Optimal Estimator Design and Properties Analysis for Interconnected Systems With Asymmetric Information Structure

Yan Wang, Junlin Xiong, Member, IEEE, Zaiyue Yang, Senior Member, IEEE, and Rong Su, Senior Member, IEEE

Abstract—This article studies the optimal state estimation problem for interconnected systems. Each subsystem can obtain its own measurement in real time, while the measurements transmitted between the subsystems suffer from random delay. The optimal estimator is analytically designed for minimizing the conditional error covariance. The boundedness of the expected error covariance (EEC) is analyzed. In particular, a new condition that is easy to verify is established for the boundedness of EEC. Further, the properties of EEC with respect to the delay probability are studied. We found that there exists a critical probability such that the EEC is bounded if the delay probability is below the critical probability. Also, a lower and upper bound of the critical probability is derived. Finally, the proposed results are applied to a power system, and the effectiveness of the designed methods is illustrated by simulations.

Index Terms—Expected error covariance (EEC), interconnected systems (ISs), optimal state estimation, random delay, subsystems.

I. INTRODUCTION

State estimation plays an important role in numerous applications, such as target tracking [1], control [2], and signal processing [3]. With the development of the wireless network and sensor technologies, the networked state estimation have received considerable attention during the past decade. The estimation performance is significantly affected by the network environment.

The network attack is one of the factors having a significant impact on the performance of the networked state estimation. The remote state estimation (RSE) under denial-of-service (DoS) attacks was studied by a stochastic game framework in [4]. The nonstationary filtering framework was designed for uncertain fuzzy Markov switching affine systems with deception attacks in [5]. Chen et al. [6] investigated the distributed dimensionality reduction fusion estimation problem for cyber-physical systems under DoS attacks. The results of [6] are only for a single system with multisensors.

The packet drops and network delays are other two major factors affecting the networked state estimation performance. The researchers tried to understand or counteract the effects of the packet drops/delays on the estimation performance. The Kalman filtering with (partial) random packet drops was investigated in [7] and [8]. The distributed Kalman filtering with multisensors in the presence of packet drops was studied in [9]. The results of [7], [8], and [9] are for the Bernoulli packet drops model. The Kalman filtering with Markovian packet drops/delays was studied in [10], [11], and [12]. Cheng et al. [13] focused on the protocol-based filtering of fuzzy Markov affine systems with uncertain packet dropouts. In general, the estimation problems with Markovian packet drops/delays are more complex than the ones with the Bernoulli packet drops/delay. However, it is difficult to analytically discuss the estimator properties for the Markovian packet drops/delays cases. Ma and Sun [14] studied the state estimation over sensor networks with mixed uncertainties of random delay, packet dropouts, and missing measurements. The state estimation problem with multiple packet losses and with the unknown varying delayed measurements were reported in [15] and [16], respectively. Only a single system with one sensor or multisensors is considered in [7], [8], [9], [10], [11], [14], [15], and [16], and the extensions to interconnected systems (ISs) are rarely reported in the literature. Numerous physical systems are modeled as ISs that have attracted lots of research attentions in the last decade [17], [18], [19], [20]. The distributed optimal estimation problem of IS with local information is studied in [21]. However, the optimal estimator is not explicitly designed, and the obtained condition of the error covariance being bounded is not easy to verify [21].
In this article, we focus on the optimal state estimator design for ISs with random delays. A condition in terms of semidefinite programming (SDP) is established to ensure the boundedness of the expected error covariance (EEC). For the IS, the measurements transmitted between the subsystems suffer from random delays. To reduce the on-line computation and save the storage space, the delayed measurements will be discarded by each subsystem. Under the above-mentioned setup, an optimal state estimator is explicitly designed. Auxiliary equations are defined to analyze the boundedness of the EEC. In addition, the relationship between the boundedness of EEC and the delay probability is studied. The existence of a critical probability is shown, where the EEC is bounded if the delay probability is less than the critical probability. Also, a lower and upper bound of the critical probability is successfully derived. Finally, the effectiveness of the proposed theories is illustrated on a power system.

Notations: Let $\gamma_a, \gamma_{a+1}, \ldots, \gamma_b$ be the sequence $\gamma_a$. The probability measure is denoted by $P(\cdot)$. The 2-norm of matrix $A$ is denoted by $||A||$. The spectral radius of matrix $A$ is denoted by $\rho(A)$. The symbol $\delta(\cdot)$ represents the minimum eigenvalue of a matrix. The symbols $\otimes$ and $\circ$ are the operators of Kronecker product and Hadamard product, respectively. Let $0_{n \times m}$ denote $n \times m$ zero matrix, and $0_{n \times n}$ is abbreviated as $0$. Let $I_{n \times n}$ be a $n \times n$ matrix whose all elements are 1, and $I_{n}$ is abbreviated as $I$. The $n \times n$ unit matrix is denoted by $I_n$. For a function $f(\cdot)$, define $f^n(\cdot) = f(f^{n-1}(\cdot))$, where $f^1(\cdot) = f(\cdot)$.

II. Problem Statement

Consider an IS composed of two subsystems. The system dynamic is given by

$$
\begin{align}
\dot{x}^1_t &= A^1 x^1_t + A^{12} x^2_t + \omega^1_t \\
\dot{x}^2_t &= A^{21} x^1_t + A^{22} x^2_t + \omega^2_t
\end{align}
$$

(1a)

(1b)

where for subsystem $i$ ($i \in \{1, 2\}$), $x^i_t \in \mathbb{R}^{n_i}$ and $\omega^i_t \in \mathbb{R}^{n_i}$ are the state and process noise, respectively. The initial state $x_0 = (x_0^1)^T (x_0^2)^T$ is a random vector satisfying $E(x_0) = 0$ and $E(x_0 x_0^T) = \Sigma_0 > 0$. The noise $\omega_t = [(\omega^1)^T (\omega^2)^T]^T$ is an i.i.d. random process with $E(\omega_t) = 0$ and $E(\omega_t \omega_t^T) = W > 0$.

Each subsystem employs sensors to measure its own subsystem state. The measurement equations are given by

$$
\gamma_i^t = C^i x^i_t + u^i_t, \quad i \in \{1, 2\}
$$

(2)

where $u^i_t \in \mathbb{R}^{m_i}$, $(i \in \{1, 2\})$ is the measurement noise, and $u_t = [(u^1)^T (u^2)^T]^T$ is an i.i.d. random process satisfying $E(u_t) = 0$ and $E(u_t u_t^T) = V > 0$; $C^i$ is the measurement matrix with a proper dimension, for $i \in \{1, 2\}$. It is assumed that $\omega_t$ is independent of $u_t$ for any $t$, $i \geq 0$.

As Fig. 1 shows, subsystem $i$ will transmit the measurement $\gamma_i^t$ to subsystem $j$ through network for $i \neq j$. The communication network between different subsystems suffers from random one-step delay (one-step delay or no delay). Define the random binary variables $\gamma_{1,1}$ and $\gamma_{1,2}$ to describe the random delay. In particular, $\gamma_{1,1} = 0$ means that the measurement $\gamma_i^t$ transmitted from subsystem $j$ to subsystem $i$ suffers from one step delay, and $\gamma_{1,2} = 1$ indicates that there is no delay, where $i, j \in \{1, 2\}$. The subsystem $i$ will broadcast the value of $\gamma_{j,i}$ to the subsystem $j$ once subsystem $i$ knows the information transmission outcomes. Because the realization of $\gamma_{j,i}$ takes a value of either 0 or 1, it is easy to broadcast the value of $\gamma_{j,i}$. Consider that the delay indicator $\gamma_{j,i}$ ($i \in \{1, 2\}$) is an i.i.d. Bernoulli process with

$$
\begin{align}
\Pr(\gamma_{j,1} = 1) &= 1 - \lambda_1, \quad \Pr(\gamma_{j,1} = 0) = \lambda_1 \\
\Pr(\gamma_{j,2} = 1) &= 1 - \lambda_2, \quad \Pr(\gamma_{j,2} = 0) = \lambda_2
\end{align}
$$

(3a)

(3b)

where $0 \leq \lambda_1, \lambda_2 \leq 1$. Due to the random delays, the real time measurements available to subsystem 1 and subsystem 2 are $\{\gamma_{1,1}^t, \gamma_{1,1}^t \gamma_{1,2}^t\}$ and $\{\gamma_{1,2}, \gamma_{2,2}\}$, respectively, which are referred to as asymmetric information sets. Using the available real-time measurements, estimator 1 and estimator 2 are designed as the Kalman-like filtering form

$$
\begin{align}
\hat{x}^1_{t-1} &= A^{11} \hat{x}^1_{t-1} + A^{12} \hat{x}^2_{t-1} - 1 \\
\hat{x}^2_{t-1} &= A^{21} \hat{x}^1_{t-1} + A^{22} \hat{x}^2_{t-1} - 1
\end{align}
$$

(4a)

(4b)

$$
\begin{align}
\hat{x}^1_{t-1} &= A^{11} \hat{x}^1_{t-1} + A^{12} \hat{x}^2_{t-1} - 1 \\
\hat{x}^2_{t-1} &= A^{21} \hat{x}^1_{t-1} + A^{22} \hat{x}^2_{t-1} - 1
\end{align}
$$

(5a)

(5b)

$$
\begin{align}
\hat{x}^1_{t-1} &= \hat{x}^1_{t-1} + \gamma_{2,1} K^{12}_{t-1} f^{1}_{t-1} \quad (5a)
\end{align}
$$

(5b)

$$
\begin{align}
\hat{x}^2_{t-1} &= \hat{x}^2_{t-1} + \gamma_{1,2} K^{21}_{t-1} f^{2}_{t-1} \quad (5b)
\end{align}
$$

(5b)

where $\phi^i = (\gamma_{1,1}^t)^T (\gamma_{1,2}^t)^T$ is an i.i.d. random process with $E(\phi^i) = 0$ and $E(\phi^i \phi^i^T) = W > 0$. The estimator of the IS is written as follows:

$$
\begin{align}
\hat{x}^i_{t-1} &= A \hat{x}^i_{t-1} - L_t (\gamma_t - C \hat{x}^i_{t-1}) \\
\hat{x}^i_{t} &= \hat{x}^i_{t-1} + L_t (\gamma_t - C \hat{x}^i_{t-1})
\end{align}
$$

(6a)

(6b)

where $A = [A_{ijkl}]_{i,j \in \{1,2\}}$, $C = \text{diag}(C^1, C^2)$, and $L_t = \begin{bmatrix} K^{11}_{t} & K^{12}_{t} \\ K^{21}_{t} & K^{22}_{t} \end{bmatrix}$. Denote $x_t = [(x^1_t)^T (x^2_t)^T]^T$. The estimation error and prediction error are defined as $e_{t|i} = x_t - \hat{x}_{t|i}$, and $e_{t|i-1} = x_t - \hat{x}_{t-1|i-1}$, respectively. The conditional error covariances are defined

$$
\begin{align}
\hat{P}_{t|i} &= E(\hat{e}_{t|i} e_{t|i-1}^T) \\
\hat{P}_{t|i-1} &= E(\hat{e}_{t|i} e_{t|i-1}^T)
\end{align}
$$

(7a)

(7b)

$$
\begin{align}
\hat{e}_{t|i} &= A e_{t-1|i} + \omega_{t-1} \\
\hat{e}_{t|i-1} &= A e_{t-1|i-1} + \omega_{t-1} - L_t (C e_{t-1|i-1} + v_t)
\end{align}
$$

(8a)

(8b)
and
\[
\begin{align*}
P_{t|t-1} &= AP_{t|t-1}A^T + W, \\
P_{0|0} &= \Sigma_0 \quad \text{(9a)} \\
P_{t|t} &= (I - L_tC)P_{t|t-1}(I - L_tC)^T + L_t\Sigma_lL_t^T \quad \text{(9b)}
\end{align*}
\]
where \(P_{t|t-1} \) and \(P_{t|t} \) are random matrices induced by the random variables \(y_{t|t} \). For the IS (1), we make the following definition and assumption:

**Definition 1:** A system with parameters \((A, C)\) is detectable with \(\Omega\) if there exists a \(K\) such that \(\rho(A - (K \circ \Omega)C) < 1\).

**Assumption 1:** In this article, we assume that \((A, C)\) is detectable with \(I_{(n_1+n_2) \times (m_1+m_2)}\), and is undetectable with \(\text{diag}(I_{n_1 \times n_1}, I_{m_2 \times m_2})\).

Under Assumption 1, \(P_{t|t-1}\) is bounded if \(\lambda_1 = \lambda_2 = 0\) and is unbounded if \(\lambda_1 = \lambda_2 = 1\), where \(P_{t|t-1}\) is viewed as a nonrandom matrix if \(\lambda_1, \lambda_2 \in [0, 1]\).

### III. Optimal Estimator Design

In this section, the optimal gains of the estimator are analytically derived, and the estimator realization algorithm is presented.

**Theorem 1:** Consider the system (1), the estimator (6) and the conditional error covariance dynamics (9). Given \(P_{t|t-1}, y_{t|1}, y_{t|2}\), the optimal \(L_t\) minimizing \(P_{t|t}\) is given by
\[
L_t = y_{t|1}y_{2,t}L_t^{[1]} + (1 - y_{t|1})y_{t|2}L_t^{[0]} + y_{t|1}(1 - y_{t|2})L_t^{[0]}(10)
\]
where
\[
\begin{align*}
L_t^{[1]} &= P_{t|t-1}C^T(V + CP_{t|t-1}C^T)^{-1} \\
L_t^{[0]} &= \left(\begin{array}{cc}
(N^4J^2I^2U^4 + U_t^3)U_3^0 & 0 \\
0_{n_2 \times m_2} & N^3U_t^2U_6^5
\end{array}\right) \\
L_t^{[0]} &= \left(\begin{array}{cc}
N^4U_t^2U_3^0 & 0_{n_1 \times m_2} \\
0_{n_2 \times m_1} & N^4 \\
0_{n_2 \times m_2} & N_3
\end{array}\right) \\
J_t &= N_4^T \left(\begin{array}{c}
U_t^3U_3^0U_t^4U_t^6 - U_t^5U_6^5
\end{array}\right) \\
J_t &= \left(I - U_t^3U_3^0U_t^4U_t^6\right)^{-1} \\
J_t &= \left(I - U_t^3U_3^0U_t^4U_t^6\right)^{-1}
\end{align*}
\]

**Remark 1:** Compared to [7], [8], [10], and [11], the derivation of the optimal estimator gain \(L_t\) in this article is more challenging, because if \(y_{t|1}y_{t|2} \neq 1\), the corresponding gains of the estimator must satisfy a certain sparse structure constraint. While, the optimal estimator gain in [7], [8], [10], and [11] is directly obtained from standard Kalman filtering formulas. Actually, for any cases of the data reception, the optimal estimator gain in [7], [8], [10], and [11] is either a zero matrix or the standard Kalman filtering gain with different measurement matrices.

### IV. Boundedness Analysis of Expected Error Covariance

In this section, we analyze the boundedness of the EEC through auxiliary matrix formulas. For simplicity, hereinafter, we denote \(P_t = P_{t|t-1}\). It follows from (9) that:
\[
\begin{align*}
P_{t+1} &= A_{L_t}P_tA_{L_t}^T + AL_tV(Al_t)^T + W_t \quad \text{(11a)} \\
P_t &= A\Sigma_tA^T + W \quad \text{(11b)}
\end{align*}
\]
where \(A_{L_t} = A - AL_tC\). In the following, we focus on studying the properties of \(E(P_t)\). According to the definition of \(P_t\), \(E(P_t)\) depends on the distributions of the random variables \(y_{t|1}, y_{t|2}\). Define the following auxiliary matrix functions:
\[
\begin{align*}
a(X, Y) &= (A - AXC)Y(A - AXC)^T \\
b(X, Y) &= a(X, Y) + AXV_n^TA^T + W \\
c(X) &= (A - AXC)^T(A - AXC) \quad \text{(14)} \\
d(X) &= (A - AXC) \circ (A - AXC) \quad \text{(15)}
\end{align*}
\]
Recall Theorem 1. Note that \(L_t^\gamma; \gamma \in \{00, 01, 10, 11\}\) depends on \(P_t\), and thus we denote \(L_t^\gamma\) by \(L_t^\gamma(P_t)\), where we take \(P_t\) as a variable. Consider the probability distribution of

**Algorithm 1 Subestimator Realization**

1. Obtain \(y_{j|1}\) and transmit \(y_{j|1}\) to subsystem \(j\), where \(i, j \in \{1, 2\}\) and \(i \neq j\).
2. Check whether \(y_{j|1}\) is obtained in real time; broadcast the value of \(y_{j|1}\) to subsystem \(j\), and determine the value of \(y_{j|2}\).
3. Compute \(L_t^{[y_{j|1}, y_{j|2}]}\) based on Theorem 1 using \(y_{j|1}, y_{j|2}\), and \(P_{t|t-1}\).
4. Compute \(z_{t|j|1-1}\) and \(z_{t|j|2-1}\) by (4a), (5a) using \(z_{t|j|1-1}\) and \(z_{t|j|2-1}\).
5. Compute \(z_{t|j|1} = z_{t|j|1-1} + N_t^{[y_{j|1}, y_{j|2}]}N_t^{[y_{j|1}, y_{j|2}]}(y_{j|2} - Cz_{t|j|1-1})\).
6. Update \(z_{t|j|1} = z_{t|j|1} + N_t^{[y_{j|1}, y_{j|2}]}N_t^{[y_{j|1}, y_{j|2}]}N_t^{[y_{j|1}, y_{j|2}]}(y_{j|2} - Cz_{t|j|1-1})\).
7. Transmit \(L_t^\gamma\) to subsystem \(j\).
8. Compute \(P_{t+1|t}\) based on (9a)–(9b).
9. Let \(t = t + 1\) and return to first step.
\text{Lemma 2:} Consider the sequence \( x, x^2, x^3, \ldots \).

Define the matrix sequence

\[ g_{\lambda_1 \lambda_2} (X) = f \left( L^{[0][0]} [Y], L^{[1]} [Y], L^{[1]} [Y], L^{[1]} [Y] \right). \]

Some useful properties of the matrix \((17)\) are proposed below (Proposition 1 and Lemmas 1 and 2) and will be used to establish the boundedness condition of EEC.

**Proposition 1:** Consider (11) and (17). The following equations hold:

\[ E(\pi_{t+1} | \pi_t) = g_{\lambda_1 \lambda_2} (\pi_t), \quad E(\pi_{t+1}) = E(g_{\lambda_1 \lambda_2} (\pi_t)). \]

**Proof:** See Appendix B.

**Lemma 1:** Define the matrix sequence \( Y_0, Y_1, Y_2, \ldots \), generated by \( Y_{t+1} = g_{\lambda_1 \lambda_2} (Y_t), \quad Y_0 = P_0 \). One has \( E(\pi_t) \leq Y_t \).  

**Proof:** This result directly follows from (19).

**Lemma 2:** Consider the sequence \( Y_0, Y_1, Y_2, \ldots \) defined in Lemma 1. The following result holds: \( \lim_{t \to +\infty} Y_t \) is bounded and only if there exists \( X^1, \ldots, X^3 \) such that \( \rho(h(X^1, \ldots, X^3)) < 1 \), where

\[ h(X^1, \ldots, X^3) = \lambda_1 \lambda_2 d \left( N^4 X^1 N^1 + N^3 X^2 N^2 \right) + \left( 1 - \lambda_2 \right) d \left( X^3 N^1 + N^3 X^4 N^2 \right) + \left( 1 - \lambda_1 \right) \lambda_2 d \left( N^4 X^1 N^1 + X^3 N^2 \right) + \left( 1 - \lambda_1 \right) \left( 1 - \lambda_2 \right) d (X^7). \]

**Proof:** See Appendix C.

In the following, we derive a sufficient condition of \( \rho(h(X^1, \ldots, X^3)) < 1 \) with the form of SDP.

**Theorem 2:** Consider (11) with (3). Given \( \lambda_1 \) and \( \lambda_2 \). If there exists \( r_1, r_2, r_3 \), and \( r_4 \) satisfying \( r_1 \lambda_1 \lambda_2 + r_2 \lambda_1 (1 - \lambda_2) + r_3 (1 - \lambda_1) \lambda_2 + r_4 (1 - \lambda_1) (1 - \lambda_2) < 1 \), where

\[
\begin{align*}
  r_1 & = \arg \min_{r,X} \left\{ r \geq 0 : p_1(r,X,\bar{X}) \geq 0 \right\} \\
  r_2 & = \arg \min_{r,X} \left\{ r \geq 0 : p_2(r,X,\bar{X}) \geq 0 \right\} \\
  r_3 & = \arg \min_{r,X} \left\{ r \geq 0 : p_3(r,X,\bar{X}) \geq 0 \right\} \\
  r_4 & = \arg \min_{r,X} \left\{ r \geq 0 : p(r,X) \geq 0 \right\}
\end{align*}
\]

\[ p(r,X) = \begin{bmatrix} \sqrt{T} I & A - AXC \\ (A - AXC)^T & \sqrt{T} I \end{bmatrix} \]

\[ p_1(r,X,\bar{X}) = p(r, N^4 X N^1 + N^3 X N^2) \]

\[ p_2(r,X,\bar{X}) = p(r, N^4 X N^1 + N^3 X N^2) \]

\[ p_3(r,X,\bar{X}) = p(r, N^4 X N^1 + \bar{X} N^2) \]

then \( \lim_{t \to +\infty} E(\pi_t) \) is bounded.

**Remark 2:** Note that \( p(r,X) \geq 0, p_1(r,X,\bar{X}) \geq 0, p_2(r,X,\bar{X}) \geq 0, \) and \( p_3(r,X,\bar{X}) \geq 0 \) are LMI problems (21a)-(21d) are SDP problems which can be effectively solved by MATLAB toolbox. Thus, the condition established in Theorem 2 can be effectively verified.

**Corollary 1:** Consider \( r_1, r_2, r_3, \) and \( r_4 \) defined in Theorem 2. It holds that \( r_4 \leq \min (r_2, r_3), r_1 \geq \max (r_2, r_3) \). In addition, if \( r_1 = r_4 \) and there exists \( r_1 < 1 \), for any \( i \in \{1, 2, 3, 4\} \), then \( \lim_{t \to +\infty} E(\pi_t) \) is bounded for any \( \lambda_1, \lambda_2 \in [0,1] \).

**Proof:** See Appendix E.

**Corollary 2:** Consider \( r_1, r_2, r_3, \) and \( r_4 \) defined in Theorem 2. Assume that the system parameters \((A, C)\) satisfy Assumption 1. If there exists a \( X \) satisfying \( c(X) < 1 \), then \( r_4 < r_1 \), where \( c(\cdot) \) is defined in (14).

**Proof:** See Appendix F.

Note that \( E(\pi_t) \) depends on \( \lambda_1 \) and \( \lambda_2 \). It is known that if the delay occurs with a bigger probability, then the estimation performance gets worse. This means that \( E(\pi_t) \) is monotone increasing with respect to \( \lambda_1 \) and \( \lambda_2 \). Additionally, if \( \lambda_1 \) and \( \lambda_2 \) are big enough, \( E(\pi_t) \) may become unbounded when the time goes to infinity. The above discussions are concluded as follows.

**Lemma 3:** Consider (11). For a fixed \( \lambda_1 \), if \( \lim_{t \to \infty} E(\pi_t) \) is unbounded for \( \lambda_2 = 1 \), but bounded for \( \lambda_2 = 0 \), then there exists a critical probability \( \lambda_{2,c} \) such that \( \lim_{t \to \infty} E(\pi_t) \) is bounded for \( \lambda_2 \leq \lambda_{2,c} \), and unbounded for \( \lambda_2 > \lambda_{2,c} \). Also, for a fixed \( \lambda_2 \), we have the similar results to \( \lambda_1 \).

**Proof:** See Appendix G.

Since the accurate critical probability of delay is not easy to obtain, the lower and upper bounds of the critical probability may be useful. A computation method is provided in the following theorem to compute the lower and upper bounds of the critical probability. For ease of notation, we first define

\[ q(X^1, X^2) = c(N^4 X N^1 + N^3 X N^2). \]

**Theorem 3:** Consider (11). For a fixed \( \lambda_1 \), the critical probability \( \lambda_{2,c} \) satisfies that if \( \lambda_2 \leq \lambda_{2,c} \), \( \lim_{t \to \infty} E(\pi_t) \) is bounded, otherwise, \( \lim_{t \to \infty} E(\pi_t) \) is unbounded. A lower and upper bound of \( \lambda_{2,c} \) is obtained, i.e., \( \lambda_{2,c} \leq \lambda_{2,c} \leq \bar{\lambda}_{2,c} \), where

\[ \bar{\lambda}_{2,c} = \begin{cases} 1, & \text{if } r_1 = r_4 < 1 \\ \min \left( \frac{1}{\alpha}, 1 \right), & \text{otherwise} \end{cases} \]

\[ \lambda_{2,c} = \begin{cases} 1 - \frac{r_2 - r_4 (1 - \alpha)}{r_1 - r_2 \lambda_1 + r_3 (1 - \alpha)}, & \text{if } r_1 \neq r_4 \\ 1, & \text{if } r_1 = r_4 < 1 \\ 0, & \text{if } r_1 = r_4 \geq 1 \end{cases} \]

where \( \alpha = \delta(\min_{X^1, X^2} q(X^1, X^2)) \). Similar upper and lower bounds for \( \bar{\lambda}_{2,c} \) can be obtained when we fix \( \lambda_2 \).

**Proof:** See Appendix H.

**Remark 3:** (Extension to Large-Scale ISs): Consider a large-scale IS composed of \( n \) subsystems. We can use a random graph \( G(V, E, p) \) to describe the communication between the subsystems, where \( V \) is the nodes set; \( E \) is the possible edges set; \( p = [p_1, \ldots, p_{|E|}] \). At time \( t \), the \( i \)th possible edge in \( E \) occurs in the graph with probability \( 0 < p_i < 1 \). The edge \((i, j)\)
occurring in the graph implies that the subsystem \( i \) can receive the information of subsystem \( j \) in real time. The continuity of \( \mathbb{E}(P_i) \) (w.r.t. \( p \)) indicates that there should exist a critical plane denoted by \( p_c \), such that the boundedness of \( \mathbb{E}(P_i) \) depends on which side of the plane \( p_c \) the point \( p \) is on. However, it is not easy to determine the upper and lower bounds of \( p_c \).

V. APPLICATIONS TO POWER SYSTEMS

In this section, a power system is used as an example to demonstrate the effectiveness of the proposed state estimation methods.

We study the state estimation problem for the distributed energy resources (DERs) connected to the power network that is shown in Fig. 2. In this example, the power network is chosen to be the IEEE 4-bus distribution network (see [22, Fig. 3]). As Fig. 2 depicts, four DERs are integrated into the main power network at the point common coupling (PCC). The voltages of the PCC are \( v_i = [v_1, v_2, v_3, v_4]^T \), where \( v_i \) is the voltage of the \( i \)-th PCC, and \( i \in \{1, 2, 3, 4\} \). Each DER is a voltage source at each bus. The input voltage of the voltage sources is denoted by \( v_p = [v_{p1}, v_{p2}, v_{p3}, v_{p4}]^T \). To maintain the proper operation of DERs, it is required to keep the PCC voltages \( v_i \) at reference values \( v_{ref} \). The PCC voltage deviation \( x_t = v_i - v_{ref} \) is chosen to be the system state. The DER control effort deviation \( u_t = v_p - v_{pref} \) is the control input, where \( v_{pref} \) is the reference of the control effort. Following the modeling process provided in [22] and [23], the dynamic of \( x_t \) is of the following form:

\[
\dot{x}_t = A_xt + B_ut + n_t
\]  

where \( n_t \) is a zero-mean Gaussian white noise whose covariance is \( Q \). The system (25) is discretized as follows:

\[
x_{t+1} = A_d x_t + B_d u_t + \omega_t
\]

where \( A_d = I + AT \), \( B_d = T_d B \), and \( \omega_t \) is the i.i.d. Gaussian process with zero-mean and covariance \( W = T_d Q \). Assume the initial state \( x_0 \) is a zero mean Gaussian variable with identity covariance. Here, we assume \( W = I \), and set the sampling time \( T_s = 0.05 \) s. In this example, \( \rho(A_d) > 1 \), which means that the system (26) under an open-loop pattern (there is no control input) is unstable. Here, the controller is of the state feedback form: \( u_t = Fx_t \), where \( F \) is the gain matrix with proper dimension. The closed-loop system is given by

\[
x_{t+1} = A_c x_t + \omega_t
\]

where \( A_c = A_d + B_d F \). As Fig. 2 shows, DERs 1 and 2 are in area 1 and DERs 3 and 4 are in area 2. Based on different areas, the system (27) can be written as an IS composed of two subsystems

\[
x_{t+1}^i = A_{d1} x_t^i + A_{d2} x_t^j + \omega_t^i, \quad i, j \in \{1, 2\}, \quad i \neq j
\]

where \( x_t^i \) is the component of the system state corresponding to area \( i \); and \( A_c = [A_{d1}^i]_{i \in \{1, 2\}}, x_t^i = [I_2 \ 0_2] x_t^i, \omega_t^1 = [I_2 \ 0_2] \omega_t, \) and \( \omega_t^2 = [0_2 \ I_2] \omega_t \). To monitor the working status of the power system (28), in each area, the sensors are employed to measure the system state. The measurement equations of area 1 and area 2 are

\[
y_t^i = C^i x_t^i + v_t^i, \quad i \in \{1, 2\}
\]

where \( v_t^i = [v_t^1, v_t^2]^T \) is measurement noise that is the zero-mean Gaussian independent processes with \( \mathbb{E}(v_t^i v_t^j) = I_2, \) \( C^1 = [I_2 \ 0_2], \) and \( C^2 = [0 \ 0 \ 1] \). In area \( i \), the estimator \( i \) is installed to estimate the state \( x_t^i \). The measurement \( y_t^i \) is transmitted to estimators 1 and 2 through network, so does \( y_t^j \). It is assumed that the data \( y_t^i \) transmitted to estimator \( j \) suffers from random delay (one step delay or no delay), where \( i \neq j \) and \( i, j \in \{1, 2\} \). Assume that the delay indicator \( y_t^{ij} \) is a Bernoulli stochastic process with \( \text{Pr}(y_t^{ij} = 0) = \lambda_1 \), where \( i \in \{1, 2\} \).

Case 1 (The Stable Closed-Loop System): The controller gain is chosen such that \( \rho(A_c) < 1 \), which means that the system (28) with the chosen \( F \) is stable. Then, our proposed method is applied for the state estimation of the system (28). The estimation performance is evaluated by the trace of the EEC \( \mathbb{E}(P_t) \). The trace of \( \mathbb{E}(P_t) \) with respect to different delay probabilities \( (\lambda_1, \lambda_2) \) is approximately computed by averaging 1000 Monte Carlo simulations, see Fig. 3. From Fig. 3, we can see that:

1) the estimation performance gets worse under larger delay probabilities;
2) the EEC \( \mathbb{E}(P_t) \) is bounded even though \( \lambda_1 = \lambda_2 = 1 \) (Note that \( \lambda_1 = \lambda_2 = 1 \) means that the delay always happens);
3) under different delay probabilities, the EEC \( \mathbb{E}(P_t) \) is within the same order of magnitude as the error covariance of standard Kalman filtering.

The simulations illustrate that our estimators can well monitor the working state of the power system (28), even though the data transmitted between different areas suffer from random delays.

Case 2 (The Unstable Closed-Loop System): It is known that to estimate the states of unstable systems is more challenging...
Fig. 4. $\lambda_1 = 1$, the value of $\text{tr}(E(P_t))$ under different $\lambda_2$.

Fig. 5. $\lambda_2 = 1$, the value of $\text{tr}(E(P_t))$ under different $\lambda_1$.

than to estimate the states of stable systems. Hence, to further illustrate the effectiveness of our estimators, we also consider the unstable system case. The controller gain is chosen such that $\rho(A_c) > 1$. Hence, the system (28) with the chosen controller gain is unstable. In this case, the values of $\text{tr}(E(P_t))$ with different ($\lambda_1, \lambda_2$) are presented in Figs. 4 and 5.

Figs. 4 and 5 show that the EEC $E(P_t)$ (with any delay probabilities) is within the same order of magnitude as the error covariance of standard Kalman filtering. This implies that even when the power systems (28) become unstable, the working states are also well monitored by our estimator under random delay. In addition, comparing Figs. 4 and 5, we find that for the power system with the parameters given in this case, the estimation performance is more sensitive to $\lambda_2$ than to $\lambda_1$.

VI. CONCLUSION

This article studied the optimal estimator design problem for IS with random delay. Due to the random delay occurring among the subsystems, the information available to different subsystem may be different. This type of IS is called IS with an asymmetric information structure. An optimal estimator has been analytically designed. The estimator realization algorithm for each subsystem was developed. In addition, some useful properties of the estimation performance were obtained. Finally, the proposed estimator was applied to a power system. The simulations showed that the designed estimator is effective and is of good performance.

In this work, we assume the delay indicator obeying the Bernoulli process. In the further work, we may extend this work to the Markov chain delay indicator case.

APPENDIX A

PROOF OF THEOREM 1

The matrices $L_t^{[1]}$, $U_t^3$, $U_t^6$, $J_t^3$, and $J_t^6$ are defined by the “inverse” operator. First, we show that these matrices are well-defined. It follows from $V + CP_{t|t-1}C^T \geq V > 0$ that $(V + CP_{t|t-1}C^T)^{-1}$ is well-defined. Thus, $L_t^{[1]}$ is well-defined. Since both $N_t^1$ and $N_t^2$ are row full rank, $N_t^1V_{N_t^1}^T + N_t^1CP_{t|t-1}C_{N_t^1}^T \geq N_t^1V_{N_t^1}^T > 0$, $N_t^2V_{N_t^2}^T + N_t^2CP_{t|t-1}C_{N_t^2}^T \geq N_t^2V_{N_t^2}^T > 0$. Thus, $U_t^3$ and $U_t^6$ are well-defined. Based on the Schur complement decomposition, it follows:

$$V + CP_{t|t-1}C^T = \begin{bmatrix} (U_t^3)^{-1} & U_t^4 \\ U_t^4 & (U_t^6)^{-1} \end{bmatrix} > 0$$

that $U_t^3U_t^3 < (U_t^6)^{-1}$ and $U_t^4U_t^6U_t^4 < (U_t^3)^{-1}$. As a result, $I - U_t^4U_t^6U_t^4 > 0$, $I - U_t^4U_t^6U_t^4 > 0$. This shows that $J_t^3$ and $J_t^6$ are well-defined.

Now, we start to prove that the optimal $L_t$ is of the form (10). For ease of notation, define $R_t^1 = \begin{bmatrix} K_t^{[1]} \\ K_t^{[2]} \end{bmatrix}$, $R_t^2 = \begin{bmatrix} \bar{K}_t^{[21]} \\ \bar{K}_t^{[22]} \end{bmatrix}$.

Case 1: If $\gamma_{t,1} = 1$ and $\gamma_{t,2} = 1$, then the optimal $L_t = L_t^{[1]}$ is the gain of the standard Kalman filtering [24].

Case 2: If $\gamma_{t,1} = 0$, $\gamma_{t,2} = 1$, then $L_t$ has the form $L_t = R_t^1N_t^1 + N_t^3K_t^{[22]}N_t^2$. Inserting $L_t = R_t^1N_t^1 + N_t^3K_t^{[22]}N_t^2$ into (9b), we have

$$P_{t|t} = \left(I - \left(R_t^1N_t^1 + N_t^3K_t^{[22]}N_t^2\right)C\right)P_{t|t-1}$$

$$\times \left(I - \left(R_t^1N_t^1 + N_t^3K_t^{[22]}N_t^2\right)C\right)^T$$

$$+ \left(R_t^1N_t^1 + N_t^3K_t^{[22]}N_t^2\right)V\left(R_t^1N_t^1 + N_t^3K_t^{[22]}N_t^2\right)^T.$$ 

Taking $R_t^1$ as a variable, $\text{tr}(P_{t|t})$ has the form $\text{tr}(P_{t|t}) = \text{tr}(R_t^1(U_t^3)^{-1}R_t^1 + \tilde{r}_{t,1})$, where $\tilde{r}_{t,1}$ is a linear function of $R_t^1$. Using the formula $\text{tr}(AXB^T) = \text{vec}(X)(B^T \otimes A)\text{vec}(X)$, one has $\text{tr}(P_{t|t}) = \text{vec}((R_t^1)^{-1})\otimes I)\text{vec}(R_t^1) + \tilde{r}_{t,1}$, where $\tilde{r}_{t,1}$ is a linear function of $\text{vec}(R_t^1)$. Thus, $[\partial^2 \text{tr}(P_{t|t})/\partial^2 \text{vec}(R_t^1)] = (U_t^3)^{-1}$ near $I > 0$. Similarly, taking $K_t^{[22]}$ as a variable, we have $[\partial^2 \text{tr}(P_{t|t})/\partial^2 \text{vec}(K_t^{[22]})] = (U_t^6)^{-1}$ near $I > 0$. Thus, $\text{tr}(P_{t|t})$ is convex with respect to both $R_t^1$ and $K_t^{[22]}$. The optimal $R_t^1$ and $K_t^{[22]}$ are given by solving $[\partial \text{tr}(P_{t|t})/\partial R_t^1] = 0$ and $[\partial \text{tr}(P_{t|t})/\partial K_t^{[22]}] = 0$. That is

$$\frac{\partial \text{tr}(P_{t|t})}{\partial R_t^1} = R_t^1(U_t^3)^{-1} + N_t^3K_t^{[22]}U_t^3 + U_t^6 = 0$$

$$\frac{\partial \text{tr}(P_{t|t})}{\partial K_t^{[22]}} = K_t^{[22]}(U_t^6)^{-1} + N_t^3R_t^1U_t^3 + N_t^3U_t^6 = 0$$

which gives

$$R_t^{1*} = -\left(N_t^3J_t^3J_t^3 + U_t^3\right)U_t^3, \quad K_t^{[22]*} = J_t^3J_t^6.$$ 

As a result, if $\gamma_{t,1} = 0$, $\gamma_{t,2} = 1$, then the optimal $L_t = R_t^{1*}N_t^1 + N_t^3K_t^{[22]*}N_t^2 = L_t^{[1]}$.

Case 3: If $\gamma_{t,1} = 1$, $\gamma_{t,2} = 0$, then $L_t$ is of the form $L_t = N_t^4K_t^{[11]}N_t^1 + R_t^{2*}N_t^2$. Following from the derivation similar to the one of case 2, we have the optimal $K_t^{[11]}$ and $R_t^2$ are given by:

$$K_t^{[11]*} = J_t^3J_t^4, \quad R_t^{2*} = -\left(N_t^4J_t^4J_t^4 + U_t^3\right)U_t^6.$$ 

Thus, if $\gamma_{t,1} = 1$, $\gamma_{t,2} = 0$, then the optimal $L_t = N_t^4K_t^{[11]*}N_t^1 + R_t^{2*}N_t^2 = L_t^{[1]}$. 

Authorized licensed use limited to the terms of the applicable license agreement with IEEE. Restrictions apply.
Case 4: If $γ_1 = 0, γ_2 = 0$, then $L_t = N^4K_1^1N^1 + N^3K_2^2N^2$. Similarly, we obtain that the optimal $K_t^1$ and $K_t^2$ are of the form

$$K_t^{1*} = -N^4T(N^3K_2^2U_t^1 + U_t^2)U_t^3 = -N^4T_1U_t^2U_t^3$$
$$K_t^{2*} = -N^3T(N^4K_1^1U_t^4 + U_t^5)U_t^8 = -N^3T_2U_t^5U_t^6.$$

Hence, if $γ_1 = 0, γ_2 = 0$, the optimal $L_t = N^4K_t^{1*}N^1 + N^3K_t^{2*}N^2 = L_t^{[00]}$. The proof is completed.

**APPENDIX B**

PROOF OF PROPOSITION 1

Equation (18) is obvious. We focus on proving (19). For any $Y^1, Y^2 ≥ 0$, let $Y = αY^1 + (1 − α)Y^2$. One has

$$g_{λ_1, λ_2}(Y) = f \left( L_0^{[00]} [Y], L_1^{[00]} [Y], L_2^{[00]} [Y], L_1^{[1]} [Y], Y \right)$$
$$= f \left( L_0^{[00]} [Y], L_1^{[00]} [Y], L_2^{[00]} [Y], L_1^{[1]} [Y], αY^1 + (1 − α)Y^2 \right)$$
$$= αf \left( L_0^{[00]} [Y], L_1^{[00]} [Y], L_2^{[00]} [Y], L_1^{[1]} [Y], Y^1 \right)$$
$$+ (1 − α)f \left( L_0^{[00]} [Y], L_1^{[00]} [Y], L_2^{[00]} [Y], L_1^{[1]} [Y], Y^2 \right)$$
$$≥ s_t ∈ A_0^{[00]}, s \in A_0^{[00]} \times A_0^{[1]}, \alpha (s^1, s^2, s^3, s^4, Y^1)$$
$$+ \min \left( \begin{array}{c}
X_{11} \\
0
\end{array} \right) \in \mathbb{R}^{n \times m_1}, \left( \begin{array}{c}
X_{22} \\
0
\end{array} \right) \in \mathbb{R}^{n \times m_2}
$$

where

$$A_0^{[00]} = \left\{ \left. \begin{array}{c}
X_{11} \\
0
\end{array} \right| X_{11} \in \mathbb{R}^{n \times m_1}, X_{22} \in \mathbb{R}^{n \times m_2} \right\}$$
$$A_0^{[1]} = \left\{ \left. \begin{array}{c}
X_{11} \\
0
\end{array} \right| X_{11} \in \mathbb{R}^{n \times m_1}, X_{22} \in \mathbb{R}^{n \times m_2} \right\}$$
$$A_0^{[2]} = \left\{ \left. \begin{array}{c}
X_{11} \\
0
\end{array} \right| X_{11} \in \mathbb{R}^{n \times m_1}, X_{22} \in \mathbb{R}^{n \times m_2} \right\}$$
$$A_0^{[1]} = \mathbb{R}^{n \times m}.$$

This shows that $g_{λ_1, λ_2}(Y)$ is concave. Using Jensen’s inequality, one has $\mathbb{E}(g_{λ_1, λ_2}(Y)) ≤ g_{λ_1, λ_2}(\mathbb{E}(Y))$. This completes the proof.

**APPENDIX C**

PROOF OF LEMMA 2

Sufficiency: Construct a sequence $Y_0, Y_1, Y_2, \ldots , \tilde{Y}_{t+1} = f \left( N^4X^1N^1 + N^3X^2N^2, X^3N^1 + N^3X^4N^2 \right)$

$$\min_{r, X, \tilde{X}} \left( r \geq 0 \rightarrow p_3(r, X, \tilde{X}) \right) \geq 0$$

is denoted by $X^{03}$, $\tilde{X}^{03}$. Then, we have $p_3(r_3, X^{03}, \tilde{X}^{03}) = p(r_3, N^4X^3N^1 + \tilde{X}^3N^2) > 0$. This implies that $r = r_3$, $X = N^4X^3N^1 + \tilde{X}^3N^2$ is a solution to $p(r, X) \geq 0$. It follows from the definition of $r_4$ [see (21d)] that $r_4 ≤ r_3$. Similarly, we can prove $r_4 ≤ r_2$, $r_2 ≤ r_1, r_3 ≤ r_1$. If $r_1 = r_4$, it follows from $r_4 ≤ r_3, r_4 ≤ r_2$, $r_2 ≤ r_1, r_3 ≤ r_1$ that $r_1 = r_2 = r_3 = r_4$. According to Theorem 2, if $r_1 = r_2 = r_3 = r_4 < 1$, $\lim_{t \rightarrow +\infty} \mathbb{E}(P_t)$ is bounded for any $λ_1, λ_2$. The proof is completed.

**APPENDIX D**

PROOF OF THEOREM 2

Using Schur complement for (21a) and (22b), we have there exists $X^1, X^2$ such that $(A − A(N^4X^1N^1 + N^3X^2N^2))C(A − A(N^4X^1N^1 + N^3X^2N^2))^\top ≤ r_1I$, which means that $\|A − A(N^4X^1N^1 + N^3X^2N^2)\| \leq \sqrt{r_1}$. It follows from the formula $\|A \otimes B\| = \|A\| \times \|B\|$ that $\|d(N^4X^1N^1 + N^3X^2N^2)\| ≤ r_1$. Similarly, there exists $X^3, X^4, X^7$ satisfying $\|d(N^4X^1N^1 + N^3X^2N^2)\| ≤ r_2, \|d(N^4X^1N^1 + N^3X^2N^2)\| ≤ r_2, \|d(X^4)\| ≤ 2, (31)$, the necessity is obvious. The proof is completed.
because \((A, C)\) is undetectable with \(\text{diag}\{1_{n_1 \times m_1}, 1_{n_2 \times m_2}\}\). This is a contradiction. As a result, \(r_1 \neq r_2\) if there exists \(X\) such that \(c(X) < I\). The proof is completed.

**APPENDIX G**

**PROOF OF LEMMA 3**

To prove Lemma 3, we need Proposition 1. It is known that Lemma 3 holds if \(\mathbb{E}(P_1)\) is monotone increasing with respect to \(\lambda_1, \lambda_2\). From Proposition 1, one has \(\mathbb{E}(P_{i+1}) = \mathbb{E}(g_{\lambda_1, \lambda_2}(P_i))\).

Hence, we only need to prove that \(g_{\lambda_1, \lambda_2}(Y)\) is monotone increasing with respect to \(\lambda_1, \lambda_2\) for any \(Y > 0\). In particular, we should show that the following.

1) If \(0 \leq \lambda_1 \leq 1\) is fixed and \(0 \leq \lambda_2^{[1]} \leq \lambda_2^{[2]} \leq 1\), then \(g_{\lambda_1, \lambda_2^{[1]}}(Y) \leq g_{\lambda_1, \lambda_2^{[2]}}(Y)\).

2) If \(0 \leq \lambda_2 \leq 1\) is fixed and \(0 \leq \lambda_1^{[1]} \leq \lambda_1^{[2]} \leq 1\), then \(g_{\lambda_1^{[1]}, \lambda_2}(Y) \leq g_{\lambda_1^{[2]}, \lambda_2}(Y)\).

From the definition of \(g_{\lambda_1, \lambda_2}(Y)\), one has

\[
g_{\lambda_1, \lambda_2^{[1]}(Y)} - g_{\lambda_1, \lambda_2^{[2]}(Y)} = \lambda_1 \left(\lambda_2^{[1]} - \lambda_2^{[2]}\right) b(L^{[0]}[Y], Y) + \lambda_1 \left(\lambda_2^{[2]} - \lambda_2^{[1]}\right) b(L^{[1]}[Y], Y)
+ (1 - \lambda_1) \left(\lambda_2^{[1]} - \lambda_2^{[2]}\right) b(L^{[0]}[Y], Y)
+ (1 - \lambda_1) \left(\lambda_2^{[2]} - \lambda_2^{[1]}\right) b(L^{[1]}[Y], Y)
+ (1 - \lambda_1) \left(\lambda_2^{[2]} - \lambda_2^{[1]}\right) b(L^{[1]}[Y], Y)
= \lambda_1 \left(\lambda_2^{[2]} - \lambda_2^{[1]}\right) b(L^{[1]}[Y], Y)
+ (1 - \lambda_1) \left(\lambda_2^{[2]} - \lambda_2^{[1]}\right) b(L^{[1]}[Y], Y).
\]

According to the definitions of \(b(X, Y), \) and \(L^r[Y], \) \(Y \in \{00, 01, 10, 11\}\) [see after (15), one has \(b(L^{[0]}[Y], Y) = \min_{X \in \Lambda^{[0]}} b(X, Y), \) \(b(L^{[1]}[Y], Y) = \min_{X \in \Lambda^{[1]}} b(X, Y), \) \(b(L^{[0]}[Y], Y) = \min_{X \in \Lambda^{[0]} \Lambda^{[1]}} b(X, Y), \) where \(\Lambda^{[0]}, \) \(\Lambda^{[1]}, \) and \(\Lambda^{[1]} \) are defined in the proof of Proposition 1. It follows from \(\Lambda^{[1]} \supseteq \Lambda^{[0]}, \) \(\Lambda^{[1]} \supseteq \Lambda^{[0]}\) that \(b(L^{[1]}[Y], Y) \leq b(L^{[0]}[Y], Y) \leq b(L^{[0]}[Y], Y).\) As a result, \(g_{\lambda_1, \lambda_2^{[1]}(Y)} - g_{\lambda_1, \lambda_2^{[2]}(Y)} \leq 0.\) Similarly, we can obtain that \(g_{\lambda_1^{[1]}, \lambda_2}(Y) - g_{\lambda_1^{[2]}, \lambda_2}(Y) \leq 0.\) This completes the proof.

**APPENDIX H**

**PROOF OF THEOREM 3**

Based on Lemma 3, we need to show two points.

**Point 1:** \(\lim_{t \to \infty} \mathbb{E}(P_t)\) is bounded if \(\lambda_2 \leq \lambda_{2,c}\).

**Point 2:** \(\lim_{t \to \infty} \mathbb{E}(P_t)\) is unbounded if \(\lambda_2 > \lambda_{2,c}\).

**Point 1:** For the case \(r_1 = r_2 \geq 1\), \(\lambda_{2,c} = 0\) is obtained directly from \(0 \leq \lambda_2 \leq 1\). For the case \(r_1 = r_2 < 1\), according to Corollary 1, \(\lim_{t \to \infty} \mathbb{E}(P_t)\) is bounded when \(\lambda_2 = 1\). Thus, \(\lambda_{2,c} = \lambda_{2,c} = 1\) if \(r_1 = r_2 < 1\). For the case \(r_1 \neq r_2\), according to (24), \(\lambda_2 \leq \lambda_{2,c}\) becomes

\[
\lambda_2 \leq \frac{1 - r_2 \lambda_1 - r_4 (1 - \lambda_1)}{(r_1 - r_2) \lambda_1 + (r_3 - r_4)(1 - \lambda_1)}.
\]

Since \(r_4 \leq \min(r_2, r_3), \) \(r_1 \geq \max(r_2, r_3),\) and \(r_1 \neq r_4,\) one has \(r_1 \neq r_2 \) or \(r_3 \neq r_4\) holds. Thus, \((r_1 - r_2) \lambda_1 + (r_3 - r_4)(1 - \lambda_1) > 0.\) Applying (33) to Theorem 2 gives that \(\lim_{t \to \infty} \mathbb{E}(P_t)\) is bounded.

**Point 2:** Define \(s_{\lambda_1, \lambda_2}(Y) = \lambda_1 \lambda_2 b(L^{[0]}[Y], Y)\). The proof is divided into two steps.

**Step 1:** To prove that \(\mathbb{E}(P_t) \geq s_{\lambda_1, \lambda_2}(P_0)\), where the definition of the composite function \(f(\cdot)\) is defined in “Notations” paragraph.

**Step 2:** To prove that \(\lim_{t \to \infty} s_{\lambda_1, \lambda_2}(P_0)\) is unbounded for \(\lambda_2 > \lambda_{2,c}\).

**Step 1:** Denote \(z_{\gamma}(Y) = b(L^r[Y], Y), \gamma \in \{00, 01, 10, 11\}\), and define

\[
z_{\gamma_{-1}, \gamma_{-2}, \ldots, \gamma_{-n}}(\cdot) = z_{\gamma_{-1}}(z_{\gamma_{-2}}(\ldots z_{\gamma_{-n}}(\cdot))).
\]

It follows from the mathematical expectation formula, (11) and (13) that

\[
\mathbb{E}[P_t] = \sum_{\gamma_0, \ldots, \gamma_{-t-1} \in \{00, 01, 10, 11\}} \Pr(\gamma_0, \ldots, \gamma_{-t-1}) \times z_{\gamma_{-1}, \gamma_{-2}, \ldots, \gamma_{-n}}(P_0).
\]

Because \(\gamma_t = [\gamma_t, \gamma_{t-1}]\) is an i.i.d. process and satisfies (3), one has

\[
\Pr(\gamma_0, \ldots, \gamma_{-t-1}) = \Pr(\gamma_0) \Pr(\gamma_1) \cdots \Pr(\gamma_{-t-1})
\]

where for any \(i \in \{0, \ldots, t - 1\}\)

\[
Pr(\gamma_t = [11]) = (1 - \lambda_1)(1 - \lambda_2), \quad Pr(\gamma_t = [01]) = (1 - \lambda_1) \lambda_2
\]

\[
Pr(\gamma_t = [10]) = \lambda_1(1 - \lambda_2), \quad Pr(\gamma_t = [00]) = \lambda_1 \lambda_2.
\]

It is known that \(s_{\lambda_1, \lambda_2}(P_0) = \lambda_1 \lambda_2^{.1}\).

From (35) and (36), we can obtain \(\mathbb{E}[P_t] = h_{\lambda_1, \lambda_2}(P_0) + \gamma, \) where \(\gamma \geq 0,\) because \(z_{\gamma_{-1}, \gamma_{-2}, \ldots, \gamma_{-n}}(P_0) \geq 0\) holds for any \(\gamma_0, \ldots, \gamma_{-t-1} \in \{00, 01, 10, 11\}.\) Thus, we have \(s_{\lambda_1, \lambda_2}(P_0) \leq \mathbb{E}[P_t].\)

**Step 2:** If \(\lambda_2 > \lambda_{2,c}\), then \(\lambda_1 \lambda_2 \lambda_1 \lambda_2 \min(\lambda_1, \lambda_2) \) is obtained directly from (15). Because \(\lambda_2 > \lambda_{2,c}\), then \(\lambda_1 \lambda_2 \lambda_1 \lambda_2 \min(\lambda_1, \lambda_2) > 0.\) Applying (33) to Theorem 2 gives that \(\lim_{t \to \infty} \mathbb{E}(P_t)\) is bounded.

**References**

[1] W. Li, K. Xiong, Y. Jia, and J. Du, “Distributed Kalman filter for multitarget tracking systems with coupled measurements,” IEEE Trans. Syst., Man, Cybern., Syst., vol. 51, no. 10, pp. 6599–6604, Oct. 2021.
[2] C. Danielson, S. A. Borotoff, and A. Chakrabarty, “Extremum seeking control with an adaptive gain based on gradient estimation error,” IEEE Trans. Syst., Man, Cybern., Syst., vol. 53, no. 1, pp. 152–164, Jan. 2023.

[3] S. Ding, Z. Ke, Z. Yue, C. Song, and L. Lu, “Noncontact multi physiological signals estimation via visible and infrared facial features fusion,” IEEE Trans. Instrum. Meas., vol. 71, no. 1, pp. 1–13, Sep. 2022.

[4] K. Ding, S. Dey, D. E. Quevedo, and L. Shi, “Stochastic game in remote estimation under DoS attacks,” IEEE Control Syst. Lett., vol. 1, no. 1, pp. 146–151, Jul. 2017.

[5] J. Cheng, Y. Wu, Z.-G. Wu, and H. Yan, “Nonstationary filtering for fuzzy Markov switching affine systems with quantization effects and deception attacks,” IEEE Trans. Syst., Man, Cybern., Syst., vol. 56, no. 10, pp. 6545–6554, Oct. 2022.

[6] B. Chen, D. W. C. Ho, W.-A. Zhang, and L. Yu, “Distributed dimensionality reduction fusion estimation for cyber-physical systems under DoS attacks,” IEEE Trans. Syst., Man, Cybern., Syst., vol. 49, no. 2, pp. 455–468, Feb. 2019.

[7] B. Sinopoli, L. Schenato, M. Franceschetti, K. Poolla, M. I. Jordan, and S. S. Sudan, “Kalman filtering with intermittent observations,” IEEE Trans. Autom. Control, vol. 49, no. 9, pp. 1453–1464, Sep. 2004.

[8] X. Liu and A. Goldsmith, “Kalman filtering with partial observation losses,” in Proc. IEEE Conf. Decis. Control, 2004, pp. 4180–4186.

[9] J. Zhou, G. Gu, and X. Chen, “Distributed Kalman filter over sensorless sensor networks in the presence of data packet drops,” IEEE Trans. Autom. Control, vol. 64, no. 4, pp. 1603–1610, Apr. 2019.

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

[11] 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.

[12] C. Han, H. Zhang, and M. Fu, “Optimal filtering for networked systems with Markovian communication delays,” Automatica, vol. 49, no. 10, pp. 3097–3104, 2013.

[13] J. Cheng, Y. Wu, H. Yan, Z.-G. Wu, and K. Shi, “Protocol-based filtering for fuzzy Markov affine systems with switching chain,” Automatica, vol. 80, pp. 1–10, Jul. 2022, Art. no. 110337.

[14] J. Ma and S. Sun, “Optimal linear estimators for systems with random sensor delays, multiple packet dropouts and uncertain observations,” IEEE Trans. Signal Process., vol. 59, no. 11, pp. 5181–5192, Nov. 2011.

[15] H. Ren, R. Lu, J. Xiong, Y. Wu, and P. Shi, “Optimal filtered and smoothed estimators for discrete-time linear systems with multiple packet dropouts under Markovian communication constraints,” IEEE Trans. Cybern., vol. 50, no. 9, pp. 4169–4181, Sep. 2020.

[16] M. Pohjola and H. Koivo, “Measurement delay estimation for Kalman filter in networked control systems,” IFAC Proc. Vol., vol. 64, no. 2, pp. 4192–4197, 2008.

[17] Y. Wang, J. Xiong, and D. W. Ho, “Decentralized control scheme for large-scale systems defined over a graph in presence of communication delays and random missing measurements,” Automatica, vol. 98, pp. 190–200, Dec. 2018.

[18] T. Souxes, I.-M. Grantis, and C. Vournas, “Effect of stochasticity on voltage stability support provided by wind farms: Application to the Hellenic interconnected system,” Electr. Power Syst. Res., vol. 170, pp. 48–56, May 2019.

[19] Y. Wang, J. Xiong, and D. W. Ho, “ Globally optimal state-feedback LQG control for large-scale systems with communication delays and correlated subsystem process noises,” IEEE Trans. Autom. Control, vol. 64, no. 10, pp. 4196–4201, Oct. 2019.

[20] W. Cai, R. Su, and B. Wang, “Optimal control of interconnected systems with time-correlated noises: Application to vehicle platoons,” Automatica, vol. 137, Mar. 2022, Art. no. 110018.

[21] Y. Wang, J. Xiong, and D. W. Ho, “Distributed LMMSE estimation for large-scale systems based on local information,” IEEE Trans. Cybern., vol. 52, no. 8, pp. 8528–8536, Aug. 2022.

[22] Y. Wang, J. Lai, and H. V. Poor, “Multicast routing for decentralized control of cyber physical systems with an application in smart grid,” IEEE J. Sel. Areas Commun., vol. 30, no. 6, pp. 1097–1107, Jul. 2012.

[23] M. M. Rana, L. Li, and S. W. Su, “ Distributed state estimation over unreliable communication networks with an application to smart grids,” IEEE Trans. Green Commun. Netw., vol. 1, no. 1, pp. 89–96, Mar. 2017.

[24] K. J. Åström, Introduction to Stochastic Control Theory. Chelmsford, MA, USA: Courier Corp., 2012.