Stochastic Output Delay Identification of Discrete-Time Gaussian Systems

Stefano Battilotti [a] and Massimiliano d'Angelo [a]

[a] *Department of Computer, Control, and Management Engineering (DIAG), Sapienza University of Rome, Rome, Italy*

Abstract

In this paper we propose a solution to the problems of detecting a generally correlated stochastic output delay sequence of a linear system driven by Gaussian noise. This is the model for uncertain observations resulting from losses in the propagation channel due to fading phenomena or packet dropouts that is common in wireless sensor networks, networked control systems, or remote sensing applications. The solution we propose consists of a nonlinear detector which identifies online the stochastic delay sequence. The solution provided is optimal in the sense that minimizes the probability of error of the delay detector. Finally, a filtering stage fed with the information given by the detector can follow to estimate the state of the system. Numerical simulations show good performance of the proposed method.

Key words: systems with time-delays; random measurement delays; identification methods; Kalman filtering; Networked control systems; Wireless communications.

1 Introduction

In the last decades a lot of effort has been made to the problem of state estimation and control of systems affected by state process, actuators or sensors delays. In this paper, we shall consider the class of discrete-time linear systems driven by state and output Gaussian noise and we shall focus on the case of random, time-varying and generally correlated measurement delays. Measurement delays can arise in uncertain observations resulting from losses in the propagation channel due to fading phenomena, limited bandwidth or packet dropouts. This situation is common in wireless sensor networks, remote sensing applications and networked control systems which are widely used because of their flexibility dealing with complex systems (Hespanha et al., 2007; Zhang and Yu, 2008). In recent years, random delays have attracted great attention both from the control and the state estimation point of view. Stabilizing controls for systems with random delays in the actuators have been addressed by Zhang and Li (2015); Zhang et al. (2005) and Xiao et al. (2000) by using a coupled Riccati-type equation and jump systems approach, respectively. In the paper of Liu et al. (2009) the stabilization problem for Markov jump linear systems is considered. Xie et al. (2009) and Shi and Yu (2009) investigate the output feedback stabilization of networked control systems in the deterministic framework but with observations affected by random delays, by adopting cone complementarity linearization methods and the design via an iterative linear matrix inequality (LMI) approach, respectively. As regards the state estimation problem, many works have been addressed in the stochastic framework (*i.e.* state and output affected by stochastic noise sequences, Gaussian in general) to filtering systems with intermittent observations or packet dropouts modeled as Bernoulli or Markov chain sequence (Sinopoli et al., 2004; Huang and Dey, 2007; Sun et al., 2008; Ma et al., 2011). Fasano et al. (2015) addressed the problem of filtering systems with intermittent observations with not necessarily Markovian packet dropouts. For what concerns specifically the state estimation problem with random measurement delays, Sahebsara et al. (2007) and Zhou and Feng (2008) consider the case of a bivalued sampling delay, modeled as a Bernoulli random variable. By using $\mathcal{H}_2$ and $\mathcal{H}_\infty$-norm, an LMI-based filter design approach has been developed. Song et al. (2009) have used an $\mathcal{H}_\infty$-norm and LMI techniques to deal with a Markov chain measurement delays by assuming deterministic square integrable state and measurement noise

⋆ This paper was not presented at any IFAC meeting.
Corresponding author: M. d'Angelo. Tel.: +39 0677274034. Fax: +39 0677274131

Email addresses: battilotti@diag.uniroma1.it (Stefano Battilotti), mdangelo@diag.uniroma1.it (Massimiliano d'Angelo).

signals. Han and Zhang (2009) give a linear optimal filtering algorithm in the case of stochastic state and output noise signal by considering a two-valued Markov measurement delay. Zhang et al. (2011) and Schenato (2008) achieve some state estimation results in the case of multiple random delayed measurements where the random delay is known online, *i.e.* time-stamped. Finally, a bayesian approximation framework for nonlinear filtering using cubature quadrature rule to evaluate numerically multivariate integral expressions has been developed recently by Singh et al. (2017). To the best of our knowledge, few attempts have been made to estimate or detect the delay affecting dynamical systems. Some techniques have been developed to identify the deterministic, fixed or time-varying state delay in the case of continuous-time linear (Orlov et al., 2009; Drakunov et al., 2006) and nonlinear systems (Loxton et al., 2010; Zheng et al., 2013; Cacace et al., 2016, 2017). For, by using the maximum a posteriori (MAP) probability decision rule, we propose a method to identify online the generally correlated multiple-valued stochastic output delay which guarantees (with an approximation) the minimum probability of error given the available observations. This technique has been recently employed in the context of distributed estimation in the presence of random communication failures by Battilotti et al. (2018). Finally, we briefly discuss two different subsequent state estimation algorithms that can cope with the delay detection strategy.

The paper is organized as follows. In Section 2 the problem statement with the description of the system model and the assumptions we made is presented. Section 3 presents the optimal delay detection strategy by using a suitable rewriting of the measurements and the MAP probability decision rule. Subsequently, in Section 4 we briefly discuss two state estimation algorithms based on the delay detection strategy. Numerical simulations show the good performance of the proposed detector with respect to the detector of the Blom and Bar-Shalom (1988) in Section 5 and finally, conclusions follow.

Notation. If $A \in \mathbb{R}^{n\times n}$, then $A^\top$ denotes its transpose and $|A|$ denotes its determinant. If $v_1, \dots, v_n$ are column vectors in $\mathbb{R}^n$, then $v = \mathrm{col}(v_1, \dots, v_n)$ denotes the vector $v = [v_1^\top, \dots, v_n^\top]^\top$. Moreover, if $A_1, \dots, A_m \in \mathbb{R}^{n\times n}$, we denote with $\mathrm{diag}(A_1, \dots, A_m) \in \mathbb{R}^{mn\times mn}$ the block diagonal matrix with entries $A_1, \dots, A_m$ on the diagonal. The set of the row vectors of the canonical base of a Euclidean space is denoted by $\mathcal{E}$. In particular, $e_i = (0, \dots, 0, 1, 0 \dots 0) \in \mathcal{E}$, with the one in the $(i+1)$-th entry, and we define the function ind $\cdot$, such that $\mathrm{ind}\, e_i = i$. If A and B are two matrices in $\mathbb{R}^{n\times n}$, then the Kroneker product is $A \otimes B$. Moreover, we indicate with $A(i, j)$ the element of the matrix A in the i-th row and j-th column. We indicate with I_n and 0_n the identity and zero matrices of dimension n, respectively. Moreover, we denote with $0_{n\times m}$ the zero matrix in $\mathbb{R}^{n\times m}$ and with δ_{ij} the Kroneker delta, namely $\delta_{ij} = 1$ if and only if $i = j$. Finally, given a random variable X in the probability space $(\Omega, \mathcal{F}, \mathbb{P})$, that will be implicit in the rest of the paper, we denote with $\mathbb{E}[X]$ its expectation. We denote with $X \sim \mathcal{N}(\mu, \sigma^2)$ a Gaussian random variable X with mean μ and variance σ^2.

2 Problem Statement and Assumptions

Consider the following class of linear discrete-time Gaussian systems

$$x_{k+1} = Ax_k + f_k, \qquad k \geq 0 \tag{1}$$
$$y_k = Cx_{k-\tau_k} + g_k, \tag{2}$$

where $x_k, f_k \in \mathbb{R}^n$, $y_k, g_k \in \mathbb{R}^m$, $A \in \mathbb{R}^{n\times n}$ and $C \in \mathbb{R}^{m\times n}$. The random sequence $\{\tau_k\}$ represents the delay affecting the measurement (2) and it takes value in the set $\mathcal{I} = \{0, 1, \dots, \Delta\}$ with Δ a positive and known integer. Clearly, if at time k we have $\tau_k = 0$, then the measurement equation (2) is not affected by any delay and it gives information on the current state x_k. Moreover, the covariance matrix of the state x_k is denoted by $\Sigma_k = \mathbb{E}[x_k x_k^\top]$. The initial state x_0 and the random sequences $\{f_k\}$, $\{g_k\}$, $\{\tau_k\}$ satisfy the following assumptions for $k \geq 0$:

1) If $\lambda, s \in \{-\Delta, \dots, 0\}$ then x_λ is a zero-mean Gaussian random variable with $\mathbb{E}\{x_\lambda x_s^\top\} = \Sigma_\lambda \delta_{\lambda s}$ symmetric,
2) $\{f_k\}$ is a white sequence of zero-mean Gaussian vectors with covariance matrix $\mathbb{E}\{f_k f_k^\top\} = Q > 0$ symmetric,
3) $\{g_k\}$ is a white sequence of zero-mean Gaussian vectors with covariance matrix $\mathbb{E}\{g_k g_k^\top\} = R > 0$ symmetric,
4) For a fixed integer $L \geq 0$ the following joint probability mass functions are known

$$\begin{cases} \mathrm{P}_k(\tau_k, \dots, \tau_0), & \text{for } 0 \leq k \leq L \\ \mathrm{P}_k(\tau_k, \dots, \tau_{k-L}), & \text{for } k > L \end{cases}$$

5) x_0, $\{f_k\}$, $\{g_k\}$, $\{\tau_k\}$ are statistically independent.

Note that Assumption 4 is a quite minimal requirement on the statistics of the delay sequence. For instance, if the output is characterized by an independent identically distributed delay sequence, namely for each $k \geq 0$ we have $\mathbb{P}(\tau_k = i) = p_i$ with $i \in \mathcal{I}$ and known p_i such that $\sum_{i=0}^{\Delta} p_i = 1$, then the joint probability mass function $\mathrm{P}_k(\tau_k, \dots, \tau_{k-L}) = \mathrm{P}(\tau_k, \dots, \tau_{k-L}) = \prod_{i=0}^{L} \mathrm{P}(\tau_{k-i})$ is fully characterized. In the paper, we shall provide the solution for a generally correlated stochastic sequence of the delay and we further exploit this solution to the particular (but rather common) case of a Markov chain. In the Markov chain case, the knowledge of the probability transition matrix of the chain and of the initial probability vector are sufficient to characterize the joint probability mass function of Assumption 4. Hence, the aforementioned assumption is quite minimal and it allows us to state the problem for a generally correlated delay sequence.

The goal of this paper is to detect the generally-correlated stochastic time-varying delay τ_k in some optimal way that we shall clarify in the next section.

We finally notice that the proposed method could be readily extended to linear time-varying systems.

3 Optimal Delay Identification

In this section we consider the problem of detecting the delay τ_k and we adopt the decision strategy exploited in Fasano et al. (2015) for the problem of state estimation in the case of intermittent observations. We choose as optimal detector the one that guarantees the minimum probability of error given the available observations at time $k \geq 0$, namely $y_0, y_1, \ldots, y_k$. This is equivalent to the maximum a posteriori (MAP) probability decision rule, namely

$$\hat{\tau}_k = \arg\max_{i \in \mathcal{I}} \mathrm{P}_k(\tau_k = i \,|\, y_0, y_1, \ldots, y_k) \tag{3}$$

where $\mathrm{P}_k(\tau_k | y_1, \ldots, y_k)$ is the probability mass function of τ_k conditional on the available observations at time $k \geq 0$. A direct consequence is that the memory and the complexity of the MAP detector (3) increase with time. For, in order to obtain a detector with finite memory, it is sufficient to carry out the decision on the last $L+1$ measurements, with $L \geq 0$, namely

$$\hat{\tau}_k = \arg\max_{i \in \mathcal{I}} \mathrm{P}_k(\tau_k = i \,|\, y_k, \ldots, y_{k-L}), \tag{4}$$

which is the optimal decision rule used in what follows. We need the following preliminary lemma.

Lemma 1 *Given the random vector*

$$\boldsymbol{x}_k := \mathrm{col}(x_k,\, x_{k-1}, \ldots,\, x_{k-\tilde{\Delta}}), \quad k \geq L \tag{5}$$

with $\tilde{\Delta} = \Delta + L$ and x_k defined recursively by (1)*, then the covariance matrix $\Xi_k = \mathbb{E}\left[\boldsymbol{x}_k \boldsymbol{x}_k^\top\right]$ with $k \geq L$ is given by*

$$\Xi_k = \begin{bmatrix} \Sigma_k & A\Sigma_{k-1} & \cdots & A^{\tilde{\Delta}}\Sigma_{k-\tilde{\Delta}} \\ \Sigma_{k-1}A^\top & \Sigma_{k-1} & \cdots & A^{\tilde{\Delta}-1}\Sigma_{k-\tilde{\Delta}} \\ \vdots & \vdots & \ddots & \vdots \\ \Sigma_{k-\tilde{\Delta}}(A^{\tilde{\Delta}})^\top & \Sigma_{k-\tilde{\Delta}}(A^{\tilde{\Delta}-1})^\top & \cdots & \Sigma_{k-\tilde{\Delta}} \end{bmatrix}$$

The proof is omitted.

We are now able to state the main theorem of this paper.

Theorem 1 *Given the system* (1)–(2) *with the Assumptions 1–5, then for $k \geq L$ the MAP delay detector* (4) *is given by*

$$\begin{aligned} \hat{\tau}_k = \mathrm{ind}\arg\min_{v \in \mathcal{E}} \sum_{v_1 \in \mathcal{E}} \sum_{v_2 \in \mathcal{E}} \cdots \\ \cdots \sum_{v_L \in \mathcal{E}} \left[\frac{1}{2} \boldsymbol{y}_k^\top S_k^{-1}(\boldsymbol{v}) \boldsymbol{y}_k - \log \mathrm{P}_k(\boldsymbol{v}) + \right. \\ \left. + \frac{1}{2} \log |S_k(\boldsymbol{v})| \right], \end{aligned} \tag{6}$$

where $\mathcal{E}$ is the set of the row vectors of the canonical basis of $\mathbb{R}^{\Delta+1}$,

$$\begin{aligned} \boldsymbol{y}_k &= \mathrm{col}(y_k, y_{k-1}, \ldots, y_{k-L}), \\ \boldsymbol{v} &= (v, v_1, \ldots, v_L), \qquad v, v_1, \ldots, v_L \in \mathcal{E}, \\ S_k(\boldsymbol{v}) &= \mathcal{C}(\boldsymbol{v}) \Xi_k \mathcal{C}^\top(\boldsymbol{v}) + \mathcal{R}, \end{aligned}$$

with Ξ_k defined in Lemma 1,

$$\mathcal{C}(\boldsymbol{v}) = \begin{bmatrix} v \otimes C & 0_{q \times nL} & \\ 0_{q \times n} & v_1 \otimes C & 0_{q \times n(L-1)} \\ & & \ddots \\ & 0_{q \times nL} & v_L \otimes C \end{bmatrix}, \tag{7}$$

$$\mathcal{R} = I_{L+1} \otimes R. \tag{8}$$

Remark 1 *We note that on the right-hand side of equation* (6)*, i.e. the delay detector, we only require the knowledge of the last $L+1$ measurements, i.e. the vector $\boldsymbol{y}_k$.*

Proof. We define the random vector $\Gamma_k = e_{\tau_k} \in \mathcal{E}$, where e_{τ_k} has the one in the $(\tau_k + 1)$-th entry, and let $\boldsymbol{\Gamma}_k = \mathrm{col}(\Gamma_k, \Gamma_{k-1}, \ldots, \Gamma_{k-L})$. Clearly, the random quantity Γ_k is characterized by $\mathrm{P}_k(\Gamma_k, \ldots, \Gamma_0) = \mathrm{P}_k(\tau_k, \ldots, \tau_0)$ and $\mathrm{P}_k(\Gamma_k, \ldots, \Gamma_{k-L}) = \mathrm{P}_k(\tau_k, \ldots, \tau_{k-L})$, for $0 \leq k \leq L$ and for $k > L$ respectively. Consequently, we can state the equivalent problem of (4) as the following minimization problem

$$\widehat{\Gamma}_k = \arg\max_{v \in \mathcal{E}} \mathrm{P}_k(\Gamma_k = v \,|\, y_k, \ldots, y_{k-L}). \tag{9}$$

We note that, bearing in mind the positions in the statement of Theorem 1, the decision rule (9) can be written as

$$\begin{aligned} \widehat{\Gamma}_k &= \arg\max_{v \in \mathcal{E}} \mathrm{P}(\Gamma_k = v \,|\, \boldsymbol{y}_k) \\ &= \arg\max_{v \in \mathcal{E}} \sum_{v_1 \in \mathcal{E}} \cdots \sum_{v_L \in \mathcal{E}} f(\boldsymbol{y}_k \,|\, \boldsymbol{\Gamma}_k = \boldsymbol{v}) \mathrm{P}_k(\boldsymbol{\Gamma}_k = \boldsymbol{v}) \end{aligned} \tag{10}$$

where $f(\boldsymbol{y}_k|\boldsymbol{\Gamma}_k)$ is the probability density function of $\boldsymbol{y}_k$ conditional on $\boldsymbol{\Gamma}_k$ and $\mathrm{P}_k(\boldsymbol{\Gamma}_k)$ is the probability mass function of $\boldsymbol{\Gamma}_k$, namely the joint probability mass function of $\Gamma_k, \Gamma_{k-1}, \ldots, \Gamma_{k-L}$ which is known by Assumption 4. Note that the probability density function $f(\boldsymbol{y}_k|\boldsymbol{\Gamma}_k)$ is Gaussian since the augmented measurement $\boldsymbol{y}_k$ obeys to

$$\boldsymbol{y}_k = \mathcal{C}(\boldsymbol{\Gamma}_k)\boldsymbol{x}_k + \boldsymbol{g}_k, \tag{11}$$

where $\boldsymbol{g}_k = \mathrm{col}(g_k, g_{k-1}, \ldots, g_{k-L})$, $\mathcal{C}(\cdot)$ is defined in (7) and $\boldsymbol{x}_k$ is defined in (5). Therefore, we have

$$\boldsymbol{y}_k|\boldsymbol{\Gamma}_k \sim \mathcal{N}\big(0, \mathcal{C}(\boldsymbol{\Gamma}_k)\Xi_k\mathcal{C}^\top(\boldsymbol{\Gamma}_k) + \mathcal{R}\big), \tag{12}$$

where $\mathcal{R}$ is defined in (8) and Ξ_k is the covariance matrix of $\boldsymbol{x}_k$ defined in Lemma 1. Finally, the decision rule (10) yields to

$$\begin{aligned}
\widehat{\Gamma}_k &= \arg\max_{v\in\mathcal{E}} \sum_{v_1\in\mathcal{E}} \sum_{v_2\in\mathcal{E}} \cdots \\
&\quad \cdots \sum_{v_L\in\mathcal{E}} \frac{\exp\left[-\frac{1}{2}\boldsymbol{y}_k^\top S_k^{-1}(\boldsymbol{v})\boldsymbol{y}_k\right] \mathrm{P}_k(\boldsymbol{\Gamma}_k = \boldsymbol{v})}{\sqrt{|S_k(\boldsymbol{v})|}} \\
&= \arg\min_{v\in\mathcal{E}} \sum_{v_1\in\mathcal{E}} \sum_{v_2\in\mathcal{E}} \cdots \\
&\quad \cdots \sum_{v_L\in\mathcal{E}} \left[\frac{1}{2}\boldsymbol{y}_k^\top S_k^{-1}(\boldsymbol{v})\boldsymbol{y}_k - \log \mathrm{P}_k(\boldsymbol{\Gamma}_k = \boldsymbol{v}) + \right. \\
&\qquad \left. + \frac{1}{2}\log|S_k(\boldsymbol{v})|\right].
\end{aligned} \tag{13}$$

It is clear that, by the definition of Γ_k, if at time $k \geq 0$ we have $\widehat{\Gamma}_k = e_i$ for some $i \in \mathcal{I}$, then we have $\hat{\tau}_k = i$, namely $\mathrm{ind}\,\widehat{\Gamma}_k = \hat{\tau}_k$, and the proof is completed.

We note that the proposed detector (6) is computed by comparing the $\Delta + 1$ hypotheses $v = e_i$ with $i \in \mathcal{I}$ and moreover, it has finite memory and a tradeoff between performance and computational cost can be easily controlled by changing the value of the memory L of the detector.

Finally, we note that in the initial transient period $k \in \{0, 1, \ldots, L\}$, it is possible to construct the optimal detector (3) by using simply the detector (6) of Theorem 1 with time-varying memory $L = k$.

3.1 The case of Markov chain delay

If the delay sequence $\{\tau_k\}$ is a Markov chain with initial probability vector

$$\pi_0 = \Big[\mathbb{P}\{\tau_0 = 0\}\ \mathbb{P}\{\tau_0 = 1\}\ \cdots\ \mathbb{P}\{\tau_0 = \Delta\}\Big] \tag{14}$$

and probability transition matrix with elements

$$\Pi_k(i+1, j+1) = \mathbb{P}\{\tau_k = j \,|\, \tau_k = i\}, \tag{15}$$

with $i, j \in \mathcal{I}$, then the probability mass function $\mathrm{P}_k(\Gamma_k)$ takes the following expression

$$\mathrm{P}_k(\Gamma_k) = \pi_0 \prod_{\ell=1}^{k} \Pi_\ell \Gamma_k^\top \tag{16}$$

which simplifies in the case of homogeneous chain in

$$\mathrm{P}_k(\Gamma_k) = \pi_0 \Pi^k \Gamma_k^\top. \tag{17}$$

Therefore, the probability mass function $\mathrm{P}_k(\boldsymbol{\Gamma}_k)$ is given by

$$\begin{aligned}
\mathrm{P}_k(\boldsymbol{\Gamma}_k) &= \prod_{\ell=0}^{L-1} \mathrm{P}_{k-i}(\Gamma_{k-\ell} \,|\, \Gamma_{k-\ell-1})\mathrm{P}_{k-L}(\Gamma_{k-L}) \\
&= \prod_{\ell=0}^{L-1} \Pi_{k-\ell}(\tau_{k-\ell-1} + 1, \tau_{k-\ell} + 1)\mathrm{P}_{k-L}(\Gamma_{k-L})
\end{aligned}$$

where $\mathrm{P}_{k-L}(\Gamma_{k-L})$ is defined in (16). Note that in the equation above we have exploited the definition of Γ_k. In the case of homogeneous chain, the probability mass function $\mathrm{P}_k(\boldsymbol{\Gamma}_k)$ simplifies into

$$\mathrm{P}_k(\boldsymbol{\Gamma}_k) = \prod_{\ell=0}^{L-1} \Pi(\tau_{k-\ell-1}+1, \tau_{k-\ell}+1)\mathrm{P}_{k-L}(\Gamma_{k-L}) \tag{18}$$

with $\mathrm{P}_{k-L}(\Gamma_{k-L})$ as in (17).

4 State Estimation

In this section we briefly discuss two state estimation algorithms to solve the problem of filtering systems of the form (1)–(2) that could cope with the proposed detector.

Kalman filtering based. The first method consists of a Kalman filter applied to the extended system [1]

$$\boldsymbol{x}_{k+1} = \mathcal{A}\boldsymbol{x}_k + \boldsymbol{f}_k \tag{19}$$

$$y_k = \mathcal{C}_{\tau_k}\boldsymbol{x}_k + g_k, \tag{20}$$

with $\mathcal{C}_{\tau_k} \doteq e_{\tau_k} \otimes C$, where the real values of the delay sequence $\{\tau_k\}$ are substituted with the detected ones $\{\hat{\tau}_k\}$. It is clear that the lower the probability of error of the detector (6) is, the closer this estimate is to the ideal one which use the real values of the delay. The algorithm can be written as follows

$$\hat{\boldsymbol{x}}_0 = 0, \qquad P_0 = \mathrm{diag}\{\Sigma_0, \ldots, \Sigma_{-\Delta}\} \tag{21}$$

[1] we note that in (19) the vector $\boldsymbol{x}_k$ is defined as (5) but with $L = 0$ (*i.e.* $\tilde{\Delta} = \Delta$), thus the initial condition $\boldsymbol{x}_0$ is properly defined by Assumption 1. Moreover, it is easy to find $\mathcal{A}$ and the random vector $\boldsymbol{f}_k$ (with covariance denoted by $\mathcal{Q}$) such that (19) holds true.

$$\hat{\tau}_k = \begin{cases} \arg\max_{\tau_k \in \mathcal{I}} \mathrm{P}_k(\tau_k \,|\, y_k, \dots, y_1), & \text{if } k < L \\ \\ \arg\max_{\tau_k \in \mathcal{I}} \mathrm{P}_k(\tau_k \,|\, y_k, \dots, y_{k-L}), & \text{if } k \geq L \end{cases} \quad (22)$$

$$\hat{\boldsymbol{x}}_{k|k-1} = \mathcal{A}\hat{\boldsymbol{x}}_{k-1} \quad (23)$$

$$P_{k|k-1} = \mathcal{A}P_{k-1}\mathcal{A}^\top + \mathcal{Q} \quad (24)$$

$$K_k = P_{k|k-1}\mathcal{C}_{\hat{\tau}_k}^\top \cdot \left(\mathcal{C}_{\hat{\tau}_k} P_{k|k-1}\mathcal{C}_{\hat{\tau}_k}^\top + R\right)^{-1} \quad (25)$$

$$\hat{\boldsymbol{x}}_k = \hat{\boldsymbol{x}}_{k|k-1} + K_k(y_k - \mathcal{C}_{\hat{\tau}_k}\hat{\boldsymbol{x}}_{k|k-1}) \quad (26)$$

$$P_k = P_{k|k-1} - K_k\mathcal{C}_{\hat{\tau}_k}P_{k|k-1}, \quad (27)$$

where the first $n \times n$ blocks of $P_{k|k-1}$ and P_{k-1} are the covariance matrices of the prediction and estimation errors, respectively. Finally, we note that the first n entries of $\hat{\boldsymbol{x}}_k$ return the estimated state $\hat{x}_k$.

IMM filtering based. For the algorithm presented in this section we assume that the delay sequence $\{\tau_k\}$ is a Markov chain. The proposed method is the combination of the delay detection strategy exposed in Section 3 and the IMM filter of Blom and Bar-Shalom (1988). In the IMM algorithm the final state estimate is provided as a weighted mean of intermediate estimates. In particular, by referring to Table I of Zhang and Li (1998), the IMM algorithm can be modified by using the following weights in the *Mode probability update* step

$$\tilde{p}_k(i) = \frac{\mathbb{P}(\tau_k = i | y_k, \dots, y_{k-L})}{\sum_{j=0}^{\Delta} \mathbb{P}(\tau_k = j | y_k, \dots, y_{k-L})}. \quad (28)$$

Note that (28) can be directly computed through the argument of the maximization on the right-hand side of (6). In fact, by setting $\zeta_i = \mathrm{col}(e_i, v_1, \dots, v_L)$, where $i \in \mathcal{I}$ and $v_1, \dots, v_L \in \mathcal{E}_\Delta$, and

$$\eta_k(i) = \sum_{v_1 \in \mathcal{E}_\Delta} \cdots \sum_{v_L \in \mathcal{E}_\Delta} \left[\frac{1}{2}\boldsymbol{y}_k^\top S_k^{-1}(\zeta_i)\boldsymbol{y}_k + - \log \mathrm{P}_k(\zeta_i) + \frac{1}{2}\log|S_k(\zeta_i)|\right],$$

the weights in (28) can be written as

$$\tilde{p}_k(i) = \frac{\eta_k(i)}{\sum_{j=0}^{\Delta} \eta_k(j)}. \quad (29)$$

The overall scheme of the proposed methods is reported in Figure 1.

5 Simulation Results

In this section we provide some numerical examples to show the performance of the detector of Theorem 1 compared with the detector of the IMM algorithm of Blom and Bar-Shalom (1988). We consider a time horizon of 150 and 300 independent realizations.

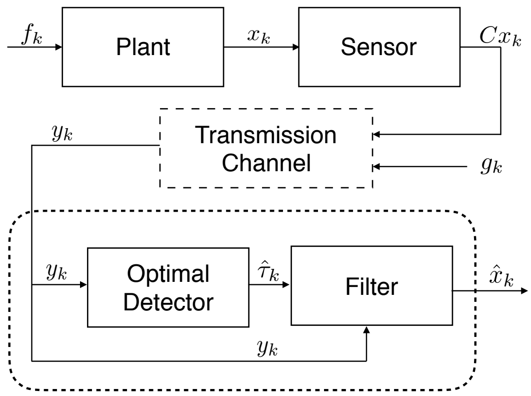

Fig. 1. Overall scheme of the proposed approach.

We consider the linear discrete-time system of the form (1)–(2) described by

$$A = \begin{bmatrix} 0.8 & 0.1 \\ 0 & 0.6 \end{bmatrix}, \qquad C = \begin{bmatrix} 1 & 0 \end{bmatrix} \quad (30)$$

with $\{f_k\}$ and $\{g_k\}$ zero-mean and Gaussian with covariance $Q = 5 \cdot 10^{-2}$ and $R = 10^{-4}$, respectively. Moreover, the initial conditions of the state are $x_\lambda \sim \mathcal{N}(0, I_2)$ for $\lambda \in \{-\Delta, \dots, 0\}$. We consider two cases for the delay sequence $\{\tau_k\}$, modeled as a Markov chain, namely two different probability transition matrices

$$\Pi_1 = \begin{bmatrix} 0.4 & 0.25 & 0.2 & 0.15 \\ 0.2 & 0.4 & 0.25 & 0.15 \\ 0.2 & 0.25 & 0.3 & 0.25 \\ 0.35 & 0.2 & 0.2 & 0.25 \end{bmatrix}, \quad \Pi_2 = \begin{bmatrix} 0.5 & 0.3 & 0.2 \\ 0.3 & 0.4 & 0.3 \\ 0.4 & 0.4 & 0.2 \end{bmatrix}.$$

Thus, in the case of Π_1 the bound Δ on the delay is $\Delta_1 = 3$, whilst in the case of Π_2 the bound on the delay is $\Delta_2 = 2$. Moreover, we assume that $\tau_0 = 0$, *i.e.* $\pi_0 = (1, 0, \dots, 0)$.

In the first example, we show the dependency of the detector of Theorem 1 on the memory L in the case with Π_1. Table 1 summarizes the results for all the values of $L \in \{0, 1, 2, 3, 4\}$. It shows the probability of error, namely

$$P_{\text{err}} \doteq \frac{1}{300}\frac{1}{150}\sum_{i=1}^{300}\sum_{k=1}^{150} \mathbb{P}\{\hat{\tau}_k^{(i)} \neq \tau_k^{(i)}\}, \quad (31)$$

where $\tau_k^{(i)}$ and $\hat{\tau}_k^{(i)}$ are the delay and the detected delay of the i-th realization of noise and delay sequence at time k, and the percentage improvement α with respect to the detector of the IMM algorithm, namely

case Π_1	P_{err}	α
IMM detector	0.633	-
Proposed detector $L = 0$	0.702	-10.90%
Proposed detector $L = 1$	0.639	-0.95%
Proposed detector $L = 2$	0.581	8.21%
Proposed detector $L = 3$	0.540	14.69%
Proposed detector $L = 4$	0.536	15.32%

Table 1
Example 1. Scenario with Π_1.Performance of the proposed detector with memory $L \in \{0, 1, 2, 3, 4\}$ and the detector of the IMM algorithm in terms of the probability of error P_{err} of (31) and the percentage index α. The time horizon and the independent realizations are 150 and 300 respectively.

$\alpha = \left(P_{\text{err}}^{\text{IMM}} - P_{\text{err}}\right) \cdot 100 / P_{\text{err}}^{\text{IMM}}$, where $P_{\text{err}}^{\text{IMM}}$ is the probability of error of the IMM detector.

We see that the slight increase in computational cost is well compensated by the improvement in performance. When the value of the memory L of the detector increases, then the probability of error decreases and it tends to a limit value. This is intuitive since, in order to estimate τ_k at time $k > 0$, further measurements convey less information than the ones close to the current instant k. In particular, we note that the proposed detector of memory $L = 4$ (*i.e.* carrying the five last measurements in the delay identification process) has a probability of error of about 0.53. We note that, the detector with memory $L = 0$ has similar performance of a trivial detector which decides for the *a priori* "most probable" value of the delay. In fact, we see as the steady state probability transition matrix, *i.e.* $\lim_{k \to +\infty} \Pi_k$, has rows $(0.2862, 0.2828, 0.2379, 0.1931)$ and thus the event $\{\tau_k = 0\}$ has the highest probability. Hence, by setting as a trivial decision of the delay $\check{\tau}_k = 0$ for any $k \geq 0$, one would obtain a probability of error $P_{\text{err}} \simeq 1 - 0.2862 = 0.7138$ which is close to the probability of error of the detector of memory $L = 0$. We finally note that the proposed detector has finite memory and a tradeoff between performance and computational cost can be easily controlled by changing the value of the memory L of the detector.

In the second example, we consider a fixed delay sequence, characterized by Π_2, across the 300 independent realizations of noises and initial conditions of the state. Figure 2 shows the real delay sequence τ_k and the mean across realizations of the detected ones, namely $\left\lfloor \frac{1}{300} \sum_{i=1}^{300} \hat{\tau}_k^{(i)} \right\rfloor$, where $\lfloor \cdot \rfloor$ is the floor function. Table 2 summarizes the results. In particular, it shows the values of the probability of error computed as (31) for the proposed detector of memory $L = 3$ and the one of the IMM algorithm. As Table 1 highlights, we see that the proposed detector with memory $L = 3$ improve the probability of error of the detected delay with respect to the IMM algorithm of about 28%. We finally stress that, in contrast with the IMM detector, the proposed detector can be applied to a generally correlated delay sequence.

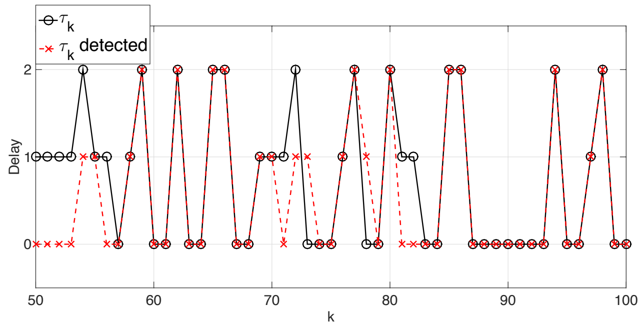

Fig. 2. Plot of 50 time steps of the real delay τ_k characterized by Π_2 and the detected one $\hat{\tau}_k$ with the detector of memory $L = 3$.

case Π_2	P_{err}	α
IMM detector	0.545	-
Proposed detector $L = 3$	0.390	28.12%

Table 2
Example 2. Scenario with Π_2. Performance of the proposed detector with $L = 3$ and the detector of the IMM algorithm in terms of the probability of error P_{err} of (31) and the percentage index α. The time horizon and the independent realizations are 150 and 300 respectively.

6 Conclusions

In this paper a solution to the problem of detecting the stochastic generally correlated output delay sequence of a linear system with Gaussian noise is given. We adopt as optimal detector the one that guarantees the minimum probability of error given the available observations, which is equivalent to the maximum a posteriori (MAP) probability decision rule. The proposed detector has finite memory and a tradeoff between performance and computational cost can be easily controlled by changing the memory of the detector. Finally, a filtering stage can cope with the proposed detector. Numerical comparisons with the detector of the IMM algorithm show good performance of the proposed method.

References

Battilotti, S., Cacace, F., d'Angelo, M., and Germani, A. (2018). Distributed kalman filtering over sensor networks with unknown random link failures. *IEEE control systems letters*, 2(4):587–592.

Blom, H. A. and Bar-Shalom, Y. (1988). The interacting multiple model algorithm for systems with markovian switching coefficients. *IEEE Transactions on Automatic Control*, 33(8):780–783.

Cacace, F., Conte, F., Germani, A., and Palombo, G. (2016). Delay identification for a class of nonlinear systems. *International Journal of Control*, 89(11):2350–2359.

Cacace, F., Conte, F., Germani, A., and Palombo, G. (2017). Joint state estimation and delay identification for nonlinear systems with delayed measurements. *IEEE Transactions on Automatic Control*, 62(9):4848–4854.

Drakunov, S. V., Perruquetti, W., Richard, J.-P., and Belkoura, L. (2006). Delay identification in time-delay systems using variable structure observers. *Annual reviews in control*, 30(2):143–158.

Fasano, A., Monteriù, A., and Villani, V. (2015). A detection-estimation approach to filtering with intermittent observations with generally correlated packet dropouts. In *2015 54th IEEE Conference on Decision and Control (CDC)*, pages 4356–4361. IEEE.

Han, C. and Zhang, H. (2009). Linear optimal filtering for discrete-time systems with random jump delays. *Signal Processing*, 89(6):1121–1128.

Hespanha, J. P., Naghshtabrizi, P., and Xu, Y. (2007). A survey of recent results in networked control systems. *Proceedings of the IEEE*, 95(1):138–162.

Huang, M. and Dey, S. (2007). Stability of kalman filtering with markovian packet losses. *Automatica*, 43(4):598–607.

Liu, M., Ho, D. W., and Niu, Y. (2009). Stabilization of markovian jump linear system over networks with random communication delay. *Automatica*, 45(2):416–421.

Loxton, R., Teo, K. L., and Rehbock, V. (2010). An optimization approach to state-delay identification. *IEEE Transactions on Automatic Control*, 55(9):2113–2119.

Ma, X., Djouadi, S. M., and Charalambous, C. D. (2011). Optimal filtering over uncertain wireless communication channels. *IEEE Signal Processing Letters*, 18(6):359–362.

Orlov, Y., Kolmanovsky, I., and Gomez, O. (2009). Adaptive identification of linear time-delay systems: From theory toward application to engine transient fuel identification. *International Journal of Adaptive Control and Signal Processing*, 23(2):150–165.

Sahebsara, M., Chen, T., and Shah, S. L. (2007). Optimal filtering with random sensor delay, multiple packet dropout and uncertain observations. *International Journal of Control*, 80(2):292–301.

Schenato, L. (2008). Optimal estimation in networked control systems subject to random delay and packet drop. *IEEE Transactions on Automatic Control*, 53(5):1311.

Shi, Y. and Yu, B. (2009). Output feedback stabilization of networked control systems with random delays modeled by markov chains. *IEEE Transactions on Automatic Control*, 54(7):1668–1674.

Singh, A. K., Date, P., and Bhaumik, S. (2017). A modified bayesian filter for randomly delayed measurements. *IEEE Transactions on Automatic Control*, 62(1):419–424.

Sinopoli, B., Schenato, L., Franceschetti, M., Poolla, K., Jordan, M. I., and Sastry, S. S. (2004). Kalman filtering with intermittent observations. *IEEE Transactions on Automatic Control*, 49(9):1453–1464.

Song, H., Yu, L., and Zhang, W.-A. (2009). h_∞ filtering of network-based systems with random delay. *Signal Processing*, 89(4):615–622.

Sun, S., Xie, L., Xiao, W., and Soh, Y. C. (2008). Optimal linear estimation for systems with multiple packet dropouts. *Automatica*, 44(5):1333–1342.

Xiao, L., Hassibi, A., and How, J. P. (2000). Control with random communication delays via a discrete-time jump system approach. In *Proceedings of the 2000 American Control Conference. ACC.*, volume 3, pages 2199–2204. IEEE.

Xie, D., Zhang, D., Han, X., Huang, H., and Wang, Z. (2009). Dynamic output feedback fault-tolerant control for networked control system with random delays. *IFAC Proceedings Volumes*, 42(8):1049–1054.

Zhang, H., Feng, G., and Han, C. (2011). Linear estimation for random delay systems. *Systems & Control Letters*, 60(7):450–459.

Zhang, H. and Li, L. (2015). Stabilization control for discrete time systems with random delay. In *2015 American Control Conference (ACC)*, pages 4180–4185. IEEE.

Zhang, L., Shi, Y., Chen, T., and Huang, B. (2005). A new method for stabilization of networked control systems with random delays. *IEEE Transactions on Automatic Control*, 50(8):1177–1181.

Zhang, W.-A. and Yu, L. (2008). Modelling and control of networked control systems with both network-induced delay and packet-dropout. *Automatica*, 44(12):3206–3210.

Zhang, Y. and Li, X. R. (1998). Detection and diagnosis of sensor and actuator failures using imm estimator. *IEEE Transactions on aerospace and electronic systems*, 34(4):1293–1313.

Zheng, G., Barbot, J.-P., and Boutat, D. (2013). Identification of the delay parameter for nonlinear time-delay systems with unknown inputs. *Automatica*, 49(6):1755–1760.

Zhou, S. and Feng, G. (2008). h_∞ filtering for discrete-time systems with randomly varying sensor delays. *Automatica*, 44(7):1918–1922.