Optimal Estimation for Discrete-Time Linear System with Communication Constraints and Measurement Quantization

Hongru Ren, Renquan Lu, Junlin Xiong, and Yong Xu

Abstract—**This paper focuses on the linear minimum mean square estimator for a networked discrete time-varying linear system subject to data quantification and communication constraints. The communication limitation is that only one transmission node can get access to the shared communication channel at each time step, and that different transmission nodes in the networked systems are scheduled to transmit information according to a Markov protocol. Then the remote estimator completes the estimation with only partially available observations, which are quantified. Suppose that the Markov chain is unknown to the remote estimator. By using orthogonal projection principle and innovation analysis method, a Kalman type filter is designed in a recurrence form. It is shown that estimation performance depends on the transition probability matrix of the Markov chain, quantization error, and the shared channel weighting parameter. Finally, an illustrative example is given to show the effectiveness of the proposed method.**

Index Terms—**Markov jump systems, networked systems, optimal estimation, Riccati equations.**

I. Introduction

RECENTLY, in many industrial and civilian applications, a lot of interesting research has been motivated by making use of communication networks to connect different spatially distributed sensors and signal estimators to build networked control systems (NCSs) [1]–[12]. Networked systems have technical advantages such as high flexibility, reduced wiring, and resources sharing. However, some new issues are brought to be addressed, e.g., packet loss, time delay and consensus, data quantization, and communication constraint. For example, consensus problems for edges of networked systems were studied in [13] and [14]. One of the usual constraints is known as medium access constraint, under which network nodes are not permitted to access the network simultaneously [15], [16]. It has been handled by the so-called time-multiplexing technique which has been worked out in various kinds of Fieldbus and controller area network-based networks [17], [18]. In time-multiplexing mechanism, time is divided into different slots, and during one slot only one transmission node is permitted to access the network according to a special communication protocol which may be deterministic or stochastic [15]. Moreover, in an ordinary way, multioutputs of a networked system are supposed to be measured via multisensors, stacked into multipackets and then transmitted via multinodes. There are two main reasons for using such a multipacket transmission strategy. First, what we consider is bandwidth limitation and packet size constraint. Second, and more significant, sensors in NCSs are always distributed over a large-scale physical region, and it is impossible to pack all measurements into one packet [19].

On the other hand, the estimation problem has been an important issue from industrial applications to research areas including optimal control, signal processing, and navigation [20]–[28]. It is well-known that a traditional Kalman filter (KF) acts as an optimal filter for linear systems with exact system model in the least mean square sense. The applications of communication networks, in particular wireless sensor networks and wireless networks, make it possible to use a KF to estimate information of distributed large-scale systems. In addition, estimation performance could be improved by data communication among various nodes connected to networks. Due to the medium access constraint and multipackets transmission, generally, remote estimators do the estimation job with only partial available observations during each sampling interval. To the best of our knowledge, many existing results with respect to communication constraints were concerned from the control design side [29]–[34], while they seldom took the problem of signal estimation into account. However, in [7], [35], and [36], the Kalman filtering problem for NCSs with communication constraints was considered. The channel accessing protocols of sensors were modeled by mutual independent Bernoulli random processes. In [37] and [38], linear minimum variance unbiased estimation was studied for discrete-time systems with uncertain

Manuscript received May 13, 2017; revised October 9, 2017; accepted December 26, 2017. This work was supported in part by the China National Funds for Distinguished Young Scientists under Grant 61425009, in part by the National Natural Science Foundation of China under Grant U1611262/61773357, in part by the Guangdong Province Higher Vocational Colleges and Schools Pearl River Scholar approved in 2015 and the China National 863 Technology Projects under Grant 2015BAF32B03-05, and in part by the Fundamental Research Funds for the Central Universities under Grant 2017FZA5010. This paper was recommended by Associate Editor W. He. *(Corresponding author: Junlin Xiong.)*

H. Ren and J. Xiong are with the Department of Automation, University of Science and Technology of China, Hefei 230026, China (e-mail: rhru@mail.ustc.edu.cn; xiong77@ustc.edu.cn).

R. Lu and Y. Xu are with the Guangdong Key Laboratory of IoT Information Processing, School of Automation, Guangdong University of Technology, Guangzhou 510006, China (e-mail: rqlu@gdut.edu.cn; xuyong809@163.com).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TSMC.2018.2792009

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

parameters in state space models and multiple sensor failures. In [39] and [40], the filtering problem was studied for some discrete-time networked systems with communication constraints. The above-mentioned results have made a lot of achievements in networked estimation systems with communication constraints. However, there are still many critical issues to be solved. For instance, when sensor measurements are inevitably influenced by information from the last sampling time, and the communication constraint of transmission networks are subject to a Markov protocol, the problem of deriving the optimal estimation to the system state under data quantization is not yet solved in the literature. In this paper, an optimal estimation system for such a new type of system model is designed. Filtering problems for Markov jump systems were also studied in some existing literature such as [41] and [42]. Some adaptive filtering algorithms were considered in [43]–[50]. To overcome modeling uncertainty and numerical error, finite-impulse response (FIR) filtering algorithms were studied in [51]–[53].

In this paper, the linear minimum mean square estimation problem for networked discrete time-varying linear system subject to data quantification and communication constraints is investigated, and a solution to the problem is provided. The communication constraint is that only one transmission node can get access to the shared communication channel at each time step, and that different transmission nodes of the networked system are scheduled to transmit information according to a Markov protocol. The remote estimator completes the estimation with only partially available quantized observations. Quantization of the signals in communication networks is also taken into consideration due to limited bandwidth. Suppose that the Markov chain is unknown to the remote estimator. By applying an augmentation transformation, the overall estimation system is remodeled as a discrete-time stochastic system with Markov jump parameters. Many results have been available in linear estimation for discrete time systems with stochastic parameters in the existing literature, such as [54]–[59]. However, the approaches above are not available to the concerned problem. By using orthogonal projection principle and innovation analysis method, we design a Kalman type filter in a recurrence form. The simulation results show that the estimation performance (specified by the trace of the estimation error covariance matrix) depends on the transition probability matrix of Markov chain, quantization error, and the shared channel weighting parameter.

Notation: $\mathbb{R}^n$ denotes the n-dimensional real Euclidean space. $E\{\cdot\}$ is the mathematical expectation. Define $\langle x, y\rangle \triangleq E\{xy^T\} = \langle y, x\rangle^T$, $\|x\|^2 \triangleq \langle x, x\rangle$, where x and y are vector-valued random variables. $x \perp y$ represents orthogonal vectors x and y. $\mathcal{L}\{x_1, x_2, \ldots\}$ is the linear subspace spanned by the vectors $\{x_1, x_2, \ldots\}$. $\mathbf{P}(A)$ denotes occurrence probability of the event A. I_m represents the identity matrix of size $m \times m$. 0 is zero matrix with appropriate dimension. For Banach spaces $\mathbb{X}$ and $\mathbb{Y}$, $\mathbb{B}(\mathbb{X}, \mathbb{Y})$ is set as the Banach space of all bounded linear operators of $\mathbb{X}$ into $\mathbb{Y}$. For convenience, we set $\mathbb{B}(\mathbb{X}) := \mathbb{B}(\mathbb{X}, \mathbb{X})$. For a matrix $M \in \mathbb{B}(\mathbb{R}^n)$, $M \geq 0$ ($M > 0$, respectively) represents that M is positive semidefinite (positive definite). Define $\mathbb{H}^{n,m} = \{\Xi = (\Xi_1, \ldots, \Xi_{\bar{m}});\ \Xi_i \in \mathbb{B}(\mathbb{R}^n, \mathbb{R}^m),\ i = 1, \ldots, \bar{m}\}$, $\mathbb{H}^n = \mathbb{H}^{n,n}$. For a matrix $M \in \mathbb{B}(\mathbb{R}^n, \mathbb{R}^m)$, $\mathrm{diag}\{M\} \in \mathbb{B}(\mathbb{R}^{\bar{m}n}, \mathbb{R}^{\bar{m}m})$ stands for a block-diagonal matrix formed by M in the diagonal and zero elsewhere. For $\Xi = (\Xi_1, \ldots, \Xi_{\bar{m}}) \in \mathbb{H}^{n,m}$, we define $\mathrm{diag}\{\Xi_i\} \in \mathbb{B}(\mathbb{R}^{\bar{m}n}, \mathbb{R}^{\bar{m}m})$ as the diagonal matrix built by Ξ_i in the diagonal and zero elsewhere.

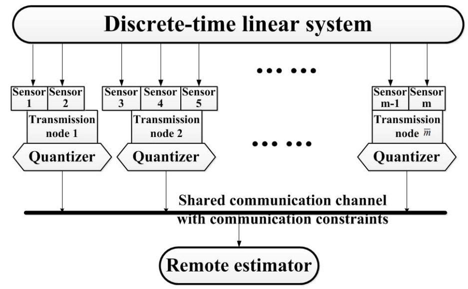


Fig. 1. Networked estimation system with communication constraints and quantization.

II. Problem Formulation and Preliminaries

Consider the following discrete-time stochastic system:

$$\begin{cases} x(k+1) = A(k)x(k) + B(k)w(k) \\ y(k) = C(k)x(k) + H(k-1)x(k-1) + v(k) \end{cases} \tag{1}$$

where $x(k) \in \mathbb{R}^n$ is the system state, $y(k) \in \mathbb{R}^m$ denotes the system output. $\{w(k)\} \in \mathbb{R}^r$ and $\{v(k)\} \in \mathbb{R}^m$ are vector-valued Gaussian white noises. The second term $H(k-1)x(k-1)$ on right hand side of the second equation in (1) indicates that measurement output is inevitably influenced by state of the last time step. $\{w(k), v(k)\}$ are uncorrelated white noises with zero means and variance matrices $Q(k) \geq 0$ and $R(k) > 0$, respectively, i.e., $E\{w(k)\} = E\{v(k)\} = 0, E\{w(i)w^T(j)\} = Q(i)\delta_{ij}, E\{v(i)v^T(j)\} = R(i)\delta_{ij}$, where $\delta_{ii} = 1$ and $\delta_{ij} = 0 (i \neq j)$. $A(k), B(k), C(k), H(k-1)$ are time-varying matrices with appropriate dimensions. Let $z(0) = [x^T(0) \quad x^T(-1)]^T$ denote the initial state vector and assume that it is uncorrelated with all $\{w(k), v(k)\}$, and satisfies $E\{z(0)\} = z_0, E\{z(0)z^T(0)\} = \Pi_{z,0}$.

Suppose that there are m sensors spatially distributed in a large-scale physical area which is shown in Fig. 1. They cannot be integrated into one transmission node. Consequently, the m measurements $y_i(k), i = 1, \ldots, m$, are collected by $\bar{m}$ transmission nodes, then transmitted to a remote estimator through a shared communication channel, where $\bar{m} \leq m$. Denote the set of sensors which transmit their measurements through transmission node j by $N_j, j \in \{1, \ldots, \bar{m}\}$. For example, $N_2 = \{3, 4\}$ means that sensors 3 and 4 transit their measurements through transmission node 2.

Because of the communication constraints, only one transmission node is permitted to gain access to the shared channel and communicate with the remote estimator at each sampling time step k. Suppose that communication protocol of the $\bar{m}$ nodes is subject to a Markov protocol. In other words, the remote estimator can choose to get access to just only one transmission node at each epoch k according to a discrete time Markov chain $\{\theta(k)\}$, which takes values in a discrete space $\{1, \ldots, \bar{m}\}$ and has transition probability matrix $P_{\bar{m}} = [p_{ij}]$, where $p_{ij} \geq 0$ denotes the probability of $\theta(k+1) = j$ conditioned on $\theta(k) = i$ and $\sum_{j=1}^{\bar{m}} p_{ij} = 1$, $(i \in \{1, 2, \ldots, \bar{m}\})$. We set $\pi_i(k) := \mathbf{P}(\theta(k) = i)$.

Remark 1: It is supposed that the actual Markov chain $\{\theta(k)\}$ is not known to the remote estimator. The only information the estimator can obtain is that the initial probability distribution $\{\pi_i(0)\}$ and transition probability matrix $P_{\bar{m}}(k)$ of the Markov chain.

Before transmission, due to the limitation of channel bandwidth, the information needs to be quantified. In this paper, an uniform quantizer installed on the shared channel is taken into account. It is assumed that the overall quantizer range is $[-M, M]$ with $M > 0$. The quantizer level length U is defined as $U = 2M/(2^b - 1)$ where b is the number of bits allowed for channel transmission. The uniform quantizer output is then given by

$$s(k) = \mathcal{Q}(y(k)) = y(k) + q(k) \tag{2}$$

where $s(k)$ is the quantized signal, $q(k) \in R^m$ is the quantization error process and is assumed to be an additive uniform distributed white noise of which each element is uniformly distributed in $[-0.5U, 0.5U]$. Apparently, the variance of this quantization error process is $\Lambda(k) = [(U^2)/(12)]I$. The above description is based on the assumption that the signals before quantization do not exceed the overall quantizer range $[-M, M]$. It is assumed that the quantization error process $\{q(k)\}$ is independent of $\{w(k), v(k), z(0)\}$ with zero mean and variance $\Lambda(k)$, i.e., $E\{q(k)\} = 0, E\{q(i)q^T(j)\} = \Lambda(i)\delta_{ij}$. In addition, the Markov chain $\{\theta(k)\}$ is independent of $\{w(k), v(k), z(0), q(k)\}$.

We first define

$$y(k) = \left[y_1(k), \ldots, y_m(k)\right]^T$$
$$s(k) = [s_1(k), \ldots, s_m(k)]^T$$
$$q(k) = \left[q_1(k), \ldots, q_m(k)\right]^T$$

and then represent the input of the remote estimator by

$$\eta(k) = [\eta_1(k), \ldots, \eta_m(k)]^T.$$

It can be observed that $\eta(k) \neq y(k)$ and only some of elements in $\eta(k)$ can be updated during each sampling period as a result of the communication constraints. During each time step k, only one transmission node j with its measurements $\{y_i(k), i \in N_j\}$ is chosen to occupy the communication channel on the basis of the Markov chain $\{\theta(k)\}$. Then, a quantization process acts on those measurement outputs before transmission. Therefore, for $i = 1, \ldots, m$, one has

$$\begin{cases} \eta_i(k) = s_i(k) = y_i(k) + q_i(k), & \text{if } \ i \in N_{\theta(k)} \\ \eta_i(k) = \gamma(k)\eta_i(k-1), & \text{otherwise} \end{cases} \tag{3}$$

where $\{\gamma(k)\}$ is a scalar sequence depending on the Markov chain $\{\theta(k)\}$, namely, $\gamma(k) = \gamma_{\theta(k)}$ and satisfies $0 \leq \gamma_{\theta(k)} \leq 1$. The second equation of (3) implies that the estimator input keeps a weighted value of the previous measurement when the current measured value is not available. Each transmission node j has its matching γ_j, where γ_j is the weighting parameter.

For $i = 1, \ldots, \bar{m}$, denote

$$\begin{cases} \Gamma_i = \sum_{j \in N_i} \text{diag}\{\delta[j-1], \ \ldots, \ \delta[j-m]\} \\ \Upsilon_i = \text{diag}\{\gamma_i, \ \ldots, \ \gamma_i\}_{m \times m} = \gamma_i I_m \end{cases} \tag{4}$$

where $\delta \in \{0, 1\}$ is the Kronecker delta function

$$\delta[n] = \begin{cases} 0, & \text{if } n \neq 0 \\ 1, & \text{if } n = 0. \end{cases} \tag{5}$$

Then it follows from (1)–(5) that the remote estimator input $\eta(k)$ can be expressed as:

$$\eta(k) = \Gamma_{\theta(k)}(y(k) + q(k)) + \Upsilon_{\theta(k)}(I_m - \Gamma_{\theta(k)})\eta(k-1). \tag{6}$$

Denote

$$z(k) = \left[x^T(k) \quad x^T(k-1)\right]^T$$
$$\xi(k) = \left[z^T(k) \quad \eta^T(k-1)\right]^T$$
$$\beta(k) = v(k) + q(k).$$

The estimation system with communication constraints and quantification can be reformulated as the following system model with Markov jump parameters:

$$\begin{cases} \xi(k+1) = \tilde{A}_{\theta(k)}(k)\xi(k) + \tilde{B}_{\theta(k)}(k)\rho(k) \\ \eta(k) = \tilde{C}_{\theta(k)}(k)\xi(k) + \Gamma_{\theta(k)}\beta(k) \end{cases} \tag{7}$$

where

$$\tilde{A}_{\theta(k)}(k) = \begin{bmatrix} A(k) & 0 & 0 \\ 0 & A(k-1) & 0 \\ \Gamma_{\theta(k)}C(k) & \Gamma_{\theta(k)}H(k-1) & \Upsilon_{\theta(k)}(I_m - \Gamma_{\theta(k)}) \end{bmatrix}$$
$$\tilde{B}_{\theta(k)}(k) = \begin{bmatrix} B(k) & 0 & 0 \\ 0 & B(k-1) & 0 \\ 0 & 0 & \Gamma_{\theta(k)} \end{bmatrix}$$
$$\tilde{C}_{\theta(k)}(k) = \begin{bmatrix} \Gamma_{\theta(k)}C(k) & \Gamma_{\theta(k)}H(k-1) & \Upsilon_{\theta(k)}(I_m - \Gamma_{\theta(k)}) \end{bmatrix}$$
$$\rho(k) = \left[w^T(k) \quad w^T(k-1) \quad \beta^T(k)\right]^T$$

with $\xi_0 = E\{\xi(0)\} = [z_0^T, 0]^T$, since set $\eta(-1) = 0$.

III. LMMSE Filter Design

First, some variables and matrices are defined in order to facilitate the subsequent derivation

$$\bar{\Gamma}(k) = E\left\{\Gamma_{\theta(k)}\right\} = \sum_{j=1}^{\bar{m}} \pi_j(k)\Gamma_j$$
$$\bar{A}(k) = E\left\{\tilde{A}_{\theta(k)}(k)\right\}, \ \bar{B}(k) = E\left\{\tilde{B}_{\theta(k)}(k)\right\}.$$

Therefore, one has that

$$\bar{B}(k) = \begin{bmatrix} B(k) & 0 & 0 \\ 0 & B(k-1) & 0 \\ 0 & 0 & \bar{\Gamma}(k) \end{bmatrix}. \tag{8}$$

By independence hypothesis, we have that

$$\begin{aligned} &\langle\beta(i), \beta(j)\rangle \\ &\quad = \langle v(i) + q(i), v(j) + q(j)\rangle \\ &\quad = \langle v(i), v(j)\rangle + \langle v(i), q(j)\rangle + \langle q(i), v(j)\rangle + \langle q(i), q(j)\rangle \\ &\quad = R(i)\delta_{ij} + 0 + 0 + \Lambda(i)\delta_{ij} \\ &\quad = (R(i) + \Lambda(i))\delta_{ij}\langle\rho(i), \rho(j)\rangle \end{aligned} \tag{9}$$

$$= \left\langle \begin{bmatrix} w(i) \\ w(i-1) \\ \beta(i) \end{bmatrix}, \begin{bmatrix} w(j) \\ w(j-1) \\ \beta(j) \end{bmatrix} \right\rangle$$
$$= \begin{bmatrix} \langle w(i), w(j)\rangle & \langle w(i), w(j-1)\rangle & \langle w(i), \beta(j)\rangle \\ \langle w(i-1), w(j)\rangle & \langle w(i-1), w(j-1)\rangle & \langle w(i-1), \beta(j)\rangle \\ \langle\beta(i), w(j)\rangle & \langle\beta(i), w(j-1)\rangle & \langle\beta(i), \beta(j)\rangle \end{bmatrix} \tag{10}$$

$$
\begin{aligned}
&= \Delta_0(i)\delta_{ij} + \Delta_{+1}(i)\delta_{i,j+1} + \Delta_{+1}^T(j)\delta_{i,j-1}\\
&\langle \beta(i), \rho(j)\rangle\\
&= \langle v(i) + q(i), \rho(j)\rangle\\
&= \langle v(i), \rho(j)\rangle + \langle q(i), \rho(j)\rangle\\
&= \left\langle v(i), \begin{bmatrix} w(j) \\ w(j-1) \\ v(j)+q(j)\end{bmatrix}\right\rangle + \left\langle q(i), \begin{bmatrix} w(j) \\ w(j-1) \\ v(j)+q(j)\end{bmatrix}\right\rangle\\
&= \Phi(i)\delta_{ij} \qquad (11)
\end{aligned}
$$

where $\Delta_0(i) = \text{diag}\{Q(i),\ Q(i-1),\ R(i)+\Lambda(i)\}$, $\Delta_{+1}(i) = \begin{bmatrix} 0 & 0 & 0 \\ Q(i-1) & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$, and $\Phi(i) = [0\ 0\ R(i)+\Lambda(i)]$.

A. Uncorrelatedness Properties

The reformulated model (7) implies that $\{\rho(i)\}$ is uncorrelated with all past states and $\{\beta(i)\}$ is uncorrelated with all past and present states, i.e.,

$$
\begin{cases} \langle \rho(i), \xi(j)\rangle = 0, & j \le i-1 \\ \langle \beta(i), \xi(j)\rangle = 0, & j \le i \end{cases}
$$

which we can write as

$$
\rho(i) \perp \xi(j), \quad j \le i-1; \quad \beta(i) \perp \xi(j), \quad j \le i. \qquad (12)
$$

The reason is that, according to (7) and independence hypothesis, $\xi(j)$ depends linearly only upon the random variables $\{\xi(0), \rho(k), k \le j-1\}$ [or briefly, $\xi(j) \in \mathcal{L}\{\xi(0), \rho(k), k \le j-1\}$], and by (10), $\rho(i)$ is uncorrelated with or orthogonal to these random variables, and so is $\beta(i)$. However, for the present state, from (10)–(12), we have

$$
\begin{aligned}
\langle \beta(i), \xi(i)\rangle &= \langle \beta(i), \xi(i-1)\rangle \bar{A}^T(i-1)\\
&\quad + \langle \beta(i), \rho(i-1)\rangle \bar{B}^T(i-1)\\
&= 0 \qquad (13)
\end{aligned}
$$

and

$$
\begin{aligned}
\langle \rho(i), \xi(i)\rangle &= \langle \rho(i), \xi(i-1)\rangle \bar{A}^T(i-1)\\
&\quad + \langle \rho(i), \rho(i-1)\rangle \bar{B}^T(i-1)\\
&= 0 + \Delta_{+1}(i)\bar{B}^T(i-1) = \Delta_{+1}(i)\bar{B}^T(i-1). \qquad (14)
\end{aligned}
$$

For the same reason, we can see that $\{\rho(i), \beta(i)\}$ are orthogonal to past outputs, i.e.,

$$
\rho(i) \perp \eta(j), \quad \beta(i) \perp \eta(j), \quad j \le i-1. \qquad (15)
$$

Next, let us define $\Pi(k) \triangleq \|\xi(k)\|^2$, the state covariance matrix with initial value $\Pi(0) = \Pi_0$.

Then $g(k,i) \triangleq \xi(k)\delta[\theta(k)-i]$, $i = 1, \ldots, \bar{m}$, is introduced to deal with the derivation on Markov chain, and $g(k) \triangleq [g^T(k,1), \ldots, g^T(k,\bar{m})]^T$. From independence hypothesis, we define

$$
\begin{aligned}
G_i(k) &\triangleq \|g(k,i)\|^2 = E\{\xi(k)\xi^T(k)\delta[\theta(k)-i]\}\\
&= \|\xi(k)\|^2 E\{\delta[\theta(k)-i]\} = \pi_i(k)\Pi(k) \qquad (16)\\
G(k) &\triangleq \|g(k)\|^2 = \text{diag}\{G_i(k)\} = \text{diag}\{\pi_i(k)\Pi(k)\} \qquad (17)
\end{aligned}
$$

and set

$$
\zeta(0) = E\{g(0)\} = \begin{bmatrix} \xi_0\pi_1(0) \\ \vdots \\ \xi_0\pi_{\bar{m}}(0)\end{bmatrix}
$$

$$
P(0) = \|g(0) - \zeta(0)\|^2 = G(0) - \zeta(0)\zeta^T(0).
$$

To derive recursive equation for matrices $G_i(k)$, recalling (7), we start with

$$
\begin{aligned}
&g(k+1,j)\\
&= \xi(k+1)\delta[\theta(k+1)-j]\\
&= \left(\tilde{A}_{\theta(k)}(k)\xi(k) + \tilde{B}_{\theta(k)}(k)\rho(k)\right)\delta[\theta(k+1)-j] \qquad (18)\\
&G_j(k+1)\\
&= \|g(k+1,j)\|^2\\
&= \|\tilde{A}_{\theta(k)}(k)\xi(k)\delta[\theta(k+1)-j]\|^2\\
&\quad + \|\tilde{B}_{\theta(k)}(k)\rho(k)\delta[\theta(k+1)-j]\|^2\\
&\quad + \left\langle \tilde{A}_{\theta(k)}(k)\xi(k)\delta[\theta(k+1)-j], \tilde{B}_{\theta(k)}(k)\rho(k)\delta[\theta(k+1)-j]\right\rangle\\
&\quad + \left\langle \tilde{B}_{\theta(k)}(k)\rho(k)\delta[\theta(k+1)-j], \tilde{A}_{\theta(k)}(k)\xi(k)\delta[\theta(k+1)-j]\right\rangle \qquad (19)
\end{aligned}
$$

where

$$
\begin{aligned}
&\|\tilde{A}_{\theta(k)}(k)\xi(k)\delta[\theta(k+1)-j]\|^2\\
&= E\left\{\tilde{A}_{\theta(k)}(k)\xi(k)\xi^T(k)\tilde{A}_{\theta(k)}^T(k)\delta[\theta(k+1)-j]\right\}\\
&= \sum_{i=1}^{\bar{m}} E\left\{\tilde{A}_i(k)\xi(k)\xi^T(k)\tilde{A}_i^T(k)\delta[\theta(k)-i]\delta[\theta(k+1)-j]\right\}\\
&= \sum_{i=1}^{\bar{m}} \tilde{A}_i(k)E\left\{\xi(k)\xi^T(k)\right\}\tilde{A}_i^T(k)E\{\delta[\theta(k)-i]\delta[\theta(k+1)-j]\}\\
&= \sum_{i=1}^{\bar{m}} \tilde{A}_i(k)E\left\{\xi(k)\xi^T(k)\right\}\tilde{A}_i^T(k)\mathbf{P}\{\theta(k)=i, \theta(k+1)=j\}\\
&= \sum_{i=1}^{\bar{m}} \tilde{A}_i(k)E\left\{\xi(k)\xi^T(k)\right\}\tilde{A}_i^T(k)\mathbf{P}\{\theta(k+1)=j|\theta(k)=i\}\\
&\quad \times \mathbf{P}\{\theta(k)=i\}\\
&= \sum_{i=1}^{\bar{m}} p_{ij}\pi_i(k)\tilde{A}_i(k)\Pi(k)\tilde{A}_i^T(k)\\
&= \sum_{i=1}^{\bar{m}} p_{ij}\tilde{A}_i(k)G_i(k)\tilde{A}_i^T. \qquad (20)
\end{aligned}
$$

Similarly, from (11) and (14), we have that

$$
\begin{aligned}
&\|\tilde{B}_{\theta(k)}(k)\rho(k)\delta[\theta(k+1)-j]\|^2\\
&= \sum_{i=1}^{\bar{m}} p_{ij}\pi_i(k)\tilde{B}_i(k)\Delta_0(k)\tilde{B}_i^T(k) \qquad (21)\\
&\left\langle \tilde{A}_{\theta(k)}(k)\xi(k)\delta\big[\theta(k+1)-j\big], \tilde{B}_{\theta(k)}(k)\rho(k)\delta\big[\theta(k+1)-j\big]\right\rangle\\
&= E\left\{\tilde{A}_{\theta(k)}(k)\xi(k)\rho^T(k)\tilde{B}_{\theta(k)}^T(k)\delta[\theta(k+1)-j]\right\}\\
&= \sum_{i=1}^{\bar{m}} p_{ij}\pi_i(k)\tilde{A}_i(k)\langle \xi(k), \rho(k)\rangle \tilde{B}_i^T(k)\\
&= \sum_{i=1}^{\bar{m}} p_{ij}\pi_i(k)\tilde{A}_i(k)\bar{B}(k-1)\Delta_{+1}^T(k)\tilde{B}_i^T(k) \qquad (22)
\end{aligned}
$$

$$\begin{aligned}&\left\langle \tilde{B}_{\theta(k)}(k)\rho(k)\delta[\theta(k+1)-j], \tilde{A}_{\theta(k)}(k)\xi(k)\delta[\theta(k+1)-j]\right\rangle\\&\quad=\left\langle \tilde{A}_{\theta(k)}(k)\xi(k)\delta[\theta(k+1)-j], \tilde{B}_{\theta(k)}(k)\rho(k)\delta\right.\\&\qquad\left.\times\ [\theta(k+1)-j]\right\rangle^T\\&\quad=\sum_{i=1}^{\bar{m}} p_{ij}\pi_i(k)\tilde{B}_i(k)\Delta_{+1}(k)\bar{B}^T(k-1)\tilde{A}_i^T(k).\end{aligned}\tag{23}$$

Define $\mathbf{G}(k)=(G_1(k),\ldots,G_{\bar{m}}(k))\in\mathbb{H}^{2n+m}$, $\mathbf{D}(k)=(D_1(k),\ldots,D_{\bar{m}}(k))\in\mathbb{H}^{2n+m}$, $\mathbf{D}^0(k)=(D_1^0(k),\ldots,D_{\bar{m}}^0(k))\in\mathbb{H}^{2n+m}$, $\mathbf{D}^{+1}(k)=(D_1^{+1}(k),\ldots,D_{\bar{m}}^{+1}(k))\in\mathbb{H}^{2n+m}$ as

$$\begin{aligned}D_j(k)&=D_j^0(k)+D_j^{+1}(k)+D_j^{+1,T}(k)\\D_j^0(k)&=\sum_{i=1}^{\bar{m}} p_{ij}\pi_i(k)\tilde{B}_i(k)\Delta_0(k)\tilde{B}_i^T(k)\\D_j^{+1}(k)&=\sum_{i=1}^{\bar{m}} p_{ij}\pi_i(k)\tilde{A}_i(k)\bar{B}(k-1)\Delta_{+1}^T(k)\tilde{B}_i^T(k)\end{aligned}$$

and set the operator $\mathcal{T}(k,\cdot)\in\mathbb{B}(\mathbb{H}^{2n+m})$ as follows. For $\Xi=(\Xi_1,\ldots,\Xi_{\bar{m}})\in\mathbb{H}^{2n+m}$, $\mathcal{T}_j(k,\Xi)$ is given by

$$\mathcal{T}_j(k,\Xi)=\sum_{i=1}^{\bar{m}} p_{ij}\tilde{A}_i(k)\Xi_i\tilde{A}_i^T(k), j=1,\ldots,\bar{m}\tag{24}$$

and $\mathcal{T}(k,\Xi)=(\mathcal{T}_1(k,\Xi),\ldots,\mathcal{T}_{\bar{m}}(k,\Xi))$. Therefore, from (18)–(23), the Lyapunov-like recursive equation for matrices $G_i(k)$ is given by

$$\mathbf{G}(k+1)=\mathcal{T}(k,\mathbf{G}(k))+\mathbf{D}(k),\quad G_i(0)=\pi_i(0)\Pi(0).\tag{25}$$

To guarantee the existence of the inverse of some matrices which will be needed in the proof of Theorem 1, some necessary assumptions are made as follows.

Assumption 1: We assume that

$$\mathcal{C}(0)P(0)\mathcal{C}^T(0)+\mathcal{D}(0)\text{diag}\{R(0)+\Lambda(0)\}\mathcal{D}^T(0)>0.$$

Assumption 2: We assume that for $k=1,2,\ldots$

$$\mathcal{D}(k)\text{diag}\{R(k)+\Lambda(k)\}\mathcal{D}^T(k)>0.$$

In addition, some matrices are given as follows before presenting our theorem:

$$\mathcal{A}(k)=\begin{bmatrix}p_{11}\tilde{A}_1(k)&\ldots&p_{\bar{m}1}\tilde{A}_{\bar{m}}(k)\\\vdots&\ddots&\vdots\\p_{1\bar{m}}\tilde{A}_1(k)&\ldots&p_{\bar{m}\bar{m}}\tilde{A}_{\bar{m}}(k)\end{bmatrix}$$

$$\mathcal{A}_\pi(k)=\begin{bmatrix}\pi_1^{1/2}(k)p_{11}\tilde{A}_1(k)&\ldots&\pi_{\bar{m}}^{1/2}(k)p_{\bar{m}1}\tilde{A}_{\bar{m}}(k)\\\vdots&\ddots&\vdots\\\pi_1^{1/2}(k)p_{1\bar{m}}\tilde{A}_1(k)&\ldots&\pi_{\bar{m}}^{1/2}(k)p_{\bar{m}\bar{m}}\tilde{A}_{\bar{m}}(k)\end{bmatrix}$$

$$\mathcal{B}(k)=\begin{bmatrix}\pi_1^{1/2}(k)p_{11}\tilde{B}_1(k)&\ldots&\pi_{\bar{m}}^{1/2}(k)p_{\bar{m}1}\tilde{B}_{\bar{m}}(k)\\\vdots&\ddots&\vdots\\\pi_1^{1/2}(k)p_{1\bar{m}}\tilde{B}_1(k)&\ldots&\pi_{\bar{m}}^{1/2}(k)p_{\bar{m}\bar{m}}\tilde{B}_{\bar{m}}(k)\end{bmatrix}$$

$$\mathcal{B}_\pi(k)=\begin{bmatrix}\pi_1(k)p_{11}\tilde{B}_1(k)&\ldots&\pi_{\bar{m}}(k)p_{\bar{m}1}\tilde{B}_{\bar{m}}(k)\\\vdots&\ddots&\vdots\\\pi_1(k)p_{1\bar{m}}\tilde{B}_1(k)&\ldots&\pi_{\bar{m}}(k)p_{\bar{m}\bar{m}}\tilde{B}_{\bar{m}}(k)\end{bmatrix}$$

$$\mathcal{C}(k)=\begin{bmatrix}\tilde{C}_1(k)&\ldots&\tilde{C}_{\bar{m}}(k)\end{bmatrix}$$

$$\mathcal{D}(k)=\begin{bmatrix}\pi_1^{1/2}(k)\Gamma_1&\ldots&\pi_{\bar{m}}^{1/2}(k)\Gamma_{\bar{m}}\end{bmatrix}$$

and

$$\begin{aligned}\varphi_{ij}(k)&=\delta[\theta(k+1)-j]-p_{ij}\\b(k)&=\sum_{i=1}^{\bar{m}}\tilde{B}_i(k)\rho(k)\delta[\theta(k)-i]\\\Omega^1(k)&=\begin{bmatrix}\sum_{i=1}^{\bar{m}}\tilde{A}_i(k)g(k,i)\varphi_{i1}(k)\\\vdots\\\sum_{i=1}^{\bar{m}}\tilde{A}_i(k)g(k,i)\varphi_{i\bar{m}}(k)\end{bmatrix}\\\Omega^2(k)&=\begin{bmatrix}b(k)\delta[\theta(k+1)-1]\\\vdots\\b(k)\delta[\theta(k+1)-\bar{m}]\end{bmatrix}.\end{aligned}$$

Let us define the linear operators $\mathcal{M}(k,\cdot):\mathbb{H}^{2n+m}\to\mathbb{B}(\mathbb{R}^{\bar{m}(2n+m)})$ and $Dg(\cdot):\mathbb{H}^{2n+m}\to\mathbb{B}(\mathbb{R}^{\bar{m}(2n+m)})$ as follows: for $\Xi=(\Xi_1,\ldots,\Xi_{\bar{m}})\in\mathbb{H}^{2n+m}$

$$\begin{aligned}\mathcal{M}(k,\Xi)&=\text{diag}\left\{\sum_{i=1}^{\bar{m}} p_{ij}\tilde{A}_i(k)\Xi_i\tilde{A}_i^T(k)\right\}-\mathcal{A}(k)\text{diag}\{\Xi_i\}\mathcal{A}^T(k)\\Dg(\Xi)&=\text{diag}\{\Xi_i\}.\end{aligned}$$

Theorem 1: Consider the system represented by (7). Then for $k=0,1,\ldots$, the linear minimum mean square estimator (LMMSE) $\hat{\xi}(k|k)$ is given by

$$\hat{\xi}(k|k)=\sum_{i=1}^{\bar{m}}\hat{g}(k,i|k)\tag{26}$$

where $\hat{g}(k|k)$ satisfies the recursive equation

$$\hat{g}(k|k)=\hat{g}(k)+P(k)\mathcal{C}^T(k)R_e^{-1}(k)\big(\eta(k)-\mathcal{C}(k)\hat{g}(k)\big)\tag{27}$$

$$\hat{g}(k+1|k)=\mathcal{A}(k)\hat{g}(k)+K(k)\big(\eta(k)-\mathcal{C}(k)\hat{g}(k)\big)\tag{28}$$

$$\hat{g}(0|-1)=\zeta(0).\tag{29}$$

The matrices $R_e(k)>0$ and $K(k)$ are given by

$$R_e(k)=\mathcal{C}(k)P(k)\mathcal{C}^T(k)+\mathcal{D}(k)\text{diag}\{R(k)+\Lambda(k)\}\mathcal{D}^T(k)\tag{30}$$

$$\begin{aligned}K(k)=\big(&\mathcal{A}(k)P(k)\mathcal{C}^T(k)+\mathcal{B}(k)\text{diag}\{\Phi^T(k)\}\mathcal{D}^T(k)\\&+\mathcal{B}_\pi(k)\text{diag}\big\{\Delta_{+1}(k)\bar{B}^T(k-1)\big\}\mathcal{C}^T(k)\big)R_e^{-1}(k).\end{aligned}\tag{31}$$

The matrices $P(k)=\|\tilde{g}(k|k-1)\|^2\geq 0$ satisfy the Riccati-like recurrent equation

$$\begin{aligned}P(k+1)=&\mathcal{A}(k)P(k)\mathcal{A}^T(k)+\mathcal{M}(k,\mathbf{G})+Dg(\mathbf{D}(k))\\&-K(k)R_e(k)K^T(k).\end{aligned}\tag{32}$$

Proof: Setting $\zeta(k)=E\{g(k)\}$, it follows that:

$$\zeta(k+1)=\mathcal{A}(k)\zeta(k).\tag{33}$$

From (7), we have

$$\begin{cases} g^c(k+1) = \mathcal{A}(k)g^c(k) + \Omega^1(k) + \Omega^2(k) \\ \eta^c(k) = \mathcal{C}(k)g^c(k) + \Gamma_{\theta(k)}\beta(k). \end{cases} \tag{34}$$

First, since one-step predicted quantities are often encountered, we shall use the following briefer notations (except when necessary for emphasis, or for some other special reason):

$$\begin{aligned} \hat{g}(k) &\triangleq \hat{g}(k|k-1), \quad \hat{g}^c(k) \triangleq \hat{g}^c(k|k-1) \\ \tilde{g}(k) &\triangleq \tilde{g}(k|k-1), \quad \tilde{g}^c(k) \triangleq \tilde{g}^c(k|k-1). \end{aligned}$$

Second, the innovation $e(k)$ and relevant $R_e(k)$ are defined as

$$e(k) \triangleq \tilde{\eta}(k|k-1) = \eta(k) - \hat{\eta}(k|k-1), \quad R_e(k) \triangleq \|e(k)\|^2. \tag{35}$$

Finally, in what follows it will be convenient to introduce the following notation. For any sequence of second order random vectors $\{\chi(k)\}$, we define the "centered" random vector $\chi^c(k)$ as $\chi(k) - E\{\chi(k)\}$. $\hat{\chi}(k|t)$ is the best affine estimator of $\chi(k)$ given $\{\eta(0), \dots, \eta(t)\}$, and $\tilde{\chi}(k|t) = \chi(k) - \hat{\chi}(k|t)$. Similarly $\hat{\chi}^c(k|t)$ is the best affine estimator of $\chi^c(k)$ given $\{\eta^c(0), \dots, \eta^c(t)\}$, and $\tilde{\chi}^c(k|t) = \chi^c(k) - \hat{\chi}^c(k|t)$. It is well known (see [60]) that

$$\hat{\chi}(k|t) = \hat{\chi}^c(k|t) + E\{\chi(k)\} \tag{36}$$

and, in particular

$$\tilde{\chi}^c(k|t) = \tilde{\chi}(k|t). \tag{37}$$

$\mathcal{L}\{(\eta^c)^t\}$ denotes the linear subspace spanned by the random vectors $\{\eta^c(0), \dots, \eta^c(t)\}$. For $t \le k$, the optimal linear estimator $\hat{g}^c(k|t)$ of the random vector $g^c(k)$, represented by

$$\hat{g}^c(k|t) = \begin{pmatrix} \hat{g}_1^c(k|t) \\ \vdots \\ \hat{g}_{\bar{m}}^c(k|t) \end{pmatrix}, \quad g^c(k) = \begin{pmatrix} g_1^c(k) \\ \vdots \\ g_{\bar{m}}^c(k) \end{pmatrix} \tag{38}$$

is the projection of $g^c(k)$ onto the linear subspace $\mathcal{L}\{(\eta^c)^t\}$ and satisfies the following properties (see [60]):

1) $\hat{g}_j^c(k|t) \in \mathcal{L}\{(\eta^c)^t\}, \ j = 1, \dots, \bar{m}$;
2) $\tilde{g}_j^c(k|t)$ is orthogonal to $\mathcal{L}\{(\eta^c)^t\}, \ j = 1, \dots, \bar{m}$;
3) if $\|(\eta)^t\|^2$ is nonsingular then

$$\hat{g}^c(k|t) = \langle g^c(k), \left(\eta^c\right)^t\rangle \|(\eta)^t\|^{-2}\left(\eta^c\right)^t \tag{39}$$

$$\begin{aligned} \hat{g}^c(k|k) = \hat{g}^c(k) &+ \langle g^c(k), \tilde{\eta}(k|k-1)\rangle \|\tilde{\eta}(k|k-1)\|^{-2} \\ &\times \left(\eta^c(k) - \hat{\eta}^c(k|k-1)\right) \end{aligned} \tag{40}$$

where

$$(\eta)^t = \begin{pmatrix} \eta(0) \\ \vdots \\ \eta(t) \end{pmatrix}, \quad (\eta^c)^t = \begin{pmatrix} \eta^c(0) \\ \vdots \\ \eta^c(t) \end{pmatrix}.$$

From (35) and (37), (40) can be rewritten as

$$\hat{g}^c(k|k) = \hat{g}^c(k) + \langle g^c(k), e(k)\rangle R_e^{-1}(k)e(k). \tag{41}$$

Denote by $\mathcal{P}^t$ the orthogonal projection onto $\mathcal{L}\{(\eta^c)^t\}$. From (39), for any null mean random vector χ

$$\mathcal{P}^{k-1}(\chi) = \langle \chi, \left(\eta^c\right)^{k-1}\rangle \|(\eta)^{k-1}\|^{-2}\left(\eta^c\right)^{k-1}. \tag{42}$$

It follows from (42) and (15) that:

$$\begin{aligned} &\mathcal{P}^{k-1}\left(\Gamma_{\theta(k)}(k)\beta(k)\right) \\ &= \left\langle \Gamma_{\theta(k)}(k)\beta(k), \left(\eta^c\right)^{k-1}\right\rangle \|(\eta)^{k-1}\|^{-2}\left(\eta^c\right)^{k-1} \\ &= \sum_{i=1}^{\bar{m}} E\left\{\Gamma_{\theta(k)}(k)\beta(k)\left(\left(\eta^c\right)^{k-1}\right)^T |\theta(k) = i\right\}\pi_i(k)\|(\eta)^{k-1}\|^{-2} \\ &\quad \times \left(\eta^c\right)^{k-1} \\ &= \sum_{i=1}^{\bar{m}} \Gamma_i(k)E\{\beta(k)\}E\left\{\left(\left(\eta^c\right)^{k-1}\right)^T\right\}\pi_i(k)\|(\eta)^{k-1}\|^{-2}\left(\eta^c\right)^{k-1} \\ &= 0. \end{aligned} \tag{43}$$

Since $E\{\beta(k)\} = 0$. From (34) and (43) it is immediate to see that

$$\hat{\eta}^c(k|k-1) = \mathcal{P}^{k-1}\left(\eta^c(k)\right) = \mathcal{C}(k)\hat{g}^c(k). \tag{44}$$

Thus, recalling (37), it follows from (34) and (44) that:

$$\begin{aligned} e(k) &= \tilde{\eta}(k|k-1) = \eta^c(k) - \hat{\eta}^c(k|k-1) \\ &= \mathcal{C}(k)g^c(k) + \Gamma_{\theta(k)}\beta(k) - \mathcal{C}(k)\hat{g}^c(k) \\ &= \mathcal{C}(k)\tilde{g}(k) + \Gamma_{\theta(k)}\beta(k). \end{aligned} \tag{45}$$

From (12) and (45) and setting $P(k) = \|\tilde{g}(k)\|^2 \ge 0$, (30) is obtained as

$$\begin{aligned} &\|\tilde{\eta}(k|k-1)\|^2 \\ &= \mathcal{C}(k)P(k)\mathcal{C}^T(k) + \sum_{i=1}^{\bar{m}} E\left\{\Gamma_i(k)\beta(k)\beta^T(k)\Gamma_i^T(k)\delta[\theta(k) - i]\right\} \\ &= \mathcal{C}(k)P(k)\mathcal{C}^T(k) + \sum_{i=1}^{\bar{m}} \Gamma_i(k)\|\beta(k)\|^2\Gamma_i^T(k)\pi_i(k) \\ &= \mathcal{C}(k)P(k)\mathcal{C}^T(k) + \mathcal{D}(k)\text{diag}\{R(k) + \Lambda(k)\}\mathcal{D}^T(k) \\ &= R_e(k) > 0 \end{aligned}$$

from Assumptions 1 and 2. Then

$$\begin{aligned} \left\langle g^c(k), \Gamma_{\theta(k)}\beta(k)\right\rangle &= \sum_{i=1}^{\bar{m}} E\left\{g^c(k)\beta^T(k)\Gamma_i^T(k)\delta[\theta(k) - i]\right\} \\ &= \sum_{i=1}^{\bar{m}} E\left\{g^c(k)\delta[\theta(k) - i]\right\}E\left\{\beta^T(k)\right\}\Gamma_i^T(k) \end{aligned} \tag{46}$$

$= 0$, and it follows from (45) and (46) that:

$$\begin{aligned} \langle g^c(k), \tilde{\eta}(k|k-1)\rangle &= \langle g^c(k), \tilde{g}(k)\rangle\mathcal{C}^T(k) \\ &= \langle \hat{g}^c(k) + \tilde{g}(k), \tilde{g}(k)\rangle\mathcal{C}^T(k) \\ &= \langle \hat{g}^c(k), \tilde{g}(k)\rangle\mathcal{C}^T(k) + \langle \tilde{g}(k), \tilde{g}(k)\rangle\mathcal{C}^T(k) \\ &= P(k)\mathcal{C}^T(k). \end{aligned} \tag{47}$$

From (30), (41), (44), and (47), it can be seen as the following form:

$$\begin{aligned} \hat{g}^c(k|k) &= \hat{g}^c(k) + \langle g^c(k), e(k)\rangle R_e^{-1}(k)e(k) \\ &= \hat{g}^c(k) + P(k)\mathcal{C}^T(k)R_e^{-1}(k)e(k) \\ &= \hat{g}^c(k) + P(k)\mathcal{C}^T(k)R_e^{-1}(k)\left(\eta^c(k) - \mathcal{C}(k)\hat{g}^c(k)\right). \end{aligned}$$

From (36) and noticing that $e(k) = \eta^c(k) - \mathcal{C}(k)\hat{g}^c(k) = \eta(k) - \mathcal{C}(k)\hat{g}(k)$, we obtain

$$\hat{g}(k|k) = \hat{g}(k) + P(k)\mathcal{C}^T(k)R_e^{-1}(k)\big(\eta(k) - \mathcal{C}(k)\hat{g}(k)\big). \quad (48)$$

Let us now derive (28). It follows from (34) that:

$$\hat{g}^c(k+1) = \mathcal{A}(k)\hat{g}^c(k|k) + \mathcal{P}^k\Big(\Omega^1(k)\Big) + \mathcal{P}^k\Big(\Omega^2(k)\Big). \quad (49)$$

It can be seen from (39) that

$$\mathcal{P}^k\Big(\Omega^1(k)\Big) = \langle \Omega^1(k), \big(\eta^c\big)^k\rangle \|(\eta)^k\|^{-2}\big(\eta^c\big)^k \quad (50)$$

and

$$\begin{aligned}
&\mathcal{P}^k\left(\sum_{i=1}^{\bar{m}} \tilde{A}_i(k)g(k,i)\varphi_{ij}(k)\right)\\
&= \sum_{i=1}^{\bar{m}} \langle \tilde{A}_i(k)g(k,i)\varphi_{ij}(k), \big(\eta^c\big)^k\rangle \|(\eta)^k\|^{-2}\big(\eta^c\big)^k\\
&= \sum_{i=1}^{\bar{m}} E\Big\{\tilde{A}_i(k)\xi(k)\delta[\theta(k)-i]\big(\delta[\theta(k+1)-j]-p_{ij}\big)\\
&\qquad \times \Big(\big(\eta^c\big)^k\Big)^T\Big\} \times \|(\eta)^k\|^{-2}\big(\eta^c\big)^k\\
&= \sum_{i=1}^{\bar{m}} \tilde{A}_i(k)E\Big\{\xi(k)\Big(\big(\eta^c\big)^k\Big)^T\Big\}\|(\eta)^k\|^{-2}\big(\eta^c\big)^k\\
&\quad \times E\big\{\delta[\theta(k)-i]\big(\delta[\theta(k+1)-j]-p_{ij}\big)\big\}\\
&= 0 \qquad (51)
\end{aligned}$$

where the last step of the equation holds because

$$\begin{aligned}
&E\big\{\delta[\theta(k)-i]\big(\delta[\theta(k+1)-j]-p_{ij}\big)\big\}\\
&\quad = \mathbf{P}(\theta(k)=i, \theta(k+1)=j) - p_{ij}\mathbf{P}(\theta(k)=i)\\
&\quad = \mathbf{P}(\theta(k+1)=j|\theta(k)=i)\mathbf{P}(\theta(k)=i) - p_{ij}\mathbf{P}(\theta(k)=i)\\
&\quad = 0. \qquad (52)
\end{aligned}$$

This shows that $\mathcal{P}^k(\Omega^1(k)) = 0$. It can be shown that from (41)

$$\mathcal{P}^k\Big(\Omega^2(k)\Big) = \mathcal{P}^{k-1}\Big(\Omega^2(k)\Big) + \langle \Omega^2(k), e(k)\rangle R_e^{-1}(k)e(k). \quad (53)$$

From the independence hypothesis and (39), it can be obtained that

$$\begin{aligned}
&\mathcal{P}^{k-1}(b(k)\delta[\theta(k+1)-j])\\
&\quad = \sum_{i=1}^{\bar{m}} \Big\langle \tilde{B}_i(k)\rho(k)\delta[\theta(k)-i]\delta[\theta(k+1)-j], \big(\eta^c\big)^{k-1}\Big\rangle\\
&\qquad \times \|(\eta)^{k-1}\|^{-2}\big(\eta^c\big)^{k-1}\\
&\quad = \sum_{i=1}^{\bar{m}} \tilde{B}_i(k)E\{\rho(k)\}E\Big\{\Big(\big(\eta^c\big)^{k-1}\Big)^T\Big\}\|(\eta)^{k-1}\|^{-2}\big(\eta^c\big)^{k-1}\\
&\qquad \times E\{\delta[\theta(k)-i]\delta[\theta(k+1)-j]\}\\
&\quad = 0 \qquad (54)
\end{aligned}$$

where the last step of the equation holds because $E\{\rho(k)\} = 0$. From (45), we have that

$$\begin{aligned}
&\langle b(k)\delta[\theta(k+1)-j], e(k)\rangle\\
&\quad = \big\langle b(k)\delta[\theta(k+1)-j], \mathcal{C}(k)\tilde{g}(k)\big\rangle + \langle b(k)\delta[\theta(k+1)-j]\\
&\qquad \times \Gamma_{\theta(k)}\beta(k)\big\rangle \qquad (55)
\end{aligned}$$

where

$$\begin{aligned}
&\langle b(k)\delta[\theta(k+1)-j], \mathcal{C}(k)\tilde{g}(k)\rangle\\
&= \left\langle \sum_{i=1}^{\bar{m}} \tilde{B}_i(k)\rho(k)\delta[\theta(k)-i]\delta[\theta(k+1)-j], \mathcal{C}(k)g(k)\right\rangle\\
&= \left\langle \sum_{i=1}^{\bar{m}} \tilde{B}_i(k)\rho(k)\delta[\theta(k)-i]\delta[\theta(k+1)-j], \sum_{i=1}^{\bar{m}} \tilde{C}_i(k)\xi(k)\right\rangle\\
&= \sum_{i=1}^{\bar{m}} p_{ij}\pi_i(k)\tilde{B}_i(k)\langle \rho(k), \xi(k)\rangle \tilde{C}_i^T\\
&= \sum_{i=1}^{\bar{m}} p_{ij}\pi_i(k)\tilde{B}_i(k)\Delta_{+1}(k)\bar{B}^T(k-1)\tilde{C}_i^T \qquad (56)
\end{aligned}$$

since $\rho(k) \perp \hat{g}(k)$, and

$$\begin{aligned}
&\big\langle b(k)\delta[\theta(k+1)-j], \Gamma_{\theta(k)}\beta(k)\big\rangle\\
&\quad = \sum_{i=1}^{\bar{m}} \tilde{B}_i(k)\big\langle \rho(k)\delta[\theta(k)-i]\delta[\theta(k+1)-j], \Gamma_{\theta(k)}\beta(k)\big\rangle\\
&\quad = \sum_{i=1}^{\bar{m}} \tilde{B}_i(k)\langle \rho(k), \beta(k)\rangle \Gamma_i^T E\{\delta[\theta(k)-i]\delta[\theta(k+1)-j]\}\\
&\quad = \sum_{i=1}^{\bar{m}} p_{ij}\pi_i(k)\tilde{B}_i(k)\Phi^T(k)\Gamma_i^T. \qquad (57)
\end{aligned}$$

Therefore

$$\begin{aligned}
\mathcal{P}^k\Big(\Omega^2(k)\Big) &= \big(\mathcal{B}_\pi(k)\text{diag}\big\{\Delta_{+1}(k)\bar{B}^T(k-1)\big\}\mathcal{C}^T(k)\\
&\quad + \mathcal{B}(k)\text{diag}\big\{\Phi^T(k)\big\}\mathcal{D}^T(k)\big)R_e^{-1}(k)e(k). \qquad (58)
\end{aligned}$$

From (27), (49) and above results, (31) is given by

$$\begin{aligned}
&\hat{g}^c(k+1)\\
&\quad = \mathcal{A}(k)\hat{g}^c(k) + \big(\mathcal{A}(k)P(k)\mathcal{C}^T(k) + \mathcal{B}(k)\text{diag}\{\Phi^T(k)\}\mathcal{D}^T(k)\\
&\qquad + \mathcal{B}_\pi(k)\text{diag}\big\{\Delta_{+1}(k)\bar{B}^T(k-1)\big\}\mathcal{C}^T(k)\big)R_e^{-1}(k)e(k)\\
&\quad = \mathcal{A}(k)\hat{g}^c(k) + K(k)e(k) \qquad (59)
\end{aligned}$$

where

$$\begin{aligned}
K(k) &= \big(\mathcal{A}(k)P(k)\mathcal{C}^T(k) + \mathcal{B}(k)\text{diag}\{\Phi^T(k)\}\mathcal{D}^T(k)\\
&\quad + \mathcal{B}_\pi(k)\text{diag}\big\{\Delta_{+1}(k)\bar{B}^T(k-1)\big\}\mathcal{C}^T(k)\big)R_e^{-1}(k)
\end{aligned}$$

is directly (31). From (59), (36), and (33), (28) is obtained. Finally, it follows from (28) and (36) that:

$$\begin{aligned}
\tilde{g}(k+1) &= g(k+1) - \hat{g}(k+1)\\
&= g^c(k+1) + \zeta(k+1) - \hat{g}(k+1)\\
&= \mathcal{A}(k)g^c(k) + \Omega^1(k) + \Omega^2(k) + \mathcal{A}(k)\zeta(k)\\
&\quad - \big(\mathcal{A}(k)\tilde{g}(k) + K(k)\big(\mathcal{C}(k)\tilde{g}(k) + \Gamma_{\theta(k)}\beta(k)\big)\big)\\
&= (\mathcal{A}(k) - K(k)\mathcal{C}(k))\tilde{g}(k)\\
&\quad + \Omega^1(k) + \Omega^2(k) - K(k)\Gamma_{\theta(k)}\beta(k). \qquad (60)
\end{aligned}$$

Set

$$\begin{aligned}
\Omega^0(k) &= (\mathcal{A}(k) - K(k)\mathcal{C}(k))\tilde{g}(k)\\
\Omega^3(k) &= -K(k)\Gamma_{\theta(k)}\beta(k).
\end{aligned}$$

Hence, it follows from (60) that:

$$\tilde{g}(k+1) = \sum_{\kappa=0}^{3} \Omega^{\kappa}(k). \tag{61}$$

From the independent hypothesis mentioned above, namely, $\beta(k) \perp \xi(k)$ and $\beta(k) \perp \tilde{g}(k)$, it can be seen that $\langle \Omega^3(k), \Omega^{\kappa}(k)\rangle = 0, \kappa = 0, 1$. Moreover, from (52) one has $\langle \Omega^1(k), \Omega^0(k)\rangle = 0$. Therefore, it can be seen that

$$\begin{aligned} P(k+1) &= \|\tilde{g}(k+1)\|^2 \\ &= \sum_{\kappa=0}^{3} \|\Omega^{\kappa}(k)\|^2 + \sum_{\kappa=0,\kappa\neq 2}^{3} \Big\langle \Omega^2(k), \Omega^{\kappa}(k)\Big\rangle \\ &\quad + \sum_{\kappa=0,\kappa\neq 2}^{3} \Big\langle \Omega^{\kappa}(k), \Omega^2(k)\Big\rangle. \end{aligned} \tag{62}$$

It can be seen that

$$\|\Omega^0(k)\|^2 = (\mathcal{A}(k) - K(k)\mathcal{C}(k))P(k)(\mathcal{A}(k) - K(k)\mathcal{C}(k))^T. \tag{63}$$

Applying the similar derivation procedures used before, we have that

$$\|\Omega^1(k)\|^2 = \mathcal{M}(k, \mathbf{G}) \tag{64}$$

$$\|\Omega^2(k)\|^2 = Dg\Big(\mathbf{D}^0(k)\Big) \tag{65}$$

$$\|\Omega^3(k)\|^2 = K(k)\mathcal{D}(k)\text{diag}\{R(k) + \Lambda(k)\}\mathcal{D}^T(k)K^T(k). \tag{66}$$

Similarly, the rest of covariance matrices are simplified as follows:

$$\begin{aligned} \Big\langle \Omega^0(k), \Omega^2(k)\Big\rangle &= (\mathcal{A}(k) - K(k)\mathcal{C}(k))\text{diag}\big\{\bar{B}(k-1)\Delta_{+1}^T(k)\big\} \\ &\quad \times \mathcal{B}_{\pi}^T(k) \end{aligned} \tag{67}$$

$$\begin{aligned} \Big\langle \Omega^1(k), \Omega^2(k)\Big\rangle &= Dg\Big(\mathbf{D}^{+1}(k)\Big) - \mathcal{A}_{\pi}(k)\text{diag} \\ &\quad \times \big\{\bar{B}(k-1)\Delta_{+1}^T(k)\big\} \times \mathcal{B}^T(k) \end{aligned} \tag{68}$$

$$\Big\langle \Omega^2(k), \Omega^3(k)\Big\rangle = -\mathcal{B}(k)\text{diag}\big\{\Phi^T(k)\big\}\mathcal{D}^T(k)K^T(k)0. \tag{69}$$

Putting all these results together, we get that

$$\begin{aligned} &P(k+1) \\ &= (\mathcal{A}(k) - K(k)\mathcal{C}(k))P(k)(\mathcal{A}(k) - K(k)\mathcal{C}(k))^T + \mathcal{M}(k, \mathbf{G}) \\ &\quad + Dg(\mathbf{D}(k)) + K(k)\mathcal{D}(k)\text{diag}\{R(k) + \Lambda(k)\}\mathcal{D}^T(k)K^T(k) \\ &\quad - \mathcal{B}(k)\text{diag}\{\Phi^T(k)\}\mathcal{D}^T(k)K^T(k) \\ &\quad - K(k)\mathcal{D}(k)\text{diag}\{\Phi(k)\}\mathcal{B}^T(k) \\ &\quad - \mathcal{B}_{\pi}(k)\text{diag}\big\{\Delta_{+1}(k)\bar{B}^T(k-1)\big\}\mathcal{C}^T(k)K^T(k) \\ &\quad - K(k)\mathcal{C}(k)\text{diag}\big\{\bar{B}(k-1)\Delta_{+1}^T(k)\big\}\mathcal{B}_{\pi}^T(k) \end{aligned} \tag{70}$$

and (32) is obtained after some algebraic manipulations.

Remark 2: From (31), it can be seen that the computational complexity of filtering algorithm in Theorem 1 mainly depends on dimensions of system matrix $A(k), B(k), C(k)$ and the number of transmitting nodes $\bar{m}$. The calculation time will increase as these parameters become larger, especially in the calculation of $R_e^{-1}(k)$.

IV. Numerical Examples

In this section, a simulation example is presented to demonstrate the effectiveness of designed LMMSE filter for discrete-time stochastic systems with communication constraint and data quantization.

The plant which is modeled by (1) is considered with the following parameters:

$$A(k) = \begin{bmatrix} 0.3 & 0.7 \\ 0.2 & 0.6 \end{bmatrix} + 2\sin(0.5\pi k)\begin{bmatrix} 0.1 & 0.05 \\ 0.2 & 0.1 \end{bmatrix}$$

$$B(k) = \begin{bmatrix} 1 & 0.5 \end{bmatrix}^T, C(k) = \begin{bmatrix} 0.5 & 1 \\ 1 & 1 \end{bmatrix}, H(k) = \begin{bmatrix} 0.25 & 0.5 \\ 0.5 & 0.5 \end{bmatrix}$$

$w(k)$ and $v(k)$ are uncorrelated white Gaussian noises with zero means and variances $Q(k) = 0.1$ and $R(k) = 0.02I_2$, respectively. It is assumed that the initial states $z(0) = z_0 = [2\ 1\ 2\ 1]^T$ and $\Pi_{z,0} = 0.1I_4$. The output $y(k)$ is measured by two sensors independently. The measurement outputs are quantified and transmitted by each sensor to the remote estimator via a shared transmission channel. The communication constraint is subject to a Markov protocol with transition probability matrix being $P_{\bar{m}} = \begin{bmatrix} 0.8 & 0.2 \\ 0.2 & 0.8 \end{bmatrix}$ and the initial probability distribution is set as $\pi_1(0) = 0.1,\ \pi_2(0) = 0.9$. Before transmitted through the shared channel, data are quantified by an uniformed quantizer with $M = 10,\ b = 8$, leading to the quantizer level length U is relatively small. After that, we choose $\gamma_1 = \gamma_2 = 0$.

In order to confirm the effectiveness of our algorithm, by applying Theorem 1, Fig. 2 shows that the estimated values tracked the states well, and the trace of the estimation error covariance matrix is convergent.

Now, we consider different scenarios with different parameters. First, we discuss the impact on state estimation effect with different weighting parameters γ_1 and γ_2 keeping other parameters unchanged. From Table I, it can be seen that as γ_1 or γ_2 becomes larger, the estimation error also becomes larger. In what follows, we shall study the influence on estimation performance under different choices of transition probability matrix $P_{\bar{m}}$ maintaining other parameters unvaried. From data in Table II, it can be inferred that the transition probability matrix $P_{\bar{m}}$ also effects the estimation performance. Therefore, one can improve estimation performance by changing $P_{\bar{m}}$. Finally, the impact of quantizer to estimation performance is shown in Fig. 3. It can be seen that as long as the quantization density is not very small, the estimation performance is satisfactory in some degree.

In summary, in addition to the noises $\{w(k), v(k)\}$, the different choice of weighting parameters $\{\gamma_1, \gamma_2\}$, transition probability matrix $P_{\bar{m}}$ and the quantization density of quantizer can also affect the state tracking performance.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

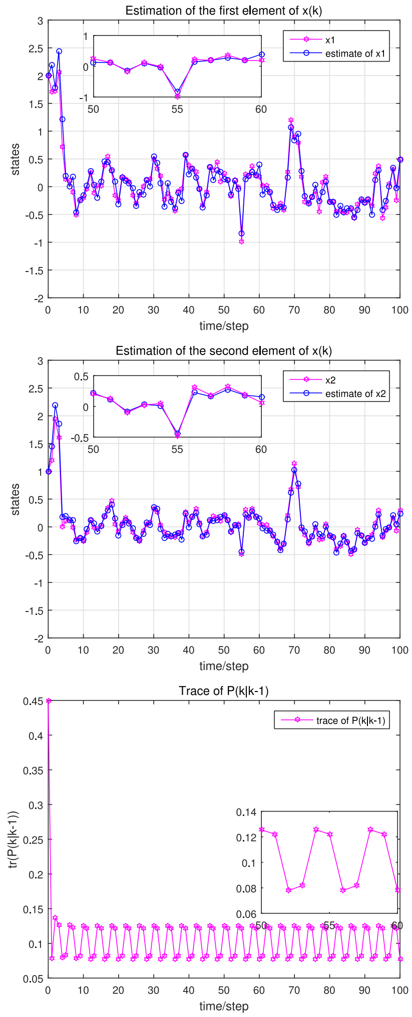


Fig. 2. LMMSE filter of system (1) with given parameters.

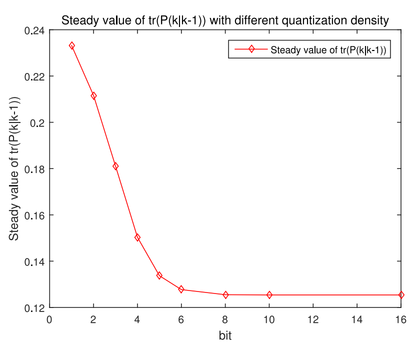


Fig. 3. Steady value of $\mathrm{tr}(P(k|k-1))$ with different quantization density.

Remark 3: Notice the fact that

$$P(k) = \|\tilde{g}(k)\|^2, \xi(k) = \begin{bmatrix} x(k) \\ x(k-1) \\ \eta(k-1) \end{bmatrix}$$

$$g(k) = \begin{bmatrix} \xi(k)\delta[\theta(k)-1] \\ \xi(k)\delta[\theta(k)-2] \\ \vdots \\ \xi(k)\delta[\theta(k)-m] \end{bmatrix}.$$

Therefore, we can find that only part of $P(k)$ reflects the estimated error of state $x(k)$. In this section, we use the trace of

TABLE I
UPPER BOUND OF STEADY VALUE OF $\mathrm{TR}(P(k|k-1))$ WITH DIFFERENT γ_1 AND γ_2

	$\gamma_1=0$	$\gamma_1=0.25$	$\gamma_1=0.5$	$\gamma_1=0.75$	$\gamma_1=1$
$\gamma_2=0$	0.1255	0.1257	0.1278	0.1340	0.1538
$\gamma_2=0.25$	0.1261	0.1262	0.1281	0.1339	0.1530
$\gamma_2=0.5$	0.1284	0.1283	0.1298	0.1350	0.1532
$\gamma_2=0.75$	0.1348	0.1343	0.1352	0.1396	0.1562
$\gamma_2=1$	0.1540	0.1528	0.1527	0.1556	0.1702

TABLE II
STEADY VALUE OF $\mathrm{TR}(P(k|k-1))$ WITH DIFFERENT $P_{\tilde{m}}$

$P_{\tilde{m}}$	$\begin{bmatrix} 0.8 & 0.2 \\ 0.8 & 0.2 \end{bmatrix}$	$\begin{bmatrix} 0.2 & 0.8 \\ 0.2 & 0.8 \end{bmatrix}$	$\begin{bmatrix} 0.8 & 0.2 \\ 0.2 & 0.8 \end{bmatrix}$	$\begin{bmatrix} 0.2 & 0.8 \\ 0.8 & 0.2 \end{bmatrix}$
$tr(P(k|k-1))$	0.1266	0.1233	0.1255	0.1255
$\pi_1(k)$	0.8	0.2	0.5	0.5
$\pi_2(k)$	0.2	0.8	0.5	0.5
$P_{\tilde{m}}$	$\begin{bmatrix} 0.95 & 0.05 \\ 0.95 & 0.05 \end{bmatrix}$	$\begin{bmatrix} 0.05 & 0.95 \\ 0.05 & 0.95 \end{bmatrix}$	$\begin{bmatrix} 0.95 & 0.05 \\ 0.05 & 0.95 \end{bmatrix}$	$\begin{bmatrix} 0.05 & 0.95 \\ 0.95 & 0.05 \end{bmatrix}$
$tr(P(k|k-1))$	0.0932	0.0866	0.0906	0.0907
$\pi_1(k)$	0.95	0.05	0.5	0.5
$\pi_2(k)$	0.05	0.95	0.5	0.5

the matrix $P(k)$ to measure the size of the error covariance matrix. Consequently, part of such trace is used to measure the size of the estimated error of $x(k)$. Therefore, any "trace of $P(k|k-1)$" appearing in this section stands for part of trace related to system state $x(k)$ only.

V. CONCLUSION

The linear minimum mean square filter design problem has been studied in this paper for a networked discrete time-varying linear system subject to data quantization and communication constraints. It has been shown that estimation performance depends on the transition probability matrix of the Markov chain, quantization error, and the shared channel weighting parameter. A Kalman-like filter is presented in Theorem 1, based on the Lyapunov and Riccati-like equations. A numerical example has demonstrated that the effectiveness and applicability of the proposed LMMSE filters. Modeling uncertainty and numerical error will be considered in our future work. To overcome this problem, an FIR filter might be used to solve the corresponding robust estimation problems.

REFERENCES

[1] H. R. Karimi and H. Gao, "New delay-dependent exponential H_∞ synchronization for uncertain neural networks with mixed time delays," *IEEE Trans. Syst., Man, Cybern. B, Cybern.*, vol. 40, no. 1, pp. 173–185, Feb. 2010.

[2] H. R. Karimi, "Robust H_∞ filter design for uncertain linear systems over network with network-induced delays and output quantization," *Model. Identification Control*, vol. 30, no. 1, pp. 27–37, 2009.

[3] H. R. Karimi, N. A. Duffie, and S. Dashkovskiy, "Local capacity H_∞ control for production networks of autonomous work systems with time-varying delays," *IEEE Trans. Autom. Sci. Eng.*, vol. 7, no. 4, pp. 849–857, Oct. 2010.

[4] J. Wu, H. R. Karimi, and P. Shi, "Network-based H_∞ output feedback control for uncertain stochastic systems," *Inf. Sci.*, vol. 232, pp. 397–410, May 2013.

[5] F. Deng *et al.*, "Energy-based sound source localization with low power consumption in wireless sensor networks," *IEEE Trans. Ind. Electron.*, vol. 64, no. 6, pp. 4894–4902, Jun. 2017, doi: 10.1109/TIE.2017.2652394.

[6] F. Deng, S. Guo, R. Zhou, and J. Chen, "Sensor multifault diagnosis with improved support vector machines," *IEEE Trans. Autom. Sci. Eng.*, vol. 14, no. 2, pp. 1053–1063, Apr. 2017.

[7] W.-A. Zhang, L. Yu, and G. Feng, "Optimal linear estimation for networked systems with communication constraints," *Automatica*, vol. 47, no. 9, pp. 1992–2000, 2011.

[8] H. Dong, Z. Wang, and H. Gao, "Robust H_∞ filtering for a class of nonlinear networked systems with multiple stochastic communication delays and packet dropouts," *IEEE Trans. Signal Process.*, vol. 58, no. 4, pp. 1957–1966, Apr. 2010.

[9] H. Gao and T. Chen, "H_∞ estimation for uncertain systems with limited communication capacity," *IEEE Trans. Autom. Control*, vol. 52, no. 11, pp. 2070–2084, Nov. 2007.

[10] X. He, Z. Wang, and D. Zhou, "Robust H_∞ filtering for networked systems with multiple state delays," *Int. J. Control*, vol. 80, no. 8, pp. 1217–1232, 2007.

[11] M. Moayedi, Y. K. Foo, and Y. C. Soh, "Adaptive Kalman filtering in networked systems with random sensor delays, multiple packet dropouts and missing measurements," *IEEE Trans. Signal Process.*, vol. 58, no. 3, pp. 1577–1588, Mar. 2010.

[12] N. Xiao, L. Xie, and M. Fu, "Kalman filtering over unreliable communication networks with bounded Markovian packet dropouts," *Int. J. Robust Nonlin. Control*, vol. 19, no. 16, pp. 1770–1786, 2009.

[13] H. Su, H. Wu, X. Chen, and M. Z. Q. Chen, "Positive edge consensus of complex networks," *IEEE Trans. Syst., Man, Cybern., Syst.*, to be published, doi: 10.1109/TSMC.2017.2765678.

[14] H. Su, H. Wu, and X. Chen, "Observer-based discrete-time nonnegative edge synchronization of networked systems," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 28, no. 10, pp. 2446–2455, Oct. 2017.

[15] A. Leon-Garcia and I. Widjaja, *Communication Networks: Fundamental Concepts and Key Architectures*. New York, NY, USA: McGraw-Hill, 1999.

[16] C. Zhou, X. Huang, X. Naixue, Y. Qin, and S. Huang, "A class of general transient faults propagation analysis for networked control systems," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 45, no. 4, pp. 647–661, Apr. 2015.

[17] D. Hristu-Varsakelis, "Short-period communication and the role of zero-order holding in networked control systems," *IEEE Trans. Autom. Control*, vol. 53, no. 5, pp. 1285–1290, Jun. 2008.

[18] D. Wang, J. Wang, and W. Wang, "H_∞ controller design of networked control systems with Markov packet dropouts," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 43, no. 3, pp. 689–697, May 2013.

[19] D. Wu, J. Wu, and S. Chen, "Robust H_∞ control for networked control systems with uncertainties and multiple-packet transmission," *IET Control Theory Appl.*, vol. 4, no. 5, pp. 701–709, May 2010.

[20] G. Tambini, G. C. Montanari, and M. Cacciari, "The Kalman filter as a way to estimate the life-model parameters of insulating materials and system," in *Proc. Int. Conf. Conduction Breakdown Solid Dielectr.*, 1992, pp. 523–527.

[21] D. Unsal and M. Dogan, "Implementation of identification system for IMUs based on Kalman filtering," in *Proc. IEEE/ION Position Location Navig. Symp. (PLANS)*, May 2014, pp. 236–240.

[22] A. Yadav, N. Naik, M. R. Ananthasayanam, A. Gaur, and Y. N. Singh, "A constant gain Kalman filter approach to target tracking in wireless sensor networks," in *Proc. IEEE Int. Conf. Ind. Inf. Syst.*, 2012, pp. 1–7.

[23] H. Chou, M. Traonmilin, E. Ollivier, and M. Parent, "A simultaneous localization and mapping algorithm based on Kalman filtering," in *Proc. Intell. Veh. Symp.*, 2004, pp. 631–635.

[24] H. Zhang *et al.*, "Codesign of event-triggered and distributed H_∞ filtering for active semi-vehicle suspension systems," *IEEE/ASME Trans. Mechatronics*, vol. 22, no. 2, pp. 1047–1058, Apr. 2017.

[25] H. Yan, Q. Yang, H. Zhang, F. Yang, and X. Zhan, "Distributed H_∞ state estimation for a class of filtering networks with time-varying switching topologies and packet losses," *IEEE Trans. Syst., Man, Cybern., Syst.*, to be published, doi: 10.1109/TSMC.2017.2708507.

[26] H. Yan, H. Zhang, F. Yang, X. Zhan, and C. Peng, "Event-triggered asynchronous guaranteed cost control for Markov jump discrete-time neural networks with distributed delay and channel fading," *IEEE Trans. Neural Netw. Learn. Syst.*, to be published, doi: 10.1109/TNNLS.2017.2732240.

[27] H. Zhang, R. Yang, H. Yan, and F. Yang, "H_∞ consensus of event-based multi-agent systems with switching topology," *Inf. Sci.*, vols. 370–371, pp. 623–635, Nov. 2016.

[28] H. Yan, H. Zhang, F. Yang, C. Huang, and S. Chen, "Distributed H_∞ filtering for switched repeated scalar nonlinear systems with randomly occurred sensor nonlinearities and asynchronous switching," *IEEE Trans. Syst., Man, Cybern., Syst.*, to be published, doi: 10.1109/TSMC.2017.2754495.

[29] D. B. Dačić and D. Nešić, "Quadratic stabilization of linear networked control systems via simultaneous protocol and controller design," *Automatica*, vol. 43, no. 7, pp. 1145–1155, 2007.

[30] D. Dacic and D. Nesic, "Observer design for wired linear networked control systems using matrix inequalities," in *Proc. IEEE Conf. Decis. Control*, 2008, pp. 3315–3320.

[31] D. Hristu and K. Morgansen, "Limited communication control," *Syst. Control Lett.*, vol. 37, no. 4, pp. 193–205, 1999.

[32] H. Ishii, "H_∞ control with limited communication and message losses," *Syst. Control Lett.*, vol. 57, no. 4, pp. 322–331, 2008.

[33] H. Rehbinder and M. Sanfridson, "Scheduling of a limited communication channel for optimal control," in *Proc. IEEE Conf. Decis. Control*, 2000, pp. 1011–1016.

[34] L. Zhang and D. Hristu-Varsakelis, "Communication and control codesign for networked control systems," *Automatica*, vol. 42, no. 6, pp. 953–958, 2006.

[35] X. Liu and A. Goldsmith, "Kalman filtering with partial observation losses," in *Proc. IEEE Conf. Decis. Control*, vol. 4, 2005, pp. 4180–4186.

[36] X. Kai, C. Wei, and L. Liu, "Robust extended Kalman filtering for nonlinear systems with stochastic uncertainties," *IEEE Trans. Syst., Man, Cybern. A, Syst., Humans*, vol. 40, no. 2, pp. 399–405, Mar. 2010.

[37] F. O. Hounkpevi and E. E. Yaz, "Robust minimum variance linear state estimators for multiple sensors with different failure rates," *Automatica*, vol. 43, no. 7, pp. 1274–1280, 2007.

[38] Z. Xing, Y. Xia, L. Yan, K. Lu, and Q. Gong, "Multisensor distributed weighted Kalman filter fusion with network delays, stochastic uncertainties, autocorrelated, and cross-correlated noises," *IEEE Trans. Syst., Man, Cybern., Syst.*, to be published, doi: 10.1109/TSMC.2016.2633283.

[39] H. Song, W.-A. Zhang, and L. Yu, "H_∞ filtering of network-based systems with communication constraints," *IET Signal Process.*, vol. 4, no. 1, pp. 69–77, Feb. 2010.

[40] W. Li, G. Wei, D. Ding, Y. Liu, and F. E. Alsaadi, "A new look at boundedness of error covariance of Kalman filtering," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 48, no. 2, pp. 309–314, Feb. 2018, doi: 10.1109/TSMC.2016.2598845.

[41] J. Qiu, Y. Wei, and H. R. Karimi, "New approach to delay-dependent H_∞ control for continuous-time Markovian jump systems with time-varying delay and deficient transition descriptions," *J. Frankl. Inst.*, vol. 352, no. 1, pp. 189–215, 2015.

[42] Y. Wei, J. Qiu, H. R. Karimi, and M. Wang, "A new design of H_∞ filtering for continuous-time Markovian jump systems with time-varying delay and partially accessible mode information," *Signal Process.*, vol. 93, no. 9, pp. 2392–2407, 2013.

[43] W. He, Y. Chen, and Z. Yin, "Adaptive neural network control of an uncertain robot with full-state constraints," *IEEE Trans. Cybern.*, vol. 46, no. 3, pp. 620–629, Mar. 2016.

[44] W. He and Y. Dong, "Adaptive fuzzy neural network control for a constrained robot using impedance learning," *IEEE Trans. Neural Netw. Learn. Syst.*, to be published, doi: 10.1109/TNNLS.2017.2665581.

[45] W. He, Y. Dong, and C. Sun, "Adaptive neural impedance control of a robotic manipulator with input saturation," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 46, no. 3, pp. 334–344, Mar. 2016.

[46] W. He, Z. Yan, C. Sun, and Y. Chen, "Adaptive neural network control of a flapping wing micro aerial vehicle with disturbance observer," *IEEE Trans. Cybern.*, vol. 47, no. 10, pp. 3452–3465, Oct. 2017.

[47] W. He and S. S. Ge, "Cooperative control of a nonuniform gantry crane with constrained tension," *Automatica*, vol. 66, no. 4, pp. 146–154, 2016.

[48] C. Wu, J. Liu, Y. Xiong, and L. Wu, "Observer-based adaptive fault-tolerant tracking control of nonlinear nonstrict-feedback systems," *IEEE Trans. Neural Netw. Learn. Syst.*, to be published, doi: 10.1109/TNNLS.2017.2712619.

[49] C. Wu, J. Liu, X. Jing, H. Li, and L. Wu, "Adaptive fuzzy control for nonlinear networked control systems," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 47, no. 8, pp. 2420–2430, Aug. 2017.

[50] C. Sun, W. He, and J. Hong, "Neural network control of a flexible robotic manipulator using the lumped spring-mass model," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 47, no. 8, pp. 1863–1874, Aug. 2017.

[51] Y. S. Shmaliy, S. Zhao, and C. K. Ahn, "Unbiased finite impluse response filtering: An iterative alternative to Kalman filtering ignoring noise and initial conditions," *IEEE Control Syst. Mag.*, vol. 37, no. 5, pp. 70–89, Oct. 2017.

[52] C. K. Ahn, P. Shi, and M. V. Basin, "Deadbeat dissipative FIR filtering," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 63, no. 8, pp. 1210–1221, Aug. 2016.

[53] J. M. Pak, C. K. Ahn, Y. S. Shmaliy, and M. T. Lim, "Improving reliability of particle filter-based localization in wireless sensor networks via hybrid particle/FIR filtering," *IEEE Trans. Ind. Informat.*, vol. 11, no. 5, pp. 1089–1098, Oct. 2015.

[54] M. Sahebsara, T. Chen, and S. L. Shah, "Optimal H_2 filtering in networked control systems with multiple packet dropout," *IEEE Trans. Autom. Control*, vol. 52, no. 8, pp. 1508–1513, Aug. 2007.

[55] L. Schenato, "Optimal estimation in networked control systems subject to random delay and packet drop," *IEEE Trans. Autom. Control*, vol. 53, no. 5, pp. 1311–1317, Jun. 2008.

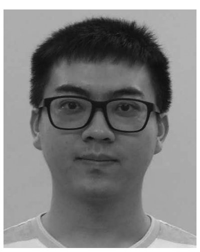
[56] B. Sinopoli *et al.*, "Kalman filtering with intermittent observations," *IEEE Trans. Autom. Control*, vol. 1, no. 9, pp. 1453–1464, Sep. 2004.

[57] S. Sun, L. Xie, W. Xiao, and Y. C. Soh, "Optimal linear estimation for systems with multiple packet dropouts," *Automatica*, vol. 44, no. 5, pp. 1333–1342, 2008.

[58] Z. Wang, D. W. C. Ho, and X. Liu, "Variance-constrained filtering for uncertain stochastic systems with missing measurements," *IEEE Trans. Autom. Control*, vol. 48, no. 7, pp. 1254–1258, Jul. 2003.

[59] Z. Wang, F. Yang, D. W. C. Ho, and X. Liu, "Robust finite-horizon filtering for stochastic systems with missing measurements," *IEEE Signal Process. Lett.*, vol. 12, no. 6, pp. 437–440, Jun. 2005.

[60] M. H. A. Davis and R. B. Vinter, "Stochastic modelling and control," *J. Oper. Res. Soc.*, vol. 37, no. 9, pp. 928–929, 1985.

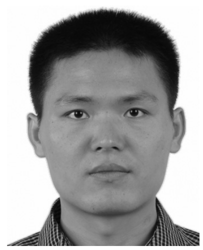


Hongru Ren received the B.E. degree in automation from the Department of Automation, University of Science and Technology of China, Hefei, China, in 2013, where he is currently pursuing the Ph.D degree in control science and engineering.

His current research interests include networked control systems and Kalman filtering.

Renquan Lu received the Ph.D. degree in control science and engineering from Zhejiang University, Hangzhou, China, in 2004.

He is currently a Professor with the School of Automation, Guangdong University of Technology, Guangzhou, China. He was supported by the National Science Fund for Distinguished Young Scientists of China in 2014, honored as the Distinguished Professor of Pearl River Scholars Program of Guangdong Province in 2015 and the Yangtze River Scholars Program by the Ministry of Education of China in 2017. His current research interests include complex systems, networked control systems, and nonlinear systems.



Junlin Xiong received the B.Eng. degree in mineral processing engineering and M.Sc. degree in mathematics from Northeastern University, Shenyang, China, in 2000 and 2003, respectively, and the Ph.D. degree in mechanical engineering from the University of Hong Kong, Hong Kong, in 2007.

From 2007 to 2010, he was a Research Associate with the University of New South Wales, Australian Defence Force Academy, Kensington, NSW, Australia. In 2010, he joined the University of Science and Technology of China, Hefei, China, where he is currently a Professor with the Department of Automation. His current research interests include negative imaginary systems, large-scale systems, and networked control systems.

Dr. Xiong is currently an Associate Editor for the *IET Control Theory and Application*.

Yong Xu received the Ph.D. degree in control science and engineering from Zhejiang University, Hangzhou, China, in 2014.

In 2013, he was a visiting internship student with the Department of Electronic and Computer Engineering, Hong Kong University of Science and Technology, Hong Kong, for five months. He is currently a Lecturer with the School of Automation, Guangdong University of Technology, Guangzhou, China. His current research interests include networked control systems, state estimation, and Markov jump systems.