,\eta_0^m\right\}, \tag{3.11}$$

$$\Delta I_s^{t_m}\left[x_{t_m},\tilde{x}_s^L;z_0^{t_m},\eta(t_m)\right] = M\left\{\ln\left[C\big(\eta(t_m),z|x_{t_m},\tilde{x}_s^L\big)/C(\eta(t_m),z)\right]\right\}, \tag{3.12}$$

$$C\left(\eta(t_m),z|x,\tilde{x}_s^L\right) = \\ = M\left\{C\left(x_{t_m},\tilde{x}_\tau^N,\eta(t_m),z\right)|x_{t_m} = x,\ \tilde{x}_s^L = \tilde{x}^L;z_0^{t_m},\eta_0^{m-1}\right\}, \tag{3.13}$$

and $I_s^{t_m-0}[x_{t_m},\tilde{x}_s^L;z_0^{t_m},\eta_0^{m-1}] = \lim I_s^t(\cdot)$ *subject to* $t\uparrow t_m$*.*

Proof. Since $p_s^t(x;\tilde{x}_N;\tilde{x}^L) = p_\tau^t(\tilde{x}_N|x,\tilde{x}^L)p_s^t(x;\tilde{x}^L)$, $p_t(x;\tilde{x}_N) = p_\tau^t(\tilde{x}_N|x)p_t(x)$, where $p_\tau^t(\tilde{x}_N|x,\tilde{x}^L) = \partial^N\mathcal{P}\{\tilde{x}_\tau^N \leqslant \tilde{x}_N|x_t = x,\tilde{x}_s^L = \tilde{x}^L,z_0^t,\eta_0^m)/\partial\tilde{x}_N$, $p_\tau^t(\tilde{x}_N|x) = \partial^N\mathcal{P}\{\tilde{x}_\tau^N \leqslant \tilde{x}_N|x_t = x,z_0^t,\eta_0^m)/\partial\tilde{x}_N$, then integrating (3.2) and (3.3) with respect to $\tilde{x}_N$ taking into account (3.11), (3.13) yields that the posterior density (2.6) on the time intervals $t_m \leqslant t < t_{m+1}$ is defined by the equation

$$\mathrm{d}_t p_s^t(x;\tilde{x}^L) = \mathcal{L}_{t,x}\left[p_s^t(x;\tilde{x}^L);p_t(x)\right]\mathrm{d}t \\ +p_s^t(x;\tilde{x}^L)\left[\overline{h(\tau_N,z|x,\tilde{x}^L)} - \overline{h(t,z)}\right]^T R^{-1}(t,z)\mathrm{d}\tilde{z}_t, \tag{3.14}$$

$$\mathrm{d}\tilde{z}_t = \mathrm{d}z_t - \overline{h(t,z)}\mathrm{d}t, \tag{3.15}$$

subject to the initial condition

$$p_s^{t_m}(x;\tilde{x}^L) = \left[C(\eta(t_m),z|x,\tilde{x}^L)/C(\eta(t_m),z)\right]p_s^{t_m-0}(x;\tilde{x}^L). \tag{3.16}$$

Since x_t is Markov process, then $p_\tau^t(\tilde{x}_N|x,\tilde{x}^L) = p_\tau^t(\tilde{x}_N|x)$. The prior density (2.7) is defined by the equation

$$\mathrm{d}_t p(t,x;\tilde{s}_L,\tilde{x}^L) = \mathcal{L}_{t,x}\left[p(t,x;\tilde{s}_L,\tilde{x}^L);p(t,x)\right]\mathrm{d}t, \tag{3.17}$$

which follows from (3.14). Innovation process $\tilde{z}_t$, differential of which has the form (3.15), is such that $\widetilde{Z}_t = (\tilde{z}_t,\mathcal{F}_t^z)$ is Wiener process with $M\{\tilde{z}_t\tilde{z}_t^T|\mathcal{F}_t^z\} = \int_0^t R(\tau,z)d\tau$ (Kallianpur, 1980; Liptser and Shiryayev 1977; 1978). Differentiation according to Ito formula taking into account (2.11), (3.14), (3.17) yields

$$\mathrm{d}_t\ln\left[\frac{p_s^t(x;\tilde{x}^L)}{p(t,x;\tilde{s}_L,\tilde{x}^L)}\right] = \left\{\frac{1}{p_t(x)}L_{t,x}[p_t(x)] - \frac{p_t(x)}{p_s^t(x;\tilde{x}^L)}L_{t,x}^*\left[\frac{p_s^t(x;\tilde{x}^L)}{p_t(x)}\right]\right\}\mathrm{d}t$$

$$-\left\{\frac{1}{p(t,x)}L_{t,x}[p(t,x)]-\frac{p(t,x)}{p(t,x;\tilde{s}_L,\tilde{x}^L)}L^*_{t,x}\left[\frac{p(t,x;\tilde{s}_L,\tilde{x}^L)}{p(t,x)}\right]\right\}\mathrm{d}\,t$$
$$-\frac{1}{2}\left[\overline{h(\tilde{\tau}_N,z|x,\tilde{x}^L)}-\overline{h(t,z)}\right]^T R^{-1}(t,z)\left[\overline{h(\tilde{\tau}_N,z|x,\tilde{x}^L)}-\overline{h(t,z)}\right]\mathrm{d}\,t$$
$$+\left[\overline{h(\tilde{\tau}_N,z|x,\tilde{x}^L)}-\overline{h(t,z)}\right]^T R^{-1}(t,z)\mathrm{d}\,\tilde{z}_t. \tag{3.18}$$

Applying to (3.18) Ito–Ventzel formula for $t_m \leqslant t < t_{m+1}$ and similar to (Liptser, 1974; Dyomin and Korotkevich, 1983) we obtain

$$\ln\left[\frac{p^t_s(x_t;\tilde{x}^L_s)}{p(t,x_t;\tilde{s}_L,\tilde{x}^L)}\right]=\ln[\cdot]|_{t=t_m}$$
$$+\frac{1}{2}\int_{t_m}^{t} tr\left[R^{-1}(\sigma,z)\left[\overline{h(\tilde{\tau}_N,z|x_\sigma,\tilde{x}^L_s)}-\overline{h(\sigma,z)}\right][\cdot]^T\right]\mathrm{d}\,\sigma$$
$$-\frac{1}{2}\int_{t_m}^{t} tr\Bigg[Q(t)\Bigg[\frac{\partial\ln p^\sigma_s(x_\sigma;\tilde{x}^L_s)}{\partial x_\sigma}\Big(\frac{\partial\ln p^\sigma_s(\cdot)}{\partial x_\sigma}\Big)^T$$
$$-\frac{\partial\ln p(\sigma,x_\sigma;\tilde{s}_L,\tilde{x}^L_s)}{\partial x_\sigma}\Big(\frac{\partial\ln p(\cdot)}{\partial x_\sigma}\Big)^T\Bigg]\Bigg]\mathrm{d}\,\sigma$$
$$+\frac{1}{2}\int_{t_m}^{t} tr\Bigg[Q(t)\Bigg[\Big(\frac{\partial\ln p^\sigma_s(x_\sigma;\tilde{x}^L_s)}{\partial x_\sigma}-\frac{\partial\ln p_\sigma(x_\sigma)}{\partial x_\sigma}\Big)\Big(\frac{\partial\ln p_\sigma(x_\sigma)}{\partial x_\sigma}\Big)^T$$
$$-\Big(\frac{\partial\ln p(\sigma,x_\sigma;\tilde{s}_L,\tilde{x}^L_s)}{\partial x_\sigma}-\frac{\partial\ln p(\sigma,x_\sigma)}{\partial x_\sigma}\Big)\Big(\frac{\partial\ln p(\sigma,x_\sigma)}{\partial x_\sigma}\Big)^T\Bigg]\Bigg]\mathrm{d}\,\sigma$$
$$+\int_{t_m}^{t} tr\left[Q(t)\left[\frac{1}{p_\sigma(x_\sigma)}\frac{\partial^2 p_\sigma(x_\sigma)}{\partial x^2_\sigma}-\frac{1}{p(\sigma,x_\sigma)}\frac{\partial^2 p(\sigma,x_\sigma)}{\partial x^2_\sigma}\right]\right]\mathrm{d}\,\sigma$$
$$+\int_{t_m}^{t} tr\Bigg[R^{-1}(\sigma,z)\left[\overline{h(\tilde{\tau}_N,z|x_\sigma,\tilde{x}^L_s)}-\overline{h(\sigma,z)}\right]$$
$$\times\left[h(\sigma,x_\sigma,\tilde{x}^N_\tau,z)-\overline{h(\tilde{\tau}_N,z|x_\sigma,\tilde{x}^L_s)}\right]^T\Bigg]\mathrm{d}\,\sigma$$
$$+\int_{t_m}^{t}\frac{\partial}{\partial x_\sigma}\ln\frac{p^\sigma_s(x_\sigma;\tilde{x}^L_s)}{p(\sigma,x_\sigma;\tilde{s}_L,\tilde{x}^L_s)}\Phi_1(\sigma)\mathrm{d}\,\omega_\sigma$$
$$+\int_{t_m}^{t}\left[\overline{h(\tilde{\tau}_N,z|x_\sigma,\tilde{x}^L_s)}-\overline{h(\sigma,z)}\right]^T R^{-1}(\sigma,z)\Phi_2(\sigma,z)\mathrm{d}\,v_\sigma. \tag{3.19}$$

Similar to (Liptser, 1974) and as well as (П.13) in (Dyomin and Korotkevich, 1983), we have

$$M\left\{\frac{1}{p_\sigma(x_\sigma)}\frac{\partial^2 p_\sigma(x_\sigma)}{\partial x_\sigma^2}-\frac{1}{p(\sigma,x_\sigma)}\frac{\partial^2 p(\sigma,x_\sigma)}{\partial x_\sigma^2}\right\}=M\left\{M\left\{\cdot|z_0^\sigma,\eta_0^m\right\}\right\}$$
$$=M\left\{\int\frac{\partial^2 p_\sigma(x)}{\partial x^2}\mathrm{d}\,x\right\}-\int\frac{\partial^2 p(\sigma,x)}{\partial x^2}\mathrm{d}\,x=O. \tag{3.20}$$

Since, in accordance with (3.5),(3.11), we have $M\{h(\sigma,x_\sigma,\tilde{x}_\tau^N,z)\}=M\{M\{M\{h(\cdot)|x_\sigma= x,\tilde{x}_s^L=\tilde{x}^L,\ z_0^\sigma,\eta_0^m\}|z_0^\sigma,\eta_0^m\}\}=M\{M\{\overline{h(\tilde{\tau}_N,z|x_\sigma,\tilde{x}_s^L)}|z_0^\sigma,\eta_0^m\}\}=M\{\overline{h(\sigma,z)}\}$ then

$$M\left\{R^{-1}(\sigma,z)\left[\overline{h(\tilde{\tau}_N,z|x_\sigma,\tilde{x}_s^L)}-\overline{h(\sigma,z)}\right]\left[h(\sigma,x_\sigma,\tilde{x}_\tau^N,z)-\overline{h(\tilde{\tau}_N,z|x_\sigma,\tilde{x}_s^L)}\right]^T\right\}$$
$$=M\left\{R^{-1}(\sigma,z)M\left\{[\cdot][\cdot]^T|z_0^\sigma,\eta_0^m\right\}\right\}=O. \tag{3.21}$$

The calculation of expectation of the left and right parts of (3.19) taking into account (3.20), (3.21), 2^0) Remark 1 followed by differentiating with respect to t gives (3.8), and substitution of (3.3) in (2.8) gives (3.9).

COROLLARY 1. The information amount (see (3.10))

$$I_t[x_t;z_0^t,\eta_0^m]=M\left\{\ln\left[p_t(x_t)/p(t,x_t)\right]\right\} \tag{3.22}$$

about the current values of the process x_t on the time intervals $t_m\leqslant t<t_{m+1}$ is defined by the equation

$$\begin{aligned}&\mathrm{d}\,I_t[x_t;z_0^t,\eta_0^m]/\mathrm{d}\,t\\&=(1/2)tr\left[M\left\{R^{-1}(t,z)\left[\overline{h(\tilde{\tau}_N,z|x_t)}-\overline{h(t,z)}\right]\left[\overline{h(\tilde{\tau}_N,z|x_t)}-\overline{h(t,z)}\right]^T\right\}\right]\\&\quad-\frac{1}{2}tr\left[Q(t)M\left\{\frac{\partial\ln p_t(x_t)}{\partial x_t}\left(\frac{\partial\ln p_t(x_t)}{\partial x_t}\right)^T-\frac{\partial\ln p(t,x_t)}{\partial x_t}\left(\frac{\partial\ln p(t,x_t)}{\partial x_t}\right)^T\right\}\right]\end{aligned} \tag{3.23}$$

subject to the initial condition

$$I_{t_m}[x_{t_m};z_0^{t_m},\eta_0^m]=I_{t_m-0}[x_{t_m};z_0^{t_m},\eta_0^{m-1}]+\Delta I_{t_m}[x_{t_m};z_0^{t_m},\eta(t_m)], \tag{3.24}$$

$$\Delta I_{t_m}[x_{t_m};z_0^{t_m},\eta(t_m)]=M\left\{\ln\left[C(\eta(t_m),z|x_{t_m})/C(\eta(t_m),z)\right]\right\}, \tag{3.25}$$

$$\overline{h(\tilde{\tau}_N,z|x)}=M\left\{h(t,x_t,\tilde{x}_\tau^N,z)|x_t=x,z_0^t,\eta_0^m\right\}, \tag{3.26}$$

$$C(\eta(t_m),z|x)=M\left\{C(x_{t_m},\tilde{x}_\tau^N,\eta(t_m),z)|x_{t_m}=x,z_0^{t_m},\eta_0^{m-1}\right\}, \tag{3.27}$$

and $I_{t_m-0}[x_{t_m};z_0^{t_m},\eta_0^{m-1}]=\lim I_t[\cdot]$ subject to $t\uparrow t_m$.

The formulated result is obtained as a limitary case from Theorem 1 subject to $s_l\downarrow t$ in (3.8) and $s_l\downarrow t_m$ in (3.12), $l=\overline{1;L}$, and defines of information amount in filtering problem. It follows from equations (3.14), (3.16) and (3.17) taking into account (2.9)–(2.11), that $p_t(x)$ on the time intervals $t_m \leqslant t < t_{m+1}$ is defined by the equation

$$\mathrm{d}\,_t p_t(x) = L_{t,x}[p_t(x)]\mathrm{d}\,t + p_t(x)[\overline{h(\tilde{\tau}_N, z|x)} - \overline{h(t,z)}]^T R^{-1}(t,z)\mathrm{d}\,\tilde{z}_t, \tag{3.28}$$

subject to the initial condition

$$p_{t_m}(x) = [C\,(\eta(t_m), z|x)\,/C(\eta(t_m), z)]\,p_{t_m-0}(x), \tag{3.29}$$

and $p(t,x)$ is defined by the equation $\mathrm{d}\,_t p(t,x) = L_{t,x}[p(t,x)]\mathrm{d}\,t$. Hence (3.23) and (3.24) can be obtained immediately by analogy with (3.8) and (3.9). Similarly the proof of Theorems 1, 3 in (Dyomin and Korotkevich, 1987) for the case $N = 1$ was made.

Along with (3.22) the information amount

$$I_s^t[\tilde{x}_s^L; z_0^t, \eta_0^m] = M\left\{\ln[p_s^t(\tilde{x}_s^L)/p(\tilde{s}_L, \tilde{x}_s^L)]\right\} \tag{3.30}$$

about the future values $\tilde{x}_s^L$ of the process x_t is of interest, i.e., information amount in generalized extrapolation problem, where

$$\begin{aligned} p_s^t(\tilde{x}^L) &= \partial^L \mathcal{P}\left\{\tilde{x}_s^L \leqslant \tilde{x}^L | z_0^t, \eta_0^m\right\}/\partial \tilde{x}^L, \\ p(\tilde{s}_L, \tilde{x}^L) &= \partial^L \mathcal{P}\left\{\tilde{x}_s^L \leqslant \tilde{x}^L\right\}/\partial \tilde{x}^L. \end{aligned} \tag{3.31}$$

Theorem 2. *The information amount* (3.30) *on the time intervals* $t_m \leqslant t < t_{m+1}$ *is defined by the equation*

$$\begin{aligned} \mathrm{d}\,I_s^t[\tilde{x}_s^L; z_0^t, \eta_0^m]/\mathrm{d}\,t \;=\; (1/2)tr\Big[M\Big\{R^{-1}(t,z)\Big[\overline{h(\tilde{\tau}_N, t, z|\tilde{x}_s^L)} - \overline{h(t,z)}\Big] \\ \times \Big[\overline{h(\tilde{\tau}_N, t, z|\tilde{x}_s^L)} - \overline{h(t,z)}\Big]^T\Big\}\Big], \end{aligned} \tag{3.32}$$

subject to the initial condition

$$I_s^{t_m}[\tilde{x}_s^L; z_0^{t_m}, \eta_0^m] = I_s^{t_m-0}[\tilde{x}_s^L; z_0^{t_m}, \eta_0^{m-1}] + \Delta I_s^{t_m}[\tilde{x}_s^L; z_0^{t_m}, \eta(t_m)], \tag{3.33}$$

$$\Delta I_s^{t_m}[\tilde{x}_s^L; z_0^{t_m}, \eta(t_m)] = M\left\{\ln\left[C(\eta(t_m), z|\tilde{x}_s^L)/C(\eta(t_m), z)\right]\right\}, \tag{3.34}$$

$$\overline{h(\tilde{\tau}_N, t, z|\tilde{x}^L)} = M\left\{h(t, x_t, \tilde{x}_\tau^N, z)|\tilde{x}_s^L = \tilde{x}^L, z_0^t, \eta_0^m\right\}, \tag{3.35}$$

$$C(\eta(t_m), z|\tilde{x}^L) = M\left\{C(x_{t_m}, \tilde{x}_\tau^N, \eta(t_m), z)|\tilde{x}_s^L = \tilde{x}^L, z_0^{t_m}, \eta_0^{m-1}\right\}, \tag{3.36}$$

and $I_s^{t_m-0}[\tilde{x}_s^L; z_0^{t_m}, \eta_0^{m-1}] = \lim I_s^t[\cdot]$ *subject to* $t \uparrow t_m$.

Proof. Since $p_s^t(x;\tilde{x}_N;\tilde{x}^L) = p_\tau^t(x;\tilde{x}_N|\tilde{x}^L)p_s^t(\tilde{x}^L)$, where $p_\tau^t(x;\tilde{x}_N|\tilde{x}^L) = \partial^{N+1}\mathcal{P}\{x_t \leqslant x;\, \tilde{x}_\tau^N \leqslant \tilde{x}_N|\tilde{x}_s^L = \tilde{x}^L, z_0^t, \eta_0^m\}/\partial x\partial\tilde{x}_N$, then integration (3.2) and (3.3) with respect to $\{x;\tilde{x}_N\}$ taking into account (2.9)–(2.11), (3.35) and (3.36), yields that $p_s^t(\tilde{x}^L)$ on the time intervals $t_m \leqslant t < t_{m+1}$ is defined by the equation

$$\mathrm{d}_t p_s^t(\tilde{x}^L) = p_s^t(\tilde{x}^L)\left[\overline{h(\tilde{\tau}_N,t,z|\tilde{x}^L)} - \overline{h(t,z)}\right] R^{-1}(t,z)\mathrm{d}\,\tilde{z}_t \tag{3.37}$$

subject to the initial condition

$$p_s^{t_m}(\tilde{x}^L) = \left[C(\eta(t_m),z|\tilde{x}^L)/C(\eta(t_m),z)\right] p_s^{t_m-0}(\tilde{x}^L). \tag{3.38}$$

Since the prior density $p(t,x;\tilde{\tau}_N,\tilde{x}_N;\tilde{s}_L,\tilde{x}^L)$ in accordance with (3.2) is defined by the equation

$$\mathrm{d}_t p\big(t,x;\tilde{\tau}_N,\tilde{x}_N;\tilde{s}_L,\tilde{x}^L\big) = \mathcal{L}_{t,x}\big[p(t,x;\tilde{\tau}_N,\tilde{x}_N;\tilde{s}_L,\tilde{x}^L);p(t,x;\tilde{\tau}_N,\tilde{x}_N)\big]\mathrm{d}t, \tag{3.39}$$

then integrating (3.39) with respect to $\{x;\tilde{x}_N\}$ taking into account (2.9)–(2.11) yields $\mathrm{d}_t p(\tilde{s}_L;\tilde{x}^L) = 0$. The further inference of (3.32) and (3.33) is similar to that of (3.8) and (3.9).

4. Conditionally-Gaussian case

The effective determination of the filtering and extrapolation estimates was obtained in (Abakumova *et al.*, 1995b; Dyomin *et al.*, 1997; Dyomin *et al.*, 2000) under the conditions (see (2.1)–(2.3), (2.5))

$$\begin{aligned} &f(\cdot) = f(t) + F(t)x_t, \quad p_0(x) = \mathcal{N}\{x;\mu_0,\Gamma_0\},\\ &h(\cdot) = h(t,z) + H_{0,N}(t,z)\tilde{x}_{t,\tau}^{N+1}, \quad g(\cdot) = g(t_m,z) + G_{0,N}(t_m,z)\tilde{x}_{t_m,\tau}^{N+1}, \end{aligned} \tag{4.1}$$

$$\begin{aligned} &H_{0,N}(\cdot) = \Big[H_0(t,z) \vdots H_1(t,z) \vdots \cdots \vdots H_N(t,z)\Big] = \Big[H_0(t,z) \vdots H_{1,N}(t,z)\Big],\\ &G_{0,N}(\cdot) = \Big[G_0(t_m,z) \vdots G_1(t_m,z) \vdots \cdots \vdots G_N(t_m,z)\Big] = \Big[G_0(t_m,z) \vdots G_{1,N}(t_m,z)\Big], \end{aligned} \tag{4.2}$$

when the posterior densities for the process $\tilde{x}_{t,\tau,s}^{N+L+1}$ are Gaussian (see (4) in (Abakumova *et al.*, 1995b), (2.15), (2.34) in (Dyomin *et al.*, 1997) and (3.3), (3.4) in (Dyomin *et al.*, 2000). Hence, if

$$\begin{aligned} &\mu(t) = M\{x_t|z_0^t,\eta_0^m\},\\ &\tilde{\mu}_N(\tilde{\tau}_N,t) = M\{\tilde{x}_\tau^N|z_0^t,\eta_0^m\}, \quad \tilde{\mu}^L(t,\tilde{s}_L) = M\{\tilde{x}_s^L|z_0^t,\eta_0^m\},\\ &\Gamma(t) = M\{[x_t-\mu(t)][\cdot]^T|z_0^t\}, \quad \widetilde{\Gamma}_N(\tilde{\tau}_N,t) = M\{[\tilde{x}_\tau^N - \tilde{\mu}_N(\tilde{\tau}_N,t)][\cdot]^T|z_0^t,\eta_0^m\},\\ &\widetilde{\Gamma}^L(t,\tilde{s}_L) = M\{[\tilde{x}_s^L - \tilde{\mu}^L(\tilde{s}_L,t)][\cdot]^T|z_0^t,\eta_0^m\},\\ &\widetilde{\Gamma}_{0N}(\tilde{\tau}_N,t) = M\{[x_t-\mu(t)][\tilde{x}_\tau^N-\tilde{\mu}_N(\tilde{\tau}_N,t)]^T|z_0^t,\eta_0^m\},\\ &\widetilde{\Gamma}_{0,N+1}^L(t,\tilde{s}_L) = M\{[x_t-\mu(t)][\tilde{x}_s^L-\tilde{\mu}^L(\tilde{s}_L,t)]^T|z_0^t,\eta_0^m\},\\ &\widetilde{\Gamma}_{N,N+1}^L(\tilde{\tau}_N,t,\tilde{s}_L) = M\{[\tilde{x}_\tau^N-\tilde{\mu}_N(\tilde{\tau}_N,t)][\tilde{x}_s^L-\tilde{\mu}^L(\tilde{s}_L,t)]^T|z_0^t,\eta_0^m\}, \end{aligned} \tag{4.3}$$

then under satistaction of conditions (4.1)

$$p_s^t(x;\tilde{x}_N;\tilde{x}^L) = \mathcal{N}\left\{\tilde{x}_{N+L+1};\tilde{\mu}_{N+L+1}(\tilde{\tau}_N,t,\tilde{s}_L),\widetilde{\Gamma}_{N+L+1}(\tilde{\tau}_N,t,\tilde{s}_L)\right\}$$
$$=\mathcal{N}\left\{\begin{bmatrix} x \\ \tilde{x}_N \\ \tilde{x}_L \end{bmatrix};\begin{bmatrix} \mu(t) \\ \tilde{\mu}_N(\tilde{\tau}_N,t) \\ \tilde{\mu}^L(t,\tilde{s}_L) \end{bmatrix},\begin{bmatrix} \Gamma(t) & \widetilde{\Gamma}_{0N}(\tilde{\tau}_N,t) & \widetilde{\Gamma}^L_{0,N+1}(t,\tilde{s}_L) \\ \widetilde{\Gamma}^T_{0N}(\cdot) & \widetilde{\Gamma}_N(\tilde{\tau}_N,t) & \widetilde{\Gamma}^L_{N,N+1}(\tilde{\tau}_N,t,\tilde{s}_L) \\ (\widetilde{\Gamma}^L_{0,N+1}(\cdot))^T & (\widetilde{\Gamma}^L_{N,N+1}(\cdot))^T & \widetilde{\Gamma}^L(t,\tilde{s}_L) \end{bmatrix}\right\}. \quad (4.4)$$

PROPOSITION 2. Subject to (4.1) for posterior density $p_s^t(x;\tilde{x}_N;\tilde{x}^L)$ of the process $\tilde{x}^{N+L+1}_{t,\tau,s}$ (see (2.5)) the condition (4.4) takes place and block parameters of this distribution is defined by the differential-reccurence equations of Theorems 1, 2 in (Abakumova *et al.*, 1995b), Theorem 3 and Colollary 2 in (Dyomin *et al.*, 1997). Gaussianity property takes place also for the posterior densities $p_t(x)$, $p_t(x;\tilde{x}_N)$, $p_s^t(\tilde{x}^L)$, $p_s^t(x;\tilde{x}^L)$ composing x_t, $\{x_t;\tilde{x}^N_\tau\}$, $\tilde{x}^L_s$, $\{x_t;\tilde{x}^L_s\}$, of the process $\tilde{x}^{N+L+1}_{t,\tau,s}$, the parameters of which are obtained obviously from (4.4), taking into account (2.5).

REMARK 3. Since the process, defined by the equation $\mathrm{d}\,x_t = [f(t)+F(t)x_t]\mathrm{d}\,t + \Phi_1(t)\mathrm{d}\,\omega_t$, is Gaussian (Liptser and Shiryayev, 1977; 1978; Meditch, 1969), then for the prior density $p(t,x;\tilde{\tau}_N,\tilde{x}_N;\tilde{s}_L,\tilde{x}^L)$ subject to (4.1) the Gaussianity property of the form (4.4) with replacement $\mu(t)$ by $a(t)$, $\tilde{\mu}_N(\tilde{\tau}_N,t)$ by $\widetilde{a}_N(\tilde{\tau}_N,t)$, $\tilde{\mu}^L(t,\tilde{s}_L)$ by $\widetilde{a}^L(t,\tilde{s}_L)$ and the letter Γ by the letter D takes place. Parameters of this density are obviously defined (Meditch, 1969). The prior densities $p(t,x)$, $p(t,x;\tilde{\tau}_N,\tilde{x}_N)$, $p(\tilde{s}_L,\tilde{x}^L)$, $p(t,x;\tilde{s}_L,\tilde{x}^L)$ are Gaussian as well.

In this paragraph the results of the previous paragraph are concretized in case of condition (4.1) fulfillment assuming that all the matrices of the second central moments are reversible.

Theorem 3. *The information amount* (2.8) *on the time intervals* $t_m \leqslant t < t_{m+1}$ *is defined by the equation*

$$\begin{aligned}&\mathrm{d}\,I_s^t[x_t,\tilde{x}^L_s;z_0^t,\eta_0^m]/\mathrm{d}\,t \\ &\quad = (1/2)tr\left[M\left\{R^{-1}(t,z)\widetilde{H}_{L+1}(t,z)(\widetilde{\Gamma}^{L+1}(\tilde{s}_L,t))^{-1}\widetilde{H}^T_{L+1}(t,z)\right\}\right] - \\ &\qquad -(1/2)tr\left[Q(t)\left[M\left\{\Gamma^{-1}(t|\tilde{s}_L)\right\} - D^{-1}(t|\tilde{s}_L)\right]\right], \end{aligned} \quad (4.5)$$

subject to the initial condition (3.9) *where (see* (4.3), (4.4) *and Remark* 3*)*

$$\widetilde{\Gamma}^{L+1}(t,\tilde{s}_L) = \begin{bmatrix} \Gamma(t) & \widetilde{\Gamma}^L_{0,N+1}(t,\tilde{s}_L) \\ (\widetilde{\Gamma}^L_{0,N+1}(\cdot))^T & \widetilde{\Gamma}^L(t,\tilde{s}_L) \end{bmatrix}, \quad (4.6)$$

$$\Gamma(t|\tilde{s}_L) = \Gamma(t) - \widetilde{\Gamma}^L_{0,N+1}(t,\tilde{s}_L)(\widetilde{\Gamma}^L(t,\tilde{s}_L))^{-1}(\widetilde{\Gamma}^L_{0,N+1}(t,\tilde{s}_L))^T, \quad (4.7)$$

$$D(t|\tilde{s}_L) = D(t) - \widetilde{D}^L_{0,N+1}(t,\tilde{s}_L)(\widetilde{D}^L(t,\tilde{s}_L))^{-1}(\widetilde{D}^L_{0,N+1}(t,\tilde{s}_L))^T, \quad (4.8)$$

$$\widetilde{H}_{L+1}(t,z) = [\widetilde{H}_0(t,z)\vdots\widetilde{H}_L(t,z)] = [\widetilde{H}_0(t,z)\vdots\widetilde{H}_{N+1}(t,z)\vdots\cdots\vdots\widetilde{H}_{N+L}(t,z)], \quad (4.9)$$

$$\widetilde{H}_0(t,z) = H_0(t,z)\Gamma(t) + H_{1,N}(t,z)\widetilde{\Gamma}^T_{0N}(\tilde{\tau}_N,t), \tag{4.10}$$

$$\widetilde{H}_{N+l}(t,z) = H_0(t,z)\Gamma^l_{0,N+1}(t,s_l) + H_{1,N}(t,z)\widetilde{\Gamma}^l_{N,N+1}(\tilde{\tau}_N,t,s_l), \quad l=\overline{1;L}, \tag{4.11}$$

$$\Delta I^{t_m}_s[\cdot] = (1/2)M\left\{\ln\left[|\widetilde{\Gamma}^{L+1}(t_m-0,\tilde{s}_L)|/|\widetilde{\Gamma}^{L+1}(t_m,\tilde{s}_L)|\right]\right\}, \tag{4.12}$$

$\widetilde{\Gamma}^{L+1}(t_m-0,\tilde{s}_L) = \lim\widetilde{\Gamma}^{L+1}(t,\tilde{s}_L)$ *subject to* $t\uparrow t_m$, $H_{1,N}(t,z)$ *is described in* (4.2), $\Gamma^l_{0,N+1}(t,s_l)$ *is the* l*-th element of the matrix* $\widetilde{\Gamma}^L_{0,N+1}(t,\tilde{s}_L)$, *and* $\widetilde{\Gamma}^l_{N,N+1}(\tilde{\tau}_N,t,s_l)$ *is* l*-th matrix column of the matrix* $\widetilde{\Gamma}^L_{N,N+1}(\tilde{\tau}_N,t,\tilde{s}_L)$.

Proof. By the property of Gaussian densities (Liptser and Shiryayev, 1977; 1978; Meditch, 1969) for $p^t_{\tau|t,s}(\tilde{x}_N|x,\tilde{x}^L) = \partial^N\mathcal{P}\{\tilde{x}^N_\tau \leqslant \tilde{x}_N|x_t = x, \tilde{x}^L_s = \tilde{x}^L, z^t_0, \eta^m_0]\}/\partial\tilde{x}_N$ in accordance with (2.5) and (4.4), we have

$$\begin{aligned}
&p^t_{\tau|t,s}(\cdot) = \mathcal{N}\left\{\tilde{x}_N;\tilde{\mu}_N(\tilde{\tau}_N|t,\tilde{s}_L),\widetilde{\Gamma}_N(\tilde{\tau}_N|t,\tilde{s}_L)\right\},\\
&\tilde{\mu}_N(\tilde{\tau}_N|t,\tilde{s}_L) = \tilde{\mu}_N(\tilde{\tau}_N,t)\\
&\quad+\widetilde{\Gamma}^{L+1}_N(\tilde{\tau}_N,t,\tilde{s}_L)(\widetilde{\Gamma}^{L+1}(t,\tilde{s}_L))^{-1}\left[\tilde{x}^{L+1} - \tilde{\mu}^{L+1}(t,\tilde{s}_L)\right], \qquad (4.13)\\
&\widetilde{\Gamma}_N(\tilde{\tau}_N|t,\tilde{s}_L) = \widetilde{\Gamma}_N(\tilde{\tau}_N,t) - \widetilde{\Gamma}^{L+1}_N(\tilde{\tau}_N,t,\tilde{s}_L)(\widetilde{\Gamma}^{L+1}(t,\tilde{s}_L))^{-1}(\widetilde{\Gamma}^{L+1}_N(\tilde{\tau}_N,t,\tilde{s}_L))^T,\\
&\tilde{\mu}^{L+1}(t,\tilde{s}_L) = \begin{bmatrix}\mu(t)\\ \tilde{\mu}^L(t,\tilde{s}_L)\end{bmatrix},\\
&\widetilde{\Gamma}^{L+1}_N(\tilde{\tau}_N,t,\tilde{s}_L) = \left[\widetilde{\Gamma}^T_{0N}(\tilde{\tau}_N,t)\vdots\widetilde{\Gamma}^L_{N,N+1}(\tilde{\tau}_N,t,\tilde{s}_L)\right]. \qquad (4.14)
\end{aligned}$$

Formulae (2.5), (3.5), (3.11), (4.1)–(4.3), and (4.13) imply that

$$\begin{aligned}
&\overline{h(\tilde{\tau}_N,z|x,\tilde{x}^L)} - \overline{h(t,z)}\\
&\quad= H_0[x-\mu(t)] + H_{1,N}\widetilde{\Gamma}^{L+1}_N(\widetilde{\Gamma}^{L+1})^{-1}\left[\tilde{x}^{L+1} - \tilde{\mu}^{L+1}(t,\tilde{s}_L)\right]. \qquad (4.15)
\end{aligned}$$

Then from (4.15) taking into account (4.4) and (4.14), we obtain

$$\begin{aligned}
&M\left\{\left[\overline{h(\tilde{\tau}_N,z|x_t,\tilde{x}^L_s)} - \overline{h(t,z)}\right][\cdot]^T|z^t_0,\eta^m_0\right\} = H_0\Gamma H^T_0\\
&\quad+\left[H_0\widetilde{\Gamma}^{L+1}_{0,N+1} + H_{1,N}\widetilde{\Gamma}^{L+1}_N\right](\widetilde{\Gamma}^{L+1})^{-1}(\widetilde{\Gamma}^{L+1}_N)^T H^T_{1,N}\\
&\quad+H_{1,N}\widetilde{\Gamma}^{L+1}_N(\widetilde{\Gamma}^{L+1})^{-1}(\widetilde{\Gamma}^{L+1}_{0,N+1})^T H^T_0, \qquad (4.16)\\
&\widetilde{\Gamma}^{L+1}_{0,N+1}(t,\tilde{s}_L) = M\left\{[x_t-\mu(t)]\left[\tilde{x}^{L+1}_{t,s} - \tilde{\mu}^{L+1}(t,\tilde{s}_L)\right]^T|z^T_0,\eta^m_0\right\}\\
&\quad= \left[\Gamma(t)\vdots\widetilde{\Gamma}^L_{0,N+1}(t,\tilde{s}_L)\right]. \qquad (4.17)
\end{aligned}$$

From (4.2), (4.9)–(4.11), (4.14), and (4.17), we obtain $H_0\widetilde{\Gamma}^{L+1}_{0,N+1} + H_{1,N}\widetilde{\Gamma}^{L+1}_N = \widetilde{H}_{L+1}$. Hence $H_{1,N}\widetilde{\Gamma}^{L+1}_N = \widetilde{H}_{L+1} - H_0\widetilde{\Gamma}^{L+1}_{0,N+1}$ and from (4.16)

$$M\left\{\left[\overline{h(\tilde{\tau}_N,z|x_t,\tilde{x}^L_s)} - \overline{h(t,z)}\right][\cdot]^T|z^t_0,\eta^m_0\right\}$$

$$= \widetilde{H}_{L+1}(\widetilde{\Gamma}^{L+1})^{-1}\widetilde{H}_{L+1}^T + H_0\left[\Gamma - \widetilde{\Gamma}_{0,N+1}^{L+1}(\widetilde{\Gamma}^{L+1})^{-1}(\widetilde{\Gamma}_{0,N+1}^{L+1})^T\right]H_0^T. \quad (4.18)$$

Assume that

$$(\widetilde{\Gamma}^{L+1})^{-1} = \begin{bmatrix} \Gamma & \widetilde{\Gamma}_{0,N+1}^L \\ (\widetilde{\Gamma}_{0,N+1}^L)^T & \widetilde{\Gamma}^L \end{bmatrix}^{-1} = \begin{bmatrix} C_{00} & C_{01} \\ C_{01}^T & C_{11} \end{bmatrix}. \quad (4.19)$$

Then, from (4.17) and (4.19), we obtain

$$\widetilde{\Gamma}_{0,N+1}^{L+1}(\widetilde{\Gamma}^{L+1})^{-1}(\widetilde{\Gamma}_{0,N+1}^{L+1})^T$$
$$= \Gamma C_{00}\Gamma + \Gamma C_{01}(\widetilde{\Gamma}_{0,N+1}^L)^T + \widetilde{\Gamma}_{0,N+1}^L C_{01}^T\Gamma + \widetilde{\Gamma}_{0,N+1}^L C_{11}(\widetilde{\Gamma}_{0,N+1}^L)^T. \quad (4.20)$$

By the Frobenius formula (Gantmakher, 1988) in accordance with (4.19), we have

$$C_{00} = \left[\Gamma - \widetilde{\Gamma}_{0,N+1}^L(\widetilde{\Gamma}^L)^{-1}(\widetilde{\Gamma}_{0,N+1}^L)^T\right]^{-1}, \quad C_{01} = -C_{00}\widetilde{\Gamma}_{0,N+1}^L(\widetilde{\Gamma}^L)^{-1},$$
$$C_{11} = (\widetilde{\Gamma}^L)^{-1} + (\widetilde{\Gamma}^L)^{-1}(\widetilde{\Gamma}_{0,N+1}^L)^T C_{00}\widetilde{\Gamma}_{0,N+1}^L(\widetilde{\Gamma}^L)^{-1}. \quad (4.21)$$

Using the (4.21) in (4.20) gives

$$\widetilde{\Gamma}_{0,N+1}^{L+1}(\widetilde{\Gamma}^{L+1})^{-1}(\widetilde{\Gamma}_{0,N+1}^{L+1})^T = \Gamma. \quad (4.22)$$

Then, from (4.18) and (4.22), we obtain

$$M\left\{\left[\overline{h(\tilde{\tau}_N, z|x_t, \tilde{x}_s^L)} - \overline{h(t,z)}\right][\cdot]^T|z_0^t, \eta_0^m\right\}$$
$$= \widetilde{H}_{L+1}(t,z)(\widetilde{\Gamma}^{L+1}(t,\tilde{s}_L))^{-1}\widetilde{H}_{L+1}^T(t,z). \quad (4.23)$$

Since $p_s^t(x;\tilde{x}^L) = \mathcal{N}\{\cdot\}$ (see Proposition 2), then for $p_{t|s}^t(x|\tilde{x}^L) = \partial\mathcal{P}\{x_t \leqslant x|\tilde{x}_s^L = \tilde{x}^L, z_0^t, \eta_0^m\}/\partial x$ similar to (4.13) from (4.4) the property of $p_{t|s}^t(x|\tilde{x}^L) = \mathcal{N}\{x;\mu(t|\tilde{s}_L), \Gamma(t|\tilde{s}_L)\}$ takes place

$$\mu(t|\tilde{s}_L) = \mu(t) + \widetilde{\Gamma}_{0,N+1}^L(\widetilde{\Gamma}^L)^{-1}[\tilde{x}^L - \tilde{\mu}^L(t,\tilde{s}_L)], \quad (4.24)$$

and $\widetilde{\Gamma}(t|\tilde{s}_L)$ is defined by the formula (4.7). Since $p_s^t(x;\tilde{x}^L) = p_{t|s}^t(x|\tilde{x}^L)p_s^t(\tilde{x}^L)$, then, we obtain

$$\partial\ln[p_s^t(x;\tilde{x}^L)]/\partial x = \partial\ln[p_{t|s}^t(x|\tilde{x}^L)]/\partial x = -\Gamma^{-1}(t|\tilde{s}_L)[x - \mu(t|\tilde{s}_L)]. \quad (4.25)$$

Thus, we have

$$M\left\{\left[\partial\ln[p_s^t(x_t;\tilde{x}_s^L)]/\partial x_t\right]\left[\partial\ln[p_s^t(x_t;\tilde{x}_s^L)]/\partial x_t\right]^T|z_0^t, \eta_0^m\right\} = \Gamma^{-1}(t|\tilde{s}_L). \quad (4.26)$$

Relation (4.4) implies that (see Proposition 2) $p_t(x) = \mathcal{N}\{x; \mu(t), \Gamma(t)\}$. Hence, we obtain

$$\partial \ln[p_t(x)]/\partial x = -\Gamma^{-1}(t)[x - \mu(t)], \tag{4.27}$$

$$M\left\{[\partial \ln[p_t(x_t)]/\partial x_t]\,[\partial \ln[p_t(x_t)]/\partial x_t]^T\,|z_0^t, \eta_0^m\right\} = \Gamma^{-1}(t). \tag{4.28}$$

Formulae (4.24), (4.25), and (4.27) imply that

$$M\left\{[\partial \ln[p_s^t(x_t; \tilde{x}_s^L)]/\partial x_t][\partial \ln[p_t(x_t)]/\partial x_t]^T|z_0^t, \eta_0^m\right\} = \Gamma^{-1}(t). \tag{4.29}$$

Then, in accordance with (4.28) and (4.29), we have

$$M\left\{\left[\frac{\partial \ln p_s^t(x_t; \tilde{x}_s^L)}{\partial x_t} - \frac{\partial \ln p_t(x_t)}{\partial x_t}\right]\left(\frac{\partial \ln p_t(x_t)}{\partial x_t}\right)^T\middle|\, z_0^t, \eta_0^m\right\} = O. \tag{4.30}$$

Analogous calculations relative to the unconditional expectation for prior densities (see Remark 3) result in the formulae

$$M\left\{\left[\partial \ln[p(t, x_t; \tilde{s}_L, \tilde{x}_s^L)]/\partial x_t\right]\left[\partial \ln[p(t, x_t; \tilde{s}_L, \tilde{x}_s^L)]/\partial x_t\right]^T\right\} = D^{-1}(t|\tilde{s}_L),$$

$$M\left\{\left[\frac{\partial \ln p(t, x_t; \tilde{s}_L, \tilde{x}_s^L)}{\partial x_t} - \frac{\partial \ln p(t, x_t)}{\partial x_t}\right]\left(\frac{\partial \ln p(t, x_t)}{\partial x_t}\right)^T\right\} = O. \tag{4.31}$$

Substitution (4.23), (4.26), (4.30), (4.31) in (3.8), taking into account the property $M\{\cdot\} = M\{M\{\cdot|z_0^t, \eta_0^m\}\}$ gives (4.5).

Relation (3.16) implies that $[p_s^{t_m}(x; \tilde{x}^L)/p_s^{t_m-0}(x; \tilde{x}^L)] = [C(\eta(t_m), z|x, \tilde{x}^L)/C(\eta(t_m), z)]$. In accordance with (4.4), (4.6), and (4.14), we have $p_s^t(x; \tilde{x}^L) = p_s^t(\tilde{x}^{L+1}) = \mathcal{N}\{\tilde{x}^{L+1}; \tilde{\mu}^{L+1}(t, \tilde{s}_L),\ \widetilde{\Gamma}^{L+1}(t, \tilde{s}_L)\}$. Taking into account $M\{\cdot\} = M\{M\{\cdot|z_0^{t_m}, \eta_0^{m-1}\}\}$ and $M\{\cdot\} = M\{M\{\cdot|z_0^{t_m}, \eta_0^m\}\}$, we obtain

$$\begin{aligned} &M\left\{\ln\left[C(\eta(t_m), z|x_{t_m}, \tilde{x}_s^L)/C(\eta(t_m), z)\right]\right\} \\ &= M\Bigg\{\ln\Big[\mathcal{N}\left\{\tilde{x}^{L+1}; \tilde{\mu}^{L+1}(t_m, \tilde{s}_L), \widetilde{\Gamma}^{L+1}(t_m, \tilde{s}_L)\right\} \\ &\qquad /\mathcal{N}\left\{\tilde{x}^{L+1}; \tilde{\mu}^{L+1}(t_m - 0, \tilde{s}_L), \widetilde{\Gamma}^{L+1}(t_m - 0, \tilde{s}_L)\right\}\Big]\Bigg\} = \\ &= (1/2)M\left\{\ln\left[|\widetilde{\Gamma}^{L+1}(t_m - 0, \tilde{s}_L)|/|\widetilde{\Gamma}^{L+1}(t_m, \tilde{s}_L)|\right]\right\}. \end{aligned} \tag{4.32}$$

Then, formulae (3.12), and (4.32) imply (4.12).

COROLLARY 2. The information amount (3.22) on the time intervals $t_m \leqslant t < t_{m+1}$ is defined by the equation

$$\begin{aligned} \mathrm{d}\, I_t[x_t; z_0^t, \eta_0^m]/\mathrm{d}\, t = {} & (1/2)tr\left[M\left\{R^{-1}(t, z)\widetilde{H}_0(t, z)\Gamma^{-1}(t)\widetilde{H}_0^T(t, z)\right\}\right] \\ & -(1/2)tr\left[Q(t)\left[M\{\Gamma^{-1}(t)\} - D^{-1}(t)\right]\right], \end{aligned} \tag{4.33}$$

subject to the initial condition (3.24), where

$$\Delta I_{t_m}[\cdot] = (1/2)M\left\{\ln\left[|\Gamma(t_m - 0)|/|\Gamma(t_m)|\right]\right\}, \tag{4.34}$$

$\Gamma(t_m - 0) = \lim \Gamma(t)$ subject to $t \uparrow t_m$ and $\widetilde{H}_0(t, z)$ is defined in (4.10).

The formulated result is obtained as a limitary case from Theorem 3 subject to $s_l \downarrow t$ in (4.5) and $s_l \downarrow t_m$ in (4.12), $l = \overline{1; L}$. Note that the same result can be proved with the use of Corollary 1, analogously to the proof of Theorem 3. Similarly proof of Theorems 2, 4 in (Dyomin and Korotkevich, 1987) for the case $N = 1$ was made.

Theorem 4. *The information amount* (3.30) *on the time intervals* $t_m \leqslant t < t_{m+1}$ *is defined by the equation*

$$\begin{aligned} &\mathrm{d}\, I_s^t[\tilde{x}_s^L; z_0^t, \eta_0^m]/\mathrm{d}\, t \\ &\quad = (1/2)tr\left[M\left\{R^{-1}(t,z)\widetilde{H}_L(t,z)(\widetilde{\Gamma}^L(t,\tilde{s}_L))^{-1}\widetilde{H}_L^T(t,z)\right\}\right], \end{aligned} \tag{4.35}$$

subject to the initial condition (3.33)*, where*

$$\Delta I_s^{t_m}[\cdot] = (1/2)M\left\{\ln\left[|\widetilde{\Gamma}^L(t_m - 0, \tilde{s}_L)|/|\widetilde{\Gamma}^L(t_m, \tilde{s}_L)|\right]\right\}, \tag{4.36}$$

$\widetilde{\Gamma}^L(t_m - 0, \tilde{s}_L) = \lim \widetilde{\Gamma}^L(t, \tilde{s}_L)$ *subject to* $t \uparrow t_m$ *and* $\widetilde{H}_L(t, z)$ *is defined in* (4.9).

Proof. For $p_{\tau,t|s}^t(x, \tilde{x}_N|\tilde{x}^L) = p_{\tau,t|s}^t(\tilde{x}_{N+1}|\tilde{x}^L) = \partial^{N+1}\mathcal{P}\{\tilde{x}_{t,\tau}^{N+1} \leqslant \tilde{x}_{N+1}|\tilde{x}_s^L = \tilde{x}^L, z_0^t, \eta_0^m\}/\partial\tilde{x}_{N+1}$ similar to (4.13) it follows that (see (4.4))

$$\begin{aligned} p_{\tau,t|s}^t(\tilde{x}_{N+1}|\tilde{x}^L) &= \mathcal{N}\{\tilde{x}_{N+1}; \tilde{\mu}_{N+1}(\tilde{\tau}_N, t|\tilde{s}_L), \widetilde{\Gamma}_{N+1}(\tilde{\tau}_N, t|\tilde{s}_L)\}, \\ \tilde{\mu}_{N+1}(\tilde{\tau}_N, t|\tilde{s}_L) &= \tilde{\mu}_{N+1}(\tilde{\tau}_N, t) \\ &\quad + \widetilde{\Gamma}_{N+1}^L(\tilde{\tau}_N, t, \tilde{s}_L)(\widetilde{\Gamma}^L(t, \tilde{s}_L))^{-1}[\tilde{x}^L - \tilde{\mu}^L(t, \tilde{s}_L)], \\ \widetilde{\Gamma}_{N+1}(\tilde{\tau}_N, t|\tilde{s}_L) &= \widetilde{\Gamma}_{N+1}(\tilde{\tau}_N, t) \\ &\quad - \widetilde{\Gamma}_{N+1}^L(\tilde{\tau}_N, t, \tilde{s}_L)(\widetilde{\Gamma}^L(t, \tilde{s}_L))^{-1}(\widetilde{\Gamma}_{N+1}^L(\tilde{\tau}_N, t, \tilde{s}_L))^T, \end{aligned} \tag{4.37}$$

$$\begin{aligned} &\tilde{\mu}_{N+1}(\tilde{\tau}_N, t) = \begin{bmatrix} \mu(\cdot) \\ \tilde{\mu}_N(\cdot) \end{bmatrix}, \quad \widetilde{\Gamma}_{N+1}(\tilde{\tau}_N, t) = \begin{bmatrix} \Gamma(\cdot) & \widetilde{\Gamma}_{0N}(\cdot) \\ \widetilde{\Gamma}_{0N}^T(\cdot) & \widetilde{\Gamma}_N(\cdot) \end{bmatrix}, \\ &\widetilde{\Gamma}_{N+1}^L(\tilde{\tau}_N, t, \tilde{s}_L) = \begin{bmatrix} \widetilde{\Gamma}_{0,N+1}^L(\cdot) \\ \widetilde{\Gamma}_{N,N+1}^L(\cdot) \end{bmatrix}. \end{aligned} \tag{4.38}$$

Formulae (3.5), (3.35), (4.1)–(4.3), and (4.37) imply that $\overline{h(\tilde{\tau}_N, t, z|\tilde{x}^L)} - \overline{h(t, z)} = H_{0,N}\widetilde{\Gamma}_{N+1}^L(\widetilde{\Gamma}^L)^{-1}[\tilde{x}^L - \tilde{\mu}^L(t, \tilde{s}_L)]$. In accordance with (4.2), (4.3), (4.9), and (4.38), we have $H_{0,N}\widetilde{\Gamma}^L_{N+1} = \widetilde{H}_L$. Therefore, we obtain $\overline{h(\tilde{\tau}_N,t,z|\tilde{x}^L)} - \overline{h(t,z)} = \widetilde{H}_L(\widetilde{\Gamma}^L)^{-1}[\tilde{x}^L - \tilde{\mu}^L(t,\tilde{s}_L)]$. Thus, we have

$$M\left\{\left[\overline{h(\tilde{\tau}_N,t,z|\tilde{x}^L_s)} - \overline{h(t,z)}\right][\cdot]^T|z^t_0,\eta^m_0\right\}$$
$$= \widetilde{H}_L(t,z)(\widetilde{\Gamma}^L(t,\tilde{s}_L))^{-1}\widetilde{H}^T_L(t,z). \tag{4.39}$$

Subsitution of (4.39) into (3.32) taking into account $M\{\cdot\} = M\{M\{\cdot|z^t_0,\eta^m_0\}\}$, gives (4.35). From (4.4) (see Proposition 2), we obtain $p^t_s(\tilde{x}^L_s) = \mathcal{N}\{\tilde{x}^L;\tilde{\mu}^L(t,\tilde{s}_L),\widetilde{\Gamma}^L(t,\tilde{s}_L)\}$. Therefore (4.36) is derived on the basis (3.34), (3.38) analogously (4.12).

COROLLARY 3. Let in (4.1) coefficients dependence on z is absent. Then Theorems 3, 4 and Corollary 2 take place, where dependence on z and operator $M\{\cdot\}$ are absent. Thus, exact calculation $I^t_s[x_t,\tilde{x}^L_s;z^t_0,\eta^m_0]$, $I_t[x_t;z^t_0,\eta^m_0]$, $I^t_s[\tilde{x}^L_s;z^t_0,\eta^m_0]$ is possible only in the conditionally-Gaussian case in the absence of feedback in the observation channels (see Remark 2).

In the next paragraphs some of the obtained results are applied to the problem investigation of stochastic process transmission on the continuous-discrete memory channels in some particular cases.

5. The Information Efficiency of the Memory Observations in Relative to the Memoryless Observations

The problem of efficiency of the memory observation, i.e., whether presence of memory increases or decreases information amount, is of interest. The given investigation is to be carried out for a particular case of the scalar stationary processes x_t, z_t, $\eta(t_m)$ defined by the equations (see (2.1)–(2.3), (4.1), (4.2))

$$\mathrm{d}\,x_t = -ax_t\mathrm{d}\,t + \sqrt{Q}\,\mathrm{d}\,\omega_t, \quad a>0, \quad p_0(x) = \mathcal{N}\{\mu_0;\gamma_0\},$$
$$\mathrm{d}\,z_t = H_0x_t\mathrm{d}\,t + \sqrt{R}\,\mathrm{d}\,v_t, \quad \eta(t_m) = G_0x_{t_m} + G_1x_\tau + \sqrt{V}\xi(t_m), \tag{5.1}$$

when continuous memoryless observation, and discrete memory observations of unit multiplicity, i.e., process x_t has the form as in (Dyomin *et al.*, 2001; item 5). As the information efficiency measure of the memory observations $\eta(t_m)$ with regard to the memoryless observations $\widetilde{\eta}(t_m)$, when $G_1 = 0$, in extrapolation problem for the case $L = 1$ $(s_1 = s)$ one can accept the value $\Delta = \Delta I^{t_m}_s[x_s;z^{t_m}_0,\eta(t_m)] - \widetilde{\Delta} I^{t_m}_s[x_s;z^{t_m}_0,\widetilde{\eta}(t_m)]$, where $\Delta I^{t_m}_s[\cdot]$ and $\widetilde{\Delta} I^{t_m}_s[\cdot]$ are information amount increments (3.30) by $L = 1$ in the time moments t_m, incoming from the observations $\eta(t_m)$ and $\widetilde{\eta}(t_m)$, respectively. Consider the case of sparse discrete time observations, when on the intervals $t\in(t_m,t_{m+1})$ solutions of the differential equations for the elements of the matrix $\widetilde{\Gamma}_3(\tau,t,s)$ (see (4.4)) attain the stationary values γ, $\gamma_{01}(t^*)$, $\gamma_{11}(t^*)$, $\gamma^{11}(T)$, $\gamma^1_0(T)$, $\gamma^1_1(t^*,T)$, defined by the formula (3.19) from (Dyomin *et al.*, 2000), where $t^* = t-\tau$ and $T = s-t$ are memory depth and extrapolation interval, respectively. Then, in accordance with (4.36) and Corollary 3 using (2.28), (2.33) from (Dyomin *et al.*, 1997)

$$\Delta = (1/2)\ln\left[\widetilde{\gamma}^{11}(s,t_m)/\gamma^{11}(s,t_m)\right], \tag{5.2}$$

$$\gamma^{11}(s,t_m) = \gamma^{11}(T) - \frac{[G_0\gamma_0^1(T) + G_1\gamma_1^1(t^*,T)]^2}{V + G_0^2\gamma + G_1^2\gamma_{11}(t^*) + 2G_0G_1\gamma_{01}(t^*)},$$
$$\widetilde{\gamma}^{11}(s,t_m) = \gamma^{11}(T) - [G_0^2(\gamma_0^1(T))^2/(V + G_0^2\gamma)]. \tag{5.3}$$

There are two marginal situations with regard to memory depth: the case of small memory depth, when $t^* \to 0$; the case of large memory depth, when $t^* \to \infty$. Assume that $\Delta_0 = \lim\Delta$ subject to $t^* \to 0$ and $\Delta_\infty = \lim\Delta$ subject to $t^* \to \infty$. From (5.2), and (5.3) taking into account (3.19) in (Dyomin *et al.*, 2000), we obtain

$$\Delta_0 = (1/2)\ln[1/(1-\delta_0)], \quad \Delta_\infty = (1/2)\ln[1/(1+\delta_\infty)], \tag{5.4}$$

$$\delta_0 = \frac{2aV\gamma^2(G_1^2 + 2G_0G_1)\exp\{-2aT\}}{[V+\gamma(G_0+G_1)^2][Q(V+\gamma G_0^2)(1-\exp\{-2aT\})+2aV\gamma\exp\{-2aT\}]}, \tag{5.5}$$

$$\delta_\infty = \frac{2aæ\gamma^3G_0^2G_1^2\exp\{-2aT\}}{[V+\gamma(G_0^2+æG_1^2)][Q(V+\gamma G_0^2)(1-\exp\{-2aT\})+2aV\gamma\exp\{-2aT\}]}. \tag{5.6}$$

Research of behavior of the $\Delta(t^*)$ as the function of the memory depth t^* basing on (5.2)–(5.6) with the use of (3.19) from (Dyomin *et al.*, 2000), gives the result.

PROPOSITION 3. Assume that

$$\mathcal{M} = \mathcal{M}^+ \cup \mathcal{M}^- = \{(G_0,G_1)\colon\ G_1^2 + 2G_0G_1 \leqslant 0\}. \tag{5.7}$$

If $(G_0,G_1) \notin \mathcal{M}$, then $\Delta(t^*)$ is monotonically diminishing function of the memory depth from the value $\Delta_0 > 0$ up to the value $\Delta_\infty < 0$, and is equal to zero at the point $t^* = t^*_{eff}$ determined the formula

$$t^*_{eff} = \frac{1}{\lambda}\ln\frac{|G_1|(V + æ\gamma G_0^2)}{|G_0|([V^2 + æ\gamma G_1^2(V + æ\gamma G_0^2)]^{1/2} \mp V)}, \tag{5.8}$$

where sign "-" if $G_0G_1 = |G_0|\cdot|G_1|$, and sign "+" if $G_0G_1 = -|G_0|\cdot|G_1|$, $\lambda = (a^2+\delta Q)^{1/2}$, $\delta = H_0^2/R$, $æ = (\lambda+a)/2\lambda$, $\gamma = (1/\delta)(\lambda - a)$, and which can be defined as an effective memory depth. If $(G_0,G_1) \in \mathcal{M}$, then $\Delta(t^*) \leqslant 0$ for all $t^* \geqslant 0$.

A physical interpretation of this result is the following. In the case of large memory depth $t^* \gg \alpha_k$, where $\alpha_k = 1/a$ is the correlation time of the process x_t, there is no correlation between x_τ, and x_{t_m}, x_s. Therefore by great t^* the signal $Y(\tau) = G_1x_\tau$ does not contain information on the current x_{t_m} and on the future values x_s of the process x_t and plays the role of additional noise in the memory channel which leads to decrease in the information amount increment as compared with the memoryless channel. Thus one can explain why $\Delta_\infty < 0$ by random values of the transmission coefficients G_0 and G_1. In the case of small memory depth, when $t^* \ll \alpha_k$, the correlation coefficient between x_τ and x_{t_m} is close to one, and therefore, the signal $Y(t_m) = G_0 x_{t_m} + G_1 x_\tau$ is accepted as $Y(t_m) = (G_0 + G_1)x_{t_m}$. Since the condition $(G_0, G_1) \notin \mathcal{M}$ means $|G_0 + G_1| > |G_0|$, then the useful signal strength $Y(\tau, t_m)$ in the memory channel is higher than the useful signal strengh $G_0 x_{t_m}$ in the memoryless channel, which provides great self-descriptiveness $Y(\tau, t_m)$ with regard to $G_0 x_{t_m}$. This explains the property $\Delta_0 > 0$ in the case $(G_0, G_1) \notin \mathcal{M}$ and an inverse property by a contrary condition. The condition $(G_0, G_1) \notin \mathcal{M}$ is an existence condition of the single positive root of the equation $\Delta(t^*) = 0$, solution of which is given by (5.8). Influence of continuous observations on the discrete observation self-descriptiveness is carried out through the parameter $\delta = H_0^2/R$, which is proportional to the signal-noise ratio by the strengh in the continuous observation channel. If $\delta \to \infty$ we obtain $\Delta I_s^{t_m}[\cdot] \to 0$ and $\widetilde{\Delta} I_s^{t_m}[\cdot] \to 0$, that yields $\Delta \to 0$. Hence, on obtaining absolutely accurate measurement in the continuous channel, the discrete observations both with memory and without memory do not introduce new information on the values x_s for all T. If $\delta = 0$ that corresponds by the case of continuous observation absence, formulae (5.2)–(5.8), where $\gamma = Q/2a$, $\lambda = a$, æ $= 1$ are correct, i.e., in this case we have evident dependence t^*_{eff} on the correlation time $\alpha_k = 1/a$ of the process x_t.

6. Optimal Transmission of the Gaussian Markov Process over the Memory Channels by the Silent Feedback

The signal x_t, an output message of the continuous transmission channel z_t and an output message of the discrete transmission channel $\eta(t_m)$ are scalar and defined in accordance with (2.1)–(2.3) in the form

$$\mathrm{d}\,x_t = F(t)x_t \mathrm{d}\,t + \Phi_1(t)\mathrm{d}\,\omega_t, \quad p_0(x) = \mathcal{N}\{x; \mu_0, \gamma_0\}, \tag{6.1}$$

$$\mathrm{d}\,z_t = h(t, x_t, x_\tau, z)\mathrm{d}\,t + \Phi_2(t)\mathrm{d}\,v_t, \quad \eta(t_m) = g(t_m, x_{t_m}, x_\tau, z) + \Phi_3(t_m)\xi(t_m). \tag{6.2}$$

Problem formulation: in the class of coding functionals $\mathcal{K} = \{\mathcal{H}; \mathcal{G}\} = \{h(\cdot); g(\cdot)\}$, satisfying energy limitation

$$M\{h^2(t, x_t, x_\tau, z)\} \leqslant \tilde{h}(t) \leqslant \tilde{h}, \quad M\{g^2(t_m, x_{t_m}, x_\tau, z)\} \leqslant \tilde{g}(t_m) \leqslant \tilde{g}, \tag{6.3}$$

the functionals $h^0(\cdot)$ and $g^0(\cdot)$, which provide the minimal decoding error $\Delta^0(t) = \inf \Delta(t)$ with regard to a filtering problem, are to be found. $\Delta(t) = M\{[x_t - \widehat{x}(t, z, \eta)]^2\}$ is the filtering estimate error $\widehat{x}(t, z, \eta)$ of the process x_t corresponding message $\{z_0^t; \eta_0^m\}$ accepted by the given $h(\cdot)$ and $g(\cdot)$.

This problem is a generalization of the problem from (Liptser, 1974) for the case continuous-discrete transmission with the memory of unit multiplicity $(N = 1, \tau_1 = \tau)$.

REMARK 4. Up to the moment τ the transmission is proceeded an optimal manner.

Since given $h(\cdot)$ and $g(\cdot)$, a posteriori mean $\mu(t) = M\{x_t|z_0^t, \eta_0^m\}$ (Liptser and Shiryayev, 1977; 1978) is optimal in root-mean-square sense filtering estimate, then $\Delta(t) \geqslant M\{\gamma(t)\}$, where $\gamma(t) = M\{[x_t - \mu(t)]^2|z_0^t, \eta_0^m\}$. Thus, we have $\Delta^0(t) = \inf M\{\gamma(t)\}$.

Theorem 5. *In the class $\mathcal{K}_l = \{\mathcal{H}_l; \mathcal{G}_l\}$ of linear functionals*

$$\mathcal{H}_l = \{h(\cdot)\colon\ h(t, x_t, x_\tau, z) = h(t, z) + H_0(t, z)x_t + H_1(t, z)x_\tau\},$$
$$\mathcal{G}_l = \{g(\cdot)\colon\ g(t_m, x_{t_m}, x_\tau, z) = g(t_m, z) + G_0(t_m, z)x_{t_m} + G_1(t_m, z)x_\tau\} \quad (6.4)$$

1^0) *optimal coding functionals $h^0(\cdot)$, $g^0(\cdot)$ are defined in the form*

$$h^0(t, z^0) = -H_0^0(t, z^0)\mu^0(t),$$
$$H_0^0(t, z^0) = [\tilde{h}(t)/\Delta^0(t)]^{1/2}, \quad H_1^0(t, z^0) = 0, \quad (6.5)$$
$$g^0(t_m, z^0) = -G_0^0(t_m, z^0)\mu^0(t_m - 0),$$
$$G_0^0(t_m, z^0) = [\tilde{g}(t_m)/\Delta^0(t_m - 0)]^{1/2}, \quad G_1^0(t_m, z^0) = 0; \quad (6.6)$$

2^0) *optimal message $\{z_t^0; \eta^0(t_m)\}$ is defined by the equations*

$$\mathrm{d}\, z_t^0 = [\tilde{h}(t)/\Delta^0(t)]^{1/2}[x_t - \mu^0(t)]\mathrm{d}\, t + \Phi_2(t)\mathrm{d}\, v_t, \quad (6.7)$$
$$\eta^0(t_m) = [\tilde{g}(t_m)/\Delta^0(t_m - 0)]^{1/2}[x_{t_m} - \mu^0(t_m - 0)] + \Phi_3(t_m)\xi(t_m); \quad (6.8)$$

3^0) *optimal decoding $\mu^0(t)$ and a minimal decoding error $\Delta^0(t)$ on the intervals $t_m \leqslant t < t_{m+1}$, are defined by the equations*

$$\mathrm{d}\, \mu^0(t) = F(t)\mu^0(t)\mathrm{d}\, t + R^{-1}(t)[\tilde{h}(t)\Delta^0(t)]^{1/2}\mathrm{d}\, z_t^0, \quad (6.9)$$
$$\mathrm{d}\, \Delta^0(t)/\mathrm{d}\, t = [2F(t) - R^{-1}(t)\tilde{h}(t)]\Delta^0(t) + Q(t), \quad (6.10)$$

subject to the initial condition

$$\mu^0(t_m) = \mu^0(t_m - 0) + [\tilde{g}(t_m)\Delta^0(t_m - 0)]^{1/2}[V(t_m) + \tilde{g}(t_m)]^{-1}\eta^0(t_m), \quad (6.11)$$
$$\Delta^0(t_m) = V(t_m)[V(t_m) + \tilde{g}(t_m)]^{-1}\Delta^0(t_m - 0), \quad (6.12)$$

where $Q(t) = \Phi_1^2(t)$, $R(t) = \Phi_2^2(t)$, $V(t_m) = \Phi_3^2(t_m)$, $\mu^0(t_m - 0) = \lim \mu(t)$, $\Delta^0(t_m - 0) = \lim \Delta(t)$ subject to $t \uparrow t_m$.

Proof. Given $\{h(\cdot); g(\cdot)\} \in \mathcal{K}_l$ on the intervals $t_m \leqslant t < t_{m+1}$ (see (Abakumova *et al.*, 1995b; Dyomin *et al.*, 1997) and Proposition 2) $\mu(t)$ and $\gamma(t)$ are defined by the equations

$$\mathrm{d}\, \mu(t) = F(t)\mu(t)\mathrm{d}\, t + R^{-1}(t)[H_0(t, z)\gamma(t) + H_1(t, z)\gamma_{01}(\tau, t)][\mathrm{d}\, z_t$$
$$-(h(t, z) + H_0(t, z)\mu(t) + H_1(t, z)\mu(\tau, t))\mathrm{d}\, t], \quad (6.13)$$
$$\mathrm{d}\, \gamma(t)/\mathrm{d}\, t = 2F(t)\gamma(t) - R^{-1}(t)[H_0(t, z)\gamma(t) + H_1(t, z)\gamma_{01}(\tau, t)]^2 + Q(t), \quad (6.14)$$

subject to the initial condition

$$\mu(t_m) = \mu(t_m - 0) + [G_0(t_m, z)\gamma(t_m - 0) + G_1(t_m, z)\gamma_{01}(\tau, t_m - 0)]\, W^{-1}(t_m)$$
$$\times [\eta(t_m) - g(t_m, z) - G_0(t_m, z)\mu(t_m - 0) - G_1(t_m, z)\mu(\tau, t_m - 0)]\,, \quad (6.15)$$
$$\gamma(t_m) = \gamma(t_m - 0) - [G_0(t_m, z)\gamma(t_m - 0) + G_1(t_m, z)\gamma_{01}(\tau, t_m - 0)]^2 W^{-1}(t_m), \quad (6.16)$$

where $\mu(\tau, t) = M\{x_\tau | z_0^t, \eta_0^m\}$, $\gamma_{01}(\tau, t) = M\{[x_t - \mu(t)][x_\tau - \mu(\tau, t)] | z_0^t, \eta_0^m\}$, $\gamma_{11}(\tau, t) = M\{[x_\tau - \mu(\tau, t)]^2 | z_0^t, \eta_0^m\}$,

$$W(t_m) = V(t_m) + G_0^2(t_m, z)\gamma(t_m - 0) + G_1^2(t_m, z)\gamma_{11}(\tau, t_m - 0)$$
$$+ 2G_0(t_m, z)G_1(t_m, z)\gamma_{01}(\tau, t_m - 0). \quad (6.17)$$

Suppose up to the moment t_m the transmission was proceeded in an optimal manner. Then, from (6.16), and (6.17), we obtain

$$\gamma(t_m) = V(t_m)\Delta^0(t_m - 0)(W^0(t_m))^{-1} + G_1^2(t_m, z^0)$$
$$\times [\Delta^0(t_m - 0)\Delta_{11}^0(\tau, t_m - 0) - (\Delta_{01}^0(\tau, t_m - 0))^2](W^0(t_m))^{-1}, \quad (6.18)$$

where $W^0(t_m)$ is defined by the formula (6.17) with replacement z by z^0, $\gamma(t_m - 0)$ by $\Delta^0(t_m - 0)$, $\gamma_{01}(\tau, t_m - 0)$ by $\Delta_{01}^0(\tau, t_m - 0)$, $\gamma_{11}(\tau, t_m - 0)$ by $\Delta_{11}^0(\tau, t_m - 0)$. For $t < t_m$ by Cauhy–Schwarz–Bunyakovskii inequality in relative to $M\{\cdot | z_0^t, \eta_0^{m-1}\}$ (Lipser and Shiryayev, 1977; 1978), we have $\gamma(t)\gamma_{11}(\tau, t) - \gamma_{01}^2(\tau, t) \geqslant 0$. Since $G_0^2\gamma(t_m - 0) + G_1^2\gamma_{11}(\tau, t_m - 0) + +2G_0G_1\gamma_{01}(\tau, t_m - 0) = M\{[G_0(x_{t_m} - \mu(t_m - 0)) + G_1(x_\tau - \mu(\tau, t_m - 0))]^2 | z_0^t, \eta_0^{m-1}\} \geqslant 0$ then $W(t_m) > 0$. Thus, relation (6.18) implies that

$$\gamma(t_m) \geqslant V(t_m)\Delta^0(t_m - 0)(W^0(t_m))^{-1}. \quad (6.19)$$

By Jensen inequality (Lipser and Shiryayev, 1977; 1978), we have $M\{(W^0(t_m))^{-1}\} \geqslant [M\{W^0(t_m)\}]^{-1}$. Then for $\Delta(t_m) = M\{\gamma(t_m)\}$ from (6.17), (6.19), we obtain

$$\Delta(t_m) \geqslant V(t_m)\Delta^0(t_m - 0)\Big[V(t_m) + M\big\{G_0^2\Delta^0(t_m - 0)$$
$$+ G_1^2\Delta_{11}^0(\tau, t_m - 0) + 2G_0G_1\Delta_{01}^0(\tau, t_m - 0)\big\}\Big]^{-1}. \quad (6.20)$$

Since $M\{\cdot\} = M\{M\{\cdot | z_0^{t_m}, \eta_0^{m-1}\}\}$, the use of (6.4) in (6.3) yields

$$M\left\{g^2(\cdot)\right\} = M\Big\{\big[g(t_m, z) + G_0\mu(t_m - 0) + G_1\mu(\tau, t_m - 0)\big]^2\Big\}$$
$$+ M\big\{G_0^2\gamma(t_m - 0) + G_1^2\gamma_{11}(\tau, t_m - 0) + 2G_0G_1\gamma_{01}(\tau, t_m - 0)\big\}$$
$$\leqslant \tilde{g}(t_m). \quad (6.21)$$

Formulae (6.20), (6.21), and (6.12) imply that

$$\Delta(t_m) \geqslant V(t_m)\Delta^0(t_m - 0)[V(t_m) + \tilde{g}(t_m)]^{-1} = \Delta^0(t_m). \quad (6.22)$$

Use of (6.6) in (6.18) yields that $\gamma^0(t_m) = V(t_m)\Delta^0(t_m-0)[V+\tilde{g}(t_m)]^{-1}$. Coincidence $\gamma^0(t_m)$ with the low bound (6.22) for $\Delta(t_m)$ proves an optimality of the coding (6.6), and (6.8), (6.11) (6.12) follow as a result of substitution (6.6) in (6.2), (6.15), (6.16) given $\{z_0^{t_m}, \eta_0^{m-1}\} = = \{(z^0)_0^{t_m}, (\eta^0)_0^{m-1}\}$.

Addition and subtraction in the right part (6.14) $R^{-1}(t)H_1^2(t,z)\gamma_{11}(\tau,t)$ yieds an equivalent (6.14) integral equation for $t_m \leqslant t < t_{m+1}$, taking into account that at the moment t_m the optimal functional $g^0(\cdot)$ is used

$$\begin{aligned}\gamma(t) = \Delta^0(t_m)\exp\Bigg\{2\int_{t_m}^{t} F(\sigma)\mathrm{d}\,\sigma \\ -\int_{t_m}^{t} R^{-1}(\sigma)\big[H_0^2(\sigma,z)\gamma(\sigma)+H_1^2(\sigma,z)\gamma_{11}(\tau,\sigma)+2H_0(\sigma,z)H_1(\sigma,z)\gamma_{01}(\tau,\sigma)\big]\mathrm{d}\sigma \\ +\int_{t_m}^{t} R^{-1}(\sigma)H_1^2(\sigma,z)\big[\gamma(\sigma)\gamma_{11}(\tau,\sigma)-\gamma_{01}^2(\tau,\sigma)\big]\gamma^{-1}(\sigma)\mathrm{d}\,\sigma\Bigg\} \\ +\int_{t_m}^{t} Q(\sigma)\exp\Bigg\{2\int_{\sigma}^{t} F(u)\mathrm{d}\,u \\ -\int_{\sigma}^{t} R^{-1}(u)\big[H_0^2(u,z)\gamma(u)+H_1^2(u,z)\gamma_{11}(\tau,u)+2H_0(u,z)H_1(u,z)\gamma_{01}(\tau,u)\big]\mathrm{d}\,u \\ +\int_{\sigma}^{t} R^{-1}(u)H_1^2(u,z)[\gamma(u)\gamma_{11}(\tau,u)-\gamma_{01}^2(\tau,u)]\gamma^{-1}(u)\mathrm{d}\,u\Bigg\}\mathrm{d}\,\sigma, \qquad (6.23)\end{aligned}$$

validity of which is proved by differetiating with respect to t. Since $M\{\cdot\} = M\{M\{\cdot|z_0^t,\eta_0^m\}\}$, then the use of (6.4) in (6.3) yields

$$\begin{aligned}M\left\{h^2(\cdot)\right\} = M\left\{[h(t,z)+H_0(t,z)\mu(t)+H_1(t,z)\mu(\tau,t)]^2\right\} \\ +M\left\{H_0^2(t,z)\gamma(t)+H_1^2(t,z)\gamma_{11}(\tau,t)+2H_0(t,z)H_1(t,z)\gamma_{01}(\tau,t)\right\} \\ \leqslant \tilde{h}(t). \qquad (6.24)\end{aligned}$$

By Cauhy–Schwarz–Bunyakovskii inequality as respects $M\{\cdot|z_0^t,\eta_0^m\}$, we have $\gamma(t)\gamma_{11}(\tau,t)-\gamma_{01}^2(\tau,t) \geqslant 0$. Then the use of Jensen inequality $M\{\varphi(\xi)\} \geqslant \varphi(M\{\xi\})$ for the convex function $\varphi(\xi) = \exp\{\xi\}$ in (6.23), taking into account (6.24) for $\Delta(t) = M\{\gamma(t)\}$ result in the inequality

$$\Delta(t) \geqslant \Delta^0(t_m)\exp\left\{\int_{t_m}^{t}\left[2F(\sigma)-R^{-1}(\sigma)\tilde{h}(\sigma)\right]\mathrm{d}\,\sigma\right\}$$

$$+\int_{t_m}^{t} Q(\sigma)\exp\left\{\int_{\sigma}^{t}\Big[2F(u)-R^{-1}(u)\tilde h(u)\Big]\,\mathrm{d}\,u\right\}\mathrm{d}\,\sigma, \tag{6.25}$$

Use of (6.5) in (6.14) for $t_m \leqslant t < t_{m+1}$ results in the equation

$$\begin{aligned}\mathrm{d}\,\gamma^0(t)/\mathrm{d}\,t &= \Big[2F(t)-R^{-1}(t)\tilde h(t)\big(\gamma^0(t)/\Delta^0(t)\big)\Big]\gamma^0(t)+Q(t),\\ \gamma^0(t_m) &= \Delta^0(t_m).\end{aligned} \tag{6.26}$$

Suppose $\Delta^0(t)$ is the right part of (6.25). Then differentiating $\Delta^0(t)$ with respect to t results in the equation (6.10) subject to the initial condition $\Delta^0(t_m)$. It is obvious that the solution (6.10), (6.26) are coicident, i.e., $\gamma^0(t)=\Delta^0(t)$. Coicidence $\gamma^0(t)$ with the low bound (6.25) for $\Delta(t)$ proves an optimal decoding (6.5), and (6.7), (6.9), (6.10) follow as a result of substitution (6.5) in (6.2), (6.13), (6.14). The validity of this result for arbitrary time interval $\tau \leqslant t_m \leqslant t < t_{m+1}$ is derived with respect to induction, taking into account Remark 4.

REMARK 5. According to (6.5) and (6.6) in the class $\mathcal{K}_l$ under the limitations (6.3) in the filtering problem, all energy $\{\tilde h(t);\tilde g(t_m)\}$ of the message $\{h(\cdot);g(\cdot)\}$ is concentrated with respect to the signal x_t in the current moment of time, since $H_1^0(t,z)=0$, $G_1^0(t_m,z)=0$. Thus, Theorem 5 provides the solution already at the time interval $[0,\tau]$, when the memory is absent, and Remark 4 losses its actuality.

REMARK 6. The proof of Theorem 5 indicates that under the energy limitations, different from (6.3) and allocating general energy of the message on the current x_t, x_{t_m} and past x_τ signal values, we obtain a different solution, when $H_1^0(t,z)\neq 0$, $G_1^0(t_m,z)\neq 0$. This problem is open for research.

Theorem 6. *Coding functionals in the class $\mathcal{K}_l$ of linear functionals* (6.4) *are optimal in the general class $\mathcal{K}$ nonlinear functionals.*

Proof. The idea of the proof is the following. Suppose $\Delta_0(t)$ is a decoding error, attained at $\{h(\cdot);g(\cdot)\}\in\mathcal{K}$. Since $\mathcal{K}_l\subset\mathcal{K}$, then $\Delta_0(t)\leqslant\Delta^0(t)$, where $\Delta^0(t)$ is defined by Theorem 5. Analogously to Theorem 16.5 in (Liptser and Shiryayev, 1977; 1978), proof by contradiction is carried out by means of proving the inequality $\Delta_0(t)\geqslant\Delta^0(t)$. Then the contradiction is excluded only by the condition that $\Delta_0(t)=\Delta^0(t)$.

Since under the conditions (4.1) $p(t,x)=\mathcal{N}\{x;a(t),D(t)\}$, then on arbitrary coding $\{h(\cdot);g(\cdot)\}\in\mathcal{K}$ with respect to Corollary 1 for $t_m\leqslant t<t_{m+1}$

$$\begin{aligned}I_t[x_t;z_0^t,\eta_0^m] = I_{t_m}[\cdot]+\frac{1}{2}\Bigg(\int_{t_m}^{t} R^{-1}(\sigma)M\left\{\Big[\overline{h(\tau,z|x_\sigma)}-\overline{h(\tau,z)}\Big]^2\right\}\mathrm{d}\,\sigma\\ -\int_{t_m}^{t} Q(\sigma)\left[M\left\{J[x_\sigma]\right\}-D^{-1}(\sigma)\right]\mathrm{d}\,\sigma\Bigg),\end{aligned} \tag{6.27}$$

where $J[x_t] = M\{[\partial \ln[p_t(x_t)]/\partial x_t]^2|z_0^t, \eta_0^m\}$ is the Fisher conditional information amount (Liptser, 1974). Since $\overline{h(t,z)} = M\{\overline{h(\tau,z|x_t)}|z_0^t,\eta_0^m\}$, then $M\{[\overline{h(\tau,z|x_t)} - \overline{h(t,z)}]^2\} = M\{M\{[\cdot]^2|z_0^t, \eta_0^m\}\} = M\{M\{\overline{h(\tau,z|x_t)}^2 + \overline{h(t,z)}^2 - 2\overline{h(\tau,z|x_t)} \cdot \overline{h(t,z)}|z_0^t, \eta_0^m\}\} = M\{\overline{h(\tau,z|x_t)}^2 - \overline{h(t,z)}^2\} \leqslant M\{\overline{h(\tau,z|x_t)}^2\}$. According to Jensen inequality, taking into account (6.3) $M\{\overline{h(\tau,z|x_t)}^2\} = M\{[M\{h(\cdot)|x_t,z_0^t,\eta_0^m\}]^2\} \leqslant M\{M\{h^2(\cdot)|x_t, z_0^t, \eta_0^m\}\} = M\{h^2(\cdot)\} \leqslant \tilde{h}(t)$. Thus $M\{[\overline{h(\tau,z|x_t)} - \overline{h(t,z)}]^2\} \leqslant \tilde{h}(t)$ and using Fisher inequality $M\{J[x_t]\} \geqslant \Delta^{-1}(t)$ (Liptser, 1974) from (6.27) it follows that

$$I_t[x_t; z_0^t, \eta_0^m] \leqslant I_{t_m}[\cdot] + \frac{1}{2}\Bigg(\int\limits_{t_m}^{t} R^{-1}(\sigma)\tilde{h}(\sigma)\mathrm{d}\,\sigma - \int\limits_{t_m}^{t} Q(\sigma)\left[\Delta^{-1}(\sigma) - D^{-1}(\sigma)\right]\mathrm{d}\,\sigma\Bigg). \tag{6.28}$$

Suppose that the transmission took place in accordance with the coding $\{h^0(\cdot); g^0(\cdot)\}$ in the form (6.5), (6.6). Since for this case $p_t(x) = \mathcal{N}\{x; \mu^0(t), \Delta^0(t)\}$ (Liptser and Shiryayev, 1977; 1978), then from (6.27), taking into account (3.5), (3.26), (6.5), (6.6)

$$I_t^0[\cdot] = I_{t_m}^0[\cdot] + \frac{1}{2}\Bigg(\int\limits_{t_m}^{t} R^{-1}(\sigma)\tilde{h}(\sigma)\mathrm{d}\,\sigma - \int\limits_{t_m}^{t} Q(\sigma)\left[(\Delta^0(\sigma))^{-1} - D^{-1}(\sigma)\right]\mathrm{d}\,\sigma\Bigg). \tag{6.29}$$

Since $[\Delta^{-1} - D^{-1}] = [\Delta^{-1} - (\Delta^0)^{-1}] + [(\Delta^0)^{-1} - D^{-1}]$, then by the transmission on the interval $t \in [0, t_m]$ in accordance with the coding (6.5), (6.6) from (6.28), (6.29) it follows that

$$I_t \leqslant I_t^0[\cdot] - \frac{1}{2}\int\limits_{t_m}^{t} Q(\sigma)\left[\Delta^{-1}(\sigma) - (\Delta^0(\sigma))^{-1}\right]\mathrm{d}\,\sigma. \tag{6.30}$$

According to (Liptser, 1974; Liptser and Shiryayev, 1977; 1978) (Ihara inequality);

$$\Delta(t) \geqslant D(t)\exp\{-2I_t[\cdot]\}, \tag{6.31}$$

then from (6.30), (6.31)

$$\Delta(t) \geqslant D(t)\exp\left\{-2I_t^0[\cdot]\right\}\exp\Bigg\{\int\limits_{t_m}^{t} Q(\sigma)\left[\Delta^{-1}(\sigma) - (\Delta^0(\sigma))^{-1}\right]\mathrm{d}\,\sigma\Bigg\}. \tag{6.32}$$

Since $\mathcal{K}_l \subset \mathcal{K}$, then $\Delta_0(t) \leqslant \Delta^0(t)$, i.e., $\Delta_0^{-1}(t) \geqslant (\Delta^0(t))^{-1}$. From (3.22) given $p(t,x) = \mathcal{N}\{x;\, a(t), D(t)\}$, $p_t(x) = \mathcal{N}\{x; \mu^0(t), \Delta^0(t)\}$ it follows $I_t^0[\cdot] = (1/2)\ln[D(t)/\Delta^0(t)]$. Thus (6.32) given $\Delta(t) = \Delta_0(t)$ result in the required contradiction $\Delta_0(t) \geqslant \Delta^0(t)$. The Theorem proof is concluded by the derivation of the contradictory inequality $\Delta_0(t_m) \geqslant \Delta^0(t_m)$ in the assumption that on the interval $t \in [0, t_m)$ the transmission took place in accordance with the coding $\{h^0(\cdot); g^0(\cdot)\}$ in the form (6.5), (6.6). From (6.31), taking into account (3.24) $\Delta(t_m) \geqslant D(t_m)\exp\{-2I_{t_m-0}^0[\cdot]\}\exp\{-2\Delta I_{t_m}[\cdot]\}$. As given $\{h(\cdot); g(\cdot)\} = \{h^0(\cdot); g^0(\cdot)\}$, $p_{t_m-0}(x) = \mathcal{N}\{x; \mu^0(t_m - 0), \Delta^0(t_m - 0)\}$ (Liptser and Shiryayev, 1977; 1978), then $I_{t_m-0}^0[\cdot] = (1/2)\ln[D(t_m)/\Delta^0(t_m - 0)]$, and consequently $\Delta(t_m) \geqslant \Delta^0(t_m - 0)$ $\exp\{-2\Delta I_{t_m}[\cdot]\}$. Multiplication of the last inequality by $V(t_m)[V(t_m) + \tilde{g}(t_m)]^{-1}$ yields, taking into account (6.12)

$$\Delta(t_m) \geqslant \Delta^0(t_m)V^{-1}(t_m)\left[V(t_m) + \tilde{g}(t_m)\right]\exp\left\{-2\Delta I_{t_m}[\cdot]\right\}. \qquad (6.33)$$

From (3.6), (3.7), (3.25), (3.27) using Jensen inequality and taking into account that $\exp\{-y\} \leqslant (1+y)^{-1}$, $\ln\{y\} \leqslant y - 1$, it follows that

$$\Delta I_{t_m}[\cdot] \leqslant (1/2)\ln\left[1 + (\tilde{g}(t_m)/V(t_m))\right]. \qquad (6.34)$$

Use of (6.34) in (6.33) given $\Delta(t_m) = \Delta_0(t_m)$ results in the required contradiction $\Delta_0(t_m) \geqslant \Delta^0(t_m)$. The validity of the proved result for the arbitrary time interval $\tau \leqslant t_m \leqslant t < t_{m+1}$ follows by induction, taking into account Remark 5.

Theorem 7. *Suppose $I_t^0[x_t; (z^0)_0^t, (\eta^0)_0^m]$ is the information amount, attained on the coding functionals* (6.5), (6.6). *The property takes place*

$$I_t^0\left[x_t; (z^0)_0^t, (\eta^0)_0^m\right] = \sup I_t[x_t; z_0^t, \eta_0^m], \qquad (6.35)$$

where the supremum is taken for all $\{h(\cdot); g(\cdot)\} \in \mathcal{K} = \{\mathcal{H}; \mathcal{G}\}$ and

$$I_t^0\left[x_t; (z^0)_0^t, (\eta^0)_0^m\right] = (1/2)\sum_{t_i \leqslant t}\ln\left[1 + (\tilde{g}(t_i)/V(t_i))\right]$$
$$+(1/2)\left[\int_0^t \left(R^{-1}(\sigma)\tilde{h}(\sigma) - Q(\sigma)\left[\left(\Delta^0(\sigma)\right)^{-1} - D^{-1}(\sigma)\right]\right)\mathrm{d}\,\sigma\right]. \qquad (6.36)$$

Proof. From (6.27), taking into account (3.24), (3.25), for $\tau \leqslant t_i \leqslant t_m \leqslant t$ it follows

$$I_t\left[x_t; z_0^t, \eta_0^m\right] = (1/2)\sum_{\tau \leqslant t_i \leqslant t} M\left\{\ln\left[C(\eta(t_i), z|x_{t_i})/C(\eta(t_i), z)\right]\right\}$$
$$+(1/2)\left(\int_\tau^t R^{-1}(\sigma)M\left\{\left[\overline{h(\tau, z|x_\sigma)} - \overline{h(\tau, z)}\right]^2\right\}\mathrm{d}\,\sigma\right.$$

$$-\int_{\tau}^{t} Q(\sigma)\left[M\left\{J[x_\sigma]\right\}-D^{-1}(\sigma)\right]\mathrm{d}\,\sigma\biggr). \tag{6.37}$$

Use of (6.28), (6.34) in (6.37) yields that $I_t[x_t; z_0^t, \eta_0^m] \leqslant I_t^0[\cdot]$, where $I_t[\cdot]$ is defined by the right part of the formula (6.36). Use of (3.25), (4.34), (6.5), (6.6), (6.12), (6.16) in (6.37) yields that the upper bound $I_t^0[\cdot]$ for $I_t[\cdot]$ is attained on the coding functionals $h^0(\cdot)$ and $g^0(\cdot)$ in the form of (6.5), (6.6). Consequently (6.35) has been proved for $\tau \leqslant t_m \leqslant t$. The validity of the result for the initial time interval $[0,\tau]$ also follows taking into account Remark 5.

REMARK 7. It is obvious that for $I_t^0[\cdot]$ is equivalent to (6.37) the differential-recurrence presentation: $I_t^0[\cdot]$ on the intervals $t_m \leqslant t < t_{m+1}$ is defined by the equation

$$\mathrm{d}\, I_t^0\left[x_t; (z^0)_0^t, (\eta^0)_0^m\right]/\mathrm{d}\, t = (1/2)\Bigl(R^{-1}(t)\tilde{h}(t) \\ -Q(t)\left[(\Delta^0(t))^{-1} - D^{-1}(t)\right]\Bigr) \tag{6.38}$$

with the initial condition $I_{t_m}^0[\cdot] = I_{t_m-0}^0[x_{t_m};\ (z^0)_0^{t_m},\ (\eta^0)_0^{m-1}] + \Delta I_{t_m}^0[\cdot]$, where

$$\Delta I_{t_m}^0[x_{t_m}; (z^0)_0^{t_m}, \eta^0(t_m)] = (1/2)\ln\left[1 + (\tilde{g}(t_m)/V(t_m))\right]. \tag{6.39}$$

Since capacity $C[0,T]$ of the transmission channel is defined in the form of $C[0,T] = \sup\{(1/T)I_T[\cdot]\}$ (Gallager, 1968; Liptser and Shiryayev, 1977; 1978) then according to Theorem 7 for the class of signals (6.1) by continuous-dicrete way of transmission (6.2), (6.3) the coding functionals (6.5), (6.6) provide the transmission of a maximum possible information amount.

7. Conclusion

1. As it follows from the considered particular problem of paragraph 5 presence of memory may both increase and decrease information efficiency of observations.
2. Obtained theoretical result can be applied for information efficiency analysis of the continuous-discrete time observation system of stochastic objects, and also for solution of information theory standard problems in the considered class of the processes x_t, z_t, $\eta(t_m)$ as an optimal of stochastic signals transmission on condituous-discrete memory channels and for research of the capacity of these channels.

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 nonlinear filtering. *Avtom. 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.

Arimoto, S. (1971). Information-theoretical considerations on estimation problem. *Inform. Control*, **19**(2), 181–194.

Dobrushin, R.L. (1963). General formulation of Shannon's main theorem in information theory. *A.M.S. Transl.*, ser. 2, **33**, 323–438.

Duncan, T.E. (1971). Mutual information for stochastic differential equations. *Inform. Control*, **19**(3), 265–271.

Dyomin, N.S., and V.I. Korotkevich (1983). Amount of information in problems of filtering the components of Markov processes. *Autom. Remote Control*, **44**(7), 899–907.

Dyomin, N.S., and V.I. Korotkevich (1987). Equations for Shannon information on transmission of Markov diffusion signals through channels with memory. *Problems of Inform. Trans.*, **23**(1), 11–21.

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.

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.

Gallager, R.G. (1968). *Information Theory and Reliable Communication.* Wiley, New–York.

Gantmakher, F.R. (1988) *Theory of Matrices.* Nauka, Moscow (in Russian).

Ihara, S. (1990). Capacity of mismatched gaussian channels with and without feedback. *Probab. Th. Relat. Fields*, **84**(4), 453–471.

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

Kalman, R.E., and R. Bucy (1961). New results in linear filtering and prediction theory. *Trans. ASME. J.Basic Eng.*, Ser. D, **83**, 95–108.

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

Kolmogorov, A.N. (1963). Theory of transmission of information. *A.M.S. Transl.*, ser. 2, **33**, 291–321.

Liptser, R.Sh. (1974). Optimal coding and decoding for transmission og Gaussian Markov signal over noiseless feedback channel. *Probl. Peredachi Inf.*, **10**(4), 3–15 (in Russian).

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

Meditch, J.S. (1969). *Stochastic Optimal Linear Estimation and Control.* Mc.Graw Hill, New York.

Ocone, D., and E. Pardoux (1989). A generelized Ito–Ventzel formula. *Ann. Inst. Henri Poincare*, **25**(1), 39–71.

Rozovsky, B.L. (1973). On the Ito–Ventzel formula. *Vest. MSU*, ser. Mathem., Mech., **28**(1), 26–32 (in Russian).

Shannon, C.E., and W. Weaver (1949). *The Mathematical Theory of Communication.* Univ. Illinois Press, Urbana, Illinois.

Stratonovitch, R.L. (1975). *Theory of Information.* Sov. Radio, Moscow (in Russian).

Tomita, Y., S. Ohmatsu and T. Soeda (1976). An application of the information theory to estimation problems. *Inform. Control*, **32**(2), 101–111.

N. Dyomin graduated from Department of Radio and Physics, Tomsk State University, Russia, in 1967. He received the PhD degree in technical cybernetics and information theory from Tomsk State University in 1987 and the Dr.Sci. degree in application of mathematical methods in science investigations from the same university in 1990. He is an full professor in Departmant of Applied Mathematics and Cybernetics of Tomsk State University. Present research interests: filtering, interpolation, extrapolation and recognition of stochastic processes; transmission of stochastic signals along continuous-discrete observation memory channels; detection of anomalous noises; mathematical theory of financial analysis and security trading.

S. Rozhkova graduated from Departmant of Applied Mathematics and Cybernetics of Tomsk State University, Russia, in 1993. She received the PhD degree in application of mathematical methods in science investigations from Tomsk State University in 1996. She is an associated professor in Departmant of Natural Sciences and Mathematics of Tomsk Politechnical University. Present research interests: filtering, interpolation, extrapolation and recognition of stochastic processes; transmission of stochastic signals along continuous-discrete observation memory channels; detection of anomalous noises.

I. Safronova graduated from Departmant of Applied Mathematics and Cybernetics of Tomsk State University, Russia, in 2000. She is a post-graduate in Departmant of Applied Mathematics and Cybernetics of Tomsk State University. Present research interests: filtering, interpolation, extrapolation of stochastic processes; transmission of stochastic signals along continuous-discrete observation memory channels.

Informacijos kiekio radimas bendrai stochastinių procesų filtracijos ir apibendrintos interpoliacijos problemai atžvilgiu tolydžios ir diskrečios atminties stebėjimų

Nikolas DYOMIN, Irina SAFRONOVA, Svetlana ROZHKOVA

Darbe nagrinėjami bendros stochastinių procesų filtracijos ir apibendrintos interpoliacijos informaciniai aspektai, kai yra stebimos jų kompomentės tolydžiame arba diskrečiame laike. Rastos Šenono informacijos kiekio evoliucijos pereinamybės. Bendri rezultatai yra taikomi informacijos kanalų efektyvumui ir stochastinių signalų perdavimo optimalumui tirti.