WAYNE STATE UNIVERSITY

Stochastic Observability and Convergent Analog State Estimation of Randomly Switched Linear Systems

Le Yi Wang
Department of Electrical and Computer Engineering
Wayne State University, USA

This is a joint work with George Yin, Feng Lin, Michael Polis, and Wen Chen

November 2021

Motivations and Focus

- General hybrid systems involve the integration and intimate interactions between the analog state processes and the discrete event systems.
- Switched linear systems are an important class of hybrid systems, such as automotive systems, autonomous vehicles, smart grids, battery networks, social networks, biological and medical systems, among many others.
- Randomly switched linear systems are very common., such as communication packet loss , machine breakdown in assembly lines, cyber attacks, power line interruption, robot malfunction in robot teams, obstacles in the line of sight of unmanned aerial vehicles, and intermittent power outputs in wind or solar generators.
- Randomly switched linear or nonlinear systems have been treated as stochastic hybrid systems, regime-switching systems, and hybrid switching diffusions

Randomly Switching Linear Systems

Consider a continuous-time single-input-single-output hybrid system

$$\dot{x}(t) = A(\alpha(t))\,x(t) + B(\alpha(t))\;u(t)$$

$$y(t) = C(\alpha(t))\;x(t)$$

$$u(t)\ \in\ R,\ x(t) \in R^n,\ y(t) \in R$$

The randomly switching process $\alpha(t)$ takes m possible values in $S = \{1, \ldots, m\}$

For given value $i \in S$, the corresponding LTI system with matrices $C(i),\ A(i),\ B(i)$ will be called the i-*th subsystem of the hybrid system.*

Assumption 1

Given a period τ,

(i) the switching process $\alpha(t)$ canswitch only at the sampling instants $\tau, k = 1, 2, \ldots,$ generating a stochastic sequence $\{\alpha_k = \alpha(k\tau)\}$ (skeleton sequence)

(ii) The sequence $\{\alpha_k\}$ isindependent and identically distributed (i.i.d.)

$$P(\alpha_k = i) = p_i > 0, i \in S, \sum_{i=1}^{m} p_i = 1.$$

For α_k, $k = 1, 2, \ldots, l,$ denote the stochastic matrices

$$A_k = A(\alpha_k) = \sum_{i=1}^{m} A(i)\ \mathrm{I}_{\{\alpha_k = i\}}, B_k = B(\alpha_k),\ C_k = C(\alpha_k)$$

Classical Observability on LTI Systems

For constant $A \in R^{n\times n}$ and $C \in R^{1\times n}$,and a finite time interval $[t_0, T)$,

$G: R^n \rightarrow C[t_0,T): y(t)=G\left(x\left(t_0\right)\right)\left(t\right)=Ce^{A(t-t_0)}x\left(t_0\right), t\in[t_0, T)$

and its kernel $Ker(G)=\{x(t_0)R^n : y\left(t\right)\equiv 0, t\in [t_0,T)$

The observability matrix

$$W = \begin{bmatrix} C \\ \vdots \\ CA^{n-1} \end{bmatrix}$$

Lemma 1

(1) Ker(G)=ker(W)

(2) The LTI system is observable if and only if W is full rank.

Objectives of this Work

(1) This paper deals with estimation of the analog states based on observations on y.

(2) It deals with hybrid systems whose subsystems are not observable.

(3) Subsystems must coordinate to achieve state estimation.

(4) We want to establish conditions and design methods to obtain convergent state observers.

Off-Line System Observability Properties

$$W(i) = \begin{bmatrix} C(i) \\ C(i)A(i) \\ \\ C(i)(A(i))^{n-1} \end{bmatrix} \{n-1\}, i = 1, \ldots, m.$$

The combined matrix for the set S of the subsystems is

$$W_S = \begin{bmatrix} W(1) \\ W(2) \\ \vdots \\ W(m) \end{bmatrix}$$

$W(i)$ and W_S are constant matrices,

there is no randomness involved.

The State Dynamics of the Hybrid System

The zero-inputcase $u(t) = 0$,

$$\dot{x}(t) = A(\alpha(t))x(t)$$

$$y(t) = C(\alpha(t))\, x(t)$$

Starting from $x(0)$, with a given switching sequence α for $k = 1, \ldots, l$,

$$y(t) = C_k e^{A_k(t-(k-1)\tau)} x((k-1)\tau),$$

for $t \in [(k-1)\tau,\ k\tau),\ k = 1\ldots,\ l$,

and $x(k\tau) = H_k\ x(0), k = 1\ldots,\ l$, with $H_k = e^{A_k\tau} \cdots e^{A_1\tau}, k = 1\ldots,\ l$

Stochastic Matrices in Operation

$$O_l = \begin{bmatrix} W(\alpha_1) \\ W(\alpha_2)\mathrm{H}_1 \\ \vdots \\ W(\alpha_l)\mathrm{H}_{l-1} \end{bmatrix}$$

Definition

For a given finite time interval $[0, l\,\tau)$ and switching sequence $\{\alpha_k$, k=1,…, l},

the hybrid system is said to be *stochastically observable*

if Ker($G_{\{\alpha_k\}}$)={0}.

Lemma 2

For a given finite time interval $[0, l\tau)$ and switching sequence $\{\alpha_k,\ k = 1,\dots,l\}$,

(1) $Ker(G) = \ker(O_l)$.

(2) The hybrid system is stochastically observable if and only if $\ker(O_l) = \{0\}$, or equivalently O_l is full column rank.

Definition

The hybrid system is said to be ***off-line collectively observable*** if W_S is full rank.

Remarks

- Since W_S contains only off-line information on subsystems, **off-line collective observability is independent of time**.
- It is easy to see that in most applications a hybrid system needs to be **off-line collectively observable** for possible **stochastic observability**.
- However, the relationship between the rank of $\boldsymbol{W_S}$ and stochastic observability is complicated.

Asymptotic Observability of
Randomly Fast Switching Linear Systems

Assumption 2

$\boldsymbol{W_S}$ is full column rank, namely the hybrid system is off-line collectively observable.

Lemma 3

Under Assumption 1, there exists $T_{\max} = \kappa\tau > 0$, where $\kappa \geq m$ is an integer, such that if there is a switching sequence in $[0, T_{\max})$ that includes the special sequence segment $\tilde{\alpha} = \{\alpha_k = 1\ldots, \alpha_{k+m-1} = m\}$ or any permutation of it, then with a positive probability, the stochastic hybrid system is stochastically observable in $[0, T_{\max})$.

Exponential Convergence in Probability

Let ρ_l be the probability of the event ``**the hybrid system remains stochastically unobservable after observing y in** *[0, $l\tau$)".*

Theorem 1

Under Assumptions 1 and 2, suppose that the decision interval $\tau = T_{\max} / m$ and $T_{\max}$ satisfies the condition of Lemma 3 so that the special string $\tilde{\alpha}$ or any of its permutations makes the hybrid system stochastically observable. Then,

$$\rho_l \to 0, \qquad l \to \infty, \text{ exponentially.}$$

Observer Design of
Randomly Switched Linear Systems

Subsystem Space Decomposition

Assume that each subsystem is not observable, $Rank(W(i)) = n_i < n$.

We construct $M_i = Base(\ker(W(i))) \in R^{n \times n - n_i}$

and select any $N_i \in R^{n \times n_i}$ 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 $G_i \in R^{n-n_i \times n}$ and $F_i \in R^{n_i \times n}$

Define $$F = \begin{bmatrix} F_1 \\ F_2 \\ \vdots \\ F_m \end{bmatrix} \in R^{n_s \times n}, \quad n_s = \sum_{i=1}^{m} n_i$$

Lemma 4

$$(i) ker(F_i) = ker(W(i)).$$

$$(ii)\ ker\left(F\right) = ker(W_S)$$

Observable Sub-State Dynamics

The state transformation $z_i = T_i^{-1}x$ can be decomposed into

$$\tilde{z}_i = T_i^{-1}x = \begin{bmatrix} G_i x \\ F_i x \end{bmatrix} = \begin{bmatrix} v_i \\ z_i \end{bmatrix}$$

where $z_i \in R^{n_i}$. This coordinate transformation

leads to the transformed matrices

$$\tilde{A}^i = T_i^{-1} A(i) T_i, \; \tilde{C}^i = C(i) T_i$$

With the structure

$$\tilde{A}^i = \begin{bmatrix} \tilde{A}_{11}^i & \tilde{A}_{12}^i \\ 0 & \tilde{A}_{22}^i \end{bmatrix}, \; \tilde{C}^i = \begin{bmatrix} 0 & \tilde{C}_2^i \end{bmatrix}$$

Dynamics of the observable sub-state z_i

$$\dot{z}_i = \tilde{A}_{22}^i z_i$$

$$y = \tilde{C}_2^i z_i$$

and $(\tilde{C}_2^i, \tilde{A}_{22}^i)$ is observable.

Integration of Observable Sub-State Dynamics

From subsystem observers,

define the true observable sub-states

$$z = \begin{bmatrix} z_1 \\ \vdots \\ z_m \end{bmatrix}$$

and their estimates, to be designed later, as

$$\hat{z} = \begin{bmatrix} \hat{z}_1 \\ \vdots \\ \hat{z}_m \end{bmatrix}$$

Integrated Error Dynamics

Under Assumption 1,by Lemma 4, F is of $\left(\text{column}\right)$ rank n.

As a result, $\Phi = \left(F'F\right)^{-1}F'$ is of $\left(\text{row}\right)$ rank n, and

$x = \Phi\, z$. Then the estimate of x can be derived $\hat{x} = \Phi\, \hat{z}$.

Conversely, since $z_i\left(t\right) = F_i x\left(t\right)$,

we have $z\left(t\right) = Fx\left(t\right)$ and $\hat{z} = F\hat{x}$.

The errors in estimating z_i and z are

$e_i = z_i - \hat{z}_i$, and $e = z - \hat{z}$.

Denote $\mu^i = \left\|e_i\right\|$, $\mu_k{}^i = \left\|e_i(k\tau)\right\|$, $\mu = \left\|e\right\|$, $\mu_k = \left\|e(k\tau)\right\|$

Observer Design at Subsystem Level

Case 1: $\alpha_k = i$

If $\alpha_k = i$, an observer for the i-thsubsystem is designed as follows:

$$\dot{\hat{z}}_i = \tilde{A}_{22}^i \hat{z}_i + L_i (y - \hat{y})$$

$$\hat{y} = \tilde{C}_2^i \ \hat{z}_i$$

where L_i is the observer feedback gain, to be designed for convergence.

Under $\alpha_k = i$, the observer error dynamics are

$$\dot{e}_i = (\tilde{A}_{22}^i \hat{z}_i - L_i C_2^i) e_i = A_c^i e_i .$$

The observer gain can be designed by the pole placement method such as Ackermann's formula such that

$$A_c^i = \tilde{A}_{22}^i \hat{z}_i - L_i C_2^i$$

has n_i eigenvalues with real part less than $-a_i$ for any $a_i > 0$

Under the given τ, for some $c > 0$,

$$\left\| A_c^i \right\| \le c\, e^{-a_i \tau}$$

which can be made arbitrarily small

by choosing sufficiently large a_i.

$\mu_{k+1}^i \le \gamma_c^i \, \mu_k^i$, where γ_c^i can be made arbitrarily small.

$$\gamma_c = \max_{i=1,\ldots,m} \gamma_c^i$$

γ_c will be selected later to ensure convergence of the integrated observer for the entire system.

Case 2: $\alpha_k \neq i$

On the other hand, if $\alpha_k \neq i$, the subsystem observer will run open loop. Since the open-loop dynamics of z_i will depend on the actual α_k, it will show interaction with other subsystems.

Since the true system is $\dot{x}(t) = A_k x(t)$ in $\left[k\tau,\ (k+1)\tau\right)$

and A_k is known, when $\alpha_k \neq i$, the i-thsubsystem will run open loop as

$$\dot{\hat{z}}_i = F_i \dot{\hat{x}}(t) = F_i A_k \hat{x}(t) = F_i A_k \Phi \hat{z}(t)$$

which implies

$$\dot{e}_i(t) = F_i A_k \Phi e(t).$$

It follows that

$$\mu_{k+1}^i \leq \gamma_o^i \mu_k \qquad \text{for some constant } \gamma_o^i > 0$$

Convergence Analysis of Observers: Independent Subspace Error Dynamics

Definition

A hybrid system is said to have **independent subspace error dynamics** if its observable sub-states dynamic system, namely the derivative of z_i, depends on z_i only.

Subsystem Error Dynamics

The situation is motivated by network systems whose dynamics are described by a large-scale constant ***A*** matrix, but each subsystem has different ***C(j)*** for different and highly limited sensing systems. Then, the subsystem state equation will be independent of α_k and

$$\dot{e}_i = A_{22}^i e_i, \quad \alpha_k \neq i$$

As a result, we have the error bound

$$\mu_{k+1}^i \leq \gamma_o^i \, \mu_k^i, \quad \alpha_k \neq i \qquad \text{Let } \gamma_0 = \max_{i=1,\ldots,m} \gamma_o^i$$

Together, we have

$$\dot{e}_i = I_{\alpha_k = i} A_c^i e_i + I_{\alpha_k \neq i} \tilde{A}_{22}^i e_i, \; t \in [k\tau, \; (k+1)\tau).$$

The errors are bounded by

$$\mu_{k+1}^i \leq \gamma_k^i \; \mu_k^i$$

with $\gamma_k^i = I_{\alpha_k = i} \gamma_c^i + I_{\alpha_k \neq i} \gamma_o^i$

Consequently,

$$\mu_{k+1}^i \leq \Pi_{j=1}^k \gamma_j^i \mu_0^i$$

and γ_j^i is i.i.d. with $P(\gamma_k^i = \gamma_c^i) = p_i > 0$ and $P(\gamma_k^i = \gamma_o^i) = 1 - p_i$

Lemma 5

Under Assumption 3, for any $\gamma_* < 1$, the pole positions in the observer design can be selected such that

$$(\gamma_c^i)^{p_i} (\gamma_o^i)^{1-p} \leq \gamma_* < 1$$

Strong Convergence of Subsystem Observers

Theorem 2

Under the observer design,

(i) μ_k^i converges strongly and exponentially to 0, as $k \to \infty$.

(ii) $\mu^i(t)$ converges strongly and exponentially to 0, as $t \to \infty$.

Strong Convergence of the Integrated Observer

Define

$$e(t) = \begin{bmatrix} e_1(t) \\ \vdots \\ e_m(t) \end{bmatrix}$$

The estimation error on x is $\varepsilon(t) = x(t) - \hat{x}(t)$ with error norm

$$\mu(t) = \|e(t)\|$$

Theorem 3

Underthe same assumptions as Theorem 2,

$\|\varepsilon(t)\|$ converges strongly and exponentially to 0, as $t \to \infty$

State Estimation Error Probability and Large Deviation Principles

The log-moment generating function

$$M(v) \; = \; \ln \, E \; e^{\gamma_j^i} = \ln \; (p \; e^{v \ln \gamma_c^i} + \; (1 - p) e^{v \ln \gamma_o^i})$$

Proposition 1

Suppose $(\gamma_c^i)^{p_i} (\gamma_o^i)^{1-p} \leq \gamma_* \; < \; 1 \; (\gamma_c^i < \; 1, \; \gamma_o^i > \; 1),$

Then under the condition of Lemma 3,

$M(v)$ is strictly convex,

$M(0) = 0, \; v^* = \arg \; \min \; M(v) \; > 0, \; and \; M(v \wedge *) < \; 0$

The LDP rate function is

$I(0) = \; \sup_v \; (- M(v)) = - \min_v \; M(v).$

Since $M(v)$ is a convex function and the random sequence is i.i.d., numerical computation of the minimum is simple.

Error Dynamics and Convergence: General Case

Lemma 5

When $\alpha_k = j$, the eigenvalues of A_c^j can be designed such that for any $0 < \gamma_c < 1$,

$$\mu_{k+1}^j \le \gamma_c \mu_k^j,$$

$$\mu_{k+1}^i \le \kappa\gamma_c \mu_k^j + \kappa \sum_{l \ne j} \mu_k^l \;, i \ne j.$$

Assumption 3

The subsystem interactions when a subsystem is running open loop are characterized by a triangular structure, namely, with possible re-arrangement of subsystem indices,

$$\dot{e}_j = \sum_{l=1}^{j} H_j^l \, e_l, \; j = 1, \ldots, m.$$

The systems of independent open dynamics are the special case of Assumption 3, when the triangular matrix is actually diagonal,

$$\dot{e}_j = \tilde{A}_{22}^j e_j$$

when it runs in open loop.

Theorem 4

Under the switching sequence $\tilde{\alpha}$, for systems satisfying Assumption 4, any value of $0 < \gamma_* < 1$ can be chosen so that $\mu_{k+m} \leq \gamma_* \mu_k$

This special is a sufficient condition. It is not a necessary conditions since there may be other switching sequences of positive probability that also allow a suitable observer design to achieve similar convergence rates. But S is uniquely suitable for employing the triangular system structure to achieve desired convergence.

Strong Convergence of the Integrated Observer

Theorem 5

If γ_* is designed such that $(\gamma_*)^{p_i}(\gamma)^{1-p} < 1$,

then $\mu_{Lm} \to 0, \quad w.p.1., \quad L \to \infty$

Thank You!