Proceedings of the 17th World Congress
The International Federation of Automatic Control
Seoul, Korea, July 6-11, 2008

Optimal Filter Design for Polynomial Systems ⋆

Michael Basin * **Peng Shi** ** **Dario Calderon-Alvarez** ***

* *Department of Physical and Mathematical Sciences, Autonomous University of Nuevo Leon, San Nicolas de los Garza, Nuevo Leon, Mexico; e-mail: mbasin@fcfm.uanl.mx*
** *Department of Computing and Mathematical Sciences, Faculty of Advanced Technology, University of Glamorgan, Pontypridd, United Kingdom; e-mail: pshi@glam.ac.uk*
*** *Department of Physical and Mathematical Sciences, Autonomous University of Nuevo Leon, San Nicolas de los Garza, Nuevo Leon, Mexico; e-mail: dcalal@hotmail.com*

Abstract: In this paper, the optimal filtering problem for polynomial system states over polynomial observations is studied proceeding from the general expression for the stochastic Ito differentials of the optimal estimate and the error variance. In contrast to the previously obtained results, the paper deals with the general case of nonlinear polynomial states and observations. 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 over observations with any polynomial drift is then established. In the example, the obtained optimal filter is applied to solve the optimal third order sensor filtering problem for a quadratic state, assuming a Gaussian initial condition for the extended third order state vector. The simulation results show that the designed filter yields a reliable and rapidly converging estimate.

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 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 that equation can be reduced to a finite-dimensional closed system of filtering equations for a certain number of lower conditional moments (see [2]–[4] for more details). 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]. There also exists an extensive bibliography on robust filtering for nonlinear stochastic systems ([6]–[18]). Apart form the "general situation," the optimal finite-dimensional filters have been designed for certain classes of polynomial system states over linear observations ([19]–[21]). However, the cited papers did not consider the optimal filtering problems for polynomial systems, where both, state and observation, equations include polynomial functions of the system state in the right-hand sides.

This paper presents the optimal finite-dimensional filter for polynomial system states over polynomial observations, continuing the research in the area of the optimal filtering for polynomial systems, which has been initiated in ([19]–[21]). In contrast to the previously obtained results, the paper deals with the general case of nonlinear polynomial states and observations. Designing the optimal filter over polynomial observations presents a significant advantage in the filtering theory and practice, since it enables one to address some filtering problems with state and observation nonlinearities, such as the optimal cubic sensor problem [22] for various polynomial systems. The optimal filtering problem is treated proceeding from the general expression for the stochastic Ito differentials of the optimal estimate and the error variance [23]. 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 any polynomial state and observation equations, additionally assuming a Gaussian initial condition for the extended state vector. This assumption is quite admissible in the filtering framework, since the real distributions of the extended vector components are actually unknown. In this case, the corresponding procedure for designing the optimal filtering equations is suggested.

As an illustrative example, 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 quadratic state and third order polynomial observations, assuming a Gaussian initial condition for the extended third order state vector. This filtering problem generalizes the optimal cubic sensor problem stated in [22]. The resulting filter yields a reliable and rapidly converging estimate, in spite of a significant difference in the initial conditions between the state and estimate, whereas the extended Kalman-Bucy filter estimate, constructed according to [24], behaves unsatisfactorily.

The paper is organized as follows. Section 2 presents the filtering problem statement for polynomial system states over

⋆ The authors thank The London Royal Society (RS) and the Mexican National Science and Technology Council (CONACyT) for financial support under an RS International Incoming Short Visits 2006/R4 Grant and CONACyT Grants 55584 and 52953.

polynomial observations. The Ito differentials for the optimal estimate and the error variance are derived in Section 3, where the procedure for obtaining a closed system of the filtering equations is suggested for any polynomial state over observations with any polynomial drift. In Section 4, the obtained optimal filter is applied to solution of the optimal third order sensor filtering problem for a quadratic state, assuming a Gaussian initial condition for the extended third order state vector.

2. PROBLEM STATEMENT

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 $(W_1(t), F_t, t \geq t_0)$ and $(W_2(t), F_t, t \geq t_0)$ be independent Wiener processes. The F_t-measurable random process $(x(t), y(t)$ is described by nonlinear polynomial differential equations for the system state

$$dx(t) = \rho(x,t)dt + \sigma(x,t)dW_1(t), \quad x(t_0) = x_0, \qquad (1)$$

and the observation process

$$dy(t) = h(x,t)dt + B(t)dW_2(t). \qquad (2)$$

Here, $x(t) \in R^n$ is the state vector and $y(t) \in R^m$ is the observation vector. The initial condition $x_0 \in R^n$ is a Gaussian vector such that x_0, $W_1(t) \in R^p$, and $W_2(t) \in R^q$ are independent. It is assumed that $B(t)B^T(t)$ is a positive definite matrix, therefore, $m \leq q$. All coefficients in (1)–(2) are deterministic functions of appropriate dimensions.

The nonlinear functions $\rho(x,t) \in R^n$, $\sigma(x,t) \in R^n$, and $h(x,t) \in R^m$ 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 $n > 1$. In accordance with [21], a p-degree polynomial of a vector $x(t) \in R^n$ is regarded as a p-linear form of n components of $x(t)$

$$\rho(x,t) = \alpha_0(t) + \alpha_1(t)x + \qquad (3)$$
$$\alpha_2(t)xx^T + \ldots + \alpha_p(t)x \ldots_{p\ times} \ldots x,$$

where $\alpha_0(t)$ is a vector of dimension n, α_1 is a matrix of dimension $n \times n$, α_2 is a 3D tensor of dimension $n \times n \times n$, α_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 (see [21] for more definition). Such a polynomial can also be expressed in the summation form

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

The estimation problem is to find the optimal estimate $\hat{x}(t)$ of the system state $x(t)$, based on the observation process $Y(t) = \{y(s), 0 \leq s \leq t\}$, 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, $E[\xi(t) \mid F_t^Y]$ means the conditional expectation of a stochastic process $\xi(t) = (x(t) - \hat{x}(t))^T (x(t) - \hat{x}(t))$ with respect to the σ - algebra F_t^Y generated by the observation process $Y(t)$ in the interval $[t_0, t]$. As known [23], this optimal estimate is given by the conditional expectation

$$\hat{x}(t) = m_x(t) = E(x(t) \mid F_t^Y)$$

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

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

is the estimation error variance.

The proposed solution to this optimal filtering problem is based on the formulas for the Ito differentials of the optimal estimate and the estimation error variance (cited after [23]) and given in the following section.

3. OPTIMAL FILTER DESIGN

The stated optimal filtering problem is solved by the following theorem.

Theorem 1. *The optimal filter for the polynomial state $x(t)$ (1) over the polynomial observations $y(t)$ (2) is given by the following equations for the optimal estimate $m(t) = [m_z(t), m_x(t)] = E([z(t), x(t)] \mid F_t^Y)$ and the estimation error variance $P(t) = E[([z(t), x(t)] - m(t))([z(t), x(t)] - m(t))^T \mid F_t^Y]$:*

$$dm(t) = E(\bar{f}(x,t) \mid F_t^Y)dt + P(t)[I, 0]^T \times \qquad (4)$$
$$(B(t)B^T(t))^{-1}(dy(t) - m_z(t)dt),$$
$$dP(t) = (E(([z(t), x(t)] - m(t))(\bar{f}(x,t))^T \mid F_t^Y) + \qquad (5)$$
$$E(\bar{f}(x,t)([z(t), x(t)] - m(t))^T) \mid F_t^Y) +$$
$$E(\bar{g}(x,t)\bar{g}^T(x,t) \mid F_t^Y) - P(t)[I, 0]^T (B(t)B^T(t))^{-1}[I, 0]P(t)),$$

with the initial conditions $m(t_0) = [m_z(t_0), m_x(t_0)] = E([z_0, x_0] \mid F_{t_0}^Y)$ and $P(t_0) = E[([z_0, x_0] - m(t_0)([z_0, x_0] - m(t_0)^T \mid F_{t_0}^Y]$. Here, $\bar{f}(x,t) = [f(x,t), \rho(x,t)]$, $\bar{g}(x,t) = [g(x,t), \sigma(x,t)]$,

$$f(x,t) = \frac{\partial h(x,t)}{\partial x}\rho(x,t)dt + \frac{\partial h(x,t)}{\partial t}dt +$$
$$\frac{1}{2}\frac{\partial^2 h(x,t)}{\partial x^2}\sigma(x,t)\sigma^T(x,t)dt, \quad g(x,t) = \frac{\partial h(x,t)}{\partial x}\sigma(x,t),$$

and the additional polynomial state $z(t) = h(x,t)$ satisfies the equation

$$dz(t) = \frac{\partial h(x,t)}{\partial x}\rho(x,t)dt + \frac{\partial h(x,t)}{\partial t}dt + \qquad (6)$$
$$\frac{1}{2}\frac{\partial^2 h(x,t)}{\partial x^2}\sigma(x,t)\sigma^T(x,t)dt + \frac{\partial h(x,t)}{\partial x}\sigma(x,t)dW_1(t),$$

with the initial condition $z(0) = z_0$. The system of filtering equations (4),(5) becomes a closed-form finite-dimensional system after expressing the superior conditional moments of the system state $x(t)$ with respect to the observations $y(t)$ as functions of only two lower conditional moments, $m(t)$ and $P(t)$.

Proof. Let us reformulate the problem, introducing the stochastic process $z(t) = h(x,t)$. Using the Ito formula (see [23]) for the stochastic differential of the nonlinear function $h(x,t)$, where $x(t)$ satisfies the equation (1), the equation (6) is obtained for $z(t)$

$$dz(t) = \frac{\partial h(x,t)}{\partial x}\rho(x,t)dt + \frac{\partial h(x,t)}{\partial t}dt +$$
$$\frac{1}{2}\frac{\partial^2 h(x,t)}{\partial x^2}\sigma(x,t)\sigma^T(x,t)dt + \frac{\partial h(x,t)}{\partial x}\sigma(x,t)dW_1(t).$$

with the initial condition $z(0) = z_0$. Note that the addition $\frac{1}{2}\frac{\partial^2 h(x,t)}{\partial x^2}\sigma(x,t)\sigma^T(x,t)$ appears in view of the second derivative in x in the Ito formula. The initial condition $z_0 \in R^n$ is considered a Gaussian random vector. This assumption is quite admissible in the filtering framework, since the real distributions of $x(t)$ and $z(t)$ are actually unknown.

A key point for further derivations is that the right-hand side of the equation (6) is a polynomial in x. Indeed, since $h(x,t)$ is a polynomial in x, the functions $\frac{\partial h(x,t)}{\partial x}$, $\frac{\partial h(x,t)}{\partial x}x(t)$, $\frac{\partial h(x,t)}{\partial t}$, and $\frac{\partial^2 h(x,t)}{\partial x^2}$ are also polynomial in x. Thus, the equation (6) is a polynomial state equation with a polynomial multiplicative noise. It can be written in the compact form

$$dz(t) = f(x,t)dt + g(x,t)dW_1(t), \quad z(t_0) = z_0, \qquad (7)$$

where

$$f(x,t) = \frac{\partial h(x,t)}{\partial x}\rho(x,t) + \frac{\partial h(x,t)}{\partial t} + \frac{1}{2}\frac{\partial^2 h(x,t)}{\partial x^2}\sigma(x,t)\sigma^T(x,t), \quad g(x,t) = \frac{\partial h(x,t)}{\partial x}\sigma(x,t).$$

In terms of the process $z(t)$, the observation equation (2) takes the form

$$dy(t) = [I,0][z(t),x(t)]^T dt + B(t)dW_2(t), \qquad (8)$$

where the matrix $[I,0]$ is the $m\times(n+m)$ matrix composed of the $m\times m$-dimensional identity matrix and $m\times n$-dimensional zero matrix.

The reformulated estimation problem is now to find the optimal estimate $[m_z(t),m_x(t)]$ of the system state $[z(t),x(t)]$, based on the observation process $Y(t) = \{y(s), 0\le s\le t\}$. This optimal estimate is given by the conditional expectation

$$m(t) = [m_z(t),m_x(t)] = [E(z(t)\mid F_t^Y), E(x(t)\mid F_t^Y)]$$

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

$$P(t) = E[([z(t),x(t)] - [m_z(t),m_x(t)])\times ([z(t),x(t)] - [m_z(t),m_x(t)])^T \mid F_t^Y]$$

is the estimation error variance for this reformulated problem.

The obtained filtering system includes the two equations, (6) (or (7)) and (1), for the partially measured state $[z(t),x(t)]$ and the equation (8) for the observations $y(t)$, where $z(t)$ is a completely measured polynomial state with a polynomial multiplicative noise, $x(t)$ is an unmeasured polynomial state, and $y(t)$ is a linear observation process directly measuring the state $z(t)$. As follows from the general optimal filtering theory ([23]), the optimal filtering equations take the following particular form for the system (7), (1), (8)

$$dm(t) = E(\bar f(x,t)\mid F_t^Y)dt + P(t)[I,0]^T\times (B(t)B^T(t))^{-1}(dy(t) - m_z(t)dt), \qquad (9)$$

$$\begin{aligned} dP(t) = {} & (E(([z(t),x(t)] - m(t))(\bar f(x,t))^T \mid F_t^Y) + \\ & E(\bar f(x,t)([z(t),x(t)] - m(t))^T)\mid F_t^Y) + \\ & E(\bar g(x,t)\bar g^T(x,t)\mid F_t^Y) - \\ & P(t)[I,0]^T(B(t)B^T(t))^{-1}[I,0]P(t))dt + \\ & E((([z(t),x(t)] - m(t))([z(t),x(t)] - m(t))\times \\ & ([z(t),x(t)] - m(t))^T \mid F_t^Y)[I,0]^T\times \\ & (B(t)B^T(t))^{-1}(dy(t) - m_z(t)dt), \end{aligned} \qquad (10)$$

where $\bar f(x,t) = [f(x,t),\rho(x,t)]$ is the polynomial drift term and $\bar g(x,t) = [g(x,t),\sigma(x,t)]$ is the polynomial diffusion (multiplicative noise) term in the entire system of the state equations (7), (1), and the last term should be understood as a 3D tensor (under the expectation sign) convoluted with a vector, which yields a matrix. The equations (9), (10) should be complemented with the initial conditions $m(t_0) = [m_z(t_0),m_x(t_0)] = E([z_0,x_0]\mid F_{t_0}^Y)$ and $P(t_0) = E[([z_0,x_0] - m(t_0)([z_0,x_0] - m(t_0)^T \mid F_{t_0}^Y]$.

Let us show that a closed system of the filtering equations can be obtained for the incompletely measured polynomial state $[z(t),x(t)]$ over the linear observations $y(t)$, in view of the polynomial properties of the functions in the right-hand side of the equation (6). Indeed, as shown in [19]–[21], a closed system of the filtering equations for polynomial system states (6) (or (7)) and (1) with polynomial multiplicative noises over linear observations can be obtained, if the observation matrix is invertible for any $t\ge t_0$. Since the observation matrix $A(t) = [I,0]\in R^{m\times(n+m)}$ in (8) is not invertible, the following transformations are introduced.

First, note that the matrix A is a matrix of complete rank, m, which is equal to the dimension of the observation process $y(t)\in R^m$. Further note that the number of Wiener processes in the observation equations can also be reduced to m, the dimension of independent observations, by summarizing and renumerating the Wiener processes in each observation equation (2). Therefore, the matrix B can always be assumed a square matrix of dimension $m\times m$, such that $B(t)B^T(t)$ is a positive definite matrix (see Section 2 for this condition). Next, the new matrices $\bar A(t)$ and $\bar B(t)$ are defined as follows. The matrix $\bar A(t)\in R^{(n+m)\times(n+m)}$ is obtained from $A(t) = [I,0]\in R^{m\times(n+m)}$ by adding n linearly independent rows such that the resulting matrix $\bar A(t)$ is invertible. The matrix $\bar B(t)\in R^{(n+m)\times(n+m)}$ is made from the matrix $B(t)\in R^{m\times m}$ by placing $B(t)$ in the upper left corner of $\bar B(t)$, defining the other n diagonal entries of $\bar B(t)$ equal to infinity, and setting to zero all other entries of $\bar B(t)$ outside the main diagonal or outside the submatrix $B(t)$. In other words, $\bar B(t) = diag[B(t),\beta I_n]$, where $\beta = \infty$, and I_n is the identity matrix of dimension $n\times n$. Thus, the new observation equation is given by

$$\bar y(t) = \bar A(t)x(t)dt + \bar B(t)dW_2(t), \qquad (11)$$

where $\bar y(t)\in R^{n+m}$.

The key point of the introduced transformation is that the new observation process $\bar y(t)$ is physically equivalent to the old one $y(t)$, since the fictitious last n components of $\bar y(t)$ consist of pure noise in view of infinite intensities of white Gaussian noises in the corresponding n equations, and the first m components of $\bar y(t)$ coincide with $y(t)$. In addition, the entire observation matrix $\bar A(t)$ is invertible, and the matrix $(\bar B(t)\bar B^T(t))^{-1}\in R^{(n+m)\times(n+m)}$ exists and equals to the $(n+m)\times(n+m)$ – dimensional square matrix, whose upper left corner is occupied by the submatrix $(B(t)B^T(t))^{-1}\in R^{m\times m}$ and all other entries are zeros.

In terms of the new observation equation (11), the optimal filtering equations (9) and (10) take the form

$$dm(t) = E(\bar f(x,t)\mid F_t^Y)dt + P(t)\bar A^T(t)\times (\bar B(t)\bar B^T(t))^{-1}(d\bar y(t) - \bar A(t)m(t)dt), \qquad (12)$$

$$\begin{aligned} dP(t) = {} & (E(([z(t),x(t)] - m(t))(\bar f(x,t))^T \mid F_t^Y) + \\ & E(\bar f(x,t)([z(t),x(t)] - m(t))^T)\mid F_t^Y) + \\ & E(\bar g(x,t)\bar g^T(x,t)\mid F_t^Y) - \\ & P(t)\bar A^T(t)(\bar B(t)\bar B^T(t))^{-1}\bar A(t)P(t))dt + \\ & E((([z(t),x(t)] - m(t))([z(t),x(t)] - m(t))\times \\ & ([z(t),x(t)] - m(t))^T \mid F_t^Y)\times \\ & \bar A^T(t)(\bar B(t)\bar B^T(t))^{-1}(d\bar y(t) - \bar A(t)m(t)dt), \end{aligned} \qquad (13)$$

with the initial conditions $m(t_0) = [m_z(t_0), m_x(t_0)] = E([z_0, x_0] \mid F^Y_{t_0})$ and $P(t_0) = E[([z_0, x_0] - m(t_0)([z_0, x_0] - m(t_0)^T \mid F^Y_{t_0}]$.

Since the new observation matrix $\bar{A}(t)$ is invertible for any $t \geq t_0$, the random variable $x(t) - m(t)$ is conditionally Gaussian with respect to the new observation process $\bar{y}(t)$, and therefore with respect to the original observation process $y(t)$, for any $t \geq t_0$ (see [19]–[21]). Hence, the following considerations are applicable to the filtering equations (12),(13).

First, since the random variable $x(t) - m(t)$ is conditionally Gaussian, the conditional third moment $E((([z(t), x(t)] - m(t))([z(t), x(t)] - m(t))([z(t), x(t)] - m(t))^T \mid F^Y_t)$ with respect to observations, which stands in the last term of the equation (13), is equal to zero, because the process $[z(t), x(t)] - m(t)$ is conditionally Gaussian. Thus, the entire last term in (13) is vanished and the following variance equation is obtained

$$dP(t) = (E(([z(t), x(t)] - m(t))(\bar{f}(x,t))^T \mid F^Y_t) + E(\bar{f}(x,t)([z(t), x(t)] - m(t))^T) \mid F^Y_t) + E(\bar{g}(x,t)\bar{g}^T(x,t) \mid F^Y_t) - P(t)\bar{A}^T(t)(\bar{B}(t)\bar{B}^T(t))^{-1}\bar{A}(t)P(t))dt \quad (14)$$

with the initial condition $P(t_0) = E[([z_0, x_0] - m(t_0)([z_0, x_0] - m(t_0)^T \mid F^Y_{t_0}]$.

Second, if the functions $\bar{f}(x,t)$ and $\bar{g}(x,t)$ are polynomial functions of the state x with time-dependent coefficients, the expression of the terms $E(\bar{f}(x,t) \mid F^Y_t)$ in (12) and $E((x(t) - m(t))\bar{f}^T(x,t)) \mid F^Y_t)$, $E(\bar{g}(x,t)\bar{g}^T(x,t) \mid F^Y_t)$ in (14) 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 $[z(t), x(t)] - m(t)$: all its odd conditional moments, $m_1 = E[([z(t), x(t)] - m(t)) \mid Y(t)]$, $m_3 = E[([z(t), x(t)] - m(t)^3 \mid Y(t)]$, $m_5 = E[([z(t), x(t)] - m(t))^5 \mid Y(t)]$, ..., are equal to 0, and all its even conditional moments $m_2 = E[([z(t), x(t)] - m(t))^2 \mid Y(t)]$, $m_4 = E[([z(t), x(t)] - m(t))^4 \mid Y(t)]$, ..., can be represented as functions of the variance $P(t)$. For example, $m_2 = P$, $m_4 = 3P^2$, $m_6 = 15P^3$,..., etc. After representing all polynomial terms in (12) and (14), that are generated upon expressing $E(\bar{f}(x,t) \mid F^Y_t)$, $E(([z(t), x(t)] - m(t))\bar{f}^T(x,t)) \mid F^Y_t)$, and $E(\bar{g}(x,t)\bar{g}^T(x,t) \mid F^Y_t)$ as functions of $m(t)$ and $P(t)$, a closed form of the filtering equations would be obtained. The corresponding representations of $E(\bar{f}(x,t) \mid F^Y_t)$, $E(([z(t), x(t)] - m(t))(\bar{f}(x,t))^T \mid F^Y_t)$, and $E(\bar{g}(x,t)\bar{g}^T(x,t) \mid F^Y_t)$ have been derived in [19–21] for certain polynomial functions $\bar{f}(x,t)$ and $\bar{g}(x,t)$.

Finally, in view of definition of the matrices $\bar{A}(t)$ and $\bar{B}(t)$ and the new observation process $\bar{y}(t)$, the filtering equations (12),(14) can be written again in terms of the original observation equation (2) using $y(t)$, $A(t) = [I, 0]$, and $B(t)$. As a result, the optimal filtering equations (4),(5) are obtained

$$dm(t) = E(\bar{f}(x,t) \mid F^Y_t)dt + P(t)[I,0]^T \times (B(t)B^T(t))^{-1}(dy(t) - m_z(t)dt),$$

$$dP(t) = (E(([z(t), x(t)] - m(t))(\bar{f}(x,t))^T \mid F^Y_t) + E(\bar{f}(x,t)([z(t), x(t)] - m(t))^T) \mid F^Y_t) + E(\bar{g}(x,t)\bar{g}^T(x,t) \mid F^Y_t) - P(t)[I,0]^T(B(t)B^T(t))^{-1}[I,0]P(t)),$$

with the initial conditions $m(t_0) = [m_z(t_0), m_x(t_0)] = E([z_0, x_0] \mid F^Y_{t_0})$ and $P(t_0) = E[([z_0, x_0] - m(t_0)([z_0, x_0] - m(t_0)^T \mid F^Y_{t_0}]$. ■

In the next example section, a closed form of the filtering equations will be obtained for a particular case of scalar second and third order polynomial functions $\rho(x,t)$, $\sigma(x,t)$, and $h(x,t)$ in the equations (1) and (2). Nonetheless, application of the same procedure would result in designing a closed system of the filtering equations for any polynomial functions $\rho(x,t)$, $\sigma(x,t)$, and $h(x,t)$ in (1),(2).

4. EXAMPLE: THIRD DEGREE SENSOR FILTERING PROBLEM FOR QUADRATIC SYSTEM

This section presents an example of designing the optimal filter for a quadratic state over third degree polynomial observations, reducing it to the optimal filtering problem for a fourth degree polynomial state with a second degree polynomial multiplicative noise over linear observations, where a Gaussian state initial condition is additionally assumed.

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

$$dx(t) = x^2(t)dt + dw_1(t), \quad x(0) = x_0, \quad (15)$$

and the observation process be given by the scalar third degree sensor equation

$$dy(t) = (x^3(t) + x(t))dt + dw_2(t), \quad (16)$$

where $w_1(t)$ and $w_2(t)$ are standard Wiener processes independent of each other and of a Gaussian random variable x_0 serving as the initial condition in (15). The filtering problem is to find the optimal estimate for the quadratic state (15), using the third degree sensor observations (16).

Let us reformulate the problem, introducing the stochastic process $z(t) = h(x,t) = x^3(t) + x(t)$. Using the Ito formula (see [23]) for the stochastic differential of the cubic function $h(x,t) = x^3(t) + x(t)$, where $x(t)$ satisfies the equation (15), the following equation is obtained for $z(t)$

$$dz(t) = (x^2(t) + 3x(t) + 3x^4(t))dt + (3x^2(t) + 1)dw_1(t), \quad (17)$$

with the initial condition $z(0) = z_0$. Here, $\frac{\partial h(x,t)}{\partial x} = 3x^2(t) + 1$, $\frac{1}{2}\frac{\partial h^2(x,t)}{\partial x^2} = 3x(t)$, and $\frac{\partial h(x,t)}{\partial t} = 0$; therefore, $f(x,t) = x^2(t) + 3x(t) + 3x^4(t)$ and $g(x,t) = 3x^2(t) + 1$. The initial condition z_0 is considered a Gaussian random variable. In terms of the process $z(t)$, the observation equation (16) takes the form

$$dy(t) = z(t)dt + dw_2(t). \quad (18)$$

The obtained filtering system includes two equations, (17) and (15), for the partially measured state $[z(t), x(t)]$ and an equation (18) for the observations $y(t)$, where $z(t)$ is a completely measured fourth degree state with a multiplicative quadratic noise, $x(t)$ is an unmeasured quadratic state, and $y(t)$ is a linear observation process directly measuring the state $z(t)$. Hence, the designed optimal filter can be applied for solving this problem. The filtering equations (4),(5) take the following particular form for the system (17),(15),(18)

$$dm_1(t) = (1 + 3m_2(t) + 3m_2^2(t) + 3P_{22}(t))dt + P_{11}(t)[dy(t) - m_1(t)dt], \quad (19)$$

$$dm_2(t) = (m_2^2(t) + P_{22}(t))dt + P_{12}(t)[dy(t) - m_1(t)dt], \quad (20)$$

with the initial conditions $m_1(0) = E(z_0 \mid y(0)) = m_{10}$ and $m_2(0) = E(x_0 \mid y(0)) = m_{20}$,

$$\dot{P}_{11}(t) = 12(P_{12}(t)m_2(t)) + 6P_{12}(t) + 27P_{22}^2(t) + 54P_{22}(t)m_2^2(t) + 9m_2^4(t) + 6P_{22}(t) + 6m_2^2 + 1 - P_{11}^2(t), \quad (21)$$

$$\dot{P}_{12}(t) = 1 + 6(P_{22}(t)m_2(t)) + 3P_{22}(t) + 3(m_2^2(t) + P_{22}(t)) - P_{11}(t)P_{12}(t), \quad (22)$$

$$\dot{P}_{22}(t) = 1 + 4P_{22}(t)m_2(t) - P_{12}^2(t), \quad (23)$$

with the initial condition $P(0) = E(([x_0,z_0]^T - m(0))([x_0,z_0]^T - m(0))^T \mid y(0)) = P_0$. Here, $m_1(t)$ is the optimal estimate for the state $z(t) = x^3(t) + x(t)$ and $m_2(t)$ is the optimal estimate for the state $x(t)$.

The estimates obtained upon solving the equations (19)–(23) are compared to the estimates satisfying the extended Kalman-Bucy filtering equations for the quadratic state (17) over the third order polynomial observations (16), which are obtained using Theorem 8.1 from [24]:

$$\dot{m}_K(t) = m_K^2(t) + P_K(t)(3m_K^2(t)+1)[y(t) - m_K^3(t) - m_K(t)], \quad (24)$$

with the initial condition $m_K(0) = E(x(0) \mid y(0)) = m_{20}$,

$$\dot{P}_K(t) = 1 + 4m_K(t)P_K(t) - (3m_K^2(t)+1)^2 P_K^2(t), \quad (25)$$

with the initial condition $P_K(0) = E((x(0) - m_K(0))(x(0) - m_K(0))^T \mid y(0)) = P_K(0) = P_{22}(0)$.

Numerical simulation results are obtained solving the systems of filtering equations (19)–(23) and (24)–(25). The obtained values of the state estimates $m_2(t)$, satisfying the equation (20), and $m_K(t)$, satisfying the equation (24), are compared to the real values of the state variable $x(t)$ in (15).

For the filters (19)–(23), (24)–(25) and the reference system (17),(15),(18) involved in simulation, the following initial values are assigned: $x_0 = z_0 = 0$, $m_{20} = m_K(0) = 10$, $m_{10} = 1000$, $P_{11}(0) = 15$, $P_{12}(0) = 3$, $P_{22}(0) = P_K(0) = 1$. Gaussian disturbances $dw_1(t)$ and $dw_2(t)$ are realized using the built-in MatLab white noise functions. The simulation interval is set to $[0, 0.7]$.

Figure 1 shows the graphs of the errors between the reference state $x(t)$ (15) and its optimal estimate $m_2(t)$ (20), and the reference state $z(t) = x^3(t) + x(t)$ (17) and its optimal estimate $m_1(t)$ (19), in the entire simulation interval from $t_0 = 0$ to $T = 0.7$. It can be observed that the optimal estimation errors converge to the real states very rapidly and then maintain zero mean value, in spite of a considerable error in the initial conditions, $m_{20} - x_0 = 10$, $m_{10} - z(0) = 1000$. The estimation error for the state $x(t)$ at $T = 0.7$ is equal to $m_2(0.7) - x(0.7) = 0.04$. Figure 2 shows the graph of the error between the reference state $x(t)$ (15) and the extended Kalman-Bucy filter estimate $m_K(t)$ (24). It can be observed that the extended Kalman-Bucy filter estimate does not converge to zero for the simulation time, admitting quite a large error at $T = 0.7$, which is equal to $m_K(0.7) - x(0.7) = 0.57$, fourteen times more than the optimal estimation error in Fig. 1.

Thus, it can be concluded that the obtained optimal filter (19)–(23) solves the optimal third order sensor filtering problem for the system (15),(16) and yields a reliable estimate of the unmeasured state.

5. CONCLUSIONS

This paper presents the optimal filter for polynomial system states over polynomial observations. It is shown that the optimal filter can be obtained in a closed form for any polynomial functions in state and observation equations. In the example, the optimal solution is obtained to the filtering problem for a quadratic state over third degree polynomial observations, assuming a Gaussian initial condition for the extended third order state vector. The resulting filter yields a reliable and rapidly converging estimate, in spite of a significant difference in the initial conditions between the state and estimate, whereas the extended Kalman-Bucy filter estimate behaves unsatisfactorily. Although this conclusion follows from the developed theory, the numerical simulation serves as a convincing illustration.

REFERENCES

[1] Kushner HJ. On differential equations satisfied by conditional probability densities of Markov processes, *SIAM J. Control* 1964; **12**:106–119.

[2] Kalman RE, Bucy RS. New results in linear filtering and prediction theory, *ASME Trans., Part D (J. of Basic Engineering)* 1961; **83**: 95–108.

[3] Wonham WM. Some applications of stochastic differential equations to nonlinear filtering, *SIAM J. 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, *J. Math. Systems, Estimation, and Control* 1994; **4**: 181–203.

[6] Xie LH, De Souza CE, Wang YY. Robust fitering for a class of discrete-time uncertain nonlinear systems, *International Journal of Robust and Nonlinear Control*, 1996; **6**: 297–312.

[7] Nguang SK, Fu MY. Robust nonlinear H_∞ filtering, *Automatica* 1996; **32**: 1195–1199.

[8] Fridman E, Shaked U. On regional nonlinear H_∞ filtering, *Systems and Control Letters* 1997; **29**: 233–240.

[9] Shi P. Filtering on sampled-data systems with parametric uncertainty, *IEEE Transactions on Automatic Control* 1998; **43**: 1022–1027.

[10] Fleming WH, McEneaney WM. Robust limits of risk sensitive nonlinear filters, *Mathematics of Control, Signals and Systems* 2001; **14**: 109–142.

[11] Yaz E, Yaz Y. State estimation of uncertain nonlinear systems with general criteria, *Applied Mathematics Letters* 2001; **14**: 605–610.

[12] Xu S., van Dooren PV. Robust H_∞ filtering for a class of nonlinear systems with state delay and parameter uncertainty, *Int. J. Control*, 2002; **75**: 766–774.

[13] Xu SY, Chen TW. Robust H_∞ filtering for uncertain impulsive stochastic systems under sampled measurements, *Automatica* 2003; **39**: 509–516.

[14] 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.

[15] Zhang WH, Chen BS, Tseng CS. Robust H_∞ filtering for nonlinear stochastic systems, *IEEE Transactions on Signal Processing* 2005; **53**: 589–598.

[16] Xu SY, Lam J, Gao HJ, Zhou Y. Robust H_∞ filtering for uncertain discrete stochastic systems with time delays, *Circuits, Systems and Signal Processing* 2005; **24**: 753–770.

[17] Gao H, Lam J, Xie L, Wang C. New approach to mixed H_2/H_∞–filltering for polytopic discrete-time systems, *IEEE Transactions on Signal Processing* 2005; **53**: 3183–3192.

[18] Zhang H, Basin MV, Skliar M. Ito-Volterra optimal state estimation with continuous, multirate, randomly sampled, and delayed measurements, *IEEE Transactions on Automatic Control* 2007; **52**: 401–416.

[19] Basin MV. On optimal filtering for polynomial system states, *ASME Trans. J. Dynamic Systems, Measurement, and Control* 2003; **125**: 123–125.

[20] 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.

[21] Basin MV, Perez J, Skliar M. Optimal filtering for polynomial system states with polynomial multiplicative noise, *International J. Robust and Nonlinear Control* 2006; **16**: 287–298.

[22] 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.

[23] Pugachev VS, Sinitsyn IN. *Stochastic Systems: Theory and Applications*. World Scientific, 2001.

[24] Jazwinski AH. *Stochastic Processes and Filtering Theory*. Academic Press: New York, 1970.

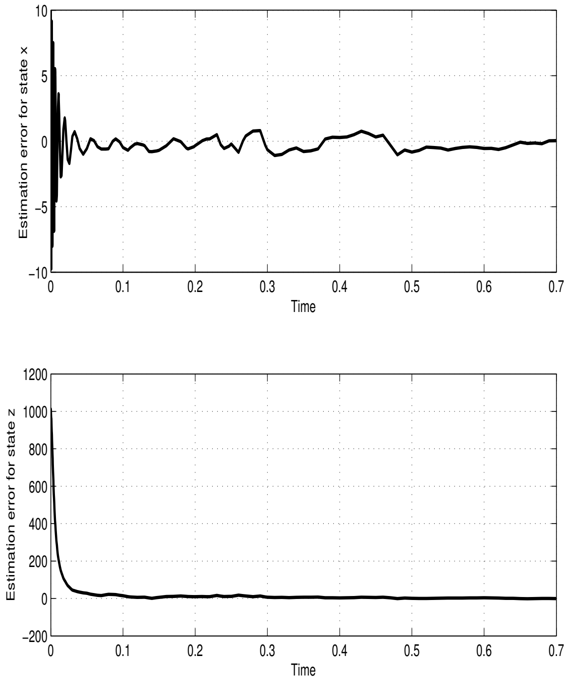

Fig. 1. **Above.** Graph of the estimation error between the reference state $x(t)$ (15) and its optimal estimate $m_2(t)$ (20) in the interval $[0, 0.7]$. **Below.** Graph of the estimation error between the reference state $z(t)$ (17) and its optimal estimate $m_1(t)$ (19) in the interval $[0, 0.7]$.

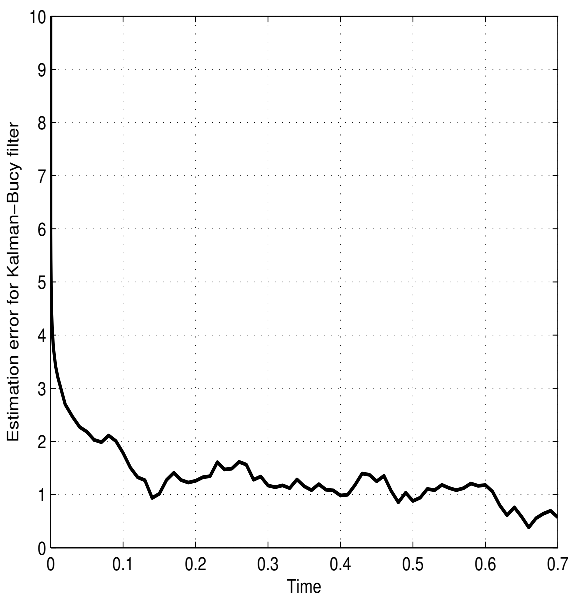

Fig. 2. Graph of the estimation error between the reference state $x(t)$ (15) and its estimate $m_K(t)$ (24) in the interval $[0, 0.7]$.