Optimal Filtering for Partially Measured Polynomial System States

Michael Basin

Autonomous University of Nuevo Leon, Mexico
mbasin@fcfm.uanl.mx

Mikhail Skliar

University of Utah, USA
Mikhail.Skliar@m.cc.utah.edu

Abstract. In this paper, the optimal filtering problem for polynomial systems with partially measured linear part over linear observations is treated proceeding from the general expression for the stochastic Ito differential of the optimal estimate and the error variance. As a result, the Ito differentials for the optimal estimate and error variance corresponding to the stated filtering problem are first derived. The procedure for obtaining a closed system of the filtering equations for any polynomial state with partially measured linear part over linear observations with delay is then established, which yields the explicit closed form of the filtering equations in the particular case of a bilinear system state. In the example, performance of the designed optimal filter is verified for a quadratic-linear state with unmeasured linear part over linear observations against the conventionally designed extended Kalman-Bucy filter.

Keywords. Filtering, stochastic system, nonlinear polynomial system, bilinear system

I. Introduction

Although the general optimal solution of the filtering problem for nonlinear state and observation equations confused with white Gaussian noises is given by the Kushner equation for the conditional density of an unobserved state with respect to observations [1], there are a very few known examples of nonlinear systems where the Kushner equation can be reduced to a finite-dimensional closed system of filtering equations for a certain number of lower conditional moments. The most famous result, the Kalman-Bucy filter [2], is related to the case of linear state and observation equations, where only two moments, the estimate itself and its variance, form a closed system of filtering equations. However, the optimal nonlinear finite-dimensional filter can be obtained in some other cases, if, for example, the state vector can take only a finite number of admissible states [3] or if the observation equation is linear and the drift term in the state equation satisfies the Riccati equation $df/dx + f^2 = x^2$ (see [4]). The complete classification of the "general situation" cases (this means that there are no special assumptions on the structure of state and observation equations and the initial conditions), where the optimal nonlinear finite-dimensional filter exists, is given in [5]. Apart form the "general situation," the optimal finite-dimensional filters have recently been designed ([6], [7]) for certain classes of polynomial system states with Gaussian initial conditions over linear observations with invertible observation matrix.

This paper presents the optimal finite-dimensional filter for polynomial systems where the observation matrix may not be invertible and only the nonlinear components of the system state are assumed to be completely measured, whereas the linear part of the state may be partially measured or even not measured at all, thus generalizing the results of ([6], [7]). The optimal filtering problem is treated proceeding from the general expression for the stochastic Ito differential of the optimal estimate and the error variance [8]. As the first result, the Ito differentials for the optimal estimate and error variance corresponding to the stated filtering problem are derived. It is then proved that a closed finite-dimensional system of the optimal filtering equations with respect to a finite number of filtering variables can be obtained if the state equation is polynomial, the observations are linear, and only the nonlinear components of the state are assumed to be completely measured. In this case, the corresponding procedure for designing the optimal filtering equations is established. Finally, the closed system of the optimal filtering equations with respect to two variables, the optimal estimate and the error variance, is derived in the explicit form for the particular case of a bilinear state equation.

In the illustrative example, performance of the designed optimal filter is verified for a quadratic-linear state with unmeasured linear part over linear observations against the conventional extended Kalman-Bucy filter. The simulation results show a definite advantage of the designed optimal

This work was supported under the US NSF Grant CTS-0117300 and the Mexican National Science and Technology Council (CONACyT) Grant 39388-A.

filter in regard to proximity of the estimate to the real state value. Moreover, it can be seen that the estimate produced by the optimal filter asymptotically converges to the real values of the reference variables as time tends to infinity, although the system state itself is unstable and the quadratic component goes to infinity for a finite time. On the contrary, the conventionally designed extended Kalman-Bucy estimates diverge from the real values.

The paper is organized as follows. Section 2 presents the filtering problem statement for a polynomial system state with partially measured linear part over linear observations. The Ito differentials for the optimal estimate and the error variance are derived in Section 3. Section 4 establishes the procedure for obtaining a closed system of the filtering equations for any polynomial state with partially measured linear part over linear observations, which yields the explicit closed form of the filtering equations in the particular case of a bilinear system state. In Section 5, performance of the obtained optimal filter is verified for a quadratic-linear state with unmeasured linear part over linear observations against the conventional extended Kalman-Bucy filter.

II. Filtering Problem for Polynomial State with Partially Measured Linear Part over Linear Observations

Let (Ω, F, P) be a complete probability space with an increasing right-continuous family of σ-algebras $F_t, t \geq t_0$, and let $(W_1(t) = [W_{11}(t), W_{12}(t)], F_t, t \geq t_0)$ and $(W_2(t) = [W_{21}(t), W_{22}(t)], F_t, t \geq t_0)$ be independent Wiener processes. The F_t-measurable random process $(x(t) = [x_1(t), x_2(t)]$, $y(t) = [y_1(t), y_2(t)])$ is described by nonlinear differential equations for the system state

$$dx_1(t) = f(x_1, x_2, t)dt + b_{11}(t)dW_{11}(t), \ x_1(t_0) = x_{10}, \quad (1)$$

$$dx_2(t) = (a_{20}(t) + a_{21}(t)x_2(t))dt + b_{12}(t)dW_{12}(t),$$
$$x_2(t_0) = x_{20}, \quad (2)$$

and linear differential equations for the observation process

$$dy_1(t) = (A_{01}(t) + A_1(t)x_1(t))dt + B_1(t)dW_{21}(t), \quad (3)$$

$$dy_2(t) = (A_{02}(t) + A_2(t)x_2(t))dt + B_2(t)dW_{22}(t). \quad (4)$$

Here, $x(t) = [x_1(t), x_2(t)] \in R^n$ is the state vector, $x_1(t) \in R^{n_1}$ is the completely measured nonlinear component and $x_2(t) \in R^{n_2}$ is the partially measured linear one, $y(t) = [y_1(t), y_2(t)] \in R^m$ is the linear observation vector, such that the component $y_1(t) \in R^{n_1}$ corresponds the completely measured nonlinear state component $x_1(t) \in R^{n_1}$, i.e., the matrix $A_1(t) \in R^{n_1 \times n_1}$ is invertible, and $y_2(t) \in R^{m_2}$ corresponds to the partially measured linear component $x_2(t) \in R^{n_2}$, $m_2 \leq n_2$, i.e., the dimension of $y_2(t)$ may be less than that of $x_2(t)$. The initial condition $x_0 \in R^n$ is a Gaussian vector such that x_0, $W_1(t) = [W_{11}(t), W_{12}(t)]$, and $W_2(t) = [W_{21}(t), W_{22}(t)]$ are independent. It is assumed that $B(t)B^T(t)$, where $B(t) = diag[B_1(t), B_2(t)]$, is a positive definite matrix. All coefficients in (1)–(4) are deterministic functions of time of appropriate dimensions.

Without loss of generality, the observation process components $y_1(t)$ and $y_2(t)$ are assumed to be uncorrelated. Indeed, if $y_1(t)$ and $y_2(t)$ are correlated *a priori*, their mutual correlation can always be set to zero by adjusting terms $A_{01}(t)$ and $A_{02}(t)$ in the equations (3) and (4) (see [9]).

The nonlinear function $f(x_1, x_2, t)$ is considered a polynomial of n variables, components of the state vector $x(t) = [x_1(t), x_2(t)] \in R^n$, with time-dependent coefficients. Since $x(t) \in R^n$ is a vector, this requires a special definition of the polynomial for $n > 1$; some of them can be found in [6], [7]. In this paper, a p-degree polynomial of a vector $x(t) \in R^n$ is regarded as a p-linear form of n components of $x(t)$

$$f(x,t) = a_{10}(t) + a_{11}(t)x + a_{12}(t)xx^T + \quad (5)$$
$$\ldots + a_{1p}(t)x \ldots_{p \ times} \ldots x,$$

where $a_{10}(t)$ is a vector of dimension n, a_{11} is a matrix of dimension, a_{12} is a 3D tensor of dimension $n \times n \times n$, a_p is an $(p+1)$D tensor of dimension $n \times \ldots_{(p+1) \ times} \ldots \times n$, and $x \times \ldots_{p \ times} \ldots \times x$ is a pD tensor of dimension $n \times \ldots_{p \ times} \ldots \times n$ obtained by p times spatial multiplication of the vector $x(t)$ by itself. Such a polynomial can also be expressed in the summation form

$$dx_k(t)/dt = a_{10\ k}(t) + \sum_i a_{11\ ki}(t)x_i(t) +$$
$$+ \sum_{ij} a_{12\ kij}(t)x_i(t)x_j(t) + \ldots$$
$$+ \sum_{i_1 \ldots i_p} a_{1p\ ki_1 \ldots i_p}(t)x_{i_1}(t) \ldots x_{i_p}(t),$$
$$k, i, j, i_1 \ldots \imath_p = 1, n.$$

The estimation problem is to find the optimal estimate $\hat{x}(t) = [\hat{x}_1(t), \hat{x}_2(t)]$ of the system state $x(t) = [x_1(t), x_2(t)]$, based on the observation process $Y(t) = \{y(s) = [y_1(s), y_2(s)], 0 \leq s \leq t\}$, that minimizes the Euclidean 2-norm

$$J = E[(x(t) - \hat{x}(t))^T (x(t) - \hat{x}(t)) \mid F_t^Y]$$

at every time moment t. Here, $E[z(t) \mid F_t^Y]$ means the conditional expectation of a stochastic process $z(t) = (x(t) - \hat{x}(t))^T (x(t) - \hat{x}(t))$ with respect to the σ - algebra F_t^Y generated by the observation process $Y(t)$ in the interval $[t_0, t]$. As known [8], this optimal estimate is given by the conditional expectation

$$\hat{x}(t) = [\hat{x}_1(t), \hat{x}_2(t)] =$$
$$= m(t) = [m_1(t), m_2(t)] = E(x(t) \mid F_t^Y)$$

of the system state $x(t) = [x_1(t), x_2(t)]$ with respect to the σ - algebra F_t^Y generated by the observation process $Y(t)$ in the interval $[t_0, t]$.

As usual, the matrix function

$$P(t)=E[(x(t)-m(t))(x(t)-m(t))^T \mid F_t^Y]$$

is the estimation error variance.

The proposed solution to this optimal filtering problem is based on the formulas for the Ito differential of the conditional expectation $E(x(t) \mid F_t^Y)$ and its variance $P(t)$ (cited after [8]) and given in the following section.

III. Optimal Filter for Polynomial State with Partially Measured Linear Part over Linear Observations

The optimal filtering equations could be obtained using the formula for the Ito differential of the conditional expectation $m(t)=E(x(t) \mid F_t^Y)$ (see [8])

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

where $\bar{f}(x,t)=[f(x,t),a_{20}(t)+a_{21}(t)x_2(t)]$ is the polynomial drift term in the entire state equation, $f(x,t)$ is the polynomial drift term in the equation (1), and $\varphi_1(x)$ is the linear drift term in the entire observation equation equal to $\varphi_1(x)=A_0(t)+A(t)x(t)$, where $A_0(t)=[A_{01}(t),A_{02}(t)]$ and $A(t)=diag[A_1(t),A_2(t)]$. Upon performing substitution, the estimate equation takes the form

$$dm(t)=E(\bar{f}(x,t) \mid F_t^Y)dt+$$
$$E(x(t)[A(t)(x(t)-m(t))]^T \mid F_t^Y)\times$$
$$(B(t)B^T(t))^{-1}(dy(t)-(A_0(t)+A(t)m(t))=$$
$$E(\bar{f}(x,t) \mid F_t^Y)dt+E(x(t)(x(t)-m(t))^T \mid F_t^Y)A^T(t)\times$$
$$(B(t)B^T(t))^{-1}(dy(t)-(A_0(t)+A(t)m(t))dt)=$$
$$E(\bar{f}(x,t) \mid F_t^Y)dt+P(t)A^T(t)\times \tag{6}$$
$$(B(t)B^T(t))^{-1}(dy(t)-(A_0(t)+A(t)m(t))dt).$$

The equation (3) should be complemented with the initial condition $m(t_0)=E(x(t_0) \mid F_{t_0}^Y)$.

To compose a closed system of the filtering equations, the equation (6) should be complemented with the equation for the error variance $P(t)$. For this purpose, the formula for the Ito differential of the variance $P(t)=E((x(t)-m(t))(x(t)-m(t))^T \mid F_t^Y)$ could be used (cited again after [8]):

$$dP(t)=(E((x(t)-m(t))(\bar{f}(x,t))^T \mid F_t^Y)+$$
$$E(\bar{f}(x,t)(x(t)-m(t))^T) \mid F_t^Y)+b(t)b^T(t)-$$
$$E(x(t)[\varphi_1(x)-E(\varphi_1(x) \mid F_t^Y)]^T \mid F_t^Y)\left(B(t)B^T(t)\right)^{-1}\times$$
$$E([\varphi_1(x)-E(\varphi_1(x) \mid F_t^Y)]x^T(t) \mid F_t^Y))dt+$$
$$E((x(t)-m(t))(x(t)-m(t))[\varphi_1(x)-E(\varphi_1(x) \mid F_t^Y)]^T \mid F_t^Y)$$
$$\times\left(B(t)B^T(t)\right)^{-1}(dy(t)-E(\varphi_1(x) \mid F_t^Y)dt),$$

where the last term should be understood as a 3D tensor (under the expectation sign) convoluted with a vector, which yields a matrix. Upon substituting the expressions for φ_1, the last formula takes the form

$$dP(t)=(E((x(t)-m(t))(\bar{f}(x,t))^T \mid F_t^Y)+$$
$$E(\bar{f}(x,t)(x(t)-m(t))^T) \mid F_t^Y)+b(t)b^T(t)-$$
$$(E(x(t)(x(t)-m(t))^T \mid F_t^Y)A^T(t)(B(t)B^T(t))^{-1}\times$$
$$A(t)E((x(t)-m(t))x^T(t)) \mid F_t^Y))dt+$$
$$E((x(t)-m(t))(x(t)-m(t))\times$$
$$(A(t)(x(t)-m(t)))^T \mid F_t^Y)(B(t)B^T(t))^{-1}\times$$
$$(dy(t)-A(t)m(t))dt).$$

Using the variance formula $P(t)=E((x(t-h)-m(t-h))x^T(t)) \mid F_t^Y)$, the last equation can be represented as

$$dP(t)=(E((x(t)-m(t))(\bar{f}(x,t))^T \mid F_t^Y)+ \tag{7}$$
$$E(\bar{f}(x,t)(x(t)-m(t))^T) \mid F_t^Y)+b(t)b^T(t)-$$
$$P(t)A^T(t)(B(t)B^T(t))^{-1}A(t)P(t))dt+$$
$$E(((x(t)-m(t))(x(t)-m(t))(x(t)-m(t))^T \mid F_t^Y)\times$$
$$A^T(t)(B(t)B^T(t))^{-1}(dy(t)-A(t)m(t-h))dt).$$

The equation (7) should be complemented with the initial condition $P(t_0)=E[(x(t_0)-m(t_0)(x(t_0)-m(t_0)^T \mid F_{t_0}^Y]$.

The equations (6) and (7) for the optimal estimate $m(t)$ and the error variance $P(t)$ form a non-closed system of the filtering equations for the nonlinear state (1),(2) over linear observations (3),(4). Let us prove now that this system becomes a closed system of the filtering equations in view of the polynomial properties of the function $f(x,t)$ in the equation (1).

As shown in [6], [7], a closed system of the filtering equations for a polynomial state over linear observations can be obtained if the observation matrix $A(t)$ is invertible for any $t \geq t_0$. This condition implies [6], [7] that the random variable $x(t)-m(t)$ is conditionally Gaussian with respect to the observation process $y(t)$ for any $t \geq t_0$.

In the considered observation equations (3),(4), only the matrix $A_1(t)$ in (3) is invertible, whereas the matrix $A_2(t)$ in (4) is not. Nonetheless, the error variable components $x_1(t)-m_1(t)$, $m_1(t)=E[x_1(t) \mid F_t^Y]$, corresponding to $A_1(t)$ in (3) form a conditionally Gaussian vector with respect to the entire observation process $y(t)=[y_1(t),y_2(t)]$, since the observation process components $y_1(t)$ and $y_2(t)$ are uncorrelated (by assumption) and the innovations process $y_1(t)-\int_{t_0}^t(A_{01}(s)+A_1(s)m_1(s))ds$ is conditionally Gaussian with respect to $y_1(t)$ ([6], [7]). The error variable components $x_2(t)-m_2(t)$, $m_2(t)=E[x_2(t) \mid F_t^Y]$, corresponding to $A_2(t)$ in (4) also form a conditionally Gaussian vector with respect to the entire observation process $y(t)=[y_1(t),y_2(t)]$, since $x_2(t)$ is Gaussian and the observation process components $y_1(t)$ and $y_2(t)$ are uncorrelated. Thus, the entire vector

$x(t)-m(t)=[x_1(t)-m_1(t),x_2(t)-m_2(t)]$ is conditionally Gaussian with respect to the entire observation process $y(t)=[y_1(t),y_2(t)]$ (see [9]), and, therefore, the following considerations outlined in [6], [7] are applicable.

First, since the random variable $x(t)-m(t)$ is conditionally Gaussian, the conditional third moment $E((x(t)-m(t))(x(t)-m(t))(x(t)-m(t))^T \mid F_t^Y)$ of $x(t)-m(t)$ with respect to observations, which stands in the last term of the equation (7), is equal to zero, because the process $x(t)-m(t)$ is conditionally Gaussian. Thus, the entire last term in (7) is vanished and the following variance equation is obtained

$$dP(t)=(E((x(t)-m(t))(\bar{f}(x,t))^T \mid F_t^Y)+ \tag{8}$$
$$E(\bar{f}(x,t)(x(t)-m(t))^T) \mid F_t^Y)+b(t)b^T(t)-$$
$$P(t)A^T(t)(B(t)B^T(t))^{-1}A(t)P(t))dt,$$

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

Second, if the function $\bar{f}(x,t)$ is a polynomial function of the state x with time-dependent coefficients, the expressions of the terms $E(\bar{f}(x,t) \mid F_t^Y)$ in (6) and $E((x(t)-m(t))(\bar{f}(x,t))^T \mid F_t^Y)$ in (8) would also include only polynomial terms of x. Then, those polynomial terms can be represented as functions of $m(t)$ and $P(t)$ using the following property of Gaussian random variable $x(t)-m(t)$: all its odd conditional moments, $m_1=E[(x(t)-m(t)) \mid Y(t)], m_3=E[(x(t)-m(t)^3 \mid Y(t)], m_5=E[(x(t)-m(t))^5 \mid Y(t)]$, ... are equal to 0, and all its even conditional moments $m_2=E[(x(t)-m(t))^2 \mid Y(t)], m_4=E[(x(t)-m(t))^4 \mid Y(t)]$,.... can be represented as functions of the variance $P(t)$. For example, $m_2=P, m_4=3P^2, m_6=15P^3, \ldots$ etc. After representing all polynomial terms in (6) and (8), that are generated upon expressing $E(\bar{f}(x,t) \mid F_t^Y)$ and $E((x(t)-m(t))(\bar{f}(x,t))^T \mid F_t^Y)$, as functions of $m(t)$ and $P(t)$, a closed form of the filtering equations would be obtained. The corresponding representations of $E(f(x,t) \mid F_t^Y)$ and $E((x(t)-m(t))(f(x,t))^T \mid F_t^Y)$ have been derived in [6], [7] for certain polynomial functions $f(x,t)$.

In the next subsection, a closed form of the filtering equations will be obtained from (6) and (8) for a bilinear function $f(x,t)$ in the equation (1). It should be noted, however, that application of the same procedure would result in designing a closed system of the filtering equations for any polynomial function $f(x,t)$ in (1).

A. Optimal Filter for Bilinear State with Partially Measured Linear Part over Linear Observations

In a particular case, if the function

$$f(x,t)=a_{10}(t)+a_{11}(t)x+a_{12}(t)xx^T \tag{9}$$

is a bilinear polynomial, where x is an n-dimensional vector, $a_{10}(t)$ is an n_1-dimensional vector, a_{11} is an $n_1 \times n$ - matrix, and a_{12} is a 3D tensor of dimension $n_1 \times n \times n$, the representations for $E(\bar{f}(x,t) \mid F_t^Y)$ and $E((x(t)-m(t))(\bar{f}(x,t))^T \mid F_t^Y)$ as functions of $m(t)$ and $P(t)$ are derived as follows (see [6])

$$E(\bar{f}(x,t) \mid F_t^Y)=a_0(t)+a_1(t)m(t)+ \tag{10}$$
$$a_2(t)m(t)m^T(t)+a_2(t)P(t),$$
$$E(\bar{f}(x,t)(x(t)-m(t))^T) \mid F_t^Y)+ \tag{11}$$
$$E((x(t)-m(t))(\bar{f}(x,t))^T \mid F_t^Y)=$$
$$a_1(t)P(t)+P(t)a_1^T(t)+$$
$$2a_2(t)m(t)P(t)+2(a_2(t)m(t)P(t))^T.$$

where $a_0(t)=[a_{10}(t),a_{20}(t)]$, $a_1(t)=diag[a_{11}(t),a_{21}(t)]$, and a_2 is a 3D tensor of dimension $n \times n \times n$ defined as $a_{2\ kij}=a_{12\ kij}$, if $k \le n_1$, and $a_{2\ kij}=0$, otherwise.

Substituting the expression (10) in (6) and the expression (11) in (8), the filtering equations for the optimal estimate $m(t)=[m_1(t),m_2(t)]$ of the bilinear state $x(t)=[x_1(t),x_2(t)]$ and the error variance $P(t)$ are obtained

$$dm(t)=(a_0(t)+a_1(t)m(t)+ \tag{12}$$
$$a_2(t)m(t)m^T(t)+a_2(t)P(t))dt+$$
$$P(t)A^T(t)(B(t)B^T(t))^{-1}[dy(t)-A(t)m(t)dt],$$
$$m(t_0)=E(x(t_0) \mid F_t^Y)),$$
$$dP(t)=(a_1(t)P(t)+P(t)a_1^T(t)+ \tag{13}$$
$$2a_2(t)m(t)P(t)+2(a_2(t)m(t)P(t))^T+$$
$$b(t)b^T(t))dt-P(t)A^T(t)(B(t)B^T(t))^{-1}A(t)P(t)dt.$$
$$P(t_0)=E((x(t_0)-m(t_0))(x(t_0)-m(t_0))^T \mid F_t^Y)),$$

where $A_0(t)=[A_{01}(t),A_{02}(t)]$, $A(t)=diag[A_1(t),A_2(t)]$, and $B(t)=diag[B_1(t),B_2(t)]$.

By means of the preceding derivation, the following result is proved.

Theorem 1. The optimal finite-dimensional filter for the bilinear state (1),(2) with partially unmeasured linear part (2), where the quadratic polynomial $f(x,t)$ is defined by (9), over the linear observations (3),(4) is given by the equation (12) for the optimal estimate $m(t)=E(x(t) \mid F_t^Y)$ and the equation (13) for the estimation error variance $P(t)=E[(x(t)-m(t))(x(t)-m(t))^T \mid F_t^Y]$.

Thus, based on the general non-closed system of the filtering equations (6),(7), it is proved that the closed system of the filtering equations (6),(8) can be obtained for any polynomial state (1),(2) with partially measured linear part (2) over linear observations (3),(4). Furthermore, the specific form (12),(13) of the closed system of the filtering equations corresponding to a bilinear state is derived. In the next section, performance of the designed optimal filter for a bilinear state with partially measured linear part over linear observations is verified against a conventional extended Kalman-Bucy filter.

IV. EXAMPLE

This section presents an example of designing the optimal filter for a quadratic-linear state with unmeasured linear part over linear observations and comparing it to a conventional extended Kalman-Bucy filter.

Let the unmeasured state $x(t) = [x_1(t), x_2(t)]$ satisfies the quadratic and linear equations

$$\dot{x}_1(t) = 0.1(x_1^2(t) + x_2(t)), \quad x_1(0) = x_{10}, \tag{14}$$

$$\dot{x}_2(t) = 0.1x_2(t), \quad x_2(0) = x_{20}, \tag{15}$$

and the observation process is given by the linear equation

$$y(t) = x_1(t) + \psi(t), \tag{16}$$

where $\psi(t)$ is a white Gaussian noise, which is the weak mean square derivative of a standard Wiener process (see [8]). The equations (14)–(16) present the conventional form for the equations (1)–(3), which is actually used in practice [10]. Note that only the state quadratic part (14) is measured, whereas the linear part (15) is not measured at all.

The filtering problem is to find the optimal estimate for the quadratic-linear state (14)–(15), using linear observations (16) confused with independent and identically distributed disturbances modelled as white Gaussian noises. Let us set the filtering horizon time to $T = 6.5$.

The filtering equations (6),(8) take the following particular form for the system (14)–(16)

$$\dot{m}_1(t) = 0.1(m_1^2(t) + P_{11}(t) + m_2(t)) + \tag{17}$$

$$P_{11}(t)[y(t) - m_1(t)],$$

$$\dot{m}_2(t) = 0.1m_2(t) + P_{12}(t)[y(t) - m_1(t)], \tag{18}$$

with the initial conditions $m_1(0) = [E(x_1(0) \mid y(0)) = m_{10}$ and $m_2(0) = [E(x_2(0) \mid y(0)) = m_{20}$,

$$\dot{P}_{11}(t) = 0.4(P_{11}(t)m_1(t)) + 0.2P_{12}(t) - P_{11}^2(t), \tag{19}$$

$$\dot{P}_{12}(t) = 0.2(P_{12}(t)m_1(t)) +$$

$$0.1(P_{12}(t) + P_2(t)) - P_{11}(t)P_{12}(t),$$

$$\dot{P}_{22}(t) = 0.2P_{22}(t) - P_{12}^2(t),$$

with the initial condition
$P(0) = E((x(0) - m(0))(x(0) - m(0))^T \mid y(0)) = P_0$.

The estimates obtained upon solving the equations (17)–(19) are compared to the estimates satisfying the following extended Kalman-Bucy filtering equations for the quadratic-linear state (14)–(15) over the linear observations (16), obtained using the direct copy of the state dynamics (14)–(15) in the estimate equation and assigning the filter gain as the solution of the Riccati equation:

$$\dot{m}_{K1}(t) = 0.1(m_{K1}^2(t) + m_{K2}(t)) + \tag{20}$$

$$P_{K11}(t)[y(t) - m_{K1}(t)],$$

$$\dot{m}_{K2}(t) = 0.1m_{K2}(t) + P_{K12}(t)[y(t) - m_{K1}(t)], \tag{21}$$

with the initial conditions $m_{K1}(0) = [E(x_1(0) \mid y(0)) = m_{10}$ and $m_{K2}(0) = [E(x_2(0) \mid y(0)) = m_{20}$,

$$\dot{P}_{K11}(t) = 0.2P_{K12}(t) - P_{K11}^2(t), \tag{22}$$

$$\dot{P}_{K12}(t) = 0.1(P_{K12}(t) + P_{K2}(t)) - P_{K11}(t)P_{K12}(t),$$

$$\dot{P}_{K22}(t) = 0.2P_{K22}(t) - P_{K12}^2(t),$$

with the initial condition
$P_K(0) = E((x(0) - m(0))(x(0) - m(0))^T \mid y(0)) = P_0$.

Numerical simulation results are obtained solving the systems of filtering equations (17)–(19) and (20)–(22). The obtained values of the estimates $m(t)$ and $m_K(t)$ satisfying the equations (17)–(18) and (20)–(21), respectively, are compared to the real values of the state variables $x(t)$ in (14)–(15).

For each of the two filters (17)–(19) and (20)–(22) and the reference system (14)–(15) involved in simulation, the following initial values are assigned: $x_{10} = x_{20} = 1.1$, $m_{10} = m_{20} = 0.1$, $P_{011} = P_{012} = P_{022} = 1$. Gaussian disturbance $\psi(t)$ in (16) is realized using the built-in MatLab white noise function.

The following graphs are obtained: graphs of the reference state variables $x_1(t)$ and $x_2(t)$ for the system (14)–(15); graphs of the optimal filter estimates $m_1(t)$ and $m_2(t)$ satisfying the equations (17)–(19); graphs of the estimates $m_{K1}(t)$ and m_{K2} satisfying the equations (20)–(22). The graphs of all those variables are shown on the entire simulation interval from $t_0 = 0$ to $T = 6.5$ (Fig. 1). Note that the gain matrix entry $P_{11}(t)$ does not converge to zero as time tends to infinity, since the polynomial dynamics of third order is stronger than the quadratic Riccati terms in the right-hand sides of the equation (19).

The following values of the reference state variables $x_1(t)$, $x_2(t)$ and the estimates $m_1(t)$, $m_2(t)$, $m_{K1}(t)$, $m_{K2}(t)$ are obtained at the reference time points $T = 5, 6, 6.5$: for $T = 5$, $x_1(5) = 6.05$, $m_1(5) = 5.69$, $m_{k1}(5) = 4.40$, $x_2(5) = 1.81$, $m_2(5) = 1.75$, $m_{K2}(5) = 2.35$; for $T = 6$, $x_1(6) = 16.05$, $m_1(6) = 16.42$, $m_{k1}(6) = 9.99$, $x_2(6) = 2.00$, $m_2(6) = 1.97$, $m_{K2}(6) = 3.47$; for $T = 6.5$, $x_1(6.5) = 96.99$, $m_1(6.5) = 96.68$, $m_{k1}(6.5) = 25.33$, $x_2(6.5) = 2.11$, $m_2(6.5) = 2.09$, $m_{K2}(6.5) = 6.26$.

Thus, it can be concluded that the obtained optimal filter (17)–(19) for a quadratic-linear state with unmeasured linear part over linear observations yield definitely better estimates than the conventional extended Kalman-Bucy filter (20)–(22). Subsequent discussion of the obtained simulation results can be found in Conclusions.

V. CONCLUSIONS

The simulation results show that the values of the estimate calculated by using the obtained optimal filter for a quadratic-linear state with unmeasured linear part over linear observations are noticeably closer to the real values of the reference variable than the values of the estimate given by the conventional extended Kalman-Bucy filter. Moreover, it can be seen that the estimate produced by the optimal filter asymptotically converges to the real values of the reference variables as time tends to infinity, although the reference system (14)–(15) itself is unstable and the nonlinear component $x_1(t)$ goes to infinity for a finite time. On the contrary, the conventionally designed extended Kalman-Bucy estimates diverge from the real values. This significant improvement in the estimate behavior is obtained due to the more careful selection of the filter gain matrix in the equations (17)–(19), as it should be in the optimal filter. Although this conclusion follows from the developed theory, the numerical simulation serves as a convincing illustration.

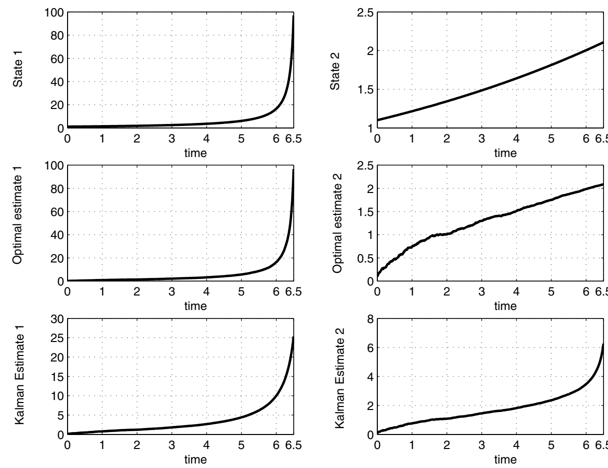

Fig. 1. Graphs of the reference state variables $x_1(t)$ (State 1) and x_2 (State 2) satisfying the equations (14)–(15), graphs of the optimal filter estimates $m_1(t)$ (Optimal estimate 1) and $m_2(t)$ (Optimal estimate 2) satisfying the equations (17)–(18), and graphs of the estimates $m_{K1}(t)$ (Kalman estimate 1) and $m_{K2}(t)$ (Kalman estimate 2) satisfying the equations (20)–(21), on the entire simulation interval $[0, 6.5]$.

REFERENCES

[1] H. J. Kushner, "On differential equations satisfied by conditional probability densities of Markov processes," *SIAM J. Control*, vol. 2, pp. 106–119, 1964.

[2] R. E. Kalman and R. S. Bucy, "New results in linear filtering and prediction theory," *ASME Journal of Basic Engineering, Ser. D*, vol. 83, pp. 95–108, 1961.

[3] W. Wonham, "Some applications of stochastic differential equations to nonlinear filtering," *SIAM J. Control*, vol. 2, pp. 347–369, 1965.

[4] V. Benes, "Exact finite-dimensional filters for certain diffusions with nonlinear drift," *Stochastics*, vol. 5, pp. 65–92, 1981.

[5] S.-T. Yau, "Finite-dimensional filters with nonlinear drift I: a class of filters including both Kalman-Bucy and Benes filters," *J. Math. Systems, Estimation, and Control*, vol. 4, pp. 181–203, 1994.

[6] M. V. Basin, "Optimal filtering for polynomial system states," *ASME Transactions. J. of Dynamic Systems, Measurement, and Control*, vol. 125, pp. 123–125, 2003.

[7] M. V. Basin and M. A. Alcorta-Garcia, "Optimal filtering and control for third degree polynomial system," *Dynamics of Continuous, Discrete, and Impulsive Systems*, vol. 10B, pp. 663–680, 2003.

[8] V. S. Pugachev and I. N. Sinitsyn, *Stochastic Systems: Theory and Applications*. Singapore: World Scientific, 2001.

[9] V. S. Pugachev, *Probability Theory and Mathematical Statistics for Engineers*. London: Pergamon Press, 1984.

[10] K. J. Åström, *Introduction to Stochastic Control Theory*, ser. Mathematics in Science and Engineering. New York: Academic Press, 1970, vol. 70.