On Stability and Convergence of Optimal Estimation for Networked Control Systems with Dual Packet Losses without Acknowledgment ⋆

Hong Lin [a,b] Hongye Su [a] Michael Z. Q. Chen [c] Zhan Shu [d] Renquan Lu [e] Zheng-Guang Wu [a]

[a] *State Laboratory of Industrial Control Technology, Institute of Cyber-Systems and Control, Zhejiang University, Yuquan Campus, Hangzhou, 310027, China.*

[b] *Department of Mechanical Engineering, The University of Hong Kong, Pokfulam, Hong Kong.*

[c] *School of Automation, Nanjing University of Science and Technology, Nanjing, Jiangsu 210094, China.*

[d] *Electro-Mechanical Engineering Group, Faculty of Engineering and The Environment, University of Southampton, Southampton SO17 1BJ, UK.*

[e] *School of Automation, Guangdong University of Technology, and Guangdong Key Laboratory of IoT Information Processing, Guangzhou 510006, China.*

Abstract

This paper studies the optimal state estimation problem for networked control systems with control and observation packet losses but without packet acknowledgment (ACK). The packet ACK is a signal sent by the actuator to inform the estimator whether control packets are lost or not. Systems with packet ACK are usually called transmission control protocol (TCP)-like systems, and those without ACK are named user datagram protocol (UDP)-like systems. For UDP-like systems, the optimal estimator is derived and it is consisted of an exponentially increasing number of terms. By developing an auxiliary estimator, it is shown that there exists a critical observation packet arrival rate determining the stability of the expected EC (EEC), and it is identical to its counterpart for TCP-like systems. It is revealed that whether there is packet ACK or not has no effect on the stability of the EEC. Furthermore, under some conditions the EEC converges exponentially.

Key words: Networked control systems; Optimal estimation; Stability; Packet losses; Packet acknowledgment

1 Introduction

Recently, significant attention has been paid to networked control systems (NCSs) as they bring numerous benefits, such as lower installation and maintenance costs, reduced network wiring, increased system flexibility, etc. However, the insertion of networks may make NCSs prone to network attacks [21] and cause some network-induced constraints, such as limited communication [3], signal quantization [4], and transmission packet losses [35]. There are two fundamental protocols in network communication for systems subject to packet losses. They are the transmission control protocol (TCP) and the user datagram protocol (UDP). For the TCP, the sending node retransmits lost data until it receives acknowledgment (ACK) from the receiving node.

⋆ The work is supported in part by the Fundamental Research Funds for the Central Universities (2017FZA5010), in part by Science Fund for Creative Research Groups of the National Natural Science Foundation of China (61621002), in part by National Natural Science Foundation of P.R. China (61374053, 61673339), in part by the China National Funds for Distinguished Young Scientists (61425009). Corresponding author: Hongye Su

Email addresses: `linhongzju@163.com` (Hong Lin), `hysu@iipc.zju.edu.cn` (Hongye Su), `mzqchen@outlook.com` (Michael Z. Q. Chen), `hustd8@gmail.com` (Zhan Shu), `rqlu@gdut.edu.cn` (Renquan Lu), `nashwzhg@126.com` (Zheng-Guang Wu).

Such retransmission mechanism guarantees the success of data transmission, but for NCSs with unreliable network communication, it would be difficult to implement the TCP, as the packet ACK cannot be transmitted without delay and random loss [7, 32, 16]. For the UDP, the ACK scheme is not used and thus no retransmission of lost data is required. The UDP, with a less transmission reliability, is able to provide more timely communication, and thus turns out to be a favorable choice for real-time NCSs [26]. The NCS without the packet ACK transmitted from the actuator to notice the estimator the status of control packet loss is usually called a UDP-like system, and the one with such packet ACK is called a TCP-like system (see Fig. 1). For the convenience of formulation, we denote by $\mathcal{S}^{u}_{UDP}$ a UDP-like system with only control packet losses, and by $\mathcal{S}^{uy}_{UDP}$ a UDP-like system with control and observation packet losses, i.e., dual packet losses. In this paper, we study the optimal estimator and its stability for the $\mathcal{S}^{uy}_{UDP}$ system.

For TCP-like systems, the optimal estimator is a time-varying Kalman filter, and the stability of the expected error covariance (EEC) has been studied in [31], in which it is pointed out that for an unstable system there exists a critical observation packet arrival rate, determining the boundedness of the EEC. NCSs with multiple packet losses were investigated in [14, 34]. Thereafter, the critical value and its upper/lower bound have been investigated in [23, 25]. For Markovian packet loss cases, significant results and techniques can be found in [2, 9, 36], and references therein.

For UDP-like systems, the literature on sub-optimal estimators is first reviewed as follows. For these systems, the linear minimum mean square error (MMSE) estimator was derived in [28], and other linear/non-linear estimators can be found in [15, 24, 32, 20]. The UDP-like system can also be viewed as a Markovian jump system (MJS) with unknown jump modes [18, 19]. Various computationally efficient estimators designed for MJSs with unknown jump modes, such as the interacting multiple model (IMM) estimator [13] and the probability hypothesis density (PHD) filter [5] also apply to UDP-like systems. However, these estimators are in fact sub-optimal. Analytic characterization on the stability and performance of these estimators is usually unavailable, and numerical approaches such as the Monte Carlo method are often used [13].

For the optimal estimator, it is shown in [19] that for the $\mathcal{S}^{u}_{UDP}$ system, it contains exponentially increasing terms, and the EEC is bounded under bounded control inputs if the $\mathcal{S}^{u}_{UDP}$ system is detectable. However, the condition for the stability of the optimal estimator for the $\mathcal{S}^{uy}_{UDP}$ system, to our best knowledge, is still unknown, due to some challenging issues: 1) A random variable γ_k, not presented in the $\mathcal{S}^{u}_{UDP}$ system, occurs not only in the Riccati equation (2e) but also in the power term (15b). As a result, equations for EC are more complicated than that in [19], and the summation part with exponentially increasing terms in (17d) will become unbounded, making analysis of the stability difficult. 2) Existing methods are not applicable to the $\mathcal{S}^{uy}_{UDP}$ system. The sequential Monte Carlo method, a simulation method, has known to be a practical tool to evaluate the EEC [6]. The hybrid approach developed in [13] is an efficient off-line algorithm for approximately computing EEC with finite mixing terms. The techniques proposed in [29, 5] to analyze the stability of the IMM and the PHD estimators, the good approximations for the optimal estimator, can be employed to approximately study the stability of the optimal estimator for UDP-like systems. However, these aforementioned methods merely render the stability and convergence results in an experimental or approximate way, and cannot determine the stability of EEC in the desired theoretical view. In [19], an auxiliary estimator method was developed to study stability of EECs for the $\mathcal{S}^{u}_{UDP}$ system. However, all the observations $\{y_1, \ldots, y_k\}$ are required in constructing this auxiliary estimator, and thus this method is not applicable due to random losses of observations in the $\mathcal{S}^{uy}_{UDP}$ system. In [18], another type of auxiliary estimator was constructed for the $\mathcal{S}^{u}_{UDP}$ system, but the relationship between the optimal and the auxiliary estimators is not established.

For UDP-like systems, the optimal estimator and its stability and convergence are studied in this paper. Main results and contributions are summarized as follows:

1) We obtain the optimal estimator for UDP-like system with dual packet losses, which is consisted of an exponentially increasing number of terms.

2) We show that the stability of the EEC is only determined by the observation packet arrival rate, and is independent of the control packet arrival rate. That is, there is a critical value for a given UDP-like system, and the EEC is stable if the observation packet arrival rate is greater than this critical value. Moreover, this critical value is identical to its counterpart for the TCP-like system corresponding this UDP-like system. It reveals the fact whether there is packet ACK or not does not affect the stability of the optimal estimator.

3) We show that the EEC, although containing exponentially increasing terms, converges if there is no observation packet loss and control inputs eventually tend to zero.

The paper is organized as follows. The system and problems are formulated in Section 2. The optimal estimator for UDP-like systems is obtained in Section 3. An auxiliary estimator is constructed in Section 4. The conditions on the stability and convergence of the optimal estimator are established in Section 5. In Section 6, numerical examples are given to illustrate the obtained results. Conclusions are presented in Section 7. The proofs of lemmas are presented in Appendix.

Notations:

- $\mathbb{P}(\cdot)$ denotes the probability measure.
- $p(\cdot)$ and $p(\cdot|\cdot)$ denote the probability density function (pdf) and the conditional pdf, respectively.
- $\mathcal{N}(\mu, P)$ denotes a Gaussian pdf with mean μ and covariance P. Both $x \sim \mathcal{N}(\mu, P)$ and $p(x) = \mathcal{N}(\mu, P)$ mean the pdf of the random variable x is a Gaussian pdf with mean μ and covariance P. $\mathcal{N}_x(\mu, P)$ is used to emphasize that the random variable of the pdf $\mathcal{N}(\mu, P)$ is x.
- $\mathbb{E}_x[\cdot]$ and $\underset{x}{\text{cov}}(\cdot)$ denote the probability expectation and the covariance with respect to x, respectively.
- $\|\cdot\|$ denotes the norm. Specifically, for a vector, $\|\cdot\|$ denotes the 2-norm; For a matrix, $\|\cdot\|$ denotes the spectral norm, i.e., the maximum singular value.
- $(\cdot)'$ denotes the transpose of a matrix or vector.
- $(\cdot)_I^2$ with the identity matrix I means $(\cdot)(\cdot)'$.
- $(\cdot)_2$ denotes the binary representation, e.g., $(101)_2 = 5$.
- $\mathbb{N}$, $\mathbb{Z}$, and $\mathbb{R}$ denote the natural number, the integer, and the real number, respectively.

2 System setup and problem formulation

Consider the following system:

$$\begin{aligned} x_{k+1} &= Ax_k + \nu_k Bu_k + \omega_k \\ y_k &= \begin{cases} Cx_k + \upsilon_k, & \text{for } \gamma_k = 1 \\ \emptyset, & \text{for } \gamma_k = 0 \end{cases} \end{aligned} \tag{1}$$

where $x_k \in \mathbb{R}^n$ is the system state, $u_k \in \mathbb{R}^q$ is the control input, and $y_k \in \mathbb{R}^p$ is the observation. ω_k and υ_k are Gaussian noises with zero means and covariances $Q \geq 0$ and $R > 0$, respectively. γ_k and ν_k are independent identically distributed (i.i.d.) Bernoulli random variables with $\mathbb{P}(\gamma_k = 1) = \gamma$ and $\mathbb{P}(\nu_k = 1) = \nu$. They describe the packet losses in the sensor-to-estimator (S/E) and the controller-to-actuator (C/A) channels, respectively. ν and γ are also known as packet arrival rates of control input and observation, respectively. This paper does not involve the design of controller, and only assume that u_k is bounded [1] and deterministic.

Assumption 1 *$x_0 \sim \mathcal{N}(\bar{x}_0, P_0)$, ω_k, υ_k, ν_k, and γ_k are mutually independent. The pair (A, C) is observable, and the pair $(A, Q^{1/2})$ is controllable.*

Define information sets $\mathcal{I}_k \triangleq \{\mathrm{y}^k, \gamma^k, \nu^{k-1}\}$ and $\mathcal{G}_k \triangleq \{\mathrm{y}^k, \gamma^k\}$ with $\mathcal{I}_0 = \mathcal{G}_0 \triangleq \emptyset$ (empty set), where $\nu^k \triangleq \{\nu_0, \cdots, \nu_k\}$, $\gamma^k \triangleq \{\gamma_1, \cdots, \gamma_k\}$, and $\mathrm{y}^k \triangleq \{y_1, \cdots, y_k\}$.

Definition 1 *(UDP-like system and TCP-like system)* **The UDP-like system**, *i.e., $\mathcal{S}_{UDP}^{uy}$, is the system described in (1) with the value of ν_k unknown to the estimator.* **The TCP-like system**, *denoted by $\mathcal{S}_{TCP}^{uy}$, is the one described in (1) with the value of ν_k available to the estimator. We also call the $\mathcal{S}_{TCP}^{uy}$ system* **the TCP-like system corresponding to $\mathcal{S}_{UDP}^{uy}$**.

[1] In practical control systems, physical actuators are subject to saturation, and thus inputs always have maximum and minimum values [8].

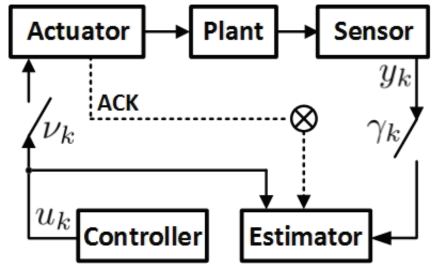

Fig. 1. The UDP-like system. The symbol $\otimes$ is used to emphasize that there is no acknowledgment signal from the actuator to the estimator.

Definition 2 *(Optimal estimation) An estimation of system state, denoted by $\hat{x}_k$, is said to be* **optimal** *in the minimum mean square error (MMSE) sense, if it minimizes $\mathbb{E}[\|x_k - \hat{x}_k|\mathcal{G}_k\|^2]$.*

It is known in [1] that the desired optimal MMSE estimation $\hat{x}_k$ is $\mathbb{E}[x_k|\mathcal{G}_k]$. Define $\bar{x}_k \triangleq \mathbb{E}[x_k|\mathcal{G}_{k-1}]$ as the prediction of system state. Denote the prediction and the estimation ECs by $\bar{P}_k \triangleq \mathbb{E}[(x_k)_I^2|\mathcal{G}_{k-1}]$ and $P_k \triangleq \mathbb{E}[(x_k)_I^2|\mathcal{G}_k]$, respectively. For the TCP-like system, define the predicted and the estimated system states as $\bar{x}_k^t \triangleq \mathbb{E}[x_k|\mathcal{I}_{k-1}]$ and $\hat{x}_k^t \triangleq \mathbb{E}[x_k|\mathcal{I}_k]$, respectively. Denote the prediction and the estimation ECs by $\bar{P}_k^t \triangleq \mathbb{E}[(x_k)_I^2|\mathcal{I}_{k-1}]$ and $P_k^t \triangleq \mathbb{E}[(x_k)_I^2|\mathcal{I}_k]$, respectively. [2]

Definition 3 *(Stability) The error covariance $\bar{P}_k$ is said to be* **stable** *in the mean sense (or stable for short), if $\mathbb{E}[\bar{P}_k]$ is bounded, i.e., $\sup_k \mathbb{E}[\bar{P}_k] < \infty$, for $k \in \mathbb{N}$. Then the optimal estimator is said to be* **stable** *in the mean sense (or stable for short), if $\bar{P}_k$ is stable.*

For UDP-like systems, our aim is to solve the following problems:

- Derive the optimal estimator and find out the relationship between it and the optimal estimator for TCP-like systems.
- Determine conditions for the stability of the optimal estimator, and analyze the impact of ACK signals and packet arrival rates.
- Find out whether there exist some conditions under which the EEC is convergent.

Preliminaries:

(i) For TCP-like systems, the optimal estimator has been obtained in [28] and is given as follows:

$$\bar{x}_{k+1}^t = A\hat{x}_k^t + \nu_k Bu_k \tag{2a}$$

[2] The superscript t of these estimates means TCP-like and is used to distinguish from those for UDP-like systems.

$$\bar{P}_{k+1}^t = AP_k^t A' + Q \quad (2b)$$
$$K_{k+1} = \bar{P}_{k+1}^t C' \left(C\bar{P}_{k+1}^t C' + R\right)^{-1} \quad (2c)$$
$$\hat{x}_{k+1}^t = \bar{x}_{k+1}^t + \gamma_{k+1} K_{k+1}(y_{k+1} - C\bar{x}_{k+1}^t) \quad (2d)$$
$$P_{k+1}^t = \bar{P}_{k+1}^t - \gamma_{k+1} K_{k+1} C\bar{P}_{k+1}^t \quad (2e)$$

with $P_0^t = P_0$.

Define a function

$$g(\gamma, P) = APA' - \gamma APC'(CPC' + R)^{-1}CPA' + Q.$$

Substituting (2e) into (2b) yields $\bar{P}_{k+1}^t = g(\gamma_k, \bar{P}_k^t)$. If $\gamma_k = 1$ for all $k \in \mathbb{N}$ then $\bar{P}_{k+1}^t = g(1, \bar{P}_k^t)$ is the standard Riccati equation.

(ii) Let X be a random variable with a Gaussian mixture pdf, i.e., $p(X) = \sum_{i=1}^N \xi_i \mathcal{N}(\mu_i, B)$. Its mean $\hat{X} = \mathbb{E}[X]$ and covariance $P_X = \mathbb{E}[(X - \hat{X})_I^2]$ can be calculated as follows ([1, p.213]):

$$\hat{X} = \sum_{i=1}^N \xi_i \mu_i \quad \text{and} \quad P_X = B + \sum_{i=1}^N \xi_i (\mu_i - \hat{X})_I^2. \quad (3)$$

(iii) Let $X \sim \mathcal{N}(m, P)$, $Y \sim \mathcal{N}(0, P_Y)$, and $Z = CX + Y$. Assume X and Y are independent. Then ([33, p. 88, (3.7), (3.8), (3.13), (3.14); p.98])

$$p(Z) = \mathcal{N}(Cm, CPC' + P_Y) \quad (4)$$
$$p(X|Z) = \mathcal{N}(m + K(Z - Cm), (I - KC)P), \quad (5)$$

where $K = PC'(CPC' + P_Y)^{-1}$.

(iv) Given two random variables Z and J, and a function $\hbar(Z, J)$. Then we have ([27, p. 117, p.119, p.180])

$$\text{cov}(Z) = \mathbb{E}[Z^2] - \left(\mathbb{E}[Z]\right)^2 \quad (6)$$
$$\mathbb{E}_{Z,J}\left[\hbar(Z, J)\right] = \mathbb{E}_J\left[\mathbb{E}_Z[\hbar(Z, J)|J]\right] \quad (7)$$
$$\text{cov}(Z) = \mathbb{E}\left[\text{cov}(Z|J)\right] + \text{cov}\left(\mathbb{E}[Z|J]\right). \quad (8)$$

3 Optimal estimator for the UDP-like system

In this section, we study the conditional pdfs $p(x_k|\mathcal{G}_{k-1})$ and $p(x_k|\mathcal{G}_k)$, and then derive the optimal estimator.

3.1 Description of the event of control packet losses

We first introduce the random events of control packet losses, which are described by the sequence of random variables $\{\nu_k\}$. At time k, the event takes the form $\{\nu_k, \ldots, \nu_0\}$ with $\nu_j \in \{0, 1\}$ for $0 \le j \le k$. The probability space denoted by Ω_k contains 2^{k+1} such elementary events. For each binary-valued sequence $(\nu_k \cdots \nu_0)$, we can determine a unique integer i by

$$i = \rho(\nu_k \cdots \nu_0) \triangleq (\nu_k \cdots \nu_0)_2 + 1. \quad (9)$$

Consequently, the event of control packet losses can be defined via ρ as follows:

$$\theta_k^{(i)} \triangleq \{\nu_k, \ldots, \nu_0 | i = \rho(\nu_k \cdots \nu_0)\}, 1 \le i \le 2^{k+1}. \quad (10)$$

For instance, $\theta_1^{(1)} = \{0, 0\}$ due to $1 = \rho(00) = (00)_2 + 1$, which means $\theta_1^{(1)} = \{\nu_1 = 0, \nu_0 = 0\}$. It is easy to verify that for $1 \le i \le 2^k$,

$$\theta_k^{(i)} = \{\nu_k = 0, \theta_{k-1}^{(i)}\} \text{ and } \theta_k^{(i+2^k)} = \{\nu_k = 1, \theta_{k-1}^{(i)}\}. \quad (11)$$

3.2 Probability density function of x_k

By the total probability law, the conditional pdfs of x_k under $\mathcal{G}_{k-1}$ and $\mathcal{G}_k$ can be presented as follows:

$$p(x_k|\mathcal{G}_{k-1}) = \sum_{i=1}^{2^k} p(x_k|\theta_{k-1}^{(i)}, \mathcal{G}_{k-1}) p(\theta_{k-1}^{(i)}|\mathcal{G}_{k-1}) \quad (12a)$$
$$p(x_k|\mathcal{G}_k) = \sum_{i=1}^{2^k} p(x_k|\theta_{k-1}^{(i)}, \mathcal{G}_k) p(\theta_{k-1}^{(i)}|\mathcal{G}_k). \quad (12b)$$

Lemma 1 *At time k with $1 \le i \le 2^k$,*

$$p(x_k|\theta_{k-1}^{(i)}, \mathcal{G}_{k-1}) = \mathcal{N}(\bar{z}_k^{(i)}, \bar{P}_k^t) \quad (13a)$$
$$p(x_k|\theta_{k-1}^{(i)}, \mathcal{G}_k) = \mathcal{N}(\hat{z}_k^{(i)}, P_k^t), \quad (13b)$$

where

$$\bar{z}_k^{(i)} = \begin{cases} A\hat{z}_{k-1}^{(i)}, \ \text{for } 1 \le i \le 2^{k-1} \\ A\hat{z}_{k-1}^{(i-2^{k-1})} + Bu_{k-1}, \ \text{for } 2^{k-1}+1 \le i \le 2^k \end{cases} \quad (14a)$$
$$\hat{z}_k^{(i)} = \bar{z}_k^{(i)} + \gamma_k K_k (y_k - C\bar{z}_k^{(i)}) \quad (14b)$$

with $\hat{z}_0^{(1)} = \bar{x}_0$. K_k, $\bar{P}_k^t$, and P_k^t are calculated by (2).

Lemma 2 *At time k with $1 \le i \le 2^k$, define $\bar{\alpha}_k^{(i)} \triangleq p(\theta_{k-1}^{(i)}|\mathcal{G}_{k-1})$ and $\hat{\alpha}_k^{(i)} \triangleq p(\theta_{k-1}^{(i)}|\mathcal{G}_k)$. Then*

$$\bar{\alpha}_k^{(i)} = \begin{cases} \bar{\nu}\hat{\alpha}_{k-1}^{(i)}, & \text{for } 1 \le i \le 2^{k-1} \\ \nu\hat{\alpha}_{k-1}^{(i-2^{k-1})}, & \text{for } 2^{k-1}+1 \le i \le 2^k \end{cases} \quad (15a)$$

$$\hat{\alpha}_k^{(i)} = \left(\frac{\phi_k^{(i)}}{\sum_{j=1}^{2^k} \phi_k^{(j)} \bar{\alpha}_k^{(j)}}\right)^{\gamma_k} \bar{\alpha}_k^{(i)} \tag{15b}$$

with $\hat{\alpha}_0^{(1)} = 1$, *where* $\bar{\nu} \triangleq 1-\nu$, $\phi_k^{(i)} \triangleq p(y_k|\theta_{k-1}^{(i)}, \mathcal{G}_{k-1}) = \mathcal{N}_{y_k}(C\bar{z}_k^{(i)}, P_k^Y)$, $P_k^Y \triangleq C\bar{P}_k^t C' + R$. $\bar{P}_k^t$, K_k, *and* P_k^t *are calculated by (2).*

Substituting these four conditional pdfs in Lemmas 1 and 2 into (12) yields

$$p(x_k|\mathcal{G}_{k-1}) = \sum_{i=1}^{2^k} \bar{\alpha}_k^{(i)} \mathcal{N}(\bar{z}_k^{(i)}, \bar{P}_k^t) \tag{16a}$$

$$p(x_k|\mathcal{G}_k) = \sum_{i=1}^{2^k} \hat{\alpha}_k^{(i)} \mathcal{N}(\hat{z}_k^{(i)}, P_k^t). \tag{16b}$$

Remark 1 *For the TCP-like system in (1), the linear system is driven by Gaussian noises, and* ν_k *is known for the estimator. Therefore, the pdf of* x_k *is Gaussian; For the UDP-like system in (1), the system is driven not only by Gaussian noises* ω_k *but also by a Bernoulli random variable* ν_k*, and this accounts for the Gaussian mixture pdf of* x_k*.*

3.3 Optimal estimator for UDP-like systems

Theorem 1 (Optimal estimator) *For the UDP-like system in (1), the optimal estimator of* x_k *is given as follows:*

$$\hat{x}_k = \sum_{i=1}^{2^k} \hat{\alpha}_k^{(i)} \hat{z}_k^{(i)} \tag{17a}$$

$$P_k = P_k^t + \sum_{i=1}^{2^k} \hat{\alpha}_k^{(i)} (\hat{z}_k^{(i)} - \hat{x}_k)_I^2 \tag{17b}$$

$$\bar{x}_{k+1} = \sum_{i=1}^{2^{(k+1)}} \bar{\alpha}_{k+1}^{(i)} \bar{z}_{k+1}^{(i)} \tag{17c}$$

$$\bar{P}_{k+1} = \bar{P}_{k+1}^t + \sum_{i=1}^{2^{(k+1)}} \bar{\alpha}_{k+1}^{(i)} (\bar{z}_{k+1}^{(i)} - \bar{x}_{k+1})_I^2, \tag{17d}$$

where $\{\bar{z}_k^{(i)}, \hat{z}_k^{(i)}\}$, $\{\bar{\alpha}_k^{(i)}, \hat{\alpha}_k^{(i)}\}$, *and* $\{\bar{P}_k^t, P_k^t\}$ *can be computed by (14), (15), and (2), respectively.*

Proof: For Gaussian mixture pdfs, by applying (3) to $p(x_k|\mathcal{G}_{k-1})$ and $p(x_k|\mathcal{G}_k)$, the optimal estimator can be readily obtained as in (17). □

4 Construction of an auxiliary estimator

Since the coefficient $\hat{\alpha}_k^{(i)}$ in (17b) is so complex that (17) cannot be expressed as recursive equations. Consequently, the conventional Riccati equation approach is not applicable to analyzing the stability of $\bar{P}_k$. In the following, an auxiliary estimator is developed to analyze $\bar{P}_k$.

4.1 Construction of an auxiliary system state

Clearly, x_k is a random variable, and $p(x_k|\mathcal{G}_{k-1})$ characterizes the pdf of x_k conditioned on $\mathcal{G}_{k-1}$. From $p(x_k|\mathcal{G}_{k-1})$, we will construct a well-defined pdf $p(x_k^a|\mathcal{G}_{k-1})$ in the following. As a pdf, $p(x_k^a|\mathcal{G}_{k-1})$ is associated with some random variable in a certain way. Similar to $p(x_k|\mathcal{G}_{k-1})$, $p(x_k^a|\mathcal{G}_{k-1})$ in fact characterizes the pdf of the associated random variable conditioned on $\mathcal{G}_{k-1}$. We denote the associated random variable by x_k^a, and call it an auxiliary system state.

From (13a), it is clear that $p(x_k|\theta_{k-1}^{(i)}, \mathcal{G}_{k-1})$ is a Gaussian pdf. Based on it, we define a function $p(x_k^a|\theta_{k-1}^{(i)}, \mathcal{G}_{k-1})$ by replacing the symbol x_k in $p(x_k|\theta_{k-1}^{(i)}, \mathcal{G}_{k-1})$ with x_k^a. As $p(x_k|\theta_{k-1}^{(i)}, \mathcal{G}_{k-1})$ is a deterministic function, $p(x_k^a|\theta_{k-1}^{(i)}, \mathcal{G}_{k-1})$ is *deterministic and well-defined*.

Note that the function ρ defined in (9) is a bijection. For each i, there is a unique sequence denoted by $(\nu_k^{(i)} \cdots \nu_0^{(i)})$ such that $i = \rho(\nu_k^{(i)} \cdots \nu_0^{(i)})$. From the definition of $\theta_k^{(i)}$ in (10) and the mutual independence of $\{\nu_k\}$, it follows that $p(\theta_k^{(i)}) = \prod_{j=0}^k p(\nu_j = \nu_j^{(i)})$. For each i, $p(\theta_k^{(i)})$ is a *deterministic* value. Moreover, it is easy to verify that $\sum_{i=1}^{2^k} p(\theta_{k-1}^{(i)}) = 1$. Based on $p(\theta_{k-1}^{(i)})$ and $p(x_k^a|\theta_{k-1}^{(i)}, \mathcal{G}_{k-1})$, we define a function

$$p(x_k^a|\mathcal{G}_{k-1}) \triangleq \sum_{i=1}^{2^k} p(x_k^a|\theta_{k-1}^{(i)}, \mathcal{G}_{k-1}) p(\theta_{k-1}^{(i)}). \tag{18}$$

Since $p(x_k^a|\theta_{k-1}^{(i)}, \mathcal{G}_{k-1})$ and $p(\theta_k^{(i)})$ are deterministic and well defined, $p(x_k^a|\mathcal{G}_{k-1})$ is *deterministic and well defined* as well. The $p(x_k^a|\mathcal{G}_{k-1})$ defined above is not only a function but also a pdf, as it satisfies two conditions: $p(x_k^a|\mathcal{G}_{k-1}) \geq 0$ and $\int_{-\infty}^{\infty} p(x_k^a|\mathcal{G}_{k-1}) \mathrm{d}x_k^a = 1$.

4.2 The auxiliary estimator and its properties

We first derive the conditional pdf of x_k^a, and then compute its estimator.

Lemma 3 $p(x_k^a|\mathcal{G}_{k-1})$ *defined in (18) can be presented* *as follows:*

$$p(x_k^a|\mathcal{G}_{k-1}) = \sum_{i=1}^{2^k} \tilde{\alpha}_k^{(i)} \mathcal{N}(\bar{z}_k^{(i)}, \bar{P}_k^t), \tag{19}$$

where

$$\tilde{\alpha}_k^{(i)} = \begin{cases} \bar{\nu}\tilde{\alpha}_{k-1}^{(i)}, & \text{for } 1 \le i \le 2^{k-1} \\ \nu\tilde{\alpha}_{k-1}^{(i-2^{k-1})}, & \text{for } 2^{k-1}+1 \le i \le 2^k \end{cases} \tag{20}$$

with $\tilde{\alpha}_0^{(1)} = 1$. $\bar{z}_k^{(i)}$ *evolves in the same way as (14), and* $\bar{P}_k^t$ *can be computed by (2).*

Applying (3) to $p(x_k^a|\mathcal{G}_{k-1})$ in (19) yields the estimator of x_k^a as follows:

$$\bar{x}_k^a = \sum_{i=1}^{2^k} \tilde{\alpha}_k^{(i)} \bar{z}_k^{(i)} \tag{21a}$$

$$\bar{P}_k^a = \bar{P}_k^t + \mathcal{M}_k, \tag{21b}$$

where $\mathcal{M}_k \triangleq \sum_{i=1}^{2^k} \tilde{\alpha}_k^{(i)} (\bar{z}_k^{(i)} - \bar{x}_k^a)_I^2$. $\bar{x}_k^a$ and $\bar{P}_k^a$ are the mean and the error covariance of x_k^a, respectively.

We will show that $\bar{P}_k^a$ is an upper bound for $\bar{P}_k$ in the mean sense. From (17d) and the computation of $\bar{\alpha}_{k+1}^{(i)}$ and $\bar{z}_{k+1}^{(i)}$, it follows that $\bar{P}_k$ contains random variables $\{\gamma^{k-1}, y^{k-1}\} = \mathcal{G}_{k-1}$. Thus, in the following lemma the mathematical expectation is taken with respect to $\mathcal{G}_{k-1}$.

Lemma 4 $\underset{\mathcal{G}_{k-1}}{\mathbb{E}}[\bar{P}_k^t] \le \underset{\mathcal{G}_{k-1}}{\mathbb{E}}[\bar{P}_k] \le \underset{\mathcal{G}_{k-1}}{\mathbb{E}}[\bar{P}_k^a]$.

5 Stability and convergence of the optimal estimator for UDP-like systems

5.1 Stability of the optimal estimator

Theorem 2 (Stability) *Consider the UDP-like system in (1) with unstable A. Then*

(*a*) $\mathbb{E}[\bar{P}_k]$ *is bounded if and only if* $\mathbb{E}[\bar{P}_k^t]$ *is bounded, where* $\bar{P}_k^t$ *and* $\bar{P}_k$ *are computed by (2) and (17).*

(*b*) *There is a critical value* $\lambda_c \in [0, 1)$ *such that*

$$\text{If } 0 \le \gamma < \lambda_c, \text{ then } \lim_{k\to\infty} \mathbb{E}[\bar{P}_k] = +\infty, \exists P_0 \ge 0$$

$$\text{If } \lambda_c < \gamma \le 1, \text{ then } \sup_k \mathbb{E}[\bar{P}_k] \le \mathcal{H}_{P_0}, \forall P_0 \ge 0$$

where $\mathcal{H}_{P_0}$ *is a constant matrix and depends on* P_0.

(*c*) *The critical value* λ_c *for the UDP-like system is identical to that for its corresponding TCP-like system.*

Before proving Theorem 2, we give some lemmas as follows. For notational simplicity, let $\mathbb{A}_k \triangleq (A - \gamma_k A K_k C)$ and $U_k \triangleq \bar{\nu}\nu B u_k u_k' B'$. Then define

$$\mathcal{A}_k^{(i)} \triangleq \begin{cases} \mathbb{A}_k \cdots \mathbb{A}_i, & \text{for } 1 \le i \le k \\ I, & \text{for } i = k+1. \end{cases} \tag{22}$$

Lemma 5 *For* $\mathcal{M}_k$ *in (21b),*

$$\mathcal{M}_{k+1} = \mathbb{A}_k \mathcal{M}_k \mathbb{A}_k' + U_k, \text{ with } \mathcal{M}_0 = 0 \tag{23}$$

$$\mathcal{M}_k = \sum_{i=1}^{k} \mathcal{A}_{k-1}^{(i)} U_{i-1} \mathcal{A}_{k-1}^{(i)}{}'. \tag{24}$$

Lemma 6 *If* $U_i \le Q$ *and* $U_0 \le \bar{P}_1^t$*, then* $\mathcal{M}_k \le \bar{P}_k^t$.

Proof of Theorem 2: Proof of part (a). From (17d), it follows that if $\bar{P}_k^t$ diverges, so does $\bar{P}_k$, which proves the necessity. To prove the sufficient condition, we consider the term U_k in $\mathcal{M}_k$. As assumed in Theorem 2 that u_k is bounded, we denote its bound by $\sqrt{\mu}$, i.e., $\|u_k\| \le \sqrt{\mu}$. Then we have $u_k u_k' \le \mu I$. Due to $\nu + \bar{\nu} = 1$, $\nu\bar{\nu} \le 1/4$. Then $U_k \triangleq \bar{\nu}\nu B u_k u_k' B' \le \frac{\mu BB'}{4}$. There is a real number $\kappa > 0$ such that both $\frac{\mu BB'}{4\kappa} \le \bar{P}_1^t$ and $\frac{\mu BB'}{4\kappa} \le Q$ are satisfied. That is, we can choose a real number $\kappa > 0$ such that, for the U_i in (24), $U_0/\kappa \le \bar{P}_1^t$ and $U_i/\kappa \le Q$ for $1 \le i \le k-1$. By (21b) and (24), $\bar{P}_k^a = \bar{P}_k^t + \sum_{i=1}^k \mathcal{A}_{k-1}^{(i)} U_{i-1} \mathcal{A}_{k-1}^{(i)}{}' = \bar{P}_k^t + \kappa \sum_{i=1}^k \mathcal{A}_{k-1}^{(i)} \frac{U_{i-1}}{\kappa} \mathcal{A}_{k-1}^{(i)}{}'$. By Lemma 6, $\sum_{i=1}^k \mathcal{A}_{k-1}^{(i)} \frac{U_{i-1}}{\kappa} \mathcal{A}_{k-1}^{(i)}{}' \le \bar{P}_k^t$. Thus, we have $\bar{P}_k^a \le \bar{P}_k^t + \kappa\bar{P}_k^t = (1+\kappa)\bar{P}_k^t$. From the hypothesis that $\mathbb{E}[\bar{P}_k^t]$ is bounded, it follows that $\mathbb{E}[\bar{P}_k^a]$ is stable (i.e., bounded). Using Lemma 4, we have that $\mathbb{E}[\bar{P}_k]$ is stable. Part (a) is proved.

Proof of part (b). According to Theorem 2 in [31], for the TCP-like system in (1), there is a critical value λ such that if $0 \le \gamma < \lambda$, $\lim_{k\to\infty} \mathbb{E}[\bar{P}_k^t] = +\infty$ for some $P_0 \ge 0$. Thus, by Theorem 2 (a), we have $\mathbb{E}[\bar{P}_k] = +\infty$ for this $P_0 \ge 0$, if $0 \le \gamma < \lambda$. Analogously, according to Theorem 2 in [31], if $\lambda < \gamma \le 1$, then $\mathbb{E}[\bar{P}_k^t] \le H_{P_0}$, $\forall k$ and $\forall P_0 \ge 0$, where H_{P_0} is a constant matrix and depends on the initial value P_0. Hence, by Theorem 2 (a), $\mathbb{E}[\bar{P}_k] \le (1+\kappa)\mathbb{E}[\bar{P}_k^t] \le (1+\kappa)H_{P_0} \triangleq \mathcal{H}_{P_0}$, for $\forall k$ and $\forall P_0 \ge 0$, if $\lambda_c < \gamma \le 1$. Consequently, the critical value λ determines the boundedness of $\mathbb{E}[\bar{P}_k^t]$ and $\mathbb{E}[\bar{P}_k]$, which proves the existence of λ_c by letting $\lambda_c = \lambda$.

Part (c) is proved by noting that λ_c is the critical value λ for the corresponding TCP-like system. □

Remark 2 *From Theorem 2 (a) and (c), it follows that most of the results on the stability and critical value of the optimal estimator for TCP-like systems are applicable to UDP-like systems.*

Remark 3 *It follows from Theorem 2(a) that the optimal estimator for the $\mathcal{S}^{uy}_{UDP}$ system is stable if and only if that for the $\mathcal{S}^{uy}_{TCP}$ system is stable, which implies that whether there exists the ACK for control packet losses or not has no effect on the stability of the optimal estimator. From a control perspective, it is reported in [10] that for a class of plants, the ACK has no effect on the critical value for stabilizing the systems.*

5.2 Convergence of the optimal estimator

For TCP-like systems with $0 < \gamma < 1$, $\mathbb{E}[\bar{P}^t_k]$ is not necessarily convergent [28, 31]. Thus, for UDP-like systems with $0 < \gamma < 1$, $\mathbb{E}[\bar{P}_k]$ in (17d)—consisting of such unconvergent $\mathbb{E}[\bar{P}^t_k]$ and a summation part with exponentially increasing number of terms—is not necessarily convergent either. To guarantee the convergence of $\mathbb{E}[\bar{P}_k]$, the first condition required would be $\gamma_k = 1$ for all $k \in \mathbb{N}$, since when $\gamma_k \equiv 1$, $\bar{P}^t_k$, recursively computed by the standard Riccati equation, is convergent under Assumption 1. However, setting $\gamma_k \equiv 1$ is not sufficient to render the convergence of $\mathbb{E}[\bar{P}_k]$. We give conditions for convergence of $\mathbb{E}[\bar{P}_k]$ as follows.

Theorem 3 (Convergence) *Consider the UDP-like system in (1) without observation packet losses.*

(*a*) *If the control input $u_k \to 0$, then $\mathbb{E}[\bar{P}_k]$ converges to $\bar{P}$, where $\bar{P}$ is the solution of the standard algebraic Riccati equation (ARE), i.e., $\bar{P} = g(1, \bar{P})$.*
(*b*) *If U_k satisfies $U_k \le \rho_u e^{-\eta_u(k-m_u)} I$ with $\rho_u, \eta_u > 0$ and $m_u \in \mathbb{Z}$, then $\mathbb{E}[\bar{P}_k]$ exponentially converges to $\bar{P}$.*

To prove Theorem 3, some preliminaries are given as follows.

Denote by $\mathcal{S}^u_{TCP}$ the TCP-like system in (1) with only control packet loss, and by $\mathcal{S}_{LTI}$ the system in (1) without observation and control packet loss, i.e., the classic linear time-invariant (LTI) system: $x_{k+1} = Ax_k + Bu_k + \omega_k$ and $y_k = Cx_k + \upsilon_k$. For this LTI system, we denote the optimal state prediction and the prediction error by $\bar{x}^l_k$ and $\bar{e}^l_k = x_k - \bar{x}^l_k$, respectively. By the Kalman filter, $\bar{x}^l_k$ and $\bar{e}^l_k$ can be calculated as follows [30, pp.131]:

$$\begin{aligned} \bar{x}^l_{k+1} &= A(I - K_kC)\bar{x}^l_k + AK_kCx_k + AK_k\upsilon_k + Bu_k \\ \bar{e}^l_{k+1} &= \mathbb{A}_k\bar{e}^l_k - AK_k\upsilon_k + \omega_k, \end{aligned} \tag{25}$$

where $\mathbb{A}_k = A(I - K_kC)$ due to $\gamma_k = 1$. K_k can be computed by (2b), (2c), and (2e) with $\gamma_k = 1$. The homogeneous part of (25) is $\bar{e}^l_{k+1} = \mathbb{A}_k\bar{e}^l_k$. The transition matrix of the prediction error equation from $\bar{e}^l_m$ to $\bar{e}^l_{k+1}$ is defined as $\mathbb{A}_k \cdots \mathbb{A}_m$. That is, $\bar{e}^l_{k+1} = \mathbb{A}_k \cdots \mathbb{A}_m\bar{e}^l_m$. Note that the transition matrix is just $\mathcal{A}^{(m)}_k$ defined in (22) with $\gamma_j = 1$, $1 \le j \le k$.

Lemma 7 *For the $\mathcal{S}^u_{TCP}$ system, there exist positive constants ρ and η such that $\|\mathcal{A}^{(m)}_k\|^2 < \rho e^{-\eta(k-m)}$, where $\mathcal{A}^{(m)}_k$ is defined in (22).*

Proof of Theorem 3: Proof of Part (a): Note that $U_k = \bar{\nu}\nu Bu_ku'_kB' \to 0$ when $u_k \to 0$. In the sequel, we show that $\mathcal{M}_{k+1} \to 0$ when $U_k \to 0$. Since $U_k \to 0$, from the knowledge of limit theory, it follows that for any $\varepsilon_u > 0$, there is an integer N_u such that $U_k < \varepsilon_u I$ for $k > N_u$. From (24),

$$\begin{aligned} \mathcal{M}_{k+1} &= \sum_{i=1}^{k+1} \mathcal{A}^{(i)}_k U_{i-1} \mathcal{A}^{(i)\prime}_k \\ &\le \sum_{i=1}^{N_u+1} \|\mathcal{A}^{(i)}_k\|^2 U_{i-1} + \sum_{i=N_u+2}^{k+1} \|\mathcal{A}^{(i)}_k\|^2 U_{i-1} \\ &\overset{(a)}{\le} \mu I \sum_{i=1}^{N_u+1} \|\mathcal{A}^{(i)}_k\|^2 + \varepsilon_u I \sum_{i=N_u+2}^{k+1} \|\mathcal{A}^{(i)}_k\|^2 \\ &\overset{(b)}{\le} (\mu - \varepsilon_u) \sum_{i=1}^{N_u+1} \rho e^{-\eta(k-i)} I + \varepsilon_u \sum_{i=1}^{k+1} \rho e^{-\eta(k-i)} I \\ &\le \frac{(\mu - \varepsilon_u)\rho e^{2\eta}}{e^\eta - 1} e^{-\eta(k-N_u)} I + \varepsilon_u \frac{\rho e^{2\eta}}{e^\eta - 1} I, \end{aligned} \tag{26}$$

where $\overset{(a)}{\le}$ is obtained by using $U_k < \varepsilon_u I$ for $k > N_u$. The inequality $\overset{(b)}{\le}$ is obtained by Lemma 7.

For any $\varepsilon > 0$, we claim that there exists an integer $N > 0$ such that by choosing a sufficiently small ε_u and by letting $k > N$, (26) $< \varepsilon I$ holds, i.e., $\mathcal{M}_{k+1} < \varepsilon I$. Then we have $\mathbb{E}[\mathcal{M}_{k+1}] < \varepsilon I$, for $k > N$, which implies that $\lim_{k\to\infty} \mathbb{E}[\mathcal{M}_{k+1}] = 0$. From Lemma 4 and (21b), it follows that $\lim_{k\to\infty} \mathbb{E}[\bar{P}^t_k] = \lim_{k\to\infty} \mathbb{E}[\bar{P}_k]$.

It is easy to check that for $\gamma_k = 1$, $k \in \mathbb{N}$, (2b) and (2e) reduce to the standard Riccati equation. Thus, $\bar{P}^t_k$ is no longer a random quantity, i.e., $\mathbb{E}[\bar{P}^t_k] = \bar{P}^t_k$. Under Assumption 1, $\bar{P}^t_k$ converges to $\bar{P}$ where $\bar{P} = g(1, \bar{P})$. From the results above, it follows that $\bar{P} = \lim_{k\to\infty} \mathbb{E}[\bar{P}_k]$. Part (a) is proved.

Proof of Part (b): Let $\zeta \triangleq \rho\rho_u e^{\eta_u(m_u+1)} I$. From (24),

$$\begin{aligned} \mathcal{M}_{k+1} &= \sum_{i=1}^{k+1} \mathcal{A}^{(i)}_k U_{i-1} \mathcal{A}^{(i)\prime}_k \le \sum_{i=1}^{k+1} \|\mathcal{A}^{(i)}_k\|^2 U_{i-1} \\ &\le \sum_{i=1}^{k+1} \rho e^{-\eta(k-i)} \cdot \rho_u e^{-\eta_u(i-1-m_u)} I \\ &\le \rho\rho_u e^{\eta_u(m_u+1)} e^{-\eta k} I \sum_{i=1}^{k+1} e^{\eta i} \cdot e^{-\eta_u i}. \end{aligned} \tag{27}$$

- If $\eta = \eta_u$, then (27) $= \zeta e^{-\eta k}(k+1)$. It is clear that there exists an integer N such that $(k+1)e^{-\frac{1}{2}\eta k} < 1$ for $k > N$. Then, $\mathcal{M}_{k+1} \leq \zeta(k+1)e^{-\eta k} = \zeta(k+1)e^{-\frac{1}{2}\eta k}e^{-\frac{1}{2}\eta k} \leq \zeta e^{-\frac{1}{2}\eta k}$ for $k > N$, which means that the convergence rate of $\mathcal{M}_k \to 0$ is exponential.
- If $\eta \neq \eta_u$, let $\tau \triangleq e^{\eta-\eta_u}$,

$$\begin{aligned}(27) &= \zeta e^{-\eta k}\sum_{i=1}^{k+1} e^{\eta i}\cdot e^{-\eta_u i} \\ &= \zeta e^{-\eta k}\frac{\tau}{1-\tau}[1-e^{(\eta-\eta_u)(k+1)}]. \qquad (28)\end{aligned}$$

  If $\eta < \eta_u$, $0 < \tau = e^{\eta-\eta_u} < 1$. From (28), we have $\mathcal{M}_{k+1} \leq (27) < \zeta e^{-\eta k}\frac{\tau}{1-\tau}$. If $\eta > \eta_u$, $1 < \tau = e^{\eta-\eta_u}$. Then, (28) can be rewritten as (27) $= \zeta e^{-\eta k}\frac{\tau}{\tau-1}[e^{(\eta-\eta_u)(k+1)}-1] < \zeta e^{-\eta k}\frac{\tau}{\tau-1}e^{(\eta-\eta_u)(k+1)} = \zeta\frac{\tau}{\tau-1}e^{\eta-\eta_u}e^{-\eta_u k}$. That is, $\mathcal{M}_{k+1} \leq (27) < \zeta\frac{\tau^2}{\tau-1}e^{-\eta_u k}$. Thus, when $\eta \neq \eta_u$, the convergence rate of $\mathcal{M}_k \to 0$ is exponential as well.

Note that the convergence rate of $\bar{P}^t_k$ is known to be exponential and that $\mathbb{E}[\bar{P}^t_k] \leq \mathbb{E}[\bar{P}_k] \leq \mathbb{E}[\bar{P}^t_k + \mathcal{M}_k]$. Therefore, $\mathbb{E}[\bar{P}_k]$ exponentially converges to $\bar{P}$. □

Remark 4 *It follows from (17d) that $\bar{P}_{k+1} \geq \bar{P}^t_{k+1}$, which suggests that the lack of ACK signals degrades the estimation performance, but this degradation may fade away if the conditions in Theorem 3 are satisfied.*

Remark 5 *By designating values of $\{u_k\}$, it would be possible to keep $\mathbb{E}[\bar{P}^t_k + \sum_{i=1}^{2^{(k)}}\bar{\alpha}^{(i)}_k(\bar{z}^{(i)}_k - \bar{x}_k)^2_I]$ constant so as to make $\mathbb{E}[\bar{P}_k]$ convergent. However, we do not consider such kind of convergence since this paper does not involve controller design, as mentioned in Section 2. Moreover, it is computationally infeasible to design u_k.*

6 Numerical Example

Consider the system in (1) with following parameters:

$$A = \begin{bmatrix}\sigma & 0\\ 0 & 0.5\end{bmatrix}, \sigma > 1, B = \begin{bmatrix}-1\\ 1\end{bmatrix}, C = \begin{bmatrix}-1 & 1\end{bmatrix},$$
$$Q = \begin{bmatrix}20 & 0\\ 0 & 20\end{bmatrix}, R = 20.$$

For the convenience in calculating the critical value, we choose A with only one unstable eigenvalue σ. If such systems are TCP-like, it follows from [31] that the critical value $\lambda = 1 - 1/\sigma^2$. It takes about 10 mins (110 mins) to compute the optimal estimator with the time step 30 (40). Here, the simulation step is set to be 30, and the EEC is computed by performing simulations 50 times with randomly generated $\{\nu_k, \gamma_k\}$ and then taking the average value of them. Main results obtained in this paper are illustrated by numerical examples as follows:

- Stability: By letting σ take different values $\{1.0541, 1.1952$, and $1.4142\}$, respectively, the corresponding critical values are $\{0.1, 0.3$, and $0.5\}$. For the UDP-like system with bounded inputs ($\|u_k\| < 10$), the relationship between the trace of $\mathbb{E}[\bar{P}_k]$ and the observation packet arrival rate γ is shown in Fig. 2. It can be seen from Fig. 2 that for a UDP-like system there does exist a critical threshold value, which is identical to that for its corresponding TCP-like system, as stated in Theorem 2.
- Convergence: By letting $u_k = 10e^{(-k^2/30)}$ and $\gamma = 1$, it is shown in Fig. 3 that for a UDP-like system, $\mathbb{E}[\bar{P}_k]$ eventually converges, regardless of what the control packet arrival rate ν is, as stated in Theorem 3.
- Impact of ACK: Let $\sigma = 1.1952$, then the corresponding critical value $\lambda = 0.2$. For different values of γ, the relationship between the trace of EEC $\mathbb{E}[\bar{P}_{30}]$ and the control packet arrival rate ν is illustrated in Fig. 4, which shows that the boundedness of EEC does not depend on ν. Note that there is an interesting phenomenon in Fig. 4. For each γ, the EEC attains its maximum value near $\nu = 0.5$.

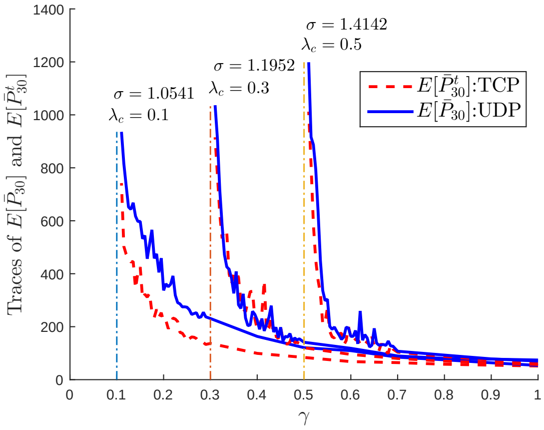

Fig. 2. Critical values for UDP-like and TCP-like systems.

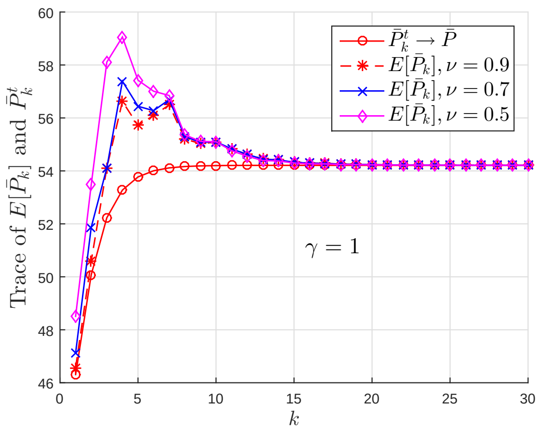

Fig. 3. The convergence of $\mathbb{E}[\bar{P}_k]$.

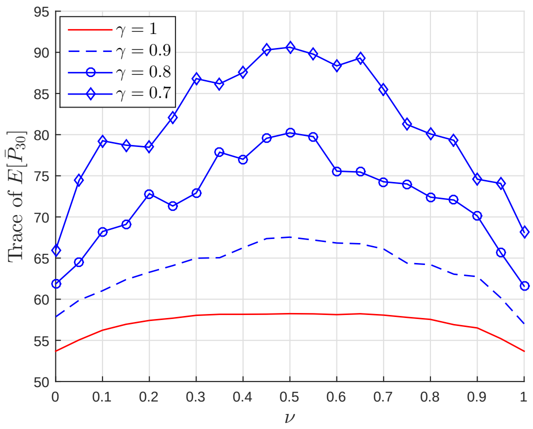

Fig. 4. Relationship between EEC and ν.

7 Conclusions

In this paper, the optimal estimator for UDP-like systems with dual packet losses has been obtained. It has been shown that its stability is only determined by the observation packet arrival rate, and that whether or not there are ACK signals does not affect its stability. If there is no observation packet loss and control inputs eventually tend to zero, then the EEC converges. Although some theoretical problems have been solved, there still exist many issues for UDP-like systems, such as designing efficient sub-optimal/approximate estimators and controllers, considering Markovian or correlated packet losses in communication channels, and studying consensus problems [37] for multi-agent UDP-like systems. Investigating estimation and control problems for UDP-like systems equipped with devices having computational ability or intelligent strategies—such as event-trigger schemes [22], smart sensors [17], and network relays [12]—may obtain some interesting results.

Appendix

Proof of Lemma 1: Lemma 1 is proved by mathematical induction in the following four steps.

Step 1: We check (13) and (14) for $k=1$. That is, we examine $p(x_1|\theta_0^{(i)},\mathcal{G}_0)$ and $p(x_1|\theta_0^{(i)},\mathcal{G}_1)$ for (13), and $\bar{z}_1^{(i)}$ and $\hat{z}_1^{(i)}$ for (14) with $i=\{1,2\}$. Note that $\mathcal{G}_0=\emptyset$, $\mathcal{G}_1=\{y_1,\gamma_1\}$, $\theta_0^{(1)}=\{\nu_0=0\}$, and $\theta_0^{(2)}=\{\nu_0=1\}$. Consider the case $i=1$. It follows from (4) that

$$\begin{aligned} p(x_1|\theta_0^{(1)}) &= p(Ax_0+\omega_0|\nu_0=0) \\ &= \mathcal{N}(A\bar{x}_0, AP_0A'+Q). \end{aligned} \tag{29}$$

By (2b), $\bar{P}_1^t = AP_0A'+Q$ due to $\bar{P}_0^t = P_0$. Letting $\bar{z}_1^{(1)} = A\bar{x}_0$, we have $p(x_1|\theta_0^{(1)}) = \mathcal{N}(\bar{z}_1^{(1)}, \bar{P}_1^t)$ with $\bar{z}_1^{(1)} = A\hat{z}_0^{(1)}$ ($\hat{z}_0^{(1)}=\bar{x}_0$). Hence, (13a) and (14a) are true when $i=1$ and $k=1$.

Step 2: For (13b), we compute $p(x_1|\theta_0^{(i)},\mathcal{G}_1)$, i.e., $p(x_1|\theta_0^{(i)})$ conditioned on $\{y_1,\gamma_1\}$. When $\gamma_1=0$, $y_1=\emptyset$. $p(x_1|\theta_0^{(1)},\gamma_1=0,y_1=\emptyset)=\mathcal{N}(\bar{z}_1^{(1)},\bar{P}_1^t)$. When $\gamma_1=1$, using (5) and (29) yields

$$\begin{aligned} &p(x_1|\theta_0^{(1)},\gamma_1=1,y_1) = p(x_1|\theta_0^{(1)},y_1) \\ &= \mathcal{N}(\bar{z}_1^{(1)}+K_1(y_1-C\bar{z}_1^{(1)}),(I-K_1C)\bar{P}_1^t). \end{aligned} \tag{30}$$

By (2e), $P_1^t = \bar{P}_1^t - \gamma_1 K_1 C\bar{P}_1^t$. By letting $\hat{z}_1^{(1)} = \bar{z}_1^{(1)} + \gamma_1K_1(y_1-C\bar{z}_1^{(1)})$, from (30), we have $p(x_1|\theta_0^1,\mathcal{G}_1) = \mathcal{N}(\hat{z}_1^{(1)},P_1^t)$. Hence, (13b) and (14b) are true when $i=1$ and $k=1$. Consequently, (13) and (14) are true when $i=1$ and $k=1$. Similarly, it is easy to show that (13) and (14) remain true for $i=2$. Hence, (13) and (14) are true when $k=1$ with $1\le i\le 2$.

Step 3: Suppose that (13) and (14) hold for $1,\ldots,k$. We check the case for $k{+}1$ with $1\le i\le 2^{k+1}$. When $1\le i\le 2^k$, from (11), $\theta_k^{(i)}=\{\nu_k=0,\theta_{k-1}^{(i)}\}$. It follows from (4) that $p(x_{k+1}|\theta_k^{(i)},\mathcal{G}_k)=p(Ax_k+\omega_k|\nu_k=0,\theta_{k-1}^{(i)},\mathcal{G}_k)$ $=\mathcal{N}(A\hat{z}_k^{(i)},AP_k^tA'{+}Q)$. By (2b), $\bar{P}_{k+1}^t = AP_k^tA'{+}Q$. By letting $\bar{z}_{k+1}^{(i)}=A\hat{z}_k^{(i)}$, $p(x_{k+1}|\theta_k^{(i)},\mathcal{G}_k)=\mathcal{N}(\bar{z}_{k+1}^{(i)},\bar{P}_{k+1}^t)$. Hence, (13a) and (14a) are true when $k+1$ with $1\le i\le 2^k$.

Step 4: When $\gamma_{k+1}=0$, $y_{k+1}=\emptyset$. $p(x_{k+1}|\theta_k^{(i)},\mathcal{G}_{k+1}) = p(x_{k+1}|\theta_k^{(i)},\mathcal{G}_k) = \mathcal{N}(\bar{z}_{k+1}^{(i)},\bar{P}_{k+1}^t)$. When $\gamma_{k+1}=1$, y_{k+1} is available. By using (5), $p(x_{k+1}|\theta_k^{(i)},\mathcal{G}_{k+1}) = p(x_{k+1}|\theta_k^{(i)},\gamma_{k+1}=1,y_{k+1},\mathcal{G}_k) = \mathcal{N}\big(\bar{z}_{k+1}^{(i)}{+}K_{k+1}(y_{k+1}-C\bar{z}_{k+1}^{(i)}),(I-K_{k+1}C)\bar{P}_{k+1}^t\big)$. By (2e), $P_{k+1}^t = \bar{P}_{k+1}^t - \gamma_{k+1}K_{k+1}C\bar{P}_{k+1}^t$. Letting $\hat{z}_{k+1}^{(i)} = \bar{z}_{k+1}^{(i)}{+}\gamma_{k+1}K_{k+1}(y_{k+1}-C\bar{z}_{k+1}^{(i)})$, we have $p(x_{k+1}|\theta_k^{(i)},\mathcal{G}_{k+1}) = \mathcal{N}(\hat{z}_{k+1}^{(i)},P_{k+1}^t)$. Hence, (13b) and (14b) are true when $k+1$, $1\le i\le 2^k$. Similarly, it is not difficult to verify that they remain true for $2^k+1\le i\le 2^{k+1}$. The derivations are not given here for saving space. Therefore, (13) and (14) are true when $k+1$, $1\le i\le 2^{k+1}$. The proof is completed. □

Proof of Lemma 2: We first check (15a). For $1\le i\le 2^k$, by (11), $\bar{\alpha}_{k+1}^{(i)} = p(\theta_k^{(i)}|\mathcal{G}_k) = p(\nu_k=0,\theta_{k-1}^{(i)}|\mathcal{G}_k) = \bar{\nu}\hat{\alpha}_k^{(i)}$, since ν_k is independent of $\theta_{k-1}^{(i)}$ and $\mathcal{G}_{k-1}$. For $2^k{+}1\le i\le 2^{k+1}$, by the same derivation, $\bar{\alpha}_{k+1}^{(i)} = p(\nu_k = 1,\theta_{k-1}^{(i-2^k)}|\mathcal{G}_k) = \nu\hat{\alpha}_k^{(i-2^k)}$. Hence, (15a) holds.

Eq.(15b) is proved by the mathematical induction method as follows. Consider $k=1$ with $i=1$ or 2. When

$\gamma_1 = 0$, $y_1 = \emptyset$. $\hat{\alpha}_1^{(i)} \triangleq p(\theta_0^{(i)}|y_1) = p(\theta_0^{(i)}) = \bar{\alpha}_1^{(i)} = (\cdot)^{\gamma_1}\bar{\alpha}_1^{(i)}|_{\gamma_1=0}$. When $\gamma_1 = 1$, y_1 is available. By Bayesian formula, we have $\hat{\alpha}_1^{(i)} \triangleq p(\theta_0^{(i)}|y_1) = c^{-1}p(y_1|\theta_0^{(i)})p(\theta_0^{(i)})$ where $c = p(y_1|\theta_0^{(1)})p(\theta_0^{(1)}) + p(y_1|\theta_0^{(2)})p(\theta_0^{(2)})$. By Lemma 1, $p(x_1|\theta_0^{(i)}) = \mathcal{N}(\bar{z}_1^{(i)}, \bar{P}_1^t)$. By using (4), $p(y_1|\theta_0^{(i)}) = \mathcal{N}_{y_1}(C\bar{z}_1^{(i)}, C\bar{P}_1^tC' + R) = \phi_1^{(i)}$. Thus, $\hat{\alpha}_1^{(i)} = \left(\frac{\phi_1^{(i)}}{\bar{\alpha}_1^{(1)}\phi_1^{(1)}+\bar{\alpha}_1^{(2)}\phi_1^{(2)}}\right)^{\gamma_1}\bar{\alpha}_1^{(i)}|_{\gamma_1=1}$. This shows that (15b) holds for $k = 1$ with $i = 1$ or 2.

Suppose that (15b) are true for $1, \ldots, k$. Consider the case $k+1$ and $1 \le i \le 2^{k+1}$. When $\gamma_{k+1} = 0$, $y_{k+1} = \emptyset$. $\hat{\alpha}_{k+1}^{(i)} \triangleq p(\theta_k^{(i)}|\mathcal{G}_{k+1}) = p(\theta_k^{(i)}|\mathcal{G}_k) = (\cdot)^{\gamma_{k+1}}\bar{\alpha}_{k+1}^{(i)}|_{\gamma_{k+1}=0}$. When $\gamma_{k+1} = 1$, by using Bayesian formula, $\hat{\alpha}_{k+1}^{(i)} \triangleq p(\theta_k^{(i)}|\mathcal{G}_{k+1}) = p(\theta_k^{(i)}|y_{k+1}, \mathcal{G}_k) = c_k^{-1}p(y_{k+1}|\theta_k^{(i)}, \mathcal{G}_k)p(\theta_k^{(i)}|\mathcal{G}_k)$, where

$$c_k = \sum_{j=1}^{2^{k+1}} p(y_{k+1}|\theta_k^{(j)}, \mathcal{G}_k)p(\theta_k^{(j)}|\mathcal{G}_k).$$

By Lemma 1, $p(x_{k+1}|\theta_k^{(i)}, \mathcal{G}_k) = \mathcal{N}(\bar{z}_{k+1}^{(i)}, \bar{P}_{k+1}^t)$. By using (4), $p(y_{k+1}|\theta_k^{(i)}, \mathcal{G}_k) = \mathcal{N}_{y_{k+1}}(C\bar{z}_{k+1}^{(i)}, P_{k+1}^Y) = \phi_{k+1}^{(i)}$, and thus $\hat{\alpha}_{k+1}^{(i)} = \left(\frac{\phi_{k+1}^{(i)}}{\sum_{j=1}^{2^{k+1}}\phi_{k+1}^{(j)}\bar{\alpha}_{k+1}^{(j)}}\right)^{\gamma_{k+1}}\bar{\alpha}_{k+1}^{(i)}$ with $\gamma_{k+1} = 1$. Therefore, (15b) is true for the case $k+1$, which completes the proof. □

Proof of Lemma 3: Note that $p(x_k^a|\theta_{k-1}^{(i)}, \mathcal{G}_{k-1})$ in (18) and $p(x_k|\theta_{k-1}^{(i)}, \mathcal{G}_{k-1})$ are the same Gaussian pdfs. It follows from Lemma 1 that $p(x_k^a|\theta_{k-1}^{(i)}, \mathcal{G}_{k-1}) = \mathcal{N}(\bar{z}_k^{(i)}, \bar{P}_k^t)$.

Define $\tilde{\alpha}_k^{(i)} \triangleq p(\theta_{k-1}^{(i)})$ in (18). Then we check (20). For $1 \le i \le 2^{k-1}$, by (11), $\theta_{k-1}^{(i)} = \{\nu_{k-1} = 0, \theta_{k-2}^{(i)}\}$. $\tilde{\alpha}_k^{(i)} \triangleq p(\theta_{k-1}^{(i)}) = p(\nu_{k-1}{=}0, \theta_{k-2}^{(i)}) = p(\nu_{k-1}{=}0)p(\theta_{k-2}^{(i)}) = \bar{\nu}\tilde{\alpha}_{k-1}^{(i)}$. By following a similar line of argument, we obtain $\tilde{\alpha}_k^{(i)} = \nu\tilde{\alpha}_{k-1}^{i-2^{k-1}}$, for $2^{k-1}{+}1 \le i \le 2^k$. Thus, (20) holds.

From (18) and the results above, it follows that (19) holds. The proof is completed. □

Proof of Lemma 4: From (17d), it is evident that $\underset{\mathcal{G}_{k-1}}{\mathbb{E}}[\bar{P}_k^t] \le \underset{\mathcal{G}_{k-1}}{\mathbb{E}}[\bar{P}_k]$. In the sequel, we prove $\underset{\mathcal{G}_{k-1}}{\mathbb{E}}[\bar{P}_k] \le \underset{\mathcal{G}_{k-1}}{\mathbb{E}}[\bar{P}_k^a]$. By (6), $\bar{P}_k = \underset{x}{\mathbb{E}}[(x_k)_I^2|\mathcal{G}_{k-1}] - (\underset{x}{\mathbb{E}}[x_k|\mathcal{G}_{k-1}])_I^2$ and $\bar{P}_k^a = \underset{x}{\mathbb{E}}[(x_k^a)^2|\mathcal{G}_{k-1}] - (\underset{x}{\mathbb{E}}[x_k^a|\mathcal{G}_{k-1}])_I^2$, where $\underset{x}{\mathbb{E}}$ denotes the expectation is taken with respect to x_k or x_k^a (the specific one is known from the context). We use $\mathcal{G}$ to denote $\mathcal{G}_{k-1}$ in the following derivations to make the formulas concise. By using the pdf in (12a), $\bar{P}_k$ can be further rewritten as

$$\begin{aligned}
\bar{P}_k &= \underset{x}{\mathbb{E}}[(x_k)_I^2|\mathcal{G}] - (\underset{x}{\mathbb{E}}[x_k|\mathcal{G}])_I^2 \\
&= \int (x_k)_I^2 p(x_k|\mathcal{G})\mathrm{d}x_k - \left(\int x_k p(x_k|\mathcal{G})\mathrm{d}x_k\right)_I^2 \\
&= \sum_{i=1}^{2^k}\int (x_k)_I^2 p(x_k|\theta_{k-1}^{(i)}, \mathcal{G})\mathrm{d}x_k p(\theta_{k-1}^{(i)}|\mathcal{G}) \\
&\quad - \left(\sum_{i=1}^{2^k}\int x_k p(x_k|\theta_{k-1}^{(i)}, \mathcal{G})\mathrm{d}x_k p(\theta_{k-1}^{(i)}|\mathcal{G})\right)_I^2 \\
&= \underset{\theta}{\mathbb{E}}\big[\underset{x}{\mathbb{E}}[(x_k)_I^2|\theta_{k-1}^{(i)}, \mathcal{G}]\big|\mathcal{G}\big] + \Big(-\underset{\theta}{\mathbb{E}}\big[\underset{x}{\mathbb{E}}[x_k|\theta_{k-1}^{(i)}, \mathcal{G}]\big|\mathcal{G}\big]\Big)_I^2 \\
&\triangleq \mathcal{W}_L + \mathcal{W}_R,
\end{aligned}$$

where $\int$ stands for $\int_{-\infty}^{\infty}$ and $\underset{\theta}{\mathbb{E}}$ denotes the expectation is taken over all $\theta_{k-1}^{(i)}$. Similarly, by using the pdf in (18),

$$\begin{aligned}
\bar{P}_k^a &= \underset{x}{\mathbb{E}}[(x_k^a)_I^2|\mathcal{G}] - (\underset{x}{\mathbb{E}}[x_k^a|\mathcal{G}])_I^2 \\
&= \sum_{i=1}^{2^k}\int (x_k)_I^2 p(x_k|\theta_{k-1}^{(i)}, \mathcal{G})\mathrm{d}x_k p(\theta_{k-1}^{(i)}) \\
&\quad - \left(\sum_{i=1}^{2^k}\int x_k p(x_k|\theta_{k-1}^{(i)}, \mathcal{G})\mathrm{d}x_k p(\theta_{k-1}^{(i)})\right)_I^2 \qquad (31) \\
&= \underset{\theta}{\mathbb{E}}\big[\underset{x}{\mathbb{E}}[(x_k)_I^2|\theta_{k-1}^{(i)}, \mathcal{G}]\big] + \big(-\big(\underset{\theta}{\mathbb{E}}\big[\underset{x}{\mathbb{E}}[x_k|\theta_{k-1}^{(i)}, \mathcal{G}]\big]\big)_I^2\big) \\
&\triangleq \mathcal{W}_L^a + \mathcal{W}_R^a.
\end{aligned}$$

For the convenience of comparing $\bar{P}_k$ with $\bar{P}_k^a$, we replace x_k^a in (31) with x_k, which does not affect the value of integration. Let

$$\begin{aligned}
\mathcal{W} &\triangleq \sum_{i=1}^{2^k}\left(\int x_k p(x_k|\theta_{k-1}^{(i)}, \mathcal{G})\mathrm{d}x_k\right)_I^2 p(\theta_{k-1}^{(i)}|\mathcal{G}) \\
&= \underset{\theta}{\mathbb{E}}[(\underset{x}{\mathbb{E}}[x_k|\theta_{k-1}^{(i)}, \mathcal{G}])_I^2|\mathcal{G}],
\end{aligned}$$

and let $\bar{P}_k = (\mathcal{W}_L - \mathcal{W}) + (\mathcal{W} + \mathcal{W}_R)$.

$$\begin{aligned}
\mathcal{W}_L - \mathcal{W} &= \underset{\theta}{\mathbb{E}}[\underset{x}{\mathbb{E}}((x_k)_I^2|\theta_{k-1}^{(i)}, \mathcal{G}) - (\underset{x}{\mathbb{E}}[x_k|\theta_{k-1}^{(i)}, \mathcal{G}])_I^2|\mathcal{G}] \\
&= \underset{\theta}{\mathbb{E}}[\underset{x}{\mathrm{cov}}(x_k|\theta_{k-1}^{(i)}, \mathcal{G})|\mathcal{G}]. \qquad (32)
\end{aligned}$$

$$\begin{aligned}
&\mathcal{W} + \mathcal{W}_R \\
&= \underset{\theta}{\mathbb{E}}[(\underset{x}{\mathbb{E}}[x_k|\theta_{k-1}^{(i)}, \mathcal{G}])_I^2|\mathcal{G}] - (\underset{\theta}{\mathbb{E}}[\underset{x}{\mathbb{E}}[x_k|\theta_{k-1}^{(i)}, \mathcal{G}]|\mathcal{G}])_I^2 \\
&= \underset{\theta}{\mathrm{cov}}(\underset{x}{\mathbb{E}}[x_k|\theta_{k-1}^{(i)}, \mathcal{G}]|\mathcal{G}), \qquad (33)
\end{aligned}$$

where (33) follows from (6) by noting that $\mathbb{E}_x[x_k|\theta_{k-1}^{(i)},\mathcal{G}]$ is a function of $\theta_{k-1}^{(i)}$. From (32) and (33), we have

$$\begin{aligned}\bar{P}_k &= \mathbb{E}_\theta\big[\operatorname{cov}_x(x_k|\theta_{k-1}^{(i)},\mathcal{G})|\mathcal{G}\big] + \operatorname{cov}_\theta\big(\mathbb{E}_x[x_k|\theta_{k-1}^{(i)},\mathcal{G}]|\mathcal{G}\big)\\ &\triangleq \bar{P}_L + \bar{P}_R, \end{aligned} \tag{34}$$

where $\bar{P}_L$ and $\bar{P}_R$ denote the first and the second terms in the right part of the first equality above, respectively.

$$\begin{aligned}\mathcal{W}^a &\triangleq \sum_{i=1}^{2^k}\left(\int x_k^a p(x_k^a|\theta_{k-1}^{(i)},\mathcal{G})\mathrm{d}x_k^a\right)_I^2 p(\theta_{k-1}^{(i)})\\ &= \mathbb{E}_\theta\Big[\big(\mathbb{E}_x[x_k^a|\theta_{k-1}^{(i)},\mathcal{G}]\big)_I^2\Big],\end{aligned}$$

and let $\bar{P}_k^a = \mathcal{W}_L^a - \mathcal{W}^a + \mathcal{W}^a + \mathcal{W}_R^a$. By following the same line above, we obtain

$$\begin{aligned}\bar{P}_k^a &= \mathbb{E}_\theta\big[\operatorname{cov}_x(x_k|\theta_{k-1}^{(i)},\mathcal{G})\big] + \operatorname{cov}_\theta\big(\mathbb{E}_x[x_k|\theta_{k-1}^{(i)},\mathcal{G}]\big)\\ &\triangleq \bar{P}_L^a + \bar{P}_R^a, \end{aligned} \tag{35}$$

where $\bar{P}_L^a$ and $\bar{P}_R^a$ denote the first and the second terms in the right part of the first equality above, respectively. For the conciseness of formulation, we replace the symbols x_k in (34) and (35) with x.

We compare $\mathbb{E}_\mathcal{G}[\bar{P}_k]$ with $\mathbb{E}_\mathcal{G}[\bar{P}_k^a]$ as follows:

$$\begin{aligned}\mathbb{E}_\mathcal{G}[\bar{P}_L^a] &= \mathbb{E}_\mathcal{G}\Big[\mathbb{E}_\theta\big[\operatorname{cov}_x(x|\theta_{k-1}^{(i)},\mathcal{G})\big]\Big]\\ &\overset{(a)}{=} \mathbb{E}_\mathcal{G}\Big[\mathbb{E}_\theta\big[\operatorname{cov}_x(x|\theta_{k-1}^{(i)},\mathcal{G})|\mathcal{G}\big]\Big] = \mathbb{E}_\mathcal{G}[\bar{P}_L], \end{aligned} \tag{36}$$

where $\overset{(a)}{=}$ is obtained by using (7). Thus, $\mathbb{E}_\mathcal{G}[\bar{P}_L^a] = \mathbb{E}_\mathcal{G}[\bar{P}_L]$.

Since $\operatorname{cov}(\mathbb{E}[Z|J]) \geq 0$, from (8), it follows that $\operatorname{cov}(Z) \geq \mathbb{E}[\operatorname{cov}(Z|J)]$. By viewing $\mathbb{E}_x[x|\theta_{k-1}^{(i)},\mathcal{G}]$ as Z and $\mathcal{G}$ as J,

$$\operatorname{cov}_\theta\Big(\mathbb{E}_x[x|\theta_{k-1}^{(i)},\mathcal{G}]\Big) \geq \mathbb{E}_\mathcal{G}\Big[\operatorname{cov}_\theta\big(\mathbb{E}_x[x|\theta_{k-1}^{(i)},\mathcal{G}]|\mathcal{G}\big)\Big]. \tag{37}$$

The term on the left-hand side of (37) is a function of $\mathcal{G}$, and the term on the right-hand side is a constant quantity. By taking the mathematical expectation to (37),

$$\mathbb{E}_\mathcal{G}\Big[\operatorname{cov}_\theta\big(\mathbb{E}_x[x|\theta_{k-1}^{(i)},\mathcal{G}]\big)\Big] \geq \mathbb{E}_\mathcal{G}\Big[\operatorname{cov}_\theta\big(\mathbb{E}_x[x|\theta_{k-1}^{(i)},\mathcal{G}]|\mathcal{G}\big)\Big],$$

where the left part equates $\mathbb{E}_\mathcal{G}[\bar{P}_R^a]$, and the right part equates $\mathbb{E}_\mathcal{G}[\bar{P}_R]$. Thus, $\mathbb{E}_\mathcal{G}[\bar{P}_R^a] \geq \mathbb{E}_\mathcal{G}[\bar{P}_R]$. By noting that $\mathbb{E}_\mathcal{G}[\bar{P}_L^a] = \mathbb{E}_\mathcal{G}[\bar{P}_L]$, we have $\mathbb{E}_\mathcal{G}[\bar{P}_k^a] \geq \mathbb{E}_\mathcal{G}[\bar{P}_k]$. □

Proof of Lemma 5: By substituting (14b) into (14a),

$$\bar{z}_{k+1}^{(i)} = \begin{cases} \mathbb{A}_k\bar{z}_k^{(i)} + \gamma_k AK_k y_k, & \text{for } 1 \leq i \leq 2^k\\ \mathbb{A}_k\bar{z}_k^{(i-2^k)} + \gamma_k AK_k y_k + Bu_k, & \\ & \text{for } 2^k{+}1 \leq i \leq 2^{k+1}. \end{cases}$$

From (20) and (21a), $\bar{x}_{k+1}^a = \sum_{i=1}^{2^{k+1}} \tilde{\alpha}_{k+1}^{(i)}\bar{z}_{k+1}^{(i)} = \sum_{i=1}^{2^k} \tilde{\alpha}_{k+1}^{(i)}\bar{z}_{k+1}^{(i)} + \sum_{i=2^k+1}^{2^{k+1}} \tilde{\alpha}_{k+1}^{(i)}\bar{z}_{k+1}^{(i)} = \mathbb{A}_k\bar{x}_k^a + \gamma_k AK_k y_k + \nu Bu_k$. According to the definition of $\mathcal{M}_k$, $\mathcal{M}_{k+1} = \sum_{i=1}^{2^{k+1}} \tilde{\alpha}_{k+1}^{(i)}(\bar{z}_{k+1}^{(i)} - \bar{x}_{k+1}^a)_I^2 = \mathbb{A}_k\sum_{i=1}^{2^k}\tilde{\alpha}_k^{(i)}(\bar{x}_k^a - \bar{z}_k^{(i)})_I^2\mathbb{A}_k' + \bar{\nu}\nu Bu_k u_k' B' = \mathbb{A}_k\mathcal{M}_k\mathbb{A}_k' + \bar{\nu}\nu Bu_k u_k' B'$. The proof of (23) is completed.

Equation (24) is proved by the mathematical induction. When $n = 1$, we have $\mathcal{M}_1 = \sum_{i=1}^{1}\mathcal{A}_0^{(1)}U_0\mathcal{A}_0^{(1)'} = U_0$ owing to $\mathcal{A}_n^{(n+1)} = 1$. From (14a), (20), and (21b), it follows that $\mathcal{M}_1 = \nu\bar{\nu}Bu_0u_0'B' = U_0$. Thus, (24) is true when $n = 1$. Suppose that (24) is true for $n = 1, \ldots, k$. We examine the case $n = k+1$ as follows. By (23), $\mathcal{M}_{k+1} = \mathbb{A}_k\mathcal{M}_k\mathbb{A}_k' + U_k = \mathbb{A}_k\sum_{i=1}^{k}\mathcal{A}_{k-1}^{(i)}U_{i-1}\mathcal{A}_{k-1}^{(i)'}\mathbb{A}_k' + \mathcal{A}_k^{(k+1)}U_k\mathcal{A}_k^{(k+1)'} = \sum_{i=1}^{k+1}\mathcal{A}_k^{(i)}U_{i-1}\mathcal{A}_k^{(i)'}$. Hence, (24) is true when $n = k+1$, which completes the proof. □

Proof of Lemma 6: . There are some different formulas for presenting the relationship between P_k^t and $\bar{P}_k^t$. One of them is (2e), and an equivalent one is $P_k^t = (I - \gamma_k K_k C)\bar{P}_k^t(I - \gamma_k K_k C)' + \gamma_k K_k R K_k'$. By (2b), $\bar{P}_{k+1}^t = \mathbb{A}_k\bar{P}_k^t\mathbb{A}_k' + \gamma_k AK_kR(AK_k)' + Q$. Let $\Delta_k \triangleq \bar{P}_k^t - \mathcal{M}_k$. From (23), $\Delta_{k+1} = \mathbb{A}_k\Delta_k\mathbb{A}_k' + \gamma_k AK_kR(AK_k)' + (Q - U_k)$. Due to $U_i \leq Q$, $\Delta_1 = \bar{P}_1^t - \mathcal{M}_1 \geq 0$ ($\mathcal{M}_1 = U_0$), and $\gamma_k AK_kR(AK_k)' \geq 0$, we have $\Delta_k \geq 0$, for all k. Therefore, $\mathcal{M}_k \leq \bar{P}_k^t$. The proof is completed. □

Proof of Lemma 7: Note that for the $\mathcal{S}_{TCP}^u$ system, $\gamma_i = 1$, $i \in \mathbb{N}$. As shown above, $\mathcal{A}_k^{(i)}$ in (22) with $\gamma_i \equiv 1$ is the transition matrix of the prediction error equation for the $\mathcal{S}_{LTI}$ system. For LTI system, it is well known [11, pp. 240] that under Assumption 1, the prediction error equation (25), also called as the filtering equation, is exponentially stable. That is, for the state transition matrix $\mathcal{A}_k^{(m)} = \mathbb{A}_k\cdots\mathbb{A}_m$, there exist positive constants ρ_1 and η_1 such that $\|\mathcal{A}_k^{(m)}\| < \rho_1 e^{-\eta_1(k-m)}$. Hence, $\|\mathcal{A}_k^{(m)}\|^2 < \rho_1^2 e^{-2\eta_1(k-m)}$. Let $\rho = \rho_1^2$ and $\eta = 2\eta_1$. We have $\|\mathcal{A}_k^{(m)}\|^2 < \rho e^{-\eta(k-m)}$. The proof is completed. □

References

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

[2] A. Censi. Kalman filtering with intermittent observations: convergence for semi-Markov chains and an intrinsic performance measure. *IEEE Trans. Automat. Control*, 56(2):376–381, 2011.

[3] W.-W. Che, J.-L. Wang, and G.-H. Yang. H_∞ control for networked control systems with limited communication. *European J. Control*, 18(2):103–118, 2012.

[4] W.-W. Che and G.-H. Yang. H_∞ filter design for continuous-time systems with quantised signals. *Int. J. Syst. Sci.*, 44(2):265–274, 2013.

[5] D. E. Clark and J. Bell. Convergence results for the particle PHD filter. *IEEE Trans. Signal Process.*, 54(7):2652–2661, 2006.

[6] A. Doucet, N. De Freitas, and N. Gordon. *An Introduction to Sequential Monte Carlo Methods*. Springer, 2001.

[7] E. Garone, B. Sinopoli, and A. Casavola. LQG control over lossy TCP-like networks with probabilistic packet acknowledgements. *Int. J. Syst., Control Commun.*, 2(1):55–81, 2010.

[8] T. Hu and Z. Lin. *Control Systems with Actuator Saturation: Analysis and Design*. Springer Science & Business Media, 2001.

[9] M. Huang and S. Dey. Stability of Kalman filtering with Markovian packet losses. *Automatica*, 43(4):598–607, 2007.

[10] H. Ishii. Limitations in remote stabilization over unreliable channels without acknowledgements. *Automatica*, 45(10):2278–2285, 2009.

[11] A. H. Jazwinski. *Stochastic Processes and Filtering Theory*. Academic Press: New York, 1970.

[12] A. S. Leong and D. E. Quevedo. Kalman filtering with relays over wireless fading channels. *IEEE Trans. Automat Control*, 61(6):1643–1648, 2016.

[13] X. R. Li and Y. Bar-Shalom. Performance prediction of the interacting multiple model algorithm. *IEEE Trans. Aerosp. Electron. Syst.*, 29(3):755–771, 1993.

[14] Y. Liang, T. Chen, and Q. Pan. Optimal linear state estimator with multiple packet dropouts. *IEEE Trans. Automat. Control*, 55(6):1428–1433, 2010.

[15] H. Lin, H. Su, P. Shi, R. Lu, and Z.-G. Wu. LQG control for networked control systems over packet drop links without packet acknowledgment. *J. Franklin Inst.*, 352(11):5042–5060, 2015.

[16] H. Lin, H. Su, P. Shi, R. Lu, and Z.-G. Wu. Estimation and LQG control over unreliable network with acknowledgment randomly lost. *IEEE Trans. Cybern.*, 47(12):4074–4085, 2017.

[17] H. Lin, H. Su, P. Shi, Z. Shu, R. Lu, and Z.-G. Wu. Optimal estimation and control for lossy network: Stability, convergence, and performance. *IEEE Trans. Automat. Control*, 62(9):4564–4579, 2017.

[18] H. Lin, H. Su, Z. Shu, Z.-G. Wu, and Y. Xu. Optimal estimation for networked control systems with intermittent inputs without acknowledgement. In *Proc. 19th IFAC World Congress*, pages 5017–5022, 2014.

[19] H. Lin, H. Su, Z. Shu, Z.-G. Wu, and Y. Xu. Optimal estimation in UDP-like networked control systems with intermittent inputs: Stability analysis and suboptimal filter design. *IEEE Trans. Automat. Control*, 61(7):1794–1809, 2016.

[20] H. Lin, Z. Xu, H. Su, Y. Xu, and Z.-G. Wu. Fast filtering algorithm for state estimation of lossy networks. *IET Control Theory & Appl.*, 8(18):2316–2324, 2014.

[21] J. Liu, J. Xia, E. Tian, and S. Fei. Hybrid-driven-based H_∞ filter design for neural networks subject to deception attacks. *Applied Mathematics and Computation*, 320:158–174, 2018.

[22] J. Liu, L. Zha, J. Cao, and S. Fei. Hybrid-driven-based stabilisation for networked control systems. *IET Control Theory & Appl.*, 10(17):2279–2285, 2016.

[23] Y. Mo and B. Sinopoli. Kalman filtering with intermittent observations: tail distribution and critical value. *IEEE Trans. Automat. Control*, 57(3):677–689, 2012.

[24] M. Moayedi, Y. K. Foo, and Y. C. Soh. Networked LQG control over unreliable channels. *Int. J. Robust Nonlin. Control*, 23(2):167–189, 2013.

[25] K. Plarre and F. Bullo. On Kalman filtering for detectable systems with intermittent observations. *IEEE Trans. Automat. Control*, 54(2):386–390, 2009.

[26] N. J. Ploplys, P. A. Kawka, and A. G. Alleyne. Closed-loop control over wireless networks. *IEEE Trans. Control Syst.*, 24(3):58–71, 2004.

[27] S. M. Ross. *Introduction to Probability Models*. Academic Press, 2006.

[28] L. Schenato, B. Sinopoli, M. Franceschetti, K. Poolla, and S. S. Sastry. Foundations of control and estimation over lossy networks. *Proc. IEEE*, 95(1):163–187, 2007.

[29] C. E. Seah and I. Hwang. Stability analysis of the interacting multiple model algorithm. In *ACC*, pages 2415–2420, 2008.

[30] D. Simon. *Optimal State Estimation: Kalman, H_∞, and Nonlinear Approaches*. John Wiley & Sons, 2006.

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

[32] B. Sinopoli, L. Schenato, M. Franceschetti, K. Poolla, and S. Sastry. Optimal linear LQG control over lossy networks without packet acknowledgment. *Asian J. Control*, 10(1):3–13, 2008.

[33] J. L. Speyer and W. H. Chung. *Stochastic Processes, Estimation, and Control*. SIAM, 2008.

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

[35] 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 Networks Learning Syst.*, DOI:10.1109/TNNLS.2017.2732240, 2017.

[36] K. You, M. Fu, and L. Xie. Mean square stability for Kalman filtering with Markovian packet losses. *Automatica*, 47(12):2647–2657, 2011.

[37] H. Zhang, R. Yang, H. Yan, and F. Yang. H_∞ consensus of event-based multi-agent systems with switching topology. *Information Sciences*, 370:623–635, 2016.