Filtering linear systems with large time-varying measurement delays

F.Cacace [a,e], F. Conte [a,b], M. d'Angelo [c], A. Germani [d], G. Palombo [e]

[a] *Università Campus Bio-Medico of Rome, Rome, Italy*

[b] *DITEN, Università di Genova, Genova, Italy*

[c] *DIAG, Università di Roma "La Sapienza", Rome, Italy*

[d] *DISIM, Università dell'Aquila, L'Aquila, Italy*

[e] *IASI CNR, Rome, Italy*

Abstract

In this paper we consider the estimation problem for linear stochastic systems affected by multiple known and time-varying delays on all the output signals. Based on a modification of a previous proposal we prove for the first time the result that a filter based on simple eigenvalue assignment of the closed-loop error system may achieve uniform performance, with respect to the delay bound and estimation variance, in presence of both constant and time-varying delays that are differentiable. A new and simple demonstration technique provides non conservative delay bounds for time-varying delays. A cascaded version of the filter can cope with arbitrarily large delays.

Key words: Time-varying delay, filtering, linear systems

1 Introduction

The filtering problem of stochastic linear systems in presence of delayed measurements is important, for example in networked digital control [5], and it has been an active research area for many years starting with the seminal work [11]. The majority of contributions in this area consider the discrete-time case, for which the optimal filtering problem has been addressed mainly through the so-called re-organized innovation analysis approach [12,16], that can be applied in the continuous case [19]. Optimal filters for the continuous-time case, in the form of modifications of the Kalman-Bucy filter (KBF), are described in [10,15] (see also [17,18] for multiplicative noise). These approaches have some important drawbacks, since they require that un-delayed measurements are also available and that the delays are constant. Some approaches consider random delays (see for example [7]), that however have values in a finite set. Time-varying delays have been considered mainly in the H_∞ framework [4,6,13,14,20], where the delay knowledge is generally not required. The resulting filters are designed through LMIs, they are robust and easier to implement. However, in all the mentioned references except [14], that considers constant delays, un-delayed measurements are needed.

Therefore, the problem of filter design in presence of purely delayed measurements and variable delays for continuous-time system is still quite open. This case is challenging because the estimation error obeys a stochastic equation with delays and it is in general difficult to ensure that its solution is stable in the appropriate sense when the system is not asymptotically stable. In the present paper, we improve the design proposed in [1], with the aim of recovering the performance for constant delays also in the case of time-varying and piece-wise differentiable measurement delays. The proposed sub-optimal filter has the following features: (i) the instantaneous value of the delay is assumed to be available but all the measures can be delayed; (ii) the filter has computational cost that is comparable to the original KBF; (iii) the design is simple and constructive

Email addresses: `f.cacace@unicampus.it` (F.Cacace), `fr.conte@unige.it` (F. Conte), `mdangelo@diag.uniroma1.it` (M. d'Angelo), `alfredo.germani@univaq.it` (A. Germani), `giovanni.palombo@iasi.cnr.it` (G. Palombo).

with non conservative and uniform delay bounds for both constant and varying delays; (iv) the filter may compensate multiple and arbitrarily large delays; (v) the performance of the estimation error tends to the optimal one when delay vanishes. The main novelty with respect to [4, 6, 10, 13, 15, 19, 20] is that there is no need of un-delayed measurements and the delay can be time-varying. With respect to [1] the new filter adds features (iii) and (iv), and has less conservative bounds on the delay and on the variance of the estimation error. We describe the problem and recall the predictor-based solution for constant delays in Section 2. The main result is proved in Section 3, and in Section 4 we extend it to the case of large and multiple delays as well as to piece-wise continuous delays.

2 Problem statement and preliminaries

2.1 Notation

$\mathbb{R}_+$ denotes non-negative reals. $\|x\|$ denotes the Euclidean norm for $x \in \mathbb{R}^n$ and $\|M\|$ the operator norm when $M \in \mathbb{R}^{n\times m}$. I_n is the identity matrix in $\mathbb{R}^n$. $\sigma(M)$ is the spectrum of the matrix M, $\mu(M) = \max \Re\{\sigma(M)\}$ is the spectral abscissa and $\mathrm{tr}(M)$ its trace. When $\mu(M) < 0$, M is said to be Hurwitz stable. $M > 0$ denotes a positive definite matrix. $\mathrm{col}_i(M)$ and $\mathrm{row}_i(M)$ are, respectively, the vertical and horizontal composition of matrices M. $\mathbb{E}[\cdot]$ denotes the expectation. On a filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_t, P)$, $L^2(\Omega; \mathbb{R}^n)$ denotes the linear space of square integrable random vectors of $\mathbb{R}^n$ endowed with the norm $\|x\|^2_{L_2} = \mathbb{E}[\|x\|^2]$ and $L^2_{t_0,t}([t_0, t] \times \Omega; \mathbb{R}^n)$ is the linear space of $\mathcal{F}_t$-adapted $\mathbb{R}^n$-valued stochastic processes in $[t_0, t]$,such that $x \in L^2_{t_0,t}([t_0, t] \times \Omega; \mathbb{R}^n)$ if $|x|_{t_0,t} < \infty$ where $|x|^2_{t_0,t} = \int_{t_0}^{t} \|x(\tau)\|^2_{L_2} \, d\tau$.

Definition 1 *A stochastic process* $\{\psi(t)\}_{t\geq 0}$ *is said*

- asymptotically centered *if* $\mathbb{E}[\psi(t)] \to 0$ *as* $t \to \infty$*;*
- exponentially centered with rate λ *if there exist positive constants* λ *and* c *such that* $\|\mathbb{E}[\psi(t)]\| \leq c \exp^{-\lambda t}$*;*
- uniformly second moment bounded *if there exist a positive constant* κ *such that* $\sup_{t\geq 0} \|\psi(t)\|_{L_2} < \kappa$*.*

2.2 Problem statement

We study the estimation problem of Gaussian linear time-invariant systems with variable output delay in the Itô formalism:

$$\mathrm{d}x(t) = (Ax(t) + Bu(t))\,\mathrm{d}t + F\,\mathrm{d}W_t, \qquad (1)$$
$$\mathrm{d}y(t) = Cx(t - \delta_t)\mathrm{d}t + G\,\mathrm{d}V_{t-\delta_t}, \qquad (2)$$

where the state $x(t) \in \mathbb{R}^n$, $\mathbb{E}[x(0)]$ and $\|x(0)\|_{L_2}$ are finite, $W_t \in \mathbb{R}^d$ and $V_t \in \mathbb{R}^\ell$ are standard mutually independent $\mathcal{F}_t$-adapted Wiener processes independent from $x(0)$. We make the standard assumption $R = GG^\top > 0$ [8]. For simplicity, let us also assume that the measurement process exists only when $t - \delta_t > 0$, which is equivalent to let $y_\tau = 0$ and $V_\tau = 0$ for $\tau \leq 0$. In addition, we make the following hypotheses.

Assumption $\mathcal{H}_1$. (a) The couple (A, F) is controllable; (b) the couple (A, C) is observable.

Assumption $\mathcal{H}_2$. The map $\delta : \mathbb{R}_+ \to [0, \bar{\delta}]$ admits a derivative for all $t \geq 0$ and the values of δ_t and $\dot{\delta}_t$ are known at $t \geq 0$.

The standard hypothesis $\mathcal{H}_1$ ensures that the optimal estimator for $\delta_t \equiv 0$ is the KBF [9] with an estimation error exponentially centered and uniformly second moment bounded. $\mathcal{H}_2$ requires that the instantaneous delay is known at the filter, a mild requirement that can be usually be fulfilled in practice, and that the delay function is bounded (with known bound) and differentiable. In particular, $\mathcal{H}_2$ implies that δ is continuous, and this in turn ensures that all the output measurements will eventually be available to the estimator. We shall consider less restrictive requirements in Section 4.5.

Notice that when $\dot{\delta}_t > 1$ the function $t - \delta_t$ moves "backward" in time, and $y(t)$ provides no new information. In this case it is possible to consider the most recent measurement by replacing δ_t with a new delay function with derivative not larger than 1. This is however not needed in the design of our filter. Let $\mathcal{F}^{y_t}$ be the σ-algebra generated by process $\{y(\tau)\}_{\tau\in[0,t]}$. $\{\mathcal{F}^{y_t}\}$ is a family of non-decreasing σ-algebras and the optimal estimate is $\hat{x}(t) = \mathbb{E}[x(t)|\mathcal{F}^{y_t}]$. Notice that $\mathcal{F}^{y_t}$ contains information on $x_{\tau-\delta_\tau}$, $W_{\tau-\delta_\tau}$, and $V_{\tau-\delta_\tau}$, $\tau \in [0, t]$. Our problem is to seek for suitable sub-optimal estimates.

Problem Given system (1)–(2) with the Assumptions $\mathcal{H}_1$ and $\mathcal{H}_2$, compute an estimate $\hat{\xi}(t)$ such that the estimation error $\varepsilon(t) = x(t) - \hat{\xi}(t)$ is exponentially centered and uniformly second moment bounded.

Remark 2 *In certain situations, for example communications delay over a noisy channel that contributes to the measurement disturbance* V_t*, it is possible that the measurements move backward in time but still provide new information. This would be the case of a packet* $y(t)$ *transmitted and received twice due to the communications protocol. We do not consider this case in the paper.*

2.3 Kalman-Bucy optimal predictor for constant delay

When $\delta_t = \delta$ (constant delay), the problem of computing $\hat{x}(t)$ can be solved in two steps: (i) compute $\hat{x}(t - \delta|t - \delta) := \mathbb{E}[x(t - \delta)|\mathcal{F}^{y_t}]$ by means of the ordinary KBF; (ii) compute the optimal prediction $\hat{x}(t)$ as (see for example [8], ch. 7, [10])

$$\hat{x}(t) = e^{A\delta}\hat{x}(t-\delta|t-\delta) + \int_{t-\delta}^{t} e^{A(t-\tau)}Bu(\tau)\mathrm{d}\tau. \quad (3)$$

However, (3) involves distributed terms and is computationally expensive. The problem is not easily overcome by transforming (3) in differential form. To see this, let us consider the asymptotic version of the KBF for $\hat{x}(t-\delta|t-\delta)$,

$$\begin{aligned}\mathrm{d}\hat{x}(t-\delta|t-\delta) =& A\hat{x}(t-\delta|t-\delta)\mathrm{d}t + Bu(t-\delta)\mathrm{d}t \\ &+ \overline{K}(dy(t) - C\hat{x}(t-\delta|t-\delta)\mathrm{d}t), \quad (4)\end{aligned}$$

where $\overline{K} = PC^\top R^{-1}$, with P solution of the Riccati equation

$$0 = AP + PA^\top + FF^\top - PC^\top R^{-1}CP. \quad (5)$$

The differential form of (3) can be obtained by using (4) and by replacing the distributed term that occurs in it with $\hat{x}(t) - e^{A\delta}\hat{x}(t-\delta|t-\delta)$. We obtain

$$\begin{aligned}\mathrm{d}\hat{x}(t) =& A\hat{x}(t)\mathrm{d}t + Bu(t)\mathrm{d}t \\ &+ e^{A\delta}\overline{K}\left(\mathrm{d}y(t) - C\hat{x}(t-\delta|t-\delta)\mathrm{d}t\right) \quad (6)\end{aligned}$$

which is a differential expression of the optimal predictor in (3) and contains both $\hat{x}(t)$ and $\hat{x}(t-\delta|t-\delta)$. Thus, it may be effectively computed together with (4). However, when A is not Hurwitz stable, the estimation error of (6), $\varepsilon(t) = x(t) - \hat{x}(t)$ is not asymptotically centered or uniformly second moment bounded. In fact, let $\eta(t) = x(t-\delta) - \hat{x}(t-\delta|t-\delta)$ be the retarded estimation error. We obtain from (6) that

$$\mathrm{d}\varepsilon(t) = (A\varepsilon(t) - e^{A\delta}\overline{K}C\eta(t))\mathrm{d}t + F\mathrm{d}W_t,$$

and $\varepsilon(t)$ is not asymptotically centered or uniformly second moment bounded when A is not Hurwitz stable.

Thus, we are left with (3) which is computationally expensive. Moreover, in the case of time-varying delays, (3) needs $\hat{x}(t-\delta_t|t-\delta_t)$, and there is no obvious way to compute it through a recursive filter of the kind (4).

Summarizing, it makes sense to look for a sub-optimal but more easily computable estimate of $x(t)$ that is exponentially centered and uniformly second moment bounded. This is particularly important for applications to output-feedback control where an exponentially centered estimate is crucial for the efficacy of the controller, and the bound on the second moment of the estimation error is linked to the energy cost of the control.

3 Structure and properties of the new filter

Let $\overline{A} = A - \overline{K}C$. When Assumption $\mathcal{H}_1$ holds, $\overline{A}$ is Hurwitz stable. The proposed filter is, for $t \geq 0$,

$$\begin{aligned}\mathrm{d}\hat{\xi}(t) =& A\hat{\xi}(t)\mathrm{d}t + Bu(t)\mathrm{d}t \\ &+ (1-\dot{\delta}_t)e^{\overline{A}\delta_t}\overline{K}\left(\mathrm{d}y(t) - C\hat{\xi}(t-\delta_t)\mathrm{d}t\right), \quad (7)\end{aligned}$$

with $\hat{\xi}(\tau) = \hat{\xi}_0(\tau)$ for $\tau \in [0,\bar{\delta}]$, where $\hat{\xi}_0 \in L^2_{-\bar{\delta},0}$ is an arbitrary initial condition. When $\dot{\delta}_t \equiv 0$, there is a certain similarity between (7) and the differential version of the optimal predictor in (6), where the estimates $\hat{x}(t)$ and $\hat{x}(t-\delta|t-\delta)$ are replaced by the estimates $\hat{\xi}(t)$, $\hat{\xi}(t-\delta)$ and $e^{A\delta}$ by $e^{\overline{A}\delta}$.

The implementation of (7) requires a buffer for the values of $\hat{\xi}$ in $[t-\bar{\delta},t]$ but it does not contain distributed terms. When $\delta_t \equiv 0$, (7) coincides with the KBF. When $\delta_t = \bar{\delta}$ is constant, (7) has the same computational complexity as a KBF, since it obtained by replacing $\overline{K}$ with $e^{\overline{A}\delta}\overline{K}$, whereas in the variable case the additional complexity is the computation of $e^{\overline{A}\delta_t}\overline{K}$. Since in a digital implementation the values of δ_t are known with good precision the computation of $\dot{\delta}_t$ is carried out as $(\delta_{k+1}-\delta_k)/\mathrm{d}t$, where $\mathrm{d}t$ is the integration step. In a numerical scheme, the denominator disappears when multiplying the derivative for the integration step $\mathrm{d}t$ and the approximation is robust. Finally, filter (7) is identical to filter (3) of [1] except for the fact that it contains term $1-\dot{\delta}_t$, which is crucial to obtain less conservative estimates for the delay bound and the variance of the estimation error. Our main result is the following.

Theorem 3 *Consider system* (1)–(2) *and assume that* $\mathcal{H}_1$ *and* $\mathcal{H}_2$ *hold. If*

$$\alpha = \int_0^{\bar{\delta}} \left\| Ce^{\overline{A}\theta}\overline{K} \right\| \mathrm{d}\theta \ < 1 \quad (8)$$

then, with $\hat{\xi}(t)$ *computed as in* (7)*, the estimation error* $\varepsilon(t) = x(t) - \hat{\xi}(t)$ *is exponentially centered and uniformly second moment bounded.*

Remark 4 *The delay condition* (8) *is the same as in [1] for constant delays. This new result shows that the same bound holds for time-varying delays too. Instead, the bound for time-varying delays of [1] is obtained by a Razumikhin arguments that leads to a much more conservative delay bound.*

PROOF. By using (1), (2) and (7), we obtain, for $t \geq 0$,

$$\begin{aligned} \mathrm{d}\varepsilon(t) &= A\varepsilon(t)\mathrm{d}t - (1-\dot{\delta}_t)e^{\overline{A}\delta_t}\overline{K}C\varepsilon(t-\delta_t)\mathrm{d}t \\ &\quad + F\mathrm{d}W_t - (1-\dot{\delta}_t)e^{\overline{A}\delta_t}\overline{K}G\mathrm{d}V_{t-\delta_t}, \end{aligned} \tag{9}$$

and, for $s \in [-\bar{\delta}, 0)$, $\varepsilon(s) = x(s) - \hat{\xi}(s) = \phi(s)$. We prove first uniform second moment boundedness of the error. Equation (9) admits the following integral form representation for $t \geq 0$:

$$\begin{aligned} \varepsilon(t) = &\int_{t-\delta_t}^{t} e^{\overline{A}(t-\tau)}\overline{K}C\varepsilon(\tau)\,\mathrm{d}\tau + \int_0^t e^{\overline{A}(t-\tau)}F\,\mathrm{d}W_\tau \\ &- \int_0^{t-\delta_t} e^{\overline{A}(t-\tau)}\overline{K}G\,\mathrm{d}V_\tau + \kappa_\varepsilon, \end{aligned} \tag{10}$$

$$\kappa_\varepsilon = \phi(0) - \int_{-\delta_0}^{0} e^{-\overline{A}\tau}\overline{K}C\phi(\tau)d\tau - \int_{-\delta_0}^{0} e^{-\overline{A}\tau}\overline{K}G\,\mathrm{d}V_\tau.$$

This can be readily checked by explicitly differentiating (10) to obtain (9) and by noticing that, with the prescribed choice of κ_ε, $\varepsilon(0) = \phi(0)$. Since the initial condition is the same in $[-\bar{\delta}, 0]$ the solution of (10) is continuous and coincides with that of (9). Moreover, it follows from (10) that when $C\varepsilon(t)$, that appears in the first integral over a bounded interval, is mean square stable then also $\varepsilon(t)$ is mean square stable. Left-multiplying by C, and taking the L_2 norm yields

$$\begin{aligned} \|C\varepsilon(t)\|_{L_2} \leq &\int_0^{\bar{\delta}} \left\| Ce^{\overline{A}\theta}\overline{K} \right\| \mathrm{d}\theta \cdot \sup_{\tau\in[t-\bar{\delta},t]} \|C\varepsilon(\tau)\|_{L_2} \\ &+ \left(\int_0^t \left\| Ce^{\overline{A}\theta}\left(F + \overline{K}G\right) \right\|^2 \mathrm{d}\theta \right)^{\frac{1}{2}} + \|C\kappa_\varepsilon\|_{L_2} \\ \leq &\ \alpha \sup_{\tau\in[t-\bar{\delta},t]} \|C\varepsilon(\tau)\|_{L_2} + \beta, \end{aligned} \tag{11}$$

where $\alpha < 1$ by hypothesis and

$$\beta = \left(\int_0^\infty \left\| Ce^{\overline{A}\theta}\left(F + \overline{K}G\right) \right\|^2 \mathrm{d}\theta \right)^{1/2} + \|C\kappa_\varepsilon\|_{L_2} < \infty \tag{12}$$

since $\overline{A} = A - \overline{K}C$ is Hurwitz stable, with $\overline{K} = PC^TR^{-1}$ and P solution of (5). Let $m(t) = \sup_{\tau\in[0,t]} \|C\varepsilon(\tau)\|_{L_2}$ and $m_0 = \sup_{\tau\in[-\bar{\delta},0]} \|C\varepsilon(\tau)\|_{L_2}$. Taking the sup of (11) in $[0, t]$ we obtain

$$m(t) \leq \alpha\, m(t) + \alpha\ m_0 + \beta, \tag{13}$$

that proves uniform second moment boundedness of $C\varepsilon(t)$ with bound $(\alpha m_0 + \beta)/(1-\alpha)$. To prove that ε is exponentially centered let $p(t) = \mathbb{E}[\varepsilon(t)]$. Taking expectations in (9) we obtain

$$\dot{p}(t) = Ap(t) - (1-\dot{\delta}_t)e^{\overline{A}\delta_t}\overline{K}Cp(t-\delta_t) \tag{14}$$

Since (8) holds, there exists a constant $\bar{\nu}$ such that,for all $0 < \nu \leq \bar{\nu}$, it holds

$$\int_0^{\bar{\delta}} \left\| Ce^{\overline{A}\theta}\overline{K} \right\| \mathrm{d}\theta\ < \alpha(\nu) = \int_0^{\bar{\delta}} \left\| Ce^{\overline{A}\theta}\overline{K} \right\| e^{\nu\theta}\mathrm{d}\theta\ < 1 \tag{15}$$

Let $p_\nu(t) = e^{\nu t}p(t)$. Clearly, if $p_\nu(t)$ is uniformly bounded then $p(t)$ is exponentially stable. The dynamics of $p_\nu(t)$ is

$$\dot{p}_\nu(t) = A_\nu p_\nu(t) - (1-\dot{\delta}_t)e^{\overline{A}_\nu\delta_t}\overline{K}Cp_\nu(t-\delta_t) \tag{16}$$

where $A_\nu = A + \nu I$ and $\overline{A}_\nu = \overline{A} + \nu I$. $p_\nu(t)$ admits the following integral representation for $t \geq 0$,

$$p_\nu(t) = \int_{t-\delta_t}^{t} e^{\overline{A}_\nu(t-\tau)}\overline{K}Cp_\nu(\tau)\,\mathrm{d}\tau + \kappa_p \tag{17}$$

$$\kappa_p = \mathbb{E}[\varepsilon(0)] - \int_{-\delta_0}^{0} e^{-\overline{A}_\nu\tau}\overline{K}Ce^{\nu\tau}\mathbb{E}[\varepsilon(\tau)], \mathrm{d}\tau. \tag{18}$$

It suffices to prove that $\tilde{p}_\nu(t) = Cp_\nu(t)$ is bounded. By taking the norm in $\mathbb{R}^n$ and proceeding similarly to (11), we get

$$\begin{aligned} \|\tilde{p}_\nu(t)\| \leq &\int_{t-\bar{\delta}}^{t} \|Ce^{\overline{A}_\nu(t-\tau)}\overline{K}\|\,\mathrm{d}\tau \cdot \sup_{\tau\in[t-\bar{\delta},t]} \|\tilde{p}_\nu(\tau)\| \\ &+ \|C\kappa_p\| \\ \|\tilde{p}_\nu(t)\| \leq &\frac{\alpha(\nu)\sup_{\tau\in[-\bar{\delta},0]}\|\tilde{p}_\nu(\tau)\| + \|C\kappa_p\|}{1-\alpha(\nu)}\ < \infty, \end{aligned} \tag{19}$$

since $1 - \alpha(\nu) > 0$ in view of (15).

4 Additional results

4.1 Multiple Delays

Consider the case when the output process $y(t) \in \mathbb{R}^q$ is affected by multiple time-varying delays $\delta_t^i : \mathbb{R}_+ \to [0, \bar{\delta}_i]$. Each delay function affect the output sub-vector $y_i(t) \in \mathbb{R}_i^q$, with $q = \sum_i^q q_i$. The output map is

$$\mathrm{d}y(t) = \mathrm{col}_i\left(C_i x(t-\delta_t^i)\right)\mathrm{d}t + \mathrm{col}_i\left(G_i\mathrm{d}V_{t-\delta_t^i}^i\right) \tag{20}$$

where $C_i \in \mathbb{R}^{q_i\times n}$, $G_i \in \mathbb{R}^{q_i\times l}$, and V_t^i is a $\mathcal{F}_t$ adapted Wiener process in $\mathbb{R}^l$. Filter (7) becomes, for $t \geq 0$,

$$\begin{aligned} d\hat{\xi}(t) = &A\hat{\xi}(t)\mathrm{d}t + Bu(t)\mathrm{d}t \\ &+ \sum_{i=1}^{q}(1-\dot{\delta}_t^i)e^{\overline{A}\delta_t^i}\overline{K}_i\left(\mathrm{d}y_i(t) - C_i\hat{\xi}(t-\delta_t^i)\mathrm{d}t\right) \end{aligned} \tag{21}$$

with $\overline{A} = A - \overline{K}C$, $C = \mathrm{col}_i(C_i)$, $\overline{K} = PC^\top R^{-1}$, $\overline{K} = \mathrm{row}_i(\overline{K}_i)$, and P solution of the Riccati equation (5). Moreover $\hat{\xi}(\tau) = \hat{\xi}_0(\tau)$ for $\tau \in [0, \bar{\delta}]$, where $\hat{\xi}_0 \in L^2_{-\bar{\delta},0}$ is an arbitrary initial condition and $\bar{\delta} = \max\{\bar{\delta}_i\}$.

Theorem 5 *Consider system* (1)–(20) *and assume that* $\mathcal{H}_1$ *and* $\mathcal{H}_2$ *hold true for all* $\delta_t^i : \mathbb{R}_+ \to [0, \bar{\delta}_i]$. *If*

$$\alpha = \sum_{i=1}^{q} \int_0^{\bar{\delta}_i} \left\| C e^{\overline{A}\theta} \overline{K}_i \right\| \mathrm{d}\theta < 1 \tag{22}$$

then the estimation error $\varepsilon(t) = x(t) - \hat{\xi}(t)$ *is exponentially centered and uniformly second moment bounded.*

PROOF. Process $\varepsilon(t) = x(t) - \hat{\xi}(t)$ obeys to

$$\begin{aligned} \mathrm{d}\varepsilon(t) = A\varepsilon(t)\mathrm{d}t - \sum_{i=1}^{q}(1-\dot{\delta}_t^i)e^{\overline{A}\delta_t^i}\overline{K}_i C_i \varepsilon(t-\delta_t^i)\mathrm{d}t \\ + F\mathrm{d}W_t - \sum_{i=1}^{q}(1-\dot{\delta}_t^i)e^{\overline{A}\delta_t^i}\overline{K}_i G_i \mathrm{d}V^i_{t-\delta_t^i}, \end{aligned} \tag{23}$$

and admits, for $t \geq 0$, the representation

$$\begin{aligned} \varepsilon(t) = & \sum_{i=1}^{q} \int_{t-\delta_t^i}^{t} e^{\overline{A}(t-\tau)}\overline{K}_i C_i \varepsilon(\tau)\,\mathrm{d}\tau \\ & + \int_0^t e^{\overline{A}(t-\tau)} F\,\mathrm{d}W_\tau - \sum_{i=1}^{q}\int_0^{t-\delta_t^i} e^{\overline{A}(t-\tau)}\overline{K}_i G_i\,\mathrm{d}V_\tau^i + \kappa_\varepsilon \end{aligned} \tag{24}$$

where κ_ε is chosen to obtain $\varepsilon(0) = \phi(0)$. Left-multiplying by C and taking the L_2 norm we obtain

$$\begin{aligned} \|C\varepsilon(t)\|_{L_2} \leq & \sum_{i=1}^{q} \int_0^{\bar{\delta}^i} \left\| C e^{\overline{A}\theta}\overline{K}_i \right\| \mathrm{d}\theta \\ & \cdot \sup_{\tau \in [t-\bar{\delta}_i, t]} \|C_i \varepsilon(\tau)\|_{L_2} + \beta \\ \leq & \alpha \sup_{\tau \in [t-\bar{\delta}, t]} \|C\varepsilon(\tau)\|_{L_2} + \beta, \end{aligned} \tag{25}$$

where $\alpha < 1$ by hypothesis, β is related to the noise terms and κ_ε, that all have finite norms and hence $\beta < \infty$, and we have used $\|C_i\varepsilon(t)\|_{L_2} \leq \|C\varepsilon(t)\|_{L_2}$. Proceeding as for Theorem 3, we conclude that $\|C_i\varepsilon(t)\|_{L_2} \leq \|C\varepsilon(t)\|_{L_2} < \infty$ that implies $\|\varepsilon(t)\|_{L_2}$ is bounded as a consequence of (24). Always as for Theorem 3, it is easy to prove that $\varepsilon(t)$ is exponentially centered.

4.2 An observer for time-varying delayed measurements

Consider the deterministic version of system (1), (2) with $F = G = 0$,

$$\dot{x}(t) = Ax(t) + Bu(t) \tag{26}$$
$$y(t) = Cx(t - \delta_t) \tag{27}$$

Our estimator becomes

$$\begin{aligned} \dot{\hat{\xi}}(t) = & A\hat{\xi}(t) + Bu(t) \\ & + (1-\dot{\delta}_t)e^{\bar{A}\delta_t}K\left(y(t) - C\hat{\xi}(t-\delta_t)\right). \end{aligned} \tag{28}$$

Under $\mathcal{H}_1$, it is possible to design K such that $\overline{A} = A - KC$ is Hurwitz.

Theorem 6 *Consider system* (26)–(27) *and assume that* $\mathcal{H}_1$ *and* $\mathcal{H}_2$ *hold. If* K *is such that* $\overline{A} = A - KC$ *is Hurwitz,* $c \in [0, -\mu(\overline{A})]$ *and*

$$\alpha_c = \int_0^{\bar{\delta}} \left\| C e^{\overline{A}\theta} K \right\| e^{c\theta}\mathrm{d}\theta < 1, \tag{29}$$

then the estimation error $\varepsilon(t) = x(t) - \hat{\xi}(t)$ *is exponentially stable with rate* $c < -\mu(\overline{A})$.

PROOF. Similarly to Theorem 3, let $\varepsilon_c(t) = e^{ct}\varepsilon(t)$. We obtain

$$\dot{\varepsilon}_c(t) = (A + cI)\varepsilon_c(t) - (1-\dot{\delta}_t)e^{\overline{A}_c\delta_t}KC\varepsilon_c(t-\delta_t), \tag{30}$$

where $\overline{A}_c = \overline{A} + cI$ is Hurwitz because $c + \mu(\overline{A}) < 0$. From

$$\varepsilon_c(t) = \int_{t-\delta_t}^{t} e^{\overline{A}_c(t-\tau)}KC\varepsilon_c(\tau)\,\mathrm{d}\tau + \kappa_\varepsilon \tag{31}$$

and proceeding as in Theorem 3, we obtain the sufficient condition (29) for the uniform boundeness of ε_c that in turn implies exponential stability of ε with rate c.

4.3 Output feedback control

The results of Section 3 can be immediately applied to the delayed output feedback control problem of system (1)–(2) by using the separation principle. For simplicity, we state the result for the case of a single delay, but clearly the extension to multiple output delays is straightforward.

Theorem 7 *Consider system* (1)-(2). *Suppose that* (A, B) *is controllable and Theorem 3 holds, and let* K_c *be such that* $A - BK_c$ *is Hurwitz stable. Then, the control law* $u(t) = -K_c\hat{\xi}(t)$*, where* $\hat{\xi}(t)$ *is computed as in* (7)*, makes the closed loop system exponentially centered and uniformly second moment bounded.*

PROOF. In the hypotheses the estimation error $\varepsilon(t)$ is exponentially centered and uniformly second moment bounded. The closed-loop system is $\mathrm{d}x(t) = (A - BK_c)\, x(t)\mathrm{d}t + BK_c\varepsilon(t) + F\,\mathrm{d}W_t$. Since $A - BK_c$ is Hurwitz stable the theorem thesis follows from standard arguments.

4.4 Large delays

When $\bar{\delta}$ is large, inequality (8) of Theorem 3 may not be satisfied. Let

$$\alpha(\delta) = \int_0^\delta \left\| Ce^{\overline{A}\theta}K \right\| \mathrm{d}\theta \tag{32}$$

and let $\delta_{\max}$ be such that $\alpha(\delta_{\max}) = 1$, or, if $\alpha(\delta) < 1$ for any $\delta > 0$, then $\delta_{\max} = +\infty$. Clearly, $\delta_{\max}$ exists and is unique since $\delta \mapsto \alpha(\delta)$ is continuous and monotonically increasing. Thus, in the case $\bar{\delta} > \delta_{\max}$, we shall design a chain of filters, called *modular filter*, that allows to achieve the same results of Theorem 3 at the expenses of the memory of the filter. Similarly to [2], we propose the following solution.

Definition 8 *Given* $\bar{\delta}, \delta^* > 0$, *a* delay partition $\mathcal{P}_{\bar{\delta},\delta^*}$ *is a set of positive reals* $\{\delta_j\}$, $j = 1, \ldots, m$, *such that* $\delta_j \le \delta^*$ *and* $\sum_{j=1}^m \delta_j = \bar{\delta}$.

We always assume for simplicity $\delta_j = \bar{\delta}/m$. We also denote $d_j = (j-1)\,\bar{\delta}/m$, hence $d_1 = 0$ and $d_{m+1} = \bar{\delta}$.

Definition 9 *Given system* (1)–(2) *and* K *such that* $\overline{A} = A - KC$ *is Hurwitz stable we say that the delay partition* $\mathcal{P}_{\bar{\delta},\delta^*}$ *is* feasible *if* $\alpha(\delta^*) < 1$.

Notice that for any feasible partition, $\alpha(\bar{\delta}/m) < 1$. In order to find m, δ_j and d_j one can proceed as follows:

- find the gain K such that $\overline{A} = A - KC$ is Hurwitz;
- choose $\epsilon \in (0, 1)$ and find δ^* such that $\alpha(\delta^*) = 1 - \epsilon$;
- compute the number of filters $m = \lceil \bar{\delta}/\delta^* \rceil$;
- set $\delta_j = \bar{\delta}/m$ and $d_j = (j-1)\bar{\delta}/m$.

With the partition $\mathcal{P}_{\bar{\delta},\delta^*}$ described above, the modular filter consists of a chain of m filters $\hat{\xi}_j(t)$, $j = 1, \ldots, m$ for $t \ge 0$ where $\hat{\xi}_j(t)$ is an estimate of $x(t-d_j)$, thus $\hat{\xi}_1(t)$ is the estimate of $x(t)$. Clearly, $\delta : \mathbb{R}_+ \to [d_0, d_{m+1}]$. At any t, let ℓ be such that $\delta_t \in [d_\ell, d_{\ell+1}]$, for some $\ell \in \{1, \ldots, m\}$. The filter equations are as follows.

For $j < \ell$,

$$\mathrm{d}\hat{\xi}_j(t) = \Big(A\hat{\xi}_j(t) + Bu(t-d_j) \Big)\,\mathrm{d}t + e^{\overline{A}\delta_j} KC\big(\hat{\xi}_{j+1}(t) - \hat{\xi}_j(t-\delta_j)\big)\mathrm{d}t \tag{33}$$

For $j = \ell$,

$$\mathrm{d}\hat{\xi}_\ell(t) = \Big(A\hat{\xi}_\ell(t) + Bu(t-d_\ell) \Big)\,\mathrm{d}t + (1-\dot{\delta}_t)e^{\overline{A}(\delta_t - d_\ell)} \cdot K\big(\mathrm{d}y(t) - C\hat{\xi}_\ell(t-\delta_t+d_\ell)\mathrm{d}t\big) \tag{34}$$

For $j > \ell$,

$$\hat{\xi}_j(t) = A\hat{\xi}_j(t)\mathrm{d}t + Bu(t-d_j)\mathrm{d}t + K\big(\mathrm{d}\bar{y}(t-d_j) - C\hat{\xi}_j(t)\mathrm{d}t\big) \tag{35}$$

$$\mathrm{d}\bar{y}(t-d_j) = Cx(t-d_j)\mathrm{d}t + GV_{t-d_j} \tag{36}$$

Remark 10 *The value of* $\mathrm{d}\bar{y}(t-d_j)$ *is a past measurement, because* $\delta(t) < d_j$. *The implementation of the modular filter requires to maintain a buffer of the measurements for the interval* $[t-\bar{\delta}, t]$. *At the same time, no artificial delay is introduced: the buffered measurements* $\mathrm{d}\bar{y}$ *are not delayed with respect to* $\hat{\xi}_j$, $j > \ell$.

Theorem 11 *Consider system* (1)–(2) *with a feasible delay partition* $\mathcal{P}_{\bar{\delta},\delta^*}$ *and assume that* $\mathcal{H}_1$ *and* $\mathcal{H}_2$ *hold true. The estimation error* $\varepsilon(t) = x(t) - \hat{\xi}_1(t)$, *where* $\hat{\xi}_1(t)$ *is defined by* (33)–(35), *is exponentially centered and uniformly second moment bounded.*

PROOF. The thesis is easily obtained by induction by noticing that $\hat{\xi}_m$ satisfies the conditions of Theorem 3 and its estimation error is exponentially centered and uniformly second moment bounded. The remaining filters $\hat{\xi}_j$ with $j < m$ satisfy the conditions of Theorem 3 too, and their equation contains an additional disturbance term corresponding to the estimation error ε_{j+1}, that is however exponentially centered and uniformly second moment bounded by inductive hypothesis.

4.5 Piece-wise continuous delays

Theorem 3 can be extended to piece-wise differentiable delays, that is, delays that in any $[0, t]$ admit a derivative except in a finite number of points. Let $T^* = \{t_i\}$ be the countable set of such points. Consider first the case of $\delta(t)$ not differentiable but still continuous.

Corollary 12 *If* $\mathcal{H}_2$ *is replaced by the requirement that* $\delta : \mathbb{R}_+ \to [0, \bar{\delta}]$ *is continuous in* $\mathbb{R}_+$ *and differentiable* *except in a countable set T^* then the results of Theorem 3, Theorem 5 and Theorem 6 continue to hold for the filter obtained by setting $\dot{\delta}_{t_i} = 0$ in* (7).

However, when $\delta(t)$ is not continuous in the set T^* there is an additional bound on the frequency of discontinuities. We prove the result for deterministic systems by extending Theorem 6. The proof of the stochastic case follows in a similar way.

Theorem 13 *If $\mathcal{H}_2$ is replaced by the requirement that $\delta : \mathbb{R}_+ \to [0, \bar{\delta}]$ is differentiable with $\dot{\delta}_t \neq 1$ except in a countable and unbounded set T^* and* (29) *is satisfied then there is a maximum frequency of jump discontinuities $\bar{f} = \sup_k\{1/(t_{k+1} - t_k)\}$ such that* (28) *with $\dot{\delta}_{t_k} = 0$ is an exponential observer of* (26)–(27).

PROOF. When setting $\dot{\delta}_{t_i^*} = 0$ the right-hand side of (7) is not continuous in t_i^* but a continuous solution still exists for $\hat{x}(t)$ and $\varepsilon(t)$. Let us denote $\delta_i^- = \lim_{t \to t_i^-} \delta(t)$ and $\delta_i^+ = \lim_{t \to t_i^+} \delta(t)$. Clearly, in $[0, t_1^*)$ the bound in (19) holds for $\|\varepsilon_c(t)\|$. In $[t_k, t_{k+1}]$ we obtain the following integral representation of the type (31)

$$\begin{aligned}\varepsilon_c(t) = &\int_{t-\delta_t}^{t} e^{\overline{A}_c(t-\tau)} KC\varepsilon_c(\tau)\, \mathrm{d}\tau + \varepsilon_c(t_k^-) \\ &- \int_{t_k-\delta_k^+}^{t_k} e^{\overline{A}_c(t_k-\tau)} KC\varepsilon_c(\tau)\, \mathrm{d}\tau \end{aligned} \tag{37}$$

where $\varepsilon_c(t_k^-) = \lim_{t \to t_k^-} \varepsilon_c(t)$. Proceeding as before we obtain

$$\begin{aligned}\sup_{\tau \in [0, t_{k+1}]} \|C\varepsilon_c(\tau)\| &\le \left(\frac{1+\alpha_c}{1-\alpha_c}\right) \sup_{\tau \in [0, t_k]} \|C\varepsilon_c(\tau)\| \\ &\le \left(\frac{1+\alpha_c}{1-\alpha_c}\right)^k \sup_{\tau \in [0, t_1]} \|C\varepsilon_c(\tau)\|. \end{aligned} \tag{38}$$

Since $k \le t_k \bar{f}$ and $\varepsilon_c(t) = e^{ct}\varepsilon(t)$, denoting $\lambda = \frac{1+\alpha_c}{1-\alpha_c}$,

$$\sup_{\tau \in [0, t_{k+1}]} \|C\varepsilon(\tau)\| \le e^{-(c - \bar{f}\log(\lambda))t_{k+1}} \sup_{\tau \in [0, t_1]} \|C\varepsilon_c(\tau)\| \tag{39}$$

that implies exponential stability when $\bar{f} \le c/\log(\lambda)$.

Corollary 14 *If $\delta : \mathbb{R}_+ \to [0, \bar{\delta}]$ is piece-wise continuous and the cardinality of T^* is finite, the delay bound expressed by* (29) *is sufficient for the conclusions of Theorem 6 to hold.*

5 Examples

We consider a planar tracking problem where system state variable $x = [p_1\ v_1\ p_2\ v_2]^\top$ is composed by planar coordinates (p_1, p_2) and velocities (v_1, v_2). The state equation is

$$\mathrm{d}x(t) = \begin{bmatrix} 0&1&0&0\\ 0&0&0&0\\ 0&0&0&1\\ 0&0&0&0 \end{bmatrix} x(t)\mathrm{d}t + \begin{bmatrix} 0&0\\ \sigma_a&0\\ 0&0\\ 0&\sigma_a \end{bmatrix} \mathrm{d}W_t, \tag{40}$$

where σ_a is the acceleration standard deviation. In all simulations, $\sigma_a = 0.1$. The available measurements are the planar coordinates, *i.e.* $y = [p_1\ p_2]^\top = [x_1\ x_3]^\top$,

$$\mathrm{d}y(t) = \begin{bmatrix} 1&0&0&0\\ 0&0&1&0 \end{bmatrix} x(t - \delta_t)\mathrm{d}t + \begin{bmatrix} \sigma_v&0\\ 0&\sigma_v \end{bmatrix} \mathrm{d}V_{t-\delta_t}, \tag{41}$$

where σ_v is the measurement standard deviation. System (40)–(41) satisfies assumption $\mathcal{H}_1$.

Three simulation setups have been analyzed. In the first one (Setup 1), we aim at showing that a filter with time-varying delays can outperform the optimal predictor with a delay fixed to the maximal value. We set $\sigma_v = 2$ and found from (8) the delay bound $\bar{\delta} = 4.967$. The delay function used in Setup 1 is shown in Fig. 1 (left) and it is such that $\max_t(\delta_t) \approx \bar{\delta}$. Setup 2 aims at showing that the new filter outperforms [1] for large and fast varying delays. We set $\sigma_v = 0.1$, and found $\bar{\delta} = 1.111$ from condition (8). In this case the maximal value of δ_t (shown in Fig. 1, right) largely exceeds the sufficient bound $\bar{\delta}$. In Setup 3, the output equation (41) is transformed in the multiple delays form (20) using the delay functions δ_t^1 and δ_t^2 in Figure 1.

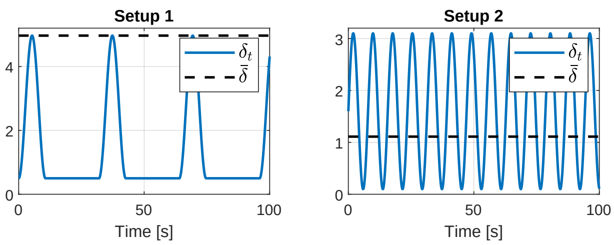

Fig. 1. Delay functions and delay bounds computed from condition (8).

We compared the sampled mean square errors (MSE) of the new filter across 100 noise realizations with (i) the Kalman-Bucy filter with no delays (KBFnd), to show the effects of the delay; (ii) the Kalman-Bucy optimal predictor (KBOP) in (3) based on the maximal delay ($\delta = \max_t(\delta_t)$); (iii) the filter in [1]. The latter has been not applied to Setup 3 since it is not designed for the multiple delay case. Similarly to Setup 2, in Setup 3, condition (22) is not satisfied ($\alpha = 1.748$).

Figure 2 shows the MSEs of the filters in the three simulation setups as a function of time. Table 1 reports the

Table 1
Average Mean Square Errors.

	KBFnd	KBOP	New Filter	[1]
Setup 1	0.723	5.885	1.704	1.722
Setup 2	0.029	0.561	0.818	43.438
Setup 3	0.732	5.885	1.993	-

time averages of the MSE for $t > 20s$ to discard the initial transient. In Setup 1, where condition (8) is satisfied, the new filter has an estimation accuracy comparable to the one of [1], significantly better than the KBOP. In Setup 2, the MSE of the new filter is bounded even if condition (8) does not hold true, whereas the filter in [1] diverges. In Setup 3, with multiple delays, the new filter is bounded, even if condition (22) is not satisfied, and it outperforms the KBOP similarly to Setup 1. A sample trajectory is shown in Fig. 3. This means that (8) is conservative, and that the estimation error of the new filter can be uniformly second moment bounded also in situations when [1] is not.

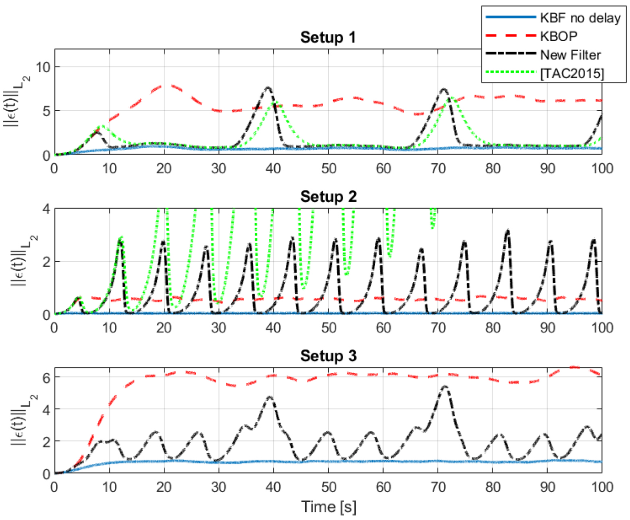

Fig. 2. Mean Square Errors $\|\varepsilon(t)\|_{L_2}$ obtained with in the three simulation setups.

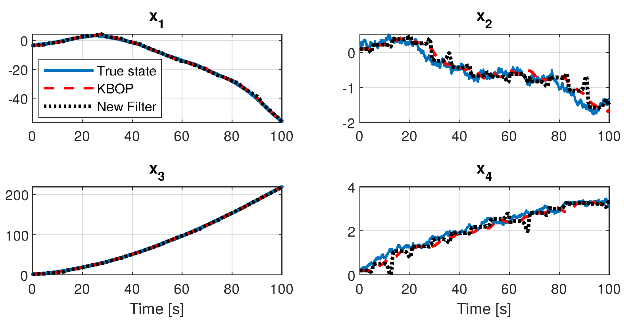

Fig. 3. One sample trajectory of Setup 2.

6 Conclusions

We have shown that including the delay derivative in the filter equations can significantly improve the filtering accuracy and lead to better performance than optimal predictors with artificial delays added. Moreover, the proposed algorithm has a computational complexity comparable to a plain KBF and it is well suited to applications with variable communication delays, for example sensor networks. Further research topics include the choice of the optimal gain $\overline{K}$ as a function of the delay bound as well as the extensions to time-varying linear systems [3], linear systems with nonlinear disturbances and distributed filtering.

References

[1] F. Cacace, F. Conte, and A. Germani. Filtering continuous-time linear systems with time-varying measurement delay. *IEEE Trans. on Automatic Control*, 60(5):1368–1373, 2015.

[2] F. Cacace, F. Conte, A. Germani, and G. Palombo. Optimal control of linear systems with large and variable input delays. *Systems & Control Letters*, 89:1–7, 2016.

[3] F. Cacace, M. d'Angelo, and A. Germani. LTV stochastic systems stabilization with large and variable input delay. *Automatica*, 123:109305, 2021.

[4] C.E. de Souza, R.M. Palhares, and P.L.D. Peres. Robust H_∞ filter design for uncertain linear systems with multiple time-varying state delays. *IEEE Trans. on Signal Processing*, 49(3):569–576, 2001.

[5] H. Dong, Z. Wang, and H. Gao. Distributed H_∞ filtering for a class of Markovian jump nonlinear time-delay systems over lossy sensor networks. *IEEE Transactions on Industrial Electronics*, 60(10):4665–4672, 2012.

[6] E. Fridman, U. Shaked, and L. Xie. Robust H_∞ filtering of linear systems with time-varying delay. *IEEE Trans. on Automatic Control*, 48(1):159–165, 2003.

[7] C. Han, H. Zhang, and G. Feng. Optimal linear estimation for continuous stochastic systems with random observation delays. *International Journal of Robust and Nonlinear Control*, 23(4):359–380, 2013.

[8] A.H. Jazwinski. *Stochastic Processes and Filtering Theory*. Academic Press, 1998.

[9] R.E. Kalman and R.S. Bucy. New results in linear filtering and prediction theory. *Journal of basic engineering*, 83(1):95–108, 1961.

[10] S. Kong, M. Saif, and H. Zhang. Optimal filtering for Itô-stochastic continuous-time systems with multiple delayed measurements. *IEEE Trans. on Automatic Control*, 58(7):1872–1877, 2013.

[11] H. Kwakernaak. Optimal filtering in linear systems with time delays. *IEEE Trans. on Automatic Control*, 12(2):169–173, 1967.

[12] X. Lu, H. Zhang, W. Wang, and K.-L. Teo. Kalman filtering for multiple time-delay systems. *Automatica*, 41(8):1455–1461, 2005.

[13] M.S. Mahmoud. New filter design for linear time-delay systems. *Linear algebra and its applications*, 434(4):1080–1093, 2011.

[14] A.W. Pila, U. Shaked, and C.E. de Souza. H_∞ filtering for continuous-time linear systems with delay. *IEEE Trans. on Automatic Control*, 44(7):1412–1417, 1999.

[15] V.S. Pugachev and I.N. Sinitsyn. *Stochastic systems: theory and applications*. World Scientific Publishing Company, 2002.

[16] S. Sun. Optimal linear filters for discrete-time systems with randomly delayed and lost measurements with/without time stamps. *IEEE Transactions on Automatic Control*, 58(6):1551–1556, 2012.

[17] Y. Sun, S. Kong, G. Cui, and Y. Zhang. Optimal filtering for time-varying stochastic system with delay and multiplicative noise. *IEEE Access*, 7, 2019.

[18] W. Wang, C. Han, and X. Wang. Optimal estimation for continuous-time stochastic systems with delayed measurements and multiplicative noise. In *2020 IEEE 16th International Conference on Control & Automation (ICCA)*, pages 961–966. IEEE, 2020.

[19] H. Zhang, X. Lu, and D. Cheng. Optimal estimation for continuous-time systems with delayed measurements. *IEEE Trans. on Automatic Control*, 51(5):823–827, 2006.

[20] X.-M. Zhang and Q.-L. Han. Robust H_∞ filtering for a class of uncertain linear systems with time-varying delay. *Automatica*, 44(1):157–166, 2008.