2008 American Control Conference
Westin Seattle Hotel, Seattle, Washington, USA
June 11-13, 2008

Optimal State Filtering and Parameter Identification for Linear Time-Delay Systems

Michael Basin Peng Shi Dario Calderon-Alvarez

Abstract— **This paper presents the optimal joint state filtering and parameter identification problem for linear stochastic time-delay systems with unknown parameters. The original identification problem is reduced to the optimal filtering problem for incompletely measured polynomial (bilinear) time-delay system states over linear observations with an arbitrary, not necessarily invertible, observation matrix, where the unknown parameters are considered standard Wiener processes and incorporated as additional states in the extended state vector. The obtained solution is based on the designed optimal filter for incompletely measured bilinear time-delay states over linear observations, taking into account that the optimal filter for the extended state vector also serves as the optimal identifier for the unknown parameters. In the example, performance of the designed optimal state filter and parameter identifier is verified for a linear time-delay system with an unknown multiplicative parameter over linear observations. Both, stable and unstable, linear systems are examined.**

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.

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.

M. Basin and D. Calderon-Alvarez 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

P. Shi is with Department of Computing and Mathematical Sciences, Faculty of Advanced Technology, University of Glamorgan, Pontypridd, United Kingdom pshi@glam.ac.uk

Another approach, based on the optimization of robust H_∞-filters, has recently been introduced in [8]–[9] for linear stochastic systems with bounded uncertainties in coefficients. The overall comment is that, despite a significant number of excellent works in the area of simultaneous state estimation and parameter identification, the optimal joint state filter and parameter identifier in the form of a closed finite-dimensional system of stochastic ODEs has not yet been obtained even for linear time-delay systems.

This paper presents the optimal joint filtering and parameter identification problem for linear stochastic time-delay systems with unknown parameters over linear observations. The solution starts with reduction of the original identification problem to the optimal filtering problem for incompletely measured bilinear time-delay system states over linear observations with an arbitrary, not necessarily invertible, observation matrix, 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 as additional states in the extended state vector, which should be estimated mean-square optimally in the optimal filtering problem for bilinear time-delay states.

To deal with the new filtering problem for the extended state vector, the paper presents the optimal finite-dimensional filter for incompletely measured bilinear time-delay system states over linear observations with an arbitrary, not necessarily invertible, observation matrix, thus generalizing the results of ([10]–[11]). The optimal filtering problem is treated proceeding from the general expression for the stochastic Ito differential of the optimal estimate and the error variance [12]. Finally, the closed system of the optimal filtering equations with respect to three variables, the optimal estimate, the error variance, and the error covariance between the current-time and delay-shifted values is derived in the explicit form in the particular case of a bilinear state equation. The paper then focuses on the original optimal joint state filtering and parameter identification problem for linear stochastic time-delay systems with unknown parameters over linear observations, whose solution is based on the obtained optimal filter for incompletely measured bilinear time-delay states. The designed optimal filter for the extended state vector also serves as the optimal identifier for the unknown parameters.

II. Optimal State Filtering 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 a nonlinear differential equation with a bilinear polynomial time-delay drift term for the system state

$$dx(t) = f(x,t)dt + b(t)dW_1(t), \tag{1}$$

where $f(x,t) = a_0(t) + a_1(t)x(t) + a_2(t)x(t)x^T(t-h)$, with the initial condition $x(s) = \phi(s)$, $s \in [t_0 - h, t_0]$, h is the state delay value, and a linear differential equation for the observation process

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

Here, $x(t) \in R^n$ is the state vector and $y(t) \in R^m$ is the linear observation vector, $m \leq n$, $a_0(t)$ is an n-dimensional vector, $a_1(t)$ is an $n \times n$ - matrix, $a_2(t)$ is 3D tensor of dimension $n \times n \times n$. 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. The system state $x(t)$ dynamics depends on the delayed state $x(t-h)$, which actually makes the system state space infinite-dimensional (see, for example, [13]). The observation matrix $A(t) \in R^{m \times n}$ is not supposed to be invertible or even square. 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 estimation problem is to find the optimal estimate $\hat{x}(t)$ of the system state $x(t)$, based on the observation process $Y(t) = \{y(s), t_0 \leq s \leq t\}$, 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 [12], 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 matrix function $P(t) = E[(x(t) - m(t))(x(t) - m(t))^T \mid F_t^Y]$ is the estimation error variance.

III. OPTIMAL FILTER DESIGN

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

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

$$dm(t) = a_0(t) + a_1(t)m(t) + a_2(t)[P(t, t-h) + m(t)m^T(t-h)]$$
$$+P(t)A^T(t)(B(t)B^T(t))^{-1}(dy(t) - (A_0(t) + A(t)m(t))dt), \tag{3}$$
$$dP(t) = (2a_2(t)m(t-h)P(t) + (2a_2(t)m(t-h)P(t))^T +$$
$$a_1(t)P(t) + P(t)a_1^T(t) + b(t)b^T(t) -$$
$$P(t)A^T(t)(B(t)B^T(t))^{-1}A(t)P(t))dt, \tag{4}$$
$$dP(t, t-h) = (2a_2(t)m(t-h)P(t, t-h) +$$
$$(2a_2(t)m(t-2h)P(t-h, t))^T + a_1(t)P(t, t-h) +$$
$$P^T(t, t-h)a_1^T(t) + 1/2[b(t)b^T(t-h) + b(t-h)b^T(t)] -$$
$$-(1/2)[P(t)A^T(t)(B(t)B^T(t))(B(t)B^T(t-h))^{-1} \times$$
$$(B(t-h)B^T(t-h))A(t-h)P(t-h)) +$$
$$P(t-h)A^T(t-h)(B(t-h)B^T(t-h)) \times$$
$$(B(t-h)B^T(t))^{-1}(B(t)B^T(t))A(t)P(t)])dt. \tag{5}$$

with the initial conditions $m(s) = E(\phi(s))$, $s \in [t_0 - h, t_0)$, $m(t_0) = E(\phi(t_0) \mid F_{t_0}^Y)$, $P(t_0) = E[(x(t_0) - m(t_0)(x(t_0) - m(t_0)^T \mid F_{t_0}^Y]$, and $P(s, s-h) = E[(x(s) - m(s)(x(s-h) - m(s-h)^T \mid F_s^Y]$ for $s \in [t_0, t_0 + h)$. The system of filtering equations (3)–(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 three lower conditional moments, $m(t)$, $P(t)$, and $P(t, t-h)$.

Proof. The optimal filtering equations could be obtained using the formula for the Ito differential of the conditional expectation $m(t) = E(x(t) \mid F_t^Y)$ (see [12])

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

where $f(x,t) = a_0(t) + a_1(t)x(t) + a_2(t)x(t)x^T(t-h)$ is the bilinear drift term in the state equation, and $\varphi_1(x) = A_0(t) + A(t)x(t)$ is the linear drift term in the observation equation. Upon performing substitution, the estimate equation takes the form

$$dm(t) = E(a_0(t) + a_1(t)x(t) + a_2(t)x(t)x^T(t-h) \mid F_t^Y)dt +$$
$$E(x(t)[A(t)(x(t) - m(t))]^T \mid F_t^Y) \times$$
$$(B(t)B^T(t))^{-1}(dy(t) - (A_0(t) + A(t)m(t)) =$$
$$(a_0(t) + a_1(t)m(t) + a_2(t)[P(t, t-h) + m(t)m^T(t-h)])dt +$$
$$P(t)A^T(t)(B(t)B^T(t))^{-1}(dy(t) - (A_0(t) + A(t)m(t))dt), \tag{6}$$

where $P(t, t-h) = E((x(t) - m(t))(x(t-h) - m(t-h))^T \mid F_t^Y)$ is the covariance of the estimation error values at the current time t and the delay-shifted moment $t-h$. The equation (6) should be complemented with the initial condition $m(s) = E(\phi(s))$, $s \in [t_0 - h, t_0)$, $m(t_0) = E(\phi(t_0) \mid F_{t_0}^Y)$.

Trying to compose a closed system of the filtering equations, the equation (6) should be complemented with the equations for the error variance $P(t)$ and covariance $P(t, t-h)$. For this purpose, the formula for the Ito differential of the variance $P(t) = E((x(t) - m(t))(x(t) - m(t))^T \mid F_t^Y)$ is used (cited again after [12]):

$$dP(t) = (E(((x(t) - m(t))(f(x,t))^T \mid F_t^Y) +$$
$$E(f(x,t)(x(t) - m(t))^T) \mid F_t^Y) +$$
$$b(t)b^T(t) - E(x(t)[\varphi_1(x,t) - E(\varphi_1(x,t) \mid F_t^Y)]^T \mid F_t^Y) \times$$

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

where the last term should be understood as a 3D tensor (under the expectation sign) convoluted with a vector, which yields a matrix. Upon substituting the expressions for $f(x,t)$ and $\varphi_1(x,t)$ and using the variance formula $P(t)=E((x(t)-m(t))x^T(t))\mid F_t^Y)$, the last equation can be represented as

$$dP(t)=(a_2(t)E(((x(t)x^T(t-h))(x(t)-m(t))^T)\mid F_t^Y)+$$
$$(a_2(t)E(((x(t)x^T(t-h))(x(t)-m(t))^T)\mid F_t^Y))^T+$$
$$a_1(t)P(t)+P(t)a_1^T(t)+b(t)b^T(t)-$$
$$P(t)A^T(t)(B(t)B^T(t))^{-1}A(t)P(t))dt+$$
$$E(((x(t)-m(t))(x(t)-m(t))(x(t)-m(t))^T\mid F_t^Y)\times$$
$$A^T(t)(B(t)B^T(t))^{-1}(dy(t)-(A_0(t)+A(t)m(t))dt). \quad (7)$$

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

Applying now the Ito differential formula to the covariance $P(t,t-h)=E((x(t)-m(t))(x(t-h)-m(t-h))^T\mid F_t^Y)$, substituting the expressions for $f(x,t)$ and $\varphi_1(x,t)$, and using the formulas for the variance $P(t)$ and covariance $P(t,t-h)=E((x(t)-m(t))(x(t-h)-m(t-h))^T\mid F_t^Y)$ yields

$$dP(t,t-h)=(a_2(t)E(((x(t)x^T(t-h))\times$$
$$(x(t-h)-m(t-h))^T)\mid F_t^Y)+$$
$$(a_2(t)E(((x(t-h)x^T(t-2h))(x(t)-m(t))^T)\mid F_t^Y))^T+$$
$$a_1(t)P(t,t-h)+P^T(t,t-h)a_1^T(t)+$$
$$1/2[b(t)b^T(t-h)+b(t-h)b^T(t)]-$$
$$-(1/2)[P(t)A^T(t)(B(t)B^T(t))(B(t)B^T(t-h))^{-1}\times$$
$$(B(t-h)B^T(t-h))A(t-h)P(t-h))$$
$$+P(t-h)A^T(t-h)(B(t-h)B^T(t-h))\times$$
$$(B(t-h)B^T(t))^{-1}(B(t)B^T(t))A(t)P(t)])dt+$$
$$(1/2)[E((x(t)-m(t))(x(t-h)-m(t-h))\times$$
$$(x(t-h)-m(t-h))^T\mid F_t^Y)A^T(t-h)(B(t-h)B^T(t-h))^{-1}\times$$
$$(dy(t-h)-(A_0(t-h)+A(t-h)m(t-h))dt)+$$
$$E((x(t-h)-m(t-h))(x(t)-m(t))(x(t)-m(t))^T\mid F_t^Y)\times$$
$$A^T(t)(B(t)B^T(t))^{-1}(dy(t)-(A_0(t)+A(t)m(t))dt)]. \quad (8)$$

The equation (8) should be complemented with the initial condition $P(s,s-h)=E[(x(s)-m(s)(x(s-h)-m(s-h)^T\mid F_s^Y]$ for $s\in[t_0,t_0+h)$.

The equations (6)–(8) for the optimal estimate $m(t)$, the error variance $P(t)$, and the error covariance $P(t,t-h)$ form a non-closed system of the filtering equations for the nonlinear state (1) over linear observations (2). The non-closeness means that the system (6)–(8) includes terms depending on x, such as $E(((x(t)x^T(t-h))(x(t)-m(t))^T)\mid F_t^Y)$ and $E(((x(t)x^T(t-h))(x(t-h)-m(t-h))^T)\mid F_t^Y)$, which are not expressed yet as functions of the system variables, $m(t)$, $P(t)$, and $P(t,t-h)$.

As shown in [11], a closed system of the filtering equations for a polynomial system state (1), without time delays, over linear observations can be obtained. Using the same technique, the optimal filtering equations (3)–(5) are finally derived. The details are omitted here due to space shortage. ■

Thus, a closed form of the filtering equations is obtained for a bilinear time-delay function $f(x,t)=a_0(t)+a_1(t)x(t)+a_2(t)x(t)x^T(t-h)$ in the equation (1).

IV. Joint State Filtering and Parameter Identification Problem

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 delay-differential equation with unknown vector parameter $\theta(t)$ for the system state

$$dx(t)=(a_0(\theta,t)+a(\theta,t)x(t-h))dt+b(t)dW_1(t), \quad (17)$$

with the initial condition $x(s)=\phi(s)$, $s\in[t_0-h,t_0]$, h is the state delay value, and a linear differential equation for the observation process

$$dy(t)=(A_0(t)+A(t)x(t))dt+B(t)dW_2(t). \quad (18)$$

Here, $x(t)\in R^n$ is the state vector, $y(t)\in R^m$ is the linear observation vector, $m\leq n$, and $\theta(t)\in R^p$, $p\leq n\times n+n$, is the vector of unknown entries of 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 $a_{ij}(t)=\theta_k(t)$, $k=p_1+1,\ldots,p\leq n\times n+n$, 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. The system state $x(t)$ dynamics depends on the delayed state $x(t-h)$, which actually makes the system state space infinite-dimensional (see, for example, [13]). The observation matrix $A(t)\in R^{m\times n}$ is not supposed to be invertible or even square. It is assumed that $B(t)B^T(t)$ is a positive definite matrix. All coefficients in (17)–(18) are deterministic functions of time of appropriate dimensions.

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 as F_t-measurable Wiener processes

$$d\theta(t)=dW_3(t), \quad (19)$$

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 $W_2(t)$.

The estimation problem is to find the optimal estimate $\hat{z}(t)=[\hat{x}(t),\hat{\theta}(t)]$ of the combined vector of the system states and unknown parameters $z(t) = [x(t), \theta(t)]$, based on the observation process $Y(t) = \{y(s), 0 \le s \le t\}$, that minimizes the Euclidean 2-norm $J = E[(z(t) - \hat{z}(t))^T (z(t) - \hat{z}(t)) \mid F_t^Y]$ at every time moment t. Here, $E[\xi(t) \mid F_t^Y]$ means the conditional expectation of a stochastic process $\xi(t) = (z(t) - \hat{z}(t))^T (z(t) - \hat{z}(t))$ with respect to the σ - algebra F_t^Y generated by the observation process $Y(t)$ in the interval $[t_0, t]$. As known [12], this optimal estimate is given by the conditional expectation $\hat{z}(t) = m(t) = E(z(t) \mid F_t^Y)$ 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 $[t_0, t]$. As usual, the matrix function $P(t) = E[(z(t) - m(t))(z(t) - m(t))^T \mid F_t^Y]$ is the estimation error variance.

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

Theorem 2. *The optimal finite-dimensional filter for the extended state vector $z(t) = [x(t), \theta(t)]$, governed by the equations (17),(19) over the linear observations (18) is given by the following equations for the optimal estimate $\hat{z}(t) = m(t) = [\hat{x}(t), \hat{\theta}(t)] = E([x(t), \theta(t)] \mid F_t^Y)$, the estimation error variance $P(t) = E[(z(t) - m(t))(z(t) - m(t))^T \mid F_t^Y]$, and the estimation error covariance $P(t, t-h) = E[(z(t) - m(t))(z(t-h) - m(t-h))^T \mid F_t^Y]$*

$$dm(t) = (c_0(t) + a_1(t)m(t) + a_2(t)[P(t, t-h) + m(t)m^T(t-h)]$$

$$+P(t)[A(t), 0_{m\times p}]^T (B(t)B^T(t))^{-1}[dy(t) - A(t)m(t)dt],$$

$$\hat{x}(s) = E(\phi(s)), s \in [t_0 - h, t_0),$$

$$\hat{x}(t_0) = E(\phi(t_0) \mid F_{t_0}^Y),\ E(\theta(t_0) \mid F_t^Y)], \qquad (20)$$

$$dP(t) = (a_1(t)P(t) + P(t)a_1^T(t) + 2a_2(t)m(t-h)P(t) +$$

$$(2a_2(t)m(t-h)P(t))^T + (diag[b(t), I_p])(diag[b(t), I_p]^T))dt -$$

$$P(t)[A(t), 0_{m\times p}]^T (B(t)B^T(t))^{-1}[A(t), 0_{m\times p}]P(t)dt,$$

$$P(t_0) = E((z(t_0) - m(t_0))(z(t_0) - m(t_0))^T \mid F_t^Y), \qquad (21)$$

$$dP(t, t-h) = (2a_2(t)m(t-h)P(t, t-h) +$$

$$(2a_2(t)m(t-2h)P(t-h, t))^T + a_1(t)P(t, t-h) +$$

$$P^T(t, t-h)a_1^T(t) + (1/2)[(diag[b(t), I_p])(diag[b(t-h), I_p]^T) +$$

$$(diag[b(t-h), I_p])(diag[b(t), I_p]^T)] -$$

$$(1/2)[P(t)[A(t), 0_{m\times p}]^T (B(t)B^T(t))(B(t)B^T(t-h))^{-1}$$

$$(B(t-h)B^T(t-h)[A(t-h), 0_{m\times p}]P(t-h) +$$

$$P(t-h)[A(t-h), 0_{m\times p}]^T (B(t-h)B^T(t-h)) \times$$

$$(B(t-h)B^T(t))^{-1}(B(t)B^T(t))[A(t), 0_{m\times p}]P(t)], \qquad (22)$$

$P(s, s-h) = E[(z(s) - m(s)(z(s-h) - m(s-h)^T \mid F_s^Y]$, for $s \in [t_0, t_0 + h)$, where $0_{m\times p}$ is the $m \times p$ - dimensional zero matrix. 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 (17), yielding the estimate subvector $\hat{\theta}(t)$ as the optimal parameter estimate.

Proof. To solve this optimal filtering problem, the following procedure is proposed for incorporating the unknown parameters as additional states in the extended state vector and writing the extended state vector equation in the polynomial form. For this purpose, a vector $c_0(t) \in R^{(n+p)}$, a matrix $a_1(t) \in R^{(n+p)\times(n+p)}$, and a cubic tensor $a_2(t) \in R^{(n+p)\times(n+p)\times(n+p)}$ are introduced as follows.

The equation for the i-th component of the state vector (17) is given by

$$dx_i(t) = (a_{0_i}(t) + \sum_{j=1}^{n} a_{ij}(t)x_j(t-h))dt + \sum_{j=1}^{n} b_{ij}(t)dW_{1_j}(t),$$

$x_i(t_0) = x_{0_i}$. 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 $(i, n+p_1+k, j)$-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 state equations (17),(19) for the vector $z(t) = [x(t), \theta(t)] \in R^{n+p}$ can be rewritten as

$$dz(t) = (c_0(t) + a_1(t)z(t) + a_2(t)z(t)z^T(t-h))dt +$$

$$diag[b(t), I_{p\times p}]d[W_1^T(t), W_3^T(t)]^T, \qquad (23)$$

where the matrix $a_1(t)$, cubic tensor $a_2(t)$, and vector $c_0(t)$ have already been defined, and $I_{p\times p}$ is the $p \times p$ - dimensional identity matrix. The equation (23) is bilinear with respect to the extended state vector $z(t) = [x(t), \theta(t)]$.

Thus, the estimation problem is now reformulated as to find the optimal estimate $\hat{z}(t) = m(t) = [\hat{x}(t), \hat{\theta}(t)]$ for the state vector $z(t) = [x(t), \theta(t)]$, governed by the bilinear equation (20), based on the observation process $Y(t) = \{y(s), 0 \le s \le t\}$, satisfying the equation (18). The solution of this problem is obtained using the optimal filtering equations (3)–(5) for incompletely measured bilinear time-delayed states over linear observations. Indeed, directly applying the optimal filter (3)–(5) for incompletely measured bilinear time-delayed states over linear observations to the bilinear state $z(t) = [x(t), \theta(t)]$, governed by (23), and incomplete linear observations (18), the filtering equations (20)–(22) are obtained for $m(t) = \hat{z}(t) = m(t) = [\hat{x}(t), \hat{\theta}(t)]$, $P(t) = E[(z(t) - m(t))(z(t) - m(t))^T \mid F_t^Y]$, and $P(t, t-h) = E[(z(t) - m(t))(z(t-h) - m(t-h))^T \mid F_t^Y]$. ■

Thus, based on the optimal filtering equations (3)–(5) for incompletely measured bilinear time-delay states over linear observations, the optimal state filter and parameter identifier is obtained for the linear time-delay system state

(17) with unknown parameters, based on the incomplete linear observations (18). In the next section, performance of the designed optimal state filter and parameter identifier is verified in an illustrative example.

V. Example

This section presents an example of designing the optimal filter and identifier for an incompletely measured linear time-delay system state with an unknown multiplicative parameter, based on linear state measurements.

Let the bi-dimensional system state $x(t)=[x_1(t),x_2(t)]$ satisfy the linear time-delay equations with unknown parameter θ

$$\dot{x}_1(t)=x_2(t), \quad x_1(0)=x_{10}, \tag{24}$$

$$\dot{x_2}(t)=\theta x_2(t-h)+\psi_1(t), \quad x_2(s)=\phi(s),\ s\in[t_0-h,t_0],$$

and the observation process be given by the linear equation

$$y(t)=x_1(t)+\psi_2(t), \tag{25}$$

where $\psi_1(t)$ and $\psi_2(t)$ are white Gaussian noises, which are the weak mean square derivatives of standard Wiener processes (see [12]). The equations (24)–(25) present the conventional form for the equations (17)–(18), which is actually used in practice [14]. The parameter θ 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 written as

$$\dot{\theta}(t)=\psi_3(t), \quad \theta(0)=\theta_0, \tag{26}$$

where $\psi_3(t)$ is a white Gaussian noise.

The filtering problem is to find the optimal estimate $m(t)=[m_1(t),m_2(t),m_3(t)]$ for the bilinear-linear state (23),(25), $[x_1(t),x_2(t),\theta]$, using linear observations (24) confused with independent and identically distributed disturbances modelled as white Gaussian noises. The filtering horizon is set to $T=10$.

The filtering equations (20)–(22) take the following particular form for the system (24)–(26)

$$\dot{m}_1(t)=m_2(t)+P_{11}(t)[y(t)-m_1(t)], \tag{27}$$

$$\dot{m}_2(t)=m_2(t-h)m_3(t)+P_{32}(t,t-h)+P_{12}(t)[y(t)-m_1(t)],$$

$$\dot{m}_3(t)=P_{13}(t)[y(t)-m_1(t)],$$

with the initial conditions $m_1(0)=E(x_{10}\mid y(0))=m_{10}$, $m_2(s)=E(\phi(s))$, $s\in[t_0-h,t_0)$, $m_2(t_0)=E(\phi(t_0)\mid F^Y_{t_0})$, and $m_3(0)=E(\theta_0\mid y(0))=m_{30}$,

$$\dot{P}_{11}(t)=2P_{12}(t)-P_{11}^2(t), \tag{28}$$

$$\dot{P}_{12}(t)=P_{22}(t)+2P_{13}(t)m_2(t-h)-P_{11}(t)P_{12}(t),$$

$$\dot{P}_{13}(t)=P_{23}(t)-P_{11}(t)P_{13}(t),$$

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

$$\dot{P}_{23}(t)=2P_{33}(t)m_2(t-h)-P_{12}(t)P_{13}(t),$$

$$\dot{P}_{33}(t)=1-P_{13}^2(t),$$

with the initial condition $P(0)=E(([x_{10},x_{20},\theta_0]-m(0))\times([x_{10},x_{20},\theta_0]-m(0))^T\mid y(0))=P_0$,

$$\dot{P}_{11}(t,t-h)=[P_{12}(t,t-h)+P_{21}(t,t-h)]-P_{11}(t)P_{11}(t-h), \tag{29}$$

$$\dot{P}_{21}(t,t-h)=P_{22}(t,t-h)+2P_{31}(t,t-h)m_2(t-h)-$$
$$(1/2)[P_{11}(t)P_{21}(t-h)+P_{11}(t-h)P_{12}(t)],$$

$$\dot{P}_{31}(t,t-h)=P_{32}(t,t-h)-$$
$$(1/2)[P_{11}(t)P_{31}(t-h)+P_{11}(t-h)P_{13}(t)],$$

$$\dot{P}_{12}(t,t-h)=P_{22}(t,t-h)+2P_{13}(t,t-h)m_2(t-2h)-$$
$$(1/2)[P_{11}(t-h)P_{12}(t)+P_{11}(t)P_{12}(t-h)],$$

$$\dot{P}_{22}(t,t-h)=1+2P_{32}(t,t-h)m_2(t-h)+$$
$$2P_{23}(t,t-h)m_2(t-2h)-$$
$$(1/2)[P_{12}(t)P_{12}(t-h)+P_{12}(t)P_{21}(t-h)],$$

$$\dot{P}_{32}(t,t-h)=2P_{33}(t,t-h)m_2(t-2h)-$$
$$(1/2)[P_{12}(t-h)P_{13}(t)+P_{12}(t)P_{31}(t-h)],$$

$$\dot{P}_{13}(t,t-h)=P_{23}(t,t-h)-$$
$$(1/2)[P_{11}(t-h)P_{13}(t)+P_{11}(t)P_{13}(t-h)],$$

$$\dot{P}_{23}(t,t-h)=2P_{33}(t,t-h)m_2(t-h)-$$
$$(1/2)[P_{12}(t)P_{13}(t-h)+P_{21}(t-h)P_{13}(t)],$$

$$\dot{P}_{33}(t,t-h)=1-(1/2)[P_{13}(t)P_{13}(t-h)+P_{13}(t)P_{31}(t-h)],$$

with the initial condition $P(s,s-h)=E(([x_1(s),x_2(s),\theta(s)]-m(s))([x_1(s-h),x_2(s-h),\theta(s-h)]-m(s-h))^T\mid F^Y_s)=R(s)$ for $s\in[t_0,t_0+h)$.

Numerical simulation results are obtained solving the system of filtering equations (27)–(29). The obtained values of the estimates $[m_1(t),m_2(t)]$ for $[x_1(t),x_2(t)]$, and $m_3(t)$, estimate for θ, are compared to the real values of the state variable $x(t)=[x_1(t),x_2(t)]$ and parameter θ in (24)–(26).

For the filter (27)–(29) and the reference system (24)–(26) involved in simulation, the following initial values are assigned: $x_{10}=x_{20}=1000$, $m_{10}=0.1$, $m_2(s)=0.1$ for any $s\in[t_0-h,t_0]$, $m_{30}=0$, $P_{110}=P_{220}=P_{330}=100$, $P_{120}=10$, $P_{130}=P_{230}=0$, $R_{11}(s)=R_{22}(s)=R_{33}(s)=100$, $R_{12}(s)=R_{21}=10$, for any $s\in[t_0,t_0+h)$, and the other entries of $R(s)$ are equal to zero for any $s\in[t_0,t_0+h)$. The delay value is set to $h=5$. The unknown parameter θ is assigned as $\theta=0.1$ in the first simulation and as $\theta=-0.1$ in the second one, thus considering the system (24) unstable and stable, respectively. Gaussian disturbances $\psi_1(t)$, $\psi_2(t)$, $\psi_3(t)$ in (24)–(26) are realized using the built-in MatLab white noise function.

The following graphs are obtained: graphs of the estimation errors between the reference state variable $x_1(t)$ and the optimal state estimate $m_1(t)$ and between the reference state variable $x_2(t)$ and the optimal state estimate $m_2(t)$, graph of the optimal parameter estimate $m_3(t)$ in the entire simulation interval $[0,10]$ for the unstable system (24) ($\theta=0.1$); graphs of the estimation errors between the reference state variable

$x_1(t)$ and the optimal state estimate $m_1(t)$ and between the reference state variable $x_2(t)$ and the optimal state estimate $m_2(t)$, graph of the optimal parameter estimate $m_3(t)$ in the entire simulation interval $[0,10]$ for the stable system (24) ($\theta = -0.1$) The graphs of all those variables are shown in the entire simulation interval from $t_0 = 0$ to $T = 10$ at the top three plots in Figs. 1 and 2 for the unstable and stable cases, respectively. The bottom plots in Figs. 1 and 2 show the graphs of the optimal parameter estimate $m_3(t)$, with more visualization details, in the simulation interval $[9.99,10]$ for the unstable and stable cases, respectively.

It can be observed that, in both cases, the state estimates $[m_1(t), m_2(t)]$ converge to the real state $[x_1(t), x_2(t)]$ and the parameter estimate $m_3(t)$ converges to the real value (0.1 or -0.1) of the unknown parameter $\theta(t)$. Thus, it can be concluded that, in both cases, the designed optimal state filter and parameter identifier (20)–(22) yields reliable estimates of the unobserved system state and the unknown parameter value.

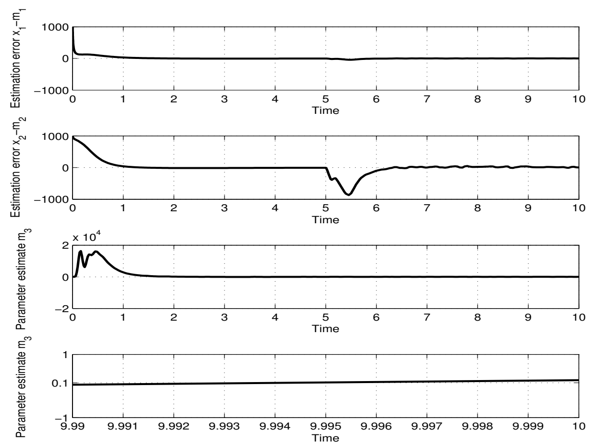

Fig. 1. Graphs of the estimation error between the reference state variable $x_1(t)$ and the optimal state estimate $m_1(t)$ (1st from the top), the estimation error between the reference state variable $x_2(t)$ and the optimal state estimate $m_2(t)$ (2nd from the top), the optimal parameter estimate $m_3(t)$ (3rd from the top) in the entire simulation interval $[0,10]$, and the optimal parameter estimate $m_3(t)$ (1st from the bottom) in the simulation interval $[9.99,10]$ for the unstable system (24).

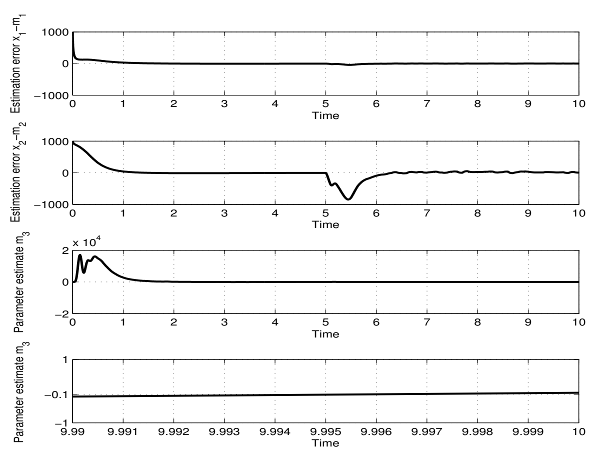

Fig. 2. Graphs of the estimation error between the reference state variable $x_1(t)$ and the optimal state estimate $m_1(t)$ (1st from the top), the estimation error between the reference state variable $x_2(t)$ and the optimal state estimate $m_2(t)$ (2nd from the top), the optimal parameter estimate $m_3(t)$ (3rd from the top) in the entire simulation interval $[0,10]$, and the optimal parameter estimate $m_3(t)$ (1st from the bottom) in the simulation interval $[9.99,10]$ for the stable system (24).

REFERENCES

[1] Y. Bar-Shalom, "Optimal simultaneous state estimation and parameter identification in linear discrete-time systems," *IEEE Transactions on Automatic Control*, vol. 17, pp. 308–319, 1972.

[2] C. R. Rao, *Linear Statistical Inference and its Applications*. New York: Wiley–Interscience, 1973.

[3] L. J. Elliott and V. Krishnamurthy, "New finite-dimensional filters for parameter estimation of discrete-time linear gaussian models," *IEEE Transactions on Automatic Control*, vol. 44, pp. 938–951, 1999.

[4] ——, "New finite dimensional filters for estimation of continuous-time linear gaussian systems," *SIAM Journal on Control and Optimization*, vol. 35, pp. 1908–1923, 1997.

[5] T. E. Duncan, P. Mandl, and B. Pasik-Duncan, "A note on sampling and parameter estimation in linear stochastic systems," *IEEE Trans. Automat. Contr.*, vol. 44, pp. 2120–2125, 1999.

[6] C. D. Charalambous and A. Logothetis, "Maximum likelihood parameter estimation from incomplete data via the sensitivity equations: The continuous-time case," *IEEE Trans. Automat. Contr.*, vol. 45, pp. 928–934, 2000.

[7] W. X. Zheng, "On unbiased parameter estimation of linear systems using noisy measurements," *Cybernetics and Systems*, vol. 34, no. 1, pp. 59–70, 2003.

[8] P. Shi, "Filtering on sampled-data systems with parametric uncertainty," *IEEE Trans. Automat. Contr.*, vol. 43, pp. 1022–1027, 1998.

[9] H. Gao, J. Lam, L. Xie, and C. Wang, "New approach to mixed H_2/H_∞–filtering for polytopic discrete-time systems," *IEEE Transactions on Signal Processing*, vol. 53, pp. 3183–3192, 2005.

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

[11] M. V. Basin and D. A. Calderon-Alvarez, "Optimal filtering for incompletely measured polynomial states over linear observations," *International Journal of Adaptive Control and Signal Processing*, p. DOI: 10.1002/acs.1004, 2008.

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

[13] M. Malek-Zavarei and M. Jashmidi, *Time-Delay Systems: Analysis, Optimization and Applications*. Amsterdam: North-Holland, 1987.

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