2009 American Control Conference
Hyatt Regency Riverfront, St. Louis, MO, USA
June 10-12, 2009

Optimal Filtering over Linear Observations with Unknown Parameters

Michael Basin Dario Calderon-Alvarez

Abstract— **This paper presents the optimal filtering and parameter identification problem for linear stochastic systems over linear observations with unknown parameters, 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 resulting filtering system is bilinear in state and linear in observations. The obtained optimal filter for the extended state vector also serves as the optimal identifier for the unknown parameters. Performance of the designed optimal 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 optimal state and parameter estimates dynamics in the "general situation," without imposing special assumptions on the system structure. Other approaches are presented by the optimal 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]–[21] for linear stochastic systems with bounded uncertainties in coefficients or missing measurements. The overall comment is that the optimal state filter and parameter identifier in the form of a closed finite-dimensional system of stochastic ODEs has not yet been obtained even in case of linear observations with unknown parameters.

A practical example where an observation matrix contains unknown parameters (matrix entries) can be encountered in the process of measuring some components of the "seeming velocity" of flying objects (see [22]). In this case, direct linear observation of vertical components of the seeming velocity is distorted by an unknown varying factor, which itself depends on the velocity fluctuations. Given that the velocity fluctuations behave similarly to Wiener processes, the modeling of unknown parameters in the observation matrix as Wiener processes is very appropriate.

This paper presents the optimal joint filtering and parameter identification problem for linear stochastic systems with unknown parameters over linear observations. The solution starts with reduction of the original identification problem to the optimal filtering problem for linear system states over bilinear (second degree 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 linear state vector, which should be estimated mean-square optimally over bilinear observations. The obtained filtering problem is then further reduced to the filtering problem for bilinear system states over direct linear observations, assuming the bilinear 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 optimal filter for bilinear polynomial states over linear observations ([23]). The designed optimal filter for the extended state vector also serves as the optimal identifier for the unknown parameters. The proposed algorithm was partially used for solving the optimal joint state filtering and parameter identification problem for linear stochastic systems with unknown parameters in the state equation ([24]).

In the illustrative example, performance of the designed optimal filter is verified for a linear system state over linear observations with a multiplicative unknown parameter. Both, positive and negative, parameter values 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.

The paper is organized as follows. Section 2 presents the filtering and parameter identification problem statement for a linear system state over linear observations with unknown multiplicative and additive parameters. In Section 3, the stated problem is reduced to the filtering problem for an extended state vector that incorporates parameters as additional states. Section 4 presents the solution to the reduced filtering problem based on the optimal filter from [23]. An example of applying the designed identification technique is

The authors thank the Mexican National Science and Technology Council (CONACyT) for financial support under Grants 52953 and 55584.

The authors are with Department of Physical and Mathematical Sciences, Autonomous University of Nuevo Leon, San Nicolas de los Garza, Nuevo Leon, Mexico mbasin@fcfm.uanl.mx dcalal@hotmail.com

given in Section 5.

II. Filtering Problem for Linear States over Linear Observations with Unknown Parameters

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 a linear differential equation for the system state

$$dx(t) = (a_0(t) + a(t)x(t))dt + b(t)dW_1(t), \quad x(t_0) = x_0, \tag{1}$$

and a linear differential equation with unknown parameters for the observation process

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

Here, $x(t) \in R^n$ is the state vector and $y(t) \in R^m$ is the observation vector, $m \leq n$, and $\theta(t) \in R^p$, $p \leq m \times n + m$, is the vector of unknown entries of the matrix $A(\theta, t)$ and unknown components of vector $A_0(\theta, t)$. The latter means that both structures contain unknown components $A_{0_i}(t) = \theta_k(t)$, $k = 1, \ldots, p_1 \leq n$ and $A_{ij}(t) = \theta_k(t)$, $k = 1, \ldots, p \leq m \times n + m$, as well as known components $A_{0_i}(t)$ and $A_{ij}(t)$, whose values are known functions of time. 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 $B(t)B^T(t)$ is a positive definite matrix. All coefficients in (1)–(2) are deterministic functions of time of appropriate dimensions.

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, this optimal estimate is given by the conditional expectation $\hat{x}(t) = m(t) = E(x(t) \mid F_t^Y)$ 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 symmetric matrix function $P(t) = E[(x(t) - m(t))(x(t) - m(t))^T \mid F_t^Y]$ is the estimation error variance.

The proposed solution is based on the results of [23] and given in the following two sections.

III. Problem Reduction

It is considered that there is no useful information on 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 (see [22] for specific examples) as F_t-measurable Wiener processes

$$d\theta(t) = \beta(t)dW_3(t), \tag{3}$$

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

Note that a practical example where an observation matrix contains unknown parameters (matrix entries) can be encountered in the process of measuring some components of the "seeming velocity" of flying objects (see [22]). In this case, direct linear observation of vertical components of the seeming velocity is distorted by an unknown varying factor, which itself depends on the velocity fluctuations. Given that the velocity fluctuations behave similarly to Wiener processes, the modeling of unknown parameters in the observation matrix as Wiener processes is very appropriate.

To apply the optimal filtering equations from [23] to the state vector $z(t) = [x(t), \theta(t)]$, governed by the equations (1) and (3), over the linear observations (2), the observation equation (2) should be transformed into the polynomial form. For this purpose, a matrix $A_1(t) \in R^{m \times (n+p)}$, a cubic tensor $A_2(t) \in R^{m \times (n+p) \times (n+p)}$, and a vector $C_0(t) \in R^m$ are introduced as follows.

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

$$dy_i(t) = (A_{0_i}(t) + \sum_{j=1}^{m} A_{ij}(t)x_j(t))dt + \sum_{j=1}^{m} B_{ij}(t)dW_{2_j}(t),$$

Then:

1. If the variable $A_{0_i}(t)$ is a known function, then the i-th component of the vector $C_0(t)$ is set to this function, $C_{0_i}(t) = A_{0_i}(t)$; otherwise, if the variable $A_{0_i}(t)$ is an unknown function, then the $(i, n+i)$-th entry of the matrix $A_1(t)$ is set to 1.

2. If the variable $A_{ij}(t)$ is a known function, then the (i, j)-th component of the matrix $A_1(t)$ is set to this function, $A_{1_{ij}}(t) = A_{ij}(t)$; otherwise, if the variable $A_{ij}(t)$ is an unknown function, then the $(i, n+k, j)$-th entry of the cubic tensor $A_2(t)$ is set to 1, where k is the number of this current unknown entry in the matrix $A_{ij}(t)$, counting the unknown entries consequently by rows from the first to n-th entry in each row.

3. All other unassigned entries of the matrix $A_1(t)$, cubic tensor $A_2(t)$, and vector $C_0(t)$ are set to 0.

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

$$dz(t) = (c_0(t) + a_1(t)z(t))dt +$$

$$diag[b(t), \beta(t)]d[W_1^T(t), W_3^T(t)]^T, \tag{4}$$

$$z(t_0) = [x_0, \theta_0],$$

$$dy(t) = (C_0(t) + A_1(t)z(t) + A_2(t)z(t)z^T(t))dt + \tag{5}$$

$$B(t)dW_2(t),$$

where $c_0(t) = [a_0(t), 0_{p\times 1}]$, $a_1(t) = diag[a(t), 0_{p\times p}]$, and the matrix $A_1(t)$, cubic tensor $A_2(t)$, and vector $C_0(t)$ have already been defined. The equation (4) is linear with respect to the extended state vector $z(t) = [x(t), \theta(t)]$ and the new observation equation (5) is a second degree polynomial with respect to the state $z(t) = [x(t), \theta(t)]$.

The reformulated estimation problem is now to find the optimal estimate $m_z(t)=[m_x(t),m_\theta(t)]$ 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 optimal estimate is given by the conditional expectation

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

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

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

is the estimation error variance for this reformulated problem.

IV. Optimal Filter for Linear State over Linear Observations with Unknown Parameters

Let us reformulate the problem again, introducing the stochastic process $z_1(t)=h(z,t)=C_0(t)+A_1(t)z(t)+A_2(t)z(t)z^T(t)$. Using the Ito formula (see [25]) for the stochastic differential of the nonlinear function $h(z,t)$, the following equation is obtained for $z_1(t)$

$$dz_1(t)=\frac{\partial(h(z,t))}{\partial z}(c_0(t)+a_1(t)z(t))dt+\frac{\partial(h(z,t))}{\partial t}dt+$$
$$\frac{1}{2}\frac{\partial^2(h(z,t))}{\partial z^2}diag[b(t),\beta(t)]diag[b(t),\beta(t)]^Tdt+$$
$$\frac{\partial(h(z,t))}{\partial z}diag[b(t),\beta(t)]d[W_1^T(t),W_3^T(t)]^T, \qquad (6)$$

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

Note that the addition

$$\frac{1}{2}\frac{\partial^2 h(z,t)}{\partial z^2}diag[b(t),\beta(t)]diag[b(t),\beta(t)]^T$$

appears in view of the second derivative in z in the Ito formula. 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 m_0 and variance P_0, of a random vector $m_0=[z_{10},z_0]$ are available, the Gaussian distribution with the same parameters, $N(m_0,P_0)$, is the best approximation for the unknown conditional distribution of $m_0=[z_{10},z_0]$ with respect to observations. This fact is also a corollary of the central limit theorem [27] in the probability theory.

Upon calculating the partial derivatives of $h(z,t)$, the equation (6) takes the form

$$dz_1(t)=(A_1(t)c_0(t)+A_1(t)\begin{bmatrix} a_1(t)x(t)\\ 0\end{bmatrix}+$$
$$A_2(t)[x(t),\theta(t)]^Tc_0(t)+A_2(t)([x(t),\theta(t)]^Tc_0(t))^T+$$
$$A_2(t)[x(t),\theta(t)]^T[x(t),\theta(t)]diag[I_n,0_p]a_1^T(t)+$$
$$A_2(t)([x(t),\theta(t)]^T[x(t),\theta(t)]diag[I_n,0_p]a_1^T(t))^T+\dot{C}_0(t)+$$
$$\dot{A}_1(t)[x(t),\theta(t)]^T+\dot{A}_2(t)[x(t),\theta(t)]^T[x(t),\theta(t)])dt+$$
$$A_1(t)diag[b(t),\beta(t)][dW_1^T,dW_3^T]+A_2(t)[x(t),\theta(t)]^T\times$$
$$[dW_1^T,dW_3^T]diag[b(t),\beta(t)]^T+A_2(t)([x(t),\theta(t)]^T\times$$
$$[dW_1^T,dW_3^T]diag[b^T(t),\beta^T(t)])^T, \qquad (7)$$

with the initial condition $z_1(0)=z_{10}$. The equation (5) can be written in the form

$$dy(t)=z_1(t)dt+B(t)dW_2(t). \qquad (8)$$

Thus, the estimation problem is now reformulated as to find the optimal estimate $[m_1(t),m_2(t),m_3(t)]$ for the state vector $[x(t),\theta(t),z_1(t)]$ governed by the linear and bilinear equations (1), (3), and (7) based on the observation process $Y(t)=\{y(s),0\le s\le t\}$, satisfying the equation (8). The solution of this problem is obtained using the optimal filtering equations for bilinear states over linear observations [23] and given by

$$dm_1(t)=(a_0(t)+a(t)m_1(t))dt+P_{13}(t)(B(t)B^T(t))^{-1}\times$$
$$(dy(t)-m_3(t)dt),$$
$$dm_2(t)=P_{23}(t)(B(t)B^T(t))^{-1}(dy(t)-m_3(t)dt),$$
$$dm_3(t)=(A_1(t)c_0(t)+A_1(t)\begin{bmatrix} a_1(t)m_1(t)\\ 0\end{bmatrix}$$
$$A_2(t)\begin{bmatrix} m_1(t)\\ m_2(t)\end{bmatrix}c_0(t)+(A_2(t)\begin{bmatrix} m_1(t)\\ m_2(t)\end{bmatrix}c_0(t))^T+$$
$$A_2(t)\begin{bmatrix} m_1(t)m_1^T(t) & 0\\ m_2(t)m_1^T(t) & 0\end{bmatrix}a_1(t)^T+$$
$$(A_2(t)\begin{bmatrix} m_1(t)m_1^T(t) & m_1(t)m_2^T(t)\\ 0 & 0\end{bmatrix}a_1(t)^T)^T+$$
$$A_2(t)\begin{bmatrix} P_{11} & 0\\ P_{12} & 0\end{bmatrix}a_1(t)+A_2(t)(\begin{bmatrix} P_{11} & P_{21}\\ 0 & 0\end{bmatrix}a_1^T(t))^T+$$
$$\dot{C}_0(t)+\dot{A}_1(t)\begin{bmatrix} m_1(t)\\ m_2(t)\end{bmatrix}+\dot{A}_2(t)\begin{bmatrix} m_1(t)\\ m_2(t)\end{bmatrix}\begin{bmatrix} m_1(t)\\ m_2(t)\end{bmatrix}^T)dt$$
$$+P_{33}(t)(B(t)B^T(t))^{-1}(dy(t)-m_3(t)dt), \qquad (9)$$

with the initial conditions

$$m_1(t_0)=[E(x(t_0)\mid F_{t_0}^Y)],\quad m_2(t_0)=E(\theta(t_0)\mid F_{t_0}^Y)],$$
$$m_3(t_0)=E(z_1(t_0)\mid F_{t_0}^Y)],$$

and

$$dP_{11}(t)=(a(t)P_{11}(t)+P_{11}(t)a^T(t)+b(t)b^T(t)-$$
$$P_{13}(t)(B(t)B^T(t))^{-1}P_{31}(t))dt,$$
$$dP_{12}(t)=(a(t)P_{12}(t)-P_{13}(t)(B(t)B^T(t))^{-1}P_{32}(t))dt,$$
$$dP_{13}(t)=(a(t)P_{13}(t)+$$
$$\left(A_1(t)\begin{bmatrix} P_{11}(t) & P_{12}(t)\\ 0 & 0\end{bmatrix}\begin{bmatrix} a(t)\\ 0\end{bmatrix}\right)^T+$$
$$(\dot{A}_1(t)[P_{11}(t)+P_{21}(t)])^T+$$

$$2(A_2(t)\begin{bmatrix} m_1(t) \\ m_2(t) \end{bmatrix} P_{21}(t)diag[I_n,0_p]a^T(t))^T+$$

$$2(A_2(t)(\begin{bmatrix} m_1(t) \\ m_2(t) \end{bmatrix} P_{21}(t)diag[I_n,0_p]a^T(t))^T)^T+$$

$$2(\dot{A}_2(t)\begin{bmatrix} m_1(t) \\ m_2(t) \end{bmatrix} P_{21}(t))^T+$$

$$2(\dot{A}_2(t)(\begin{bmatrix} m_1(t) \\ m_2(t) \end{bmatrix} P_{21}(t))^T)^T+(A_1(t)\begin{bmatrix} b(t)b^T(t) \\ 0 \end{bmatrix})^T+$$

$$\begin{bmatrix} b(t) & 0 \end{bmatrix}\left(A_2(t)\begin{bmatrix} m_1(t) \\ m_2(t) \end{bmatrix} diag[b^T(t),\beta^T(t)]\right)^T-$$

$$P_{13}(t)(B(t)B^T(t))^{-1}P_{33}(t))dt,$$

$$dP_{22}(t)=(\beta(t)-P_{23}(t)(B(t)B^T(t))^{-1}P_{32}(t))dt,$$

$$dP_{23}(t)=(P_{21}(t)A_1(t)\begin{bmatrix} a(t) \\ 0 \end{bmatrix}^T+$$

$$(\dot{A}_1(t)[P_{12}(t)+P_{12}(t)])^T+$$

$$2(A_2(t)\begin{bmatrix} m_1(t) \\ m_2(t) \end{bmatrix} P_{22}(t)diag[I_n,0_p]a^T(t))^T+$$

$$2(A_2(t)(\begin{bmatrix} m_1(t) \\ m_2(t) \end{bmatrix} P_{22}(t)diag[I_n,0_p]a^T(t))^T)^T+$$

$$\left(A_1(t)\begin{bmatrix} 0 \\ \beta(t) \end{bmatrix}\right)^T+2(\dot{A}_2(t)\begin{bmatrix} m_1(t) \\ m_2(t) \end{bmatrix} P_{22}(t))^T+$$

$$2(\dot{A}_2(t)(\begin{bmatrix} m_1(t) \\ m_2(t) \end{bmatrix} P_{22}(t))^T+$$

$$+\begin{bmatrix} 0 & \beta(t) \end{bmatrix}\left(A_2(t)\begin{bmatrix} m_1(t) \\ m_2(t) \end{bmatrix} diag[b^T(t),\beta^T(t)]\right)^T$$

$$-P_{23}(t)(B(t)B^T(t))^{-1}P_{33}(t))dt,$$

$$dP_{33}(t)=(A_1(t)\begin{bmatrix} a(t) \\ 0 \end{bmatrix} P_{13}(t)+P_{31}(t)(A_1(t)\begin{bmatrix} a(t) \\ 0 \end{bmatrix})^T+$$

$$2(\dot{A}_1(t)[P_{13}(t)+P_{23}(t)])^T+$$

$$2A_2(t)\begin{bmatrix} m_1(t) \\ m_2(t) \end{bmatrix} P_{23}(t)diag[I_n,0_p]a^T(t)+$$

$$2(A_2(t)\begin{bmatrix} m_1(t) \\ m_2(t) \end{bmatrix} P_{23}(t)diag[I_n,0_p]a^T(t))^T+$$

$$2A_2(t)(\begin{bmatrix} m_1(t) \\ m_2(t) \end{bmatrix} P_{23}(t)diag[I_n,0_p]a^T(t))^T+$$

$$2((A_2(t)\begin{bmatrix} m_1(t) \\ m_2(t) \end{bmatrix} P_{23}(t)diag[I_n,0_p]a^T(t))^T)^T+$$

$$2\dot{A}_2(t)\begin{bmatrix} m_1(t) \\ m_2(t) \end{bmatrix} P_{23}(t)+2(\dot{A}_2(t)\begin{bmatrix} m_1(t) \\ m_2(t) \end{bmatrix} P_{23}(t))^T+$$

$$+2\dot{A}_2(t)(\begin{bmatrix} m_1(t) \\ m_2(t) \end{bmatrix} P_{23}(t))^T+2((\dot{A}_2(t)\begin{bmatrix} m_1(t) \\ m_2(t) \end{bmatrix} P_{23}(t))^T)^T$$

$$+A_1(t)diag[b(t)b^T(t),\beta(t)\beta^T(t)]A_1^T(t)+$$

$$A_1(t)diag[b(t),\beta(t)](A_2(t)\begin{bmatrix} m_1(t) \\ m_2(t) \end{bmatrix} diag[b^T(t),\beta^T(t)])^T+$$

$$(A_2(t)\begin{bmatrix} m_1(t) \\ m_2(t) \end{bmatrix} diag[b^T(t),\beta^T(t)])(A_1(t)diag[b(t),\beta(t)])^T$$

$$+(A_2(t)diag[b^T(t),\beta^T(t)])\begin{bmatrix} P_{11}(t) & 0 \\ P_{12}(t) & 0 \end{bmatrix}\times$$

$$(A_2^T(t)diag[b(t),\beta(t)])^T+$$

$$A_2(t)diag[b^T(t),\beta^T(t)]\begin{bmatrix} m_1(t)m_1^T(t) & 0 \\ m_2(t)m_1^T(t) & 0 \end{bmatrix}\times$$

$$(A_2^T(t)diag[b(t),\beta(t)])^T-P_{33}(t)(B(t)B^T(t))^{-1}P_{33}(t))dt, \tag{10}$$

with the initial condition

$$P(t_0)=E(([x(t_0),\theta(t_0),z_1(t_0)]-[m_1(t_0),m_2(t_0),m_3(t_0)])\times$$
$$([x(t_0),\theta(t_0),z_1(t_0)]-[m_1(t_0),m_2(t_0),m_3(t_0)])^T \mid F_{t_0}^Y).$$

Theorem 1. The optimal finite-dimensional filter for the extended state vector $[x(t),\theta(t),z_1(t)]$, governed by the equations (1),(3), and (7), over the linear observations (8) is given by the equations (9) for the optimal estimate $m(t)=[m_1(t),m_2(t),m_3(t)]=E([x(t),\theta(t),z_1(t)] \mid F_t^Y)$ and the equations (10) for the estimation error variance $P(t)=E(([x(t),\theta(t),z_1(t)]-[m_1(t),m_2(t),m_3(t)])\times([x(t),\theta(t),z_1(t)]-[m_1(t),m_2(t),m_3(t)])^T \mid F_t^Y)$. This filter, applied to the subvector $\theta(t)$, also serves as the optimal identifier for the vector of unknown parameters $\theta(t)$ in the equation (2), yielding the estimate subvector $\hat{\theta}(t)$ as the optimal parameter estimate.

Proof. The proof directly follows from the steps 1-3 for designing the coefficients in the equation (5), the new extended state equations (4),(6), and the optimal filtering equations (9),(10) for bilinear states over linear observations, which were obtained in [23].

V. Example

This section presents an example of designing the optimal filter for a linear state over linear observations with a multiplicative unknown parameter, where a conditionally Gaussian state initial condition for the extended state vector is additionally assumed.

Let the state $x(t)$ satisfy the linear equation

$$\dot{x}(t)=1+x(t)+\psi_1(t) \quad x(0)=x_0, \tag{11}$$

and the observation process with unknown multiplicative parameter be given by the equation

$$y(t)=\theta(t)x(t)+\psi_2(t), \tag{12}$$

where $\psi_1(t)$ and $\psi_2(t)$ are white Gaussian noises, which are the weak mean square derivatives of standard Wiener processes (see [25]) independent of each other and of a Gaussian random variable x_0 serving as the initial condition in (11). The parameter $\theta(t)$ is modelled as a standard Wiener process, i.e., satisfies the equation

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

which can also be rewritten as

$$\dot{\theta}(t)=\psi_3(t), \quad \theta(0)=\theta_0, \tag{13}$$

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 optimal estimate $m_z(t) = [m_1(t), m_2(t)]$ for the linear states (11),(13), $z(t) = [x(t), \theta(t)]$, using linear observations (12) with an unknown multiplicative parameter $\theta(t)$ confused with independent and identically distributed disturbances modeled as white Gaussian noises.

Let us reformulate the problem, introducing the stochastic process $z_1(t) = h(x,t) = \theta(t)x(t)$. Using the Ito formula (see [25]) for the stochastic differential of the function $\theta(t)x(t)$, where $x(t)$ satisfies the equation (11) and $\theta(t)$ satisfies the equation (13), the following equation is obtained for $z_1(t)$

$$dz_1(t) = \theta(t)(1+x(t))dt + \theta(t)dw_1(t) + x(t)dw_3(t),$$
$$z_1(0) = z_{10}. \qquad (14)$$

The initial condition $z_{10} \in R$ is considered a conditionally Gaussian random vector with respect to observations (see the paragraph following (6) for details). 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 (12) takes the form

$$dy(t) = z_1(t)dt + dw_2(t). \qquad (15)$$

The obtained filtering system includes three equations, (11), (13) and (14), for the partially measured state $[z(t), z_1(t)]$ and an equation (15) for the observations $y(t)$, where $z_1(t)$ is a completely measured bilinear state with multiplicative noise, $z(t) = [x(t), \theta(t)]$ is an unmeasured linear state, and $y(t)$ is a linear observation process directly measuring the state $z_1(t)$. The filtering equations (9),(10) take the following particular form for the system (11),(13),(14), and (15)

$$dm_1(t) = (1+m_1(t))dt + P_{13}(t)[dy(t) - m_3(t)dt], \qquad (16)$$
$$dm_2(t) = P_{23}(t)[dy(t) - m_3(t)dt],$$
$$dm_3(t) = (m_2(t)+m_3(t))dt + P_{33}(t)[dy(t) - m_3(t)dt],$$

with the initial conditions $m_1(0) = E(x_0 \mid y(0)) = m_{10}$, $m_2(0) = E(\theta_0 \mid y(0)) = m_{20}$, and $m_3(0) = E(\theta_0 x_0 \mid y(0)) = m_{30}$, and

$$\dot{P}_{11}(t) = 2P_{11}(t) + 1 - P_{13}^2(t), \qquad (17)$$
$$\dot{P}_{12}(t) = P_{12}(t) - P_{13}(t)P_{23}(t),$$
$$\dot{P}_{13}(t) = P_{12}(t) + 2P_{13}(t) + m_2(t) - P_{13}(t)P_{33}(t),$$
$$\dot{P}_{22}(t) = 1 - P_{23}^2(t),$$
$$\dot{P}_{23}(t) = P_{22}(t) + P_{23}(t) + m_1(t) - P_{23}(t)P_{33}(t),$$
$$\dot{P}_{33}(t) = 2P_{23}(t) + 2P_{33}(t) + P_{11}(t) + P_{22}(t) +$$
$$m_1^2(t) + m_2^2(t) - P_{33}^2(t).$$

with the initial condition

$$P(0) = E(([x(0), \theta(0), z_1(0)] - [m_1(0), m_2(0), m_3(0)]) \times$$
$$([x(0), \theta(0), z_1(0)] - [m_1(0), m_2(0), m_3(0)])^T \mid y(0)).$$

Here, $m_1(t)$ is the optimal estimate for the state $x(t)$, $m_2(t)$ is the optimal estimate for the state $\theta(t)$, and $m_3(t)$ is the optimal estimate for the state $z_1(t) = \theta(t)x(t)$.

Numerical simulation results are obtained solving the systems of filtering equations (16)–(17). For the filter (16)–(17) and the reference system (11),(13),(14), (15), involved in simulation, the following initial values are assigned: $x(0) = 1$, $m_1(0) = 100$, $m_2(0) = 0$, $m_3(0) = 0$, $P_{11}(0) = 100$, $P_{12}(0) = 0$, $P_{13}(0) = 10$, $P_{22}(0) = 100$, $P_{23}(0) = 10$, $P_{33}(0) = 100$, the unknown parameter θ is assigned as $\theta = -10$ in the first simulation and as $\theta = 10$ in the second one, thus considering negative and positive parameter values. Gaussian disturbances $dw_1(t)$, $dw_2(t)$ and $dw_3(t)$ are realized using the built-in MatLab white noise functions. The simulation interval is $[0, 3.5]$.

Figure 1 shows the graphs of the reference state variable $x(t)$ (11) and its optimal estimate $m_1(t)$ (16), the state $z_1(t) = \theta(t)x(t)$ and its optimal estimate $m_3(t)$, as well as the optimal parameter estimate $m_2(t)$ (θ=-10) for the negative parameter value, in the entire simulation interval [0,3.5]. For better visualization, Figure 2 shows the graph of the state $z_1(t) = \theta(t)x(t)$ and its optimal estimate $m_3(t)$ for the negative parameter value in detail in the interval [0.0.5]. Figure 3 shows the graphs of the reference state variable $x(t)$ (11) and its optimal estimate $m_1(t)$ (16), the state $z_1(t) = \theta(t)x(t)$ and its optimal estimate $m_3(t)$, as well as the optimal parameter estimate $m_2(t)$ ($\theta = 10$) for the positive parameter value, in the entire simulation interval [0,3.5]. For better visualization, Figure 4 shows the graph of the state $z_1(t) = \theta(t)x(t)$ and its optimal estimate $m_3(t)$ for the positive parameter value in detail in the interval [0,0.5].

It can be observed that the optimal estimate $m_1(t)$ converges to the real state $x(t)$ very rapidly, in spite of a considerable difference in the initial conditions, $m_1(0) - x(0) = 99$, the optimal estimate $m_3(t)$ converges and then remains very close to the state $z_1(t)$, and the optimal parameter estimate $m_2(t)$ definitely converges to the real parameter value θ in both positive and negative cases.

VI. Conclusions

This paper presents the optimal mean-square solution to the simultaneous state filtering and parameter identification problem for linear stochastic systems over linear observations with unknown parameters, where unknown parameters are considered Wiener processes. The optimal state filter and parameter identifier is designed in the form of a closed finite-dimensional system of stochastic ODEs. This result is theoretically proved based on the previously obtained optimal filter for bilinear system states over linear observations and numerically verified. The simulation results show very reliable behavior of the designed filter and parameter identifier, which works equally well for both positive and negative parameter values.

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 C.R. *Linear Statistical Inference and its Applications*, New York: Wiley-Interscience, 1973.

[3] Elliott L.J., 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 L.J., Krishnamurthy V. New finite–dimensional filters for estimation of continuous–time linear gaussian systems, *SIAM. J. Contr. Optim.* 1997; **35**: 1908-1923.

[5] Duncan T. E., 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 C. D., 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 W. X. 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] 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.

[11] Shin DR, Verriest E. Optimal access control of simple integrated networks with incomplete observations, *Proc. American Control Conf.* 1994; Baltimore, MD, USA, 3487-3488.

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

[13] Wang Z. Filtering on nonlinear time-delay stochastic systems, *Automatica* 2003, **39**: 101-109.

[14] 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.

[15] 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.

[16] Sheng J, Chen T, Shah SL. Optimal filtering for multirate systems. *IEEE Transactions on Circuits and Systems* 2005; **52**: 228-232.

[17] Sheng J. Optimal filtering for multirate systems based on lifted models. *Proc. American Control Conf.* 2005; Portland, OR, USA, 3459-3461.

[18] 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.

[19] 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.

[20] 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.

[21] Wang Z, Liu Y, Liu Y, H-infinity filtering for uncertain stochastic time-delay systems with sector-bounded nonlinearities, *Automatica*; **44**: 1268-1277.

[22] Åström KJ. *Introduction to Stochastic Control Theory*. Academic Press: New York, 1970.

[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 state filtering and parameter identification for linear systems, *Optimal Control Applications and Methods* 2008; **29**: 159–166.

[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.

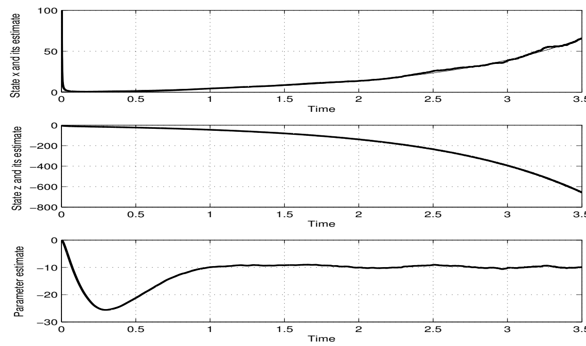

Fig. 1. Negative parameter value. **Above.** Graphs of the real state $x_1(t)$ (thin line) and its optimal estimate $m_1(t)$ (thick line) in the interval $[0,3.5]$. **Middle.** Graph of the state $z_1(t)$ (thin line) and its optimal estimate $m_3(t)$ (thick line) in the interval $[0,3.5]$. **Below.** Graph of the optimal parameter estimate $m_2(t)$ in the interval $[0,3.5]$.

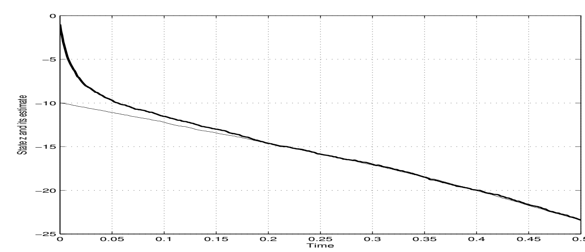

Fig. 2. Negative parameter value. Graph of the state $z_1(t)$ (thin line) and its optimal estimate $m_3(t)$ (thick line) in the interval $[0,0.5]$.

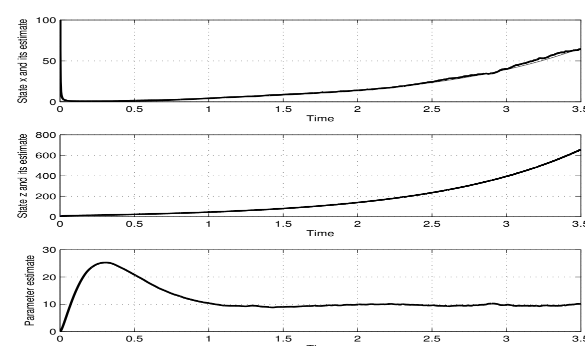

Fig. 3. Positive parameter value. **Above.** Graphs of the real state $x_1(t)$ (thin line) and its optimal estimate $m_1(t)$ (thick line) in the interval $[0,3.5]$. **Middle.** Graph of the state $z_1(t)$ (thin line) and its optimal estimate $m_3(t)$ (thick line) in the interval $[0,3.5]$. **Below.** Graph of the optimal parameter estimate $m_2(t)$ in the interval $[0,3.5]$.

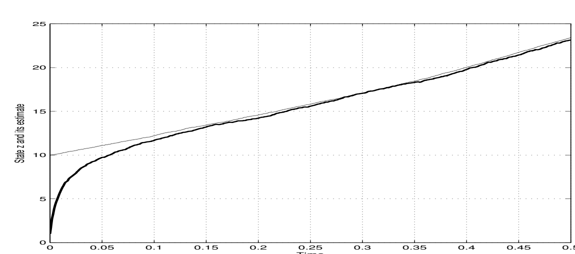

Fig. 4. Positive parameter value. Graph of the state $z_1(t)$ (thin line) and its optimal estimate $m_3(t)$ (thick line) in the interval $[0,0.5]$.