Proceedings of the 8th
World Congress on Intelligent Control and Automation
July 6-9 2010, Jinan, China

# Linear Estimation for Discrete-time Systems with Markov Jump Delays

Chunyan Han and Huanshui Zhang,
*School of Control Science and Engineering*
*Shandong University*
*Jingshi Road 73, Jinan 250061, China*
*cyhan823@hotmail.com, hszhang@sdu.edu.cn*

Minyue Fu
*School of Electrical Engineering and Computer Science*
*University of Newcastle*
*Callaghan, NSW 2308, Australia*
*minyue.fu@newcastle.edu.au*

*Abstract*— **This paper is concerned with the linear minimum mean square error (MMSE) estimation for discrete-time systems with random delays in the observations. It is assumed that the delay process is modeled as a finite state Markov chain and only its transition probability matrix is known. To overcome the difficulty of estimation caused by random delays, the random delay system is firstly rewritten as a constant delay system with multiplicative noises. By applying the measurement reorganization approach, the system is further transformed into the delay-free one with Markov jump parameters. Then the estimator is derived by using the innovation analysis method in the Hilbert space, and the solution is given in terms of Riccati difference equations.**

***Index Terms*— *Linear estimation, discrete-time system, Markov jump delay, innovation analysis method, Riccati equations.***

## I. INTRODUCTION

Linear estimation for time-delay systems has been an active research area since the 1960's, see e.g. [1]. There have been many different kinds of estimators designed under different conditions. For the systems with constant delays, many effective techniques have been developed, such as the classical state augmentation method [1], the polynomial approach [2], the linear matrix inequality (LMI) algorithms [3], and the reorganization innovation analysis method [4]. It is worth pointing out that most of the estimation problem in systems with constant delays have been well studied. However, in the case of systems involving random delays, the estimation problem becomes very difficult and it remains to be challenging.

Time delays occur in a random way, rather than a deterministic way, for a number of engineering applications. Examples include real-time distributed decision making systems, multiplexed communication networks [5], [6]. Hence, there is a great need to develop new estimator approaches for the system with randomly varying delays, and some efforts have been made in this regard so far. For the case of observations with irregular times, the minimum variance state estimators were designed in [7]. On the filtering problems with intermittent observations, the initial work can be tracked back to Nahi [8] and Hadidi [9]. Recently, this problem has been studied in [10] and [11], respectively. For the situation that the one-step sensor delay was described as a binary white noise sequence, a reduced-order linear unbiased estimator was designed via state augmentation in [12]. When the random delay was characterized by a set of Bernoulli variables, the unscented filtering algorithms [13], the linear and quadratic least-square estimation method [14] and $H_\infty$ filter [15] have been presented. The rationality of modeling the random delay as Bernoulli variable sequences has been justified in the references above. In addition, modeling the random delay as a finite state Markov chain is also a reasonable way. The concerning estimation results for this type of modeling can be found in [16], [17], and the reference therein.

Although estimation problems with random delays have been studied for several years, there are still some interesting problems that deserve further research. For example, most of the proposed results focus on the case that the random delay is known online via time-stamped data, while few results study the case that the random delay is unknown. And further, most existing results employ the state augmentation method to deal with time delays. In fact, there exist another efficient way to handle the delay terms, which reorganizes delayed observations as delay-free ones [4].

Motivated by [4], the purpose of this paper is to investigate linear MMSE estimation for systems with unknown random delays, where the delay process is characterized by a finite state Markov chain space.The key technique applied for treating the random delay is the re-organization analysis method. In order to solve the estimation problem, the random delayed measurement is firstly rewritten as a two-channel constant delayed measurement system with multiplicative noises, where the noises are jump variables and are mutually independent. Then, with the application of measurement reorganization, the delayed measurement system is further transformed into a two channel delay-free system, and thus

*This work is supported by the Taishan Scholar Construction Engineering by Shandong Government, the National Natural Science Foundation for Distinguished Young Scholars of China (No. 60825304), and the Major State Basic Research Development Program of China (973 Program) (No. 2009cb320600)

978-1-4244-6712-9/10/$26.00 

the proposed estimation problem can be reformulated as the one for a class of Markov jump linear systems without delays. Finally, the estimator is derived by using the innovation analysis method in the Hilbert space, and the solution is given in terms of two Riccati equations.

*Notation:* The notation in this paper is fairly standard. For $L \in \mathcal{R}^{n\times n}$, $L^T$ denotes the transpose of $L$, and $L \geq 0$ ($L > 0$) means that the symmetric matrix $L$ is positive semi-definite (positive definite). For a collection of $N$ matrices $D_1, \cdots, D_N$, with $D_j \in \mathcal{R}^{n\times m}$, diag$\{D_j\} \in \mathcal{R}^{Nn\times Nm}$ represents the diagonal matrix formed by $D_j$ in the diagonal and zero elsewhere. In addition, $\mathbf{1}_{\{.\}}$ stands for the Dirac measure, and E(.) represents the operator expected value.

## II. Problem Formulation

Consider the following discrete-time systems with random jump delays

$$x(k+1) = \Phi(k)x(k) + \Gamma(k)w(k), \tag{1}$$
$$y_{d(k)}(k) = H(k)x(k-d(k)) + v(k-d(k)), \tag{2}$$

where $x(k)$ denotes the $\mathcal{R}^n$-valued state sequence, $w(k)$ and $v(k)$ are random disturbances in $\mathcal{R}^{q_1}$ and $\mathcal{R}^m$, respectively, $y_.(k)$ is the $\mathcal{R}^m$-valued output sequence, and $d(k)$ is the random time delay of the system. $\Phi(k), \Gamma(k)$ and $H(k)$ are matrices of appropriate dimension. We first make the following assumptions for the above system.

*Assumption 1:* $w(k)$ and $v(k)$ are zero-mean, second-order, independent wide sense stationary sequences with covariance matrices $Q_w$ and $Q_v$. Here the covariance matrix $Q_v$ is positive definite.

*Assumption 2:* The initial state $x(0)$ is also a second-order, independent wide sense stationary sequence with zero mean and covariance matrix $\mathrm{E}[x(0)x^T(0)] = V$.

*Assumption 3:* $d(k)$ is a discrete-time Markov chain with finite state space $\{0, d\}$, and transition probability matrix $P = [p_{ij}]$. We set $p_i(k) = \mathrm{Prob}(d(k) = i), i = 0, d$ and denote $p_k = [p_0(k)\ p_d(k)]^T$, which satisfies the difference equation $p_k = P^T p_{k-1}$.

*Assumption 4:* $x(0)$ and $\{d(k)\}$ are independent of $\{w(k)\}$ and $\{v(k)\}$.

Given the above statements, we know that the measurement output may or may not experience sensor delays, and thus it can be written as two equivalent random measurement channels

$$y_0(k) = \gamma_{k,0}H(k)x(k-0) + \gamma_{k,0}v(k-0), \tag{3}$$
$$y_d(k) = \gamma_{k,d}H(k)x(k-d) + \gamma_{k,d}v(k-d), \tag{4}$$

where

$$\gamma_{k,i} = \begin{cases} 1, & \text{if } d(k) = i; \\ 0, & \text{if } d(k) \neq i, \end{cases}$$

for $i = 0, d$. As remarked in [6], $\gamma_{k,0}$ as well as $\gamma_{k,d}$ is a Markov process which represents the jumping characteristics of $d(k)$. Namely,

$$\begin{aligned} \mathrm{Prob}(\gamma_{k,0} = 1) &= \mathrm{Prob}(d(k) = 0) = p_0(k), \\ \mathrm{Prob}(\gamma_{k,d} = 1) &= \mathrm{Prob}(d(k) = d) = p_d(k), \\ \mathrm{Prob}(\gamma_{k,0} = 0) &= 1 - p_0(k) = p_d(k), \\ \mathrm{Prob}(\gamma_{k,d} = 0) &= 1 - p_d(k) = p_0(k). \end{aligned}$$

In addition, $\gamma_{k,0}$ has the same transition probability matrix as $d(k)$, while the one of $\gamma_{k,d}$ is described as

$$\begin{bmatrix} p_{dd} & p_{d0} \\ p_{0d} & p_{00} \end{bmatrix}.$$

Let $y(k)$ denote all possible observation sequences of the systems (1), (3) and (4) at the time $k$, then we have that

$$y(k) = \begin{cases} y_0(k), & 0 \leq k < d; \\ \left[y_0^T(k)\ y_d^T(k)\right]^T, & k \geq d. \end{cases} \tag{5}$$

So the estimation problem can be stated as: Given the observations $\{y(s)|_{0\leq s\leq k}\}$, find a linear minimum mean square error estimator $\hat{x}(k|k)$ of $x(k)$.

## III. Optimal Estimation

In this section, we shall present an analytical solution to the above optimal estimation by reorganizing the observation sequences and applying the innovation analysis method. For the convenience of latter discussion, we denote $k_1 = k - d$.

### A. Re-organized Observations

Define a new observation sequence

$$\bar{y}_d(s) \triangleq \begin{bmatrix} y_0(s) \\ y_d(s+d) \end{bmatrix}, 0 \leq s \leq k_1, \tag{6}$$
$$\bar{y}_0(s) \triangleq y_0(s), k_1 < s \leq k, \tag{7}$$

where $y_0(s)$ and $y_d(s+d)$ are as in (3) and (4). Then the re-organized observations $\bar{y}_0(s)$ and $\bar{y}_d(s)$ satisfy the following difference equations:

$$\bar{y}_d(s) = \begin{bmatrix} \gamma_{s,0}H(s) \\ \gamma_{s+d,d}H(s+d) \end{bmatrix} x(s) + \begin{bmatrix} \gamma_{s,0}v(s) \\ \gamma_{s+d,d}v(s) \end{bmatrix}$$
$$0 \leq s \leq k_1, \tag{8}$$
$$\bar{y}_0(s) = \gamma_{s,0}H(s)x(s) + \gamma_{s,0}v(s), \quad k_1 < s \leq k. \tag{9}$$

Seen from (8), $\bar{y}_d(s)$ is composed of different observations associated with the same state $x(s)$. Obviously, there no longer exist delays in the observation equations (8) and (9). Moreover, it is apparent that the following lemma is true.

*Lemma 1:* For the given time instant $k$, the linear space generated by $\{y(s)|_{0\leq s\leq k}\}$ is equivalent to the linear space of

$$\mathcal{L}\{\bar{y}_d(s)|_{0\leq s\leq k_1}; \bar{y}_0(s)|_{k_1<s\leq k}\}. \tag{10}$$

*Proof:* This follows the same arguments as in [4], and is omitted here. ■

### B. Re-organized Markov Chains

The purpose of this subsection is to reconstruct a set of new Markov chains associated with the re-organized observations (6) and (7). For any given time $k$, the state of the jumping process at time $s$ $(0 \le s \le k)$ is determined by the vector of the binary variables $\gamma_{s,0}$ and $\gamma_{s,d}$. Let

$$\begin{aligned} \bar{\theta}(s) &= [\gamma_{s,0}\ \gamma_{s+d,d}]^T,\ 0 \le s \le k_1, && (11)\\ \theta(s) &= \gamma_{s,0},\quad k_1 < s \le k. && (12) \end{aligned}$$

For $0 \le s \le k_1$, $\bar{\theta}(s)$ takes a random walk on the finite set

$$\begin{aligned} \bar{S}_v &= \left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \end{bmatrix} \right\} \\ &\triangleq \{\bar{e}_1, \bar{e}_2, \bar{e}_3, \bar{e}_4\}. \end{aligned}$$

Set $\bar{\lambda}_{ij} \triangleq \text{Prob}(\bar{\theta}(s+1) = \bar{e}_j | \bar{\theta}(s) = \bar{e}_i)$, and denote $\bar{\Lambda} \triangleq [(\bar{\lambda}_{ij})]$, then $\bar{\Lambda}$ can be obtained via the transition probability matrices of $\gamma_{s,0}$ and $\gamma_{s+d,d}$, that is

$$\bar{\Lambda} = \begin{bmatrix} p_{00} & p_{0d} \\ p_{d0} & p_{dd} \end{bmatrix} \otimes \begin{bmatrix} p_{dd} & p_{d0} \\ p_{0d} & p_{00} \end{bmatrix}, \qquad (13)$$

where $\bar{\Lambda}$ is a matrix all of whose entries are non-negative and each of whose rows sum to 1. Hence $\bar{\Lambda}$ is a transition probability matrix. Alternatively, we define $\bar{\pi}_i(s) \triangleq \text{Prob}(\bar{\theta}(s) = \bar{e}_i)$, and further denote $\bar{\pi}(s) \triangleq [\bar{\pi}_1(s) \cdots \bar{\pi}_4(s)]^T$, then the expression of $\bar{\pi}(s)$ is given by

$$\bar{\pi}(s) = \begin{bmatrix} p_0(s) \\ p_d(s) \end{bmatrix} \otimes \begin{bmatrix} p_d(s+d) \\ p_0(s+d) \end{bmatrix}, \qquad (14)$$

where the initial value

$$\bar{\pi}(0) = [p_0(0)p_d(d),\ p_0(0)p_0(d),\ p_d(0)p_d(d),\ p_d(0)p_0(d)]^T.$$

Recalling that $p_{s+1} = P^T p_s$, and in view of (13) and (14), we show that

$$\bar{\pi}(s+1) = \bar{\Lambda}^T \bar{\pi}(s). \qquad (15)$$

From (13)-(15), we know that the distribution of $\bar{\theta}(s)$ depends only where it was at the time it jumped and not on where it was in the past. Thus $\bar{\theta}(s)$ is indeed a Markov chain, while (13)-(15) give a complete description of its Markovian characteristics.

Next, for $k_1 < s \le k$, we can also show that $\theta(s)$ is a Markov chain with finite state $S_v = \{1, 0\} \triangleq \{e_1, e_2\}$ and transition probability matrix

$$\Lambda \triangleq \begin{bmatrix} \lambda_{11} & \lambda_{12} \\ \lambda_{21} & \lambda_{22} \end{bmatrix} = \begin{bmatrix} p_{00} & p_{0d} \\ p_{d0} & p_{dd} \end{bmatrix}, \qquad (16)$$

where $\lambda_{ij} \triangleq \text{Prob}(\theta(s+1) = e_j | \theta(s) = e_i), i, j = 1, 2$. Let the probability distribution be $\pi(s) \triangleq [\pi_1(s)\ \pi_2(s)]^T$, which is determined by $\pi_1(s) = p_0(s)$ and $\pi_2(s) = p_d(s)$. We have

$$\pi(s+1) = \Lambda^T \pi(s), k_1 < s \le k. \qquad (17)$$

We define $\bar{\theta}(s)$ and $\theta(s)$ as the re-organized Markov chains. With these preparations, it is easy to complete the description of (8) and (9), which are given by

$$\begin{aligned} \bar{y}_d(s) &= \bar{H}_{\bar{\theta}(s)}(s)x(s) + \bar{v}_{\bar{\theta}(s)}(s), 0 \le s \le k_1, && (18)\\ \bar{y}_0(s) &= H_{\theta(s)}(s)x(s) + v_{\theta(s)}(s), k_1 < s \le k, && (19) \end{aligned}$$

where

$$\begin{aligned} \bar{H}_{\bar{\theta}(s)}(s) &= \begin{bmatrix} \gamma_{s,0}H(s) \\ \gamma_{s+d,d}H(s+d) \end{bmatrix}, \\ \bar{v}_{\bar{\theta}(s)}(s) &= \begin{bmatrix} \gamma_{s,0}v(s) \\ \gamma_{s+d,d}v(s) \end{bmatrix}, \\ H_{\theta(s)}(s) &= \gamma_{s,0}H(s), \\ v_{\theta(s)}(s) &= \gamma_{s,0}v(s). \end{aligned}$$

It can be seen that

$$\begin{aligned} \bar{H}_{\bar{\theta}(s)} &\in \left\{ \begin{bmatrix} H(s) \\ H(s+d) \end{bmatrix}, \begin{bmatrix} H(s) \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ H(s+d) \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \end{bmatrix} \right\} \\ &\triangleq \{\bar{H}_1(s), \bar{H}_2(s), \bar{H}_3(s), \bar{H}_4(s)\}, \quad 0 \le s \le k_1, \\ H_{\theta(s)} &\in \{H(s), 0\} \triangleq \{H_1(s), H_2(s)\}, \quad k_1 < s \le k. \end{aligned}$$

Meanwhile, $\bar{v}_{\bar{\theta}(s)}(s)$ and $v_{\theta(s)}(s)$ are independent wide sense stationary sequences with zero mean and covariance matrices

$$\begin{aligned} \bar{\mathcal{Q}}_{\bar{v}}(s) &= \text{diag}\{p_0(s)Q_v,\ p_d(s+d)Q_v\}, 0 \le s \le k_1, && (20)\\ \mathcal{Q}_v(s) &= p_0(s)Q_v,\ k_1 < s \le k. && (21) \end{aligned}$$

### C. Re-organized State Variables

It follows from Assumption 3 that the random communication delay $d(k)$ is an unknown variable, and thus the new jumping parameters $\bar{\theta}(s)$ and $\theta(s)$ which represent the same jumping properties as $d(s)$ can not be determined at the time $s$. In this paper, we shall estimate $(x(s), \bar{\theta}(s))$ and $(x(s), \theta(s))$ simultaneously by introducing the new stochastic variables

$$\begin{aligned} z_d(s,i) &= x(s)\mathbf{1}_{\{\bar{\theta}(s)=\bar{e}_i\}}, 0 \le s \le k_1, i = 1, \cdots, 4 && (22)\\ z_0(s,i) &= x(s)\mathbf{1}_{\{\theta(s)=e_i\}}, k_1 < s \le k, i = 1, 2. && (23) \end{aligned}$$

Let

$$\begin{aligned} z_d(s) &= [z_d(s,1)^T\ \cdots\ z_d(s,4)^T]^T \in \mathcal{R}^{4n}, \\ z_0(s) &= [z_0(s,1)^T\ z_0(s,2)^T]^T \in \mathcal{R}^{2n}, \end{aligned}$$

then it is obvious that

$$\begin{aligned} x(s) &= \sum_{i=1}^{4} z_d(s,i), 0 \le s \le k_1, && (24)\\ x(s) &= \sum_{i=1}^{2} z_0(s,i), k_1 < s \le k. && (25) \end{aligned}$$

Given (24) and (25), we will show that the estimator $\hat{x}(t|t)$ can be obtained directly from the estimators of $z_d(s,i)(0 \leq s \leq k_1; i = 1, \cdots, 4)$ and $z_0(s,i)(k_1 < s \leq k; i = 1,2)$. Consider the definition of (22) and (23), the state representation (1) can be reformulated as

$$\begin{aligned} z_d(s+1) &= \bar{M}(s)z_d(s) + \bar{\xi}(s), 0 \leq s \leq k_1, & (26)\\ z_0(s+1) &= M(s)z_0(s) + \xi(s), k_1 < s \leq k. & (27)\end{aligned}$$

where

$$\begin{aligned} \bar{M}(s,j) &= [\overbrace{\mathbf{1}_{\{\bar{\theta}(s+1)=\bar{e}_j\}}\Phi(s) \cdots \mathbf{1}_{\{\bar{\theta}(s+1)=\bar{e}_j\}}\Phi(s)}^{4\ blocks}]\\ \bar{M}(s) &= \left[\begin{array}{ccc} \bar{M}^T(s,1) & \cdots & \bar{M}^T(s,4) \end{array}\right],\\ \bar{\xi}(s) &= \left[\begin{array}{c} \mathbf{1}_{\{\bar{\theta}(s+1)=\bar{e}_1\}}\Gamma(s)w(s) \\ \vdots \\ \mathbf{1}_{\{\bar{\theta}(s+1)=\bar{e}_4\}}\Gamma(s)w(s) \end{array}\right], \end{aligned}$$

and

$$\begin{aligned} M(s,j) &= [\mathbf{1}_{\{\theta(s+1)=e_j\}}\Phi(s)\ \mathbf{1}_{\{\theta(s+1)=e_j\}}\Phi(s)],\\ M(s) &= \left[\begin{array}{cc} M^T(s,1) & M^T(s,2) \end{array}\right],\\ \xi(s) &= \left[\begin{array}{c} \mathbf{1}_{\{\theta(s+1)=e_1\}}\Gamma(s)w(s) \\ \mathbf{1}_{\{\theta(s+1)=e_2\}}\Gamma(s)w(s) \end{array}\right], \end{aligned}$$

Meanwhile, the observation equations (18) and (19) can be reformulated as

$$\begin{aligned} \bar{y}_d(s) &= \bar{\mathcal{H}}(s)z_d(s) + \bar{v}_{\bar{\theta}(s)}(s), 0 \leq s \leq k_1, & (28)\\ \bar{y}_0(s) &= \mathcal{H}(s)z_0(s) + v_{\theta(s)}(s), k_1 < s \leq k, & (29)\end{aligned}$$

where

$$\begin{aligned} \bar{\mathcal{H}}(s) &\triangleq [\bar{H}_1(s)\ \bar{H}_2(s)\ \bar{H}_3(s)\ \bar{H}_4(s)],\\ \mathcal{H}(s) &\triangleq [H_1(s)\ H_2(s)]. \end{aligned}$$

In view of (28) and (29), the new stochastic variable $z_d(s)(0 \leq s \leq k_1)$ and $z_0(s)(k_1 < s \leq k)$ can be estimated directly from the observation sequences $\{\bar{y}_d(\tau)|_{0\leq\tau\leq s}\}$ and $\{\bar{y}_d(\tau)|_{0\leq\tau\leq k_1};\ \bar{y}_0(\tau)|_{k_1<\tau<s}\}$. Also, as in the derivation of the Kalman filter, the new innovation processes need to be defined further.

### *D. Re-organized Innovation Sequences*

According to the re-organized observations (28) and (29), we introduce the following stochastic sequences

$$\begin{aligned} \varepsilon_d(s) &= \bar{y}_d(s) - \hat{\bar{y}}_d(s|s-1), 0 \leq s \leq k_1, & (30)\\ \varepsilon_0(s) &= \bar{y}_0(s) - \hat{\bar{y}}_0(s|s-1), k_1 < s \leq k. & (31)\end{aligned}$$

where $\hat{\bar{y}}_d(s|s-1)$ is the projection of $\bar{y}_d(s)$ onto the linear space of

$$\mathcal{L}\{\bar{y}_d(\tau)|_{0\leq\tau\leq s-1}\}, \quad (32)$$

and $\hat{\bar{y}}_0(s|s-1)$ is the projection of $\bar{y}_0(s)$ onto the linear space of

$$\mathcal{L}\{\bar{y}_d(\tau)|_{0\leq\tau\leq k_1}; \bar{y}_0(\tau)|_{k_1<\tau\leq s-1}\}. \quad (33)$$

We then have the following equations

$$\begin{aligned} \varepsilon_d(s) &= \bar{\mathcal{H}}(s)\tilde{z}_d(s|s-1) + \bar{v}_{\bar{\theta}(s)}(s), 0 \leq s \leq k_1, & (34)\\ \varepsilon_0(s) &= \mathcal{H}(s)\tilde{z}_0(s|s-1) + v_{\theta(s)}(s), k_1 < s \leq k, & (35)\end{aligned}$$

where

$$\begin{aligned} \tilde{z}_d(s|s-1) &= z_d(s) - \hat{z}_d(s|s-1), & (36)\\ \tilde{z}_0(s|s-1) &= z_0(s) - \hat{z}_0(s|s-1), & (37)\end{aligned}$$

with $\hat{z}_d(s|s-1)$ is the projection of $z_d(s)$ onto the linear space of (32), and $\hat{z}_0(s|s-1)$ is the projection of $z_0(s)$ onto the linear space of (33). The following lemma shows that $\varepsilon_i(s)(i = 0, d)$ are in fact the innovation sequences for the new re-organized observations $\bar{y}_d(s)$ and $\bar{y}_0(s)$, respectively.

*Lemma 2:* The elements of the sequence

$$\{\varepsilon_d(s)|_{0\leq s\leq k_1}; \varepsilon_0(s)|_{k_1<s\leq k}\} \quad (38)$$

are mutually uncorrelated, and spans the same linear space (10).

*Proof:* This follows from the same arguments as in *Lemma 2.1* in [4], and thus is omitted here. ■

Based on Lemma 2, the estimation problem proposed in this paper can be further converted into the one of estimating $z_d(s)$ and $z_0(s)$ based on the innovation process (38), which due to its orthogonality, can significantly simply the calculation of the projection.

### *E. Optimal Estimator $\hat{x}(k|k)$*

In order to calculate the covariance matrices of the estimator errors and thus compute the optimal estimator, we first define the following matrices associated with the second moment of the above variables. For $0 \leq s \leq k_1$, we define

$$\begin{aligned} Z_d(s) &\triangleq \mathrm{E}[z_d(s)z_d^T(s)],\\ Z_d(s,i) &\triangleq \mathrm{E}[z_d(s,i)z_d^T(s,i)],\\ \hat{Z}_d(s|s-1) &\triangleq \mathrm{E}[\hat{z}_d(s|s-1)\hat{z}_d^T(s|s-1)],\\ P_d(s|s-1) &\triangleq \mathrm{E}[\tilde{z}_d(s|s-1)\tilde{z}_d^T(s|s-1)], \end{aligned}$$

and for $t_1 < s \leq t$, we denote

$$\begin{aligned} Z_0(s) &\triangleq \mathrm{E}[z_0(s)z_0^T(s)],\\ Z_0(s,i) &\triangleq \mathrm{E}[z_0(s,i)z_0^T(s,i)],\\ \hat{Z}_0(s|s-1) &\triangleq \mathrm{E}[\hat{z}_0(s|s-1)\hat{z}_0^T(s|s-1)],\\ P_0(s|s-1) &\triangleq \mathrm{E}[\tilde{z}_0(s|s-1)\tilde{z}_0^T(s|s-1)]. \end{aligned}$$

Based on the preceding notations, we present the following Lyapunov-equation results which are associated with $z_d(s)$ and $z_0(s)$ respectively, and will be used latter in the derivation of the estimator.

*Lemma 3:* For $0 \leq s \leq k_1$, the covariance matrix of $z_d(s,i)$ satisfies the following difference equation

$$\begin{aligned} Z_d(s+1,j) &= \sum_{i=1}^{4} \bar{\lambda}_{ij}\Phi(s)Z_d(s,i)\Phi^T(s) + \sum_{i=1}^{4}\bar{\lambda}_{ij} \\ &\times \bar{\pi}_i(s)\Gamma(s)Q_w\Gamma^T(s), j=1,\cdots,4 \quad (39) \end{aligned}$$

with the initial value $Z_d(0,i) = \bar{\pi}_i(0)V, i = 1,\cdots,4$.

For $k_1 < s \leq k$, the covariance matrix of $z_0(s,i)$ can be calculated by the following recursive equation

$$\begin{aligned} Z_0(s+1,j) &= \sum_{i=1}^{2} \lambda_{ij}\Phi(s)Z_0(s,i)\Phi^T(s) + \sum_{i=1}^{2}\lambda_{ij} \\ &\times \pi_i(s)\Gamma(s)Q_w\Gamma^T(s),\ j=1,2 \quad (40) \end{aligned}$$

with the given initial value

$$\begin{aligned} Z_0(k_1+1,1) &= Z_d(k_1+1,1) + Z_d(k_1+1,2), \quad (41) \\ Z_0(k_1+1,2) &= Z_d(k_1+1,3) + Z_d(k_1+1,4). \quad (42) \end{aligned}$$

*Proof:* See Appendix A. ■

In the following, we shall present the Riccati difference equations for $P_d(s|s-1)$ and $P_0(s|s-1)$, respectively.

*Theorem 1:* For a given time instant $k$, the covariance matrices $P_d(s|s-1)(0 \leq s \leq k_1)$ and $P_0(s|s-1)(k_1 < s \leq k)$ can be calculated as follows:

- For $0 \leq s \leq k_1$, the matrix $P_d(s|s-1)$ can be derived by the following Riccati difference equation

$$\begin{aligned} P_d(s+1|s) &= \bar{\mathcal{A}}(s)P_d(s|s-1)\bar{\mathcal{A}}^T(s) - \bar{\mathcal{A}}(s)K_d(s) \\ &\times \bar{\mathcal{H}}(s)P_d(s|s-1)\bar{\mathcal{A}}^T(s) + \mathcal{Z}_d(s) \\ &+ \text{diag}\{\sum_{i=1}^{4}\bar{\lambda}_{ij}\bar{\pi}_i(s)\Gamma(s)Q_w\Gamma^T(s)\}, \end{aligned} \quad (43)$$

  where the initial value $P_d(0|-1) = \text{diag}\{\bar{\pi}_i(0)V\}$, and

$$\begin{aligned} \bar{\mathcal{A}}(s) &= \bar{\Lambda}^T \otimes \Phi(s), \quad (44) \\ K_d(s) &= P_d(s|s-1)\bar{\mathcal{H}}^T(s)(\bar{\mathcal{H}}(s)P_d(s|s-1)\bar{\mathcal{H}}^T(s) \\ &+ \bar{\mathcal{Q}}_{\bar{v}}(s))^{-1}, \quad (45) \\ \mathcal{Z}_d(s) &= \text{diag}\{\sum_{i=1}^{4}\bar{\lambda}_{ij}\Phi(s)Z_d(s,i)\Phi^T(s)\} - \bar{\mathcal{A}}(s) \\ &\times \text{diag}\{Z_d(s,i)\}\bar{\mathcal{A}}^T(s). \quad (46) \end{aligned}$$

- For $k_1 < s \leq k$, the matrix $P_0(s|s-1)$ can be calculated by the difference equation as

$$\begin{aligned} P_0(s+1|s) &= \mathcal{A}(s)P_0(s|s-1)\mathcal{A}^T(s) - \mathcal{A}(s)K_0(s) \\ &\times \mathcal{H}(s)P_0(s|s-1)\mathcal{A}^T(s) + \mathcal{Z}_0(s) \\ &+ \text{diag}\{\sum_{i=1}^{2}\lambda_{ij}\pi_i(s)\Gamma(s)Q_w\Gamma^T(s)\}, \end{aligned} \quad (47)$$

  where the given initial value $P_0(k_1+1,i,j|k_1) = \sum_{l=2i-1}^{2i}\sum_{m=2j-1}^{2j} P_d(k_1+1,l,m|k_1)(i,j=1,2)$, and

$$\begin{aligned} \mathcal{A}(s) &= \Lambda^T \otimes \Phi(s), \quad (48) \\ K_0(s) &= P_0(s|s-1)\mathcal{H}^T(s)(\mathcal{H}(s)P_0(s|s-1)\mathcal{H}^T(s) \\ &+ \mathcal{Q}_v(s))^{-1}, \quad (49) \\ \mathcal{Z}_0(s) &= \text{diag}\{\sum_{i=1}^{2}\lambda_{ij}\Phi(s)Z_0(s,i)\Phi^T(s)\} - \mathcal{A}(s) \\ &\times \text{diag}\{Z_0(s,i)\}\mathcal{A}^T(s). \quad (50) \end{aligned}$$

  *Proof:* See Appendix B. ■

Now, we present the solution to the optimal filtering.

*Theorem 2:* Consider the system (1) and (2), the optimal linear mean square error estimator $\hat{x}(k|k)$ is given by

$$\hat{x}(k|k) = \sum_{i=1}^{2}\hat{z}_0(k,i|k), \quad (51)$$

where the estimator $\hat{z}_0(k,i|k)$ is the $i$th block element of $\hat{z}_0(k \mid k)$ which can be computed as the following steps:

- For $0 \leq s \leq k_1$, $\hat{z}_d(s|s)$ can be calculated by the following difference equation with the initial value of $\hat{z}_d(0|-1) = \text{E}(z(0)) = 0$,

$$\begin{aligned} \hat{z}_d(s|s) &= \hat{z}_d(s|s-1) + K_d(s)(y_d(s) - \bar{\mathcal{H}}(s) \\ &\times \hat{z}_d(s|s-1)), \quad (52) \\ \hat{z}_d(s|s-1) &= \bar{\mathcal{A}}(s-1)\hat{z}_d(s-1|s-1), \quad (53) \end{aligned}$$

  where $K_d(s)$ can be obtained by (43) and (45).
- For $k_1 < s \leq k$, $\hat{z}_0(s|s)$ can be computed by the following recursive equation

$$\begin{aligned} \hat{z}_0(s|s) &= \hat{z}_0(s|s-1) + K_0(s)(y_0(s) - \mathcal{H}(s) \\ &\times \hat{z}_0(s|s-1)), \quad (54) \\ \hat{z}_0(s|s-1) &= \mathcal{A}(s-1)\hat{z}_0(s-1|s-1), \quad (55) \end{aligned}$$

  where the initial value $\hat{z}_0(k_1+1|k_1) = \text{col}\{\hat{z}_d(k_1+1,1|k_1) + \hat{z}_d(k_1+1,2|k_1), \hat{z}_d(k_1+1,3|k_1) + \hat{z}_d(k_1+1,4|k_1)\}$, and $K_0(s)$ can be obtained by (47) and (49).

  *Proof:* The proof of this theorem is omitted here. ■

## IV. Numerical Examples

In this section, we present a simple example to illustrate the previous theoretical results. Consider the dynamic system described by (1) and (2) with the following specifications:

$$\begin{aligned} \Phi(k) &= \begin{bmatrix} 0.7+0.1\cos(2\pi k/T) & 0 \\ 0.1\sin(2\pi k/T) & 0.95 \end{bmatrix}, \\ \Gamma(k) &= \begin{bmatrix} 1 \\ 0.5 \end{bmatrix}, H(k) = \begin{bmatrix} 2 & 2 \end{bmatrix}, \end{aligned}$$

where $T$ is the time horizon and set to 200 in this simulation. $d(k)$ is the random jump delay. To present the efficiency of the proposed results, we shall consider different $d(k)'s$ with different transition probabilities. For simplicity, we consider three different cases in this simulation, see the following table.

| Case | $p_{00}$ | $p_{dd}$ | $d(k) \in \{0, d\}$ |
|---|---|---|---|
| 1 | 0.975 | 0.05 | $d = 4$ |
| 2 | 0.85 | 0.3 | $d = 4$ |
| 3 | 0.95 | 0.5 | $d = 8$ |

The initial condition are as follows: $x_{1,2}(0) \sim N(0,1)$, $p_0(0) = P(d(0) = 0) = 0.5$, $p_d(0) = P(d(0) = d) = 0.5$. In the actual system we use $x(0) = [1,1]'$ for the simulation to generate $z_0(k)$ and $z_d(k)$. Furthermore, $\{w(k)\}$ and $\{v(k)\}$ are mutually independent zero-mean white noise sequences with covariance matrices $Q_w = 1$ and $Q_v = 1$, respectively. Fig. 1 and Fig. 2 show the estimation error variances of $x_1(k)$ and $x_2(k)$ under the three different cases. It can be seen from the simulation results that the obtained linear estimator for systems with Markov time delays are tracking well to the real state value and the proposed method is efficient to different delays and different transition probabilities, that is, the estimation scheme proposed in this paper produces good performance.

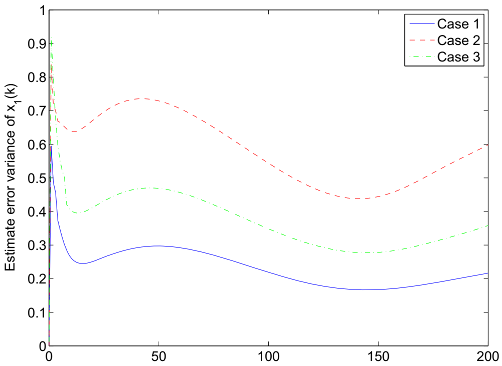


Fig. 1. Estimation error variance of $x_1(k)$ under the three different cases

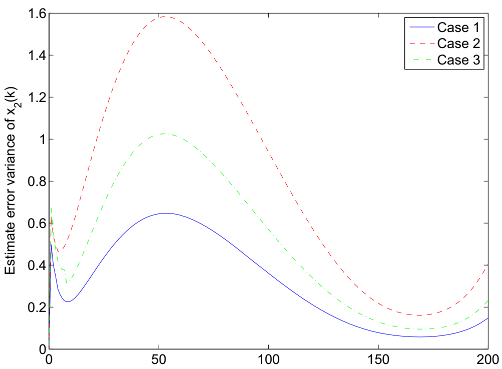


Fig. 2. Estimation error variance of $x_2(k)$ under the three different cases

## V. Conclusion

This paper has addressed the linear MMSE estimation problem for discrete-time linear systems with random observation delays, while the delays are modeled as a finite-state Markov Chain. A complete analytical solution has been obtained by solving two Riccati difference equations. The key to our development in the estimation for the systems with random jump delay is the reorganization of the Markov chains, which transforms the problem into one with two-channel delay-free measurements. It is worth highlighting that the results of the presented paper are applicable to the optimal estimation for systems with multiple random jump delays.

## Appendix A Proof of Lemma 3

For $0 \leq s \leq k_1$, in view of (1) and using the definition of $z_d(s+1,j)$, we get that

$$\begin{aligned} Z_d(s+1,j) &= \mathrm{E}[(\Phi(s)x(s) + \Gamma(s)w(s))(\Phi(s)x(s) \\ &+ \Gamma(s)w(s))^T \mathbf{1}_{\{\bar{\theta}(s+1)=\bar{e}_j\}}] \\ &= \sum_{i=1}^{4} \bar{\lambda}_{ij}\Phi(s)Z_d(s,i)\Phi^T(s) + \sum_{i=1}^{4} \bar{\lambda}_{ij} \\ &\times \bar{\pi}_i(s)\Gamma(s)Q_w\Gamma^T(s). \end{aligned}$$

Recalling that $\mathrm{E}[x(0)x^T(0)] = V$, we get the initial covariance matrix $Z_d(0,i) = \pi_i(0)V, i = 1,2,\cdots,4$.

Similarly, from the definition of $z_0(s+1,j)(j = 1,2)$ and system (1), we get the recursive difference equation (40) with the initial value given by (41) and (42). The proof is accomplished.

## Appendix B Proof of Theorem 1

The proof is divided into two stages. We start with $\mathrm{E}[x(0)] = 0$, namely, $\mathrm{E}[z_d(0,i)] = 0 (i = 1,\cdots,4)$, then all random variables $z_d(s), z_0(s), \varepsilon_d(s)$ and $\varepsilon_0(s)$ have zero expectation.

Firstly, for $0 \leq s \leq k_1$, the optimal linear estimate $\hat{z}_d(s|s)$ can be written as a linear operation on the innovations

$$\hat{z}_d(s|s) = \sum_{l=0}^{s} F(s,l)\varepsilon_d(l), \tag{56}$$

where $F(s,l)$ is an $4n \times m$ matrix, such that $\hat{z}_d(s|s)$ as in (56) satisfies the orthogonal property [18]

$$(z_d(s) - \hat{z}_d(s|s)) \perp \{\varepsilon_d(\tau)|_{0\leq\tau\leq s}\}. \tag{57}$$

Use of (57) with $\hat{z}_d(s|s)$ as in (56) yields

$$F(s,l) = \mathrm{E}[z_d(s)\varepsilon_d^T(l)]\bar{Q}_{\varepsilon_d}^{-1}(l), \tag{58}$$

where $\bar{Q}_{\varepsilon_d}(l)$ is the covariance matrix of $\varepsilon_d(l)$. Then (56) can be written as

$$\hat{z}_d(s|s) = \hat{z}_d(s|s-1) + \mathrm{E}[z_d(s)\varepsilon_d^T(s)]\bar{Q}_{\varepsilon_d}^{-1}(s)\varepsilon_d(s). \tag{59}$$

It follows from (34) that

$$\mathrm{E}[z_d(s)\varepsilon_d^T(s)] = P_d(s|s-1)\bar{\mathcal{H}}^T(s) + \mathrm{E}[z_d(s)\bar{v}_{\bar{\theta}(s)}^T(s)]. \quad (60)$$

From Assumption 4, we know that $\mathrm{E}[z_d(s)\bar{v}_{\bar{\theta}(s)}^T(s)] = 0$. Thus,

$$\mathrm{E}[z_d(s)\varepsilon_d^T(s)] = P_d(s|s-1)\bar{\mathcal{H}}^T(s). \quad (61)$$

Meanwhile, it follows from (34) that

$$\bar{\mathcal{Q}}_{\varepsilon_d}(s) = \bar{\mathcal{H}}(s)P_d(s|s-1)\bar{\mathcal{H}}^T(s) + \bar{\mathcal{Q}}_{\bar{v}}(s). \quad (62)$$

Then using of (61) and (62), we obtain the recursive equation

$$\begin{aligned}\hat{z}_d(s|s) &= \hat{z}_d(s|s-1) \\ &\quad +K_d(s)[\bar{\mathcal{H}}(s)\tilde{z}_d(s|s-1) + \bar{v}_{\bar{\theta}(s)}(s)] \quad (63)\end{aligned}$$

where

$$\begin{aligned}K_d(s) &= P_d(s|s-1)\bar{\mathcal{H}}^T(s) \\ &\quad \times[\bar{\mathcal{H}}(s)P_d(s|s-1)\bar{\mathcal{H}}^T(s) + \bar{\mathcal{Q}}_{\bar{v}}(s)]^{-1}.\end{aligned}$$

Similarly, the one-step ahead prediction $\hat{z}_d(s+1|s)$ can be expressed as

$$\hat{z}_d(s+1|s) = \bar{\mathcal{A}}(s)\hat{z}_d(s|s), \quad (64)$$

where the initial value $\hat{z}_d(0|-1) = 0$, and $\bar{\mathcal{A}}(s) = \bar{\Lambda}^T \otimes \Phi(s)$.

Then from (63), and after noticing that $\hat{z}_d(s|s-1), \tilde{z}_d(s|s-1)$ and $v_{\bar{\theta}(s)}(s)$ are orthogonal among themselves, we get that

$$\begin{aligned}&\mathrm{E}[\hat{z}_d(s|s)\hat{z}_d^T(s|s)] \\ &= \hat{Z}_d(s|s-1) + K_d(s)\bar{\mathcal{H}}(s)P_d(s|s-1), \quad (65)\end{aligned}$$

and in view of (64), we get that

$$\hat{Z}_d(s+1|s) = \bar{\mathcal{A}}(s)\hat{Z}_d(s|s)\bar{\mathcal{A}}^T(s), \hat{Z}_d(0|-1) = 0. \quad (66)$$

Alternatively, with regard to Lemma 3, the equation (39) can be reformulated as

$$\begin{aligned}Z_d(s+1) &= \bar{\mathcal{A}}(s)Z_d(s)\bar{\mathcal{A}}^T(s) + \mathcal{Z}_d(s) \\ &\quad +\mathrm{diag}\{\sum_{i=1}^{4}\bar{\lambda}_{ij}\bar{\pi}_i(s)\Gamma(s)Q_w\Gamma^T(s)\}, (67)\end{aligned}$$

where

$$\begin{aligned}\mathcal{Z}_d(s) &= \mathrm{diag}\{\sum_{i=1}^{4}\bar{\lambda}_{ij}\Phi(s)Z_d(s,i)\Phi^T(s)\} \\ &\quad -\bar{\mathcal{A}}(s)Z_d(s)\bar{\mathcal{A}}^T(s), \\ Z_d(0) &= \mathrm{diag}\{\bar{\pi}_i(0)V\},\ i = 1, \cdots, 4.\end{aligned}$$

Consider (65)-(67), and note that $P_d(s|s-1) = Z_d(s) - \hat{Z}_d(s|s-1)$, we get (43) immediately.

Secondly, for $k_1 < s \leq k$, following the same step as above, we can obtain the Riccati equation for $P_0(s|s-1)$. The desired result is obtained.

## References


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

[2] L. Chisci and E. Mosca, "Polynomial equations for the linear MMSE state estimation," *IEEE Trans. Autom. Control,* vol. 37, pp. 623-626, May 1992.

[3] L. Xie, C. Du, C. Zhang and Y. C. Soh, "$H_\infty$ deconvolution filtering of 2-D digital systems," *IEEE Trans. Signal Processing,* vol.50, no. 9, pp. 2319-2332, 2002.

[4] H. Zhang, L. Xie, D. Zhang and Y. C. Soh, "A reorganized innovation approach to linear estimation," *IEEE Trans. Autom. Control,* vol. 49, no. 10, pp. 1810-1814, 2004.

[5] W. Zhang and L. Yu, "Modeling and control of networked control systems with both network-induced delay and packet-dropout," *Automatica,*vol. 44, pp. 3206-3210, 2008.

[6] C. Lin, Z. Wang and F. Yang, "Observer-based networked control for continuous-time systems with random sensor delays," *Automatica,* vol. 45, pp. 578-584, 2009.

[7] A. S. Matveev and A. V. Savkin, "The problem of state estimation via asynchronous communication channels with irregular transmission times," *IEEE Trans. Automat. Contr.,* vol. 48, no. 4, pp. 670-676, 2003.

[8] N. E. Nahi, "Optimal recursive estimation with uncertain observation," *IEEE Trans. Inf. Theory,* vol. 15, no. 4, pp. 457-462, 1969.

[9] M. T. Hadidi and C. S. Schwartz, "Linear recursive state estimators under uncertain observations," *IEEE Trans. Automat. Contr.,* vol. 24, no. 6, pp. 944-948, 1979.

[10] B. Sinopoli, L. Schenato, M. Franceschetti, K. Poolla, M. I. Jordan, and S. S. Sastry, "Kalman filtering with intermittent observations," *IEEE Trans. Automat. Contr.,* vol. 49, no. 9, pp. 1453-1463, Sep. 2004.

[11] K. Plarre and F. Bullo, "On Kalman filtering for detectable systems with intermittent observations," *IEEE Trans. Automat. Contr.,* vol. 54, no. 2, pp. 386-390, 2009.

[12] E. Yaz, A. Ray, Linear Unbiased State Estimation Under Randomly Varying Bounded Sensor Delay, *Appl. Math. Letters*, vol. 11, 1998, pp. 27-32.

[13] A. H. Carazo and J. L. Pérez, "Extended and unscented filtering algorithms using one-step randomly delayed observations," *Appl. Math. Comput.,* vol. 190, pp. 1375-1393, 2007.

[14] A. H. Carazo and J. L. Pérez, "Linear and quadratic least-squares estimation using measurements with correlated one-step random delay," *Digital Signal Processing,* vol. 18, pp. 450-464, 2008.

[15] B. Shena, Z. Wang, H. Shu and G. Wei, "$H_\infty$ filtering for nonlinear discrete-time stochastic systems with randomly varying sensor delays," *Automatica,* vol. 45, no. 4, pp. 1032-1037, 2009.

[16] J. Evans and V. Krishnamurthy, "Hidden Markov model state estimation with randomly delayed observations," *IEEE Trans. Automat. Contro.,* vol. 47, no. 8, pp. 2157-2166, 1999.

[17] H. Song, L. Yu and W. Zhang, "$H_\infty$ filtering of network-based systems with random delay," *Signal Processing,* vol. 89, no. 4, pp. 615-622, 2009.

[18] M. H. A. Davis and R. B. Vinter, *Stochastic Modeling and Control*. New York: Chapman and Hall, 1985.