INTERNATIONAL JOURNAL OF ROBUST AND NONLINEAR CONTROL
Int. J. Robust Nonlinear Control 2006; **16**:303–314
Published online 23 January 2006 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/rnc.1055

Optimal filtering for polynomial system states with polynomial multiplicative noise

Michael Basin[1,*,†], Joel Perez[1,‡] and Mikhail Skliar[2,§]

[1] *Department of Physical and Mathematical Sciences, Autonomous University of Nuevo Leon, Apdo postal 144-F, C.P. 66450, San Nicolas de los Garza, Nuevo Leon, Mexico*
[2] *Department of Chemical and Fuels Engineering, University of Utah, 50 S. Central Campus Dr., Salt Lake City, UT 84112, U.S.A.*

SUMMARY

In this paper, the optimal filtering problem for polynomial system states with polynomial multiplicative noise 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 polynomial multiplicative noise over linear observations is then established, which yields the explicit closed form of the filtering equations in the particular cases of a linear state equation with linear multiplicative noise and a bilinear state equation with bilinear multiplicative noise. In the example, performance of the designed optimal filter is verified for a quadratic state with a quadratic multiplicative noise over linear observations against the optimal filter for a quadratic state with a state-independent noise and a conventional extended Kalman–Bucy filter. Copyright © 2006 John Wiley & Sons, Ltd.

KEY WORDS: filtering; stochastic system; nonlinear polynomial system; polynomial multiplicative noise; bilinear system

1. 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

*Correspondence to: Michael Basin, Department of Physical and Mathematical Sciences, Autonomous University of Nuevo Leon, Apdo postal 144-F, C.P. 66450, San Nicolas de los Garza, Nuevo Leon, Mexico.
†E-mail: mbasin@fcfm.uanl.mx
‡E-mail: joelperezp@yahoo.com
§E-mail: Mikhail.Skliar@m.cc.utah.edu

Contract/grant sponsor: U.S. National Science Foundation; contract/grant number: CTS-0117300
Contract/grant sponsor: Mexican National Science and Technology Council; contract/grant number: 39388-A

Received 29 August 2005
Revised 14 November 2005
Accepted 5 December 2005

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 $\mathrm{d}f/\mathrm{d}x + f^2 = x^2$ (see Reference [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 Reference [5]. The last two papers actually deal with specific types of polynomial filtering systems. There also exists a considerable bibliography on robust filtering for the 'general situation' systems (see, for example, References [6–9]). Apart form the 'general situation,' the optimal finite-dimensional filters have recently been designed [10, 11] 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 system states with polynomial multiplicative noise over linear observations with invertible observation matrix, thus generalizing the results of References [10–12] obtained for polynomial system states with state-independent noise. Designing the optimal filter with polynomial multiplicative noise presents a significant advantage in the filtering theory and practice, since it enables one to address filtering problems with polynomial observation nonlinearities, such as the optimal cubic sensor problem [13]. The optimal filtering problem is treated proceeding from the general expression for the stochastic Ito differential of the optimal estimate and the error variance [14]. 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 for a polynomial state equation with polynomial multiplicative noise and linear observations with invertible observation matrix. 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 cases of a linear state equation with linear multiplicative noise and a bilinear state equation with bilinear multiplicative noise.

In the illustrative example, performance of the designed optimal filter is verified for a quadratic state with a quadratic multiplicative noise over linear observations against the optimal filter for a quadratic state with a state-independent noise and a conventional extended Kalman–Bucy filter. The simulation results show a definite advantage of the designed optimal filter in regard to proximity of the estimate to the real state value. Moreover, it can be seen that the estimation error produced by the optimal filter rapidly reaches and then maintains the zero mean value even in a close vicinity of the asymptotic time point, although the system state itself is unstable and the quadratic component goes to infinity for a finite time. On the contrary, the estimation errors given by the other two applied filters diverge to infinity near the asymptotic time point.

The paper is organized as follows. Section 2 presents the filtering problem statement for a polynomial system state with multiplicative noise over linear observations. The Ito differentials for the optimal estimate and the error variance are derived in Section 3. Section 3 also establishes the procedure for obtaining a closed system of the filtering equations for any polynomial state, which yields the explicit results for linear and bilinear state equations. Performance of the obtained optimal filter is verified in Section 4.

2. FILTERING PROBLEM FOR POLYNOMIAL STATE WITH MULTIPLICATIVE NOISE OVER LINEAR OBSERVATIONS

Let (Ω, F, P) be a complete probability space with an increasing right-continuous family of σ-algebras $F_t, t \geqslant t_0$, and let $(W_1(t), F_t, t \geqslant t_0)$ and $(W_2(t), F_t, t \geqslant t_0)$ be independent Wiener processes. The F_t-measurable random process $(x(t), y(t))$ is described by a nonlinear differential equation with both polynomial drift and diffusion terms for the system state

$$dx(t) = f(x, t)\,dt + g(x, t)\,dW_1(t), \quad x(t_0) = x_0 \tag{1}$$

and a linear differential equation for the observation process

$$dy(t) = (A_0(t) + A(t)x(t))\,dt + B(t)\,dW_2(t) \tag{2}$$

Here, $x(t) \in R^n$ is the state vector and $y(t) \in R^n$ is the linear observation vector, such that the matrix $A(t) \in R^{n \times n}$ is invertible. The initial condition $x_0 \in R^n$ is a Gaussian vector such that x_0, $W_1(t)$, and $W_2(t)$ are independent. It is assumed that $B(t)B^{\mathrm{T}}(t)$ is a positive definite matrix. All coefficients in (1)–(2) are deterministic functions of time of appropriate dimensions. The nonlinear diffusion function $g(x, t)$ forms a state-dependent multiplicative noise in the state equation (1).

The nonlinear functions $f(x, t)$ and $g(x, t)$ are considered polynomials of n variables, components of the state vector $x(t) \in R^n$, 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 References [10–12]. 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 $x(t)$

$$f(x, t) = a_0(t) + a_1(t)x + a_2(t)xx^{\mathrm{T}} + \cdots + a_p(t)x \ldots_{p \text{ times}} \ldots x \tag{3}$$

where $a_0(t)$ is a vector of dimension n, a_1 is a matrix of dimension, a_2 is a 3D tensor of dimension $n \times n \times n$, a_p is an $(p+1)$D tensor of dimension $n \times \ldots_{(p+1) \text{ times}} \ldots \times n$, and $x \times \ldots_{p \text{ times}} \ldots \times x$ is a pD tensor of dimension $n \times \ldots_{p \text{ times}} \ldots \times n$ obtained by p times spatial multiplication of the vector $x(t)$ by itself. Such a polynomial can also be expressed in the summation form

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

The estimation problem is to find the optimal estimate $\hat{x}(t)$ of the system state $x(t)$, based on the observation process $Y(t) = \{y(s), t_0 \leqslant s \leqslant t\}$, that minimizes the Euclidean 2-norm

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

at every time moment t. Here, $E[z(t)|F_t^Y]$ means the conditional expectation of a stochastic process $z(t) = (x(t) - \hat{x}(t))^{\mathrm{T}}(x(t) - \hat{x}(t))$ with respect to the σ-algebra F_t^Y generated by the observation process $Y(t)$ in the interval $[t_0, t]$. As known [14], this optimal estimate is given by the conditional expectation

$$\hat{x}(t) = m(t) = E(x(t)|F_t^Y)$$

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

$$P(t) = E[(x(t) - m(t))(x(t) - m(t))^{\mathrm{T}}|F_t^Y]$$

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 $E(x(t)|F_t^Y)$ and its variance $P(t)$ (cited after Reference [14]) and given in the following section.

3. OPTIMAL FILTER FOR POLYNOMIAL STATE WITH MULTIPLICATIVE NOISE OVER LINEAR OBSERVATIONS

The optimal filtering equations could be obtained using the formula for the Ito differential of the conditional expectation $m(t) = E(x(t)|F_t^Y)$ (see Reference [14])

$$\mathrm{d}m(t) = E(f(x,t)|F_t^Y)\,\mathrm{d}t + E(x[\varphi_1(x) - E(\varphi_1(x)|F_t^Y)]^{\mathrm{T}}|F_t^Y)$$
$$\times (B(t)B^{\mathrm{T}}(t))^{-1}(\mathrm{d}y(t) - E(\varphi_1(x)|F_t^Y)\,\mathrm{d}t)$$

where $f(x,t)$ is the polynomial drift term in the state equation, and $\varphi_1(x)$ is the linear drift term in the observation equation equal to $\varphi_1(x,t) = A_0(t) + A(t)x(t)$. Upon performing substitution, the estimate equation takes the form

$$\begin{aligned}\mathrm{d}m(t) &= E(f(x,t)|F_t^Y)\,\mathrm{d}t + E(x(t)[A(t)(x(t) - m(t))]^{\mathrm{T}}|F_t^Y)\\ &\quad\times (B(t)B^{\mathrm{T}}(t))^{-1}(\mathrm{d}y(t) - (A_0(t) + A(t)m(t))\\ &= E(f(x,t)|F_t^Y)\,\mathrm{d}t + E(x(t)(x(t) - m(t))^{\mathrm{T}}|F_t^Y)A^{\mathrm{T}}(t)\\ &\quad\times (B(t)B^{\mathrm{T}}(t))^{-1}(\mathrm{d}y(t) - (A_0(t) + A(t)m(t))\,\mathrm{d}t)\\ &= E(f(x,t)|F_t^Y)\,\mathrm{d}t + P(t)A^{\mathrm{T}}(t)(B(t)B^{\mathrm{T}}(t))^{-1}(\mathrm{d}y(t) - (A_0(t) + A(t)m(t))\,\mathrm{d}t)\end{aligned} \tag{4}$$

Equation (4) should be complemented with the initial condition $m(t_0) = E(x(t_0)|F_{t_0}^Y)$.

Trying to compose a closed system of the filtering equations, Equation (4) should be complemented with the equation for the error variance $P(t)$. For this purpose, the formula for the Ito differential of the variance $P(t) = E((x(t) - m(t))(x(t) - m(t))^{\mathrm{T}}|F_t^Y)$ could be used (cited again after Reference [14])

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

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

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

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

$$\begin{aligned} dP(t) = {} & (E((x(t)-m(t))(f(x,t))^{\mathrm{T}}|F_t^Y) + E(f(x,t)(x(t)-m(t))^{\mathrm{T}})|F_t^Y) \\ & + E(g(x,t)g^{\mathrm{T}}(x,t)|F_t^Y) - P(t)A^{\mathrm{T}}(t)(B(t)B^{\mathrm{T}}(t))^{-1}A(t)P(t))\,dt \\ & + E(((x(t)-m(t))(x(t)-m(t))(x(t)-m(t))^{\mathrm{T}}|F_t^Y) \\ & \times A^{\mathrm{T}}(t)(B(t)B^{\mathrm{T}}(t))^{-1}(dy(t) - A(t)m(t-h))\,dt) \end{aligned} \tag{5}$$

Equation (5) should be complemented with the initial condition $P(t_0) = E[(x(t_0) - m(t_0)(x(t_0) - m(t_0)^{\mathrm{T}}|F_{t_0}^Y]$.

Equations (4) and (5) 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) over linear observations (2). The non-closeness means that the system (4), (5) includes terms depending on 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 not expressed yet as functions of the system variables, $m(t)$ and $P(t)$. Let us prove now that this system becomes a closed system of the filtering equations in view of the polynomial properties of the functions $f(x,t)$ and $g(x,t)$ in Equation (1).

As shown in References [10, 11], a closed system of the filtering equations for a system state with polynomial drift and state-independent diffusion over linear observations can be obtained if the observation matrix $A(t)$ is invertible for any $t \geqslant t_0$. The last condition, also assumed for the observation process (2), implies [10, 11] that the random variable $x(t) - m(t)$ is conditionally Gaussian with respect to the observation process $y(t)$ for any $t \geqslant t_0$. Hence, the following considerations outlined in References [10, 11] are applicable to the filtering equations (4), (5).

First, since the random variable $x(t) - m(t)$ is conditionally Gaussian, the conditional third moment $E((x(t) - m(t))(x(t) - m(t))(x(t) - m(t))^{\mathrm{T}}|F_t^Y)$ of $x(t) - m(t)$ with respect to observations, which stands in the last term of Equation (5), is equal to zero, because the process $x(t) - m(t)$ is conditionally Gaussian. Thus, the entire last term in (5) is vanished and the following variance equation is obtained:

$$\mathrm{d}P(t) = (E((x(t) - m(t))(f(x,t))^{\mathrm{T}}|F_t^Y) + E(f(x,t)(x(t) - m(t))^{\mathrm{T}})|F_t^Y) + E(g(x,t)g^{\mathrm{T}}(x,t)|F_t^Y) - P(t)A^{\mathrm{T}}(t)(B(t)B^{\mathrm{T}}(t))^{-1}A(t)P(t))\,\mathrm{d}t \tag{6}$$

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, if the functions $f(x,t)$ and $g(x,t)$ are polynomial functions of the state x with time-dependent coefficients, the expressions of the terms $E(f(x,t)|F_t^Y)$ in (4) and $E((x(t) - m(t))f^{\mathrm{T}}(x,t))|F_t^Y)$ and $E(g(x,t)g^{\mathrm{T}}(x,t)|F_t^Y)$ in (6) 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, $m_1 = E[(x(t) - m(t))|Y(t)], m_3 = E[(x(t) - m(t)^3|Y(t)], m_5 = E[(x(t) - m(t))^5|Y(t)], \ldots$ are equal to 0, and all its even conditional moments $m_2 = E[(x(t) - m(t))^2|Y(t)], m_4 = E[(x(t) - m(t))^4|Y(t)], \ldots$ 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 (4) and (6), that are generated upon expressing $E(f(x,t)|F_t^Y)$, $E((x(t) - m(t))f^{\mathrm{T}}(x,t))|F_t^Y)$, and $E(g(x,t)g^{\mathrm{T}}(x,t)|F_t^Y)$, as functions of $m(t)$ and $P(t)$, a closed form of the filtering equations would be obtained. The corresponding representations of $E(f(x,t)|F_t^Y)$ and $E((x(t) - m(t))(f(x,t))^{\mathrm{T}}|F_t^Y)$ have been derived in References [10, 11] for certain polynomial functions $f(x,t)$.

In the next subsections, a closed form of the filtering equations will be obtained from (4) and (6) for linear and bilinear functions $f(x,t)$ and $g(x,t)$ in 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 functions $f(x,t)$ and $g(x,t)$ in (1).

3.1. Optimal filter for linear state with linear multiplicative noise

In a particular case, if the functions $f(x,t) = a_0(t) + a_1(t)x(t)$ and $g(x,t) = b_0(t) + b_1(t)x(t)$ are linear, where b_1 is a 3D tensor of dimension $n \times n \times n$, the representations for $E(f(x,t)|F_t^Y)$, $E((x(t) - m(t))(f(x,t))^{\mathrm{T}}|F_t^Y)$, and $E(g(x,t)g^{\mathrm{T}}(x,t)|F_t^Y)$ as functions of $m(t)$ and $P(t)$ are derived as follows:

$$E(f(x,t)|F_t^Y) = a_0(t) + a_1(t)m(t) \tag{7}$$

$$E(f(x,t)(x(t) - m(t))^{\mathrm{T}})|F_t^Y) + E((x(t) - m(t))(f(x,t))^{\mathrm{T}}|F_t^Y) = a_1(t)P(t) + P(t)a_1^{\mathrm{T}}(t) \tag{8}$$

$$E(g(x,t)g^{\mathrm{T}}(x,t)|F_t^Y) = b_0(t)b_0^{\mathrm{T}}(t) + b_0(t)(b_1(t)m(t))^{\mathrm{T}} + (b_1(t)m(t))b_0^{\mathrm{T}}(t) + b_1(t)P(t)b_1^{\mathrm{T}}(t) + b_1(t)m(t)m^{\mathrm{T}}(t)b_1^{\mathrm{T}}(t) \tag{9}$$

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

Substituting expression (7) in (4) and the expressions (8), (9) in (6), the filtering equations for the optimal estimate $m(t)$ and the error variance $P(t)$ are obtained

$$\begin{aligned} \mathrm{d}m(t) &= (a_0(t) + a_1(t)m(t))\,\mathrm{d}t \\ &\quad + P(t)A^{\mathrm{T}}(t)(B(t)B^{\mathrm{T}}(t))^{-1}[\mathrm{d}y(t) - A(t)m(t)\,\mathrm{d}t], \quad m(t_0) = E(x(t_0)|F_t^Y)) \end{aligned} \tag{10}$$

$$\begin{aligned} \mathrm{d}P(t) &= (a_1(t)P(t) + P(t)a_1^{\mathrm{T}}(t) \\ &\quad + b_0(t)b_0^{\mathrm{T}}(t) + b_0(t)(b_1(t)m(t))^{\mathrm{T}} + (b_1(t)m(t))b_0^{\mathrm{T}}(t) + b_1(t)P(t)b_1^{\mathrm{T}}(t) \\ &\quad + b_1(t)m(t)m^{\mathrm{T}}(t)b_1^{\mathrm{T}}(t))\,\mathrm{d}t - P(t)A^{\mathrm{T}}(t)(B(t)B^{\mathrm{T}}(t))^{-1}A(t)P(t)\,\mathrm{d}t \end{aligned} \tag{11}$$

$$P(t_0) = E((x(t_0) - m(t_0))(x(t_0) - m(t_0))^{\mathrm{T}}|F_t^Y))$$

Note that the observation matrix $A(t)$ should not even be necessarily invertible to obtain the filtering equations (10)–(11). 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 variable with finite second moments, not only Gaussian.

3.2. Optimal filter for bilinear state with bilinear multiplicative noise

Let the functions

$$f(x,t) = a_0(t) + a_1(t)x + a_2(t)xx^{\mathrm{T}} \tag{12}$$

and

$$g(x,t) = b_0(t) + b_1(t)x + b_2(t)xx^{\mathrm{T}} \tag{13}$$

be bilinear polynomials, where x is an n-dimensional vector, $a_0(t)$ is an n-dimensional vector, $a_1(t)$ and $b_0(t)$ are $n \times n$-matrices, $a_2(t)$ and $b_1(t)$ are 3D tensors of dimension $n \times n \times n$, and $b_2(t)$ is a 4D tensor of dimension $n \times n \times n \times n$. In this case, the representations for $E(f(x,t)|F_t^Y)$, $E((x(t) - m(t))(f(x,t))^{\mathrm{T}}|F_t^Y)$, and $E(g(x,t)g^{\mathrm{T}}(x,t)|F_t^Y)$ as functions of $m(t)$ and $P(t)$ are derived as follows (see References [10, 11]):

$$E(f(x,t)|F_t^Y) = a_0(t) + a_1(t)m(t) + a_2(t)m(t)m^{\mathrm{T}}(t) + a_2(t)P(t) \tag{14}$$

$$\begin{aligned} &E(f(x,t)(x(t) - m(t))^{\mathrm{T}})|F_t^Y) + E((x(t) - m(t))(f(x,t))^{\mathrm{T}}|F_t^Y) \\ &\quad = a_1(t)P(t) + P(t)a_1^{\mathrm{T}}(t) + 2a_2(t)m(t)P(t) + 2(a_2(t)m(t)P(t))^{\mathrm{T}} \end{aligned} \tag{15}$$

$$\begin{aligned} E(g(x,t)g^{\mathrm{T}}(x,t)|F_t^Y) &= b_0(t)b_0^{\mathrm{T}}(t) + b_0(t)(b_1(t)m(t))^{\mathrm{T}} \\ &\quad + (b_1(t)m(t))b_0^{\mathrm{T}}(t) + b_1(t)P(t)b_1^{\mathrm{T}}(t) + b_1(t)m(t)m^{\mathrm{T}}(t)b_1^{\mathrm{T}}(t) \\ &\quad + b_0(t)(P(t) + m(t)m^{\mathrm{T}}(t))b_2^{\mathrm{T}}(t) + b_2(t)(P(t) + m(t)m^{\mathrm{T}}(t))b_0^{\mathrm{T}}(t) \\ &\quad + b_1(t)(3m(t)P(t) + m(t)(m(t)m^{\mathrm{T}}(t)))b_2^{\mathrm{T}}(t) + b_2(t)(3P(t)m^{\mathrm{T}}(t) \end{aligned}$$

$$+ (m(t)m^{\mathrm{T}}(t))m^{\mathrm{T}}(t))b_1^{\mathrm{T}}(t) + 3b_2(t)P^2(t)b_2^{\mathrm{T}}(t) + 3b_2(t)(P(t)m(t)m^{\mathrm{T}}(t)$$
$$+ m(t)m^{\mathrm{T}}(t)P(t))b_2^{\mathrm{T}}(t) + b_2(t)(m(t)m^{\mathrm{T}}(t))^2 b_2^{\mathrm{T}}(t) \tag{16}$$

where $b_2^{\mathrm{T}}(t)$ denotes the tensor obtained from $b_2(t)$ by transposing its two rightmost indices.

Substituting expression (14) in (4) and expressions (15), (16) in (6), the filtering equations for the optimal estimate $m(t)$ and the error variance $P(t)$ are obtained

$$\mathrm{d}m(t) = (a_0(t) + a_1(t)m(t) + a_2(t)m(t)m^{\mathrm{T}}(t) + a_2(t)P(t))\,\mathrm{d}t$$
$$+ P(t)A^{\mathrm{T}}(t)(B(t)B^{\mathrm{T}}(t))^{-1}[\mathrm{d}y(t) - A(t)m(t)\,\mathrm{d}t], \quad m(t_0) = E(x(t_0)|F_t^Y)) \tag{17}$$

$$\mathrm{d}P(t) = (a_1(t)P(t) + P(t)a_1^{\mathrm{T}}(t) + 2a_2(t)m(t)P(t) + 2(a_2(t)m(t)P(t))^{\mathrm{T}}$$
$$+ b_0(t)b_0^{\mathrm{T}}(t) + b_0(t)(b_1(t)m(t))^{\mathrm{T}} + (b_1(t)m(t))b_0^{\mathrm{T}}(t) + b_1(t)P(t)b_1^{\mathrm{T}}(t)$$
$$+ b_1(t)m(t)m^{\mathrm{T}}(t)b_1^{\mathrm{T}}(t) + b_0(t)(P(t) + m(t)m^{\mathrm{T}}(t))b_2^{\mathrm{T}}(t)$$
$$+ b_2(t)(P(t) + m(t)m^{\mathrm{T}}(t))b_0^{\mathrm{T}}(t) + b_1(t)(3m(t)P(t) + m(t)(m(t)m^{\mathrm{T}}(t)))b_2^{\mathrm{T}}(t)$$
$$+ b_2(t)(3P(t)m^{\mathrm{T}}(t) + (m(t)m^{\mathrm{T}}(t))m^{\mathrm{T}}(t))b_1^{\mathrm{T}}(t) + 3b_2(t)P^2(t)b_2^{\mathrm{T}}(t)$$
$$+ 3b_2(t)(P(t)m(t)m^{\mathrm{T}}(t) + m(t)m^{\mathrm{T}}(t)P(t))b_2^{\mathrm{T}}(t)$$
$$+ b_2(t)(m(t)m^{\mathrm{T}}(t))^2 b_2^{\mathrm{T}}(t))\,\mathrm{d}t - P(t)A^{\mathrm{T}}(t)(B(t)B^{\mathrm{T}}(t))^{-1}A(t)P(t)\,\mathrm{d}t \tag{18}$$

$$P(t_0) = E((x(t_0) - m(t_0))(x(t_0) - m(t_0))^{\mathrm{T}}|F_t^Y))$$

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

Theorem 1

The optimal finite-dimensional filter for the bilinear state with bilinear multiplicative noise (1), where the bilinear polynomials $f(x,t)$ and $g(x,t)$ are defined by (12), (13), over the linear observations (2), is given by Equation (17) for the optimal estimate $m(t) = E(x(t)|F_t^Y)$ and Equation (18) for the estimation error variance $P(t) = E[(x(t) - m(t))(x(t) - m(t))^{\mathrm{T}}|F_t^Y]$.

Thus, based on the general non-closed system of the filtering equations (4), (6), it is proved that the closed system of the filtering equations can be obtained for any polynomial state with a polynomial multiplicative noise (1) over linear observations (2). Furthermore, the specific form (17), (18) of the closed system of the filtering equations corresponding to a bilinear state with a bilinear multiplicative noise is derived. In the next section, performance of the designed optimal filter for a bilinear state with a bilinear multiplicative noise over linear observations is verified against the optimal filter for a bilinear state with a state-independent noise and a conventional extended Kalman–Bucy filter.

4. EXAMPLE

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

Let the scalar real state $x(t)$ satisfy the quadratic equation

$$\dot{x}(t) = 0.1x^2(t) + 0.1x^2(t)\psi_1(t), \quad x(0) = x_0 \tag{19}$$

and the scalar observation process be given by the linear equation

$$y(t) = x(t) + \psi_2(t) \tag{20}$$

where $\psi_1(t)$ and $\psi_2(t)$ are white Gaussian noises, which are the weak mean square derivative of standard Wiener processes (see Reference [14]). Equations (19), (20) present the conventional form for Equations (1), (2), which is actually used in practice [15].

The filtering problem is to find the optimal estimate for the quadratic state with quadratic noise (19), using linear observations (20) confused with independent and identically distributed disturbances modelled as white Gaussian noises. Let us set the filtering horizon time to $T = 9.2$.

The filtering equations (17), (18) take the following particular form for the system (19), (20):

$$\dot{m}(t) = 0.1(m^2(t) + P(t)) + P(t)[y(t) - m(t)] \tag{21}$$

with the initial conditions $m(0) = E(x(0)|y(0)) = m_0$

$$\dot{P}(t) = 0.4P(t)m(t) + 0.03P^2(t) + 0.06P(t)m^2(t) + 0.01m^4(t) - P^2(t) \tag{22}$$

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

The estimates obtained upon solving Equations (21)–(22) are compared first to the estimates satisfying the optimal filtering equations for a quadratic state with a state-independent noise (see Reference [10]), based on system (19) where the quadratic multiplicative noise $x^2(t)\psi_1(t)$ is replaced by the standard additive noise $\psi_1(t)$. The corresponding filtering equations are given by

$$\dot{m}_1(t) = 0.1(m_1^2(t) + P_1(t)) + P_1(t)[y(t) - m_1(t)] \tag{23}$$

with the initial conditions $m(0) = E(x(0)|y(0)) = m_0$

$$\dot{P}_1(t) = 0.4P_1(t)m(t) + 0.01 - P_1^2(t) \tag{24}$$

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

The estimates obtained upon solving Equations (21)–(22) are also compared to the estimates satisfying the following extended Kalman–Bucy filtering equations for the quadratic state (19) over the linear observations (20), obtained by replacing the quadratic multiplicative noise $x^2(t)\psi_1(t)$ by the standard additive noise $\psi_1(t)$, using the direct copy of the state dynamics (19) in the estimate equation, and assigning the filter gain as the solution of the Riccati equation

$$\dot{m}_K(t) = 0.1m_K^2(t) + P_K(t)[y(t) - m_K(t)] \tag{25}$$

with the initial conditions $m_K(0) = E(x(0)|y(0)) = m_0$

$$\dot{P}_K(t) = 0.4P_K(t) + 0.01 - P_K^2(t) \tag{26}$$

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

Numerical simulation results are obtained solving the systems of filtering equations (21)–(22), (23)–(24), and (25)–(26). The obtained values of the estimates $m(t)$, $m_1(t)$, and $m_K(t)$ satisfying Equations (21), (23), and (25), respectively, are compared to the real values of the state variables $x(t)$ in (19).

For each of the three filters (21)–(22), (23)–(24), and (25)–(26) and the reference system (19)–(20) involved in simulation, the following initial values are assigned: $x_0 = 1.1$, $m_0 = 0.1$, $P_0 = 1$. Gaussian disturbances $\psi_1(t)$ in (19) and $\psi_2(t)$ in (20) are realized using the built-in MatLab white noise function.

The following graphs are obtained: graphs of the error between the reference state variable $x(t)$ satisfying Equation (19) and the optimal filter estimate $m(t)$ satisfying Equation (21); graph of the error between the reference state variable $x(t)$ satisfying Equation (19) and the estimate $m_1(t)$ satisfying Equation (23); graph of the error between reference state variable $x(t)$ satisfying Equation (19) and the estimate $m_K(t)$ satisfying Equation (25). The graphs of all estimation errors are shown on the simulation interval from $t_0 = 0$ to $T = 7.3$ (Figure 1) and the entire simulation interval from $t_0 = 0$ to $T = 9.2$ (Figure 2). It can be observed that the error given by the optimal filter estimate (21) rapidly reaches and then maintains the zero mean value even in a close vicinity of the asymptotic time point $T = 9.205$, where the reference quadratic state variable (19) goes to infinity. Evidently, there is oscillatory behaviour of the estimation error

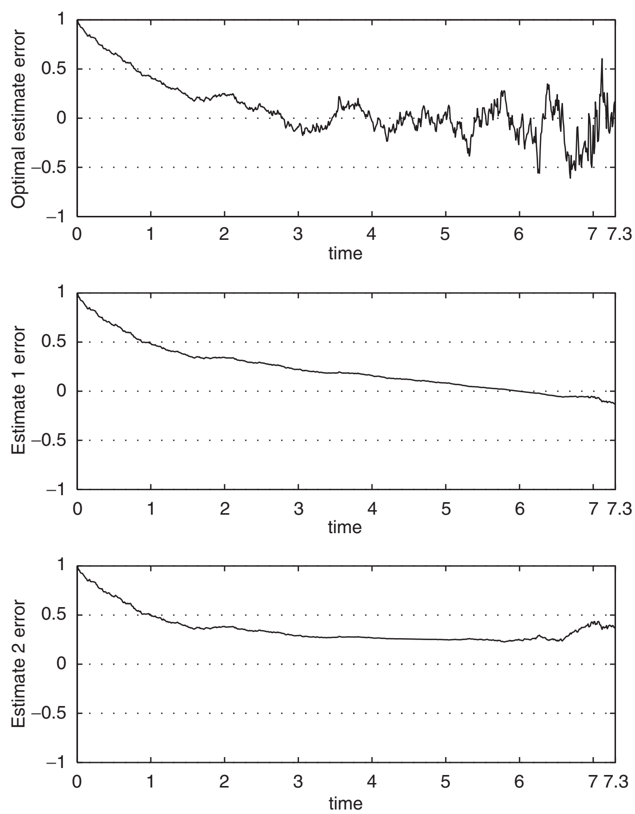

Figure 1. Graph of the error between the real state $x(t)$ satisfying Equation (19) and the optimal filter estimate $m(t)$ satisfying Equation (21) (Optimal estimate error), graph of the error between the real state $x(t)$ satisfying Equation (19) and the estimate $m_1(t)$ satisfying Equation (23) (Estimate 1 error), graph of the error between the real state $x(t)$ satisfying Equation (19) and the estimate $m_K(t)$ satisfying Equation (25) (Estimate 2 error), on the simulation interval $[0, 7.3]$.

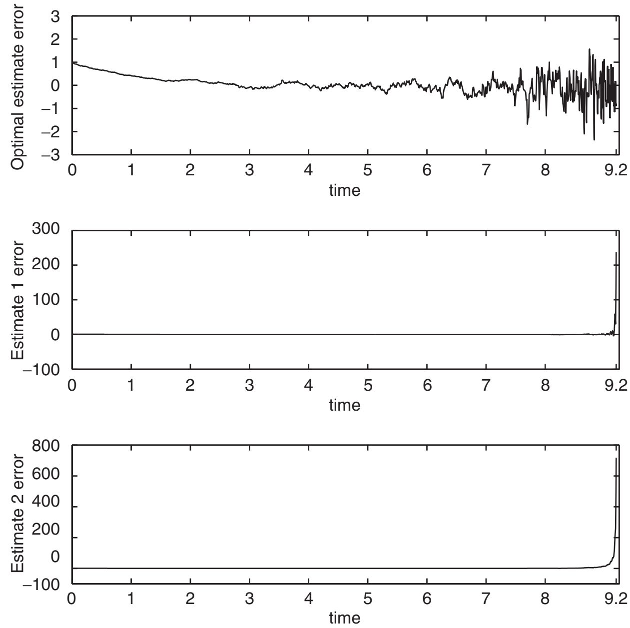

Figure 2. Graph of the error between the real state $x(t)$ satisfying Equation (19) and the optimal filter estimate $m(t)$ satisfying Equation (21) (Optimal estimate error), graph of the error between the real state $x(t)$ satisfying Equation (19) and the estimate $m_1(t)$ satisfying Equation (23) (Estimate 1 error), graph of the error between the real state $x(t)$ satisfying Equation (19) and the estimate $m_K(t)$ satisfying Equation (25) (Estimate 2 error), on the entire simulation interval $[0, 9.2]$.

around the zero mean value near the asymptotic time point $T = 9.205$, since the multiplicative noise intensity in (19) diverges to infinity as time tends to the asymptotic time point. Nevertheless, the peak error absolute values do not exceed 2.5. On the contrary, the errors given by the other considered filters reach zero more slowly or do not reach it at all, have systematic (biased) deviations from zero, and clearly diverge to infinity near the asymptotic time point, taking values that exceed 250. Note that the optimal filtering error variance $P(t)$ does not converge to zero as time tends to the asymptotic time point, since the polynomial dynamics of fourth-order is stronger than the quadratic Riccati terms in the right-hand side of Equation (22).

Thus, it can be concluded that the obtained optimal filter (21)–(22) for a quadratic state with a quadratic multiplicative noise over linear observations yield definitely better estimates than the optimal filter for a quadratic state with a state-independent noise or a conventional extended Kalman–Bucy filter. Subsequent discussion of the obtained simulation results can be found in Section 5.

5. CONCLUSIONS

The simulation results show that the values of the estimate calculated by using the obtained optimal filter for a quadratic state with a quadratic multiplicative noise over linear observations are noticeably closer to the real values of the reference variable than the values of the estimates given by the optimal filter for a quadratic state with a state-independent noise or a conventional extended Kalman–Bucy filter. Moreover, it can be seen that the estimation error produced by the optimal filter rapidly reaches and then maintains the zero mean value even in a close vicinity of the asymptotic time point, where the reference quadratic state variable (19) goes to infinity for a finite time. On the contrary, the estimation errors given by the other two applied filters diverge to infinity near the asymptotic time point. This significant improvement in the estimate behaviour is obtained due to the more careful selection of the filter gain matrix in Equations (21)–(22), as it should be in the optimal filter. Although this conclusion follows from the developed theory, the numerical simulation serves as a convincing illustration. Study of the steady-state behaviour of the designed filter for time-invariant polynomial systems is viewed as a feasible direction of future research.

ACKNOWLEDGEMENTS

The authors thank the U.S. National Science Foundation (NSF) and the Mexican National Science and Technology Council (CONACyT) for financial support under Grants CTS-0117300 and 39388-A, respectively.

REFERENCES

1. Kushner HJ. On differential equations satisfied by conditional probability densities of Markov processes. *SIAM Journal on Control* 1964; **12**:106–119.
2. Kalman RE, Bucy RS. New results in linear filtering and prediction theory. *Journal of Basic Engineering* (ASME) 1961; **83**:95–108.
3. Wonham WM. Some applications of stochastic differential equations to nonlinear filtering. *SIAM Journal on Control* 1965; **2**:347–369.
4. Benes VE. Exact finite-dimensional filters for certain diffusions with nonlinear drift. *Stochastics* 1981; **5**:65–92.
5. Yau SST. Finite-dimensional filters with nonlinear drift I: a class of filters including both Kalman–Bucy and Benes filters. *Journal of Mathematical Systems, Estimation, and Control* 1994; **4**:181–203.
6. Shi P. Filtering on sampled-data systems with parametric uncertainty. *IEEE Transactions on Automatic Control* 1998; **43**:1022–1027.
7. Mahmoud M, Shi P. Robust Kalman filtering for continuous time-lag systems with Markovian jump parameters. *IEEE Transactions on Circuits and Systems* 2003; **50**:98–105.
8. Sheng J, Chen T, Shah SL. Optimal filtering for multirate systems. *IEEE Transactions on Circuits and Systems* 2005; **52**:228–232.
9. Sheng J. Optimal filtering for multirate systems based on lifted models. *Proceedings of the American Control Conference*, Portland, OR, U.S.A., 2005; 3459–3461.
10. Basin MV. On optimal filtering for polynomial system states. *Journal of Dynamic Systems, Measurement, and Control* (ASME) 2003; **125**:123–125.
11. Basin MV, Alcorta-Garcia MA. Optimal filtering and control for third degree polynomial systems. *Dynamics of Continuous, Discrete, and Impulsive Systems* 2003; **10B**:663–680.
12. Basin MV, Skliar M. Optimal filtering for partially measured polynomial system states. *Proceedings of the American Control Conference*, Portland, OR, U.S.A., 2005; 4022–4027.
13. Hazewinkel M, Marcus SI, Sussmann HJ. Nonexistence of exact finite-dimensional filters for conditional statistics of the cubic sensor problem. *Systems and Control Letters* 1983; **5**:331–340.
14. Pugachev VS, Sinitsyn IN. *Stochastic Systems: Theory and Applications*. World Scientific: Singapore, 2001.
15. Åström KJ. *Introduction to Stochastic Control Theory*. Academic Press: New York, 1970.

