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

Optimal Filtering for Itô-Stochastic Continuous-time Systems with Multiple Delayed Measurements

Shulan Kong Mehrdad Saif and Huanshui Zhang

***Abstract*— This paper focuses on the problem of Kalman filtering for Itô stochastic continuous-time systems with multiple delayed measurements, for which very little work exist to date. For an Itô-stochastic system, its stochastic differential and integral have a significant place and are different from other stochastic systems owing to the Wiener or the Brownian process. In this paper, an Itô stochastic continuous-time system with multiple delayed measurements is first reduced to a system with delay free measurements by applying the stochastic analysis and calculus of stochastic variables. Next, the Itô differentials for the optimal filter and its error variance are derived. Finally, through an illustrative example, the performance of the designed optimal filter is verified.**

I. INTRODUCTION

The problem of optimal filter design for both deterministic and stochastic systems has been the subject of many systematic research studies for the last few decades. As a dual problem of optimal control, the optimal filter design problem with no delay in the measurements, has been well studied, see references [1]-[5]. Standard Kalman or Kalman-Bucy filter for systems with perfect model were studies in the aforementioned references. The most famous result is related to the case of linear state and observation equations, which indicate the current moments of the estimate itself and its variance. However, great many practical engineering systems exhibit variety of random phenomenon as well as delays. Yet, a significantly smaller number of authors have studied the problem of optimal filter design for systems with time delays and stochastic noise. Some of the publications related to the robust filter problem for time delay systems are [6]-[9]. Additionally, comprehensive reviews of theory and algorithms for stochastic or time delay systems are given in [5] and [10]-[13].

In recent years there have been two approaches for solving the optimal filtering problem in stochastic systems with time delays. One is to use reorganization innovation analysis approach along with the orthogonal projection lemma. Optimal estimations over observations with multiple delays are treated for discrete or continuous time system in [14] and [15] by using this same idea. The Kalman filter is derived according to the solution of standard Riccati equations. As a result, the method has also been applied successfully to H_∞ fixed-lag smoothing for time delay linear systems in [16] and [17]. The other idea is to use the general expression for the stochastic Ito differential of the optimal estimate and the error variance. The general expression for the stochastic Ito differential of the optimal estimate and the error variance is first given for the optimal filtering problem in [5]. With the application of that result, optimal linear filtering over an observation with multiple delays is investigated and the optimal filtering equations similar to the traditional Kalman-Bucy ones [18], are obtained in [19] and [20]. Finally, the optimal filtering problem for polynomial system states with polynomial multiplicative noise over linear observations with an arbitrary observation matrix is treated in [21]-[22].

This paper presents an optimal filter for linear continuous time varying Itô stochastic system with current and multiple delayed measurements in multiple channels, thus generalizing and developing the results in [15],[19] and [20] based on the standard Kalman filter problem in [1]-[2] and [18]. It is believed that this result has important applications to many fields. The paper applies stochastic analysis and stochastic calculus to solve the problem instead of reorganization innovation analysis approach. The optimal filtering problem is treated proceeding from the general expression for stochastic Itô differential of the optimal estimate and the error variance [5]. At first, multiple delayed measurements are changed into delay free measurements by solving stochastic linear equations, a transformation of the observation equations makes the original problem reduce to a new solvable one. Then the optimal filter and the error variance equations are derived by the formulas for Itô differential of the conditional expectation of system state and the error variance based on the given observations. It is concluded that time delays in observations just cause a new stochastic item for the system. Compared to the conventional Kalman-Bucy filter of the standard system with free delay, there is no intrinsic difference between the optimal filter except the adjustments for delays in the estimate and variance equations. The effect of delays on noise is considered in the adjustments rather than in [19] and [20]. The performance of the designed optimal filter is illustrated and verified by an example with a time delay in the observation equation.

The rest of the paper is organized as follows. In Section 2, the optimal filter problem is formulated for a continuous-time system with multiple channels and measurement delays. A transformation of the observation equation is introduced and a new delay free observation equation is obtained. Then the Ito differentials for the optimal filter and the error

This work was supported by Shandong Province Outstanding Youth Foundation under Grant (2007BS01014) and Nature Science Foundation under Grant (ZR2009GM022), China

Shulan Kong is with School of Mathematics Science, Qufu Normal University, Qufu, 273165, China shulkong@gmail.com

Mehrdad Saif is with School of Engineering Science, Simon Fraser University, Burnaby, BC V5A 1S6, Canada saif@ensc.sfu.ca

Mehrdad Saif is with School of Control Science and Engineering, Shandong University, Jianan, 250061, China hszhang@sdu.edu.cn

variance equations are established by stochastic analysis and stochastic calculus in section 3. The performance of the presented optimal filter is illustrated and verified through an example in Section 4. Finally, some conclusions are drawn in Section 5.

II. PROBLEM FORMULATION

Let $(\Omega, \mathcal{F}, P)$ be a complete probability space with an increasing right continuous family of $\sigma-$algebra $\mathcal{F}_t, t \geq t_0$ and let $(\beta(t), \mathcal{F}_t, t \geq t_0)$ and $(\eta_i(t), \mathcal{F}_t, t \geq t_0)$ be independent Wiener processes. The $\mathcal{F}_t-$measurable random processes $(x(t), y_i(t), i = 0, 1, \cdots, l)$ are determined by a linear continuous-time Itô stochastic differential equation for the system state

$$dx(t) = F(t)x(t)dt + G(t)d\beta(t), x(t_0) = x_0 \tag{1}$$

and linear stochastic differential equations with time delays for the observation processes

$$dy_i(t) = H_i(t)x(t - d_i)dt + d\eta_i(t), d_i \geq 0, i = 0, \cdots, l \tag{2}$$

where random processes, $x(t) \in R^n$ represents the state, $y_0(t) \in R^{m_0}$ and $y_i(t) \in R^{m_i}$ are the current and the delayed measurements respectively. $F(t)$, $G(t)$, and $H_i(t)$ are bounded time-varying matrices with appropriate dimensions. $\beta(t)$ and $\eta_i(t)(i = 0, 1, \cdots, l)$ are independent. Assume that the delays are in an increasing order: $t_0 = 0 = d_0 < d_1 < \cdots < d_l$, the initial condition $x_0 \in R^n$ is a Gaussian vector such that $x_0, \beta(t)$, and $\eta_i(t), i = 0, 1, \cdots, l$ are independent, $E[x_0] = 0, E[x_0 x_0^T] = \Pi_0$, where ^{T} is the transpose of a vector or a matrix. Further, let

$$\begin{aligned} E[d\beta(t)] &= 0, \qquad E[d\beta(t)d\beta(t)^T] = W(t)dt, \\ E[d\eta_i(t)] &= 0, \qquad E[d\eta_i(t)d\eta_i(t)^T] = V_i(t)dt. \end{aligned}$$

Thus the vector random process $[x(t)^T, y_0(t)^T, \cdots, y_l(t)^T]^T$ is determined by the stochastic differential equations as above.

Let $\mathbf{y}(t)$ be the observation of systems (2) at time t, then $\mathbf{y}(t)$ is given by

$$\mathbf{y}(t) = \begin{cases} \mathbf{y}_i = col\{y_0(t), \cdots, y_i(t)\}, & d_i \leq t < d_{i+1}; \\ \mathbf{y}_l = col\{y_0(t), \cdots, y_l(t)\}, & t \geq d_l. \end{cases} \tag{3}$$

The optimal filtering problem is to find the optimal estimate $\hat{x}(t|t)$ of the system state $x(t)$, based on the observation process $Y(t) = \{\mathbf{y}(\tau)|t_0 \leq \tau \leq t\}$, that minimizes

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

at every time moment t. Here $\mathcal{F}_t^Y$ means a $\sigma-$algebra generated by the observation process $Y(t)$ in the interval $[0, t]$ and J means the conditional expectation of a stochastic process $(x(t) - \hat{x}(t))^T(x(t) - \hat{x}(t))$ with respect to the $\sigma-$algebra $\mathcal{F}_t^Y$ generated by the observation process $Y(t)$ in the interval $[t_0, t]$.

III. OPTIMAL FILTERING WITH DELAYED OBSERVATIONS

For systems with delayed observations, to deal with the time delay, a natural idea is to change the observation with time delay into a new delay free observation. At moment t, the relation between $y_i(t)$ and $x(t-d_i)$ is known through (2). Using the state equation (1), we aim to find a relationship between $y_i(t)$ and $x(t)$. Then the optimal estimate is easily obtained by the conditional expectation $\hat{x}(t|t) = E[x(t)|\mathcal{F}_t^Y]$ of the system state $x(t)$ with respect to the $\sigma-$algebra $\mathcal{F}_t^Y$ generated by the observation process $Y(t)$ in the interval $[t_0, t]$.

A. Problem Reduction

From [13], for all $t_0 \geq 0$ and ξ measurable with respect to $\mathcal{F}_t$ and $E|\xi|^2 < \infty$ there exists a unique continuous solution $x(t), t > t_0$, of the system (1), verifying $x(t_0) = \xi$. For $\xi = x_0$, the solution of (1) with $x(t_0) = x_0$ is

$$x(t) = \Phi(t, t_0)x_0 + \Phi(t, t_0) \int_{t_0}^{t} \Phi^{-1}(s, t_0)G(s)d\beta(s) \tag{4}$$

where $\Phi(t, t_0), t \geq t_0$ is the fundamental matrix of solution of the system (1), that is the solution of the matrix equation

$$\frac{d\Phi(t, t_0)}{dt} = F(t)\Phi(t, t_0)$$

and $\Phi(t_0, t_0)$ is an unit matrix.

Then there is a unique continuous solution $x(t-d_i)$ of the system (1), verifying $t - d_i > 0$ and $x(t_0) = x_0$ as follows:

$$\begin{aligned} x(t - d_i) &= \Phi(t - d_i, t_0)x_0 \\ &\quad + \Phi(t - d_i, t_0) \int_{t_0}^{t-d_i} \Phi^{-1}(s, t_0)G(s)d\beta(s). \end{aligned}$$

Then we can obtain the following equation:

$$\begin{aligned} x(t) &= \Phi(t, t - d_i)x(t - d_i) \\ &\quad + \Phi(t, t - d_i) \int_{t-d_i}^{t} \Phi^{-1}(s, t - d_i)G(s)d\beta(s) \end{aligned} \tag{5}$$

the fundamental matrix $\Phi(t, t - d_i)(t \geq t - d_i \geq 0)$ of solution of the system (1) is invertible, $\Phi(t, t - d_i) = exp\left(\int_{t-d_i}^{t} F(s)ds\right)$, then we can obtain

$$x(t - d_i) = \Phi^{-1}(t, t - d_i)x(t) - \int_{t-d_i}^{t} \Phi^{-1}(s, t - d_i)G(s)d\beta(s)$$

and the observation processes (2) can also be rewritten as

$$\begin{aligned} dy_i(t) &= H_i(t)\Phi^{-1}(t, t - d_i)x(t)dt + d\eta_i(t) \\ &\quad - H_i(t) \int_{t-d_i}^{t} \Phi^{-1}(s, t - d_i)G(s)d\beta(s)dt, \\ &\quad i = 0, \cdots, l. \end{aligned} \tag{6}$$

Then introducing the following notations:

$$b_1(t) = \begin{cases} b_{1i} = \begin{bmatrix} H_0(t)\Phi^{-1}(t,t-d_0) \\ H_1(t)\Phi^{-1}(t,t-d_1) \\ \cdots \\ H_i(t)\Phi^{-1}(t,t-d_i) \end{bmatrix}, \\ \quad d_i \le t < d_{i+1}; \\ b_{1l} = \begin{bmatrix} H_0(t)\Phi^{-1}(t,t-d_0) \\ H_1(t)\Phi^{-1}(t,t-d_1) \\ \cdots \\ H_l(t)\Phi^{-1}(t,t-d_l) \end{bmatrix}, \\ \quad t \ge d_l, \end{cases} \tag{7}$$

$$b_0(t) = \begin{cases} b_{0i} = \begin{bmatrix} -H_0(t)\int_{t-d_0}^{t}\Phi^{-1}(s,t-d_0)G(s)d\beta(s) \\ -H_1(t)\int_{t-d_1}^{t}\Phi^{-1}(s,t-d_1)G(s)d\beta(s) \\ \cdots \\ -H_i(t)\int_{t-d_i}^{t}\Phi^{-1}(s,t-d_i)G(s)d\beta(s) \end{bmatrix}, \\ \quad d_i \le t < d_{i+1}; \\ b_{0l} = \begin{bmatrix} -H_0(t)\int_{t-d_0}^{t}\Phi^{-1}(s,t-d_0)G(s)d\beta(s) \\ -H_1(t)\int_{t-d_1}^{t}\Phi^{-1}(s,t-d_1)G(s)d\beta(s) \\ \cdots \\ -H_l(t)\int_{t-d_l}^{t}\Phi^{-1}(s,t-d_l)G(s)d\beta(s) \end{bmatrix}, \\ \quad t \ge d_l, \end{cases} \tag{8}$$

$$dV(t) = \begin{cases} \begin{bmatrix} d\eta_0 \\ d\eta_1 \\ \cdots \\ d\eta_i \end{bmatrix}, d_i \le t < d_{i+1}; \\ \begin{bmatrix} d\eta_0 \\ d\eta_1 \\ \cdots \\ d\eta_l \end{bmatrix}, t \ge d_l \end{cases} \tag{9}$$

and using the notation in (3), then the equations (2) are changed into

$$d\mathbf{y}(t) = (b_1(t)x(t) + b_0(t))\, dt + dV(t) \tag{10}$$

where

$$\begin{aligned} & E[dV(t)dV(t)^T] = Rdt \\ & = \begin{cases} R_i dt = diag\{V_0 dt, \cdots, V_i dt\}, d_i \le t < d_{i+1}; \\ R_l dt = diag\{V_0 dt, \cdots, V_l dt\}, t \ge d_l. \end{cases} \end{aligned} \tag{11}$$

Now we consider the optimal filtering problem for the state equation (1) and the observation system (10) over the observation process $Y(t)$ in the interval $[t_0, t]$ in the following subsection.

B. Calculation of optimal filtering

As known [5], the best approximation of a random variable using the results of measurements is given by the conditional expectation relative to the results of measurements. The optimal estimate $\hat{x}(t|t)$ in section 2 is obtained by the conditional expectation $E[x(t)|\mathcal{F}_t^Y]$ of the system state $x(t)$ with respect to the $\sigma-$algebra $\mathcal{F}_t^Y$ generated by the observation process $Y(t)$ in the interval $[t_0, t]$, that is, $\hat{x}(t|t) = E[x(t)|\mathcal{F}_t^Y]$. And it is based on the formulas for the Itô differential of the conditional expectation and estimation error variance.

Let us begin by denoting the estimation error variance

$$P(t) = E[(x(t) - \hat{x}(t|t))(x(t) - \hat{x}(t|t))^T|\mathcal{F}_t^Y]$$

for the problem. Using the formula for the Itô differential of the conditional expectation $\hat{x}(t|t) = E(x(t)|\mathcal{F}_t^Y)$ in [5] the optimal filtering equations can be obtained

$$\begin{aligned} d\hat{x}(t|t) &= E[dx(t)|\mathcal{F}_t^Y] + E\left[x(t)\ (b_1(t)x(t) \right. \\ &\quad + b_0(t) - E[b_1(t)x(t)|\mathcal{F}_t^Y])^T|\mathcal{F}_t^Y] \\ &\quad \times R^{-1}(d\mathbf{y}(t) - (E[b_1(t)x(t) + b_0(t)|\mathcal{F}_t^Y])dt), \\ \hat{x}(t_0|t_0) &= E[x(t_0)|\mathcal{F}_{t_0}^Y] = 0. \end{aligned} \tag{12}$$

And the formula for the Itô differential of the variance $P(t)$ in [5] can be used:

$$\begin{aligned} dP(t) &= E[(x(t) - \hat{x}(t|t))(F(t)x(t))^T|\mathcal{F}_t^Y]dt \\ &+ E[F(t)x(t)(x(t) - \hat{x}(t|t))^T|\mathcal{F}_t^Y]dt \\ &\quad + G(t)W(t)G(t)^T dt - E[x(t) \\ &\times (b_1(t)x(t) + b_0(t) - E[b_1(t)x(t)|\mathcal{F}_t^Y])^T|\mathcal{F}_t^Y] \\ &\times R^{-1}E[b_1(t)x(t) + b_0(t) \\ &\quad - E[b_1(t)x(t)|\mathcal{F}_t^Y]x(t)^T|\mathcal{F}_t^Y]dt, \\ P(t_0) &= \Pi_0. \end{aligned} \tag{13}$$

Taking into account that

$$\begin{aligned} & E\left[x(t)(x^T(t) - \hat{x}(t|t)^T)|\mathcal{F}_t^Y)\right] \\ = \ & E\left[(x(t) - \hat{x}(t|t))x(t)^T|\mathcal{F}_t^Y)\right] \\ = \ & E\left[(x(t) - \hat{x}(t|t))(x(t) - \hat{x}(t|t))^T|\mathcal{F}_t^Y)\right] = P(t), \\ & E\left[x(t) - \hat{x}(t|t)|\mathcal{F}_t^Y)\right] = 0, \quad E[b_0(t)|\mathcal{F}_t^Y] = 0, \end{aligned}$$

we transform formulae (12), (13) into the forms

$$\begin{aligned} d\hat{x}(t|t) &= F(t)\hat{x}(t|t)dt \\ &\quad + (P(t)b_1(t)^T + E[x(t)b_0(t)^T|\mathcal{F}_t^Y]) \\ &\quad \times R^{-1}(d\mathbf{y}(t) - b_1(t)\hat{x}(t|t)dt), \\ \hat{\mathbf{x}}(t_0|t_0) &= E[\mathbf{x}(t_0)|\mathcal{F}_{t_0}^Y] = 0, \end{aligned} \tag{14}$$

$$\begin{aligned} \dot{P}(t) &= P(t)F(t)^T + F(t)P(t) + G(t)W(t)G(t)^T \\ &\quad - (P(t)b_1(t)^T + E\left[x(t)b_0(t)^T|\mathcal{F}_t^Y\right]) \\ &\quad \times R^{-1}(P(t)b_1(t)^T + E\left[x(t)b_0(t)^T|\mathcal{F}_t^Y\right])^T, \\ P(t_0) &= \Pi_0. \end{aligned} \tag{15}$$

Then we obtain the optimal filtering $\hat{x}(t|t)$ for the state vector $x(t)$ governed by the equation (1) based on the observation process $Y(t) = \{\mathbf{y}(\tau)|t_0 \le \tau \le t\}$, satisfying the equation (2): when $d_i \le t < d_{i+1}$,

$$\begin{aligned} d\hat{x}(t|t) &= F(t)\hat{x}(t|t)dt + (P(t)b_{1i}^T + E[x(t)b_{0i}(t)^T|\mathcal{F}_t^Y]) \\ &\quad \times R_i^{-1}(d\mathbf{y}_i - b_{1i}\hat{x}(t|t)dt), \\ \hat{x}(t_0|t_0) &= 0; \end{aligned}$$

$$
\begin{aligned}
\dot{P}(t) &= P(t)F(t)^T + F(t)P(t) + G(t)W(t)G(t)^T \\
&\quad -(P(t)b_{1i}^T + E\left[x(t)b_{0i}(t)^T|\mathcal{F}_t^Y\right]) \\
&\quad \times R_i^{-1}(P(t)b_{1i}^T + E\left[x(t)b_{0i}(t)^T|\mathcal{F}_t^Y\right])^T, \\
P(t_0) &= \Pi_0,
\end{aligned}
$$

when $t \geq d_l$,

$$
\begin{aligned}
d\hat{x}(t|t) &= F(t)\hat{x}(t|t)dt + (P(t)b_{1l}^T + E\left[x(t)b_{0l}(t)^T|\mathcal{F}_t^Y\right]) \\
&\quad \times R_l^{-1}(d\mathbf{y}_l - b_{1l}\hat{x}(t|t)dt), \\
\hat{x}(t_0|t_0) &= 0; \\
\dot{P}(t) &= [P(t)F(t)^T + F(t)P(t) + G(t)W(t)G(t)^T \\
&\quad -(P(t)b_{1l}^T + E\left[x(t)b_{0l}(t)^T|\mathcal{F}_t^Y\right]) \\
&\quad \times R_l^{-1}(P(t)b_{1l}^T + E\left[x(t)b_{0l}(t)^T|\mathcal{F}_t^Y\right])^T, \\
P(t_0) &= \Pi_0
\end{aligned}
$$

where b_{1i} and b_{1l}, R_i and R_l are satisfied (7) and (11).

Let

$$
\begin{aligned}
& A(j) \\
= \; & P(t)\Phi^{-1}(t, t-d_j)^T \\
& -\int_{t-d_j}^{t} \Phi^{-1}(s,t)G(s)W(s)G(s)^T\Phi^{-1}(s,t-d_j)^T ds.
\end{aligned}
$$

That is, when $d_i \leq t < d_{i+1}$,

$$
\begin{aligned}
d\hat{x}(t|t) &= F(t)\hat{x}(t|t)dt + \sum_{j=0}^{i} A(j)H_j(t)^T V_j(t)^{-1} \\
&\quad \times \left(dy_j(t) - H_j(t)\Phi^{-1}(t,t-d_j)^T\hat{x}(t|t)dt\right), \\
\hat{x}(t_0|t_0) &= 0; \qquad (16) \\
\dot{P}(t) &= P(t)F(t)^T + F(t)P(t) + G(t)W(t)G(t)^T \\
&\quad -\sum_{j=0}^{i} A(j)H_j(t)^T V_j(t)^{-1} H_j(t)A(j)^T, \\
P(t_0) &= \Pi_0, \qquad (17)
\end{aligned}
$$

when $t \geq d_l$,

$$
\begin{aligned}
d\hat{x}(t|t) &= F(t)\hat{x}(t|t)dt + \sum_{j=0}^{l} A(j)H_j(t)^T V_j(t)^{-1} \\
&\quad \times \left(dy_j(t) - H_j(t)\Phi^{-1}(t,t-d_j)^T\hat{x}(t|t)dt\right), \\
\hat{x}(t_0|t_0) &= 0; \qquad (18) \\
\dot{P}(t) &= P(t)F(t)^T + F(t)P(t) + G(t)W(t)G(t)^T \\
&\quad -\sum_{j=0}^{l} A(j)H_j(t)^T V_j(t)^{-1} H_j(t)A(j)^T, \\
P(t_0) &= \Pi_0. \qquad (19)
\end{aligned}
$$

We can summarize the results obtained so far in the following theorem.

Theorem: The optimal filter for the state with (1), over the linear observations (2), is given by (14) for the optimal estimate $\hat{\mathbf{x}}(t|t) = E[\mathbf{x}(t)|\mathcal{F}_t^Y]$ and (15) for the estimation error variance $P(t) = E[(\mathbf{x}(t) - \hat{\mathbf{x}}(t|t))(\mathbf{x}(t) - \hat{\mathbf{x}}(t|t))^T|\mathcal{F}_t^Y]$. Specially, when $d_i \leq t < d_{i+1}$ and $t \geq d_l$, $\hat{\mathbf{x}}(t|t)$ and $P(t)$ are given by (16)-(17) and (18)-(19).

Proof: The proof directly follows from the above.

Remark: In equations (16)-(19), the adjustments for delays include a item that relates to state noise $d\beta(s)$. The effect of delays on state noise is considered in the paper, however it did not appear in the aforementioned references.

IV. NUMERICAL EXAMPLE

This section presents an example of designing the optimal filter for a linear state over linear observations with time delays. It illustrates the results obtained in previous section.

Consider the system

$$
\begin{aligned}
dx(t) &= Fx(t)dt + Gdw(t) \qquad (20) \\
dy_i(t) &= H_i x(t-d_i)dt + d\eta_i(t), i = 0,1 \qquad (21)
\end{aligned}
$$

with

$$
\begin{aligned}
x(t) &= \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix}, F = \begin{bmatrix} -6 & 1 \\ 0 & -5 \end{bmatrix}, G = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \\
H_0 &= [0 \quad -1], H_1 = [1 \quad 0]
\end{aligned}
$$

where $w(t)$ and $\eta_i(t), i = 0,1$ are white Gaussian noises, which are the weak mean square derivatives of standard Wiener processes independent of each other. $E[dw(t)] = 0$, $E[dw(t)dw(t)^T] = dt$; $E[d\eta_i(t)] = 0$, $E[d\eta_i(t)d\eta_i(t)^T] = dt$, $t_0 = 0, x_0$ is a Gaussian random variable, $d_0 = 0, d_1 = d, d > 0$.

The initial values are assigned: $x_1(0) = 1, x_2(0) = 0.5, \hat{x}_1(0|0) = 0, \hat{x}_2(0|0) = 0$, $p_{11}(0) = 1, p_{12}(0) = 0, p_{22}(0) = 0.25$.

The filter problem is to find the optimal estimate for the state (20) and observation (21).

When $0 \leq t < d$, the optimal filter $\hat{x}(t|t)$ of the system

$$
\begin{aligned}
dx(t) &= Fx(t)dt + Gdw(t) \\
dy(t) &= H_0x(t)dt + dv(t)
\end{aligned}
$$

is the standard Kalman-Bucy filter

$$
\begin{aligned}
d\hat{x}(t|t) &= F\hat{x}(t|t)dt \\
&\quad + P(t){H_0}^T{V_0}^{-1}(t)[dy_0(t) - H_0\hat{x}(t|t)dt], \\
\hat{x}(0|0) &= 0; \\
\dot{P}(t) &= P(t)F^T + FP(t) + GWG^T \\
&\quad -P(t)H_0^T{V_0}^{-1}(t)H_0P(t), \\
P(0) &= \begin{bmatrix} 1 & 0 \\ 0 & 0.25 \end{bmatrix}.
\end{aligned}
$$

That is

$$
\begin{aligned}
d\hat{x}_1(t|t) &= \left[-6\hat{x}_1(t|t) + (1 - p_{12}(t))\hat{x}_2(t|t)\right]dt \\
&\quad -p_{12}(t)dy_0(t), \\
d\hat{x}_2(t|t) &= -(5 + p_{22}(t))\hat{x}_2(t|t)dt - p_{22}(t)dy_0(t), \\
\hat{x}_1(0|0) &= 0, \hat{x}_2(0|0) = 0; \\
\dot{p}_{11}(t) &= -p_{12}(t)^2 - 12p_{11}(t) + 2p_{12} + 1, \\
\dot{p}_{12}(t) &= -p_{12}(t)p_{22}(t) - 11p_{12}(t) + p_{22}(t), \\
\dot{p}_{22}(t) &= -p_{22}(t)^2 - 10p_{22}(t), \\
p_{11}(0) &= 1, p_{12}(0) = 0, p_{22}(0) = 0.25.
\end{aligned}
$$

When $t \geq d$, we can obtain the system as follows:

$$\begin{aligned} dx(t) &= Fx(t)dt + Gdw(t) \\ dy(t) &= (b_1 x(t) + b_0)dt + dv(t) \end{aligned}$$

where

$$\begin{aligned} y(t) &= \begin{bmatrix} y_0(t) \\ y_1(t) \end{bmatrix}, b_1 = \begin{bmatrix} 1 & -1 \\ e^{6d} & -e^{6d} + e^{5d} \end{bmatrix}, \\ b_0 &= \begin{bmatrix} 0 \\ -H_1 \int_{t-d_1}^{t} \Phi^{-1}(s, t-d) G dw(s) \end{bmatrix}, \\ \Phi(t) &= \begin{bmatrix} e^{-6t} & e^{-5t} \\ 0 & e^{-5t} \end{bmatrix}, \Phi^{-1}(t) = \begin{bmatrix} e^{6t} & -e^{6t} \\ 0 & e^{5t} \end{bmatrix}, \end{aligned}$$

the optimal filter $\hat{x}(t|t)$ of the system (20)-(21) is

$$\begin{aligned} d\hat{x}_1(t|t) &= \left[-6\hat{x}_1(t|t) + (1 - p_{12}(t))\hat{x}_2(t|t)\right] dt \\ &\quad -p_{12}(t) dy_0(t) \\ &\quad +[p_{11}e^{6d} - p_{12}(e^{6d} - e^{5d}) - \frac{1}{12}(e^{6d} - e^{-6d})] \\ &\quad \times \left(dy_1(t) - e^{6d}\hat{x}_1(t|t)dt\right), \\ d\hat{x}_2(t|t) &= -(5 + p_{22}(t))\hat{x}_2(t|t)dt - p_{22}(t)dy_0(t) \\ &\quad +[p_{12}e^{6d} - p_{22}(e^{6d} - e^{5d})] \\ &\quad \times \left(dy_1(t) - e^{6d}\hat{x}_1(t|t)dt\right), \\ \hat{x}_1(0|0) &= 0, \hat{x}_2(0|0) = 0; \end{aligned}$$

where

$$\begin{aligned} \dot{p}_{11}(t) &= -p_{12}(t)^2 - 12p_{11}(t) + 2p_{12} + 1 \\ &\quad -[p_{11}e^{6d} - p_{12}(e^{6d} - e^{5d}) - \frac{1}{12}(e^{6d} - e^{-6d})]^2, \\ \dot{p}_{12}(t) &= -p_{12}(t)p_{22}(t) - 11p_{12}(t) + p_{22}(t) \\ &\quad -[p_{11}e^{6d} - p_{12}(e^{6d} - e^{5d}) - \frac{1}{12}(e^{6d} - e^{-6d})] \\ &\quad \times[p_{12}e^{6d} - p_{22}(e^{6d} - e^{5d})], \\ \dot{p}_{22}(t) &= -[p_{12}e^{6d} - p_{22}(e^{6d} - e^{5d})]^2 - p_{22}(t)^2 \\ &\quad -10p_{22}(t), \\ p_{11}(0) &= 1, p_{12}(0) = 0, p_{22}(0) = 0.25. \end{aligned}$$

The example displays the process of solving the filter problem of the stochastic system. It can be observed that the gain of the optimal estimate $\hat{x}(t|t)$ has two parts: one is the same as that of the standard Kalman-Bucy filter, the others is a new one generated by the delays. The optimal filtering can be expressed by differential equation. The error variance is a solution of an ordinary differential equation. So the filter of the stochastic system is figured out easily.

V. CONCLUSIONS AND FUTURE WORK

The paper presents the optimal filtering for the It$\hat{o}$ stochastic continuous-time system with multiple delayed measurements. The optimal filtering is designed by the conditional expectation of the system state over the observation processes. The optimal filter and the error variance equations are derived by the formulas for It$\hat{o}$ differential of the conditional expectation of system state and the error variance based on the given observations. This result is theoretically proved based on the stochastic analysis and Calculate and numerically verified. Compared to the conventional Kalman-Bacy filter of the standard system with free delay in [18], the optimal filter of the system with multiple delayed measurements just adds one item involving a stochastic noise caused by time delays in structure, but there is no intrinsic difference except the adjustments of the delays in the estimate and variance equations. An example is well explained to solve the optimal estimate. Moreover the obtained results could be applied to the controller problem of systems with time delays.

REFERENCES

[1] R. E. Kalmam, "A new approach to linear filtering and prediction problems", *Transactions of the ASME-Journal of Basic Engineering*, vol. 82, pp. 35-45, 1960.

[2] R. E. Kalmam and R. S. Bucy, "New results in linear filtering and prediction theory", *Transactions of the ASME-Journal of Basic Engineering*, vol. 83, pp. 95-108, 1961.

[3] B. D. O. Anderson and J. B. Moore, *Optimal filtering,* Prentice-Hall, Englewood Cliffs, N.J., 1979.

[4] T. Kailath, A.H. Sayed and B. Hassibi, *Linear estimation*. Prentice-Hall, Englewood Cliffs, N.J, 1999.

[5] V. S. Pugachev and I. N. Sinitsyn, *Stochastic systems: theory and applications*. World Scientific Publishing, Singapore, 2001.

[6] H. Kwakernaak, "Optimal filtering in linear systems with time delays", *IEEE Trans. Autom. Control*, vol. 12, pp. 169-173, 1967.

[7] H. Kwong and A. S. Willsky, "Estimation and filter stability of stochastic delay systems", *SIAM J. Control and Optimiz.* , vol. 16, pp. 660-681, 1978.

[8] Y. Theodor, Shaked. "Robust discrete time minimum variance filtering", *IEEE Transactions on Signal Processing*, vol. 44, pp. 181-189, 1996.

[9] M. S. Mahmoud. *Robust control and filtering for time-delay systems*, Marcel Dekker: New York, 2000.

[10] V. B. Kolmanovskii, A. D. Myshkis, *Introduction to the theory and applications of functional differential equations*, Kluwer: New York. 1999

[11] M. Malek-Zavarei, M. Jamshidi, *Time-delay systems: analysis, optimization and applications*, North-Holland: Amsterdam, 1987.

[12] L. Dugard, E. I. Verriest. *Stability and control of time-delay systems*, Springer-Verlag, London, 1998.

[13] V. Dragan, T. Morozan, *Mathematical methods in robust control of linear stochastic systems*, Springer Science+Business Media LLC. 2006.

[14] X. Lu and H. Zhang, "Kalman filtering for multiple time-delay systems", *Automatica*, vol. 41, pp. 1455-1461, 2005.

[15] H. S. Zhang, X. Lu and D. Z. Cheng, "Optimal estimation for continuous-time system with delayed measurements". *IEEE Transactions on Automatic Control*, vol. 51, pp. 823-827, 2006.

[16] H. S. Zhang, L. H. Xie, C. Y. Soh, D. Zhang, "H_∞ fixed-lag smoothing for discrete linear time-varying systems", *Automatica*, vol. 41, pp. 839-846, 2005.

[17] H. Zhang, L. Xie and G. Duan, "H_∞ control of discrete-time systems with multiple input delays", *IEEE Transaction on Automatic Control*, vol. 52, pp. 271-283, 2007.

[18] J. L. Speyer and W. H. Chung, *Stochastic processes, estimation, and control*. SIAM Society for Industrial and Applied Mathematics, Phliadelphia, USA, 2008.

[19] M. V. Basin and R. Martiner Zuniga, "Optimal filtering over observations with multiple delays", it Int. J. Roubst Nonlinear control, vol. 14, pp. 685-696, 2004.

[20] M. V. Basin, J Rodriguez-Gonzalez, "Optimal control for linear systems with multiple time delays in control input", *IEEE Transaction on Automatic Control,* vol. 51, pp. 91-97, 2006.

[21] M. Basin, C. A. Dario, and S. Mikhail, "Optimal filtering for incompletely measured polynomial states over linear observations", *Int. J. Adapt. Control Signal Process*, vol. 22, pp. 482-494, 2008.

[22] M. Basin and C. A. Dario, "Optimal filtering for incompletely measured polynomial systems with multiplicative Noise", *Circuits, Systems, and Signal Processing*, vol. 28, pp. 223-239, 2009.