ISSN 0005-1179, Automation and Remote Control, 2016, Vol. 77, No. 2, pp. 242–260. © Pleiades Publishing, Ltd., 2016.
Original Russian Text © M.V. Basin, 2016, published in Avtomatika i Telemekhanika, 2016, No. 2, pp. 69–93.

TOPICAL ISSUE

Root-Mean-Square Filtering of the State of Polynomial Stochastic Systems with Multiplicative Noise

M. V. Basin

Autonomous University of Nuevo Leon, Mexico
St. Petersburg State University of Information Technologies, Mechanics, and Optics (ITMO), St. Petersburg, Russia
e-mail: mbasin2007@gmail.com

Received March 30, 2015

Abstract—Some results obtained by the present author in the field of designing the finite-dimensional root-mean-square filters for stochastic systems with polynomial equations of state and multiplicative noise from the linear observations were overviewed. A procedure to derive the finite-dimensional system of approximate filtering equations for a polynomial arbitrary-order equation of state was presented. The closed system of filtering equations for the root-mean-square estimate and covariance matrix error was deduced explicitly for special cases of linear and quadratic coefficients of drift and diffusion in the equation of state. For linear stochastic systems with unknown parameters, the problem of joint root-mean-square state filtering and identification of the parameters from linear observations was considered in the Appendix.

DOI: 10.1134/S000511791602003X

1. INTRODUCTION

In the general case, the optimal solution of the problem of filtering for the stochastic systems obeying the nonlinear equations of state and observation with the Gaussian white noise is given by the Stratonovich–Kushner equations [1, 2]. However, some examples are known of the nonlinear systems where the Stratonovich–Kushner equations can be reduced to a finite-dimensional closed system of equations for the conditional moments of lower order. The most popular result represented by the Kalman–Bucy filter [3] was obtained for the case of linear equations of state and observation. The Kalman–Bucy equations make up a closed system of filtering equations with respect to the two lower conditional moments, expectation, and covariance matrix. Nevertheless, the finite-dimensional system of equations of nonlinear filtering can be obtained in some other cases, provided that the state vector can assume only a finite number of permissible states [4] or the observation equation is linear and the drift vector f in the equation of state satisfies the Riccati equation $\frac{df}{dx} = f^2 = x^2$ [5]. The reader is referred to [6] for complete classification of the "general situation" cases,—which means that there are no special assumptions about the structure of the equations of state and observation and the initial conditions,—where the optimal nonlinear finite-dimensional filter exists. In fact, the two latter references consider the problem of filtering for some polynomial systems. Apart from the aforementioned results, the finite-dimensional filters for some classes of systems with invertible measurement matrix, polynomial equations of state, and Gaussian initial conditions from linear observations were suggested in [7–10].

The present paper overviews some results established by the author in the field of designing the finite-dimensional root-mean-square (rms) filters for the stochastic systems with polynomial equations of state and multiplicative noise from the linear observations with an arbitrary—not necessarily invertible—observation matrices, thus generalizing the findings of [7–9]. Yet, in distinction to [7–9], it removes the requirement of invertibility of the observation matrix. Additionally, obtained was a procedure to deduce approximate rms equations of filtering for an arbitrary-order polynomial equation of state.

The aforementioned procedure for derivation of the approximate rms equations of filtering for the polynomial stochastic systems with multiplicative noise was later used to solve the problem of optimal control in the polynomial stochastic systems with unknown parameters [11], problem of filtering from observations with polynomial nonlinearities [12], problems of filtering and control for the polynomial stochastic systems in the sliding modes [13, 14], as well as the problems of filtering and control for the polynomial stochastic systems with Poisson noise [15, 16]. Among other publications devoted to determining the general approximate solution of the problem of nonlinear filtering, the references [17–19] deserve special mentioning (see also their bibliographies).

The problem of filtering for the polynomial state of a system with multiplicative noise from the linear observations with arbitrary measurement matrix is formulated in Section 2. The following Section 3 derives relations for the Ito differentials of the rms estimate and the error covariance matrix. The observation equation is transformed so as to reduce the original problem of filtering to the problem with invertible observation matrix that was solved before. Described here is a procedure for determining the closed system of approximate rms filtering equations for the arbitrary-order polynomial equations of state from the linear observations. The closed system is deduced explicitly for special cases of linear and quadratic coefficients of drift and diffusion in the equation of state. An example given in Section 4 illustrates efficiency of the obtained filter. The following Section 5 formulates the problem of joint rms filtering of state and identification of the parameters for the linear stochastic systems with unknown parameters from the linear observations. Section 6 reduces the posed problem to the considered problem of rms filtering for the extended state vector involving unknown parameters as additional components of the state vector. The equations of joint filtering and identification are sought on the basis of the equations of filtering from Section 3. Efficiency of the resulting filter-identifier is illustrated by way of an example in Section 7.

2. FORMULATION OF THE PROBLEM OF FILTERING THE STATE OF POLYNOMIAL STOCHASTIC SYSTEMS FROM LINEAR OBSERVATIONS

Let (Ω, F, P) be a full probabilistic space with an increasing right-continuous family of σ-algebras F_t, $t \geq t_0$ where given are independent standard Wiener processes $(W_1(t), F_t, t \geq t_0)$ and $(W_2(t), F_t, t \geq t_0)$. Let also the F_t-measurable random process $(x(t), y(t))$ be described by the nonlinear polynomial differential equation of state of the system

$$dx(t) = f(x,t)dt + g(x,t)dW_1(t), \quad x(t_0) = x_0 \tag{2.1}$$

and the linear differential equation of the observation process

$$dy(t) = (A_0(t) + A(t)x(t))dt + B(t)dW_2(t), \tag{2.2}$$

where $x(t) \in R^n$ is the state vector and $y(t) \in R^l$ is the vector of linear observations. The Gaussian vector of initial state $x_0 \in R^n$ and the standard Wiener processes $W_1(t) \in R^p$ and $W_2(t) \in R^q$ are assumed to be independent. It is assumed that $B(t)B^{\mathrm{T}}(t)$ is a positive definite matrix, and, consequently, $l \leq q$. We also notice that, in distinction to the results presented in [7–9], it is not required here that $A(t) \in R^{l\times n}$ be an invertible or even square matrix. All coefficients in (2.1) and (2.2) are determinate functions of corresponding sizes. It is assumed that here and in what follows meant are the stochastic differential Ito equations.

The nonlinear functions $f(x,t)$ and $g(x,t)$ represent polynomials of n variables, the components of the state vector $x(t) \in R^n$ with time-dependent coefficients. Since $x(t) \in R^n$ is a vector, the polynomial of size $n > 1$ needs a special definition. According to [9], the polynomial of degree p of the vector $x(t) \in R^n$ is understood as a p-linear form of n components of the vector $x(t)$

$$f(x,t) = a_0(t) + a_1(t)x + a_2(t)xx^{\mathrm{T}} + \cdots + a_p(t)x \ldots_{p \text{ times}} \ldots x, \tag{2.3}$$

where $a_0(t)$ is the vector of size n, a_1 is an $n \times n$ matrix, a_2 is a 3D $n \times n \times n$ tensor, a_p is a $(p+1)$D $n \times \ldots_{(p+1) \text{ time}} \ldots \times n$ tensor, and $x \times \ldots_{p \text{ times}} \ldots \times x$ is a pD $n \times \ldots_{p \text{ times}} \ldots \times n$ tensor obtained by the p-fold spatial multiplication of the vector $x(t)$ by itself. Such polynomial is also representable as the sum

$$f_k(x,t) = a_{0\,k}(t) + \sum_i a_{1\,ki}(t)x_i(t) + \sum_{ij} a_{2\,kij}(t)x_i(t)x_j(t) + \ldots$$
$$+ \sum_{i_1 \ldots i_p} a_{p\,ki_1\ldots i_p}(t)x_{i_1}(t)\ldots x_{i_p}(t), \quad k,i,j,i_1,\ldots,i_p = 1,\ldots,n.$$

The tensor operations are executed according to [20].

The problem of rms filtering lies in determining the optimal estimate $\hat{x}(t)$ of the state of system $x(t)$ from all observations from the initial time instant to the current time instant $Y(t) = \{y(s), t_0 \le s \le t\}$ which minimizes the rms criterion

$$J = E\left[(x(t) - \hat{x}(t))^{\mathrm{T}}(x(t) - \hat{x}(t))|F_t^Y\right]$$

at each instant t. Here, $E[\xi(t)|F_t^Y]$ denotes the conditional expectation of the random process $\xi(t) = (x(t) - \hat{x}(t))^{\mathrm{T}}(x(t) - \hat{x}(t))$ with respect to the α-algebra F_t^Y generated by the process of observation $Y(t)$ over the interval $[t_0, t]$. This optimal estimate is known [21] to be defined by the conditional expectation $\hat{x}(t) = m(t) = E(x(t)|F_t^Y)$ of the state of system $x(t)$ relative to the α-algebra F_t^Y. The matrix function $P(t) = E[(x(t) - m(t))(x(t) - m(t))^{\mathrm{T}}|F_t^Y]$ is the covariance matrix of the estimation error.

The proposed solution of the formulated problem of filtering which relies on the formulas for the Ito differential conditional expectation $E(x(t)|F_t^Y)$ and the error covariance matrix $P(t)$ (see [21]) is given in the following section.

3. ROOT-MEAN-SQUARE FILTERING OF THE STATE OF POLYNOMIAL STOCHASTIC SYSTEMS FROM LINEAR OBSERVATIONS

The rms filtering equations can be derived using the Ito formula of the conditional differential expectation $m(t) = E(x(t)|F_t^Y)$ (see [21])

$$dm(t) = E\left(f(x,t)|F_t^Y\right)dt + E\left(x[\varphi_1(x) - E(\varphi_1(x)|F_t^Y)]^{\mathrm{T}}|F_t^Y\right)$$
$$\times \left(B(t)B^{\mathrm{T}}(t)\right)^{-1}\left(dy(t) - E(\varphi_1(x)|F_t^Y)dt\right),$$

where $\varphi_1(x)$ is the linear term in the observation equation equal to $\varphi_1(x,t) = A_0(t) + A(t)x(t)$. With regard for this expression, we establish the estimate equation

$$dm(t) = E\left(f(x,t)|F_t^Y\right)dt + E\left(x(t)[A(t)(x(t) - m(t))]^{\mathrm{T}}|F_t^Y\right) \tag{3.1}$$
$$\times \left(B(t)B^{\mathrm{T}}(t)\right)^{-1}(dy(t) - (A_0(t) + A(t)m(t))dt)$$
$$= E\left(f(x,t)|F_t^Y\right)dt + E\left(x(t)(x(t) - m(t))^{\mathrm{T}}|F_t^Y\right)A^{\mathrm{T}}(t)$$
$$\times \left(B(t)B^{\mathrm{T}}(t)\right)^{-1}(dy(t) - (A_0(t) + A(t)m(t))dt)$$
$$= E\left(f(x,t)|F_t^Y\right)dt + P(t)A^{\mathrm{T}}(t)\left(B(t)B^{\mathrm{T}}(t)\right)^{-1}(dy(t) - (A_0(t) + A(t)m(t))dt).$$

Equation (3.1) is completed by the initial condition $m(t_0) = E(x(t_0) \mid F_{t_0}^Y)$ calculated by processing the measurement at the initial time instant.

At generating the closed system of filtering equations, Eq. (3.1) should be completed by the equation for covariance of the error matrix $P(t)$ based on the formula for the Ito differential of the covariance matrix [21]:

$$\begin{aligned} dP(t) = &\Big(E\left((x(t)-m(t))f^{\mathrm{T}}(x,t)|F_t^Y\right) + E\left(f(x,t)(x(t)-m(t))^{\mathrm{T}}\right)|F_t^Y\Big) \\ &+ E\left(g(x,t)g^{\mathrm{T}}(x,t)|F_t^Y\right) - E\left(x(t)\left[\varphi_1(x) - E\left(\varphi_1(x)|F_t^Y\right)\right]^{\mathrm{T}}|F_t^Y\right) \\ &\times \left(B(t)B^{\mathrm{T}}(t)\right)^{-1} E\left(\left[\varphi_1(x) - E\left(\varphi_1(x)|F_t^Y\right)\right] x^{\mathrm{T}}(t)|F_t^Y\right) dt \\ &+ E\left((x(t)-m(t))(x(t)-m(t))\left[\varphi_1(x) - E\left(\varphi_1(x)|F_t^Y\right)\right]^{\mathrm{T}}|F_t^Y\right) \\ &\times \left(B(t)B^{\mathrm{T}}(t)\right)^{-1}\left(dy(t) - E\left(\varphi_1(x)|F_t^Y\right)dt\right), \end{aligned}$$

where the nonlinear polynomial $g(x,t)$ is the diffusion term in the state equation. The last term in this expression (under the sign of expectation) must be understood as the 3D tensor [20]. Substitution of the expression for φ_1, gives rise to the following formula:

$$\begin{aligned} dP(t) = &\Big(E\left((x(t)-m(t))f^{\mathrm{T}}(x,t)|F_t^Y\right) + E\left(f(x,t)(x(t)-m(t))^{\mathrm{T}}\right)|F_t^Y\Big) \\ &+ E\left(g(x,t)g^{\mathrm{T}}(x,t)|F_t^Y\right) - \Big(E\left(x(t)(x(t)-m(t))^{\mathrm{T}}|F_t^Y\right)A^{\mathrm{T}}(t) \\ &\times \left(B(t)B^{\mathrm{T}}(t)\right)^{-1} A(t)E\left((x(t)-m(t))x^{\mathrm{T}}(t)\right)|F_t^Y\Big)\,dt \\ &+ E\left((x(t)-m(t))(x(t)-m(t))(A(t)x(t)m(t))^{\mathrm{T}}|F_t^Y\right) \\ &\times \left(B(t)B^{\mathrm{T}}(t)\right)^{-1}(dy(t) - (A_0(t) + A(t)m(t))dt). \end{aligned}$$

Using the formula of the covariance matrix

$$P(t) = E\left(x(t) - m(t))x^{\mathrm{T}}(t)|F_t^Y\right),$$

the last equation is representable as

$$\begin{aligned} dP(t) = &\Big(E\left((x(t)-m(t))f^{\mathrm{T}}(x,t)|F_t^Y\right) + E\left(f(x,t)(x(t)-m(t))^{\mathrm{T}}\right)|F_t^Y\Big) \\ &+ E\left(g(x,t)g^{\mathrm{T}}(x,t)|F_t^Y\right) - P(t)A^{\mathrm{T}}(t)\left(B(t)B^{\mathrm{T}}(t)\right)^{-1}A(t)P(t)dt \\ &+ E\left((x(t)-m(t))(x(t)-m(t))(x(t)-m(t))^{\mathrm{T}}|F_t^Y\right) \\ &\times A^{\mathrm{T}}(t)\left(B(t)B^{\mathrm{T}}(t)\right)^{-1}(dy(t) - (A_0(t) + A(t)m(t))dt). \end{aligned} \tag{3.2}$$

Equation (3.2) is completed by the initial condition

$$P(t_0) = E\left[(x(t_0) - m(t_0))(x(t_0) - m(t_0))^{\mathrm{T}}|F_{t_0}^Y\right]$$

which also is calculated as the result of processing the measurement at the initial time instant.

Equations (3.1) and (3.2) for the optimal estimate $m(t)$ and the covariance error matrix $P(t)$ make up an open system of filtering equations for the nonlinear state vector (2.1) from the linear observations (2.2). Openness implies that (3.1) and (3.2) includes the expectations of the nonlinear functions of x such as $E(f(x,t)|F_t^Y)$, $E((x(t)-m(t))f^{\mathrm{T}}(x,t)|F_t^Y)$ and $E(g(x,t)g^{\mathrm{T}}(x,t)|F_t^Y)$ which are still not expressed as the functions of the variables $m(t)$ and $P(t)$.

As was shown in [7–9], the closed system of filtering equations for the state vector of the polynomial system with multiplicative noise (2.1) can be established from the linear observations if the observation matrix $A(t)$ is invertible for all $t \geq t_0$. Since in (2.2) this requirement is removed for the observation matrix $A(t)$, the following transformations are done first.

We begin by noticing that it is always possible to assume that the matrix A is that of full rank l which is equal to the dimension of the independent linear observations. Otherwise, the dependent linear observation corresponding to the dependent linear rows of the matrix A must be disregarded. At that, by adding and renumbering the Wiener processes in each observation Eq. (2.2), the number of Wiener processes in the observation equations can also be reduced to l. Consequently, B can always be assumed to be an $l \times l$ matrix such that $B(t)B^{\mathrm{T}}(t)$ is a positive definite matrix (see Section 2). The matrices $\overline{A}(t)$ and $\overline{B}(t)$ are determined as follows. The matrix $\overline{A}(t) \in R^{n\times n}$ is obtained from the matrix $A(t) \in R^{l\times n}$ by adding $n-l$ independent linear rows so as to make the resulting matrix $\overline{A}(t)$ invertible. The matrix $\overline{B}(t) \in R^{n\times n}$ is obtained from the matrix $B(t) \in R^{l\times l}$ by placing $B(t)$ at the upper left corner of $\overline{B}(t)$, equating the remaining $n-m$ diagonal elements of $\overline{B}(t)$ to infinity, and zeroing all other elements of $\overline{B}(t)$ outside the main diagonal or the submatrix $B(t)$. Stated differently, $\overline{B}(t) = \mathrm{diag}[B(t), \beta I_{(n-l)\times(n-l)}]$, where $\beta = \infty$ and $I_{(n-l)\times(n-l)}$ is the identity $(n-l)\times(n-l)$ matrix. Therefore, the new observation equation is representable as

$$\overline{y}(t) = (\overline{A}_0(t) + \overline{A}(t)x(t))dt + \overline{B}(t)dW_2(t), \tag{3.3}$$

where $\overline{y}(t) \in R^n$, $\overline{A}_0(t) = [A_0^{\mathrm{T}}(t), 0_{n-l}]^{\mathrm{T}} \in R^n$, and 0_{n-l} is the $n-l$ zero vector.

The key point of the introduced transformation is the fact that the new process of observation $\overline{y}(t)$ physically is equivalent to the original process of observation $y(t)$ because the $n-l$ last dummy components of the process $\overline{y}(t)$ consist of pure noise because the infinite intensities of the white Gaussian noise in the corresponding $n-l$ equations and the m first components of the process $\overline{y}(t)$ coincide with $y(t)$. Additionally, the entire observation matrix $\overline{A}(t)$ is invertible and the matrix $(\overline{B}(t)\overline{B}^{\mathrm{T}}(t))^{-1} \in R^{n\times n}$ exists and represents the $n \times n$ square matrix whose upper left corner is occupied by the submatrix $(B(t)B^{\mathrm{T}}(t))^{-1} \in R^{l\times l}$ and the rest of the elements are zero.

Taking into consideration the new observation Eq. (3.3), the filtering Eqs. (3.1) and (3.2) take on form

$$dm(t) = E\left(f(x,t)|F_t^Y\right)dt + P(t)\overline{A}^{\mathrm{T}}(t)\left(\overline{B}(t)\overline{B}^{\mathrm{T}}(t)\right)^{-1} \times (d\overline{y}(t) - (\overline{A}_0(t) + \overline{A}(t)m(t))dt), \tag{3.4}$$

$$\begin{aligned} dP(t) = &\left(E\left((x(t)-m(t))f^{\mathrm{T}}(x,t)|F_t^Y\right) + E\left(f(x,t)(x(t)m(t))^{\mathrm{T}}\right)|F_t^Y\right) \\ &+ E\left(g(x,t)g^{\mathrm{T}}(x,t)|F_t^Y\right) - P(t)\overline{A}^{\mathrm{T}}(t)\left(\overline{B}(t)\overline{B}^{\mathrm{T}}(t)\right)^{-1}\overline{A}(t)P(t)dt \\ &+ E\left((x(t)-m(t))(x(t)-m(t))(x(t)-m(t))^{\mathrm{T}}|F_t^Y\right) \\ &\times \overline{A}^{\mathrm{T}}(t)\left(\overline{B}(t)\overline{B}^{\mathrm{T}}(t)\right)^{-1}(d\overline{y}(t) - (\overline{A}_0(t) + \overline{A}(t)m(t))dt), \end{aligned} \tag{3.5}$$

with the initial conditions

$$m(t_0) = E\left(x(t_0)|F_{t_0}^Y\right) \text{ and } P(t_0) = E\left[(x(t_0)-m(t_0))(x(t_0)-m(t_0))^{\mathrm{T}}|F_{t_0}^Y\right].$$

Since the new observation matrix $\overline{A}(t)$ is invertible for any $t \geq t_0$, the density of distribution of the random vector $x(t) - m(t)$ may be assumed to be conditionally Gaussian relative to the new observation process $\overline{y}(t)$, and, consequently, to the original process of observation $y(t)$ for any $t \geq t_0$ (see [7–9]). Therefore, the following considerations set forth in [7–9] are applicable to the filtering Eqs. (3.4) and (3.5).

First, since it was assumed that the random vector $x(t) - m(t)$ has conditionally Gaussian distribution density in the above sense, in the last addend of Eq. (3.5) the conditional central moment with respect to the observations $E((x(t) - m(t))(x(t) - m(t))(x(t) - m(t))^{\mathrm{T}}|F_t^Y)$ is zero. Therefore, the last addend in Eq. (3.5) is zero and the equation of the covariance matrix is given by

$$dP(t) = \Big(E\left((x(t) - m(t))f^{\mathrm{T}}(x,t)|F_t^Y\right) + E\left(f(x,t)(x(t)m(t))^{\mathrm{T}}|F_t^Y\right) \\ + E\left(g(x,t)g^{\mathrm{T}}(x,t)F_t^Y\right) - P(t)\overline{A}^{\mathrm{T}}(t)\left(\overline{B}(t)\overline{B}^{\mathrm{T}}(t)\right)^{-1}\overline{A}(t)P(t)\Big)\,dt, \tag{3.6}$$

with the initial condition $P(t_0) = E[(x(t_0) - m(t_0))(x(t_0) - m(t_0))^{\mathrm{T}}|F_{t_0}^Y]$.

Second, since $f(x,t)$ and $g(x,t)$ are polynomial functions of x with time-dependent coefficients, expressions comprising the terms $E(f(x,t)|F_t^Y)$ in (3.4) and $E((x(t) - m(t))f^{\mathrm{T}}(x,t)|F_t^Y)$, $E(g(x,t)g^{\mathrm{T}}(x,t)|F_t^Y)$ in (3.6) also will comprise only the polynomial functions of x. Using the following characteristic of the scalar Gaussian centered random variable $x(t) - m(t)$ which states that all of its odd conditional moments $m_1 = E[(x(t) - m(t))|Y(t)]$, $m_3 = E[(x(t) - m(t))^3|Y(t)]$, $m_5 = E[(x(t) - m(t))^5|Y(t)], \ldots$ are zero and the even conditional moments $m_2 = E[(x(t) - m(t))^2|Y(t)]$, $m_4 = E[(x(t) - m(t))^4|Y(t)], \ldots$ are representable as the functions $P(t)$, these polynomial terms can be then represented as the functions $m(t)$ and $P(t)$. For example, $m_2 = P$, $m_4 = 3P^2$, $m_6 = 15P^3$, …and so on. We notice that, as was shown in Section 2 [20], in the case of a random vector variable these equalities must be established for each component of the corresponding tensor by representing the operation as a sum. By expressing in this way all polynomial terms in (3.4) and (3.6) that result after representing $E(f(x,t)|F_t^Y)$, $E((x(t) - m(t))f^{\mathrm{T}}(x,t)|F_t^Y)$ and $E(g(x,t)g^{\mathrm{T}}(x,t)|F_t^Y)$ as functions of $m(t)$ and $P(t)$, we arrive to the closed form of the approximate filtering equations. The corresponding representations for $E(f(x,t)|F_t^Y)$, $E((x(t) - m(t))(f(x,t))^{\mathrm{T}}|F_t^Y)$ and $E(g(x,t)g^{\mathrm{T}}(x,t)|F_t^Y)$ were established in [7–9] for some polynomial functions $f(x,t)$ and $g(x,t)$.

Finally, by taking into account the definitions of the matrices $\overline{A}(t)$ and $\overline{B}(t)$ and the observation process $\overline{y}(t)$, the filtering Eqs. (3.4) and (3.6) can be again represented in terms of the original observation Eq. (2.2) by using $y(t)$, $A(t)$, and $B(t)$

$$dm(t) = E\left(f(x,t)|F_t^Y\right)dt + P(t)A^{\mathrm{T}}(t)\left(B(t)B^{\mathrm{T}}(t)\right)^{-1} \\ \times(dy(t) - (A_0(t) + A(t)m(t))dt), \tag{3.7}$$

$$dP(t) = \Big(E\left((x(t) - m(t))f^{\mathrm{T}}(x,t)|F_t^Y\right) + E\left(f(x,t)(x(t) - m(t))^{\mathrm{T}}\right)|F_t^Y\Big) \\ + E\left(g(x,t)g^{\mathrm{T}}(x,t)|F_t^Y\right) - P(t)A^{\mathrm{T}}(t)\left(B(t)B^{\mathrm{T}}(t)\right)^{-1}A(t)P(t))dt, \tag{3.8}$$

with the initial conditions

$$m(t_0) = E\left(x(t_0)|F_{t_0}^Y\right) \text{ and } P(t_0) = E\left[(x(t_0) - m(t_0))(x(t_0) - m(t_0))^{\mathrm{T}}|F_{t_0}^Y\right].$$

In the following sections, the closed forms of the approximate filtering equation will be obtained from Eqs. (3.7) and (3.8) for the linear and quadratic functions $f(x,t)$ and $g(x,t)$ in Eq. (2.1). One can prove that the same procedure leads to a closed system of approximate filtering equations for any polynomial functions $f(x,t)$ and $g(x,t)$ in (2.1).

3.1. Root-Mean-Square Filtering for Stochastic Systems with Linear Drift and Diffusion

In a special case of $f(x,t)=a_0(t)+a_1(t)x(t)$ and $g(x,t)=b_0(t)+b_1(t)x(t)$, the linear functions where b_1 is understood as a 3D $n\times n\times p$ tensor, $E(f(x,t)|F_t^Y)$, $E((x(t)-m(t))f^{\mathrm{T}}(x,t)|F_t^Y)$, and $E(g(x,t)g^{\mathrm{T}}(x,t)|F_t^Y)$ are represented as the following functions $m(t)$ and $P(t)$:

$$E\left(f(x,t)|F_t^Y\right)=a_0(t)+a_1(t)m(t), \tag{3.9}$$

$$E\left(\left(f(x,t)(x(t)-m(t))^{\mathrm{T}}\right)|F_t^Y\right)+E\left((x(t)-m(t))(f(x,t))^{\mathrm{T}}|F_t^Y\right)=a_1(t)P(t)+P(t)a_1^{\mathrm{T}}(t), \tag{3.10}$$

$$E\left(g(x,t)g^{\mathrm{T}}(x,t)|F_t^Y\right)=b_0(t)b_0^{\mathrm{T}}(t)+b_0(t)(b_1(t)m(t))^{\mathrm{T}}$$
$$+(b_1(t)m(t))b_0^{\mathrm{T}}(t)+b_1(t)P(t)b_1^{\mathrm{T}}(t)+b_1(t)m(t)m^{\mathrm{T}}(t)b_1^{\mathrm{T}}(t), \tag{3.11}$$

where $b_1^{\mathrm{T}}(t)$ denotes the tensor obtained from $b_1(t)$ by transposing two of its extreme right indices.

Substitution of (3.9) into (3.7) and (3.10), (3.11) into (3.8) provides the following filtering equations for the rms estimate $m(t)$ and the covariance error matrix $P(t)$:

$$dm(t)=(a_0(t)+a_1(t)m(t))dt+P(t)A^{\mathrm{T}}(t)\left(B(t)B^{\mathrm{T}}(t)\right)^{-1}[dy(t)-(A_0(t)+A(t)m(t))dt], \tag{3.12}$$
$$m(t_0)=E\left(x(t_0)|F_{t_0}^Y\right),$$

$$dP(t)=\Big(a_1(t)P(t)+P(t)a_1^{\mathrm{T}}(t)+b_0(t)b_0^{\mathrm{T}}(t)+b_0(t)(b_1(t)m(t))^{\mathrm{T}}+(b_1(t)m(t))b_0^{\mathrm{T}}(t) \tag{3.13}$$
$$+b_1(t)P(t)b_1^{\mathrm{T}}(t)+b_1(t)m(t)m^{\mathrm{T}}(t)b_1^{\mathrm{T}}(t)\Big)dt-P(t)A^{\mathrm{T}}(t)\left(B(t)B^{\mathrm{T}}(t)\right)^{-1}A(t)P(t)dt,$$
$$P(t_0)=E\left[(x(t_0)-m(t_0))(x(t_0)-m(t_0))^{\mathrm{T}}|F_{t_0}^Y\right].$$

We note that to determine the filtering Eqs. (3.12) and (3.13) the observation matrix $A(t)$ even needs not to be invertible. Indeed, the only used polynomial equality $E(x(t)x^{\mathrm{T}}(t)|F_t^Y)=P(t)+m(t)m^{\mathrm{T}}(t)$ is valid for any random vector with, not necessarily Gaussian, finite second moment.

3.2. Root-Mean-Square Filtering for the Stochastic Systems with Quadratic Drift and Diffusion

Let

$$f(x,t)=a_0(t)+a_1(t)x+a_2(t)xx^{\mathrm{T}} \tag{3.14}$$

and

$$g(x,t)=b_0(t)+b_1(t)x+b_2(t)xx^{\mathrm{T}}, \tag{3.15}$$

where x and $a_0(t)$ are vectors of size n, $a_1(t)$ and $b_0(t)$ are, respectively, $n\times n$ and $n\times p$ matrices, $a_2(t)$ and $b_1(t)$ are 3D $n\times n\times n$ and $n\times n\times p$ tensors, and $b_2(t)$ is a 4D $n\times n\times n\times p$ tensor. In this case, $E(f(x,t)|F_t^Y)$, $E((x(t)-m(t))f^{\mathrm{T}}(x,t)|F_t^Y)$, and $E(g(x,t)g^{\mathrm{T}}(x,t)|F_t^Y)$ are representable as the functions $m(t)$ and $P(t)$ (see [8, 9]):

$$E\left(f(x,t)|F_t^Y\right)=a_0(t)+a_1(t)m(t)+a_2(t)m(t)m^{\mathrm{T}}(t)+a_2(t)P(t), \tag{3.16}$$

$$E\left(\left(f(x,t)(x(t)-m(t))^{\mathrm{T}}\right)|F_t^Y\right)+E\left((x(t)-m(t))f^{\mathrm{T}}(x,t)|F_t^Y\right) \tag{3.17}$$
$$=a_1(t)P(t)+P(t)a_1^{\mathrm{T}}(t)+2a_2(t)m(t)P(t)+2(a_2(t)m(t)P(t))^{\mathrm{T}},$$

$$E\left(g(x,t)g^{\mathrm{T}}(x,t)|F_t^Y\right)=b_0(t)b_0^{\mathrm{T}}(t)+b_0(t)(b_1(t)m(t))^{\mathrm{T}}+(b_1(t)m(t))b_0^{\mathrm{T}}(t)+b_1(t)P(t)b_1^{\mathrm{T}}(t) \tag{3.18}$$
$$+b_1(t)m(t)m^{\mathrm{T}}(t)b_1^{\mathrm{T}}(t)+b_0(t)\left(P(t)+m(t)m^{\mathrm{T}}(t)\right)b_2^{\mathrm{T}}(t)+b_2(t)\left(P(t)+m(t)m^{\mathrm{T}}(t)\right)b_0^{\mathrm{T}}(t)$$
$$+b_1(t)\left(3m(t)P(t)+m(t)\left(m(t)m^{\mathrm{T}}(t)\right)\right)b_2^{\mathrm{T}}(t)+b_2(t)\left(3P(t)m^{\mathrm{T}}(t)+\left(m(t)m^{\mathrm{T}}(t)\right)m^{\mathrm{T}}(t)\right)b_1^{\mathrm{T}}(t)$$
$$+3b_2(t)P^2(t)b_2^{\mathrm{T}}(t)+3b_2(t)\left(P(t)m(t)m^{\mathrm{T}}(t)+m(t)m^{\mathrm{T}}(t)P(t)\right)b_2^{\mathrm{T}}(t)+b_2(t)\left(m(t)m^{\mathrm{T}}(t)\right)^2b_2^{\mathrm{T}}(t).$$

Substitution of (3.16) into (3.7) and (3.17) and (3.18) into (3.8) leads to the following approximate—in virtue of the assumption that the random vector $x(t) - m(t)$ has a conditionally Gaussian distribution density—filtering equations for the rms estimate $m(t)$ and the covariance error matrix $P(t)$:

$$dm(t) = \Big(a_0(t) + a_1(t)m(t) + a_2(t)m(t)m^{\mathrm{T}}(t) + a_2(t)P(t)\Big)\,dt \tag{3.19}$$
$$+ P(t)A^{\mathrm{T}}(t)\left(B(t)B^{\mathrm{T}}(t)\right)^{-1}[dy(t) - (A_0(t) + A(t)m(t))dt],$$
$$m(t_0) = E\left(x(t_0)|F^Y_{t_0}\right),$$

$$dP(t) = \Big(a_1(t)P(t) + P(t)a_1^{\mathrm{T}}(t) + 2a_2(t)m(t)P(t) + 2(a_2(t)m(t)P(t))^{\mathrm{T}} + b_0(t)b_0^{\mathrm{T}}(t) \tag{3.20}$$
$$+ b_0(t)(b_1(t)m(t))^{\mathrm{T}} + (b_1(t)m(t))b_0^{\mathrm{T}}(t) + b_1(t)P(t)b_1^{\mathrm{T}}(t) + b_1(t)m(t)m^{\mathrm{T}}(t)b_1^{\mathrm{T}}(t)$$
$$+ b_0(t)\left(P(t) + m(t)m^{\mathrm{T}}(t)\right)b_2^{\mathrm{T}}(t) + b_2(t)\left(P(t) + m(t)m^{\mathrm{T}}(t)\right)b_0^{\mathrm{T}}(t)$$
$$+ b_1(t)\Big(3m(t)P(t)m(t)\left(m(t)m^{\mathrm{T}}(t)\right)\Big)b_2^{\mathrm{T}}(t) + b_2(t)\Big(3P(t)m^{\mathrm{T}}(t) + \left(m(t)m^{\mathrm{T}}(t)\right)m^{\mathrm{T}}(t)\Big)b_1^{\mathrm{T}}(t)$$
$$+ 3b_2(t)P^2(t)b_2^{\mathrm{T}}(t) + 3b_2(t)\left(P(t)m(t)m^{\mathrm{T}}(t) + m(t)m^{\mathrm{T}}(t)P(t)\right)b_2^{\mathrm{T}}(t)$$
$$+ b_2(t)\left(m(t)m^{\mathrm{T}}(t)\right)^2 b_2^{\mathrm{T}}(t)\Big)dt - P(t)A^{\mathrm{T}}(t)\left(B(t)B^{\mathrm{T}}(t)\right)^{-1}A(t)P(t)dt,$$
$$P(t_0) = E\left[(x(t_0) - m(t_0))(x(t_0) - m(t_0))^{\mathrm{T}}|F^Y_{t_0}\right].$$

The above reasoning gives rise to the statement that the rms finite-dimensional filter for the stochastic system with the equation of state (2.1), where the polynomials $f(x,t)$ and $g(x,t)$ are defined in (3.14) and (3.15), is determined from the linear observations (2.2) by the approximate Eq. (3.19) for the rms estimate $m(t) = E(x(t)|F^Y_t)$ and approximate Eq. (3.20) for the covariance error matrix $P(t) = E[(x(t) - m(t))(x(t) - m(t))^{\mathrm{T}}|F^Y_t]$.

Therefore, on the basis of the general open system of filtering Eqs. (3.7) and (3.8) a procedure is proposed for determining the closed system of approximate filtering equations for any polynomial state (2.1) from the linear observations (2.2). Additionally, obtained was a particular form of the closed system of approximate filtering Eqs. (3.19) and (3.20) for the state vector with quadratic drift and diffusion.

4. EXAMPLE OF SOLUTION OF THE FILTERING PROBLEM WITH QUADRATIC DRIFT AND DIFFUSION

In this section we consider an example of constructing an rms filter for the state vector with quadratic drift and diffusion from linear observations which is compared with the rms filter for the quadratic state vector with state-independent noise and one of the generalized linear Kalman–Bucy filters.

Let the two-dimensional state vector $x(t)$ satisfy the quadratic equation system (state $x(t)$, time and the system coefficients are assumed to be dimensionless)

$$\dot{x}_1(t) = x_2(t) + 0.1x_1^2(t)\psi_1(t), \quad x_1(0) = x_{10},$$
$$\dot{x}_2(t) = 0.1x_2^2(t), \quad x_2(0) = x_{20}, \tag{4.1}$$

and the scalar process of observation be defined by the linear equation

$$y_1(t) = x_1(t) + \psi_2(t), \tag{4.2}$$

where $\psi_1(t)$ and $\psi_2(t)$ are white Gaussian noise of unit intensity which are weak rms derivatives of the standard Wiener processes (see [21]). Equations (4.1) and (4.2) represent an alternative form of Eqs. (2.1) and (2.2) which is used usually in practice [22] and where $dy(t)/dt$ is redenoted as $y_1(t)$.

The problem of filtering lies in determining the rms estimate of the state vector (4.1) from the linear observations (4.2) with independent and identically distributed perturbations in the form of white Gaussian noise. The solution duration is assumed to be $T = 0.92$.

The filtering Eqs. (3.19) and (3.20) take on the following form for system (4.1) and (4.2):

$$\dot m_1(t) = m_2(t) + P_{11}(t)[y(t) - m_1(t)], \qquad (4.3)$$
$$\dot m_2(t) = 0.1m_2^2 + 0.1P_{22}(t) + P_{12}(t)[y(t) - m_1(t)]$$

with the initial condition

$$m(0) = E(x(0)|y(0)) = m_0$$

and

$$\dot P_{11}(t) = 2P_{12}(t) - 0.97P_{11}^2(t) + 0.03P_{12}^2(t) + 0.06P_{12}(t)m_1(t)m_2(t) + 0.01m_1^4 + 0.01m_1^2m_2^2, \qquad (4.4)$$
$$\dot P_{12}(t) = P_{22}(t) + 0.2m_2(t)P_{12}(t) - P_{11}(t)P_{12}(t),$$
$$\dot P_{22}(t) = 0.4m_2(t)P_{22}(t) - P_{12}^2(t)$$

with the initial condition

$$P(0) = E\Big((x(0) - m(0))(x(0) - m(0))^{\mathrm T}|y(0)\Big) = P_0.$$

The estimates satisfying Eqs. (4.3) and (4.4) are first compared with the estimates satisfying the rms filtering equations for the polynomial system with state-independent diffusion [7]:

$$\dot m_{I1}(t) = m_{I2}(t) + P_{I11}(t)[y(t) - m_{I1}(t)], \qquad (4.5)$$
$$\dot m_{I2}(t) = 0.1m_{I2}^2 + 0.1P_{I22}(t) + P_{I12}(t)[y(t) - m_{I1}(t)]$$

with the initial condition

$$m_I(0) = E(x(0)|y(0)) = m_0$$

and

$$\dot P_{I11}(t) = 2P_{I12}(t) + 0.01 - P_{I11}^2(t), \qquad (4.6)$$
$$\dot P_{I12}(t) = P_{I22}(t) + 0.2m_{I2}(t)P_{I12}(t) - P_{I11}(t)P_{I12}(t),$$
$$\dot P_{I22}(t) = 0.4m_{I2}(t)P_{I22}(t) - P_{I12}^2(t)$$

with the initial condition

$$P_I(0) = E\Big((x(0) - m(0))(x(0) - m(0))^{\mathrm T}|y(0)\Big) = P_0.$$

The estimates satisfying Eqs. (4.3) and (4.4) are also compared with the estimates satisfying the equations of the generalized Kalman–Bucy filter where the matrix $P_K(t)$ satisfies the Riccati equation:

$$\dot m_{K1}(t) = m_{K2}(t) + P_{K11}(t)[y(t) - m_{K1}(t)], \qquad (4.7)$$
$$\dot m_{K2}(t) = 0.1m_{K2}^2 + 0.1P_{K22}(t) + P_{K12}(t)[y(t) - m_{K1}(t)]$$

with the initial condition

$$m_K(0) = E(x(0)|y(0)) = m_0$$

and

$$\dot P_{K11}(t) = 2P_{K12}(t) + 0.01 - P_{K11}^2(t), \qquad (4.8)$$
$$\dot P_{K12}(t) = P_{K22}(t) + 0.2P_{K12}(t) - P_{K11}(t)P_{K12}(t),$$
$$\dot P_{K22}(t) = 0.4P_{K22}(t) - P_{K12}^2(t)$$

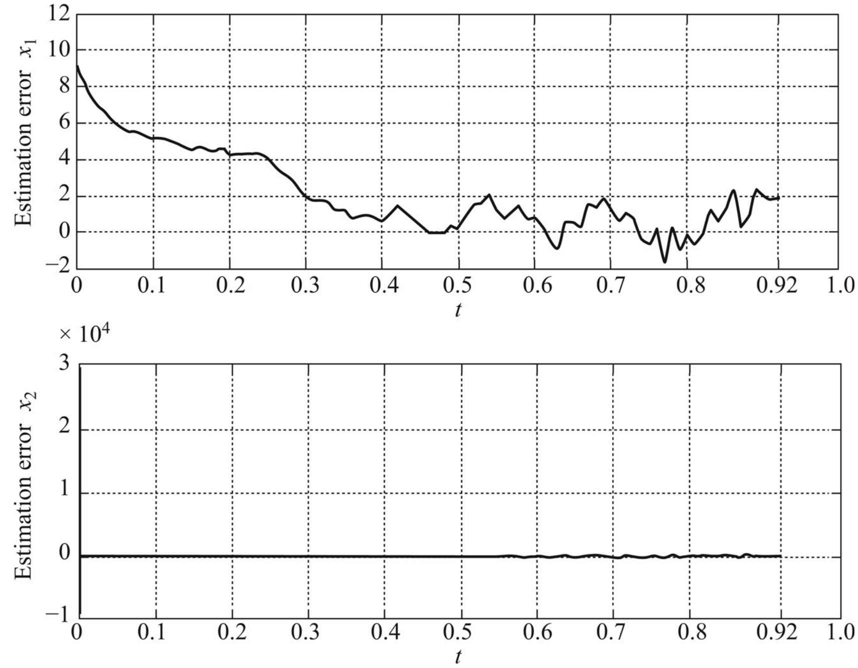

Fig. 1. Errors of estimation of the state variables $x_1(t)$ and $x_2(t)$ satisfying Eqs. (4.1) by the estimates $m_1(t)$ and $m_2(t)$ satisfying Eqs. (4.3) over the interval [0; 0.92].

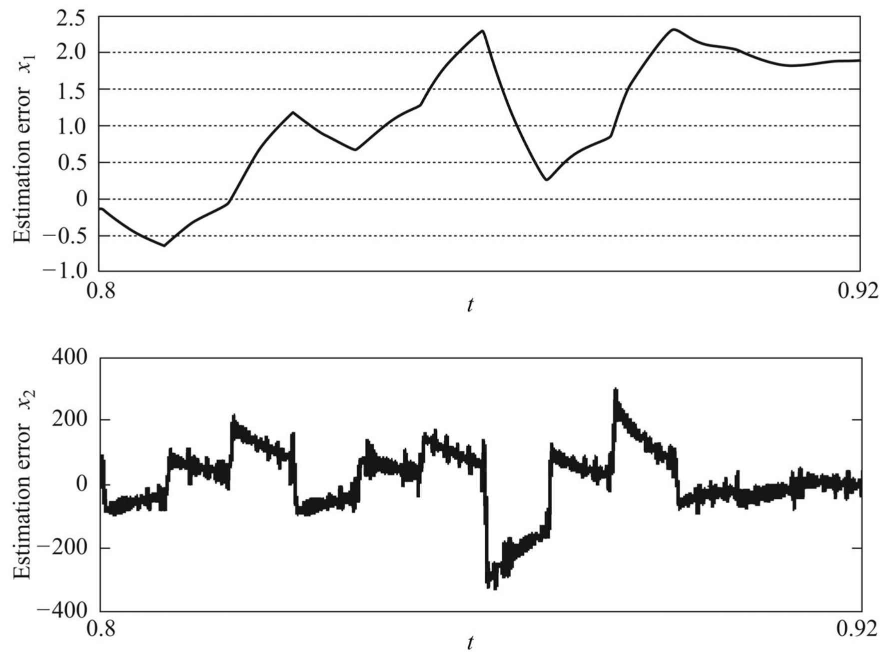

Fig. 2. Errors of estimation of the state variables $x_1(t)$ and $x_2(t)$ satisfying Eqs. (4.1), by the estimates $m_1(t)$ and $m_2(t)$ and satisfying Eqs. (4.3) over interval [0.80; 0.92].

with the initial condition

$$P_K(0) = E\left((x(0) - m(0))(x(0) - m(0))^{\mathrm{T}} | y(0)\right) = P_0.$$

The numerical results of modeling were obtained as the result of solving the equation systems (4.3) and (4.4), (4.5) and (4.6), (4.7) and (4.8). The values of estimates $m_1(t)$, $m_2(t)$, $m_{I1}(t)$, $m_{I2}(t)$,

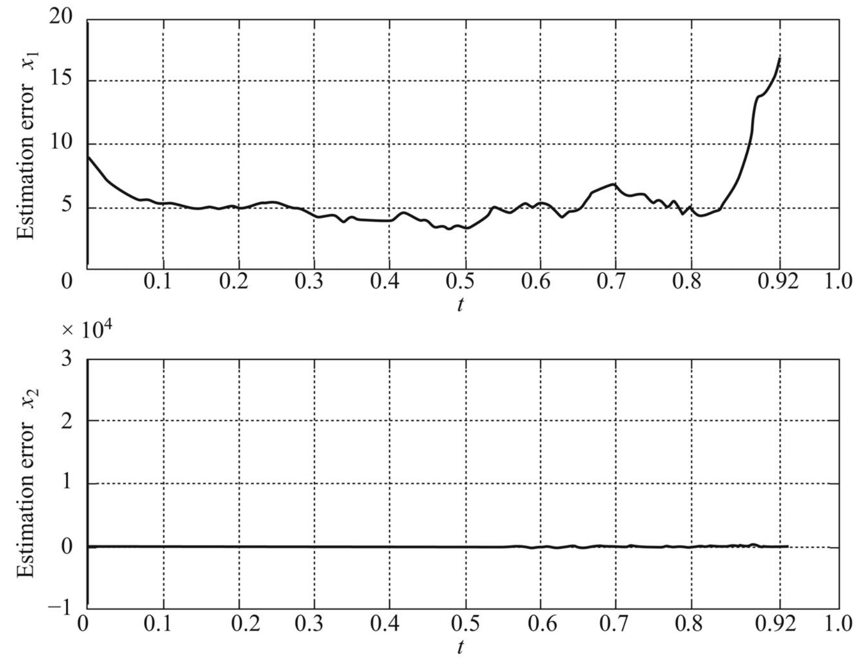

Fig. 3. Errors of estimation of the state variables $x_1(t)$ and $x_2(t)$ satisfying Eqs. (4.1) by the estimates $m_{I1}(t)$ and $m_{I2}(t)$ satisfying Eqs. (4.5) over the interval [0; 0.92].

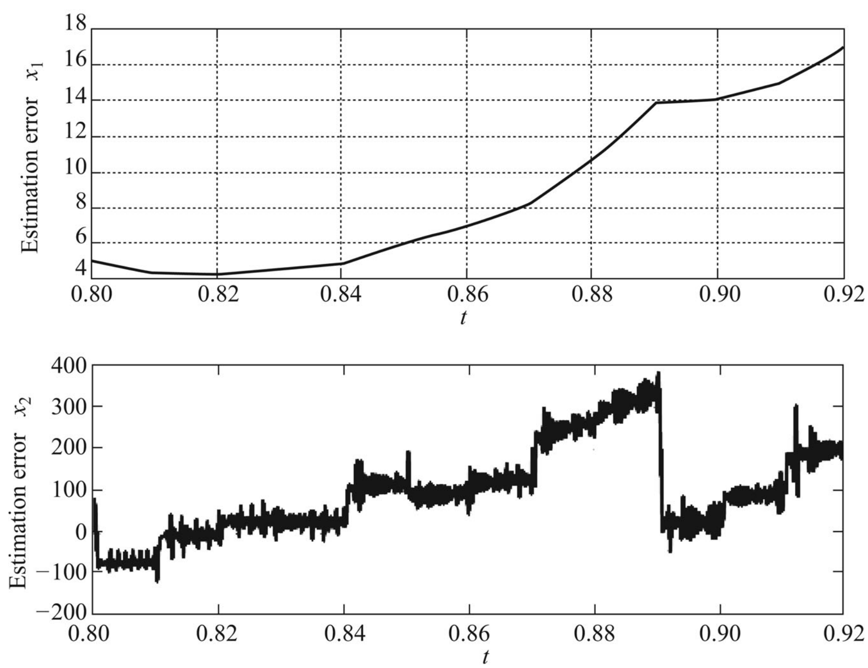

Fig. 4. Errors of estimation of the state variables $x_1(t)$ and $x_2(t)$ satisfying Eqs. (4.1) by the estimates $m_{I1}(t)$ and $m_{I2}(t)$ satisfying Eqs. (4.5) over the interval [0.80; 0.92].

$m_{K1}(t)$, and $m_{K2}(t)$, satisfying Eqs. (4.3), (4.5), and (4.7), respectively, are compared with the actual values of the state variables $x_1(t)$ and $x_2(t)$ in (4.1).

The initial conditions $x_{10} = 10.1$, $x_{20} = 10.1$, $m_{10} = 1.1$, $m_{20} = 1.1$, $P_{110} = 10$, $P_{120} = 1$, $P_{220} = 10$ are used for each of the three filters (4.3) and (4.4), (4.5) and (4.6), (4.7) and (4.8) and the original system (4.1) and (4.2). The Gaussian noise $\psi_1(t)$ in (24) and $\psi_2(t)$ in (4.2) are realized using the built-in MatLab function.

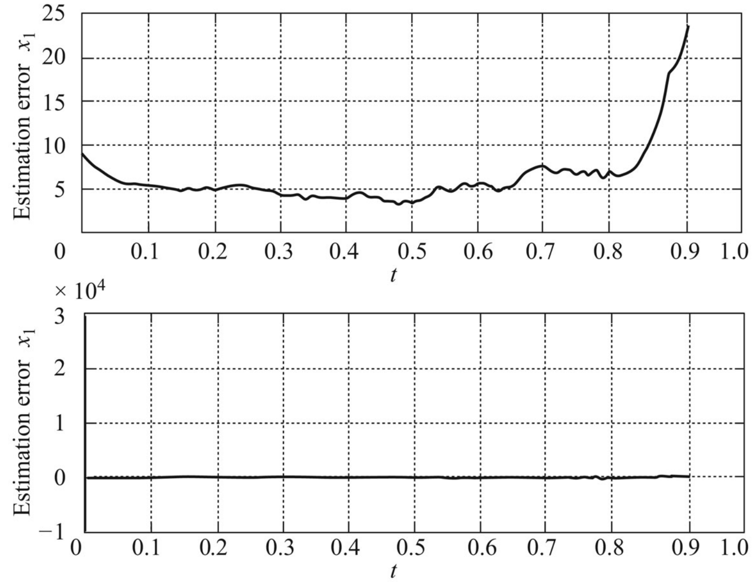

Fig. 5. Errors of estimation of the state variables $x_1(t)$ and $x_2(t)$ satisfying Eqs. (4.1) by the estimates $m_{K1}(t)$ and $m_{K2}(t)$ satisfying (4.7) over the interval [0; 0.92].

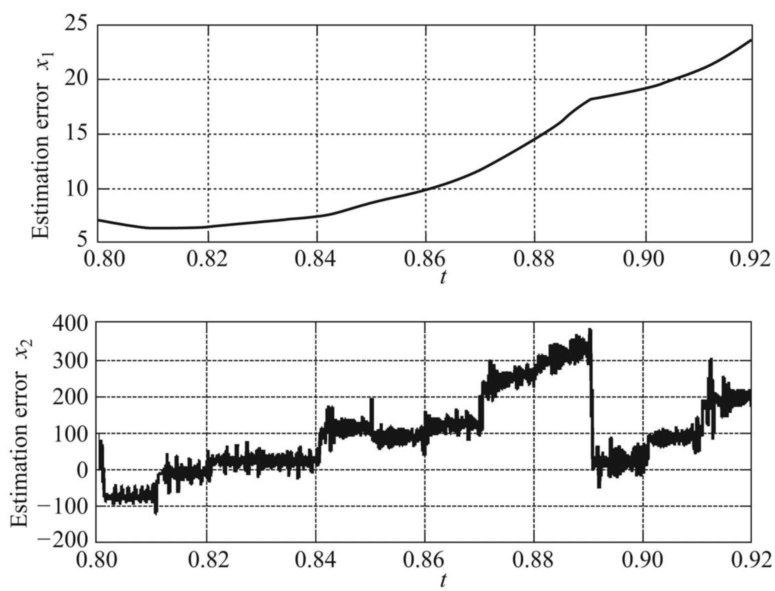

Fig. 6. Errors of estimation of the state variables $x_1(t)$ and $x_2(t)$ satisfying Eqs. (4.1) by the estimates $m_{K1}(t)$ and $m_{K2}(t)$ satisfying Eqs. (4.7) over the interval [0.80; 0.92].

Figures 1, 3, and 5 depict the graphs of the estimate errors of the state variables $x_1(t)$ and $x_2(t)$ obtained using Eqs. (4.3) (Fig. 1), Eqs. (4.5) (Fig. 3), and Eqs. (4.7) (Fig. 5). All graphs correspond to the interval from $t_0 = 0$ to $T = 0.92$. Detailed elaboration of the aforementioned graphs over the interval from $t = 0.80$ to $T = 0.92$ is shown in Figs. 2, 4, and 6, respectively. As can be seen from them, the estimation error corresponding to the filter (4.3) reach rapidly the neighborhood of zero and then maintain the zero men even in the immediate vicinity of the asymptotic time moment $T = 0.99$ where the component $x_1(t)$ of the state vector (4.1) directs towards the infinity. Nevertheless, for the unmeasurable component $x_2(t)$ there exists a short transient period during which the estimation error deviates appreciably from zero. The errors corresponding to other considered filters reach zero much slower, if reach at all, deviate systematically from zero and, obviously, go to infinity in the neighborhood of the asymptotic time instant. We note that the approximate covariance matrix of the estimation error $P(t)$ does not remain near zero in the neighborhood of the asymptotic time instant in virtue of the available fourth-order terms in the right side of Eq. (4.4), which suggests that the exact variances of the estimation error components have good quality.

The calculations for this realization suggest that filter (4.3) and (4.4) obtained for the quadratic state vector with quadratic multiplicative noise certainly gives a better estimate as compared with the filter corresponding to the case of state-independent noise or the generalized Kalman–Bucy filter.

5. FORMULATION OF THE PROBLEM OF JOINT RMS STATE FILTERING AND PARAMETER IDENTIFICATION FOR THE LINEAR STOCHASTIC SYSTEMS WITH UNKNOWN PARAMETERS

Let (Ω, F, P) be a complete probabilistic space with the increasing right-continuous family of σ-algebras F_t, $t \geq t_0$, where independent standard Wiener processes $(W_1(t), F_t, t \geq t_0)$ and $(W_2(t), F_t, t \geq t_0)$ are defined. The F_t-measurable random process $(x(t), y(t))$ obeys a linear differential equation of state with an unknown vector parameter $\theta(t)$

$$dx(t) = (a_0(\theta,t) + a(\theta,t)x(t))dt + b(t)dW_1(t), \quad x(t_0) = x_0, \tag{5.1}$$

and linear differential equation of the observation process

$$dy(t) = (A_0(t) + A(t)x(t))dt + B(t)dW_2(t), \tag{5.2}$$

where $x(t) \in R^n$ is the state vector, $y(t) \in R^l$, $l \leq n$, is the vector of linear observations, and $\theta(t) \in R^p$, $p \leq n \times n + n$, is the vector of unknown elements of the matrix $a(\theta,t)$ and the components of the vector $a_0(\theta,t)$ comprising both unknown components $a_{0_i}(t) = \theta_k(t)$, $k = 1, \ldots, p_1 \leq n$, and $a_{ij}(t) = \theta_k(t)$, $k = p_1 + 1, \ldots, p \leq n \times n + n$, and certain components $a_{0_i}(t)$ and $a_{ij}(t)$ representing the given time functions. The conditions imposed on Eqs. (2.1) and (2.2) in Section 2 are satisfied as well.

It is assumed that there is no useful information about the unknown parameters $\theta_k(t)$, $k = 1, \ldots, p$, and this uncertainty increases with time tending to infinity. Stated differently, the unknown parameters can be described as the F_t-measurable standard Wiener processes

$$d\theta(t) = dW_3(t) \tag{5.3}$$

with unknown initial conditions $\theta(t_0) = \theta_0 \in R^p$, where $(W_3(t), F_t, t \geq t_0)$ is a Wiener process independent of x_0, $W_1(t)$, and $W_2(t)$.

The problem of estimation lies in determining the rms estimate $\hat{z}(t) = [\hat{x}(t), \hat{\theta}(t)]$ of the augmented vector of system state and the unknown parameters $z(t) = [x(t), \theta(t)]$ from all observations since the initial instant to the current one $Y(t) = \{y(s), t_0 \leq s \leq t\}$ which minimizes the quadratic criterion

$$J = E\left[(z(t) - \hat{z}(t))^{\mathrm{T}}(z(t) - \hat{z}(t)) | F_t^Y\right]$$

at each instant t. Here, $E[\xi(t)|F_t^Y]$ denotes the conditional expectation of the random process $\xi(t) = (z(t) - \hat{z}(t))^{\mathrm{T}}(z(t) - \hat{z}(t))$ relative to the σ-algebra F_t^Y. Solution of this problem is based on the rms finite-dimensional filter for the quadratic equation of state from the linear observations determined in Section 3.

6. JOINT ROOT-MEAN-SQUARE STATE FILTERING AND PARAMETER IDENTIFICATION FOR THE LINEAR STOCHASTIC SYSTEMS WITH UNKNOWN PARAMETERS

The equation of state (5.1) must be presented in a polynomial form to enable application of the filtering Eqs. (3.19) and (3.20) to the state vector $z(t) = [x(t), \theta(t)]$ consistent with Eqs. (5.1) and (5.3) in the linear observations (5.2). For that, the matrix $a_1(t) \in R^{(n+p)\times(n+p)}$, cubic tensor $a_2(t) \in R^{(n+p)\times(n+p)\times(n+p)}$, and vector $c_0(t) \in R^{(n+p)}$ in the following Eq. (6.1) are introduced as follows.

The equation for the ith component of the state vector is given by

$$dx_i(t) = \left(a_{0_i}(t) + \sum_{j=1}^{n} a_{ij}(t)x_j(t)\right) dt + \sum_{j=1}^{n} b_{ij}(t)dW_{1_j}(t), \quad x_i(t_0) = x_{0_i}.$$

Then,

1. If $a_{0_i}(t)$ is a certain function, then the ith component of the vector $c_0(t)$ is assumed to be equal to this function $c_{0_i}(t) = a_{0_i}(t)$. Otherwise, if $a_{0_i}(t)$ is an unknown function, the $(i, n+i)$th element of the matrix $a_1(t)$ is assumed to be equal to 1.

2. If $a_{ij}(t)$ is a certain function, then the (i, j)th element of the matrix $a_1(t)$ is assumed to be equal to this function, $a_{1_{ij}}(t) = a_{ij}(t)$. Otherwise, if $a_{ij}(t)$ is an unknown function, then the $(i, n+p_1+k, j)$th element of the cubic tensor $a_2(t)$ is assumed to be equal to 1, where k is the number of the unknown current element of the matrix $a_{ij}(t)$ on condition that the unknown elements are counted successively row-by-row from the first to the nth element of each row.

3. The rest of the elements of the matrix $a_1(t)$, cubic tensor $a_2(t)$, and vector $c_0(t)$ that were not determined earlier are assumed to be zero.

With the notation introduced, the equations of state (5.1), (5.3) for the vector $z(t) = [x(t), \theta(t)] \in R^{n+p}$ are representable as

$$dz(t) \left(c_0(t) + a_1(t)z(t) + a_2(t)z(t)z^{\mathrm{T}}(t)\right) dt + \mathrm{diag}[b(t), I_{p\times p}]d \left[W_1^{\mathrm{T}}(t), W_3^{\mathrm{T}}(t)\right]^{\mathrm{T}}, \quad z(t_0) = [x_0, \theta_0], \tag{6.1}$$

where the matrix $a_1(t)$, cubic tensor $a_2(t)$, and vector $c_0(t)$ already have been determined, and $I_{p\times p}$ is the $p \times p$ identity matrix. Equation (6.1) is quadratic relative to the extended state vector $z(t) = [x(t), \theta(t)]$.

Therefore, the problem of estimation now lies in determining an rms estimate $\hat{z}(t) = m(t) = [\hat{x}(t), \hat{\theta}(t)]$ of the extended state vector $z(t) = [x(t), \theta(t)]$ meeting Eq. (6.1) in the linear observations $Y(t) = \{y(s), 0 \le s \le t\}$ satisfying Eq. (5.2). This problem is solved using the filtering Eqs. (3.19) and (3.20) and is given by

$$dm(t) = \left(c_0(t) + a_1(t)m(t) + a_2(t)m(t)m^{\mathrm{T}}(t) + a_2(t)P(t)\right) dt + P(t)[A(t), 0_{m\times p}]^{\mathrm{T}} \left(B(t)B^{\mathrm{T}}(t)\right)^{-1} [dy(t) - A(t)m(t)dt], \tag{6.2}$$

$$m(t_0) = \left[E\left(x(t_0)|F_t^Y\right), E\left(\theta(t_0)|F_{t_0}^Y\right)\right],$$

$$dP(t) = \Big(a_1(t)P(t) + P(t)a_1^{\mathrm{T}}(t) + 2a_2(t)m(t)P(t) + 2(a_2(t)m(t)P(t))^{\mathrm{T}} + \mathrm{diag}[b(t), I_p]\mathrm{diag}[b(t), I_p]^{\mathrm{T}}\Big) dt - P(t)[A(t), 0_{m\times p}]^{\mathrm{T}} \left(B(t)B^{\mathrm{T}}(t)\right)^{-1} [A(t), 0_{m\times p}]P(t)dt, \tag{6.3}$$

$$P(t_0) = E\left((z(t_0) - m(t_0))(z(t_0) - m(t_0))^{\mathrm{T}}|F_{t_0}^Y\right),$$

where $0_{m\times p}$ is an $m \times p$ zero matrix and $P(t)$ is the covariance matrix of the estimation error $z(t) - m(t)$ from observations $Y(t)$.

The following assertion can be proved using the above results. The rms finite-dimensional filter for the extended state vector $[x(t), \theta(t)]$ satisfying Eq. (6.1) in linear observations (5.2) is defined by the approximate Eq. (6.2) for the rms estimate $\hat{z}(t) = m(t) = [\hat{x}(t), \hat{\theta}(t)] = E([x(t), \theta(t)]|F_t^Y)$ and the approximate Eq. (6.3) of the error covariance matrix $P(t) = E[(z(t) - m(t))(z(t) - m(t))^{\mathrm{T}}|F_t^Y]$. In particular, the generated filter gives an approximate rms estimate of the vector of unknown parameters $\theta(t)$ in Eq. (5.3).

7. EXAMPLE OF JOINT ROOT-MEAN-SQUARE STATE FILTERING AND PARAMETER IDENTIFICATION

Here we give an example of constructing a joint rms filter and parameter identifier for a linear system with an unknown multiplicative parameter in the equation of state. Let the two-dimensional state vector $x(t)$ satisfy the linear system of equations with an unknown parameter θ (state $x(t)$, time, system coefficients, and the parameter θ are assumed to be dimensionless)

$$\dot{x}_1(t) = x_2(t), \quad x_1(0) = x_{10},$$
$$\dot{x}_2(t) = \theta x_2(t) + \psi_1(t), \quad x_2(0) = x_{20} \tag{7.1}$$

and the scalar process of observation be given by the linear equation

$$y_1(t) = x_1(t) + \psi_2(t), \tag{7.2}$$

where $\psi_1(t)$ and $\psi_2(t)$ are white Gaussian noise of unit intensity which are weak rms derivatives of the standard Wiener processes (see [21]). The parameter θ is modeled as a standard Wiener process, that is, satisfies the equation

$$d\theta(t) = dW_3(t), \quad \theta(0) = \theta_0,$$

representable also as

$$\dot{\theta}(t) = \psi_3(t), \quad \theta(0) = \theta_0, \tag{7.3}$$

where $\psi_3(t)$ is white Gaussian noise of unit intensity. It can be noted that problems of this kind are encountered at processing the navigation information under imprecisely known system parameters [23, 24].

The problem of filtering lies in determining the rms estimate $m(t) = [m_1(t), m_2(t), m_3(t)]$ of the state vector (7.1), (7.3), $[x_1(t), x_2(t), \theta]$ from linear observations (7.2). The solution duration is taken to be $\mathrm{T} = 4$.

For the system (7.1)–(7.3), the equations of filtering (6.2) and (6.3) go over to

$$\dot{m}_1(t) = m_2(t) + P_{11}(t)[y(t) - m_1(t)], \tag{7.4}$$
$$\dot{m}_2(t) = m_2(t)m_3(t) + P_{23}(t) + P_{12}(t)[y(t) - m_1(t)],$$
$$\dot{m}_3(t) = P_{13}(t)[y(t) - m_1(t)]$$

with the initial conditions $m_1(0) = E(x_{10}|y(0)) = m_{10}$, $m_2(0) = E(x_{20}|y(0)) = m_{20}$ and $m_3(0) = E(\theta_0|y(0)) = m_{30}$,

$$\dot{P}_{11}(t) = 2P_{12}(t) - P_{11}^2(t), \tag{7.5}$$
$$\dot{P}_{12}(t) = P_{22}(t) + 2P_{13}(t)m_2(t) - P_{11}(t)P_{12}(t),$$
$$\dot{P}_{13}(t) = P_{23}(t) - P_{11}(t)P_{13}(t),$$
$$\dot{P}_{22}(t) = 1 + 4P_{23}(t)m_2(t) - P_{12}^2(t),$$
$$\dot{P}_{23}(t) = 2P_{33}(t)m_2(t) - P_{12}(t)P_{13}(t),$$
$$\dot{P}_{33}(t) = 1 - P_{13}^2(t)$$

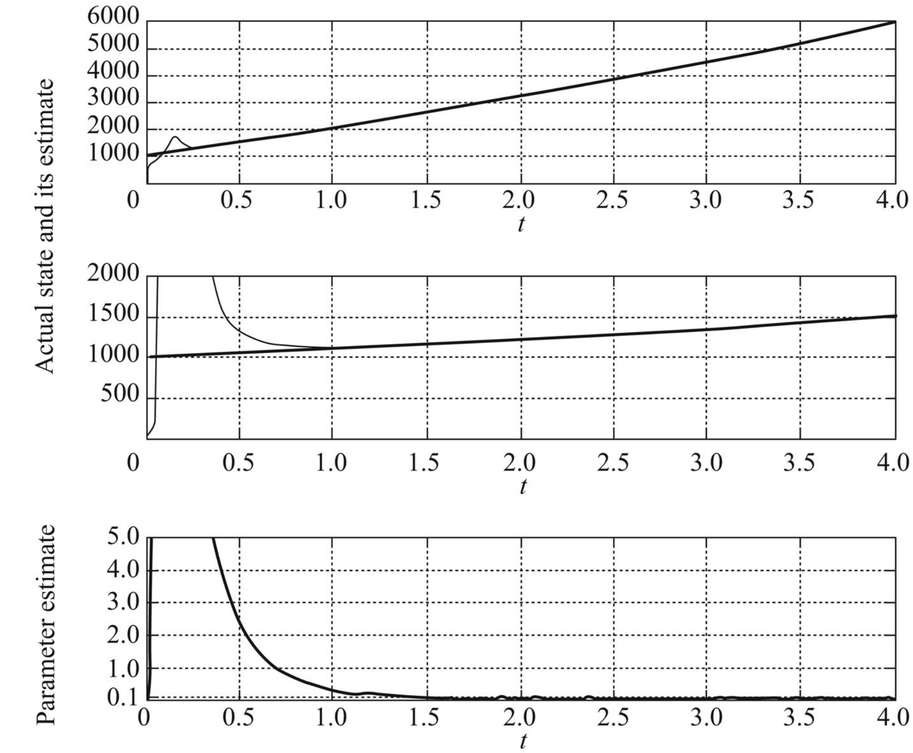

Fig. 7. State variables $x_1(t)$ (solid upper line), $x_2(t)$ (solid central line), their estimates $m_1(t)$ (thin upper line), $m_2(t)$ (thin central line), and the parameter estimate $m_3(t)$ (solid line below) for $\theta = 0.1$ over the interval $[0,\ 4]$.

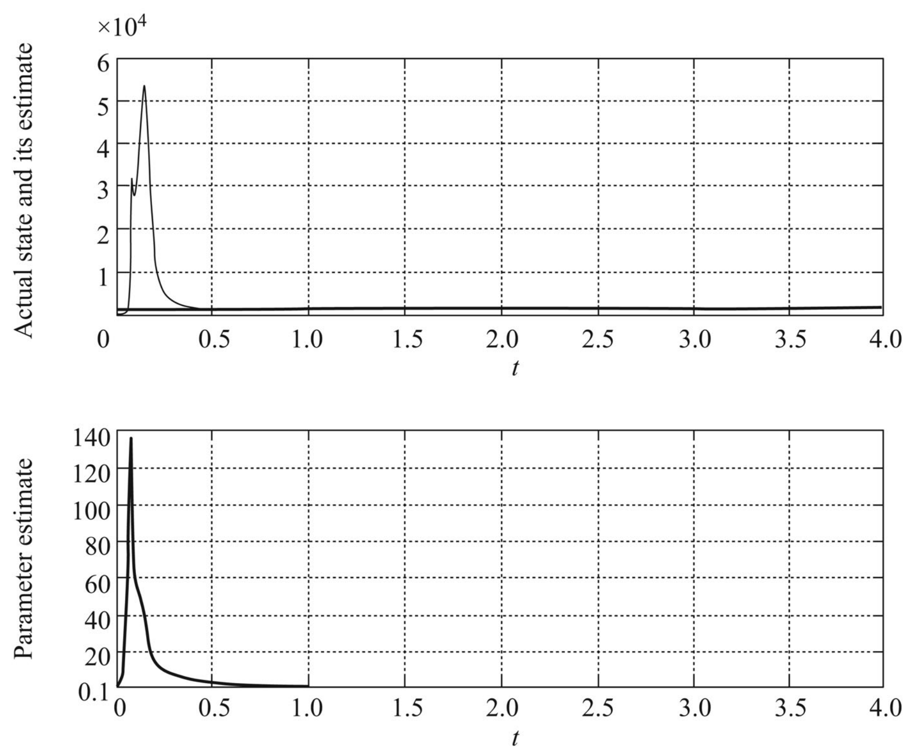

Fig. 8. State variable $x_2(t)$ (solid upper line), its estimate $m_2(t)$ (thin upper line), and parameter estimate $m_3(t)$ (solid line below) for $\theta = 0.1$ over the interval $[0,\ 4]$.

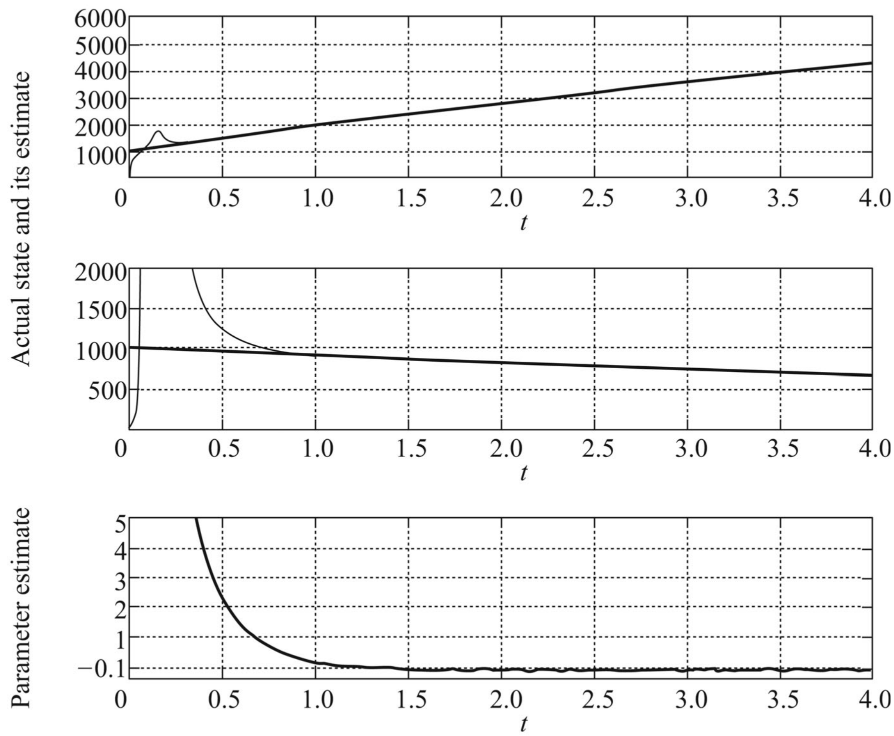

Fig. 9. State variables $x_1(t)$ (solid upper line), $x_2(t)$ (solid central line), their estimates $m_1(t)$ (thin upper line), $m_2(t)$ (thin central line) and the parameter estimate $m_3(t)$ (solid line below) for $\theta = -0.1$ over interval [0, 4].

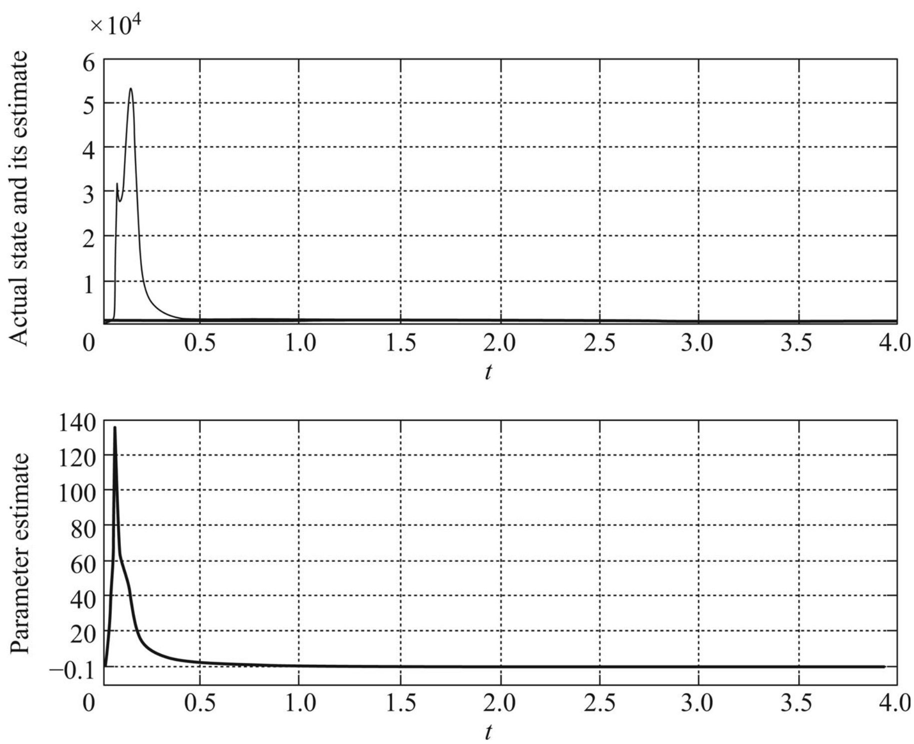

Fig. 10. State variable $x_2(t)$ (solid upper line), its estimate $m_2(t)$ (thin upper line), and parameter estimate $m_3(t)$ (solid lower line) for $\theta = -0.1$ over the interval [0, 4].

with the initial condition

$$P(0) = E\Big(([x_{10}, x_{20}, \theta_0] - m(0))([x_{10}, x_{20}, \theta_0] - m(0))^{\mathrm{T}} | y(0)\Big) = P_0.$$

The numerical results of modeling were established by solving the system of filtering Eqs. (7.4) and (7.5). The resulting estimates $[m_1(t), m_2(t)]$ for $[x_1(t), x_2(t)]$ and $m_3(t)$ for $\theta(t)$ were compared with the actual values of the state vector $x(t) = [x_1(t), x_2(t)]$ and the parameter θ in (7.4), (7.5).

The following initial values were used for filter (7.4), (7.5) and the original system (7.1), (7.2): $x_{10} = x_{20} = 1000$, $m_{10} = m_{20} = 0.1$, $m_{30} = 0$, $P_{110} = P_{220} = P_{330} = 100$, $P_{120} = 10$ and $P_{130} = P_{230} = 0$. The unknown parameter θ was taken $\theta = 0.1$ in the first modeling and $\theta = -0.1$ in the second. Thus, consideration was given both to the stable and unstable cases in Eq. (7.1).

Figures 7–10 depict the graphs of the state vector $x(t) = [x_1(t), x_2(t)]$, state vector estimates $[m_1(t), m_2(t)]$, and the estimates of the parameter $m_3(t)$ for the negative ($\theta = -0.1$) and positive ($\theta = 0.1$) values of the parameter. The results of modeling are shown for the positive (Figs. 7 and 8) and negative (Figs. 9 and 10) cases over the interval from $t_0 = 0$ to T = 4. The full views of the graphs for estimation of the unknown parameter is shown in Figs. 8 and 10. As can be seen, in both cases the state estimates $[m_1(t), m_2(t)]$ converge to the state vector $[x_1(t), x_2(t)]$ and the estimates of the parameter $m_3(t)$ converge to the actual value (0.1 or -0.1) of the unknown parameter $\theta(t)$. We note that, as it was expected, the diagonal elements of the approximate covariance matrix $P_{ii}(t)$, $i = 1, 2, 3$, converge to the finite values that are near 1 when time tends to infinity, which allows one to conclude that the precise variances of the components of the estimation errors behave well.

Therefore, one can conclude that in both cases the constructed consistent state filter and parameter identifier (7.4) and (7.5) provide reliable estimates of system state and the unknown value of the parameter. The results obtained demonstrate that the estimates of state and parameters as computed using the constructed consistent identifier filter sufficiently rapidly converge to the actual values of state and parameters less than in four time units. Such behavior can be classified as very reliable, especially with regard for great departures of the estimation error from zero at the initial time instant.

8. CONCLUSIONS

Some results of the present author in the field of constructing the approximate finite-dimensional rms filters for stochastic systems with polynomial equations of state and multiplicative noise from linear observations are presented. Described are, in particular, a filtering algorithm for stochastic systems with polynomial equation of state and the methodology of applying it to solve the problem of joint filtering of state and parametric identification. The results obtained open way to their further use in the problems of filtering and identification encountered in navigation and trajectory tracking, in electronic and mechatronic applications, as well as in the chemical industry where the plant equations are set down or approximated with the use of nonlinear polynomial functions.

ACKNOWLEDGMENTS

This work was supported by the Russian Foundation for Basic Research, project no. 14-08-00347.

REFERENCES

1. Stratonovich, R.L., On the Theory of Optimal Nonlinear Filtering of Random Functions, *Teor. Veroyat. Primen.*, 1959, no. 4, pp. 239–242.

2. Kushner, H.J., On the Differential Equations Satisfied by Conditional Probability Densities of Markov Processes, with Applications, *SIAM J. Control*, 1964, vol. 2, pp. 106–119.

3. Kalman, R.E. and Bucy, R.S., New Results in Linear Filtering and Prediction Theory, *ASME Trans., Part D (J. Basic Eng.)*, 1961, vol. 83, pp. 95–108.

4. Wonham, W.M., Some Applications of Stochastic Differential Equations to Nonlinear Filtering, *SIAM J. Control*, 1965, vol. 2, pp. 347–369.

5. Benes, V.E., Exact Finite-dimensional Filters for Certain Diffusions with Nonlinear Drift, *Stochastics*, 1981, vol. 5, pp. 65–92.

6. Yau, S.S.-T., Finite-dimensional Filters with Nonlinear Drift I: A Class of Filters Including Both Kalman–Bucy and Benes Filters, *J. Math. Syst. Estimat. Control*, 1994, vol. 4, pp. 181–203.

7. Basin, M.V., On Optimal Filtering for Polynomial System States, *ASME Trans., J. Dynam. Syst. Meas. Control*, 2003, vol. 125, pp. 123–125.

8. Basin, M.V. and Alcorta García, M.A., Optimal Filtering and Control for Third Degree Polynomial Systems, *Dynam. Contin. Discrete Impulsive Syst.*, 2003, vol. 10B, pp. 663–680.

9. Basin, M.V., Perez, J., and Skliar, M., Optimal Filtering for Polynomial System States with Polynomial Multiplicative Noise, *Int. J. Robust Nonlin. Control*, 2006, vol. 16, pp. 287–298.

10. Basin, M.V., Calderón Álvarez, D., and Skliar, M., Optimal Filtering for Incompletely Measured Polynomial States over Linear Observations, *Int. J. Adaptive Control Signal Proc.*, 2008, vol. 22. pp. 482–494.

11. Basin, M.V. and Calderón Álvarez, D., Optimal Controller for Uncertain Stochastic Polynomial Systems with Deterministic Disturbances, *Int. J. Control*, 2009, vol. 82, pp. 1435–1447.

12. Basin, M.V., Shi, P., and Calderón Álvarez, D., Approximate Finite-dimensional Filtering for Polynomial States over Polynomial Observations, *Int. J. Control*, 2010, vol. 83, pp. 724–730.

13. Basin, M.V., Rodríguez Ramírez, P., Sliding Mode Filter Design for Nonlinear Polynomial Systems with Unmeasured States, *Inform. Sci.*, 2012, vol. 204, pp. 82–91.

14. Basin, M.V. and Rodríguez Ramírez, P., Sliding Mode Controller Design for Stochastic Polynomial Systems with Unmeasured States, *IEEE Trans. Indust. Electron.*, 2014, vol. 61, pp. 387–396.

15. Basin, M.V. and Maldonado Ortiz, J.J., and Karimi, H.R., Mean-square Filtering for Polynomial System States Confused with Poisson Noises over Polynomial Observations, *Model. Identificat. Control*, 2011, vol. 32, pp. 47–55.

16. Basin, M.V. and Maldonado Ortiz, J.J., Optimal Controller for Uncertain Stochastic Linear Systems with Poisson Noises, *IEEE Trans. Indust. Inform.*, 2014, vol. 10, pp. 267–275.

17. Germani, A., Manes, C., and Palumbo, P., Filtering of Stochastic Nonlinear Differential Systems via Carleman Approximation Approach, *IEEE Trans. Automat. Control*, 2007, vol. 52, pp. 2166–2172.

18. Luo, X. and Yau, S.S.-T., Complete Real-time Solution of the General Nonlinear Filtering Problem without Memory, *IEEE Trans. Automat. Control*, 2013, vol. 58, pp. 2563–2578.

19. Luo, X., On Recent Advance of Nonlinear Filtering Theory: Emphasis on Global Approaches, *Pure Appl. Math. Quart.*, 2014, vol. 10, pp. 685–721.

20. Sedov, L.I., *Mekhanika sploshnoi sredy* (Machanics of Continuous Media), Moscow: Nauka, 1970.

21. Pugachev, V.S. and Sinitsyn, I.N., *Stokhasticheskie differentsial'nye sistemy* (Stochastic Differential Systems), Moscow: Nauka, 1985.

22. Åström, K.J., *Introduction to Stochastic Control Theory*, New York: Academic, 1970.

23. Stepanov, O.A., Koshaev, D.A., and Motorin, A.V., Designing Models for Signals and Errors of Sensors in Airborne Gravimetry Using Nonlinear Filtering Methods, in *Proc. Inst. Navigat. 2015 Int. Technic. Meeting (ITM 2015)*, pp. 279–284.

24. Stepanov, O.A., Motorin, A.V., and Vasilyev, V.A., Identification of Sensor Errors by Using of Nonlinear Filtering, in *Proc. 1st Conf. Model. Identificat. Control Nonlinear Syst. (MICNON 2015)*, St. Petersburg, 2015.

This paper was recommended for publication by O.A. Stepanov, a member of the Editorial Board