

Performance and Stability Analysis of Interacting Multiple Model Estimator Under Unobservable Packet Loss

Hong Lin, James Lam, Zidong Wang, Zhan Shu

***Abstract*—For a system with packet loss, if the estimator cannot observe the packet-loss status (PLS), the system is called an unobservable-packet-loss (UPL) system. Otherwise, it is called an observable-packet-loss (OPL) system. This paper studies the interacting multiple model (IMM) estimator for UPL systems, and the main contributions are twofold. (i) Estimation accuracy of the unobservable PLS. For an *unstable* UPL system, we prove that the UPL system will become an OPL one with time, since the PLS can be exactly estimated with time. For a *stable* UPL system, there exists an accuracy threshold such that the estimation accuracy of the PLS cannot be better than this threshold. (ii) Stability of the IMM estimator. For an *unstable* UPL system, we establish a necessary and sufficient condition: there exists a threshold such that the IMM estimator is stable almost everywhere if and only if the packet-arrival rate is greater than this threshold. For a *stable* UPL system, we show that the IMM estimator is stable, no matter what value the packet-arrival rate is.**

***Index Terms*—Interacting multiple model estimator; Stability; Unobservable packet loss;**

I. Introduction

A. Research background

In the past two decades, the rapid development of network techniques allows the components of a system to be connected via networks. The introduction of networks facilitates information sharing, but it also causes packet loss [1, 2] and communication delay [3]. A large number of state estimation and control techniques have been developed for networked systems to deal with packet loss such as distributed filtering [4–6], neural-network-based filtering [7], event-based estimation [8, 9], feedback control [10], particle filtering [11], optimal H_2/H_∞ filtering [12], and least squares estimation [13].

According to the observability of packet-loss status (PLS), systems can be divided into two classes: systems with unobservable packet loss (UPL) and systems with observable packet loss (OPL).

Typical applications of the UPL system model can be found in the literature on radar-based target tracking. Subject to the environmental shelter or intermittent appearance of targets (see Figure 1), radar waves cannot always reach the targets, and the reflected radar waves may consist of noises alone [17, 18]. However, the radar system cannot identify whether the target is detected or missed from the reflected wave, and the loss of the waves reflected from targets is unobservable.

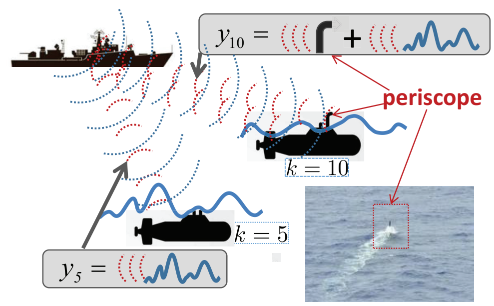


Fig. 1. Detect a periscope in a region of the ocean's surface [18]. The periscope will be exposed every 10 seconds. The high-intensity clutter of ocean waves may produce a clutter in the reflected radar waves that looks like the waves reflected from the target. Hence, subject to intermittent appearance of the periscope and ocean waves, the reflected radar waves may come from the periscope and ocean waves (y_{10}) or from ocean waves alone (y_5).

For UPL systems, approximate optimal estimators (AOEs), instead of the optimal estimator (OE), are commonly used in practice. It was reported in [15] that the OE for UPL systems cannot be implemented in practice, since the computational complexity of the OE grows exponentially with time. Since then, various computationally efficient AOEs have been developed, including linear estimators [19–22] and nonlinear ones [16, 23]. The Interacting Multiple Model (IMM) algorithm was first developed in [24]. Thanks to its excellent estimation performance and low computational complexity, it has been widely used in state estimation and target tracking [25–27]. These successful and widespread applications have proven that the IMM algorithm is a valid and credible technique to compute the optimal state estimate, and thus this paper adopts the IMM estimator for UPL systems.

B. Underlying issues and limitations of existing research

The estimator stability for *unstable* systems with packet loss is a hot research topic. Since the pioneering work [28]

Hong Lin is with Institute of Intelligence Science and Engineering, Shenzhen Polytechnic University, Shenzhen 518055, China. E-mail: linhongzju@163.com

James Lam is with Department of Mechanical Engineering, The University of Hong Kong, Pokfulam, Hong Kong. E-mail: james.lam@hku.hk

Zidong Wang is with Department of Computer Science, Brunel University London, Uxbridge, Middlesex, UB8 3PH, U.K. E-mail: Zidong.wang@brunel.ac.uk

Zhan Shu is with Department of Electrical and Computer Engineering, Faculty of Engineering, University of Alberta, Edmonton, Alberta, Canada. E-mail: zhan.shu@ualberta.ca

The work was supported in part by National Natural Science Foundation of China under Grant (62273286) and Grant (61873148), in part by the Scientific Research Startup Fund for Shenzhen High-Caliber Personnel of SZPT (6021310027K), in part Shenzhen Polytechnic University Research Fund (602331015PQ), in part Industry-University-Research Innovation Fund of Chinese University (2021ZYB05002), and in part by GRF (17203922).

established a condition on estimator stability for unstable OPL systems in 2004, fruitful results on estimator stability have been obtained for unstable OPL systems [29–32] in the past two decades. However, there are few advances for unstable UPL systems. The works on estimator stability for UPL systems are introduced as follows. *Linear estimators*: It was early known in [19, 20] that the condition "the system is stable (that is, the spectrum radius $\rho(A) < 1$)" is a sufficient one for the optimal linear estimator (OLE) is to be stable. Our recent works showed that $\rho(A) < 1$ is also a necessary condition for the stability of the OLE with multi-sensors [21, 33] or with Markovian packet loss [22]. *Nonlinear estimators*: Bayesian-filtering-based AOEs [16] and a Kullback-Leibler-Divergence (KLD)-based AOE [23] were developed for UPL systems. It was proved that the KLD-based AOE is stable for a stable system. It can be seen from above that **(Problem 1)** "for an *unstable* UPL system, under what condition, can an estimator remain stable" is unsolved.

The observability of PLS (that is, the value of γ_k in (2)) greatly affects estimator design and performance. It is known that (i) that the loss of observability of γ_k makes the estimators for a UPL system significantly differ from the estimators for an OPL system, and (ii) that for an OPL system, knowing the PLS facilitates the theoretical analysis of estimator stability and improves estimation performance. However, **(Problem 2)** how accurate γ_k can be estimated for a UPL system is unknown?

From Problems 1 and 2 above, the limitations of the existing bodies of research on UPL systems are twofold: (1) Estimator stability for unstable UPL systems remains unknown. Existing works did not study the estimation accuracy of PLS. The research challenges are twofold. Although some notations below are defined in the following sections, they would not affect the explanations. (i) For Problem 2, the averaged estimation accuracy of γ_k requires computing an integral like $\int \eta$ in (23), where $\eta \triangleq \frac{\phi_k \psi_k}{\phi_k + \psi_k}$, and ϕ_k and ψ_k are Gaussian probability density functions. There is no analytical expression of the integral, and thus it is difficult to determine its limit as in (24) and its lower and upper boundedness as in (25) and (48). (ii) For Problem 1, the boundedness of $\int \eta(\overline{m}_k - m_k)_I^2$ is hard to be determined, which makes it difficult to analyze the stability of $\mathbb{E}[P_k]$ in (9).

C. Main results and contributions

To the best of our knowledge, few results have been reported on Problems 1 and 2, which motivated our research of this paper. The main results and contributions are summarized as follows:

(i) **Stability of IMM estimator.**
- When $\rho(A) \geq 1$, a **necessary and sufficient condition** is established for the stability of the IMM estimator: there exists a threshold λ_γ such that the IMM estimator is stable almost everywhere if and only if the packet-arrival rate is greater than λ_γ.
- When $\rho(A) < 1$, the IMM estimator is stable, no matter what value the packet-arrival rate is.

(ii) **Estimation accuracy of γ_k.**
- When $\rho(A) \geq 1$, the estimation errors of γ_k are proved to converge to 0. It means that γ_k can be exactly estimated with time. This result suggests that an unstable UPL system will become an OPL one with time.
- When $\rho(A) < 1$, there exists an **accuracy threshold** Γ_γ such that the estimation accuracy of γ_k cannot exceed (be better than) this threshold. This result suggests that an unstable UPL system has a better estimation accuracy of γ_k than a stable UPL system.

D. Novelties and comparison with existing works

- The main result (i) above is novel, since it establishes a necessary and sufficient condition for the proposed IMM estimator. Existing works on UPL systems do not study estimator stability for unstable UPL systems. Specifically, linear estimators have been proved to be unstable for unstable UPL systems [21, 22]. Whether the non-linear estimation methods [16, 22, 23] are stable for unstable systems is uncertain.
- The main result (ii) on packet-loss-status estimation is new, as it reveals the relationship between the estimation accuracy of the PLS and the stability of the system, which has not yet been reported in existing published pieces of research.

E. Paper Organization

The rest of the paper is organized as follows: The system setup is introduced in Section II. The IMM estimator for UPL systems is derived in Sections III. Estimation accuracy of γ_k and stability of the IMM estimator are studied in Sections IV and V, respectively. In Section VI, numerical examples are given to illustrate the main results of this paper. The conclusions are presented in Section VII.

Notation

$x \sim \mathcal{N}_x(\mu, P)$	Random variable x follows a Gaussian probability density function (pdf) with mean μ and covariance P.
$p(\cdot)$, $p(\cdot\|\cdot)$	pdf and conditional pdf, respectively.
$\mathbb{E}[\cdot]$, $\mathbb{E}_x[\cdot]$	$\mathbb{E}[\cdot]$ stands for mathematical expectation. $\mathbb{E}_x[\cdot]$ is used to emphasize that $\mathbb{E}[\cdot]$ is taken with respect to random variable x.
$\mathbb{P}(\cdot)$	Probability measure
$\mathbb{Cov}[\cdot]$	Covariance
$(\cdot)'$, $(\cdot)^*$	Transpose and conjugate transpose of a matrix or a vector, respectively
$(\cdot)_I^2$	$(\cdot)(\cdot)'$
$\rho(M)$	Spectral radius of matrix M
$\overline{\sigma}(M)$, $\underline{\sigma}(M)$	Maximum and minimum singular values of matrix M, respectively.
$M > (\geq) 0$	M is a positive definite (semi-definite) matrix.
$\mathbb{Z}_+$	Set of non-negative integers

II. System Setup and Research Objectives

Consider the following discrete-time linear system:

- **Plant:**
$$x_k = Ax_{k-1} + \omega_k, \tag{1}$$
where $A \in \mathbb{R}^{n\times n}$, $x_k \in \mathbb{R}^n$ is the system state, $x_0 \sim \mathcal{N}_{x_0}(\widehat{x}_0, P_0)$, and $\omega_k \sim \mathcal{N}_{\omega_k}(0, Q)$ is the system process noise, where $P_0 > 0$ and $Q > 0$.
- **Observations:**
$$y_k = \gamma_k C x_k + \upsilon_k, \tag{2}$$
Cx_k is the output of the sensor, where $C \in \mathbb{R}^{m\times n}$. It is assumed throughout this paper that C has full row rank. It is a mild assumption, since when $\mathsf{rank}(C) < m$, there are linear dependent rows in C, which are redundant as they can be represented by the linear independent rows. Thus, they can be removed such that $\mathsf{rank}(C) = m$. $\upsilon_k \sim \mathcal{N}_{\upsilon_k}(0, R)$ is the noise at the estimator side, where $R > 0$. $\{\gamma_k\}$ is a sequence of i.i.d. Bernoulli random variables with $\mathbb{P}(\gamma_k = 1) = \gamma$ and $\mathbb{P}(\gamma_k = 0) = \overline{\gamma} = 1 - \gamma$. γ_k models the packet loss: $\gamma_k = 1$ stands for that the sensor output is successfully received by the estimator. Otherwise, $\gamma_k = 0$. γ is called the **packet-arrival rate**. Denote all the observations received at the estimator side by $Y_k \triangleq \{y_1, \ldots, y_k\}$, where $y_j \in \mathbb{R}^m$.
- **IMM estimator:** As previously mentioned, the optimal estimate for a UPL system cannot be computed in practice. Hence, in this paper, the optimal estimate is computed by the standard IMM algorithm [34, 35].

This paper studies the IMM estimator for UPL systems. The research objectives are listed as follows:

(i) Determine the stability conditions of the IMM estimator for UPL systems.
(ii) Determine the estimation accuracy of the unobservable packet-loss status γ_k.

It is assumed in this paper that $(A, Q^{1/2})$ is controllable and (A, C) is observable. ω_k, υ_k, γ_k, and x_0 are mutually independent.

III. IMM Estimator for UPL Systems

In this section, we first give the definition of optimal estimation and estimation accuracy. Then, we compute optimal estimates $\widehat{x}_k$ and $\widehat{\gamma}_k$ by the IMM algorithm (that is, the IMM estimator) for UPL systems. Finally, we design auxiliary IMM estimators, which will be used to study the estimation accuracy of $\widehat{\gamma}_k$ and the stability of the IMM estimator in Sections IV and V, respectively.

A. Definitions and preliminaries

Definition 1 (Optimal estimation)
The optimal estimation of γ_k and x_k are defined as
$$\widehat{\gamma}_k \triangleq \arg\min_{\gamma^\sharp \in [0,1]} \mathbb{E}[(\gamma_k - \gamma^\sharp)^2 | Y_k]$$
$$\widehat{x}_k \triangleq \arg\min_{x^\sharp \in \mathbb{R}^n} \mathbb{E}[(x_k - x^\sharp)^2_I | Y_k].$$
The estimation accuracy of γ_k is defined as
$$\Gamma_k \triangleq \mathbb{Cov}(\gamma_k | Y_k) = \mathbb{E}[(\gamma_k - \widehat{\gamma}_k)^2 | Y_k].$$
The estimation error covariance of x_k is defined as
$$P_k \triangleq \mathbb{Cov}(x_k | Y_k) = \mathbb{E}[(x_k - \widehat{x}_k)^2_I | Y_k].$$

Denote an m-dimensional cube with edge length $l^{\frac{1}{4m}}$ by
$$\mathbb{C}_l \triangleq \left\{ y = [y(1), \ldots, y(m)]' \in \mathbb{R}^m \,\middle|\, |y(i)| \le \frac{l^{\frac{1}{4m}}}{2}, 1 \le i \le m \right\}.$$
The volume of $\mathbb{C}_l$ is $(l^{\frac{1}{4m}})^m = l^{\frac{1}{4}}$.

Define $s_k \triangleq \{y_k \in \mathbb{C}_k\}$. Then, $s_k^c = \{y_k \notin \mathbb{C}_k\}$. Define a random variable S_k as follows:
$$S_k = \begin{cases} s_k, & \text{if } \; y_k \in \mathbb{C}_k \\ s_k^c, & \text{otherwise.} \end{cases}$$
S_k takes the value s_k and s_k^c and is a function of y_k.

B. IMM estimator for UPL systems

First, we derive the IMM estimator for UPL systems.

Theorem 1 (IMM estimator for UPL systems) *The optimal estimates $\widehat{\gamma}_k$ and $\widehat{x}_k$ computed by using the IMM algorithm are given as follows:*
$$\overline{m}_k = A\widehat{x}_{k-1} \tag{3}$$
$$\overline{M}_k = AP_{k-1}A' + Q \tag{4}$$
$$K_k = \overline{M}_k C'(C\overline{M}_k C' + R)^{-1}$$
$$m_k = \overline{m}_k + K_k(y_k - C\overline{m}_k)$$
$$M_k = \overline{M}_k - \overline{M}_k C'(C\overline{M}_k C' + R)^{-1} C'\overline{M}_k \tag{5}$$
and
$$\phi_k = \mathcal{N}_{y_k}(0, R), \quad \psi_k = \mathcal{N}_{y_k}(C\overline{m}_k, C\overline{M}_k C' + R)$$
$$\widehat{\gamma}_k = \frac{\gamma\psi_k}{\overline{\gamma}\phi_k + \gamma\psi_k} \tag{6}$$
$$\Gamma_k = \frac{\overline{\gamma}\gamma\phi_k\psi_k}{(\overline{\gamma}\phi_k + \gamma\psi_k)^2} \tag{7}$$
$$\widehat{x}_k = (1 - \widehat{\gamma}_k)\overline{m}_k + \widehat{\gamma}_k m_k \tag{8}$$
$$P_k = (1 - \widehat{\gamma}_k)\big(\overline{M}_k + (\widehat{x}_k - \overline{m}_k)^2_I\big) + \widehat{\gamma}_k\big(M_k + (\widehat{x}_k - m_k)^2_I\big). \tag{9}$$

Proof: It is known that the optimal estimate $\widehat{x}_k$ is computed by $\widehat{x}_k = \mathbb{E}[x_k|Y_k] = \int_{\mathbb{R}^n} x_k p(x_k|Y_k)\mathsf{d}x_k$. By the total probability law,
$$p(x_k|Y_k) = \textstyle\sum_{j=0}^1 p(x_k|\gamma_k = j, Y_k)\mathbb{P}(\gamma_k = j|Y_k). \tag{10}$$
In the standard IMM algorithm [34, 35], $p(x_k|\gamma_k = j, Y_k)$ is computed by letting $p(x_{k-1}|\gamma_k = j, Y_{k-1}) = \mathcal{N}_{x_{k-1}}(\widehat{x}_{k-1}, P_{k-1})$, where and its mean and covariance are $\mathbb{E}[x_{k-1}|\gamma_k = j, Y_{k-1}]$ and $\mathbb{Cov}[x_{k-1}|\gamma_k = j, Y_{k-1}]$, respectively. γ_k is independent of x_{k-1} and Y_{k-1}, and thus $p(x_{k-1}|\gamma_k = j, Y_{k-1}) = p(x_{k-1}|Y_{k-1})$, $\mathbb{E}[x_{k-1}|\gamma_k = j, Y_{k-1}] = \mathbb{E}[x_{k-1}|Y_{k-1}] = \widehat{x}_{k-1}$, and $\mathbb{Cov}[x_{k-1}|\gamma_k = j, Y_{k-1}] = \mathbb{Cov}[x_{k-1}|Y_{k-1}] = P_{k-1}$. Consequently,

$p(x_{k-1}|Y_{k-1}) = \mathcal{N}_{x_{k-1}}(\widehat{x}_{k-1}, P_{k-1})$. From $x_k = Ax_{k-1}+\omega_k$ in (1), it follows that

$$p(x_k|Y_{k-1}) = \mathcal{N}_{x_k}(\overline{m}_k, \overline{M}_k), \tag{11}$$

where $\overline{m}_k = A\widehat{x}_{k-1}$ and $\overline{M}_k = AP_{k-1}A' + Q$.

Computation of $p(x_k|\gamma_k = j, Y_k)$ in (10): When $\gamma_k = 0$, $y_k = \upsilon_k$ is independent of x_k and Y_{k-1}. $p(x_k|y_k, Y_{k-1}, \gamma_k = 0) = p(x_k|Y_{k-1}, \gamma_k = 0) = p(x_k|Y_{k-1})$. By (11), $p(x_k|Y_k, \gamma_k = 0) = \mathcal{N}_{x_k}(\overline{m}_k, \overline{M}_k)$, which means

$$\overline{M}_k = \mathbb{E}[(x_k - \overline{m}_k)_I^2|\gamma_k = 0, Y_k]. \tag{12}$$

When $\gamma_k = 0$, $y_k = \upsilon_k$. $p(y_k|Y_{k-1}, \gamma_k = 0) = \mathcal{N}_{y_k}(0, R) = \phi_k$.

When $\gamma_k = 1$, $y_k = Cx_k + \upsilon_k$. For $p(x_k|Y_{k-1})$ in (11), it is well known that

$$p(y_k|Y_{k-1}, \gamma_k = 1) = \mathcal{N}_{y_k}(C\overline{m}_k, C\overline{M}_kC' + R) = \psi_k \tag{13}$$
$$p(x_k|Y_k, \gamma_k = 1) = \mathcal{N}_{x_k}(m_k, M_k),$$

where m_k and M_k are computed as above, and

$$M_k = \mathbb{E}[(x_k - m_k)_I^2|\gamma_k = 1, Y_k]. \tag{14}$$

By the total probability law and the independence of γ_k and Y_{k-1},

$$\begin{aligned} p(y_k|Y_{k-1}) &= \sum_{j=0}^{1} p(y_k|Y_{k-1}, \gamma_k = j)\mathbb{P}(\gamma_k = j|Y_{k-1}) \\ &= \psi_k\gamma + \phi_k\overline{\gamma}. \end{aligned} \tag{15}$$

Computation of $\mathbb{P}(\gamma_k = j|Y_k)$, $\widehat{\gamma}_k$ in (6), and Γ_k in (7): It is known that the optimal estimation $\widehat{\gamma}_k = \mathbb{E}[\gamma_k|Y_k]$. Then, $\mathbb{E}[\gamma_k|Y_k] = 0 \cdot \mathbb{P}(\gamma_k = 0) + 1 \cdot \mathbb{P}(\gamma_k = 1) = \mathbb{P}(\gamma_k = 1)$, which is computed as follows. By Bayesian formula, (13), and (15),

$$\begin{aligned} \mathbb{P}(\gamma_k = 1|Y_k) &= \frac{p(y_k|Y_{k-1}, \gamma_k = 1)p(\gamma_k = 1|Y_{k-1})}{p(y_k|Y_{k-1})} \\ &= \frac{\psi_k\gamma}{\psi_k\gamma + \phi_k\overline{\gamma}} = \widehat{\gamma}_k \\ \mathbb{P}(\gamma_k = 0|Y_k) &= 1 - \mathbb{P}(\gamma_k = 1|Y_k) = 1 - \widehat{\gamma}_k. \end{aligned}$$

Clearly, for a Bernoulli random variable γ_k, $\Gamma_k = \widehat{\gamma}_k(1-\widehat{\gamma}_k)$, and thus (7) holds.

Computation of $\widehat{x}_k$ in (8) and P_k in (9): Based on the results above and (10), we have $p(x_k|Y_k) = (1 - \widehat{\gamma}_k)\mathcal{N}_{x_k}(\overline{m}_k, \overline{M}_k) + \widehat{\gamma}_k\mathcal{N}_{x_k}(m_k, M_k)$. For such a Gaussian mixture $p(x_k|Y_k)$, it has been obtained in [36, p. 213] that $\widehat{x}_k = \mathbb{E}[x_k|Y_k]$ and $P_k = \mathbb{E}[(x_k - \widehat{x}_k)_I^2|Y_k]$, and they equals (8) and (9), respectively. The proof is completed. ■

The computational complexity (CC) of the proposed IMM estimator in Theorem 1 is analyzed as follows. Denote the CC of Kalman filtering (KF) equations from (3) to (5) by c_k. Denote the CC of a pdf $\mathcal{N}_x(\cdot)$ by c_p, and denote the CC of $a_1\times w$, $a_1\times(W+(v_1+v_2)_I^2)$, and $a_1\times a_2/(a_3\times a_4+a_5\times a_6)$ by c_1, c_2, and c_3, respectively, where $a_i \in \mathbb{R}$, w and v_i $(i = 1, 2)$ are vectors, and W is a matrix.

Corollary 1 *For time instant k, the CC of the IMM estimator is $c_k + 2(c_p + c_1 + c_2 + c_3)$.*

Proof: First, the CC of the KF equations from (3) to (5) is c_k. Then, two Gaussian pdfs ϕ_k and ψ_k are computed with the CC $2c_p$. Since $1-\widehat{\gamma}_k = \frac{\overline{\gamma}\phi_k}{\overline{\gamma}\phi_k+\gamma\psi_k}$, $\widehat{\gamma}_k$ and $1-\widehat{\gamma}_k$ are calculated with the same CC c_3. Finally, the CC of $\widehat{x}_k$ and P_k in (8) and (9) is $2(c_1+c_2)$. It needs to mention that Γ_k is used when analyzing the estimation accuracy of $\widehat{\gamma}_k$, and there is no need to calculate it in practice. Therefore, the overall CC of the IMM estimator is $c_k + 2(c_p + c_1 + c_2 + c_3)$. The proof is completed. ■

C. Auxiliary IMM estimators

Based on $\overline{m}_k$ and m_k in Theorem 1, we construct two auxiliary IMM estimators $\widetilde{x}_k$ and $\widetilde{x}_k^\sharp$ as follows:

$$\widetilde{\gamma}_k = \begin{cases} 0, & \text{if} \quad S_k = s_k \\ 1, & \text{if} \quad S_k = s_k^c. \end{cases} \tag{16}$$

$$\widetilde{x}_k \triangleq (1 - \widetilde{\gamma}_k)\overline{m}_k + \widetilde{\gamma}_k m_k \tag{17}$$

$$\widetilde{x}_k^\sharp \triangleq (1 - \gamma_k)\overline{m}_k + \gamma_k m_k. \tag{18}$$

Their properties presented in the following Lemma 1 will be used in the next section to study the stability of the IMM estimator.

Lemma 1 *The following properties hold.*

$$\mathbb{E}[\mathbb{Cov}(\gamma_k|Y_k)] \leq \mathbb{E}[(\gamma_k - \widetilde{\gamma}_k)^2|S_k] \tag{19}$$

$$\widetilde{P}_k \triangleq \mathbb{E}[(x_k - \widetilde{x}_k)_I^2|Y_k] \geq P_k. \tag{20}$$

$$P_k^\sharp \triangleq \mathbb{E}[(x_k - \widetilde{x}_k^\sharp)_I^2|\gamma_k, Y_k] = (1 - \gamma_k)\overline{M}_k + \gamma_k M_k. \tag{21}$$

Proof of (19): It is shown in [37, Proposition 3.1] that for two random quantities X and Y, $\mathbb{Cov}(X) = \mathbb{E}[\mathbb{Cov}(X|Y)] + \mathbb{Cov}(\mathbb{E}[X|Y])$. Due to $\mathbb{Cov}(\mathbb{E}[X|Y]) \geq 0$,

$$\mathbb{Cov}(X) \geq \mathbb{E}[\mathbb{Cov}(X|Y)]. \tag{22}$$

By this result, $\mathbb{Cov}(\gamma_k|S_k) \geq \mathbb{E}[\mathbb{Cov}(\gamma_k|S_k, y_k, Y_{k-1})]$. Note that S_k is a function of y_k, and therefore knowing y_k is equivalent to knowing $\{S_k, y_k\}$. Thus, $\mathbb{E}[\mathbb{Cov}(\gamma_k|S_k, y_k, Y_{k-1})] = \mathbb{E}[\mathbb{Cov}(\gamma_k|y_k, Y_{k-1})]$. Then, we have $\mathbb{Cov}(\gamma_k|S_k) \geq \mathbb{E}[\mathbb{Cov}(\gamma_k|Y_k)]$.

Let $\widetilde{\gamma}_k^\sharp \triangleq \mathbb{E}[\gamma_k|S_k]$. By the definition of covariance, $\mathbb{Cov}(\gamma_k|S_k) = \mathbb{E}[(\gamma_k - \widetilde{\gamma}_k^\sharp)^2|S_k]$. It is known that $\widetilde{\gamma}_k^\sharp = \mathbb{E}[\gamma_k|S_k]$ is the one that minimizes $\mathbb{E}[(\gamma_k - \widetilde{\gamma}_k^\sharp)^2|S_k]$. Hence, $\mathbb{E}[(\gamma_k - \widetilde{\gamma}_k^\sharp)^2|S_k] \leq \mathbb{E}[(\gamma_k - \widetilde{\gamma}_k)^2|S_k]$. Therefore, $\mathbb{Cov}(\gamma_k|S_k) \leq \mathbb{E}[(\gamma_k - \widetilde{\gamma}_k)^2|S_k]$. Thus, (19) holds due to $\mathbb{Cov}(\gamma_k|S_k) \geq \mathbb{E}[\mathbb{Cov}(\gamma_k|Y_k)]$ obtained above.

Proof of (20): It is well known that $\widehat{x}_k = \mathbb{E}[x_k|Y_k]$ minimizes $\mathbb{E}[(x_k - x^\sharp)_I^2|Y_k]$ for $x^\sharp \in \mathbb{R}^n$, which means $P_k = \mathbb{E}[(x_k - \widehat{x}_k)_I^2|Y_k] \leq \mathbb{E}[(x_k - \widetilde{x}_k)_I^2|Y_k] = \widetilde{P}_k$, which proves (20).

Proof of (21): When $\gamma_k = 0$, $\widetilde{x}_k^\sharp = \overline{m}_k$ in (8). Then, by (12), $\mathbb{E}[(x_k - \widetilde{x}_k^\sharp)_I^2|\gamma_k = 0, Y_k] = \mathbb{E}[(x_k - \overline{m}_k)_I^2|\gamma_k = 0, Y_k] = \overline{M}_k$. Similarly, $\gamma_k = 1$, $\widetilde{x}_k^\sharp = m_k$ in (8). Then, by (14), $\mathbb{E}[(x_k - \widetilde{x}_k^\sharp)_I^2|\gamma_k = 1, Y_k] = M_k$. Consequently, $P_k^\sharp = (1 - \gamma_k)\overline{M}_k + \gamma_k M_k$.

The proof is completed. ■

IV. Estimation Accuracy of γ_k for UPL Systems

This section aims at solving Problem 2. The results on the estimation accuracy of γ_k for stable and unstable UPL systems are given in Theorem 2.

$\mathbb{Cov}(\gamma_k|Y_k)$ contains random quantity Y_k, and thus the estimation accuracy is accessed in the mean sense, that is,

$$\mathbb{E}[\mathbb{Cov}(\gamma_k|Y_k)] = \int_{\mathbb{R}^m} \frac{\overline{\gamma}\gamma\psi_k\phi_k}{\overline{\gamma}\phi_k + \gamma\psi_k} \mathsf{d}y_k, \tag{23}$$

where due to the independence of γ_k and Y_{k-1}, the integral is computed with respect to y_k rather than Y_k.

Theorem 2 (Estimation accuracy of γ_k) *For a UPL system,*

(i) when $\rho(A) \geq 1$,

$$\lim_{k\to\infty} \mathbb{E}[\mathbb{Cov}(\gamma_k|Y_k)] = 0. \tag{24}$$

(ii) When $\rho(A) < 1$, there is an **accuracy threshold** Γ_γ *such that*

$$\mathbb{E}[\mathbb{Cov}(\gamma_k|Y_k)] \geq \Gamma_\gamma > 0. \tag{25}$$

Theorem 2(i) and (ii) are proved in the following two subsections, respectively.

Remark 1 An unstable UPL system will become an OPL one, *since the estimation of γ_k becomes accurate with time when $\rho(A) > 1$, as shown in Theorem 2. In other words, γ_k can be exactly estimated without errors, and the packet-loss status become unknown (observable), that is, an OPL system. However, for a stable UPL system, there is an accuracy threshold such that the estimation accuracy cannot be better than this threshold. The reason for such a phenomenon is explained in Remark 2 at the end of this section.*

A. Proof of Theorem 2(i): The case $\rho(A) \geq 1$

The proof of Theorem 2(i) involves the convergence of the probabilities of some events (e.g. (29c)), and formula (26) will be involved in analysing the convergence. The following Lemmas 2–4 are given to provide the relevant properties for proving Theorem 2(i).

Lemma 2 *Let $W_k \triangleq \sum_{j=0}^{k-1} A^j Q (A^j)'$. When $\rho(A) \geq 1$,*

$$\lim_{k\to+\infty} \frac{k^{\frac{1}{4}}}{\sqrt{\det(CW_kC')}} = 0. \tag{26}$$

Proof: There exists an invertible matrix H such that $A = HJH^{-1}$, where J is the Jordan canonical form of A. For simplicity, we assume that $J = \left[\begin{smallmatrix} J_1 & 0 \\ 0 & J_2 \end{smallmatrix}\right]$ has two Jordan blocks, where J_1 and J_2 correspond to the eigenvalues λ_1 and λ_2, respectively. $|\lambda_1| = \rho(A) \geq 1$, $\lambda_1 \neq \lambda_2$, and $|\lambda_1| \geq |\lambda_2|$. For $j = 1$ or 2, J_j may contain multiple Jordan sub-blocks with λ_j on the diagonal. The case that the number of distinct eigenvalues is greater than two can be dealt with in the same way as follows.

Denote $\underline{\sigma}_q \triangleq \underline{\sigma}(Q)$. $W_k \geq \underline{\sigma}_q \sum_{j=0}^{k-1} A^j(A^j)'$. Since $A \in \mathbb{R}^{n\times n}$, $A^j(A^j)' = A^j(A^j)^* = HJ^jH^{-1}(HJ^jH^{-1})^*$. Let $\underline{\sigma}_H \triangleq \underline{\sigma}(H^{-1}(H^{-1})^*)$. $A^j(A^j)' \geq \underline{\sigma}_H HJ^j(J^*)^jH^*$. Note that $J^0(J^0)^* = I$, $J^j = \left[\begin{smallmatrix} J_1^j & 0 \\ 0 & J_2^j \end{smallmatrix}\right]$, and $J^j(J^*)^j = \left[\begin{smallmatrix} J_1^j(J_1^j)^* & 0 \\ 0 & J_2^j(J_2^j)^* \end{smallmatrix}\right]$. Let

$$\mathcal{J}_k \triangleq \begin{bmatrix} I_1 + \sum_{j=1}^{k-1} J_1^j(J_1^j)^* & 0 \\ 0 & I_2 + \sum_{j=1}^{k-1} J_2^j(J_2^j)^* \end{bmatrix},$$

where I_j is the identity matrix with the same dimension as J_j, $j = 1, 2$.

$$\begin{aligned}\sum_{j=0}^{k-1} A^j(A^j)' &\geq \underline{\sigma}_H H\Big(I + \sum_{j=1}^{k-1} J^j(J^*)^j\Big)H^* \\ &\geq \underline{\sigma}_H H\mathcal{J}_kH^*.\end{aligned} \tag{27}$$

By the property [38, Corollary 8.4.15] that $\det(X+Y) \geq \det(X) + \det(Y)$ for $X \geq 0, Y \geq 0$,

$$\begin{aligned}\det(\mathcal{J}_k) &= \det\Big(I_1 + \sum_{j=1}^{k-1} J_1^j(J_1^j)^*\Big)\det\Big(I_2 + \sum_{j=1}^{k-1} J_2^j(J_2^j)^*\Big) \\ &\geq \Big(1 + \sum_{j=1}^{k-1}\det(J_1^j(J_1^j)^*)\Big)\Big(1 + \sum_{j=1}^{k-1}\det(J_2^j(J_2^j)^*)\Big) \\ &\geq 1 + \sum_{j=1}^{k-1}\det\Big(J_1^j(J_1^j)^*\Big),\end{aligned} \tag{28}$$

where the last inequality is obtained by noting that $\sum_{j=1}^{k-1}\det(J_2^j(J_2^j)^* \geq 0$.

By the property $\det(XY) = \det(X)\det(Y)$ [38, 18, Proposition 2.7.3],

$$\sum_{j=1}^{k-1}\det\Big(J_1^j(J_1^j)^*\Big) = \sum_{j=1}^{k-1}\Big(\det(J_1)\Big)^j\Big(\det(J_1^*)\Big)^j.$$

J_1 is an upper triangular matrix with diagonal entries λ_1. Suppose that its dimension is r. Then, $\sum_{j=1}^{k-1}\det\Big(J_1^j(J_1^j)^*\Big) = \sum_{j=1}^{k-1}(|\lambda_1|^r)^{2j}$. Since $|\lambda_1| = \rho(A) \geq 1$, $(|\lambda_1|^2)^{rj} \geq 1$ and $\sum_{j=1}^{k-1}(|\lambda_1|^r)^{2j} \geq k-1$. Thus, $\sum_{j=1}^{k-1}\det\big(J_1^j(J_1^j)^*\big) \geq k-1$. By (28), $\det(\mathcal{J}_k) \geq k$.

By the property $\det(XY) = \det(X)\det(Y)$ and (27), $\det(\sum_{j=0}^{k-1} A^j(A^j)') \geq \det(\underline{\sigma}_H H\mathcal{J}_kH^*) = \underline{\sigma}_H^n \det(HH^*)\det(\mathcal{J}_k)$. Similarly, by $W_k \geq \underline{\sigma}_q \sum_{j=0}^{k-1} A^j(A^j)'$, $\det(CW_kC') \geq \underline{\sigma}_q^n \det(C\sum_{j=0}^{k-1} A^j(A^j)'C') \geq \xi\det(\mathcal{J}_k)$, where $\xi \triangleq \det(CC')\underline{\sigma}_q^n\underline{\sigma}_H^n\det(HH^*)$, and $\det(CC') > 0$ since $C \in \mathbb{R}^{m\times n}$ is assumed to have full row rank. Due to $\det(\mathcal{J}_k) \geq k$, $\det(CW_kC') \geq \xi k$. Thus, $\lim_{k\to+\infty}\frac{k^{\frac{1}{4}}}{\sqrt{\det(CW_kC')}} \leq \lim_{k\to+\infty}\frac{k^{\frac{1}{4}}}{\sqrt{\xi k}} = 0$. Thus, Lemma 2 holds. ■

Lemma 3 *The following facts hold.*

$$\lim_{k\to\infty}\mathbb{P}(S_k = s_k|\gamma_k = 0) = 1 \tag{29a}$$

$$\lim_{k\to\infty}\mathbb{P}(S_k = s_k|\gamma_k = 1) = 0 \tag{29b}$$

$$\lim_{k\to\infty}\mathbb{P}(S_k = s_k^c|\gamma_k = 0) = 0 \tag{29c}$$

$$\lim_{k\to\infty}\mathbb{P}(S_k = s_k^c|\gamma_k = 1) = 1. \tag{29d}$$

Proof: For the case $S_k = s_k$,

$$\mathbb{P}(S_k = s_k|\gamma_k = 0) = \mathbb{P}(y_k \in \mathbb{C}_k|\gamma_k = 0).$$

When $\gamma_k = 0$, $y_k = \upsilon_k$ and $p(y_k|\gamma_k = 0) = \mathcal{N}_{y_k}(0, R)$. $\mathbb{P}(y_k \in \mathbb{C}_k|\gamma_k = 0) = \int_{\mathbb{C}_k} \mathcal{N}_{y_k}(0, R)\mathsf{d}y_k$. As $k \to \infty$, $\mathbb{C}_k \to \mathbb{R}^m$ and $\int_{\mathbb{R}^m} \mathcal{N}_{y_k}(0, R)\mathsf{d}y_k = 1$. Therefore, $\lim_{k\to\infty} \mathbb{P}(S_k = s_k|\gamma_k = 0) = 1$, which proves (29a).

x_k in (1) can be recursively calculated as $x_k = A^k x_0 + \omega_k^\sharp$, where $\omega_k^\sharp = \sum_{j=0}^{k-1} A^j \omega_{k-1-j}$, $\omega_k^\sharp \sim \mathcal{N}_{\omega_k^\sharp}(0, W_k)$, and W_k is defined in Lemma 2. Consequently,

$$p(x_k) = \mathcal{N}_{x_k}\Big(A^k \overline{x}_0,\ A^k P_0 (A^k)' + W_k\Big). \tag{30}$$

When $\gamma_k = 1$, $y_k = Cx_k + \upsilon_k$. Then, $p(y_k|\gamma_k = 1) = \mathcal{N}_{y_k}(\overline{y}_k, Z_k)$, where $\overline{y}_k \triangleq CA^k\overline{x}_0$ and $Z_k \triangleq CA^k P_0 (A^k)'C' + CW_kC + R$.

$$\begin{aligned}\mathbb{P}(S_k = s_k|\gamma_k = 1) &= \mathbb{P}(y_k \in \mathbb{C}_k|\gamma_k = 1)\\ &= \int_{\mathbb{C}_k} \mathcal{N}_{y_k}(\overline{y}_k, Z_k)\mathsf{d}y_k.\end{aligned}$$

Since $\mathcal{N}_{y_k}(\overline{y}_k, Z_k) = \frac{1}{\sqrt{(2\pi)^m \det(Z_k)}} \exp(-\frac{1}{2}(\cdot)'Z_k^{-1}(\cdot))$ where $\exp(-\frac{1}{2}(\cdot)'Z_k^{-1}(\cdot)) \le 1$, $\mathbb{P}(S_k = s_k|\gamma_k = 1) \le \frac{1}{\sqrt{(2\pi)^m \det(Z_k)}} \int_{\mathbb{C}_k} \mathsf{d}y_k$. Note that $Z_k \ge CW_kC'$ and that $\int_{\mathbb{C}_k} \mathsf{d}y_k = k^{\frac{1}{4}}$ is the volume of $\mathbb{C}_k$. $\frac{1}{\sqrt{\det(Z_k)}} \int_{\mathbb{C}_k} \mathsf{d}y_k \le \frac{k^{\frac{1}{4}}}{\sqrt{\det(CW_kC')}}$. By Lemma 2 that $\lim_{k\to\infty} \frac{k^{\frac{1}{4}}}{\sqrt{\det(CW_kC')}} = 0$, $\lim_{k\to\infty} \mathbb{P}(S_k = s_k|\gamma_k = 1) = 0$. (29b) is proved.

For the case $S_k = s_k^c$, since $s_k \cap s_k^c = \emptyset$, $\mathbb{P}(S_k = s_k^c|\gamma_k = 0) = 1 - \mathbb{P}(S_k = s_k|\gamma_k = 0)$ and $\mathbb{P}(S_k = s_k^c|\gamma_k = 1) = 1 - \mathbb{P}(S_k = s_k|\gamma_k = 1)$. From (29a) and (29b), it follows that (29c) and (29d) hold. ∎

Lemma 4 *The mean squared estimation error of $\widetilde{\gamma}_k$ converges to zero, and the estimation accuracy $\mathbb{P}(\gamma_k = \widetilde{\gamma}_k|S_k)$ converges to* 1*. Specifically,*

$$\lim_{k\to\infty} \mathbb{E}[(\gamma_k - \widetilde{\gamma}_k)^2|S_k] = 0 \tag{31}$$

$$\lim_{k\to\infty} \mathbb{P}(\gamma_k = \widetilde{\gamma}_k|S_k) = 1. \tag{32}$$

Proof: By the law of total probability,

$$\begin{aligned}&\mathbb{E}[(\gamma_k - \widetilde{\gamma}_k)^2|S_k]\\ &= (0 - \widetilde{\gamma}_k)^2\mathbb{P}(\gamma_k = 0|S_k) + (1 - \widetilde{\gamma}_k)^2\mathbb{P}(\gamma_k = 1|S_k)\\ &= \widetilde{\gamma}_k\mathbb{P}(\gamma_k = 0|S_k) + (1 - \widetilde{\gamma}_k)\mathbb{P}(\gamma_k = 1|S_k),\end{aligned} \tag{33}$$

where (33) is obtained by noting that $\widetilde{\gamma}_k$ takes the value 0 or 1.

When $S_k = s_k$, $\widetilde{\gamma}_k = 0$ in (16), and then (33) $= \mathbb{P}(\gamma_k = 1|S_k = s_k)$. By using Bayesian formula,

$$\begin{aligned}&\mathbb{P}(\gamma_k = 1|S_k = s_k)\\ &= \frac{\gamma\mathbb{P}(S_k = s_k|\gamma_k = 1)}{\overline{\gamma}\mathbb{P}(S_k = s_k|\gamma_k = 0) + \gamma\mathbb{P}(S_k = s_k|\gamma_k = 1)}.\end{aligned}$$

By (29a) and (29b), $\lim_{k\to\infty} \mathbb{E}[(\gamma_k - \widetilde{\gamma}_k)^2|S_k = s_k] = 0$. Similarly, when $S_k = s_k^c$, $\widetilde{\gamma}_k = 1$. Then,

$$\begin{aligned}(33) &= \mathbb{P}(\gamma_k = 0|S_k = s_k^c)\\ &= \frac{\overline{\gamma}\mathbb{P}(S_k = s_k^c|\gamma_k = 0)}{\overline{\gamma}\mathbb{P}(S_k = s_k^c|\gamma_k = 0) + \gamma\mathbb{P}(S_k = s_k^c|\gamma_k = 1)}.\end{aligned}$$

By (29c) and (29d), $\lim_{k\to\infty} \mathbb{E}[(\gamma_k - \widetilde{\gamma}_k)^2|S_k = s_k^c] = 0$.

As proved as above, no matter $S_k = s_k$ or $S_k = s_k^c$, $\lim_{k\to\infty} \mathbb{E}[(\gamma_k - \widetilde{\gamma}_k)^2|S_k] = 0$, which proves that (31) holds.

Let $d_k = (\gamma_k - \widetilde{\gamma}_k)^2$. When $\gamma_k = \widetilde{\gamma}_k$, $d_k = 0$. When $\gamma_k \ne \widetilde{\gamma}_k$, $d_k = 1$. $\mathbb{E}[(\gamma_k - \widetilde{\gamma}_k)^2|S_k] = 0 \cdot \mathbb{P}(\gamma_k = \widetilde{\gamma}_k|S_k) + 1 \cdot \mathbb{P}(\gamma_k \ne \widetilde{\gamma}_k|S_k)$. From (31), it follows that $\lim_{k\to\infty} \mathbb{P}(\gamma_k \ne \widetilde{\gamma}_k|S_k) = 0$, which means $\lim_{k\to\infty} \mathbb{P}(\gamma_k = \widetilde{\gamma}_k|S_k) = 1$. The proof of (32) is completed. ∎

Proof of Theorem 2(i): By $\mathbb{E}[(\gamma_k - \widehat{\gamma}_k)^2|Y_k] \le \mathbb{E}[(\gamma_k - \widetilde{\gamma}_k)^2|S_k]$ in (19) and $\lim_{k\to\infty} \mathbb{E}[(\gamma_k - \widetilde{\gamma}_k)^2|S_k] = 0$ in (31), it is clear that Theorem 2(i) holds. ∎

B. Proof of Theorem 2(ii): The case $\rho(A) < 1$

Some preliminaries and lemmas are given as follows.

Steady pdf of x_k: By (1), it is easy to obtain $p(x_k) = \mathcal{N}_{x_k}(A^k\overline{x}_0, P_k)$ with $P_k = AP_{k-1}A' + Q$. When $\rho(A) < 1$, it is well known that $A^k\overline{x}_0$ converges to zero and P_k converges to P_x, where P_x is the unique solution of $P_x = AP_xA' + Q$. Since they converge exponentially fast, we assume that $p(x_k)$ has reached the steady pdf, that is,

$$p(x_k) = \mathcal{N}_{x_k}(0, P_x). \tag{34}$$

Probability cumulative function $\Phi(\cdot)$: For $\mathcal{N}_x(0, P_x)$ and $\mathcal{N}_y(0, R)$ with $x \in \mathbb{R}^n$ and $y \in \mathbb{R}^m$, clearly, $f(x, y) \triangleq \mathcal{N}_y(0, R)\mathcal{N}_x(0, P_x)$ with $(x, y) \in \mathbb{R}^n \times \mathbb{R}^m$ still is a Gaussian pdf. Define the probability cumulative function of f by

$$\Phi(\Omega) \triangleq \int_{\Omega \subseteq \mathbb{R}^n \times \mathbb{R}^m} \mathcal{N}_y(0, R)\mathcal{N}_x(0, P_x)\mathsf{d}y\mathsf{d}x. \tag{35}$$

Lemmas 5 and 6 in the following establish the relationship between $\mathbb{E}[\mathbb{C}\mathrm{ov}(\gamma_k|Y_k)]$ and $\Phi(\Omega)$, which plays an important role in proving the existence of Γ_γ in (25) in Theorem 2(ii).

Lemma 5 *For a given $l > 0$, let $\kappa \triangleq (\underline{\sigma}(R))^{-1}$ and $\Omega_l \triangleq \big\{(x, y) \in \mathbb{R}^n \times \mathbb{R}^m \;\big|\; \|y - Cx\| \le l\big\}$. Then,*

$$\int_{\mathbb{R}^n \times \mathbb{R}^m} \frac{\overline{\gamma}\gamma\mathcal{N}_y(0, R)\mathcal{N}_y(Cx, R)}{\overline{\gamma}\mathcal{N}_y(0, R) + \gamma\mathcal{N}_y(Cx, R)}\mathcal{N}_x(0, P_x)\mathsf{d}y\mathsf{d}x \tag{36}$$

$$\ge \frac{\overline{\gamma}\gamma}{\overline{\gamma}\exp(0.5\kappa l^2) + \gamma}\Phi(\Omega_l). \tag{37}$$

Proof: In (36),

$$\begin{aligned}&\frac{\mathcal{N}_y(Cx, R)}{\overline{\gamma}\mathcal{N}_y(0, R) + \gamma\mathcal{N}_y(Cx, R)}\\ &= \frac{1}{\overline{\gamma}\exp(-\frac{y'R^{-1}y}{2})\exp(\frac{(y-Cx)'R^{-1}(y-Cx)}{2}) + \gamma}.\end{aligned} \tag{38}$$

Note that $\exp(\frac{(y-Cx)'R^{-1}(y-Cx)}{2}) \le \exp(0.5\kappa l^2)$ for $(x, y) \in \Omega_l$ and that $\exp(-\frac{y'R^{-1}y}{2}) \le 1$, we have (38) $\ge \frac{1}{\overline{\gamma}\exp(0.5\kappa l^2)+\gamma}$ for $(x, y) \in \Omega_l$. Thus, (36) $\ge \frac{\overline{\gamma}\gamma}{\overline{\gamma}\exp(0.5\kappa l^2)+\gamma} \int_{\Omega_l} \mathcal{N}_y(0, R)\mathcal{N}_x(0, P_x)\mathsf{d}y\mathsf{d}x$. The proof is completed. ∎

Lemma 6 *For an arbitrary given $l > 0$,*

$$\mathbb{E}_{x_k}\mathbb{E}_{y_k}[\mathbb{C}\mathrm{ov}(\gamma_k|y_k, x_k)] \ge \frac{\gamma\overline{\gamma}}{\overline{\gamma}\exp(0.5\kappa l^2) + \gamma}\Phi(\Omega_l) > 0.$$

Proof: It is known that for a Bernoulli random variable, say t, $\mathbb{Cov}(t) = \mathbb{P}(t=0)\mathbb{P}(t=1)$. Thus, $\mathbb{Cov}(\gamma_k|y_k,x_k) = \mathbb{P}(\gamma_k=0|y_k,x_k)\mathbb{P}(\gamma_k=1|y_k,x_k)$, and the two probability quantities on the right-hand side are calculated as follows.

Recall that $y_k = \gamma_k C x_k + \upsilon_k$. It is easy to obtain that

$$\begin{aligned} p(y_k|\gamma_k=0,x_k) &= \mathcal{N}_{y_k}(0,R) \\ p(y_k|\gamma_k=1,x_k) &= \mathcal{N}_{y_k}(Cx_k,R). \end{aligned}$$

Then, using the law of total probability and noting that γ_k is independent of x_k, $p(y_k|x_k) = p(y_k|\gamma_k=0,x_k)\mathbb{P}(\gamma_k=0|x_k) + p(y_k|\gamma_k=1,x_k)\mathbb{P}(\gamma_k=1|x_k) = \overline{\gamma}\mathcal{N}_{y_k}(0,R) + \gamma\mathcal{N}_{y_k}(Cx_k,R)$.

By Bayesian formula,

$$\begin{aligned} \mathbb{P}(\gamma_k=0|y_k,x_k) &= \frac{p(y_k|\gamma_k=0,x_k)\mathbb{P}(\gamma_k=0|x_k)}{p(y_k|x_k)} \\ &= \frac{\overline{\gamma}\mathcal{N}_{y_k}(0,R)}{\overline{\gamma}\mathcal{N}_{y_k}(0,R)+\gamma\mathcal{N}_{y_k}(Cx_k,R)}, \\ \mathbb{P}(\gamma_k=1|y_k,x_k) &= 1-\mathbb{P}(\gamma_k=0|y_k,x_k). \end{aligned}$$

Then, by $\mathbb{Cov}(\gamma_k|y_k,x_k) = \mathbb{P}(\gamma_k=0|y_k,x_k)\mathbb{P}(\gamma_k=1|y_k,x_k)$, $p(y_k|x_k)$ obtained above, $p(x_k)$ in (34), and using Lemma 5,

$$\begin{aligned} &\mathbb{E}_{x_k}\mathbb{E}_{y_k}[\mathbb{Cov}(\gamma_k|y_k,x_k)] \\ &= \int_{\mathbb{R}^n}\int_{\mathbb{R}^m} \mathbb{Cov}(\gamma_k|y_k,x_k)p(y_k|x_k)p(x_k)\mathsf{d}y_k\mathsf{d}x_k \\ &= \int_{\mathbb{R}^n\times\mathbb{R}^m} \frac{\overline{\gamma}\gamma\mathcal{N}_{y_k}(0,R)\mathcal{N}_{y_k}(Cx_k,R)\mathcal{N}_{x_k}(0,P_x)}{\overline{\gamma}\mathcal{N}_{y_k}(0,R)+\gamma\mathcal{N}_{y_k}(Cx_k,R)}\mathsf{d}y_k\mathsf{d}x_k \\ &\geq \frac{\gamma\overline{\gamma}}{\overline{\gamma}\exp(0.5\kappa l^2)+\gamma}\Phi(\Omega_l) > 0. \end{aligned}$$

■

Proof of Theorem 2(ii): By using (22), $\mathbb{Cov}(\gamma_k|Y_k) \geq \mathbb{E}_{x_k}[\mathbb{Cov}(\gamma_k|y_k,Y_{k-1},x_k)]$. Since γ_k in (2) is independent of Y_{k-1}, $\mathbb{E}_{x_k}[\mathbb{Cov}(\gamma_k|y_k,Y_{k-1},x_k)] = \mathbb{E}_{x_k}[\mathbb{Cov}(\gamma_k|y_k,x_k)]$.

For an arbitrary given l and by Lemma 6,

$$\begin{aligned} \mathbb{E}_{y_k}[\mathbb{Cov}(\gamma_k|Y_k)] &\geq \mathbb{E}_{y_k}\left[\mathbb{E}_{x_k}[\mathbb{Cov}(\gamma_k|y_k,x_k)]\right] \\ &= \mathbb{E}_{x_k}\left[\mathbb{E}_{y_k}[\mathbb{Cov}(\gamma_k|y_k,x_k)]\right] \\ &\geq \frac{\gamma\overline{\gamma}}{\overline{\gamma}\exp(0.5\kappa l^2)+\gamma}\Phi(\Omega_l) > 0. \end{aligned}$$

Due to $\mathbb{E}_{Y_k}[\cdot] = \mathbb{E}_{Y_{k-1}}[\mathbb{E}_{y_k}[\cdot]]$, $\mathbb{E}_{Y_k}[\mathbb{Cov}(\gamma_k|Y_k)] \geq \frac{\gamma\overline{\gamma}}{\overline{\gamma}\exp(0.5\kappa l^2)+\gamma}\Phi(\Omega_l) > 0$, which means $\inf\mathbb{E}_{Y_k}[\mathbb{Cov}(\gamma_k|Y_k)] > 0$. Let $\Gamma_\gamma \triangleq \inf\mathbb{E}_{Y_k}[\mathbb{Cov}(\gamma_k|Y_k)]$. According to the property of infimum, there is no $k\in\mathbb{N}$ such that $\mathbb{E}_{Y_k}[\mathbb{Cov}(\gamma_k|Y_k)] < \Gamma_\gamma$. The proof of Theorem 2(ii) is completed. ■

Remark 2 (Explanation on Theorem 2) *For an unstable UPL system, it follows from $p(x_k)$ in (30) that x_k leaves the original point further and further away as $k\to\infty$, and so does $y_k = \gamma_k C x_k + \upsilon_k$ when $\gamma_k = 1$. When $\gamma_k = 0$, $y_k = \upsilon_k$ stays nearby the original point. Then, $\gamma_k = 0$ or 1 can be estimated by observing the distance between y_k and the original point. y_k goes further away from the original point as $k\to\infty$, and therefore the estimation becomes more and more accurate.*

For a stable UPL system, it follows from $p(x_k)$ in (34) that x_k stays nearby the original point. If R is significantly larger than P_x in (34), the noise υ_k has a dominant position in $y_k = \gamma_k C x_k + \upsilon_k$, which makes it difficult to identify from y_k the existence of $\gamma_k C x_k$. That is why the estimation of γ_k is limited and cannot be better than Γ_γ.

V. IMM Estimator Stability for UPL Systems

This section aims at solving Problem 1. Based on the results established in Section III-B and C on the IMM and auxiliary IMM estimators, this section studies the stability of the IMM estimator for UPL systems. The results are given in the following Theorem 3.

For a given probability measure space $(\Omega,\mathscr{F},\mathbb{P})$, where Ω, $\mathscr{F}$, and $\mathbb{P}$ denote the sample space, σ-field of Ω, and the probability measure, respectively. For a property $\mathcal{E}$, denote N as the set on which $\mathcal{E}$ holds. By convention, $\mathbb{P}(\mathcal{E})$ means the probability that $\mathcal{E}$ holds, and thus $\mathbb{P}(\mathcal{E}) = \mathbb{P}(N)$.

Definition 2 (Almost everywhere) *[39, p.185]*
A property $\mathcal{E}$ is said to hold almost everywhere (a. e.), if there exists a set $N\in\mathscr{F}$ such that $\mathcal{E}$ holds on N and $\mathbb{P}(N^c) = 0$, that is, $\mathbb{P}(\mathcal{E}) = \mathbb{P}(N) = 1-\mathbb{P}(N^c) = 1$.

According to the definition of estimator stability in [40, 41], the stability of the IMM estimator is given as follows.

Definition 3 (Estimator stability) *[41]*
The IMM estimator is said to be stable, if

$$\limsup_{k\to\infty}\mathbb{E}[P_k] < \infty, \forall P_0 > 0.$$

It is said to be stable a. e., if

$$\mathbb{P}\Big(\limsup_{k\to\infty}\mathbb{E}[P_k] < \infty\Big) = 1, \forall P_0 > 0.$$

Theorem 3 (Stability of IMM estimator)
For a UPL system,

(i) **(Necessary and sufficient condition)** *when $\rho(A) \geq 1$, there exists a* **packet-arrival-rate threshold** *λ_γ such that the IMM estimator is stable a. e., that is,*

$$\mathbb{P}\Big(\limsup_{k\to\infty}\mathbb{E}[P_k] < \infty\Big) = 1, \forall P_0 > 0,$$

if and only if $\gamma > \lambda_\gamma$.

(ii) When $\rho(A) < 1$, the IMM estimator is stable, that is,

$$\limsup_{k\to\infty}\mathbb{E}[P_k] < +\infty, \forall P_0 > 0,$$

no matter what value γ is.

Theorem 3(i) and (ii) are proved in the following Section V-A and V-B, respectively.

A. Proof of Theorem 3(i): The case $\rho(A) \geq 1$

Define the following functions:

$$\Psi(X,K) \triangleq (I-KC)X(I-KC)' + KRK'.$$

$$K_X \triangleq XC'(CXC'+R)^{-1}.$$
$$\mathcal{L}(X,\gamma) \triangleq (1-\gamma)AXA' + \gamma A\Psi(X,K_X)A' + Q.$$

The above defined $\Psi(\cdot)$ and $\mathcal{L}(\cdot)$ are commonly-used functions for analysing the covariance P_k. Thus, their properties derived in the following two lemmas will be used for studying estimator stability.

Lemma 7 *The following facts hold.*

$$\Psi(X,K_X) \le \Psi(X,K), \forall K. \tag{39}$$
$$\Psi(X_1,K) \le \Psi(X_2,K), \text{ if } X_1 \le X_2. \tag{40}$$
$$\Psi(X,K_X) = X - K_XCX. \tag{41}$$

Proof: (39) and (40) can be obtained by letting $A=1$ and $Q=0$ in Lemma 1(a)(b)(h) of [28]. (41) is an existing result in [42, p. 73 Eq. (2.58)]. ■

Lemma 8 *For random matrix $X>0$ and $Y>0$, if $\mathbb{E}[X] \ge \mathbb{E}[Y]$, then $\mathbb{E}[\Psi(X,K_X)] \ge \mathbb{E}[\Psi(Y,K_Y)]$.*

Proof: By (40), $\Upsilon \triangleq \mathbb{E}[\Psi(X,K_X) - \Psi(Y,K_Y)] \ge \mathbb{E}[\Psi(X,K_X) - \Psi(Y,K_X)] = \mathbb{E}[(I-K_XC)(X-Y)(I-K_XC)']$.

If $\underline{\sigma}(I-K_XC)=0$, then $\Upsilon \ge 0$. If $\underline{\sigma}(I-K_XC) \ne 0$, then there exists a real number κ such that $(\underline{\sigma}(I-K_XC))^2 > \kappa > 0$. Hence, $\Upsilon \ge \kappa(\mathbb{E}[X]-\mathbb{E}[Y]) \ge 0$. ■

The following properties will be used in analysing probability-related convergence (e.g. (44)) in Lemmas 10 and 11.

Lemma 9 *For events A, B and C_k,*

$$\mathbb{P}(A) \ge \mathbb{P}(A|B)\mathbb{P}(B) \tag{42}$$
$$\lim_{k\to\infty} \mathbb{P}(C_k) = 1, \text{ if } \lim_{k\to\infty} \mathbb{P}(C_k) \ge 1. \tag{43}$$

Proof: For events A and B, $\mathbb{P}(A) = \mathbb{P}(A|B)\mathbb{P}(B) + \mathbb{P}(A|B^c)\mathbb{P}(B^c)$. Then, (42) is obtained due to $\mathbb{P}(A|B^c)\mathbb{P}(B^c) \ge 0$. (43) is the result of $1 \ge \lim_{k\to\infty}\mathbb{P}(C_k) \ge 1$. ■

Lemmas 10 and 11 in the following study the convergence of the probability of some estimator-stability-related conditions (such the following $\mathcal{M}_k$). They are essential for proving Theorem 3(i).

Lemma 10 *Denote the following condition by $\mathcal{M}_k$:*

$$\mathcal{M}_k \triangleq \Big\{\mathbb{E}[\overline{M}_k] \le \mathbb{E}[\mathcal{L}(\overline{M}_{k-1},\gamma_{k-1})]\Big\},$$

where $\overline{M}_{k-1}$ is computed in Theorem 1. Then,

$$\lim_{k\to\infty}\mathbb{P}(\mathcal{M}_k) = 1. \tag{44}$$

Proof: By viewing the event $\{\gamma_k = \widetilde{\gamma}_k\}$ as B in (42),

$$\mathbb{P}(\widetilde{x}_k = \widetilde{x}_k^\sharp|S_k) \ge \mathbb{P}(\widetilde{x}_k = \widetilde{x}_k^\sharp|\gamma_k=\widetilde{\gamma}_k, S_k)\mathbb{P}(\gamma_k=\widetilde{\gamma}_k|S_k).$$

Under the condition $\gamma_k = \widetilde{\gamma}_k$, $\widetilde{x}_k$ in (17) equals $\widetilde{x}_k^\sharp$ in (18), and thus $\mathbb{P}(\widetilde{x}_k = \widetilde{x}_k^\sharp|\gamma_k=\widetilde{\gamma}_k,S_k) = 1$. By (32) and (43), we have $\lim_{k\to\infty}\mathbb{P}(\widetilde{x}_k = \widetilde{x}_k^\sharp|S_k) = 1$.

Note that S_k is determined by y_k. That is to say, knowing Y_k is equivalent to knowing $\{S_k, Y_k\}$. Thus, in (21), $P_k^\sharp = \mathbb{E}[(x_k-\widetilde{x}_k)_I^2|S_k,Y_k,\gamma_k]$. Since $\widetilde{\gamma}_k$ is determined by S_k, knowing S_k is equivalent to knowing $\{S_k,\widetilde{\gamma}_k\}$, which means $\widetilde{P}_k = \mathbb{E}[(x_k-\widetilde{x}_k)_I^2|S_k,Y_k,\widetilde{\gamma}_k]$ in (20). By $\lim_{k\to\infty}\mathbb{P}(\gamma_k = \widetilde{\gamma}_k|S_k) = 1$ in (32) and $\lim_{k\to\infty}\mathbb{P}(\widetilde{x}_k = \widetilde{x}_k^\sharp|S_k) = 1$ obtained above, it is known that under condition S_k, $\widetilde{x}_k$ and $\widetilde{\gamma}_k$ converges with probability 1 to $\widetilde{x}_k^\sharp$ and γ_k, respectively. Consequently, under condition S_k, $\mathbb{E}[(x_k-\widetilde{x}_k)_I^2|S_k,\widetilde{\gamma}_k]$ converges with probability 1 to $\mathbb{E}[(x_k-\widetilde{x}_k^\sharp)_I^2|S_k,\gamma_k]$, It is known that for random variables X and Y, $\mathbb{E}_Y[\mathbb{E}_X[X|Y]] = \mathbb{E}_X[X]$. By using this property, we have $\mathbb{E}_{Y_k}[P_k] = \mathbb{E}_{Y_k}[\mathbb{E}[(x_k-\widetilde{x}_k)_I^2|S_k,\widetilde{\gamma}_k,Y_k]] = \mathbb{E}[(x_k-\widetilde{x}_k)_I^2|S_k,\widetilde{\gamma}_k]$ converges with probability 1 to $\mathbb{E}[(x_k-\widetilde{x}_k^\sharp)_I^2|S_k,\gamma_k] = \mathbb{E}_{Y_k}[\mathbb{E}[(x_k-\widetilde{x}_k^\sharp)_I^2|S_k,\gamma_k,Y_k]] = \mathbb{E}_{Y_k}[P_k^\sharp]$, that is, $\lim_{k\to\infty}\mathbb{P}(\mathbb{E}[\widetilde{P}_k] = \mathbb{E}[P_k^\sharp]) = 1$.

By (20),

$$\lim_{k\to\infty}\mathbb{P}(\mathbb{E}[P_k] \le \mathbb{E}[P_k^\sharp]) = 1. \tag{45}$$

By (4) and (21),

$$\lim_{k\to\infty}\mathbb{P}\Big(\mathbb{E}[\overline{M}_{k+1}] \le A\big(\mathbb{E}[(1-\gamma_k)\overline{M}_k + \gamma_k M_k]\big)A' + Q\Big) = 1,$$

that is, $\lim_{k\to\infty}\mathbb{P}(\mathcal{M}_{k+1}) = 1$, which proves (44). ■

Lemma 11 *Let $Z_{k+1} = \mathcal{L}(Z_k,\gamma_k)$ with $Z_1 = \overline{M}_1$. Then,*

(i) $\lim_{k\to\infty}\mathbb{P}(\mathbb{E}[Z_k] \ge \mathbb{E}[\overline{M}_k]) = 1$.
(ii) $\lim_{k\to\infty}\mathbb{P}(\mathbb{E}[Z_k] \le \mathbb{E}[\overline{M}_k]) = 1$.
(iii) $\lim_{k\to\infty}\mathbb{P}(\mathbb{E}[Z_k] = \mathbb{E}[\overline{M}_k]) = 1$.

Proof of (i): To prove Part (i), we first prove the following equality by the mathematical induction method:

$$\mathbb{P}(\mathbb{E}[Z_k] \ge \mathbb{E}[\overline{M}_k]|\mathcal{M}_k) = 1. \tag{46}$$

Clearly, $\mathbb{E}[Z_k] \ge \mathbb{E}[\overline{M}_k]$ holds when $k=1$. Suppose that it holds for $1,\ldots,k-1$. Then, for the case k, under the condition $\mathcal{M}_k$, $\mathbb{E}[Z_k - \overline{M}_k] \ge \mathbb{E}[(1-\gamma_{k-1})A[Z_{k-1} - \overline{M}_{k-1}]A' + \gamma_{k-1}A(\Psi(Z_{k-1},K_{Z_{k-1}}) - \Psi(\overline{M}_{k-1},K_{\overline{M}_{k-1}}))A']$.

By the hypothesis $\mathbb{E}[Z_{k-1}] \ge \mathbb{E}[\overline{M}_{k-1}]$ and Lemma 8, $\mathbb{E}[\Psi(Z_{k-1},K_{Z_{k-1}})] \ge \mathbb{E}[\Psi(\overline{M}_{k-1},K_{\overline{M}_{k-1}})]$. Thus, $\mathbb{E}[Z_k] \ge \mathbb{E}[\overline{M}_k]$. This result shows that the probability of the event $\{\mathbb{E}[Z_k] \ge \mathbb{E}[\overline{M}_k]\}$ under the condition $\mathcal{M}_k$ is 1, that is, $\mathbb{P}(\mathbb{E}[Z_k] \ge \mathbb{E}[\overline{M}_k]|\mathcal{M}_k) = 1$. Thus, (46) holds.

By (42), $\mathbb{P}(\mathbb{E}[Z_k] \ge \mathbb{E}[\overline{M}_k]) \ge \mathbb{P}(\mathbb{E}[Z_k] \ge \mathbb{E}[\overline{M}_k]|\mathcal{M}_k)\mathbb{P}(\mathcal{M}_k)$, where $\mathbb{P}(\mathbb{E}[Z_k] \ge \mathbb{E}[\overline{M}_k]|\mathcal{M}_k) = 1$ is proved in Lemma 11(ii). Thus, $\mathbb{P}(\mathbb{E}[Z_k] \ge \mathbb{E}[\overline{M}_k]) \ge \mathbb{P}(\mathcal{M}_k)$. By (43) and Lemma 11, $\lim_{k\to\infty}\mathbb{P}(\mathbb{E}[Z_k] \ge \mathbb{E}[\overline{M}_k]) = 1$, which proves Part (i).

Proof of (ii): We prove Part (ii) by the mathematical induction method. It holds for $k=1$, as $Z_1 = \overline{M}_1$. Suppose that it holds for $1,\ldots,k$, and consider the case $k+1$ as follows.

From (8) and (9), and due to $(\widehat{x}_k - \overline{m}_k)_I^2 \ge 0$ and $(\widehat{x}_k - m_k)_I^2 \ge 0$, $P_k \ge (1-\widehat{\gamma}_k)\overline{M}_k + \widehat{\gamma}_k\Psi(\overline{M}_k,K_{\overline{M}_k})$. By $p(y_k|Y_{k-1}) = \overline{\gamma}\phi_k + \gamma\psi_k$ in (15),

$$\mathbb{E}_{y_k}[\widehat{\gamma}_k] = \textstyle\int_{\mathbb{R}^m}\frac{\gamma\psi_k}{\overline{\gamma}\phi_k+\gamma\psi_k}p(y_k|Y_{k-1})\mathrm{d}y_k = \int_{\mathbb{R}^m}\gamma\psi_k\mathrm{d}y_k = \gamma.$$

Since $\overline{M}_k$ does not contain y_k, $\mathbb{E}_{y_k}[P_k] \ge \mathbb{E}_{y_k}[(1-\widehat{\gamma}_k)]\overline{M}_k + \mathbb{E}_{y_k}[\widehat{\gamma}_k]\Psi(\overline{M}_k,K_{\overline{M}_k}) = \overline{\gamma}\overline{M}_k + \gamma\Psi(\overline{M}_k,K_{\overline{M}_k})$. Let $Z_k^\sharp \triangleq (1-\gamma_k)Z_k + \gamma_k\Psi(Z_k,K_{Z_k})$. Then, $Z_{k+1} = AZ_k^\sharp A' + Q$.

Thus, based on this result, by taking mathematical expectation with respect to all the random quantities in P_k and $\overline{M}_k$, we have

$$\begin{aligned}\mathbb{E}[P_k] &\geq \overline{\gamma}\mathbb{E}[\overline{M}_k] + \gamma\mathbb{E}\big[\Psi(\overline{M}_k, K_{\overline{M}_k})\big]\\ &\geq \overline{\gamma}\mathbb{E}[Z_k] + \gamma\mathbb{E}\big[\Psi(Z_k, K_{Z_k})\big] \qquad (47)\\ &= \mathbb{E}[Z_k^{\sharp}],\end{aligned}$$

where (47) is obtained by the hypothesis $\mathbb{E}[Z_k] \leq \mathbb{E}[\overline{M}_k]$ and using Lemma 8. Consequently, $\mathbb{E}[\overline{M}_{k+1}] = \mathbb{E}[AP_kA' + Q] \geq \mathbb{E}[AZ_k^{\sharp}A' + Q] = \mathbb{E}[Z_{k+1}]$, which proves that $\mathbb{E}[Z_k] \leq \mathbb{E}[\overline{M}_k]$ holds. This result implies that $\mathbb{P}(\mathbb{E}[Z_k] \leq \mathbb{E}[\overline{M}_k]) = 1$, and clearly Part (ii) holds.

Proof of (iii): Part (iii) holds due to Parts (i) and (ii). ∎

Proof of Theorem 3(i): By the definition of $\mathcal{L}(\cdot)$ and $\Psi(\cdot)$, $Z_{k+1} = AZ_kA' - \gamma_k AZ_kC'(CZ_kC' + R)^{-1}CZ_kA' + Q$. For this modified Riccati equation, it has been reported in [28, Theorems 2] that there exists a threshold value, denoted by λ_d, such that $\limsup_{k\to\infty}\mathbb{E}[Z_k]$ is bounded if and only if $\gamma > \lambda_d$. Let $\lambda_\gamma = \lambda_d$.

Proof of sufficiency: when $\gamma > \lambda_\gamma$, $\limsup_{k\to\infty}\mathbb{E}[Z_k] \leq +\infty$. Then, due to $\lim_{k\to\infty}\mathbb{P}(\mathbb{E}[Z_k] = \mathbb{E}[\overline{M}_k]) = 1$ in Lemma 11(iii), $\mathbb{P}(\limsup_{k\to\infty}\mathbb{E}[\overline{M}_k] \leq +\infty) = 1$. By (5) and (21), $P_k^{\sharp} \leq \overline{M}_k$. From (45), it follows that $\mathbb{P}(\mathbb{E}[P_k] \leq \mathbb{E}[P_k^{\sharp}] \leq \mathbb{E}[\overline{M}_k]) = 1$. Consequently, $\mathbb{P}(\limsup_{k\to\infty}\mathbb{E}[P_k] \leq +\infty) = 1$ when $\gamma > \lambda_\gamma$.

Proof of necessity: When $\gamma \leq \lambda_\gamma$, $\limsup_{k\to\infty}\mathbb{E}[Z_k] = +\infty$. By Lemma 11(iii), $\mathbb{P}(\limsup_{k\to\infty}\mathbb{E}[\overline{M}_k] = +\infty) = 1$. By (4), $\mathbb{E}[\overline{M}_k] = A\mathbb{E}[P_{k-1}]A' + Q$. Thus, it must have $\mathbb{P}(\limsup_{k\to\infty}\mathbb{E}[P_k] = +\infty) = 1$, when $\gamma \leq \lambda_\gamma$.

B. Proof of Theorem 3(ii): The case $\rho(A) < 1$

Proof of Theorem 3(ii): In (9), $\mathcal{W}_k \triangleq (1-\widehat{\gamma}_k)\widehat{\gamma}_k(m_k - \overline{m}_k)_I^2 = \frac{\gamma\overline{\gamma}\psi_k\phi_k}{(\gamma\psi_k+\overline{\gamma}\phi_k)^2}(m_k - \overline{m}_k)_I^2$. Due to $m_k = \overline{m}_k + K_k(y_k - C\overline{m}_k)$ in Theorem 1, $(m_k - \overline{m}_k)_I^2 = K_k(y_k - C\overline{m}_k)_I^2K_k'$. By $p(y_k) = \gamma\psi_k + \overline{\gamma}\phi_k$ in (15),

$$\mathbb{E}_{y_k}[\mathcal{W}_k] = K_k\int_{\mathbb{R}^m}\frac{\gamma\overline{\gamma}\psi_k\phi_k}{\gamma\psi_k + \overline{\gamma}\phi_k}(y_k - C\overline{m}_k)_I^2\mathrm{d}y_kK_k'. \quad (48)$$

Note $\frac{\overline{\gamma}\phi_k}{\gamma\psi_k+\overline{\gamma}\phi_k} \leq 1$ and $\psi_k = \mathcal{N}_{y_k}(C\overline{m}_k, (C\overline{M}_kC' + R)^{-1})$, $\mathbb{E}_{y_k}[\mathcal{W}_k] \leq \gamma K_k\int_{\mathbb{R}^m}\psi_k(y_k - C\overline{m}_k)_I^2\mathrm{d}y_kK_k' = \gamma K_k(C\overline{M}_kC' + R)^{-1}K_k'$.

Since $K_k = \overline{M}_kC'(C\overline{M}_kC' + R)^{-1}$, $\mathbb{E}_{y_k}[\mathcal{W}_k] \leq \gamma\overline{M}_kC'(C\overline{M}_kC' + R)^{-1}C\overline{M}_k = \gamma K_{\overline{M}_k}C'\overline{M}_k$.

$\mathbb{E}_{y_k}[P_k] = \overline{\gamma}\overline{M}_k + \gamma\Psi(\overline{M}_k, K_{\overline{M}_k}) + \mathbb{E}_{y_k}[\mathcal{W}_k] = \overline{M}_k - \gamma K_{\overline{M}_k}C'\overline{M}_k + \mathbb{E}_{y_k}[\mathcal{W}_k] \leq \overline{M}_k$, and thus $\mathbb{E}_{Y_k}[P_k] \leq \mathbb{E}_{Y_{k-1}}[\overline{M}_k]$.

By $\overline{M}_{k+1} = AP_kA' + Q$ in Theorem 1, we have $\mathbb{E}_{Y_k}[\overline{M}_{k+1}] = A\mathbb{E}_{Y_k}[P_k]A' + Q \leq A\mathbb{E}_{Y_{k-1}}[\overline{M}_k]A' + Q$. Due to $\rho(A) < 1$, $\mathbb{E}_{Y_k}[\overline{M}_{k+1}]$ is convergent and bounded, thus $\mathbb{E}_{Y_k}[P_k]$ is bounded for arbitrary $\gamma \in [0, 1]$. The proof is completed. ∎

VI. Numerical Examples

Some numerical examples are presented to illustrate the main results of this paper.

Consider system (1) and (2) with the following parameters:

$$A = \begin{bmatrix}\sigma & 0\\ 0 & 0.3\end{bmatrix}, C = \begin{bmatrix}1 & 0.15\\ 0 & 1\end{bmatrix},\ Q = R = \begin{bmatrix}4 & 0\\ 0 & 4\end{bmatrix}.$$

Estimation accuracy of γ_k. It is shown in Figure 2 that for an unstable UPL system (that is, $\sigma = 1.2910$ in A), $\mathbb{Cov}(\gamma_k|Y_k)$ converges to zero, which means that the estimation γ_k becomes accurate with time, and the error $|\gamma_k - \widehat{\gamma}_k|$ converges to zero as shown in the third sub-figure. This result also indicates that an unstable UPL system will become an OPL system with time. However, as shown in Figure 3, for a stable UPL system (that is, $\sigma = 0.25$ in A), the estimation accuracy $\mathbb{Cov}(\gamma_k|Y_k)$ cannot be better than a threshold. These phenomena verify the results of Theorem 2.

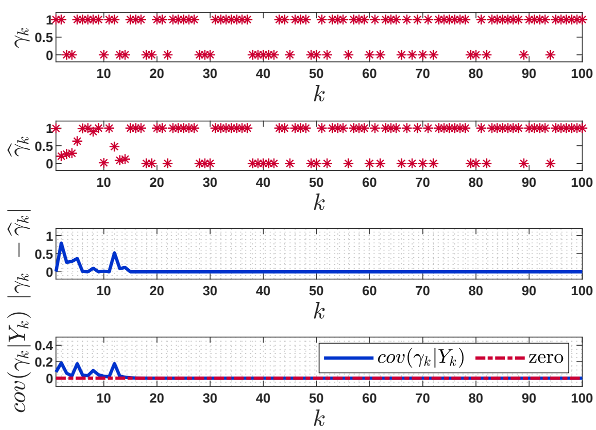


Fig. 2. Estimation of γ_k for the case $\rho(A) \geq 1$ and $\gamma = 0.7$.

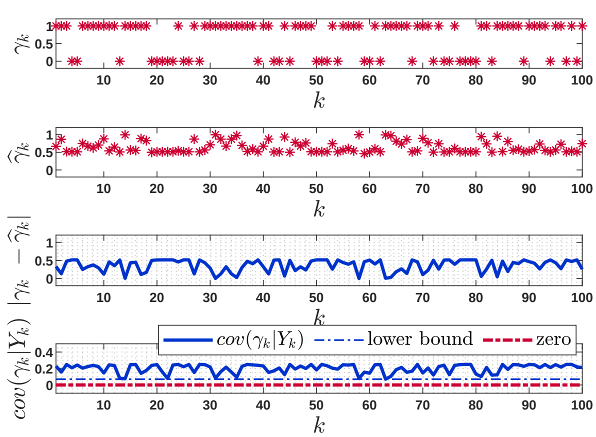


Fig. 3. Estimation of γ_k for the case $\rho(A) < 1$ and $\gamma = 0.7$.

Stability of IMM estimator. For unstable UPL systems, let $\sigma = \{1.0541, 1.1952, 1.4142, 1.8257\}$, and the thresholds $\lambda_\gamma = 1 - 1/\sigma^2$ are $\{0.1, 0.3, 0.5, 0.7\}$, respectively. Figure 4 shows that $\mathbb{E}[P_k]$ is stable for $\gamma > \lambda_\gamma$. For a stable UPL system (that is, $\sigma = 0.25$), Figure 5 shows that $\mathbb{E}[P_k]$ is bounded for different $\gamma = \{0, 0.1, 0.3, 0.5, 0.7, 0.9, 1\}$, which verifies the results of Theorem 3.

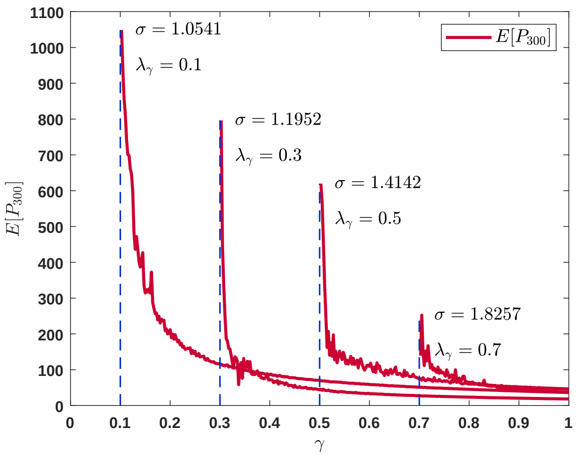


Fig. 4. Relationship between $\mathbb{E}[P_{300}]$ and γ for the case $\rho(A) > 1$.

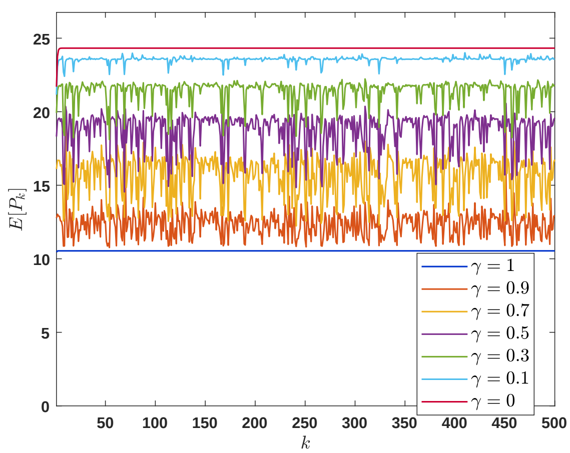


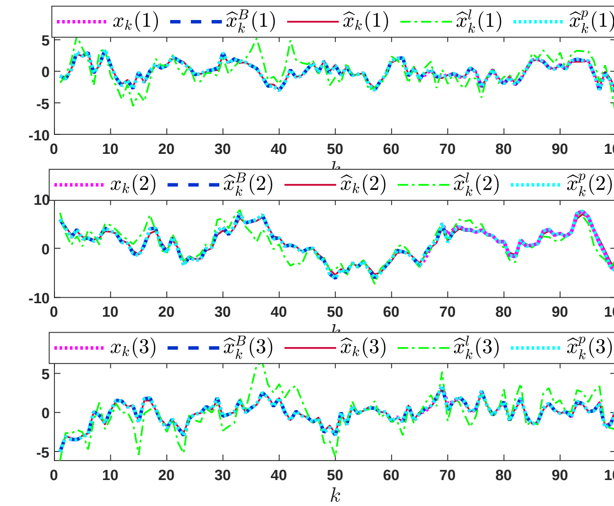
Fig. 5. $\mathbb{E}[P_k]$ for the case $\rho(A) < 1$.

Comparison with existing estimators. In this part, the performance of four estimators is compared on a real gas turbine system, called GE-F404 engine [43]. This gas turbine system takes the form of (1) and (2) with the following parameters and initial conditions:

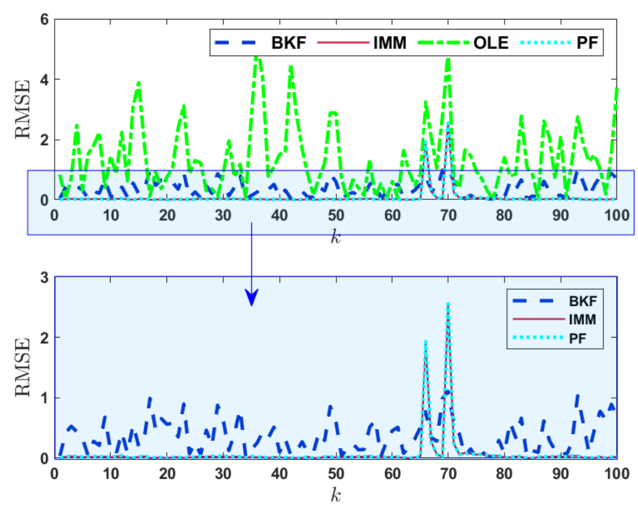
$$A = \begin{bmatrix} 0.6474 & 0 & 0.0429 \\ 0.0339 & 0.8869 & -0.0764 \\ 0.0538 & 0 & 0.5141 \end{bmatrix}, C = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix},$$

$Q = 4I_3$, $R = 0.4I_2$, $\gamma = 0.8, x_0 = [-1, 7, -5]'$, and $P_0 = 5I_3$. In the state of this gas turbine system $x_k = [x_k(1), x_k(2), x_k(3)]'$, $(x_k(1), x_k(2))$ is the horizontal position and $x_k(3)$ is the altitude.

The four estimators are the optimal linear estimator (OLE) [21], the Bayesian Kalman filter (BKF) [16], the particle filter (PF) [11], and the proposed IMM estimator in Theorem 1. In Figure 6(a), $\widehat{x}_k^B, \widehat{x}_k, \widehat{x}_k^l$, and $\widehat{x}_k^p$ denote the estimates computed by the BKF, the IMM estimator, the OLE, and PF, respectively. The root mean squared errors (RMSE) between the estimates and the real states are shown in Figure 6(b). Among these four estimators, overall, the IMM estimator and PF have the best performance, followed by the BKF, and the OLE has the worst performance. The performance of IMM estimator and PF is about the same, however, the advantage of the IMM estimator is that its stability can be theoretically determined as in Theorem 3.



(a) State state and its estimates computed by OLE, BKF, PF, and IMM estimator.



(b) RMSE of OLE, BKF, PF, and IMM estimator.

Fig. 6. State estimates and RMSE of OLE, BKF, PF, and IMM estimator.

Robustness analysis of IMM estimator. In this section, the robustness of the IMM estimator against system parameter uncertainties is studied. Specifically, we investigate how the uncertainties of system parameters (A, γ and R) affect the estimation performance, where the performance is measured by the RMSE. Nominal system parameters are given as follows:

$$A^* = \begin{bmatrix} 0.5 & 0 \\ 0 & 0.3 \end{bmatrix}, C^* = \begin{bmatrix} 1 & 0.5 \\ 0 & 1 \end{bmatrix}, Q^* = R^* = \begin{bmatrix} 10 & 0 \\ 0 & 10 \end{bmatrix},$$

$\gamma^* = 0.8$. Take the parameter A^* with an uncertainty of 5% of A^* for example. Let $A = A^* + 5\%A^*$. Denote the IMM estimator designed for the system with A (IMM with A for short) by IMM_A, and denote the IMM estimator designed for the system with A^* (IMM with A^* for short) by IMM_{A^*}. The RMSEs of IMM_A and IMM_{A^*} are shown in the first one of Figure 7, which are denoted by RMSE_A and RMSE_{A^*}, respectively.

Note that IMM_{A^*} is designed for the system with the parameter A^*. When an uncertainty of $5\%A^*$ occurs, IMM_{A^*} is used for the system with the parameter A. The performance RMSE_{A^*} is inferior to RMSE_A, since IMM_A is specifically designed for A. The performance loss is denoted by $\Delta\mathsf{RMSE} = \mathsf{RMSE}_{A^*} - \mathsf{RMSE}_A$, and the performance loss rate is denoted by $\frac{\Delta\mathsf{RMSE}}{\mathsf{RMSE}_{A^*}}$. The average performance loss rate is 3.45%, which corresponds to the column 5% and the row A^* of Table I. The RMSEs of the IMM estimator under the parameter uncertainties $10\%A^*, 15\%A^*$ and $20\%A^*$ are shown in the sequential sub-figures of Figure 7. The corresponding performance loss rates are listed in the second row of Table I, which marked with blue background colour.

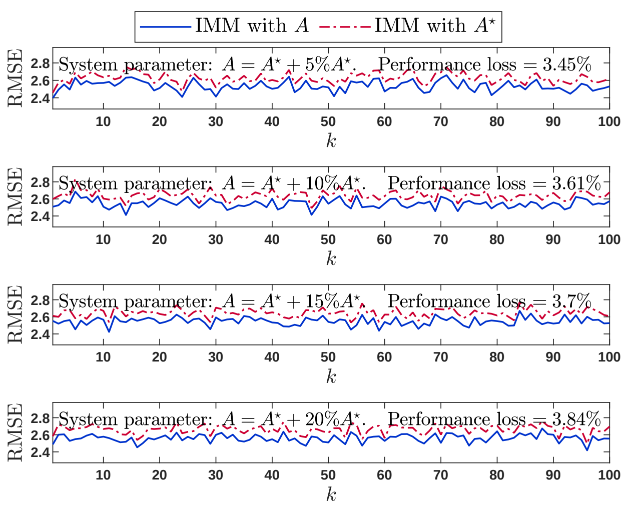


Fig. 7. Estimation performance for A^* with different uncertainties.

TABLE I
ESTIMATION PERFORMANCE LOSS WITH RESPECT TO DIFFERENT PARAMETER UNCERTAINTIES.

	-20%	-15%	-10%	-5%	5%	10%	15%	20%
A^*	3.62%	3.57%	3.41%	3.36%	3.45%	3.61%	3.37%	3.84%
R^*	4.04%	3.82%	3.70%	3.61%	3.43%	3.51%	3.63%	3.72%
γ^*	3.66%	3.57%	3.43%	3.33%	3.49%	3.69%	3.78%	3.83%

The results on the estimator robustness against parameter uncertainties are shown in Table I. The first row indicates the degree of uncertainties, and the data in the remaining rows means the performance loss rate (e.g. 3.45% in the second row) caused by the parameter located on its leftmost side (that is, A^*) under the uncertainty located at its topmost (that is, $5\%A^*$). For the performance loss rates caused by the different parameter uncertainties of γ^* and R^* are presented in the third and forth rows. From which, it can be seen that the IMM estimator has a good robustness, since the performance loss rates are about $3\% - 5\%$ under parameter uncertainties from 5% to 20%.

VII. Conclusions

For a UPL system, we have studied the estimation accuracy of the packet loss and the stability of the IMM state estimator. Compared with existing estimators for UPL systems, the advantages of the proposed IMM estimator are twofold. It is not only applicable to unstable UPL systems, but also the necessary and sufficient condition stability can be theoretically determined. We also reveal the relationship between the estimation accuracy of packet loss status and the system's stability. However, we also noticed some limitations of the proposed method in this paper, e.g., the method was designed for Bernoulli packet loss and the packet loss rate was assumed to be known. In practice, packet loss may follow a Markovian chain, and the packet loss rate is time-varying and unknown. Our subsequent research will address these issues. Another research direction is to design controllers for UPL systems and study the stability of the closed-loop systems.

References

[1] A. Rajagopal and S. Chitraganti, "State estimation and control for networked control systems in the presence of correlated packet drops," *International Journal of Systems Science*, vol. 54, no. 11, pp. 2352–2365, 2023.

[2] Y. Guo, X. Fang, and Y. Fan, "Asynchronous dynamic event-triggered control for network systems with dual triggers," *IET Control Theory & Applications*, vol. 17, no. 12, pp. 1625–1636, 2023.

[3] K. Ding and Q. Zhu, "Extended dissipative anti-disturbance control for delayed switched singular semi-Markovian jump systems with multi-disturbance via disturbance observer," *Automatica*, vol. 128, p. 109556, 2021.

[4] Y. Shen and S. Sun, "Distributed recursive filtering for multi-rate uniform sampling systems with packet losses in sensor networks," *International Journal of Systems Science*, vol. 54, no. 8, pp. 1729–1745, 2023.

[5] Y.-A. Wang, B. Shen, L. Zou, and Q.-L. Han, "A survey on recent advances in distributed filtering over sensor networks subject to communication constraints," *International Journal of Network Dynamics and Intelligence*, vol. 2, no. 2, p. 10007, 2023.

[6] X. Meng, Y. Chen, L. Ma, and H. Liu, "Protocol-based variance-constrained distributed secure filtering with measurement censoring," *International Journal of Systems Science*, vol. 53, no. 15, pp. 3322–3338, 2022.

[7] S. Feng, X. Li, S. Zhang, Z. Jian, H. Duan, and Z. Wang, "A review: state estimation based on hybrid models of Kalman filter and neural network," *Systems Science & Control Engineering*, vol. 11, no. 1, p. 2173682, 2023.

[8] C. Zhu, Z. Su, Y. Xia, L. Li, and J. Dai, "Event-triggered state estimation for networked systems with correlated noises and packet losses," *ISA Transactions*, vol. 104, pp. 36–43, 2020.

[9] J. Hu, Z. Wang, G.-P. Liu, C. Jia, and J. Williams, "Event-triggered recursive state estimation for dynamical networks under randomly switching topologies and multiple missing measurements," *Automatica*, vol. 115, p. 108908, 2020.

[10] K. Ding and Q. Zhu, "Intermittent static output feedback control for stochastic delayed-switched positive systems with only partially measurable information," *IEEE Transactions on Automatic Control*, vol. 68, no. 12, pp. 8150–8157, 2023.

[11] E. Gasmi, M. A. Sid, and O. Hachana, "Nonlinear event-based state estimation using particle filter under packet loss," *ISA Transactions*, vol. 144, pp. 176–187, 2024.

[12] N. Wang and G.-H. Yang, "Distributed optimal H_2/H_∞ filtering over unreliable wireless sensor networks," *Journal of the Franklin Institute*, vol. 361, no. 4, p. 106641, 2024.

[13] Y. Zhao, X. He, L. Ma, and H. Liu, "Unbiasedness-constrained least squares state estimation for time-varying systems with missing measurements under round-robin protocol," *International Journal of Systems Science*, vol. 53, no. 9, pp. 1925–1941, 2022.

[14] N. E. Nahi, "Optimal recursive estimation with uncertain observation," *IEEE Transactions on Information Theory*, vol. 15, no. 4, pp. 457–462, 1969.

[15] A. Jaffer and S. Gupta, "On estimation of discrete processes under multiplicative and additive noise conditions," *Information Sciences*, vol. 3, no. 3, pp. 267–276, 1971.

[16] J. Zhang, K. You, and L. Xie, "Bayesian filtering with unknown sensor measurement losses," *IEEE Transactions on Control of Network Systems*, vol. 6, no. 1, pp. 163–175, 2018.

[17] S. Wang and Q. Bao, "Single target tracking for noncooperative bistatic radar with unknown signal illumination," *Signal Processing*, vol. 183, p. 107991, 2021.

[18] L. D. Stone, R. L. Streit, T. L. Corwin, and K. L. Bell, *Bayesian Multiple Target Tracking*. Artech House, 2013.

[19] N. E. Nahi and E. Knobbe, "Optimal linear recursive estimation with uncertain system parameters," *IEEE Transactions on Automatic Control*, vol. 21, no. 2, pp. 263–266, 1976.

[20] X. Wang, "Recursive algorithms for linear LMSE estimators under uncertain observations," *IEEE Transactions on Automatic control*, vol. 29, no. 9, pp. 853–854, 1984.

[21] H. Lin, Y. Li, J. Lam, and Z.-G. Wu, "Multi-sensor optimal linear estimation with unobservable measurement losses," *IEEE Transactions on Automatic Control*, vol. 67, no. 1, pp. 481–488, 2021.

[22] Y. Li, J. Lam, and H. Lin, "On stability and performance of optimal linear filter over Gilbert-Elliott channels with unobservable packet losses," *IEEE Transactions on Control of Network Systems*, vol. 9, no. 2, pp. 1029–1039, 2021.

[23] H. Lin, J. Lam, Z. Wang, and Z. Shu, "State estimation for systems with unobservable packet losses: Approximate estimation, stability, and performance analysis," *International Journal of Robust and Nonlinear Control*, vol. 32, no. 2, pp. 545–566, 2022.

[24] H. A. Blom and Y. Bar-Shalom, "The interacting multiple model algorithm for systems with Markovian switching coefficients," *IEEE Transactions on Automatic Control*, vol. 33, no. 8, pp. 780–783, 1988.

[25] S. Zhang, S. Cheng, and Z. Jin, "Visual measurement method and application of mobile manipulator pose estimation based on PPMCC-IMM filtering," *IEEE Transactions on Instrumentation and Measurement*, vol. 72, pp. 1–12, 2023.

[26] M. Sun, Q. Duan, W. Xia, Q. Bao, and Y. Mao, "Multiple adaptive factors based interacting multiple model estimator," *IET Control Theory & Applications*, vol. 18, no. 8, pp. 1059–1069, 2024.

[27] H. Li, L. Yan, Y. Xia, and J. Zhang, "Distributed multiple model filtering for Markov jump systems with communication delays," *Journal of the Franklin Institute*, vol. 360, no. 4, pp. 3407–3435, 2023.

[28] B. Sinopoli, L. Schenato, M. Franceschetti, K. Poolla, M. I. Jordan, and S. S. Sastry, "Kalman filtering with intermittent observations," *IEEE Transactions on Automatic Control*, vol. 49, no. 9, pp. 1453–1464, 2004.

[29] M. Huang and S. Dey, "Stability of Kalman filtering with Markovian packet losses," *Automatica*, vol. 43, no. 4, pp. 598–607, 2007.

[30] K. You, M. Fu, and L. Xie, "Mean square stability for Kalman filtering with Markovian packet losses," *Automatica*, vol. 47, no. 12, pp. 2647–2657, 2011.

[31] L. Shi, M. Epstein, and R. M. Murray, "Kalman filtering over a packet-dropping network: A probabilistic perspective," *IEEE Transactions on Automatic Control*, vol. 55, no. 3, pp. 594–604, 2010.

[32] J. Wei and D. Ye, "Transmission schedule for jointly optimizing remote state estimation and wireless sensor network lifetime," *Neurocomputing*, vol. 514, pp. 374–384, 2022.

[33] H. Lin, S. Lu, P. Lu, H. Que, and P. Sun, "Centralized fusion estimation over wireless sensor-actuator networks with unobservable packet dropouts," *Journal of the Franklin Institute*, vol. 359, no. 2, pp. 1569–1584, 2022.

[34] E. Mazor, A. Averbuch, Y. Bar-Shalom, and J. Dayan, "Interacting multiple model methods in target tracking: A survey," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 34, no. 1, pp. 103–123, 1998.

[35] X. R. Li and Y. Bar-Shalom, "Performance prediction of the interacting multiple model algorithm," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 29, no. 3, pp. 755–771, 1993.

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

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

[38] D. S. Bernstein, *Matrix Mathematics*. Princeton University Press, New York, 2009.

[39] A. N. Shiryaev and A. Lyasoff, *Probability,* 2ed. Springer, New York, USA, 1996.

[40] E. R. Rohr, D. Marelli, and M. Fu, "Kalman filtering with intermittent observations: On the boundedness of the expected error covariance," *IEEE Transactions on Automatic Control*, vol. 59, no. 10, pp. 2724–2738, 2014.

[41] H. Lin, J. Lam, M. Z. Q. Chen, Z. Shu, and Z.-G. Wu, "Interacting multiple model estimator for networked control systems: Stability, convergence, and performance," *IEEE Transactions on Automatic Control*, vol. 64, no. 3, pp. 928–943, 2019.

[42] F. L. Lewis, L. Xie, and D. Popa, *Optimal and Robust Estimation: With an Introduction to Stochastic Control Theory*. CRC Press, New York, USA, 2017.

[43] X. Ge, Q.-L. Han, M. Zhong, and X.-M. Zhang, "Distributed Krein space-based attack detection over sensor networks under deception attacks," *Automatica*, vol. 109, p. 108557, 2019.