



Mean-square convergent continuous state estimation of randomly switched linear systems with unobservable subsystems and stochastic output noises[☆]

Le Yi Wang^{a,*}, George Yin^b, Feng Lin^a, Michael P. Polis^c, Wen Chen^d

^a Department of Electrical and Computer Engineering, Wayne State University, Detroit, MI 48202, USA

^b Department of Mathematics, University of Connecticut, Storrs, CT 06269, USA

^c School of Engineering and Computer Science, Oakland University, Rochester, MI 48309, USA

^d Division of Engineering Technology, Wayne State University, Detroit, MI 48202, USA

ARTICLE INFO

Article history:

Received 7 December 2021

Received in revised form 6 June 2023

Accepted 16 June 2023

Available online 1 August 2023

Keywords:

Randomly switched linear systems

State estimation

Mean-square convergence

Observer design

Kalman–Bucy filters

Stochastic hybrid systems

ABSTRACT

This paper studies mean-square (MS) convergent observers for estimating continuous states of randomly switched linear systems (RSLs) with unobservable subsystems that are subject to stochastic output observation noises. When subsystems are unobservable and switching sequences are random, the classical Kalman–Bucy filters that are applied to observable sub-states are shown to be potentially divergent. It is also shown that unless the switching interval T can be selected to be sufficiently small from the outset, MS convergence may never be achieved, regardless of how the observers for the subsystems are designed. The critical threshold $T_{\max}$ on T is derived for MS convergent observers to be achievable. Under the condition $T < T_{\max}$, this paper introduces design algorithms for subsystem observers to generate a globally MS convergent observer for the entire continuous state. A fundamental design tradeoff between convergence speeds and steady-state estimation errors is analyzed. This paper extends our recent new framework and algorithms for strong convergent observer design in RSLs by including observation noises, considering multi-output systems, establishing new algorithms for MS convergence, and developing design tradeoff analysis. Examples and a practical case study are presented to illustrate the design procedures, main convergence properties, and error analysis.

Published by Elsevier Ltd.

1. Introduction

This paper studies continuous state estimation of randomly switched linear systems (RSLs) that are subject to stochastic output observation noises and whose subsystems may not be observable individually. Hybrid systems cover diversified applications that consist of interacting continuous dynamics and discrete events (Ezzine & Haddad, 1989; Lunze & Lamnabhi-Lagarrigue, 2009; Nerode & Kohn, 1992; Sun & Ge, 2005). Existing technical results on observability and observer designs for continuous states of hybrid systems are mostly for the class of hybrid systems called switched linear systems (Babaali & Egerstedt, 2004; Fliess,

Join, & Perruquetti, 2008; Sellami & Abderrahim, 2015; Wang, Khargonekar, & Beydoun, 1997; Zhao, Liu, Zhang, & Li, 2013).

In a deterministic framework, a switched linear system can be treated as a time-varying linear system when the switching sequence is known *a priori*. Extensive research has been conducted on observability, observer design, and other related properties in deterministic frameworks, such as observability and controllability (Ezzine & Haddad, 1989), various notions of observability and their testing conditions (Haddadi, Gazzam, & Benalia, 2019; Johnson, 2016; Küsters & Trenn, 2018), the concepts of distinguishability (Vidal, Chiuso, Soatto, & Sastry, 2003), geometric subspace characterization (Gomez-Gutierrez, Ramirez-Trevino, Ruiz-Leon, & Di Gennaro, 2010), hybrid observability under input probing (Babaali & Pappas, 2005), almost always observability (Arbib & De Santis, 2020), and parameter estimation (Farina, Garulli, & Giannitrapani, 2022). Recent results on observability, detectability, attractivity, observer design, and related convergence analysis (Bernard & Sanfelice, 2020; Goebel, Sanfelice, & Teel, 2012; Ríos, Davila, & Teel, 2019; Ríos, Dávila, & Teel, 2020; Sanfelice, Goebel, & Teel, 2007) accommodate hybrid systems with certain predictable or known jumps.

[☆] This research was supported in part by the National Science Foundation, United States under Grants DMS-2229108 and DMS-2229109. The material in this paper was not presented at any conference. This paper was recommended for publication in revised form by Associate Editor Paolo Frasca under the direction of Editor Sophie Tarbouriech.

* Corresponding author.

E-mail addresses: lywang@wayne.edu (L.Y. Wang), gyin@uconn.edu (G. Yin), flin@wayne.edu (F. Lin), polis@oakland.edu (M.P. Polis), wchenc@wayne.edu (W. Chen).

On the other hand, when the switching sequence is random, the techniques and approaches depart significantly from deterministic switched linear systems. The RSLs represent many real-world hybrid systems due to stochastic natures of diverse physical systems. RSLs are exemplified by power line interruptions in power systems, machine breakdowns in manufacturing systems, physical and cyber attacks in networked systems, random communication packet losses in communication networks, among many others (Li & Zhang, 2010; Yin, Sun and Wang, 2011). In principle, randomly switched systems can be modeled and treated as stochastic hybrid systems, stochastic systems with time-varying parameters (Dragan & Aberkane, 2020), or hybrid switching diffusions (Cassandras & Lygeros, 2018; Lygeros & Prandini, 2010; Yin, Wang and Sun, 2011; Yin & Zhu, 2010).

Typical large-scale complex systems involve many sensors for monitoring their internal states, however any single sensor or a local sensor cluster often cannot provide sufficient information for state estimation. The model of unobservable subsystems reflects this scenario. Within each time interval, only the observable subspace of the active subsystem can be estimated. The same subspace may become unobservable when the system is switched to another subsystem. As a result, its estimation error may grow exponentially. This paper introduces new methodologies for designing subsystem observers to achieve a globally mean-square (MS) convergent state observer for the entire RSLs.

This paper contains the following original contributions: (1) We formulate observer design problems under MS convergence for general multi-output RSLs with unobservable subsystems and observation noises. (2) We show that for RSLs, the classical Kalman–Bucy filters that are applied to observable sub-states do not guarantee convergence of the resulting RSLS observer. (3) We prove that there is a critical switching time interval limit, beyond which subsystem observers may become divergent regardless of how observer feedback gains are designed. (4) We introduce new algorithms to design subsystem observers and prove the MS convergence of the estimator for the entire state. (5) A fundamental design tradeoff is studied and demonstrated. To balance the conflicting objectives of convergence speeds and steady-state errors, a useful method for parameter selection is proposed.

In comparison to our recent papers (Wang, Yin, Lin, Polis, & Chen, 2022, 2023), which cover single-output systems under noise-free observations and target strong convergence, the current paper adds stochastic observation noises, leading to two interacting stochastic processes (observation noises and random switching), and treats general multi-output systems. By using MS errors, the switching time interval T can no longer be arbitrary. An upper bound on T is derived for convergent observers to be achievable. In Wang et al. (2022, 2023), there were several technical constraints that confined subsystem interactions. These constraints have been removed by using a different approach in this paper.

The rest of the paper is organized as follows. Section 2 describes notations and basic descriptions of RSLs. Section 3 develops observer structures for unobservable subsystems and derives the overall observation error dynamics for the entire system. Section 4 covers the design procedures. To motivate our new design algorithms, Section 4.1 first shows that the classical Kalman–Bucy filters applied to subsystem observer design may result in divergent observers when switching is random. Then new algorithms are introduced to design subsystem observers in Section 4.2. Section 5 discusses a fundamental design tradeoff between convergence speeds and steady-state estimation errors, and proposes an optimal parameter selection method. Section 6 employs a commonly used IEEE bus system in power systems to demonstrate model development of RSLs in practical systems, observer design procedures, convergence properties, and design tradeoffs. Finally, the main findings and their potential extensions are summarized in Section 7.

2. Preliminaries

For a column vector $v \in \mathbb{R}^n$, $\|v\|$ is its Euclidean norm. For a matrix $M \in \mathbb{R}^{n \times m}$, M' is its transpose, $\lambda(M)$ an eigenvalue of M , $\sigma(M) = \sqrt{\lambda(M'M)}$ a singular value of M , and $\sigma_{\max}(M)$ its largest singular value. The value $\sigma_{\max}(M)$ is also its operator norm induced by the Euclidean norm $\sigma_{\max}(M) = \|M\| = \sup_{\|v\|=1} \|Mv\|$. For a square matrix $M = [a_{ij}] \in \mathbb{R}^{n \times n}$, $\text{Tr}(M) = \sum_{i=1}^n a_{ii}$ is its trace. The kernel or null space of $M \in \mathbb{R}^{n \times m}$ is $\ker(M) = \{x \in \mathbb{R}^m : Mx = 0\}$ and its range is $\text{Range}(M) = \{y = Mx : x \in \mathbb{R}^m\}$. I_n is the n -dimensional identity matrix. For a subspace $\mathbb{U} \subseteq \mathbb{R}^n$ of dimension m , a matrix $M \in \mathbb{R}^{n \times m}$ is said to be a base matrix of $\mathbb{U}$, written as $M = \text{Base}(\mathbb{U})$, if the column vectors of M are linearly independent, and $\text{Range}(M) = \mathbb{U}$.

A function $y(t) \in \mathbb{R}$ in a time interval $[0, T)$ is piecewise continuously differentiable if $[0, T)$ can be divided into a finite number of subintervals $[t_{k-1}, t_k)$, $k = 1, \dots, \ell$, $t_0 = 0$, $t_\ell = T$ such that $y(t)$ is right continuous in $[t_{k-1}, t_k)$ and continuously differentiable, to any order as needed, in (t_{k-1}, t_k) . The space of such functions is denoted by $\mathbb{C}[0, T)$.

For a random variable q , $E(q)$ is its expectation. For a subset $S_0 \subseteq S = \{1, \dots, m\}$, the indicator function of the set S_0 is $\mathbb{1}_{q \in S_0} = 1$ if $q \in S_0$; and $\mathbb{1}_{q \in S_0} = 0$ otherwise. $\Pr\{\cdot\}$ is the probability. $\mathcal{N}(0, I_\rho)$ is the standard ρ -dimensional Gaussian distribution of mean zero and variance I_ρ .

2.1. Systems

For state observer design, we ignore the input to the system and consider a continuous-time RSLS with output observation noises

$$\begin{cases} \dot{x}(t) &= A(\alpha(t))x(t) \\ dy(t) &= C(\alpha(t))x(t)dt + \mathcal{E}(\alpha(t))dw \end{cases} \quad (1)$$

where $x(t) \in \mathbb{R}^n$ is the state, $y(t) \in \mathbb{R}^r$ is the output. The system matrices $A(\alpha(t)) \in \mathbb{R}^{n \times n}$, $C(\alpha(t)) \in \mathbb{R}^{r \times n}$, and $\mathcal{E}(\alpha(t)) \in \mathbb{R}^{r \times \rho}$ depend on the randomly switching process $\alpha(t)$ that takes m possible values in a finite discrete state space $S = \{1, \dots, m\}$. The noise process $w \in \mathbb{R}^\rho$ is the ρ -dimensional standard real-valued Brownian motion with $\rho \leq r$ and $(\mathcal{E}(i))' \mathcal{E}(i) > 0$, $i \in S$.

Remark 1. The output equation in (1) is a stochastic differential equation (SDE) that has been ubiquitously used as observation equations in stochastic systems. However, in the engineering literature, the output equation $\bar{y}(t) = C(\alpha(t))x(t) + n(t)$, in which $\bar{y}(t)$ is $\dot{y}$ in (1), is often used with $n(t)$ being a white noise and viewed as the derivative of the Brownian motion w . Since a Brownian motion is nowhere differentiable, $\bar{y}(t)$ is not well defined in the time domain and hence the expression is symbolic. This symbolic expression is useful in mean square estimation problems for linear time invariant (LTI) systems. However, it lacks the rigorous foundation for more complicated time-domain analysis. The rigorous SDE forms are now commonly used under the Ito sense stochastic calculus. Since this paper must treat two interacting stochastic processes (noise and switching), it is beneficial to have a mathematically rigorous expression so that future developments of more sophisticated interacting stochastic processes will have a solid foundation.

For each $i \in S$, the corresponding LTI system in (1) with matrices $C(i)$, $A(i)$, $\mathcal{E}(i)$ will be called the *ith subsystem of the RSLs* in (1).

Assumption 1. For a given time interval T , (a) the switching process $\alpha(t)$ can switch only at the instants kT , $k = 0, 1, 2, \dots$, generating a stochastic sequence $\{\alpha_k = \alpha(kT)\}$ (the skeleton

sequence). (b) The sequence $\{\alpha_k\}$ is independent and identically distributed (i.i.d.) with probability $\Pr\{\alpha_k = i\} = p_i > 0, i \in \mathcal{S}$, and $\sum_{i=1}^m p_i = 1$. (c) α_k is independent of $x(0)$ and the Brownian motion w . (d) $\alpha(t)$ can be directly measured, but it is not known before its occurrence.

Remark 2. Although T is not a design variable in constructing observers, it will be shown in Section 4 that there is a critical threshold $T_{\max}$ on T such that if $T > T_{\max}$, MS convergence of state observers may not be achievable no matter how the subsystem observers are designed. For this reason, T will be selected below $T_{\max}$. The critical upper bound $T_{\max}$ will be derived in Section 4.

As functions of α_k , the matrix sequences

$$\begin{aligned} A_k &= A(\alpha_k) = \sum_{i=1}^m \mathbb{1}_{\{\alpha_k=i\}} A(i), \\ C_k &= C(\alpha_k) = \sum_{i=1}^m \mathbb{1}_{\{\alpha_k=i\}} C(i), \\ \mathcal{E}_k &= \mathcal{E}(\alpha_k) = \sum_{i=1}^m \mathbb{1}_{\{\alpha_k=i\}} \mathcal{E}(i) \end{aligned}$$

are stochastic. Denote $x_k = x(kT)$, $k = 0, 1, \dots$. From $x_{k+1} = e^{A_k T} x_k$, the state transition mapping from x_0 to x_k is

$$x_k = e^{A_{k-1}T} \dots e^{A_0T} x_0 = H_k x_0, k = 1, \dots \quad (2)$$

where $H_k = e^{A_{k-1}T} \dots e^{A_0T}$.

2.2. Stochastic observability

For constant $A \in \mathbb{R}^{n \times n}$ and $C \in \mathbb{R}^{r \times n}$, and a finite time interval $[0, T)$, consider the noise-free output $\tilde{y}$, defined by the mapping $\mathcal{G} : \mathbb{R}^n \rightarrow \mathbb{C}[0, T)$, $\tilde{y}(t) = \mathcal{G}(x_0)(t) = Ce^{At}x_0$, $t \in [0, T)$. The kernel of the time function $\mathcal{G}$ is defined as $\text{Ker}(\mathcal{G}) = \{x_0 \in \mathbb{R}^n : \tilde{y}(t) = 0, t \in [0, T)\}$. Let W be the observability matrix of (C, A) :

$$W = \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{n-1} \end{bmatrix} \in \mathbb{R}^{m \times n}.$$

Lemma 1 (Kailath, 1980). $\text{Ker}(\mathcal{G}) = \text{ker}(W)$.

For the i th subsystem in $\mathcal{S}$, the matrices $A(i)$ and $C(i)$ are constant matrices, and its observability matrix is

$$W(i) = \begin{bmatrix} C(i) \\ C(i)A(i) \\ \vdots \\ C(i)A(i)^{n-1} \end{bmatrix}, i = 1, \dots, m \quad (3)$$

and the combined observability matrix for the set $\mathcal{S}$ is

$$W_{\mathcal{S}} = \begin{bmatrix} W(1) \\ W(2) \\ \vdots \\ W(m) \end{bmatrix}. \quad (4)$$

We note that $W(i)$ and $W_{\mathcal{S}}$ are deterministic matrices.

Remark 3. During system operation, the stochastic switching process α_k naturally induces stochastic matrix sequences $A_k \in \{A(1), \dots, A(m)\}$, $C_k \in \{C(1), \dots, C(m)\}$, and $\mathcal{E}_k \in \{\mathcal{E}(1), \dots, \mathcal{E}(m)\}$. Under the notation of this paper, all quantities involving time index k are stochastic quantities. On the other hand, the value sets, such as $\{A(1), \dots, A(m)\}$, are the finite sets induced by the space $\mathcal{S} = \{1, \dots, m\}$ and hence are not stochastic processes.

In this paper, we consider RSLs whose subsystems may be unobservable.

Assumption 2. (a) Subsystems may be unobservable, namely $\text{Rank}(W(i)) = n_i \leq n, i = 1, \dots, m$. (b) The combined observability matrix $W_{\mathcal{S}}$ is full rank.

Remark 4. In our subsequent technical treatment, we will focus on the more difficult case in which all subsystems are unobservable, including the potential case $n_i = 0$. The case $n_i = 0$ is a common engineering scenario that represents a total loss of sensing capability due to sensor failures, channel interruptions, packet losses, etc. The condition (b) of Assumption 2 ensures that their collective observable subspaces cover $\mathbb{R}^n$. The reader is referred to an example in Wang et al. (2023) that shows this condition to be a necessary condition for designing convergent observers for the entire state.

Definition 1 (Wang et al., 2023). For a given finite time interval $[0, \ell T)$ and switching sequence $\{\alpha_k, k = 0, \dots, \ell - 1\}$, the RSLs in (1) is said to be *stochastically observable* if the kernel of the mapping $\mathcal{G}$ satisfies $\text{Ker}(\mathcal{G}) = \{0\}$.

The mapping $\mathcal{G}$ is conditioned on α_k , and hence it is a random quantity. Define the matrix sequence

$$\mathcal{O}_{\ell} = \begin{bmatrix} W_0 \\ W_1 H_0 \\ \vdots \\ W_{\ell} H_{\ell-1} \end{bmatrix}, \ell = 0, 1, \dots, \quad (5)$$

where H_k is defined in (2) and $H_{-1} = I_n$. Since W_k is random, $\mathcal{O}_{\ell}$ is also a matrix-valued random variable.

The following basic lemma from our recent paper (Wang et al., 2023) will be used in this paper.

Lemma 2 (Wang et al., 2023). For a given finite time interval $[0, \ell T)$ and switching sequence $\{\alpha_k, k = 0, \dots, \ell - 1\}$, (a) $\text{Ker}(\mathcal{G}) = \text{ker}(\mathcal{O}_{\ell})$. (b) The RSLs in (1) is stochastically observable if and only if $\text{ker}(\mathcal{O}_{\ell}) = \{0\}$, or equivalently, $\mathcal{O}_{\ell}$ is full column rank.

3. Observer structures and error dynamics

The observer design procedure involves many factors. The main ideas of the design procedure can be summarized as follows.

- (1) For each subsystem i , which may be unobservable, the state space is decomposed into its observable and unobservable subspaces by using the Kalman decomposition. The observer for the i th subsystem only estimates the observable sub-state z^i when $\alpha_k = i$. The observer feedback gain L_i for estimating z^i must be designed. This observer structure is discussed in Section 3.1.
- (2) Since the design of L_i involves further coordination that depends on the switching process, it must take into consideration the error analysis for the entire system. As a result, the overall error dynamics must be derived. This is presented in Section 3.2.
- (3) Section 4 discusses the design procedures for L_i . We first show by a counterexample in Section 4.1 that using Kalman filter designs for L_i can potentially lead to divergent observers for the entire state. Then our new and convergent observer design methods and algorithms are introduced in Section 4.2. This design procedure generates a collection of subsystem observers for observable sub-states. During implementation, these observers are used according to α_k that occurs randomly. MS convergence of the combined observer is established for the entire state. Section 4.3 derives a critical threshold $T_{\max}$ such that when $T < T_{\max}$ convergent observers can be designed.

3.1. Observer structure for subsystems

We first review the subspace decomposition on each subsystem, developed in Wang et al. (2023). Feedback-based linear observer design for subsystems is used so that the observers can be robust against errors, the design step can be simple and constructive, and convergence analysis can utilize properties of LTI systems.¹

Let $W(i)$ be the observability matrix of the i th subsystem defined in (3). If the i th subsystem is unobservable, then $\text{Rank}(W(i)) = n_i < n$. We construct the base of its kernel as $M_i = \text{Base}(\ker(W(i))) \in \mathbb{R}^{n \times (n-n_i)}$, and select any $N_i \in \mathbb{R}^{n \times n_i}$ such that $T_i = [M_i, N_i]$ is invertible. The inverse of T_i is decomposed into $T_i^{-1} = \begin{bmatrix} G_i \\ F_i \end{bmatrix}$, where $G_i \in \mathbb{R}^{(n-n_i) \times n}$ and $F_i \in \mathbb{R}^{n_i \times n}$.

Define $\tilde{z}_i = T_i^{-1}x = \begin{bmatrix} G_i x \\ F_i x \end{bmatrix} = \begin{bmatrix} v^i \\ z^i \end{bmatrix}$, where $z^i \in \mathbb{R}^{n_i}$

represents the observable sub-state of the i th subsystem. We focus on constructing a subsystem observer for estimating the sub-state z^i when $\alpha_k = i$. Denote $A^i = T_i^{-1}A(i)T_i$, $C^i = C(i)T_i$, which have the structure

$$A^i = \begin{bmatrix} A_{11}^i & A_{12}^i \\ 0 & A_{22}^i \end{bmatrix}, \quad C^i = [0, C_2^i]$$

with $A_{22}^i \in \mathbb{R}^{n_i \times n_i}$, $C_2^i \in \mathbb{R}^{r \times n_i}$, and (C_2^i, A_{22}^i) is observable with respect to the sub-state z^i .

The observable part z^i of the i th subsystem can be estimated when $\alpha_k = i$, but in general it may belong to the unobservable subspace of the j th subsystem if $\alpha_k = j \neq i$. As a result, estimation errors on z^i are fundamentally different. Consequently, the error dynamics on z^i estimation are divided into the following two cases.

Case 1: $\alpha_k = i$. When $\alpha_k = i$, the open-loop dynamics of z^i are

$$\begin{cases} \dot{z}^i &= A_{22}^i z^i \\ dy &= C_2^i z^i dt + \mathcal{E}_k dw \end{cases} \quad (6)$$

and (C_2^i, A_{22}^i) is observable. The observer for the i th subsystem in $[kT, (k+1)T)$ for estimating the sub-state z^i is

$$d\tilde{z}^i = A_{22}^i \tilde{z}^i dt + L_i(dy - C_2^i \tilde{z}^i dt) \quad (7)$$

where $L_i \in \mathbb{R}^{n_i \times r}$ is the constant observer feedback gain. Denote $A_c^i = A_{22}^i - L_i C_2^i$ and $e^i = \tilde{z}^i - z^i$. Then

$$de^i = A_c^i e^i dt + L_i \mathcal{E}_k dw \quad (8)$$

where L_i is designed such that $A_c^i = A_{22}^i - L_i C_2^i$ is stable in the continuous-time domain, i.e. all eigenvalues of A_c^i are in the open left half plane.

For convergence analysis and error variance computation, we need to obtain $E(e_k^i e_k^{i'})$. Note that $e^i \in \mathbb{R}^{n_i}$, $A_c^i \in \mathbb{R}^{n_i \times n_i}$, $L_i \mathcal{E}_k \in \mathbb{R}^{n_i \times \rho}$, and $w(\cdot)$ is the ρ -dimensional standard Brownian motion. Consider the stochastic differential equation (8) for $t \in [kT, kT + T)$ with initial data e_k^i . Obviously $e^i(t)$ is normally distributed and its distribution is completely specified by its mean $m^i(t) = E(e^i(t))$ and covariance $R^i(t, s) = E(e^i(t) - m^i(t))(e^i(s) - m^i(s))'$. Denote $m_k^i = m^i(kT)$ and $V_k^i = E((e_k^i - m_k^i)(e_k^i - m_k^i)')$.

Let $e^i(t)$ be a solution of the stochastic differential equation given by (8) together with initial data e_k^i .

Proposition 1. Suppose that A_c^i is Hurwitz, i.e., all of its eigenvalues are in the open left half plane of the complex plane. Then the mean vector $m^i(t)$ and covariance matrix $R^i(t, s)$ of $e^i(t)$ are given by $m^i(t) = e^{A_c^i(t-kT)} m_k^i$ and

$$\begin{aligned} R^i(t, s) &= e^{A_c^i(t-kT)} V_k e^{(A_c^i)'(s-kT)} \\ &+ \int_{kT}^{s \wedge t} e^{A_c^i((s \wedge t) - \tau)} (L_i \mathcal{E}_k)(L_i \mathcal{E}_k)' e^{(A_c^i)'((s \wedge t) - \tau)} d\tau, \end{aligned} \quad (9)$$

where $s \wedge t = \min(s, t)$.

Proof. It is easily seen that $e^i(t) = e^{A_c^i(t-kT)} e_k^i + \int_{kT}^t e^{A_c^i(t-\tau)} L_i \mathcal{E}_k dw(\tau)$. By taking the expectation, $m^i(t)$ is obtained. Furthermore,

$$\begin{aligned} R^i(t, s) &= E([e^{A_c^i(t-kT)}(e_k^i - m_k^i) + \int_{kT}^t e^{A_c^i(t-\tau)} L_i \mathcal{E}_k dw(\tau)] \\ &\times [e^{A_c^i(s-kT)}(e_k^i - m_k^i) + \int_{kT}^s e^{A_c^i(s-\nu)} L_i \mathcal{E}_k dw(\nu)]') \\ &= [e^{A_c^i(t-kT)} V_k e^{(A_c^i)'(s-kT)}] \\ &+ E \left[\int_{kT}^t e^{A_c^i(t-\tau)} L_i \mathcal{E}_k dw(\tau) \right] \left[\int_{kT}^s e^{(A_c^i)'(s-\nu)} L_i \mathcal{E}_k dw(\nu) \right]'. \end{aligned} \quad (10)$$

For $kT < t \leq s$,

$$\begin{aligned} E \left[\int_{kT}^t e^{A_c^i(t-\tau)} L_i \mathcal{E}_k dw(\tau) \right] \left[\int_{kT}^s e^{(A_c^i)'(s-\nu)} L_i \mathcal{E}_k dw(\nu) \right]' \\ = \int_{kT}^t e^{A_c^i(t-\tau)} L_i \mathcal{E}_k (L_i \mathcal{E}_k)' e^{(A_c^i)'(t-\tau)} d\tau. \end{aligned} \quad (11)$$

Likewise, for $kT < s \leq t$,

$$\begin{aligned} E \left[\int_{kT}^t e^{A_c^i(t-\tau)} L_i \mathcal{E}_k dw(\tau) \right] \left[\int_{kT}^s e^{(A_c^i)'(s-\nu)} L_i \mathcal{E}_k dw(\nu) \right]' \\ = \int_{kT}^s e^{A_c^i(s-\tau)} L_i \mathcal{E}_k (L_i \mathcal{E}_k)' e^{(A_c^i)'(s-\tau)} d\tau. \end{aligned} \quad (12)$$

Combining the above two expressions and using (10), we obtain the last expression in (9). $\square$

By Proposition 1, for $t = s = (k+1)T$, we have $m_{k+1}^i = e^{A_c^i T} m_k^i$, and $V_{k+1}^i = e^{A_c^i T} V_k e^{(A_c^i)' T} + \int_0^T e^{A_c^i \tau} L_i \mathcal{E}_k (L_i \mathcal{E}_k)' e^{(A_c^i)' \tau} d\tau$. Denote $Q_k^i = (V_{k+1}^i)^{1/2}$. Since e_{k+1}^i is Gaussian, which is completely determined by m_{k+1}^i and Q_k^i , e_{k+1}^i can be expressed as

$$e_{k+1}^i = e^{A_c^i T} e_k^i + Q_k^i d_k \quad (13)$$

where $\{d_k\}$ is a sequence of i.i.d. random variables, and $d_k \sim N(0, I_\rho)$.

Case 2: $\alpha_k = j \neq i$. When $\alpha_k = j \neq i$, in general z^i may not be observable in the j th subsystem, namely the observed output y may not contain any information on the sub-state z^i . Since the observer is feedback-based by using y , in this case, the subsystem observer for z^i runs open-loop, using only the system model without feedback correction. To derive the dynamics, we need the mapping from the estimate for the entire state, which is presented next.

$$\text{Define } n_s = \sum_{i=1}^m n_i \text{ and } F = \begin{bmatrix} F_1 \\ \vdots \\ F_m \end{bmatrix} \in \mathbb{R}^{n_s \times n}.$$

Lemma 3 (Wang et al., 2023). (a) $\ker(F_i) = \ker(W(i))$. (b) $\ker(F) = \ker(W_S)$.

¹ For more details on feedback-based linear observer design such as pole placement design of full-order observers and Luenberger observers, the Kalman decomposition, and observer stability analysis, the reader is referred to control engineering textbooks (Kailath, 1980; Kuo & Golnaraghi, 1994).

From subsystem observers, define $z = \begin{bmatrix} z^1 \\ \vdots \\ z^m \end{bmatrix} \in \mathbb{R}^{n_s}$, and its estimate $\hat{z} = \begin{bmatrix} \hat{z}^1 \\ \vdots \\ \hat{z}^m \end{bmatrix} \in \mathbb{R}^{n_s}$. Then, $z = Fx$ and $\hat{z} = F\hat{x}$.

Under [Assumption 2](#), F is of (column) rank n . As a result, $\Phi = (F'F)^{-1}F' \in \mathbb{R}^{n \times n_s}$ is of (row) rank n , and $x = \Phi z$. Consequently, their sampled values are $x_k = \Phi z_k$, $\hat{x}_k = \Phi \hat{z}_k$.

Remark 5. [Assumption 2](#) means that although each subsystem may be unobservable individually, the span of all observable subspaces covers the whole state space collectively. This is reflected by the rank of F , resulting in the bounded mapping $x_k = \Phi z_k$. As a result, convergence of $\hat{z}_k$ implies that of $\hat{x}_k$.

Since the true system is $x_{k+1} = e^{A_k T} x_k$ and A_k is known at kT , the true sampled value of the subsystem state is $z_{k+1}^i = F_i x_{k+1} = F_i e^{A_k T} x_k$. Consequently, the subsystem observer runs open-loop with $\hat{z}_{k+1}^i = F_i \hat{x}_{k+1} = F_i e^{A_k T} \hat{x}_k$ if $\alpha_k = j \neq i$. The observer for estimating x is $\hat{x}_k = \Phi \hat{z}_k$. Denote the estimation errors $\epsilon_k = \hat{x}_k - x_k$ and $e_k = \hat{z}_k - z_k$. We have $\epsilon_k = \Phi e_k$ and when $\alpha_k = j \neq i$

$$e_{k+1}^i = F_i e^{A_k T} \epsilon_k = F_i e^{A_k T} \Phi e_k. \quad (14)$$

3.2. Observer error analysis

Since these errors are related by the bounded mappings $\epsilon_k = \Phi e_k$ and $e_k = F \epsilon_k$, the following analysis on state estimation errors and their convergence will concentrate on e_k .

Theorem 1. *The observation error dynamics can be expressed as*

$$e_{k+1} = \Lambda_k e_k + \Gamma_k d_k \quad (15)$$

where $\Lambda_k = \Lambda^1 + \Lambda^2 F e^{A_k T} \Phi$, $\Lambda^1 = \text{diag}[\mathbb{1}_{\{\alpha_k=i\}} e^{A_k^i T}]$, $\Lambda^2 = \text{diag}[\mathbb{1}_{\{\alpha_k \neq i\}} I_{n_i}]$, and $\Gamma_k = \begin{bmatrix} \mathbb{1}_{\{\alpha_k=1\}} Q_k^1 \\ \vdots \\ \mathbb{1}_{\{\alpha_k=m\}} Q_k^m \end{bmatrix}$.

Proof. For the i th subsystem, by (13) and (14),

$$e_{k+1}^i = \begin{cases} e^{A_k^i T} e_k^i + Q_k^i d_k, & \text{if } \alpha_k = i \\ F_i e^{A_k T} \Phi e_k, & \text{if } \alpha_k \neq i \end{cases}$$

and

$$\begin{aligned} e_{k+1} &= \begin{bmatrix} e_{k+1}^1 \\ \vdots \\ e_{k+1}^m \end{bmatrix} \\ &= \begin{bmatrix} \mathbb{1}_{\{\alpha_k=1\}} (e^{A_k^1 T} e_k^1 + Q_k^1 d_k) + \mathbb{1}_{\{\alpha_k \neq 1\}} F_1 e^{A_k T} \Phi e_k \\ \vdots \\ \mathbb{1}_{\{\alpha_k=m\}} (e^{A_k^m T} e_k^m + Q_k^m d_k) + \mathbb{1}_{\{\alpha_k \neq m\}} F_m e^{A_k T} \Phi e_k \end{bmatrix} \\ &= \Lambda^1 e_k + \Lambda^2 F e^{A_k T} \Phi e_k + \Gamma_k d_k = \Lambda_k e_k + \Gamma_k d_k. \quad \square \end{aligned}$$

This section has established the observer structure for estimating observable sub-states of subsystems, the combined observer for z (and hence for $x = \Phi z$), and estimation error dynamics. To achieve convergence, the error dynamics must be stable in the MS sense. This stability condition depends on the design of L_i . The next section will discuss design methods and present convergence results.

4. Convergent observer design

4.1. Kalman-Bucy filters and instability

For the i th subsystem, when $\alpha_k = i$, the dynamics of the sub-state z^i are governed by (6) and (C_2^i, A_{22}^i) is observable. One possible approach is to design the observer gain L_i in (7) by using the Kalman-Bucy filters. The classical Kalman-Bucy filters, see [Anderson and Moore \(1979\)](#), [Kalman \(1963\)](#) and [Zarchan and Musof \(2005\)](#) for detail, deal with systems with both state and output stochastic noises,

$$\begin{cases} dx(t) = Ax(t)dt + Q^{1/2}dv \\ dy(t) = Cx(t)dt + R^{1/2}dw \end{cases} \quad (16)$$

where $Q^{1/2}$ and $R^{1/2}$ satisfy $Q^{1/2}(Q^{1/2})' = Q \geq 0$ and $R^{1/2}(R^{1/2})' = R > 0$, and v and w are mutually independent standard Brownian motions. The initial value $x(0)$ is independent of v and w , and Gaussian distributed $x(0) \sim \mathcal{N}(x_0, P_0)$ with $P_0 > 0$. The Kalman-Bucy filters start with the initial variance $P(0) = P_0$ and initial estimate $\hat{x}(0) = \hat{x}_0$, and is represented by

$$\dot{P}(t) = AP(t) + P(t)A' + Q - K(t)RK'(t)$$

$$K(t) = P(t)C'R^{-1}$$

$$d\hat{x}(t) = A\hat{x}(t)dt + K(t)(dy(t) - C\hat{x}(t)dt),$$

with the related limiting Algebraic Riccati Equation (ARE)

$$AP + PA' + Q - PC'R^{-1}CP = 0.$$

Since switched unobservable linear systems involve multiple subsystems and stochastic switching, the switching process will have critical impact on MS convergence of observers. We now show by a counterexample that the Kalman-Bucy filters designed to estimate the observable sub-state at each subsystem by using the ARE may lead to a divergent observer for the entire state of the RSLs.

Example 1. Consider a second-order system without state noise

$$\begin{cases} \dot{x}_1(t) = a_1 x_1(t) \\ \dot{x}_2(t) = a_2 x_2(t) \\ dy = (\mathbb{1}_{\{\alpha(t)=1\}} x_1(t) + \mathbb{1}_{\{\alpha(t)=2\}} x_2(t))dt + \sqrt{\zeta} dw \end{cases}$$

where $a_1 > 0$, $a_2 > 0$, $\alpha(t) \in \{1, 2\}$ and w is the standard Brownian motion. Since $A = \begin{bmatrix} a_1 & 0 \\ 0 & a_2 \end{bmatrix}$, $C(1) = [1, 0]$, $C(2) =$

$[0, 1]$, the observability matrices are $W(1) = \begin{bmatrix} 1 & 0 \\ a_1 & 0 \end{bmatrix}$, $W(2) =$

$\begin{bmatrix} 0 & 1 \\ 0 & a_2 \end{bmatrix}$. As a result, both subsystems $(C(1), A)$ and $(C(2), A)$ are unobservable.

Within $[kT, (k+1)T)$, if $\alpha_k = 1$, only x_1 is observed with noise. The filter takes the form

$$\begin{cases} d\hat{x}_1(t) = a_1 \hat{x}_1(t)dt + L_1(dy - \hat{x}_1(t)dt) \\ \hat{x}_2(t) = a_2 \hat{x}_2(t). \end{cases}$$

If $\alpha_k = 2$, only x_2 is observed with noise. The filter takes the form

$$\begin{cases} \hat{x}_1(t) = a_1 \hat{x}_1(t) \\ d\hat{x}_2(t) = a_2 \hat{x}_2(t)dt + L_2(dy - \hat{x}_2(t)dt). \end{cases}$$

To calculate the Kalman gains L_1 and L_2 , we use the ARE. If $\alpha_k = 1$, the Kalman gain is $L_1 = P_1/\zeta$, and P_1 satisfies $2a_1 P_1 - \frac{1}{\zeta} P_1^2 = 0$, resulting in $P_1 = 2a_1 \zeta$ and $L_1 = 2a_1$. Similarly, if $\alpha_k = 2$, $L_2 = 2a_2$. The estimation error dynamics for $e^1 = \hat{x}_1 - x_1$ and $e^2 = \hat{x}_2 - x_2$ are

$$\begin{cases} de^1 = (\mathbb{1}_{\{\alpha(t)=1\}}(-a_1) + \mathbb{1}_{\{\alpha(t)=2\}}a_1)e_1(t)dt + L_1\sqrt{\zeta}dw, \\ de^2 = (\mathbb{1}_{\{\alpha(t)=1\}}a_2 + \mathbb{1}_{\{\alpha(t)=2\}}(-a_2))e_2(t)dt + L_2\sqrt{\zeta}dw. \end{cases} \quad (17)$$

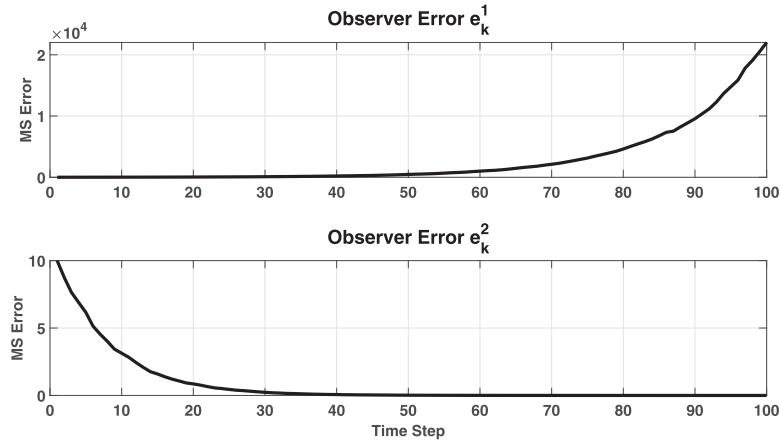


Fig. 1. Observer error dynamics under Kalman-Bucy design.

Suppose that $\Pr\{\alpha_k = 1\} = p > 0$ and $\Pr\{\alpha_k = 2\} = 1 - p$. For stability analysis, we may consider only the noise-free part for e^1 , whose sampled values are

$$e_{k+1}^1 = \gamma_k^1 e_k^1 \quad (18)$$

where $\gamma_k^1 = e^{-a_1 T} \mathbb{1}_{\{\alpha_k=1\}} + e^{a_1 T} \mathbb{1}_{\{\alpha_k=2\}}$.

To evaluate the MS errors, denote $\sigma_k^i = E((e_k^i)^2)$, $i = 1, 2$. By [Assumption 1](#), $\{\alpha_k\}$ is a sequence of i.i.d. random variables and independent of $x(0)$. This implies that γ_k^i and σ_k^i are independent, and $\sigma_{k+1}^i = E_{\alpha_k}((\gamma_k^i)^2 (e_k^i)^2) = \bar{\gamma}^i \sigma_k^i$, where $\bar{\gamma}^1 = E_{\alpha_k}((\gamma_k^1)^2) = pe^{-2a_1 T} + (1-p)e^{2a_1 T}$.

If $\bar{\gamma}^1 > 1$, $\sigma_k^1 \rightarrow \infty$, as $k \rightarrow \infty$, namely, the filter's error dynamics are divergent in MS. For a numerical example, let $a_1 = 2$, $a_2 = 4$, $T = 0.03$, $p = 0.2$. Then, we can calculate that $\bar{\gamma}^1 = pe^{-2a_1 T} + (1-p)e^{2a_1 T} = 1.0794 > 1$, $\bar{\gamma}^2 = (1-p)e^{-2a_2 T} + pe^{2a_2 T} = 0.8836 < 1$. These imply that when we use MS errors, the error dynamics for e_k^1 are unstable although the error dynamics for e_k^2 are stable. [Fig. 1](#) verifies this conclusion.

This example demonstrates that in general the Kalman-Bucy filters designed on a given sample path of $\{\alpha_k\}$ and applied to the observable sub-states are not convergent in RSLs with unobservable subsystems. This is mainly due to excluding the stochastic information of the switching process α_k in design.

We now introduce new algorithms to overcome the divergence issues of Kalman-Bucy filters, develop MS convergent observers, and derive their main convergence properties.

4.2. Design of the observer feedback gain L_i

The design of L_i aims to achieve the MS convergence of e_k . The following theorem provides a design criterion.

Theorem 2. *If the feedback gains L_i , $i = 1, \dots, m$, are designed such that $\gamma_k = E(\|A_k\|^2) \leq \gamma < 1$, then e_k is MS convergent.*

Proof. From the error dynamics $e_{k+1} = A_k e_k + \Gamma_k d_k$, the solution is

$$e_k = (\Pi_{j=0}^{k-1} A_j) e_0 + \sum_{j=0}^{k-1} A_{k-j} \Gamma_j d_j.$$

For convergence analysis of the error dynamics, we first focus on the noise-free dynamics $e_{k+1} = A_k e_k$. From $\|e_{k+1}\|^2 \leq \|A_k\|^2 \|e_k\|^2$ and independence of A_k and e_k , we have

$$\begin{aligned} E(\|e_{k+1}\|^2) &\leq E(\|A_k\|^2) E(\|e_k\|^2) \\ &= \gamma_k E(\|e_k\|^2) \\ &\leq \gamma E(\|e_k\|^2). \end{aligned}$$

Since $\gamma < 1$, the noise-free system is MS exponentially stable.

Furthermore, since A_k , e_k , and d_k are mutually independent, and Γ_k and d_k are independent,

$$\begin{aligned} E(\|e_{k+1}\|^2) &= E(e_k' A_k' A_k e_k) + E(d_k' \Gamma_k' \Gamma_k d_k) \\ &\leq \gamma E(\|e_k\|^2) + \Xi_k \end{aligned}$$

for some $\Xi_k \geq 0$. Since $\gamma < 1$, this is a stable system. It follows that $E(\|e_k\|^2)$ is convergent. $\square$

To achieve the design criterion $E(\|A_k\|^2) \leq \gamma < 1$, there is a constraint on T . We first establish a sufficient condition showing the existence of design algorithms for L_i , $i = 1, \dots, m$, that can satisfy this condition. Then we use a counterexample to explain why this constraint on T is necessary in general.

Theorem 3. *Under [Assumption 2](#), there exists $T_{\max} > 0$ such that for any $T < T_{\max}$, L_i , $i = 1, \dots, m$, can be designed to satisfy $\gamma_k \leq \gamma < 1$.*

Proof. By the definitions of Λ^1 and Λ^2 , we have $(\Lambda^1)' \Lambda^2 = 0$, and as a result,

$$\begin{aligned} A_k' A_k &= (\Lambda^1 + \Lambda^2 F e^{A_k T} \Phi)' (\Lambda^1 + \Lambda^2 F e^{A_k T} \Phi) \\ &= (\Lambda^1)' \Lambda^1 + (\Lambda^2 F e^{A_k T} \Phi)' \Lambda^2 F e^{A_k T} \Phi. \end{aligned}$$

By Singular Value Decomposition, $\|A_k\|^2 = \|(\Lambda_k)' \Lambda_k\| \leq \|(\Lambda^1)' \Lambda^1\| + \|(\Lambda^2 F e^{A_k T} \Phi)' \Lambda^2 F e^{A_k T} \Phi\| = \|\Lambda^1\|^2 + \|\Lambda^2 F e^{A_k T} \Phi\|^2$. It follows that $E(\|A_k\|^2) \leq E(\|\Lambda^1\|^2) + E(\|\Lambda^2 F e^{A_k T} \Phi\|^2)$.

Since the poles of $e^{A_k T}$ can be arbitrarily placed, $\gamma^i = \|e^{A_k T}\|^2$ can be made arbitrarily small. As a result, $\gamma_{\max} = \max_{i=1, \dots, m} \gamma^i$ can be made arbitrarily small. Denote

$$p_{\max} = \max_{i=1, \dots, m} p_i, \quad q_{\max} = \max_{i=1, \dots, m} (1 - p_i).$$

By [Assumption 1](#), $0 < p_{\max} < 1$ and $0 < q_{\max} < 1$. Then, we have $\|\Lambda^1\|^2 \leq \gamma_{\max}$. Consequently, subsystem observers can be designed to satisfy

$$E(\|\Lambda^1\|^2) \leq p_{\max} \gamma_{\max} = \gamma_1 < \varepsilon, \quad (19)$$

for any $\varepsilon > 0$.

On the other hand,

$$h(T) = \max_{i \in S} \|F e^{A(i)T} \Phi\|^2 = \max_{i \in S} \|F e^{A(i)T} (F' F)^{-1} F'\|^2$$

is a continuous function of T and for $T = 0$

$$\begin{aligned} h(0) &= \max_{i \in S} \|Fe^{A(i)0}(F'F)^{-1}F'\|^2 \\ &= \lambda_{\max}(F(F'F)^{-1}F'F(F'F)^{-1}F') \\ &= \lambda_{\max}(F(F'F)^{-1}F') \\ &= \lambda_{\max}((F'F)^{-1}F'F) \text{ by Jacobson's Lemma} \\ &= \lambda_{\max}(I_S) \\ &= 1, \end{aligned}$$

where $\lambda_{\max}$ is the largest eigenvalue. Since $q_{\max} < 1$, by continuity of $h(T)$, there exists $T_{\max} > 0$ such that

$$\begin{aligned} \gamma_2 &= \max_{T < T_{\max}} E(\|A^2 Fe^{A_k T} \Phi\|^2) \\ &\leq q_{\max} h(T_{\max}) < 1 - \varepsilon \end{aligned} \quad (20)$$

for some $\varepsilon > 0$.

Together, by selecting T first to satisfy $T < T_{\max}$ in (20), and then designing subsystem observers to satisfy (19), we obtain $\gamma \leq \gamma_1 + \gamma_2 < \varepsilon + 1 - \varepsilon = 1$. $\square$

4.3. Fundamental limitations

In Wang et al. (2023), without observation noises, we showed that for any given $T > 0$, L_i could be designed by using pole placement such that the observer for the entire state x converges strongly. This paper deals with MS observation errors and convergence. Intriguingly, the same design as in Wang et al. (2023) may not converge in MS. In fact, a fundamental limitation exists for MS convergence. We will use an example to demonstrate this property.

Example 2. Consider the system in Example 1. To illustrate this property, we note that by Theorem 3, for stability we need to choose feedback gains such that $pe^{2(a_1-L_1)T} + (1-p)e^{2a_1T} < 1$ and $pe^{2a_2T} + (1-p)e^{2(a_2-L_2)T} < 1$. Since $a_1 > 0$, for any given $0 < p < 1$, there exists $T_{\max}$ such that $(1-p)e^{2a_1T} > 1$ when $T > T_{\max}$. This results in the scenario

$$\min_{L_1}(pe^{2(a_1-L_1)T} + (1-p)e^{2a_1T}) > 1,$$

implying a divergent state observer, regardless of how L_1 is selected. It is similar for L_2 design.

Due to unobservable subsystems, there is always a time interval in which the observer must run open-loop. When this open-loop system is unstable, if T is too big, it will become apparent that the average value γ will always be greater than 1, resulting in an unstable observer, no matter how the observer gain is designed. This fundamental limitation indicates that T must be selected first to be small. Then L_i can be designed to achieve MS convergent observers.

For the system in Example 1, if $a_1 = 2$, $a_2 = 4$, $p = 0.2$, then the thresholds are $T_{\max}^1 = \ln(1/(1-p))/(2a_1) = 0.0558$ for the error dynamics of e_k^1 and $T_{\max}^2 = \ln(1/p)/(2a_2) = 0.2012$ for the error dynamics of e_k^2 . As a result, T must be selected to be smaller than $\min\{0.0558, 0.2012\} = 0.0558$.

Suppose that we select $T = 0.03 < 0.0558$ as in Example 1. Instead of using the Kalman-Bucy filter design, we select pole positions such that $pe^{2(a_1-L_1)T} + (1-p)e^{2a_1T} < 1$ and $(1-p)e^{2(a_2-L_2)T} + pe^{2a_2T} < 1$. This leads to the bounds on selection of L_1 and L_2 as

$$L_1 > 25.7775, \quad L_2 > 6.3407.$$

By choosing $L_1 = 40$ and $L_2 = 15$, we have $\gamma = \max\{0.9225, 0.6677\} = 0.9225 < 1$, resulting in a convergent observer design. This is demonstrated in Fig. 2, showing that both e_k^1 and e_k^2 are now convergent.

To demonstrate the critical importance of selecting T correctly first, we now select $T = 0.25 > 0.0558$. Under the same design of $L_1 = 40$ and $L_2 = 15$, Fig. 3 shows that observers for both e_k^1 and e_k^2 are now unstable.

5. Steady state error and design tradeoff

We now study the steady-state estimation error of the error dynamics (15), $e_{k+1} = \Lambda_k e_k + \Gamma_k d_k$, and discuss some design tradeoffs that are inherent between convergence speeds and steady-state error variances.

5.1. Steady-state error variance

Denote $V_k = E_{\alpha_k}(e_k e_k')$, $\mu_k = E_{\alpha_k}(e_k' e_k) = \text{Tr}(V_k)$. Select $T < T_{\max}$ and design L_i so that the matrix $M = E(\Lambda_k' \Lambda_k)$ is stable, namely all of its eigenvalues are inside the open unit circle. Write $M = P'P$ and denote $\Psi = E(\Gamma_k' \Gamma_k)$. Since Λ_k is induced by α_k , it is i.i.d. and M is a constant matrix. By Assumption 1, Λ_k , e_k , and d_k are mutually independent. Consequently, noting that Tr and E commute,

$$\begin{aligned} \mu_{k+1} &= \text{Tr}V_{k+1} \\ &= \text{Tr}(E(\Lambda_k e_k e_k' \Lambda_k')) + \text{Tr}(E(\Gamma_k d_k d_k' \Gamma_k')) \\ &= E(\text{Tr}(\Lambda_k e_k e_k' \Lambda_k')) + E(\text{Tr}(\Gamma_k d_k d_k' \Gamma_k')) \\ &= \text{Tr}(E(\Lambda_k' \Lambda_k e_k e_k')) + \text{Tr}(E(\Gamma_k' \Gamma_k d_k d_k')) \\ &= \text{Tr}(MV_k + \Psi_k) \\ &= \text{Tr}(P'V_k P + \Psi_k). \end{aligned}$$

Let $\Psi = \lim_{k \rightarrow \infty} \Psi_k$. The steady-state variance is V_{∞} , satisfying $\mu_{\infty} = \text{Tr}V_{\infty} = \text{Tr}(P'V_{\infty}P + \Psi)$. As a result, μ_{∞} can be obtained by first solving the Lyapunov equation

$$P'V_{\infty}P - V_{\infty} + \Psi = 0 \quad (21)$$

and then compute $\mu_{\infty} = \text{Tr}V_{\infty}$.

There is an inherent tradeoff in designing L_i . Intuitively, from the expression of Γ_k , it is clear that the larger the feedback gains, the larger the steady-state error μ_{∞} . On the other hand, to increase convergence speed, it is desirable to place poles with larger absolute values of their negative real parts. In other words, fast convergence and small steady-state errors are in conflict. We use an example to explain the tradeoff and discuss some potential approaches for choosing suitable design parameters.

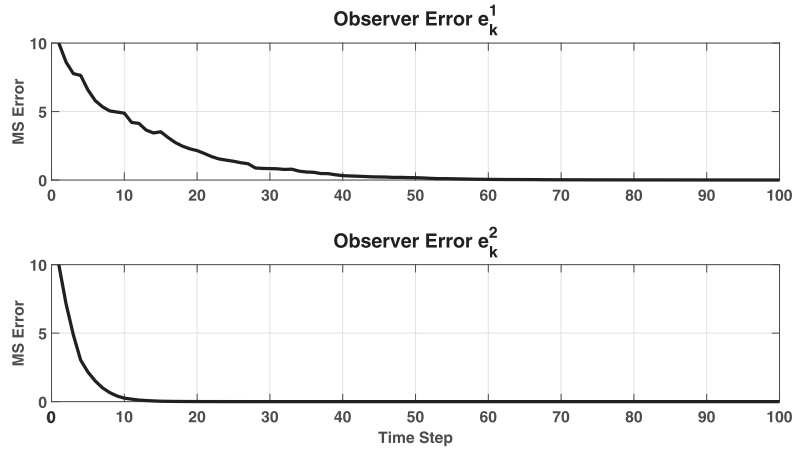
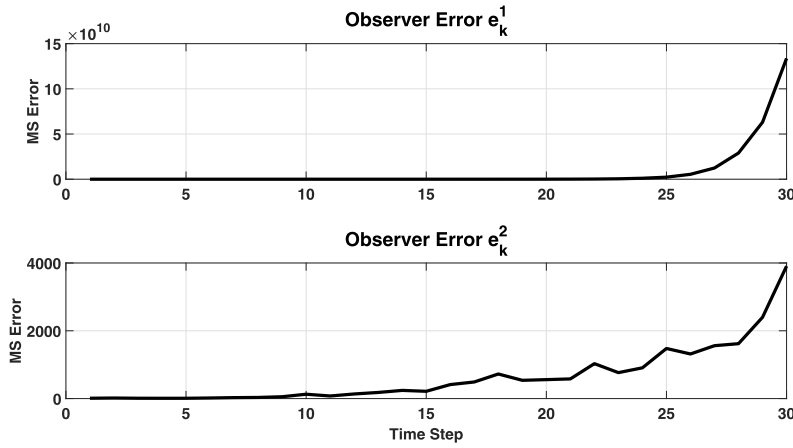
5.2. Discussion on design tradeoff

We now use the system in Example 1 to illustrate a tradeoff in filter design.

Example 3. Let the parameters in Example 1 be $c_1 = c_2 = 1$, $a_1 = a_2 = 10$, $\Pr\{\alpha_k = 1\} = 0.5$, $\zeta = 1$. When $\alpha_k = 1$, we observe x_1 only, and $a_k^1 = a_1 - L_1$. For stability, it is required that $0.5e^{2a_1T} + 0.5e^{2(a_1-L_1)T} = 0.5e^{20T} + 0.5e^{20T}e^{-2L_1T} < 1$, which implies $T < \frac{\ln 2}{20}$. Under this condition, L_1 must satisfy $e^{-2L_1T} < \frac{1-0.5e^{20T}}{0.5e^{20T}} = 2e^{-20T} - 1$, or $L_1 > -\frac{1}{2T} \ln(2e^{-20T} - 1)$.

From the viewpoint of convergence speed, the larger the gain L_1 is, the smaller the factor $\bar{\gamma}_1$ is, and in turn the faster the convergence of e_k^1 . On the other hand, from (15), a larger L_1 contributes to a larger equivalent variance on the diffusion term, which leads to a larger steady-state variance of the state estimation error.

Because of the constraints on T and L_1 , the transient convergence speed and the steady-state error variance form a fundamental tradeoff in the design consideration.

Fig. 2. Observer error dynamics under small T and new observer gain design.Fig. 3. Observer error dynamics under large T .

This tradeoff indicates that some optimization may be used for parameter selection. While a comprehensive treatment of this problem is beyond the scope of this paper, we now use an example to illustrate the main aspects of parameter selection.

Example 4. Consider the system in [Example 1](#), with $a_1 = 2$, $a_2 = 4$, $T = 0.03$, $p = 0.2$, $\zeta = 0.01$.

State Estimation Error Dynamics: The state estimation error dynamics are

$$\begin{aligned} e_{k+1}^1 &= \gamma_k^1 e_k^1 + \mathbb{1}_{\{\alpha_k=1\}} Q_k^1 d_k \\ e_{k+1}^2 &= \gamma_k^2 e_k^2 + \mathbb{1}_{\{\alpha_k=2\}} Q_k^2 d_k. \end{aligned} \quad (22)$$

where

$$\begin{aligned} \gamma_k^1 &= e^{(a_1-L_1)T} \mathbb{1}_{\{\alpha_k=1\}} + e^{a_1 T} \mathbb{1}_{\{\alpha_k=2\}}, \\ \gamma_k^2 &= e^{(a_2-L_2)T} \mathbb{1}_{\{\alpha_k=2\}} + e^{a_2 T} \mathbb{1}_{\{\alpha_k=1\}}, \end{aligned}$$

and by (15) for $i = 1, 2$, the limiting values of Q_k^1 and Q_k^2 (only these values are used in steady-state error calculations) are

$$\begin{aligned} Q_i &= \left(\int_0^T e^{2(a_i-L_i)\tau} L_i^2 \zeta d\tau \right)^{1/2} \\ &= \frac{L_i \sqrt{\zeta}}{\sqrt{2(L_i - a_i)}} (1 - e^{-2(L_i - a_i)T}). \end{aligned}$$

To analyze the dynamics of error variances, define $\sigma_k^1 = E((e_k^1)^2)$, $\sigma_k^2 = E((e_k^2)^2)$, and $\sigma_k = \sigma_k^1 + \sigma_k^2$. Under $\Pr\{\alpha_k = 1\} = p$

with $0 < p < 1$,

$$\begin{aligned} \sigma_{k+1}^1 &= \gamma^1 \sigma_k^1 + p Q_1^2 \\ \sigma_{k+1}^2 &= \gamma^2 \sigma_k^2 + (1-p) Q_2^2 \end{aligned} \quad (23)$$

with

$$\begin{aligned} \gamma^1 &= E((\gamma_k^1)^2) = p e^{2(a_1-L_1)T} + (1-p) e^{2a_1 T}, \\ \gamma^2 &= E((\gamma_k^2)^2) = (1-p) e^{2(a_2-L_1)T} + p e^{2a_2 T}. \end{aligned}$$

T is selected such that

$$\max\{(1-p) e^{2a_1 T}, p e^{2a_2 T}\} < 1. \quad (24)$$

For an illustration, we consider only σ_k^1 . To show the impact of L_1 on convergence rates and steady-state error variances, we select three different values in the permissible range $L_1 > 25.7775$, see [Example 2](#). For $L_1 = 30$, $\gamma^1 = 0.9883$; for $L_1 = 60$, $\gamma^1 = 0.9371$; for $L_1 = 100$, $\gamma^1 = 0.9126$. [Fig. 4](#) shows the error variance trajectories under these three different L_1 values. When L_1 is small ($L_1 = 30$), the error trajectories converge slowly to a relative high value. After increasing L_1 to 60, we see a faster convergence and smaller steady state variance. However, if L_1 is further increased to 200, the steady-state variance becomes bigger. These curves indicate that there is potentially an optimal L_1 value.

As a demonstration, for $L_1 = 60$, one sample-path error trajectory for e_k^1 from (22) is shown in [Fig. 5](#).

Selection of Design Parameters: For stability, L_1 and L_2 must be designed to satisfy the condition

$$\max\{\gamma^1, \gamma^2\} < 1. \quad (25)$$

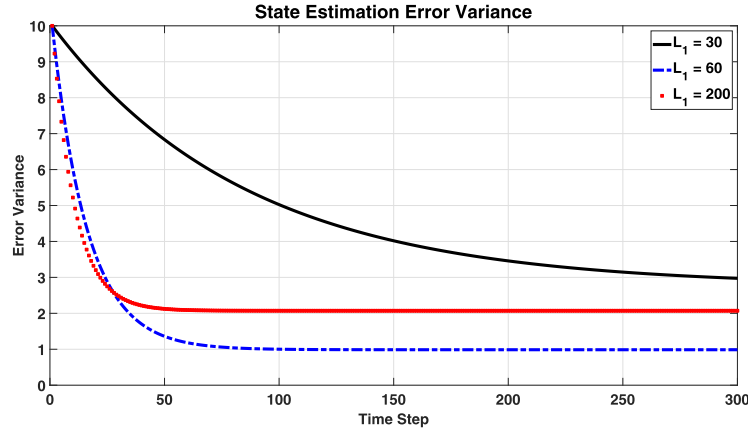


Fig. 4. Observer error variances under different L_1 for design tradeoff.

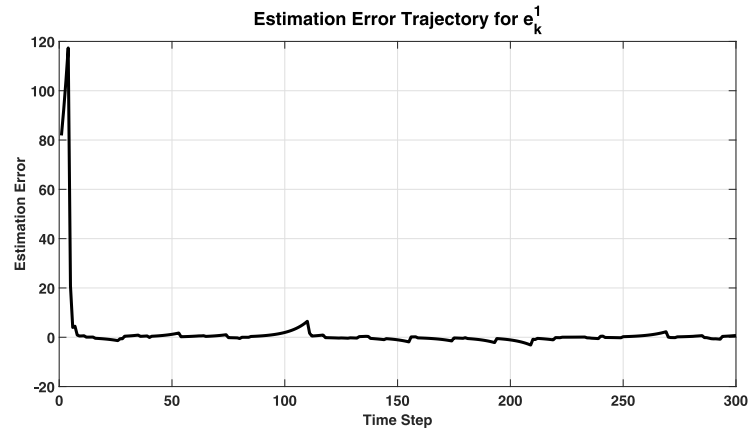


Fig. 5. Estimation error trajectories for e_k^1 under $L_1 = 60$.

Then, $\sigma_\infty^1(L_1, T) = \frac{pQ_1^2}{1-\gamma^1}$, $\sigma_\infty^2(L_2, T) = \frac{(1-p)Q_2^2}{1-\gamma^2}$, and the total steady-state MS error is

$$\sigma_\infty(L_1, L_2, T) = \sigma_\infty^1(L_1, T) + \sigma_\infty^2(L_2, T).$$

Optimal design of T , L_1 , L_2 is a constrained optimization problem

$$\min_{L_1, L_2, T} \sigma_\infty^1(L_1, T) + \sigma_\infty^2(L_2, T),$$

$$\text{Subject to: } (1-p)e^{2a_1T} < 1, pe^{2a_2T} < 1, \gamma^1 < 1, \gamma^2 < 1.$$

For a fixed T that satisfies (24), both $\sigma_\infty^1(L_1, T)$ and $\sigma_\infty^2(L_2, T)$ are strictly convex functions of L_1 and L_2 , respectively, and can be individually optimized with the unique solutions $L_1^*(T)$ and $L_2^*(T)$. Then T is optimized under the constraint (24). Since exponential functions are involved in this optimization problem, in general it is a highly nonlinear problem that can be solved numerically. For $T = 0.03$, Fig. 6 shows $\sigma_\infty^1(L_1, T)$ as a function of L_1 . The optimal value is $L_1^* = 61.3$ with the corresponding minimum variance $\sigma_\infty^1(L_1^*, T) = 0.9864$.

In conclusion, in practical applications of the design method introduced in this paper, design of the observer feedback gains should be carefully evaluated to achieve a balance between convergence speeds and steady-state errors.

6. A case study

In this section, we use a practical case of power systems to illustrate the design process, performance evaluation, and related issues. We use a common IEEE 5-Bus system that has been

widely used in performing power system optimization, power flow analysis, and contingency detection.

Example 5. Consider the IEEE 5-Bus system shown in Fig. 7. The bus structure and data are from the open-source information in Tan (2023). Bus 1 and Bus 2 are generator buses and Buses 3–5 are load buses.

Buses 1 and 2 are PV buses (generator buses) with bus voltage magnitudes controlled to its rated values by Var condensers or compensators.² Similarly, in light of rapid advancement in Var compensation technology such as flexible AC transmission systems (FACTS), we assume that all load buses have their voltage magnitudes maintained near the rated values.

Dynamic Systems. The two generator buses are dynamic buses with synchronous generators (Kundur, 1994). Denote $\omega_1 = \delta_1$, $z_1 = [\delta_1, \omega_1]'$, $\omega_2 = \delta_2$, $z = [\delta_2, \omega_2]'$. The dynamic systems are

$$M_1 \dot{\omega}_1 + g_1(\omega_1) = P_{in}^1 - P_L^1 + P_{21}^1 + P_{31}^1$$

$$M_2 \dot{\omega}_2 + g_2(\omega_2) = P_{in}^2 - P_L^2 + P_{12}^2 + P_{32}^2 + P_{42}^2 + P_{52}^2$$

where the real power flow from Bus i into Bus j on Bus j is

$$P_{ij}^j = \frac{V_j^2}{|Z_{ij}|} \cos(\theta_{ij}) - \frac{V_i V_j}{X_{ij}} \cos(\theta_{ij} + \delta_{ij}), \quad (26)$$

² In power systems, a PQ bus has its real power (P) and reactive power (Q) specified, often by loads; a PV bus has its real power (P) and the voltage magnitude (V) specified, often for generators; and a slack bus has its voltage phasor specified (as a reference point) and its real and reactive powers are used to balance powers in a grid.

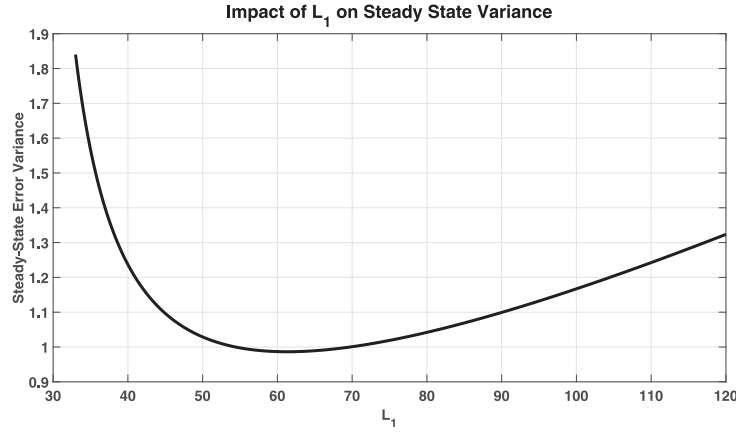
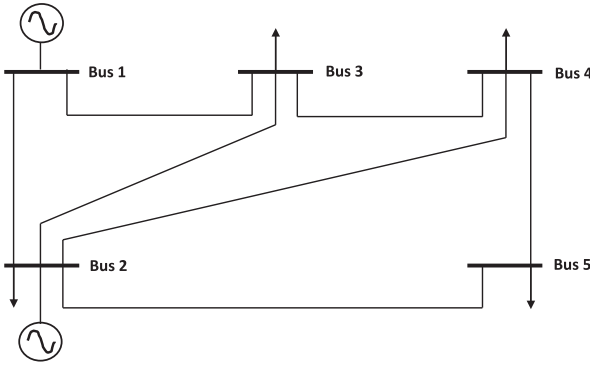
Fig. 6. Impact of L_1 on $\sigma_\infty^1(L_1, T)$ when $T = 0.03$.

Fig. 7. IEEE 5-Bus system.

and $\delta_{ij} = \delta_i - \delta_j$. The damping term $g_i(w_i)$ has the linear part

$$b_i \omega_1 \text{ with } b_i > 0, i = 1, 2. \text{ Denote } \xi_d = \begin{bmatrix} \delta_1 \\ \delta_2 \end{bmatrix}, \xi_{nd} = \begin{bmatrix} \delta_3 \\ \delta_4 \\ \delta_5 \end{bmatrix},$$

and their perturbations from the nominal values as $x_d = \Delta \xi_d$, $x_{nd} = \Delta \xi_{nd}$. The dynamic systems can be expressed as a nonlinear state equation

$$\dot{\xi}_d = F(\xi_d, \xi_{nd}) + B_1 v + B_2 e$$

$$\text{where } v = \begin{bmatrix} p_{in}^1 \\ p_{in}^2 \end{bmatrix}, e = \begin{bmatrix} p_L^3 \\ p_L^4 \\ p_L^5 \end{bmatrix}, B_1 = \begin{bmatrix} 0 & 0 \\ 1/M_1 & 0 \\ 0 & 0 \\ 0 & 1/M_2 \end{bmatrix}, B_2 =$$

$$\begin{bmatrix} 0 & 0 \\ -1/M_1 & 0 \\ 0 & 0 \\ 0 & -1/M_2 \end{bmatrix}. \text{ The dynamic systems can be linearized near}$$

the nominal operating points as $\dot{x}_d = A_1 x_d + A_2 x_{nd}$.

The three load buses have real-power equations $p_L^3 = p_{13}^3 + p_{23}^3 + p_{43}^3$, $p_L^4 = p_{24}^4 + p_{34}^4 + p_{54}^4$, $p_L^5 = p_{25}^5 + p_{45}^5$, where p_{ij}^j is given in (26). These nonlinear equations can be linearized near the nominal operating points as $0 = H_1 x_{nd} + H_2 x_d$ where H_1 and H_2 are the corresponding Jacobian matrices. In power systems under permitted operating ranges, H_1 is invertible, leading to $x_{nd} = H x_d$ with $H = -H_1^{-1} H_2$. It follows that

$$\dot{x}_d = (A_1 + A_2 H) x_d = A x_d$$

with $A = A_1 + A_2 H$. The bus line parameters are defined in Tan (2023) and listed in Table 1, with R = Resistance, X = Reactance, $Z = |Z| \angle \theta$ = Impedance.

Table 1

IEEE 5-Bus system line parameters.

Line	R (p.u.)	X (p.u.)	Z (p.u. $ Z \angle \theta$ rad)
1–2	0.02	0.06	$0.06 \angle 1.25$
1–3	0.08	0.24	$0.25 \angle 1.25$
2–3	0.06	0.25	$0.26 \angle 1.33$
2–4	0.06	0.18	$0.19 \angle 1.25$
2–5	0.04	0.12	$0.13 \angle 1.25$
3–4	0.01	0.03	$0.03 \angle 1.25$
4–5	0.08	0.24	$0.25 \angle 1.25$

Table 2

IEEE 5-Bus system bus data.

Bus	V (p.u. $\angle$ rad)	P	Q	P_L	Q_L
1	$1.06 \angle 0$	129	-7.42	0	0
2	$1.0474 \angle -2.8063$	40	30	20	10
3	$1.0242 \angle -4.997$	0	0	45	15
4	$1.0236 \angle -5.3291$	0	0	40	5
5	$1.0179 \angle -6.1503$	0	0	60	10

The nominal operating condition defined in Bhandakkar and Mathew (2018) and Tan (2023) is used here with the nominal bus voltages, generation powers and load powers listed in Table 2 with real power P (MW) and reactive power Q (MVar). The base MVA is $S_B = 100$ MVA and base voltage is $V_B = 230$ kV.

Under the per unit system, the normalized generator parameters are $M_1 = 1.9$ and $b_1 = 0.2$ with equivalent time constant $T_1 = M_1/b_1 = 9.5$ s for Generator 1, and $M_2 = 0.9$, $b_1 = 0.16$ with equivalent time constant $T_2 = M_2/b_2 = 5.625$ s for Generator 2, see Bhandakkar and Mathew (2018) for the computational results.

Under these operating conditions, we obtain

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 7.7926 & -0.1053 & -7.7926 & 0 \\ 0 & 0 & 0 & 1 \\ -20.3866 & 0 & 20.3866 & -0.1778 \end{bmatrix}.$$

Since eigenvalues of A are $\{-5.3880, 5.2302, 0, -0.1253\}$, it is an unstable system.

Sensor Systems. In energy management systems (EMS), frequencies and voltage phasors on some buses are measured by frequency sensors and synchrophasor measurement units (PMUs). Each sensor represents an output equation. For cost reduction and maintenance simplification, it is highly desirable to reduce sensor complexity. Table 3 lists some common sensor systems and their observability under normal and contingency conditions. S_i = the i th sensor, $i = 1, 2$. Yes = Observable; No = Unobservable.

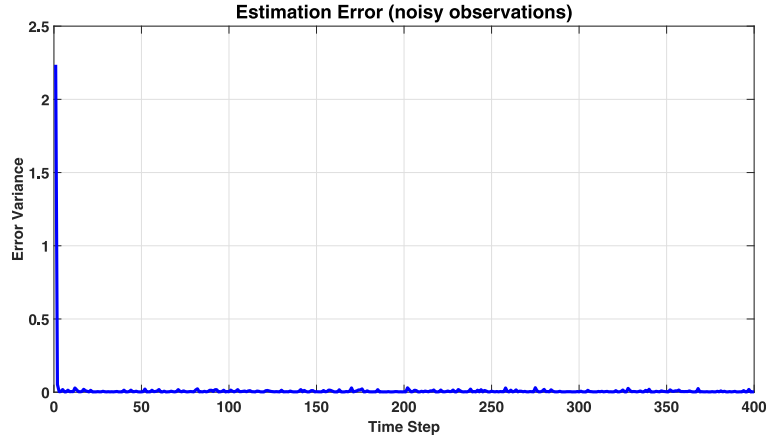


Fig. 8. State estimation error variance trajectories under noisy observations.

Table 3

Sensor systems and observability.

S_1	S_2	Normal	S_1 fails	S_2 fails	Both fail
δ_1	δ_2	Yes	Yes	Yes	No
δ_1	ω_1	Yes	No	Yes	No
ω_1	ω_2	No	No	No	No
δ_1	ω_2	Yes	No	Yes	No

Since the sensor system with (δ_1, δ_2) has better observability properties, we will use this as an example here.

Contingency and Stochastic Hybrid Systems. In modern microgrids, the measured signals are communicated to the supervisory control and data acquisition (SCADA) and EMS via wireless communication networks. Due to communication packet loss, the packet delivery ratio for each channel is given by $\xi_j > 0$ for successful data transfer on channel j .

Suppose that δ_1 and δ_2 are measured. By adding measurement and communication noises in PMU measurement errors (Castello, Muscas, & Pegoraro, 2022; Salls, Torres, Varghese, Patterson, & Pal, 2021), we have $C(1) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$ for normal operation,

$C(2) = [1, 0, 0, 0]$ for failure of Sensor 2 (δ_1 measurement only), $C(3) = [0, 0, 1, 0]$ for failure of Sensor 1 (δ_2 measurement only), and $C(4) = [0, 0, 0, 0]$ for failure on both sensors. Suppose that the packet delivery ratio for Sensor 1 is $\xi_1 = 0.99$ and for Sensor 2 is $\xi_2 = 0.995$. This data acquisition scheme can be modeled by an i.i.d. stochastic process $\alpha_k \in \mathcal{S} = \{1, \dots, 4\}$ with $p_1 = \xi_1 \xi_2 = 0.985$, $p_2 = \xi_1(1 - \xi_2) = 0.005$, $p_3 = (1 - \xi_1)\xi_2 = 0.01$, $p_4 = (1 - \xi_1)(1 - \xi_2) = 5 \cdot 10^{-5}$.

Observer Design. The pole placement design is used for designing observer feedback gains for $\alpha_k = 1$ and $\alpha_k = 2$. Since $C(4) = [0, 0, 0, 0]$, the observer can only run open-loop. For example, if we choose the desired closed-loop poles to be $\lambda = [-6, -4.5, -5 - 5.5]$, then the Matlab function $L_i = \text{place}(A', C'(i), \lambda)$, $i = 1, 2, 3, 4$, yields the suitable feedback gains and the closed-loop error dynamics with $A_c^i = A - L_i C(i)$, $i = 1, 2, 3, 4$.

For this example, the upper bound on T can be computed as $T_{\max} = 0.7365$, which is a sufficient condition for feasible observer design. Select $T = 0.6997$. Denote $\Lambda(i) = e^{A_c^i T}$, $i = 1, 2, 3, 4$. The closed-loop error dynamics of e_k are a stochastic system with noise-free dynamics $e_{k+1} = \Lambda_k e_k$, where $\Lambda_k = \sum_{j=1}^4 \mathbb{1}_{\{\alpha_k=j\}} \Lambda(i)$. For MS stability of the combined observer, we

verify that the eigenvalues of $p_1(\Lambda(1))' \Lambda(1) + p_2(\Lambda(2))' \Lambda(2) + p_3(\Lambda(3))' \Lambda(3) + p_4(\Lambda(4))' \Lambda(4)$ are $\{0.9904, 0.54, 0.0032, 0.0356\}$ that are stable.

The initial estimation error is selected to be $e(0) = [2, 0, 1, 0]'$ with error norm 2.2361. Suppose that the standard deviations are $\mathcal{E}(1) = 0.01 \cdot [1, 1]'$; $\mathcal{E}(2) = 0.011$; $\mathcal{E}(3) = 0.012$. To simulate the MS errors, the simulations for error trajectories are repeated 400 times, and the sample MS value at each k is computed as an approximation of $E(e_k' e_k)$. Fig. 8 shows estimation error variance trajectories.

Design Tradeoff. We now illustrate the design tradeoff discussed in Section 5. Suppose that we now select the designed pole positions to be more negative at $[-12, -9, -10, -11]$ which represents a more aggressive observer design. It can be computed that the eigenvalues of $p_1(\Lambda(1))' \Lambda(1) + p_2(\Lambda(2))' \Lambda(2) + p_3(\Lambda(3))' \Lambda(3) + p_4(\Lambda(4))' \Lambda(4)$ become $\{0.6891, 0.0046, 0.0015, 0.0001\}$. Since the eigenvalues are smaller than before, it converges faster. However, the steady-state errors become bigger due to much larger feedback gains. For example, L_1 is increased from $L_1 = [20.7170, 187.0468, -123.7698, -568.3264]'$ to $L_1 = [41.7, 675.4, -681.6, -3181.7]'$. Fig. 9 shows estimation error variance trajectories, showing larger persistent errors.

7. Conclusions

This paper has shown that continuous state estimation of RSLs with unobservable subsystems and observation noises must jointly consider features of continuous subsystems and stochastic switching sequences in order to achieve MS convergence of the observers. Due to a tradeoff between convergence speeds and steady-state estimation errors, a constrained optimization problem may be used to minimize estimation errors.

There are some open issues and potential extensions for the problems treated in this paper. More sophisticated switching processes such as Markov chains can be studied to accommodate broader classes of RSLs. Modeling errors that are of practical importance are not considered in this paper. Applications of the methodologies and algorithms to emerging technology areas, such as autonomous systems, modern power systems, cyber-physical systems will be highly valuable.

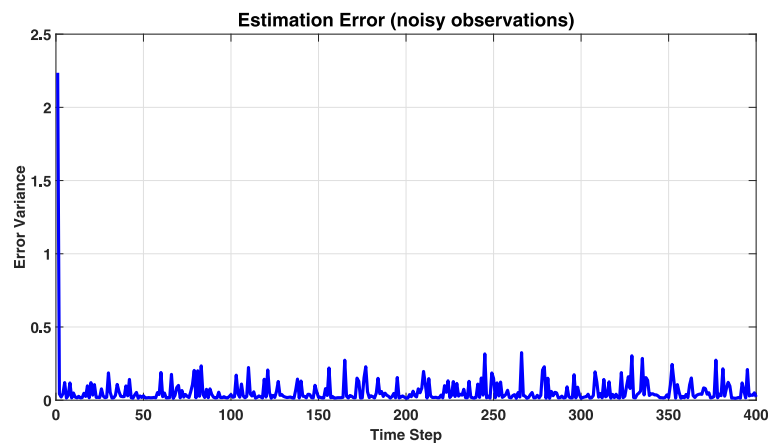


Fig. 9. State estimation error variance trajectories under a more aggressive observer design.

References

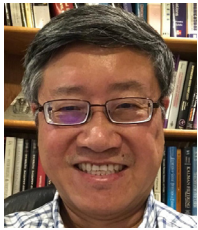
- Anderson, B. D. O., & Moore, J. B. (1979). *Optimal filtering*. New Jersey: Prentice-Hall, Inc..
- Arbib, C., & De Santis, E. (2020). Almost always observable hybrid systems. *Nonlinear Analysis. Hybrid Systems*, 36.
- Babaali, M., & Egerstedt, M. (2004). Observability of switched linear systems. In *International workshop on hybrid systems: computation and control* (pp. 48–63). Springer.
- Babaali, M., & Pappas, G. J. (2005). Observability of switched linear systems in continuous time. In *International workshop on hybrid systems: computation and control* (pp. 103–117). Springer.
- Bernard, P., & Sanfelice, R. G. (2020). On notions of detectability and observers for hybrid systems. In *2020 59th IEEE conference on decision and control* (pp. 5767–5772). IEEE.
- Bhandakkar, A. A., & Mathew, L. (2018). Real-time-simulation of IEEE-5-bus network on OPAL-RT-OP4510 simulator. *IOP Conference Series: Materials Science and Engineering*, 331(1), Article 012028.
- Cassandras, C. G., & Lygeros, J. (2018). *Stochastic hybrid systems*. CRC Press.
- Castello, P., Muscas, C., & Pegoraro, P. A. (2022). Statistical behavior of PMU measurement errors: An experimental characterization. *IEEE Open Journal of Instrumentation and Measurement*, 1, <http://dx.doi.org/10.1109/OJIM.2022.3205678>.
- Dragan, V., & Aberkane, S. (2020). Exact detectability and exact observability of discrete-time linear stochastic systems with periodic coefficients. *Automatica*, 112, <http://dx.doi.org/10.1016/j.automatica.2019.108660>.
- Ezzine, J., & Haddad, A. H. (1989). Controllability and observability of hybrid systems. *International Journal of Control*, 49(6), 2045–2055.
- Farina, F., Garulli, A., & Giannitrapani, A. (2022). An interpolatory algorithm for distributed set membership estimation in asynchronous networks. *IEEE Transactions on Automatic Control*, 67(10), 5464–5470.
- Fliess, M., Join, C., & Perruquetti, W. (2008). Real-time estimation for switched linear systems. In *2008 47th IEEE conference on decision and control* (pp. 941–946). IEEE.
- Goebel, R., Sanfelice, R. G., & Teel, A. R. (2012). *Hybrid dynamical systems*. Princeton University Press.
- Gomez-Gutierrez, D., Ramirez-Trevino, A., Ruiz-Leon, J., & Di Gennaro, S. (2010). Observability of switched linear systems: A geometric approach. In *49th IEEE conf. decis. control* (pp. 5636–5642). IEEE.
- Haddadi, K., Gazzam, N., & Benalia, A. (2019). Observability of switched linear systems based geometrical conditions: An application to dc/dc multilevel converters. In *2019 international conference on advanced electrical engineering* (pp. 1–5). IEEE.
- Johnson, S. C. (2016). *Observability and observer design for switched linear systems* (Ph.D. thesis), Purdue University.
- Kailath, T. (1980). *Linear systems*, Vol. 156. Englewood Cliffs, NJ: Prentice-Hall.
- Kalman, R. E. (1963). New methods in Wiener filtering theory. In J. L. Bogdanoff, & F. Kozin (Eds.), *Proceedings of the first symposium on engineering applications of random function theory and probability* (pp. 270–388). New York: John Wiley & Sons.
- Kundur, P. S. (1994). *Power system stability and control* (1st ed.). McGraw Hill.
- Kuo, B. C., & Golnaraghi, F. (1994). *Automatic control systems* (7th ed.). Prentice Hall.
- Kusters, F., & Trenn, S. (2018). Switch observability for switched linear systems. *Automatica*, 87, 121–127.
- Li, T., & Zhang, J.-F. (2010). Consensus conditions of multi-agent systems with time-varying topologies and stochastic communication noises. *IEEE Transactions on Automatic Control*, 55(9), 2043–2057.
- Lunze, J., & Lamnabhi-Lagarrigue, F. (2009). *Handbook of hybrid systems control: theory, tools, applications*. Cambridge University Press.
- Lygeros, J., & Prandini, M. (2010). Stochastic hybrid systems: a powerful framework for complex, large scale applications. *European Journal of Control*, 16(6), 583–594.
- Nerode, A., & Kohn, W. (1992). Models for hybrid systems: Automata, topologies, controllability, observability. In *Hybrid systems* (pp. 317–356). Springer.
- Ríos, H., Davila, J., & Teel, A. R. (2019). Linear hybrid systems with periodic jumps: a notion of strong observability and strong detectability. *IEEE Transactions on Automatic Control*, 65(6), 2640–2646.
- Ríos, H., Dávila, J., & Teel, A. R. (2020). State estimation for linear hybrid systems with periodic jumps and unknown inputs. *International Journal of Robust and Nonlinear Control*, 30(15), 5966–5988.
- Salls, D., Torres, J. R., Varghese, A. C., Patterson, J., & Pal, A. (2021). Statistical characterization of random errors present in synchrophasor measurements. In *Proceedings of 2021 IEEE-PESGM conference*.
- Sanfelice, R. G., Goebel, R., & Teel, A. R. (2007). Invariance principles for hybrid systems with connections to detectability and asymptotic stability. *IEEE Transactions on Automatic Control*, 52(12), 2282–2297.
- Sellami, L., & Abderrahim, K. (2015). State estimation for switched linear systems. In *2015 IEEE 12th international multi-conference on systems, signals & devices (SSD15)* (pp. 1–5). IEEE.
- Sun, Z., & Ge, S. S. (2005). Analysis and synthesis of switched linear control systems. *Automatica*, 41(2), 181–195.
- Tan, R. (2023). IEEE 5-bus system model. In *MATLAB central file exchange*.
- Vidal, R., Chiasso, A., Soatto, S., & Sastry, S. (2003). Observability of linear hybrid systems. In *International workshop on hybrid systems: computation and control* (pp. 526–539). Springer.
- Wang, L. Y., Khargonekar, P. P., & Beydoun, A. (1997). Robust control of hybrid systems: performance guided strategies. In *International hybrid systems workshop* (pp. 356–390). Springer.
- Wang, L. Y., Yin, G., Lin, F., Polis, M. P., & Chen, W. (2022). Joint estimation of continuous and discrete states in randomly switched linear systems with unobservable subsystems. *IEEE Transactions on Automatic Control*, 9, <http://dx.doi.org/10.1109/TAC.2022.3233289>.
- Wang, L. Y., Yin, G., Lin, F., Polis, M. P., & Chen, W. (2023). Stochastic observability and convergent analog state estimation of randomly switched linear systems with unobservable subsystems. *IEEE Transactions on Automatic Control*, 68, 898–911.
- Yin, G., Sun, Y., & Wang, L. Y. (2011). Asymptotic properties of consensus-type algorithms for networked systems with regime-switching topologies. *Automatica*, 47(7), 1366–1378.

Yin, G., Wang, L. Y., & Sun, Y. (2011). Stochastic recursive algorithms for networked systems with delay and random switching: Multiscale formulations and asymptotic properties. *SIAM Journal of Multiscale Modeling & Simulation*, 9(3), 1087–1112.

Yin, G., & Zhu, C. (2010). *Hybrid switching diffusions*. Springer.

Zarchan, P., & Musof, H. (2005). *Fundamentals of Kalman filtering: a practical approach*. AIAA.

Zhao, X., Liu, H., Zhang, J., & Li, H. (2013). Multiple-mode observer design for a class of switched linear systems. *IEEE Transactions on Automation Science and Engineering*, 12(1), 272–280.



Le Yi Wang received the Ph.D. degree in electrical engineering from McGill University, Montreal, Canada, in 1990. Since 1990, he has been with Wayne State University, Detroit, Michigan, where he is currently a Professor in the Department of Electrical and Computer Engineering. His research interests are in the areas of complexity and information, system identification, robust control, information processing and learning, as well as medical, automotive, communications, power systems, and computer applications of control methodologies. He was a keynote speaker in several international conferences. He serves on the IFAC Technical Committee on Modeling, Identification and Signal Processing. He was an Associate Editor of the *IEEE Transactions on Automatic Control* and several other journals. He is a Life Fellow of IEEE.



George Yin received the B.S. degree in mathematics from the University of Delaware in 1983, and the M.S. degree in electrical engineering and the Ph.D. degree in applied mathematics from Brown University in 1987. He joined Wayne State University in 1987, became Professor in 1996, and University Distinguished Professor in 2017. He moved to the University of Connecticut in 2020. His research interests include stochastic processes, stochastic systems theory, and applications. He served as Co-chair for a number of conferences, was on the Board of Directors of the American Automatic Control Council, and was Chair of the SIAM Activity Group on Control and Systems Theory. He is Editor-in-Chief of *SIAM Journal on Control and Optimization*. He was an Associate Editor of *Automatica* 1995–2011, *IEEE Transactions on Automatic Control* 1994–1998 and Senior Editor of *IEEE Control Systems Letters* 2017–2019. He is a Life Fellow of IEEE, a Fellow of IFAC, and a Fellow of SIAM.



Feng Lin received his B.Eng. degree in electrical engineering from Shanghai Jiao Tong University, Shanghai, China, in 1982, and the M.A.Sc. and Ph.D. degrees in electrical engineering from the University of Toronto, Toronto, ON, Canada, in 1984 and 1988, respectively. He was a Post-Doctoral Fellow with Harvard University, Cambridge, MA, USA, from 1987 to 1988. Since 1988, he has been with the Department of Electrical and Computer Engineering, Wayne State University, Detroit, MI, USA, where he is currently a Professor. His current research interests include discrete event systems, hybrid systems, robust control, and their applications in alternative energy, biomedical systems, and automotive control. He authored a book entitled “Robust Control Design: An Optimal Control Approach” and coauthored a paper that received a George Axelby outstanding paper award from the IEEE Control Systems Society. He was an associate editor of *IEEE Transactions on Automatic Control*. He is a Fellow of IEEE.



Michael P. Polis received the B.S. degree in electrical engineering from the University of Florida, and the M.S.E.E. and Ph.D. degrees from Purdue University. From 1972–1983 he was a faculty member in Electrical Engineering, at Ecole Polytechnique de Montreal. From 1983–1987 he directed the Systems Theory and Operations Research Program at NSF. In 1987 he joined Wayne State University as Chair of Electrical and Computer Engineering. From 1993–2001 he was Dean of Engineering and Computer Science, Oakland University, Rochester, MI, where he is currently with the Department of Industrial and Systems Engineering. He has been a consultant for several companies, and an Associate Editor of the *IEEE Transactions on Automatic Control*. His research interests are in energy and transportation systems, and the identification, and control of distributed parameter systems. He co-authored the paper named the “Best Paper IEEE Transactions on Automatic Control 1974–1975”. He is a Life Senior Member of IEEE.



Wen Chen received his Ph.D. degree from Simon Fraser University, Burnaby, BC, Canada in 2004. From 2005 through 2007, he pursued his postdoctoral research at University of Louisiana at Lafayette, USA. His research interest was in the areas of Control and Fault Diagnosis. From Aug. 2007 to Dec. 2008, he was a Control Systems Engineer at Paton Controls and Triconex, Houston, Texas, working on design of control and fault-diagnostic systems. In 2009, he joined Division of Engineering Technology at Wayne State University as an assistant professor. Currently, he is a professor and his research is focused on monitoring and diagnostics of industrial Systems. He is a Senior Member of IEEE.