2011 American Control Conference
on O'Farrell Street, San Francisco, CA, USA
June 29 - July 01, 2011

Mean-Square Joint State and Parameter Estimation for Uncertain Nonlinear Polynomial Stochastic Systems

Michael Basin Alexander Loukianov Miguel Hernandez-Gonzalez

Abstract— **This paper presents the mean-square joint state filtering and parameter identification problem for uncertain nonlinear polynomial stochastic systems with unknown parameters in the state equation over nonlinear polynomial observations, where the unknown parameters are considered Wiener processes. The original problem is reduced to the filtering problem for an extended state vector that incorporates parameters as additional states. The obtained mean-square filter for the extended state vector also serves as the mean-square identifier for the unknown parameters. Performance of the designed mean-square state filter and parameter identifier is verified for both, positive and negative, parameter values.**

I. INTRODUCTION

The problem of the optimal simultaneous state estimation and parameter identification for stochastic systems with unknown parameters has been receiving systematic treatment beginning from the seminal paper [1]. The optimal result was obtained in [1] for a linear discrete-time system with constant unknown parameters within a finite filtering horizon, using the maximum likelihood principle (see, for example, [2]), in view of a finite set of the state and parameter values at time instants. The application of the maximum likelihood concept was continued for linear discrete-time systems in [3] and linear continuous-time systems in [4]. Nonetheless, the use of the maximum likelihood principle reveals certain limitations in the final result: a. the unknown parameters are assumed constant to avoid complications in the generated optimization problem and b. no direct dynamical (difference) equations can be obtained to track the state and parameter estimates dynamics in the "general situation," without imposing special assumptions on the system structure. Other approaches are presented by the parameter identification methods without simultaneous state estimation, such as designed in [5], [6], [7], which are also applicable to nonlinear stochastic systems. Robust approximate identification in nonlinear systems using various approaches, such as H_∞ filtering, is studied in a variety of papers [8]–[22] for stochastic systems with bounded uncertainties in coefficients.

This paper presents the mean-square joint filtering and parameter identification problem for uncertain nonlinear polynomial stochastic systems with unknown parameters in the state equation over nonlinear polynomial observations. The solution starts with reduction of the original identification problem to the mean-square filtering problem for nonlinear polynomial system states over nonlinear polynomial observations, upon considering the unknown parameters as additional system states satisfying linear stochastic Ito equations with zero drift and unit diffusion, i.e., standard Wiener processes. In doing so, the unknown parameters are incorporated into the extended polynomial state vector, which should be mean-square estimated over polynomial observations. The obtained filtering problem is then further reduced to the filtering problem for polynomial system states over direct linear observations, assuming the nonlinear drift components in the observation equation as more additional states and including them in the extended state vector. The latter filtering problem is solved using the mean-square filter for nonlinear polynomial states over linear observations ([23]). The designed mean-square filter for the extended state vector also serves as the identifier for the unknown parameters.

In the illustrative example, performance of the designed mean-square filter is verified for a nolinear system state over nonlinear polynomial observations with multiplicative unknown parameter in the state equation. Both, positive and negative, values of the parameter in the state equation are examined. The simulation results demonstrate reliable performance of the filter: in both cases, the state estimate converges to the real state and the parameter estimate converges to the real parameter value rapidly.

II. MEAN-SQUARE JOINT STATE AND PARAMETER ESTIMATION PROBLEM FOR NONLINEAR POLYNOMIAL SYSTEMS

Let $(\Omega, \mathcal{F}, P)$ be a complete probability space with an increasing right-continuous family of σ-algebras $\mathcal{F}_t, t \geq t_0$, and let $(W_1(t), \mathcal{F}_t, t \geq t_0)$ and $(W_2(t), \mathcal{F}_t, t \geq t_0)$ be independent standard Wiener processes. The $\mathcal{F}_t$-measurable random process $(x(t), y(t))$ is described by a nonlinear polynomial differential equation for the system state

$$dx(t) = F(x(t), \theta(t), t)dt + b(t)dW_1(t) \quad (1)$$

and a nonlinear polynomial equation for the observation process

$$dy(t) = h(x(t), t)dt + G(t)dW_2(t) \quad (2)$$

here $x(t) \in \Re^n$ is the state vector and $y(t) \in \Re^m$ is the observation vector, $m \leq n$. The expressions $F(x(t), \theta(t), t)$ and $h(x(t), t)$ are considered polynomials of n variables,

The authors thank the Mexican National Science and Technology Council (CONACyT) for financial support under Grants 55584 and 46069 and Sabbatical Fellowship for the first author.

M. Basin is with the Department of Physical and Mathematical Sciences, Autonomous University of Nuevo Leon, mbasin@fcfm.uanl.mx

M. Basin, A.G. Loukianov, M. Hernandez-Gonzalez is with Center for Research and Graduate Studies (CINVESTAV), Campus Guadalajara, Jalisco, Mexico {mbasin,louk,mglez}@gdl.cinvestav.mx

components of the state vector $x(t) \in \Re^n$. Since $x(t)$ is a vector, this requires a special definition of the polynomials for $n > 1$. In accordance with [24], a p-degree and an r-degree polynomials of a vector $x(t) \in \Re^n$ are regarded as p-linear and r-linear forms of n components of $x(t)$

$$\begin{aligned} F(x(t),\theta(t),t) = F_0(\theta(t),t) + F_1(\theta(t),t)x(t) \\ +F_2(\theta(t),t)x(t)x^\top(t) + \ldots \\ +F_p(\theta(t),t)x(t)x(t)\ldots_{p \text{ times}} \ldots x^\top(t), \\ h(x(t),t) = H_0(t) + H_1(t)x(t) + \ldots \\ +H_r(t)x(t)x(t)\ldots_{r \text{ times}} \ldots x^\top(t) \end{aligned} \tag{3}$$

where $F_0(\theta(t),t)$ is a vector of dimension n, $F_1(\theta(t),t)$ is a matrix of dimension $n \times n$, $F_2(\theta(t),t)$ is a 3D tensor of dimension $n \times n \times n$, $F_p(\theta(t),t)$ is a $(p+1)$D tensor of dimension $n \times n \ldots_{(p+1) \text{ times}} \ldots \times n$ and $x \times \ldots_{p \text{ times}} \ldots x$ 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 polynomial can also be expressed in the sumation form

$$\begin{aligned} F_k(x(t),\theta(t),t) = F_{0k}(\theta(t),t) + \sum_i F_{1\,ki}(\theta(t),t)x_i(t) \\ +\sum_{ij} F_{2\,kij}(\theta(t),t)x_i(t)x_j(t) + \ldots \\ +\sum_{i_1 i_2 \ldots i_p} F_{p\,k i_1 i_2 \ldots i_p}(\theta(t),t)x_{i1}(t)x_{i2}(t)\ldots x_{ip}(t), \\ k,i,j,i_1 \ldots i_p = 1,\ldots,n \end{aligned} \tag{4}$$

and the vector $H_0(t)$, the matrix $H_1(t)$, the $(r+1)$D tensor $H_r(t)$ are defined in a similar way but of dimension m. $\theta(t) \in \Re^p$, $p \le n + n \times n + \ldots + n \times n \ldots_{p \text{ times}} \ldots n$ in (3) is the state vector of unknown entries of $F_0(\theta(t),t) \in \Re^n$, $F_1(\theta(t),t) \in \Re^{n \times n}$, ..., $F_p(\theta(t),t) \in \Re^{n \times n \ldots_{p+1 \text{ times}} \ldots n}$. The unknown entries in $F_j(\theta(t),t)$, $j = 0,\ldots,p$ are such that $F_{0_{i_1}} = \theta_l(t)$, $l = 1,\ldots,p_1 \le n$, $F_{1_{i_1 i_2}} = \theta_l(t)$, $l = p_1+1,\ldots,p_2 \le n + n \times n$, ..., $F_{p_{i_1 i_2 \ldots i_p}} = \theta_l(t)$, $l = p_1 + p_2 + \ldots + p_{p-1} + 1,\ldots,p \le n + n \times n + \ldots + n \times n \times \ldots_{p \text{ times}} \ldots n$, $i_1, i_1 i_2, \ldots, i_1 i_2 \ldots i_p$ represent known and unknown parameters in $F_0, F_1, \ldots, F_p$. The initial condition $x_0 \in \Re^n$ is a Gaussian vector such that x_0, $W_1(t)$ and $W_2(t)$ are independent. It is assumed that $G(t)G^\top(t)$ is a positive definite matrix. The coefficients in (1)-(2) are deterministic functions of appropriate dimensions.

The estimation problem is to find the mean-square estimate $\hat{x}(t)$ of the system state $x(t)$, based on the observation process $Y(t) = \{y(s), 0 \le s \le t\}$, that minimizes the conditional expectation of the Euclidean 2-norm

$$J = E\left[(x(t) - \hat{x}(t))^\top (x(t) - \hat{x}(t)) \mid \mathcal{F}_t^Y\right]$$

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

$$\hat{x}(t) = E\left[x(t) \mid \mathcal{F}_t^Y\right]$$

of the system state $x(t)$ with respect to the σ-algebra $\mathcal{F}_t^Y$ generated by the observation process $Y(t)$ in the interval of time $[t_0, t]$. As usual, the symmetric matrix function

$$P(t) = E\left[(x(t) - \hat{x}(t))(x(t) - \hat{x}(t))^\top \mid \mathcal{F}_t^Y\right]$$

is the estimation error variance.

The solution is based on the results of [23] and is given as follows.

III. Problem Reduction

It is considered that there is no useful information on the values of the unknown parameters $\theta_k(t)$, $k = 1,\ldots,p$ and this uncertainty even grows as time tends to infinity. In other words, the unknown parameters can be modeled as $\mathcal{F}_t$-measurable Wiener processes

$$d\theta(t) = \beta(t)dW_3(t), \tag{5}$$

with unknown initial condition $\theta(t_0) = \theta_0 \in R^p$, where $(W_3(t),\ \mathcal{F}_t,\ t \ge t_0)$ is a Wiener processes independent of x_0, $W_1(t)$ and of $W_2(t)$, $\beta(t) \in \Re^{p \times p}$ is an intensity matrix.

To apply the mean-square filtering equations from [23] to the state vector $z(t) = [x(t), \theta(t)]$, governed by equations (1) and (5) over the nonlinear polynomial observations (2), the state equation (1) should be transformed into a polymonial form and the observation equation (2) into a linear form. For this purpose, a vector $A_0(t) \in \Re^{n+p}$, a matrix $A_1(t) \in R^{(n+p)\times(n+p)}$, an $A_{p+1}(t)$ tensor of dimension $(n+p) \times (n+p) \ldots_{p+2} \ldots (n+p)$, a vector $C_0 \in \Re^m$, a matrix $C_1(t) \in \Re^{m \times (n+p)}$, an $(r+1)$D tensor $C_r(t) \in \Re^{m \times (n+p) \cdots_{(r+1) \text{ times}} \cdots (n+p)}$ are introduced as follows.

The equation for the i-th component of the state vector is given by

$$\begin{aligned} dx_i(t) = (F_{0_i}(t) + \sum_{j=1}^n F_{1_{ij}}(t)x_j(t) \\ +\sum_{j_1=1}^n \sum_{j_2=1}^n F_{2_{ij_1j_2}}(t)x_{j_1}(t)x_{j_2}(t) \\ +\sum_{j_1=1}^n \sum_{j_2=1}^n \cdots \sum_{j_p=1}^n F_{p_{ij_1j_2\cdots j_p}}(t)x_{j_1}(t)x_{j_2}(t)\ldots x_{j_p}(t))dt \\ +\sum_{j=1}^n b_{ij}(t)dW_{1_j}(t), \\ x_i(t_0) = x_{0_i}. \quad i = 1,\ldots,n \end{aligned}$$

Then: 1.

1) If the variable $F_{0_i}(t)$ is a known function, then the i-th component of the vector $A_0(t)$ is set to this function, $A_{0_i}(t) = F_{0_i}(t)$; otherwise, if the variable $F_{0_i}(t)$ is an unknown function, then the $(i, n+k_1)$-th entry of the matrix $A_1(t)$ is set to 1, where k_1 is the number of the current unknown parameter in the vector $F_0(t)$.
2) If the variable $F_{l_{ij_1j_2\ldots j_l}}(t)$ is a known function, then the $(i, j_1, j_2, \ldots, j_l)$-th component of the tensor $A_l(t)$ is set to this function, $A_{l_{ij_1j_2\ldots j_l}}(t) = F_{l_{ij_1j_2\ldots j_l}}(t)$; otherwise, if the variable $F_{l_{ij_1j_2\ldots j_l}}(t)$ is an unknown function, then the $(i, j_1, j_2, \ldots, j_l, n+k)$-th entry of the $(l+1)$D tensor $A_{l+1}(t)$ is set to 1, where k is the number of the current unknown entry in the matrix $F_{1_{ij_1j_2\ldots j_l}}(t)$, and other matrices of lower dimension

counting the unknown entries by rows from the first to the n-th entry in each row.

3) All other unassigned entries of the vector $A_0(t)$, matrix $A_1(t)$, lD-tensor $A_{l_{j_1 j_2 \dots j_l}}(t)$, are set to 0.

2.

1) $C_{0_i}(t) = H_{0_i}(t), \quad i = 1, \dots, m,$
2) $C_{1_{ij_1 j_2 \dots j_r}}(t) = H_{1_{ij_1 j_2 \dots j_r}}(t), \quad i = 1, \dots, m, \quad j_1, j_2, \dots, j_r = 1, \dots, n,$
3) All other unassigned entries of matrix $C_1(t)$, $(r+1)$D tensor $C_r(t)$ are set to 0.

Using the introduced notation, the equations for the extended state vector $z(t) = [x(t), \theta(t)] \in R^{n+p}$ and the polynomial observation process (2) can be rewritten like

$$\begin{gathered} dz(t) = (A_0(t) + A_1(t)z(t) + A_2(t)z(t)z^\top(t) + \dots \\ + A_{p+1}(t)z(t)z(t)\dots_{p+1 \text{ times}}\dots z^\top(t))dt \\ + \text{diag}[b(t), \beta(t)][dW_1^\top(t), dW_3^\top(t)]^\top, \\ z(t_0) = [x_0, \theta_0], \end{gathered} \tag{6}$$

$$\begin{gathered} dy(t) = (C_0(t) + C_1(t)z(t) + C_2(t)z(t)z^\top(t) \\ + \dots + C_r(t)z(t)z(t)\dots_{r \text{ times}}\dots z^\top(t))dt \\ + G(t)dW_2(t), \end{gathered} \tag{7}$$

where the vector $A_0(t)$, matrix $A_1(t)$, $\dots$, $(p+2)$D tensor $A_{p+1}(t)$, vector $C_0(t)$, $(r+1)$D tensor $C_r(t)$ have already been defined. The right-hand sides of (6) and (7) are polynomials with respect to the extended state vector $z(t) = [x(t), \theta(t)]$.

The estimation problem is now reformulated as to find the mean-square estimate $\hat{x}_z(t) = [\hat{x}_x(t), \hat{x}_\theta(t)]$ of the system state $z(t) = [x(t), \theta(t)]$, based on the observation process $Y(t) = \{y(s), 0 \le s \le t\}$. This estimate is given by the conditional expectation

$$\hat{x}_z(t) = [\hat{x}_x(t), \hat{x}_\theta(t)] = \left[E(x(t) \mid \mathcal{F}_t^Y), E(\theta(t) \mid \mathcal{F}_t^Y)\right]$$

of the system state $z(t) = [x(t), \theta(t)]$ with respect to the σ - algebra $\mathcal{F}_t^Y$ generated by the observation process $Y(t)$ in the interval $[t_0, t]$. The symmetric matrix function

$$\begin{gathered} P(t) = E\Big[\, ([x(t), \theta(t)] - [\hat{x}_x(t), \hat{x}_\theta(t)]) \\ \times ([x(t), \theta(t)] - [\hat{x}_x(t), \hat{x}_\theta(t)])^\top \mid \mathcal{F}_t^Y \Big] \end{gathered}$$

is the estimation error variance for this reformulated problem.

IV. Mean-Square Joint State Filter and Parameter Identifier Design

Let us reformulate the problem by introducing the stochastic process

$$\begin{gathered} z_1(t) = h(z(t), t) = C_0(t) + C_1(t)z(t) + C_2(t)z(t)z^\top(t) \\ + \dots + C_r(t)z(t)z(t)\dots_{r \text{ times}}\dots z^\top(t) \end{gathered}$$

Using the Ito formula (see [25], Section 5.10) for the stochastic differential of the nonlinear function $h(z,t)$, the following equation is obtained for $z_1(t)$

$$\begin{gathered} dz_1(t) = \tfrac{\partial h(z,t)}{\partial t}dt + \tfrac{\partial h(z,t)}{\partial z}(A_0(t) + A_1(t)z(t) + A_2(t) \\ \times z(t)z^\top(t) + \dots + A_{p+1}(t)z(t)z(t)\dots_{p+1 \text{ times}}\dots \\ z^\top(t))dt + \tfrac{\partial h(z,t)}{\partial z}(\text{diag}[b(t), \beta(t)][dW_1^\top(t), dW_3^\top(t)]^\top)dt \\ + \tfrac{1}{2}\text{tr}\tfrac{\partial^2 h(z,t)}{\partial z^2}\text{diag}[b(t), \beta(t)]\text{diag}[b(t), \beta(t)]^\top dt \end{gathered} \tag{8}$$

with the initial condition $z_1(0) = z_{10}$.

The initial condition $z_{10} \in R^m$ is considered a conditionally Gaussian random vector with respect to observations. This assumption is quite admissible in the filtering framework, since the real distributions of $z(t)$ and $z_1(t)$ are actually unknown. Indeed, as follows from [26], if only two lower conditional moments, expectation $\hat{x}_0$ and variance P_0, of a random vector $\hat{x}_0 = [z_{10}, z_0]$ are available, the Gaussian distribution with the same parameters, $N(\hat{x}_0, P_0)$, is the best approximation for the unknown conditional distribution of $\hat{x}_0 = [z_{10}, z_0]$ with respect to observations. This fact is also a corollary of the central limit theorem [27] in the probability theory.

A. Case study: Second degree polynomial state and second degree polynomial observations

Let us consider the second degree polynomial functions

$$\begin{gathered} F(x(t), \theta(t), t) = F_0(\theta(t), t) + F_1(\theta(t), t)x(t) \\ + F_2(\theta(t), t)x(t)x^\top(t) \\ h(x(t), t) = H_0(t) + H_1(t)x(t) + H_2(t)x(t)x^\top(t) \end{gathered} \tag{9}$$

where $x(t)$ is an n-dimensional vector and $F_0(\theta(t), t)$, $F_1(\theta(t), t)$, $F_2(\theta(t), t)$, $H_0(t)$, $H_1(t)$ and $H_2(t)$ were previously defined.

In this case, equations (6)-(7) take the following form

$$\begin{gathered} dz(t) = (A_0(t) + A_1(t)z(t) + A_2(t)z(t)z^\top(t) \\ + A_3(t)z(t)z(t)z^\top(t))dt \\ + \text{diag}[b(t), \beta(t)][dW_1^\top(t), dW_3^\top(t)]^\top, \\ z(t_0) = [x_0, \theta_0], \end{gathered} \tag{10}$$

$$\begin{gathered} dy(t) = \left(C_0(t) + C_1(t)z(t) + C_2(t)z(t)z^\top(t)\right) dt \\ + G(t)dW_2(t), \end{gathered} \tag{11}$$

and $z_1(t) = C_0(t) + C_1(t)z(t) + C_2(t)z(t)z^\top(t)$.

Upon calculating the partial derivatives of $h(z,t)$ in equation (9), equation (8) takes the form

$$\begin{gathered} dz_1(t) = (\dot{C}_0(t) + \dot{C}_1(t)z(t) + \dot{C}_2(t)z(t)z^\top(t))dt \\ + C_1(t)[A_0(t) + A_1(t)z(t) + A_2(t)z(t)z^\top(t) \\ + A_3(t)z(t)z(t)z^\top(t)]dt + C_2(t)z(t)[A_0(t) \\ + A_1(t)z(t) + A_2(t)z(t)z^\top(t) + A_3(t)z(t)z(t)z^\top(t)]^\top dt \\ + C_2[A_0(t) + A_1(t)z(t) + A_2(t)z(t)z^\top(t) \\ + A_3(t)z(t)z(t)z^\top(t)]z^\top(t)dt + C_1(t) \\ \times \text{diag}[b(t), \beta(t)][dW_1^\top(t), dW_3^\top(t)]^\top + C_2(t) \\ \times z(t)(\text{diag}[b(t), \beta(t)][dW_1^\top(t), dW_3^\top(t)]^\top)^\top \\ + C_2(t)\text{diag}[b(t), \beta(t)][dW_1^\top(t), dW_3^\top(t)]^\top z^\top(t) \end{gathered} \tag{12}$$

with the initial condition $z_1(0) = z_{10}$. Equation (11) can be written in the form

$$dy(t) = z_1(t)dt + G(t)dW_2(t). \tag{13}$$

Thus, the estimation problem is reformulated as to find the mean-square estimate $\hat{x}(t) = [\hat{x}_z(t), \hat{x}_{z_1}(t)] = [\hat{x}_1(t) = (\hat{x}_x(t), \hat{x}_\theta(t)), \hat{x}_2(t)]$, for the state vector $[z(t) = (x(t), \theta(t)), z_1(t)]$ governed by the polynomial equations (10),(12), that is based on the observation process $Y(t) = \{y(s), 0 \le s \le t\}$ satisfying the equation (13). The solution of this problem is obtained using the mean-square filtering equations for fourth degree polynomial states over linear observations [23] and given by

$$d\hat{x}_1(t) = (A_0(t) + A_1(t)\hat{x}_1(t) + A_2(t)[\hat{x}_1(t)\hat{x}_1^\top(t) + P_{11}(t)] + A_3(t)[3\hat{x}_1(t)P(t) + \hat{x}_1(t)\hat{x}_1(t)\hat{x}_1^\top(t)])dt + P_{12}(t)(G(t)G^\top(t))^{-1}(dy(t) - \hat{x}_2(t)dt), \tag{14}$$

$$d\hat{x}_2(t) = (\dot{C}_0(t) + \dot{C}_1(t)\hat{x}_1(t) + \dot{C}_2(t)(\hat{x}_1(t)\hat{x}_1^\top(t) + P_{11}(t)))dt + C_1(t)(A_0(t) + A_1(t)\hat{x}_1(t) + A_2(t)(\hat{x}_1(t)\hat{x}_1^\top(t) + P_{11}(t)) + A_3(t)(3\hat{x}_1(t)P(t) + \hat{x}_1(t)\hat{x}_1(t)\hat{x}_1^\top(t)))dt + C_2(t)(\hat{x}_1(t)A_0^\top(t) + [\hat{x}_1(t)\hat{x}_1^\top(t) + P_{11}(t)]A_1^\top(t) + [3\hat{x}_1(t)P_{11}(t) + \hat{x}_1(t)\hat{x}_1(t)\hat{x}_1^\top(t)] \times A_2^\top(t) + 3[P_{11}(t)P_{11}(t) + \hat{x}_1(t)\hat{x}_1^\top(t)P_{11}(t) + P_{11}(t)\hat{x}_1(t)\hat{x}_1^\top(t) + (\hat{x}_1(t)\hat{x}_1^\top(t))^2]A_3^\top(t))dt + C_2(t)(\hat{x}_1(t)A_0^\top(t) + [\hat{x}_1(t)\hat{x}_1^\top(t) + P_{11}(t)]A_1^\top(t) + [3\hat{x}_1(t)P_{11}(t) + \hat{x}_1(t)\hat{x}_1(t)\hat{x}_1^\top(t)]A_2^\top(t) + 3[P_{11}(t)P_{11}(t) + \hat{x}_1(t)\hat{x}_1^\top(t)P_{11}(t) + P_{11}(t)\hat{x}_1(t) \times\hat{x}_1^\top(t) + (\hat{x}_1(t)\hat{x}_1^\top(t))^2]A_3^\top(t))^\top dt + P_{22}(t)(G(t)G^\top(t))^{-1}(dy(t) - \hat{x}_2(t)dt), \tag{15}$$

with the initial conditions

$$\hat{x}_1(t_0) = E\left[z(t_0) \mid \mathcal{F}_{t_0}^Y\right], \quad \hat{x}_2(t_0) = E\left[z_1(t_0) \mid \mathcal{F}_{t_0}^Y\right],$$

and

$$dP_{11}(t) = (A_1(t)P_{11}(t) + 2A_2(t)\hat{x}_1(t)P_{11}(t) + 2(A_2(t)\hat{x}_1(t)P_{11}(t))^\top + 3(\hat{x}_1(t)\hat{x}_1^\top(t) \times P_{11}(t) + P_{11}(t)P_{11}(t)) + P_{11}(t)A_1^\top(t) + 3(\hat{x}_1(t)\hat{x}_1^\top(t)P_{11}(t) + P_{11}(t)P_{11}(t))^\top + L(t)L^\top(t) - P_{12}(t)(G(t)G^\top(t))^{-1}P_{21}(t))dt \tag{16}$$

$$dP_{12}(t) = (A_1(t)P_{12}(t) + 2A_2(t)\hat{x}_1(t)P_{12}(t) + 3A_3(t) \times(P_{11}(t)P_{12}(t) + \hat{x}_1\hat{x}_1^\top(t)P_{12}(t)) + (\dot{C}_1(t)P_{11}(t) + 2\dot{C}_2(t)\hat{x}_1(t)P_{11}(t))^\top + (C_1(t)(A_1(t)P_{11}(t) + 2A_2(t) \times\hat{x}_1(t)P_{11}(t)) + 3A_3(t)(P_{11}(t)P_{11}(t) + \hat{x}_1(t) \times\hat{x}_1^\top(t)P_{11}(t)))^\top + (C_2(t)(A_0(t)P_{11} + 2A_1(t)\hat{x}_1(t) \times P_{11}(t) + 3A_2(t)(P_{11}(t)P_{11}(t) + \hat{x}_1(t)\hat{x}_1^\top(t)P_{11}(t)) + 4A_3(t)(\hat{x}_1(t)\hat{x}_1(t)\hat{x}_1^\top(t)P_{11}(t) + 3\hat{x}_1(t)P_{11}(t) \times P_{11}(t))))^\top + (C_2(t)(A_0(t)P_{11} + 2A_1(t)\hat{x}_1(t)P_{11}(t) + 3A_2(t)(P_{11}(t)P_{11}(t) + \hat{x}_1(t)\hat{x}_1^\top(t)P_{11}(t)) + 4A_3(t)(\hat{x}_1(t)\hat{x}_1(t)\hat{x}_1^\top(t)P_{11}(t) + 3\hat{x}_1(t) \times P_{11}(t)P_{11}(t)))^\top)^\top + L(t)L^\top(t)C_1^\top(t) + L(t)L^\top(t) \times(C_2(t)\hat{x}_1(t))^\top + (L(t)L^\top(t)(C_2(t)\hat{x}_1(t))^\top)^\top - P_{12}(t)(G(t)G^\top(t))^{-1}P_{22}(t))dt, \tag{17}$$

$$dP_{22}(t) = (\dot{C}_1P_{12}(t) + 2\dot{C}_2\hat{x}_1P_{12}(t) + (\dot{C}_1P_{12}(t) + 2\dot{C}_2\hat{x}_1P_{12}(t))^\top + C_1(t)(A_1(t)P_{12}(t) + 2A_2(t)\hat{x}_1(t) \times P_{12}(t) + 3A_3(t)(P_{11}(t)P_{12}(t) + \hat{x}_1(t)\hat{x}_1^\top(t)P_{12}(t))) + (C_1(t)(A_1(t)P_{12}(t) + 2A_2(t)\hat{x}_1(t)P_{12}(t) + 3A_3(t) \times(P_{11}(t)P_{12}(t) + \hat{x}_1(t)\hat{x}_1^\top(t)P_{12}(t))))^\top + C_2(t)(A_0(t) \times P_{12}(t) + 2A_1(t)\hat{x}_1(t)P_{12}(t) + 3A_2(t)(P_{11}(t)P_{12}(t) + \hat{x}_1(t)\hat{x}_1^\top(t)P_{12}(t)) + 4A_3(t)(\hat{x}_1(t)\hat{x}_1(t)\hat{x}_1^\top(t) \times P_{12} + 3\hat{x}_1(t)P_{11}P_{12}(t))) + C_2(t)(A_0(t)P_{12}(t) + 2A_1(t) \times\hat{x}_1(t)P_{12}(t) + 3A_2(t)(P_{11}(t)P_{12}(t) + \hat{x}_1(t)\hat{x}_1^\top(t) \times P_{12}(t)) + 4A_3(t)(\hat{x}_1(t)\hat{x}_1(t)\hat{x}_1^\top(t)P_{12} + 3\hat{x}_1(t) \times P_{11}(t)P_{12}(t)))^\top + (C_2(t)(A_0(t)P_{12}(t) + 2A_1(t)\hat{x}_1(t) \times P_{12}(t) + 3A_2(t)(P_{11}(t)P_{12}(t) + \hat{x}_1(t)\hat{x}_1^\top(t)P_{12}(t)) + 4A_3(t)(\hat{x}_1(t)\hat{x}_1(t)\hat{x}_1^\top(t)P_{12} + 3\hat{x}_1(t)P_{11}(t)P_{12}(t))))^\top + (C_2(t)(A_0(t)P_{12}(t) + 2A_1(t)\hat{x}_1(t)P_{12}(t) + 3A_2(t)(P_{11}(t)P_{12}(t) + \hat{x}_1(t)\hat{x}_1^\top(t)P_{12}(t)) + 4A_3(t)(\hat{x}_1(t)\hat{x}_1(t)\hat{x}_1^\top(t)P_{12} + 3\hat{x}_1(t)P_{11} \times P_{12}(t)))^\top)^\top + C_1(t)L(t)L^\top(t)C_1^\top + C_1(t)L(t)L(t) \times(C_2(t)\hat{x}_1(t))^\top + (C_1(t)L(t)L(t)(C_2(t)\hat{x}_1(t))^\top)^\top + C_1(t)L(t)(C_2(t)\hat{x}(t))L^\top(t) + (C_1(t)L(t)(C_2(t)\hat{x}(t)) \times L^\top(t))^\top + C_2(t)\hat{x}_1(t)L(t)P_{11}(t)C_2(t)\hat{x}_1(t)L^\top(t) + (C_2(t)\hat{x}_1(t)L(t)P_{11}(t)C_2(t)\hat{x}_1(t)L^\top(t))^\top + C_2(t)\hat{x}_1(t)L^\top(t)P_{11}(t)C_2(t)\hat{x}_1(t)L^\top(t) + (C_2(t)\hat{x}_1(t)L^\top(t)P_{11}(t)C_2(t)\hat{x}_1(t)L^\top(t))^\top - P_{22}(t)G(t)G^\top(t)P_{22}(t))dt \tag{18}$$

with the initial condition

$$P(t_0) = E[\big([z(t_0), z_1(t_0)] - [\hat{x}_1(t_0), \hat{x}_2(t_0)]\big) \times\big([z(t_0), z_1(t_0)] - [\hat{x}_1(t_0), \hat{x}_2(t_0)]\big)^\top \mid \mathcal{F}_{t_0}^Y].$$

and

$$L(t) = \mathrm{diag}[b(t), \beta(t)]$$

Theorem 1. The mean-square filter for the extended state vector $[x(t), \theta(t), z_1(t)]$, governed by the equations (1),(5),(12), over the linear observations (13) is given by the equations (14)-(15) for the mean-square estimate $\hat{x}(t) = [\hat{x}_1(t) = (\hat{x}_x(t), \hat{x}_\theta(t)), \hat{x}_2(t) = \hat{x}_{z_1}] = E([z(t) = (x(t), \theta(t)), z_1(t)] \mid \mathcal{F}_t^Y)$ and the equations (16)–(18) for the estimation error variance $P(t) = E(([z(t), z_1(t)] - [\hat{x}_1(t), \hat{x}_2(t)])([z(t), z_1(t)] - [\hat{x}_1(t), \hat{x}_2(t)])^\top \mid \mathcal{F}_t^Y)$. This filter, applied to the subvector $\theta(t)$, also serves as the identifier for the vectors of unknown parameters $\theta(t)$ in the equation (1), yielding the estimate subvector $\hat{x}_\theta(t)$ as the parameter estimates.

Proof: The proof directly follows from the steps 1–3 for designing the coefficients in the extended state equations (10),(12) and the mean-square filtering equations (14)-(18) for fourth degree polynomial states over linear observations which were obtained in [23]. ■

V. Example

This section presents an example of designing the mean-square filter for a nonlinear polynomial stochastic system with a multiplicative unknown parameter in the state equation, where a conditionally Gaussian state initial condition for the extended state vector is additionally assumed.

Let the real scalar state variable $x(t)$ satisfy the nonlinear equation with an unknown multiplicative parameter

$$dx(t) = (1 + \theta x^2(t))dt + dW_1(t) \quad x(0) = x_0, \tag{19}$$

and the scalar observation process be given by the nonlinear equation

$$dy(t) = x^2(t)dt + dW_2(t) \tag{20}$$

where $W_1(t)$ and $W_2(t)$ are Wiener processes, independent of each other and of a Gaussian random variable x_0 serving as the initial condition in (19). Equations (19) and (20) represent the conventional form for the equations (1) and (2), where $F(\theta(t), x(t))$ and $h(x(t))$ satisfy the equations given in (9). The parameter $\theta(t)$ is modeled as a standard Wiener process, i.e., satisfy the equation ($\beta = 1$)

$$d\theta(t) = dW_3(t), \quad \theta(0) = \theta_0, \tag{21}$$

which can also be rewritten as

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

where $\psi_3(t)$ is a white Gaussian noise. The Wiener process $W_3(t)$ is independent of x_0, $W_1(t)$, and $W_2(t)$.

The filtering problem is to find the mean-square estimate $\hat{x}_z(t) = [\hat{x}_x(t), \hat{x}_\theta(t)]$ for the nonlinear state (19) and (20), $z(t) = [x(t), \theta(t)]$, using the nonlinear observation (20), confused with independent and identically distributed disturbances modeled as white Gaussian noises.

Let us reformulate the problem by introducing the stochastic process $z_1(t) = h(x,t) = x^2(t)$. Ito formula (see [25]) is used for the stochastic differential of $x^2(t)$, where $x(t)$ satisfies the equation (19), the following equation is obtained for $z_1(t)$

$$\begin{aligned} dz_1(t) = &\quad (1 + 2\theta(t)x(t)z_1(t))dt + 2x(t)dW_1(t), \\ z_1(0) = &\quad z_{10}. \end{aligned} \tag{22}$$

The initial condition $z_{10} \in R$ is considered a conditionally Gaussian random variable with respect to observations. This assumption is quite admissible in the filtering framework, since the real distributions of $z(t)$ and $z_1(t)$ are unknown. In terms of the process $z_1(t)$, the observation equation (20) takes the form

$$dy(t) = z_1(t)dt + dW_2(t). \tag{23}$$

The state equation (19) can be written as

$$dx(t) = 1 + \theta(t)z_1(t)dt + dW_1(t) \tag{24}$$

The obtained filtering system includes three equations, (24), (21) and (22), for the partially measured state $[z(t) = (x(t), \theta(t), z_1(t)]$ and an equation (23) for the observations $y(t)$, where $z_1(t)$ is a completely measured third degree state, $z(t) = [x(t), \theta(t)]$ is an unmeasured second degree state, and $y(t)$ is a linear observation process directly measuring the state $z_1(t)$. Thus, the estimation problem is reformulated as to find the mean-square estimate $\hat{x}(t) = [\hat{x}_z(t), \hat{x}_{z_1}(t)] = [\hat{x}_1(t) = (\hat{x}_x(t), \hat{x}_\theta(t)), \hat{x}_2(t)]$, for the state vector $[z(t) = (x(t), \theta(t)), z_1(t)]$ governed by equations (24), (21) and (22), that is based on the observation process $Y(t) = \{y(s), 0 \leq s \leq t\}$ satisfying the equation (23).

The filtering equations (14)–(18) take the following form for the system (24), (21) and (23)

$$\begin{aligned} d\hat{x}_x(t) = &\quad (1 + \hat{x}_2(t)\hat{x}_\theta(t) + P_{\theta 2}(t))dt \\ &\quad + P_{x2}(t)(dy(t) - \hat{x}_2(t)dt) \\ d\hat{x}_\theta(t) = &\quad P_{\theta 2}(t)(dy - d\hat{x}_2(t)) \\ d\hat{x}_2(t) = &\quad (1 + 2\hat{x}_x(t)\hat{x}_2(t)\hat{x}_\theta(t) + 6\hat{x}_x(t) \\ &\quad \times P_{\theta 2}(t))dt + P_{22}(t)(dy(t) - \hat{x}_2(t)dt) \end{aligned} \tag{25}$$

with the initial conditions $\hat{x}_x(0) = E(x_0 \mid y(0))$, $\hat{x}_\theta(0) = E(\theta_0 \mid y(0))$ and $\hat{x}_2(0) = E(x_0^2 \mid y(0))$, and

$$\begin{aligned} dP_{xx}(t) = &\quad (1 + 4\hat{x}_2(t)P_{x\theta}(t))dt - P_{x2}^2(t)dt \\ dP_{x\theta}(t) = &\quad (2\hat{x}_2 P_{\theta\theta}(t) - P_{x2}P_{\theta 2}(t))dt \\ dP_{x2}(t) = &\quad (2\hat{x}_2 P_{\theta 2}(t) + 6\hat{x}_x(t)\hat{x}_2(t)P_{x\theta}(t))dt \\ &\quad + (6P_{x2}(t)P_{x\theta}(t) + 2\hat{x}_x(t) - P_{x2}(t)P_{22}(t))dt \\ dP_{\theta\theta}(t) = &\quad (\beta^2 - P_{\theta 2}^2(t))dt \\ dP_{\theta 2}(t) = &\quad (6\hat{x}_x(t)\hat{x}_2(t)P_{\theta\theta}(t) + 6P_{x2}(t)P_{\theta\theta}(t))dt \\ &\quad - P_{22}(t)P_{\theta 2}(t)dt \\ dP_{22}(t) = &\quad 12\hat{x}_x(t)\hat{x}_2(t)P_{\theta 2}(t)dt + 12P_{x2}(t)P_{\theta 2}(t)dt \\ &\quad + (4\hat{x}_x^2(t) + 4P_{xx}(t) - P_{22}^2(t))dt \end{aligned} \tag{26}$$

with the initial condition

$$\begin{aligned} P(0) = E\Big(&\big([x_0, \theta_0, z_{10}] - [\hat{x}_x(0), \hat{x}_\theta(0), \hat{x}_2(0)] \big) \\ &\times \big([x_0, \theta_0, z_{10}] - [\hat{x}_x(0), \hat{x}_\theta(0), \hat{x}_2(0)] \big)^\top \mid y(0) \Big). \end{aligned}$$

Here, $\hat{x}_x(t)$ is the estimate for the state $x(t)$, $\hat{x}_\theta(t)$ is the estimate for the state $\theta(t)$ and $\hat{x}_2(t)$ is the estimate for the state $z_1(t) = x^2(t)$.

Numerical simulation results are obtained by solving the systems of filtering equations (25)–(26). For the filter (25)–(26) and the reference system (21)–(24), involved in simulation, the following initial values are assigned: $x(0) = .4$, $\hat{x}_2(0) = 1$, $P_{xx}(0) = 180$, $P_{x\theta}(0) = 12$, $P_{x2}(0) = 12$, $P_{\theta\theta}(0) = 50$, $P_{\theta 2}(0) = 56$, $P_{22}(0) = 150$. The unknown parameter θ in the state equation is assigned as $\theta = -0.5$ in the first simulation and as $\theta = 0.5$ in the second one, thus considering stable and unstable cases. Gaussian disturbances $dW_1(t)$, $dW_2(t)$, $dW_3(t)$ are realized using the built-in MatLab white noise functions.

Figure 1 shows the graphs of the estimate for the parameter $\hat{x}_\theta(t)$, (for $\theta = -0.5$), the reference state variable $x(t)$ and its estimate $\hat{x}_x(t)$, the state $z_1(t) = x^2(t)$ and its estimate $\hat{x}_2(t)$, in the simulation interval [0,1]. Figure 2 shows the graphs of the estimate for the parameter $\hat{x}_\theta(t)$, (for $\theta = 0.5$), the reference state variable $x(t)$ and its estimate $\hat{x}_x(t)$, the state $z_1(t) = x^2(t)$ and its estimate $\hat{x}_2(t)$, in the simulation interval [0,0.7]. The simulation results show very reliable behavior of the designed filter and parameter identifier in both cases.

REFERENCES

[1] Bar–Shalom Y. Optimal simultaneous state estimation and parameter identification in linear discrete-time systems, *IEEE Trans. Automat. Contr.* 1972; **17**: 308-319.

[2] Rao CR. *Linear Statistical Inference and its Applications*, New York: Wiley-Interscience, 1973.

[3] Elliott RJ, Krishnamurthy V. New finite–dimensional filters for parameter estimation of discrete–time linear gaussian models, *IEEE Trans. Automat. Contr.* 1999; **44**: 938-951.
[4] Elliott RJ, Krishnamurthy V. New finite–dimensional filters for estimation of continuous–time linear gaussian systems, *SIAM. J. Contr. Optim.* 1997; **35**: 1908-1923.
[5] Duncan TE, Mandl P, Pasik–Duncan B. A note on sampling and parameter estimation in linear stochastic systems, *IEEE Transactions on Automatic Control*, 1999; **44**: 2120-2125.
[6] Charalambous CD, Logothetis A. Maximum likelihood parameter estimation from incomplete data via the sensitivity equations: The continuous–time case, *IEEE Transactions on Automatic Control* 2000; **45**: 928–934.
[7] Zheng WX. On unbiased parameter estimation of linear systems using noisy measurements, *Cybernetics and Systems* 2003; **34**: 59-70.
[8] Shi P. Filtering on sampled-data systems with parametric uncertainty, *IEEE Trans. Automat. Contr.* 1998; **43**: 1022–1027.
[9] Shi P., Boukas E., Agarwal R. K. Control of markovian jump discrete-time systems with norm-bounded uncertainty and unknown delay, *IEEE Trans. Automat. Contr.* 1999; **44**: 2139–2144.
[10] Xu S., van Dooren P. V. Robust H_∞–filtering for a class of nonlinear systems with state delay and parameter uncertainty, *Int. J. Control*, 2002; **75**: 766–774.
[11] Yang F., Wang W., Hung Y. S. Robust Kalman filtering for discrete time-varying uncertain systems with multiplicative noises, *IEEE Transactions on Automatic Control*, 2002; **47**: 1179–1183.
[12] Germani A, Manes P., Palumbo P. Linear filtering for bilinear stochastic differential systems with unknown inputs, *IEEE Transactions on Automatic Control* 2002; **47**: 1726-1730.
[13] Shin DR, Verriest E. Optimal access control of simple integrated networks with incomplete observations, *Proc. American Control Conf.* 1994; Baltimore, MD, USA, 3487-3488.
[14] Zhang WH, Chen BS, Tseng CS. Robust H_∞ filtering for nonlinear stochastic systems, *IEEE Transactions on Signal Processing* 2005; **53**: 589–598.
[15] Wang Z. Filtering on nonlinear time-delay stochastic systems, *Automatica* 2003, **39**: 101-109.
[16] 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.
[17] Sheng J, Chen T, Shah SL. Optimal filtering for multirate systems. *IEEE Transactions on Circuits and Systems* 2005; **52**: 228-232.
[18] Sheng J. Optimal filtering for multirate systems based on lifted models. *Proc. American Control Conf.* 2005; Portland, OR, USA, 3459-3461.
[19] 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.
[20] Jeong CS, Yaz E, Bahakeem A, Yaz Y. Nonlinear observer design with general criteria, *International Journal of Innovative Computing, Information and Control* 2006; **2**: 693-704.
[21] Shen B., Wang Z., Shu H., Wei G. On nonlinear H-infinity filtering for discrete-time stochastic systems with missing measurements, *IEEE Transactions on Automatic Control* 2008; **53**: 2170-2180.
[22] Wang Z., Liu Y., Liu Y. H-infinity filtering for uncertain stochastic time-delay systems with sector-bounded nonlinearities, *Automatica*; 2008; **44**: 1268-1277.
[23] Basin MV, Calderon-Alvarez D, Skliar M. Optimal filtering for incompletely measured polynomial states over linear observations, *International J. Adaptive Control and Signal Processing* 2008; **22**: 482–494.
[24] Basin MV, Perez J, Skliar M. Optimal filtering for polynomial system states with polynomial multiplicative noise, *International J. of Robust and Nonlinear Control* 2006; **16**: 287–298.
[25] Pugachev VS, Sinitsyn IN. *Stochastic Systems: Theory and Applications*. World Scientific, 2001.
[26] Pugachev VS. *Probability Theory and Mathematical Statistics for Engineers*. Pergamon, 1984.
[27] Tucker HG. *A Graduate Course in Probability*. Academic Press, 1967.

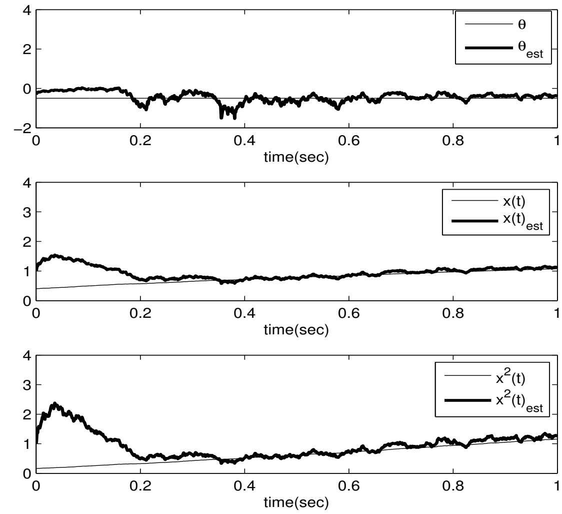

Fig. 1. Negative parameter value of θ. **Above.** Graphs of the parameter θ (thin line) and its estimate $\hat{x}_\theta(t)$ (dotted line). **Middle.** Graphs of the state $x(t)$ (thin line) and its estimate $\hat{x}_x(t)$ (dotted line). **Below.** Graphs of the state $z_1(t)$ (thin line) and its estimate $\hat{x}_{z_1}(t)$ (dotted line) in the interval $[0, 1]$.

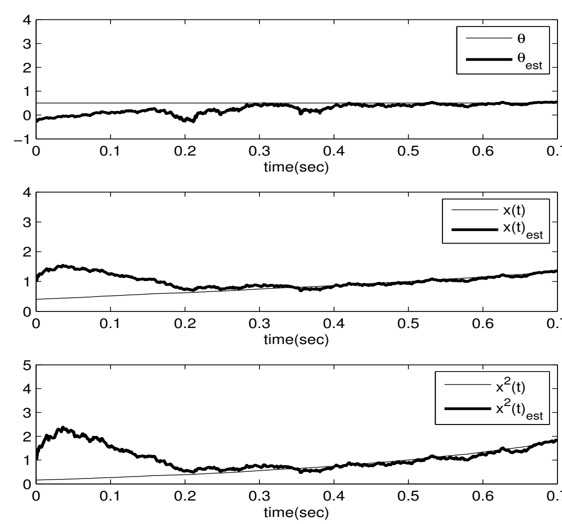

Fig. 2. Positive parameter value of θ. **Above.** Graphs of the parameter θ (thin line) and its estimate $\hat{x}_\theta(t)$ (dotted line). **Middle.** Graphs of the state $x(t)$ (thin line) and its estimate $\hat{x}_x(t)$ (dotted line). **Below.** Graphs of the state $z_1(t)$ (thin line) and its estimate $\hat{x}_{z_1}(t)$ (dotted line) in the interval $[0, 0.7]$.