SCIENCE CHINA
Information Sciences

• Supplementary File •

ADP-based composite learning framework for asynchronous event-triggered control of singularly perturbed systems and its applications

Jianguo ZHAO[1,2], Yangmin LI[2*] & Chunyu YANG[1]

[1]*School of Information and Control Engineering, China University of Mining and Technology, Xuzhou 221116, China*
[2]*Department of Industrial and Systems Engineering, The Hong Kong Polytechnic University, Hong Kong 999077, China*

Appendix A Literature review and motivation

Many practical systems possess fast and slow phenomena, including industrial processes [1, 2], power systems [3, 4], chopper circuit [5], and mechatronic systems [6, 7]. Such plants are called singularly perturbed systems (SPSs) or multi-time-scale systems in the control community [8]. In particular, a singular perturbation parameter μ is leveraged to indicate the degree of separation between slow and fast states, which could denote quantities such as small moments of inertia, time constant, or parasitic capacitance. Since μ is usually relatively small and perturbed, the conventional control designs are unsuitable for SPSs. To alleviate the inherent numerical stiffness issues, a common approach, which is independent of μ, is to decompose SPS into reduced-order slow and fast subsystems, and then design a composite controller to stabilize the original full-order system [9, 10]. However, this model reduction technique heavily relies on the precise knowledge of plant dynamics, no matter if decomposing a system or implementing a composite controller.

This letter focuses on the event-triggered composite optimal control problem of unknown linear SPSs through adaptive dynamic programming (ADP), which is a computational intelligence art, that can iteratively learn the optimal control strategy in an unknown environment [11–13]. In the past decade, ADP technique had been leveraged to address numerous real-world engineering problems, for instance, wastewater treatment process [14], underactuated autonomous vehicles [15], mixed traffic systems [16], and motor speed servo systems [7]. Also, ADP has been widely applied in the power system; see, for instance [17–20] and the references therein. In order to address the limited wireless resources, an ADP-based control torque quantization approach was presented in [21] for flexible spacecraft. This approach cleverly combines the integral sliding surface to stabilize the attitude control system and guarantee the near-optimal performance. Mu et al [22] investigated the optimal avoidance control problem for multiplayer systems, in which a safe ADP scheme is proposed to avoid both multiple static and dynamic no-entry regions. By recasting the minimal polynomial of exosystem, an output-feedback controller was designed in [23] for output regulation problem. This method employs the state reconstruction methodology to reformulate the system state by means of input and output measurements. In [24], an novel value iteration algorithm was proposed for two-player zero-sum game. This algorithm has the ability to learn a stabilizing policy pair and could be initialed by an arbitrary control. By using a robust policy iteration, Wei et al [25] developed an efficient optimal self-learning control algorithm for Vertical Take Off and Landing aircraft systems. Nevertheless, these mentioned methods are limited to single-time-scale systems.

Extending ADP to optimal control designs of SPSs has recently received considerable attention. When the system dynamics are unknown but μ is known, some researchers developed full-order optimal controller design algorithms for linear [26], nonlinear [27], and Markov jump SPSs [28]. But, this full-order design may suffer from high dimensionality [8]. In [29] and [30], the ADP-based model-free reduced-dimensional control methods were proposed for the case where the fast dynamics are unknown but asymptotically stable. Some practical applications were also reported, such as wide-area damping control [4] and motor speed regulation [7]. However, this reduced feedback controller can lead the system far from its desired performance due to ignorance of fast dynamics [8]. For SPSs with known fast dynamics and unknown slow dynamics, some authors of [31–35] imposed different ADP-based composite optimal control methodologies to optimize the overall system. In order to achieve completely model-free composite designs, the authors of [2] and [36] used the least-squares algorithm to identify the information related to fast dynamics for industrial processes. But such an identification procedure has a negative estimation error because of ignoring the measurement of the fast subsystem. Hence, how to design the ADP-based composite optimal controller without identification remains open to SPSs with unknown slow and fast dynamics, which is also our main motivation.

On the other hand, the event-triggered control has proven its advantages in reducing the network bandwidth and saving communication resources [37]. Roughly speaking, the signal is sampled, and the measurement is updated only when some triggering rules are satisfied. Recently, the ADP-based event-triggered optimal control has also been studied to provide a tradeoff between resources and performance. In [38], a novel decentralized event-triggered ADP control approach was presented for mismatched interconnected systems. Both constrained inputs and uncertainties for unknown multiplayer nonlinear systems were addressed in [39]. To tackle the difficulty arising from the unmeasurable state, the authors of [40] designed an optimal controller by event-triggered output-feedback. However, most recent studies focus on single-time-scale systems. For SPSs, the slow state varies much slower than the fast state. The existing event-triggered works [10, 41, 42] not only rely on model parameters but are computationally inefficient because of the coupling of slow

* Corresponding author (email: yangmin.li@polyu.edu.hk)

and fast information in the triggering mechanism. In the absence of the knowledge of plant dynamics, it shall be thus of significance to develop asynchronous triggering conditions related to decoupled slow and fast modes for the ability to strengthen the usage of limited resources. This is the second motivation of this letter.

In this letter, we propose a novel composite learning approach to address the asynchronous event-triggered optimal control problem of linear SPSs subject to unknown slow and fast dynamics through ADP. Existing researches on ADP-based composite optimal control of SPSs suppose that the fast dynamics are known *a priori*. To obviate this limitation, we first introduce an overall performance index with a convergence factor, wherein the solutions to two separate optimal control problems associated with slow and fast subsystems are guaranteed under some rationale conditions. Then, an ADP-based learning algorithm is developed to find the composite optimal control strategy without requiring the knowledge of both slow and fast dynamics. Besides, we introduce the event-triggering mechanism to save communication resources and achieve asynchronous measurement updates for its slow and fast states. The closed-loop stability is rigorously analyzed using singular perturbation theory, and the existence of a positive lower bound of the execution time between two triggers is ensured to exclude the Zeno behavior. Finally, a benchmark example and a two-area power system are leveraged to manifest the efficacy of our theoretical results.

Appendix B Composite optimal control of SPSs

Consider the system (1) and the performance index (2), then $J(x_0, u)$ is minimized with

$$u^* = K_1 x + K_2 z \tag{B1}$$

where $[K_1, K_2] \triangleq -\mathcal{R}^{-1}\mathcal{B}_\mu^T P_\mu$ and P_μ satisfies the following algebraic Riccati equation (ARE)

$$\mathcal{A}_\mu^T P_\mu + P_\mu \mathcal{A}_\mu + \mathcal{Q}_\mu - P_\mu \mathcal{B}_\mu \mathcal{R}^{-1}\mathcal{B}_\mu^T P_\mu + 2\alpha P_\mu = 0 \tag{B2}$$

with

$$\mathcal{A}_\mu = \begin{bmatrix} \mathcal{A}_{11} & \mathcal{A}_{12} \\ \frac{\mathcal{A}_{21}}{\mu} & \frac{\mathcal{A}_{22}}{\mu} \end{bmatrix},\ \mathcal{B}_\mu = \begin{bmatrix} \mathcal{B}_1 \\ \frac{\mathcal{B}_2}{\mu} \end{bmatrix},\ \mathcal{Q}_\mu = \begin{bmatrix} \mathcal{C}^T\mathcal{Q}\mathcal{C} & 0 \\ 0 & 0 \end{bmatrix}.$$

However, owing to the existence of singular perturbation parameter μ, the conventional algorithms to solve ARE (B2) might face numerical stiffness issues [8, 26]. For this reason, we introduce composite control designs for solving the optimal control problem associated with (1) and (2).

Inspired by [8], we shall decompose the system (1) into two different time-scale subsystems and then give the near-optimal composite controller designs via two separate optimal control problems. Setting $\mu = 0$ in (1), we get

$$\dot{x}_s = \mathcal{A}_{11} x_s + \mathcal{A}_{12} z_s + \mathcal{B}_1 u_s \tag{B3a}$$

$$0 = \mathcal{A}_{21} x_s + \mathcal{A}_{22} z_s + \mathcal{B}_2 u_s. \tag{B3b}$$

Since $\mathcal{A}_{22}$ is nonsingular, (B3b) yields

$$z_s = -\mathcal{A}_{22}^{-1}\mathcal{A}_{21} x_s - \mathcal{A}_{22}^{-1}\mathcal{B}_2 u_s. \tag{B4}$$

Substituting (B4) into (B3a) gives the slow time-scale subsystem

$$\dot{x}_s(t) = \mathcal{A}_0 x_s(t) + \mathcal{B}_0 u_s(t) \tag{B5}$$

where $\mathcal{A}_0 = \mathcal{A}_{11} - \mathcal{A}_{12}\mathcal{A}_{22}^{-1}\mathcal{A}_{21}$ and $\mathcal{B}_0 = \mathcal{B}_1 - \mathcal{A}_{12}\mathcal{A}_{22}^{-1}\mathcal{B}_2$. From (2), the corresponding performance index for (B5) is

$$J_s(x_s(0), u_s) = \int_0^\infty \mathrm{e}^{2\alpha\tau}(x_s^T \mathcal{C}^T \mathcal{Q}\mathcal{C} x_s + u_s^T \mathcal{R} u_s) d\tau. \tag{B6}$$

Under Assumption 1, by [43], the optimal control for (B5), (B6) can be obtained as

$$u_s^* = K_0 x_s \triangleq -\mathcal{R}^{-1}\mathcal{B}_0^T P_s x_s \tag{B7}$$

where P_s is the unique positive definite solution of the ARE

$$\mathcal{A}_0^T P_s + P_s \mathcal{A}_0 + \mathcal{C}^T\mathcal{Q}\mathcal{C} - P_s \mathcal{B}_0 \mathcal{R}^{-1}\mathcal{B}_0^T P_s + 2\alpha P_s = 0. \tag{B8}$$

Let $z_f = z - z_s$ and $u_f = u - u_s$. From (1b) and (B4), the fast time-scale subsystem of (1) is defined as

$$\mu \dot{z}_f(t) = \mathcal{A}_{22} z_f(t) + \mathcal{B}_2 u_f(t). \tag{B9}$$

Also, the corresponding performance index from (2) is given by

$$J_f(z_f(0), u_f) = \int_0^\infty \mathrm{e}^{2\alpha\tau} u_f^T \mathcal{R} u_f d\tau. \tag{B10}$$

Under Assumptions 2 and 3, the control input of (B9) that minimizes (B10) in the fast time scale [8] is

$$u_f^* = K_2 z_f \triangleq -\mathcal{R}^{-1}\mathcal{B}_2^T P_f z_f \tag{B11}$$

where P_f is the unique positive definite solution of the ARE [44]

$$\mathcal{A}_{22}^T P_f + P_f \mathcal{A}_{22} - P_f \mathcal{B}_2 \mathcal{R}^{-1}\mathcal{B}_2^T P_f + 2\alpha P_f = 0. \tag{B12}$$

Lemma B1 ([8]). Let K_0 and K_2 be designed such that $\mathrm{Re}(\lambda(\mathcal{A}_0-\mathcal{B}_0K_0))<0$ and $\mathrm{Re}(\lambda(\mathcal{A}_{22}-\mathcal{B}_2K_2))<0$. If the controls (B7), (B11), and (B1) are applied to systems (B5), (B9), and (1), respectively, then there exists a $\mu^*>0$ such that

$$x(t)=x_s(t)+\mathcal{O}(\mu) \tag{B13a}$$
$$z(t)=z_s(t)+z_f(t)+\mathcal{O}(\mu) \tag{B13b}$$
$$u^*(t)=u_s^*(t)+u_f^*(t)+\mathcal{O}(\mu) \tag{B13c}$$

for all $\mu\in(0,\mu^*]$.

The results of Lemma B1 imply that we can construct the composite controller

$$\begin{aligned}u&=u_s^*+u_f^*\\&=K_0x_s+K_2z_f\\&=K_0x+K_2\left(z+\mathcal{A}_{22}^{-1}\mathcal{A}_{21}x_s+\mathcal{A}_{22}^{-1}\mathcal{B}_2u_s^*\right)\\&=K_1x+K_2z\end{aligned} \tag{B14}$$

as a near-optimal control input to solve the optimal problem (1), (2), where $K_1=(I+K_2\mathcal{A}_{22}^{-1}\mathcal{B}_2)K_0+K_2\mathcal{A}_{22}^{-1}\mathcal{A}_{21}$.

Appendix C Data-driven composite learning algorithm

At the outset, for $\mathscr{A}=[a_i]\in\mathbb{R}^n$, $\mathscr{B}=[b_1,b_2,\ldots,b_m]\in\mathbb{R}^{n\times m}$, and $\mathscr{C}=[c_{ij}]\in\mathbb{R}^{n\times n}$, we define the notations

$$\begin{aligned}\mathrm{vev}(\mathscr{A})&\triangleq[a_1^2,a_1a_2,\ldots a_1a_n,a_2^2,a_2a_3,\ldots,a_{n-1}a_n,a_n^2]^T\in\mathbb{R}^{0.5n(n+1)}\\\mathrm{vec}(\mathscr{B})&\triangleq[b_1^T,b_2^T,\ldots,b_n^T]^T\in\mathbb{R}^{nm}\\\mathrm{ves}(\mathscr{C})&\triangleq[c_{11},2c_{12},\ldots,2c_{1n},c_{22},2c_{23},\ldots,2c_{n-1,n},c_{nn}]^T\in\mathbb{R}^{0.5n(n+1)}.\end{aligned}$$

For any two vectors β,ϑ and a sufficiently large number $h>0$, we further define

$$\begin{aligned}\Gamma_\beta&\triangleq[\mathrm{vev}(\bar\beta(t_1))-\mathrm{vev}(\bar\beta(t_0)),\mathrm{vev}(\bar\beta(t_2))-\mathrm{vev}(\bar\beta(t_1)),\ldots,\mathrm{vev}(\bar\beta(t_h))-\mathrm{vev}(\bar\beta(t_{h-1}))]^T\\\Pi_{\beta\vartheta}&\triangleq[\int_{t_0}^{t_1}\bar\beta\otimes\bar\vartheta d\tau,\int_{t_1}^{t_2}\bar\beta\otimes\bar\vartheta d\tau,\ldots,\int_{t_{h-1}}^{t_h}\bar\beta\otimes\bar\vartheta d\tau]^T\end{aligned}$$

where $\bar\beta=\mathrm{e}^{\alpha t}\beta$, $\bar\vartheta=\mathrm{e}^{\alpha t}\vartheta$, $0\leqslant t_0<t_1<\cdots<t_{h-1}<t_h$ are time constants, and $\otimes$ represents the Kronecker product operator.

By Kronecker product representation, the integral equation (5) is reorganized as

$$\mathcal{X}_s^k\begin{bmatrix}\mathrm{ves}(P_s^k)\\\mathrm{vec}(K_0^{k+1})\end{bmatrix}=\mathcal{W}_s^k \tag{C1}$$

with $\mathcal{X}_s^k=[\Gamma_{x_s},-2\Pi_{x_sx_s}(I\otimes(K_0^k)^T\mathcal{R})+2\Pi_{x_su_s^0}(I\otimes\mathcal{R})]$ and $\mathcal{W}_s^k=-\Pi_{y_sy_s}\mathrm{vec}(\mathcal{Q})-\Pi_{x_sx_s}\mathrm{vec}((K_0^k)^T\mathcal{R}K_0^k)$. Note that the linear matrix equation (C1) can be used to compute P_s^k,K_0^{k+1} without knowledge of system dynamics. But, the information of x_s cannot be directly measured from the actual plant (1). Thanks to Lemma B1, we will utilize x instead of x_s during learning. As a result, we rewrite (C1) as

$$\mathcal{X}^k\begin{bmatrix}\mathrm{ves}(P_s^k)\\\mathrm{vec}(K_0^{k+1})\end{bmatrix}=\mathcal{W}^k \tag{C2}$$

with $v_s^0=K_0^0x+u_{se}$, $\mathcal{X}^k=[\Gamma_x,-2\Pi_{xx}(I\otimes(K_0^k)^T\mathcal{R})+2\Pi_{xv_s^0}(I\otimes\mathcal{R})]$, and $\mathcal{W}^k=-\Pi_{yy}\mathrm{vec}(\mathcal{Q})-\Pi_{xx}\mathrm{vec}((K_0^k)^T\mathcal{R}K_0^k)$. Similarly, we replace x_s of (7) with x by Lemma 1. By (B13a), we rewrite (7) as

$$\begin{aligned}&\mu\mathrm{e}^{2\alpha(t+\delta t)}\phi^T(t+\delta t)\bar P_f^j\phi(t+\delta t)-\mu\mathrm{e}^{2\alpha t}\phi^T(t)P_f^j\phi(t)\\=&-2\int_t^{t+\delta t}\mathrm{e}^{2\alpha\tau}(v_f^0-\bar K_2^j\phi)\mathcal{R}\bar K_2^{j+1}\phi d\tau-\int_t^{t+\delta t}\mathrm{e}^{2\alpha\tau}\phi^T(\bar K_2^j)^T\mathcal{R}\bar K_2^j\phi d\tau\end{aligned} \tag{C3}$$

where $\phi=[z^T,x^T,(v_s^0)^T]^T$ and $v_f^0=\bar K_2^0\phi+u_{fe}$. By Kronecker product representation, the integral equation (C3) is reorganized as

$$\mathcal{Z}^j\begin{bmatrix}\mathrm{ves}(\bar P_f^j)\\\mathrm{vec}(\bar K_2^{j+1})\end{bmatrix}=\mathcal{Y}^j \tag{C4}$$

with $\mathcal{Z}^j=[\mu\Gamma_\phi,-2\Pi_{\phi\phi}(I\otimes(\bar K_2^j)^T\mathcal{R})+2\Pi_{\phi v_f^0}(I\otimes\mathcal{R})]$ and $\mathcal{Y}^j=-\Pi_{\phi\phi}\mathrm{vec}((\bar K_2^j)^T\mathcal{R}\bar K_2^j)$. Thus, we can invoke the linear matrix equation (C4) to compute $\bar P_f^j$ and $\bar K_2^{j+1}=[K_2^{j+1},K_2^{j+1}\mathcal{A}_{22}^{-1}\mathcal{A}_{21},K_2^{j+1}\mathcal{A}_{22}^{-1}\mathcal{B}_2]$ based on state and input data. The following lemma guarantees the uniqueness of solutions to (C2), (C4). We omit its proof that adopts the same line of proofs as in [11].

Lemma C1. For all $k, j \in \mathbb{N}$, if there exists an $h^* \in \mathbb{N}$ so that for all $h > h^*$,

$$\text{rank}([\Pi_{\phi\phi}, \Pi_{\phi v_f^0}]) = 0.5(n_x + n_z + n_u)(n_x + n_z + 3n_u + 1) \tag{C5}$$

then matrices $\mathcal{X}^k, \mathcal{Z}^j$ have full column rank for all $k, j \in \mathbb{N}$.

We summarize the data-driven ADP learning routine in Algorithm C1 that iteratively learns the composite optimal feedback gains, where $L^j \triangleq K_2^j \mathcal{A}_{22}^{-1}\mathcal{A}_{21}$ and $G^j \triangleq K_2^j \mathcal{A}_{22}^{-1}\mathcal{B}_2$. As shown in Algorithm C1, we need to apply an input with appropriate control gains to excite system (1), so that the collected data satisfy (C5). These data are used to learn the optimal control gains of both fast and slow subsystems through (C2) and (C4), respectively. Then, the optimal policy is constructed based on the learned gain matrices. It is interesting to point out that Algorithm C1 enjoys parallel computing.

Algorithm C1 Data-Driven Composite Learning Algorithm

1: Choose initial gains K_0^0 and $\bar{K}_2^0$ such that $\mathcal{A}_0 + \alpha I + \mathcal{B}_0 K_0^0$ and $\mathcal{A}_{22} + \alpha I + \mathcal{B}_2 K_2^0$ are Hurwitz, and small thresholds σ_s, σ_f
2: Apply $u = v_s^0 + v_f^0$ to excite system (1) on $[t_0, t_h]$ and collect data x, z, y, v_s^0, v_f^0 such that (C5) holds
3: $k, j = 0$
4: **repeat**
5: Solve P_s^k and K_0^{k+1} from (C2)
6: $k = k + 1$
7: **until** $||P_s^k - P_s^{k-1}|| < \sigma_s$
8: **repeat**
9: Solve $\bar{P}_f^j$ and $\bar{K}_2^{j+1}$ from (C4)
10: $j = j + 1$
11: **until** $||\bar{P}_f^j - \bar{P}_f^{j-1}|| < \sigma_f$
12: Output $K_1 = (I + G^j)K_0^k + L^j$ and $K_2 = K_2^j$, where $\bar{K}_2^j = [K_2^j, L^j, G^j]$

Remark 1. To eliminate the dependence on model parameters in the composite control designs, the existing results [2, 36] invoked (B4) to identify the information of $\mathcal{A}_{22}^{-1}\mathcal{A}_{21}$ and $\mathcal{A}_{22}^{-1}\mathcal{B}_2$, in which the signal of z_s is replaced by the data of z during identification. However, it can be found that because of $z(t) = z_s(t) + z_f(t) + \mathcal{O}(\mu)$ from (B13b), this way has an unacceptable estimation error caused by $z_f(t) \neq 0$. In contrast, our algorithm meaningfully avoids the identification process and thus has better learning performance. This analysis is consistent with our simulation results.

Remark 2. The data required in Algorithm C1 are collected off-line. Algorithm C1 provides a data-driven method to design optimal controller (4). The data regarding x, z, y, and u are needed in Algorithm C1. In the data collection phase, the probing noise in the control input can ensure that the rank condition (C5) in Lemma C1 is satisfied. These datasets are collected off-line and are cyclically employed in successive iterations.

Appendix D Proof of Theorem 1

Dividing (5) by δt and taking limit gives its model-based algorithm. This means that Algorithm C1 is in fact equivalent to Kleinman's algorithm. Hence, by [45, Theorem], $\lim_{k\to\infty} P_s^k = P_s$ and $\lim_{k\to\infty} K_0^k = K_0$ from (C1). By (B13a) of Lemma B1, we can rewrite (C2) as

$$\mathcal{X}_s^k \begin{bmatrix} \text{ves}(P_s^k + \mathcal{O}(\mu)) \\ \text{vec}(K_0^{k+1} + \mathcal{O}(\mu)) \end{bmatrix} = \mathcal{W}_s^k.$$

Therefore, comparing the above equation and (C1), the property (8a) holds. The proof of the property (8b) can be shown by the same line, which is thus omitted.

Appendix E Proof of Theorem 2

In order to analyze closed-loop stability, we introduce the Chang transformation [10]

$$\begin{bmatrix} \varsigma_s(t) \\ \varsigma_f(t) \end{bmatrix} = \begin{bmatrix} I - \mu\mathcal{H}\mathcal{L} & -\mu\mathcal{H} \\ \mathcal{L} & I \end{bmatrix} \begin{bmatrix} x(t) \\ z(t) \end{bmatrix} \tag{E1a}$$

$$y_{\varsigma_s}(t) = \mathcal{C}\varsigma_s(t) \tag{E1b}$$

where $\varsigma_s \in \mathbb{R}^{n_x}$ and $\varsigma_f \in \mathbb{R}^{n_z}$ are the so-called exact slow and fast modes, $y_{\varsigma_s} \in \mathbb{R}^{n_y}$ is a nominal output, matrices $\mathcal{L} \in \mathbb{R}^{n_z \times n_x}$ and $\mathcal{H} \in \mathbb{R}^{n_x \times n_z}$ are found from the following equations

$$\mu\mathcal{L}\Lambda_{11} + \Lambda_{21} - \Lambda_{22}\mathcal{L} - \mu\mathcal{L}\Lambda_{12}\mathcal{L} = 0 \tag{E2a}$$

$$\mu\Lambda_{11}\mathcal{H} - \mu\Lambda_{12}\mathcal{L}\mathcal{H} + \Lambda_{12} - \mathcal{H}\Lambda_{22} - \mu\mathcal{H}\mathcal{L}\Lambda_{12} = 0 \tag{E2b}$$

where $\Lambda_{ij} = \mathcal{A}_{ij} + \mathcal{B}_i K_j$ for $i, j = 1, 2$. Meanwhile, in view of (4), the closed-loop system (1) becomes

$$\begin{bmatrix} \dot{x}(t) \\ \mu\dot{z}(t) \end{bmatrix} = \begin{bmatrix} \Lambda_{11} & \Lambda_{12} \\ \Lambda_{21} & \Lambda_{22} \end{bmatrix} \begin{bmatrix} x(t) \\ z(t) \end{bmatrix} + \begin{bmatrix} \mathcal{B}_1 \\ \mathcal{B}_2 \end{bmatrix} Ke(t) \tag{E3}$$

where $K=[K_1,K_2]$ and $e=[e_x^T,e_z^T]^T$. Using (E1), we arrange the system (E3) in $(\varsigma_s,\varsigma_f)$-coordinates

$$\begin{bmatrix}\dot{\varsigma}_s(t)\\ \dot{\varsigma}_f(t)\end{bmatrix}=\begin{bmatrix}\mathcal{A}_s+\mathcal{B}_sK_s & 0\\ 0 & \frac{\mathcal{A}_f+\mathcal{B}_fK_2}{\mu}\end{bmatrix}\begin{bmatrix}\varsigma_s(t)\\ \varsigma_f(t)\end{bmatrix}+\begin{bmatrix}\mathcal{B}_a\\ \frac{\mathcal{B}_f}{\mu}\end{bmatrix}Ke(t) \tag{E4}$$

where

$$\begin{aligned}\mathcal{A}_s&=\mathcal{A}_0-\mu\mathcal{A}_{12}\mathcal{A}_{22}^{-1}\mathcal{L}(\mathcal{A}_{11}-\mathcal{A}_{12}\mathcal{L})\\ \mathcal{B}_s&=\mathcal{B}_0-\mu\mathcal{A}_{12}\mathcal{A}_{22}^{-1}\mathcal{L}\mathcal{B}_1\\ K_s&=K_1-K_2\mathcal{L}\\ \mathcal{B}_a&=\mathcal{B}_1-\mathcal{H}\mathcal{B}_2-\mu\mathcal{H}\mathcal{L}\mathcal{B}_1\\ \mathcal{A}_f&=\mathcal{A}_{22}+\mu\mathcal{L}\mathcal{A}_{12}\\ \mathcal{B}_f&=\mathcal{B}_2+\mu\mathcal{L}\mathcal{B}_1.\end{aligned} \tag{E5}$$

By (E1), we observe that the stability of $[\varsigma_s^T,\varsigma_f^T]^T$ suggests the stability of $[x^T,z^T]^T$. Thus, we shall resort to the system (E4) instead of (E3) for proving the stability. For $\mathcal{O}(\mu)$ approximations, by (E2) and (E5), we have

$$\mathcal{A}_s+\mathcal{B}_sK_s=\mathcal{A}_0+\mathcal{B}_0K_0+\mathcal{O}(\mu) \tag{E6a}$$
$$\mathcal{A}_f+\mathcal{B}_fK_2=\mathcal{A}_{22}+\mathcal{B}_2K_2+\mathcal{O}(\mu) \tag{E6b}$$
$$\mathcal{B}_f=\mathcal{B}_2+\mathcal{O}(\mu). \tag{E6c}$$

Consider the candidate Lyapunov function $V(\varsigma_s,\varsigma_f)=\varsigma_s^TP_s\varsigma_s+\varsigma_f^T\mu P_f\varsigma_f$. The derivative of $V(t)$ along the trajectories of (E4) is

$$\begin{aligned}\dot{V}(\varsigma_s,\varsigma_f)=&\,\varsigma_s^T((\mathcal{A}_s+\mathcal{B}_sK_s)^TP_s+P_s(\mathcal{A}_s+\mathcal{B}_sK_s))\varsigma_s+2\varsigma_s^TP_s\mathcal{B}_aKe\\ &+\varsigma_f^T((\mathcal{A}_f+\mathcal{B}_fK_2)^TP_f+P_f(\mathcal{A}_f+\mathcal{B}_fK_2))\varsigma_f+2\varsigma_f^TP_f\mathcal{B}_fKe\\ =&-\varsigma_s^T(K_0^T\mathcal{R}K_0+\mathcal{C}^T\mathcal{Q}\mathcal{C}+2\alpha P_s)\varsigma_s+2\varsigma_s^TP_s\mathcal{B}_aK_1e_x+2\varsigma_s^TP_s\mathcal{B}_aK_2e_z\\ &-\varsigma_f^T(K_2^T\mathcal{R}K_2+2\alpha P_f)\varsigma_f+2\varsigma_f^TK_2^T\mathcal{R}K_1e_x+2\varsigma_f^TK_2^T\mathcal{R}K_2e_z+\mathcal{O}(\mu)\end{aligned} \tag{E7}$$

where (B7), (B8), (B11), (B12), and (E6) are used. By matrix operations, we have

$$\varsigma_s^TP_s\mathcal{B}_aK_1e_x\leqslant c\,\|P_s\varsigma_s\|\,\|K_1e_x\| \tag{E8a}$$
$$\begin{aligned}\varsigma_s^TP_s\mathcal{B}_aK_2e_z&\leqslant c\,\|P_s\varsigma_s\|\,\|K_2e_z\|\\ &\leqslant\frac{c}{2\gamma_s}\|P_s\varsigma_s\|^2+\frac{c\gamma_s}{2}\|K_2e_z\|^2\end{aligned} \tag{E8b}$$
$$\begin{aligned}\varsigma_f^TK_2^T\mathcal{R}K_1e_x&\leqslant\|\mathcal{R}\|\,\|K_2\varsigma_f\|\,\|K_1e_x\|\\ &\leqslant\frac{\|\mathcal{R}\|}{2\gamma_f}\|K_2\varsigma_f\|^2+\frac{\gamma_f\,\|\mathcal{R}\|}{2}\|K_1e_x\|^2\end{aligned} \tag{E8c}$$

for any $\gamma_s,\gamma_f>0$. Substituting (E8) into (E7) yields

$$\dot{V}\leqslant\Phi_s(e_x)+\Phi_f(e_z)+\mathcal{O}(\mu) \tag{E9}$$

where Φ_s and Φ_f are defined in (9). For $\gamma_s>\frac{c\|P_s\|^2}{2\alpha\min\{\lambda(P_s)\}}$ and $\gamma_f>\frac{\|\mathcal{R}\|\|K_2\|^2}{2\alpha\min\{\lambda(P_f)\}}$, it follows from (9) that

$$\begin{aligned}\Phi_s(e_x(t)=0)=&-\varsigma_s^T(K_0^T\mathcal{R}K_0+2\alpha P_s)\varsigma_s-y_{\varsigma_s}^T\mathcal{Q}y_{\varsigma_s}+c\gamma_s^{-1}\|P_s\varsigma_s\|^2\\ \leqslant&-2\alpha\min\{\lambda(P_s)\}\|\varsigma_s\|^2+c\gamma_s^{-1}\|P_s\|^2\|\varsigma_s\|^2\\ \leqslant&\,0\\ \Phi_f(e_z(t)=0)=&-\varsigma_f^T(K_2^T\mathcal{R}K_2+2\alpha P_f)\varsigma_f+\gamma_f^{-1}\,\|\mathcal{R}\|\,\|K_2\varsigma_f\|^2\\ \leqslant&-2\alpha\min\{\lambda(P_f)\}\|\varsigma_f\|^2+\gamma_f^{-1}\,\|\mathcal{R}\|\,\|K_2\|^2\|\varsigma_f\|^2\\ \leqslant&\,0\end{aligned}$$

which lead to $\dot{V}\leqslant 0$ for sufficiently small μ. As a result, there exists a $\mu^*>0$ such for all $\mu\in(0,\mu^*]$, (E4) is asymptotically stable. This concludes the proof.

Appendix F Proof of Theorem 3

By applying inequality to $2c\,\|P_s\varsigma_s\|\,\|K_1e_x\|$, one has

$$2c\,\|P_s\varsigma_s\|\,\|K_1e_x\|\leqslant c\varepsilon_s^{-1}\|P_s\varsigma_s\|^2+c\varepsilon_s\|K_1e_x\|^2$$

Table G1 Real Parts of Eigenvalues of the Closed-Loop System under Composite Control

α	1.1	5	20	100
$\mathrm{Re}(\lambda_1)$	-1.7064	-8.6194	-35.1163	-176.3877
$\mathrm{Re}(\lambda_2)$	-2.5184	-11.6619	-47.0764	-236.0261
$\mathrm{Re}(\lambda_3)$	-108.4659	-777.7384	-3319.9383	-16868.7311
$\mathrm{Re}(\lambda_4)$	-180.8092	-1055.4804	-4451.3691	-22572.3552

Table G2 Performance Comparison of Our Method and Existing Methods for Benchmark Example

Algorithm	Our method	Existing methods [2], [36]
Identification step	No	Yes
Error	0.0392	1.0266
CPU time (s)	0.7734	1.3453

where $\varepsilon_s > \frac{c\|P_s\|^2}{2\alpha\min\{\lambda(P_s)\}-c\gamma_s^{-1}\|P_s\|^2} > 0$. Then, it follows from (9a) that

$$\varsigma_s^T(K_0^T\mathcal{R}K_0+2\alpha P_s+\mathcal{C}^T\mathcal{Q}\mathcal{C})\varsigma_s - c(\gamma_s^{-1}+\varepsilon_s^{-1})\|P_s\|^2\|\varsigma_s\|^2 \leqslant (\gamma_f\|\mathcal{R}\|+c\varepsilon_s)\|K_1e_x\|^2$$

which implies that

$$\frac{\left\|e_x(t_x^{i+1})\right\|}{\left\|\varsigma_s(t_x^{i+1})\right\|} \geqslant g_s,\ \forall i\in\mathbb{N} \tag{F1}$$

where $g_s = \sqrt{\frac{\min\{\lambda(K_0^T\mathcal{R}K_0+2\alpha P_s+\mathcal{C}^T\mathcal{Q}\mathcal{C})\}-c(\gamma_s^{-1}+\beta_s^{-1})\|P_s\|^2}{(\gamma_f\|\mathcal{R}\|+c\varepsilon_s)\|K_1\|^2}} > 0$. Now, we continue to look at the dynamics of $h_s(t) \triangleq \frac{\|e_x(t)\|}{\|\varsigma_s(t)\|}$. Since the parameters are bounded, there exists a constant $f_s > 0$, such that the evolution of the slow mode in (E4) satisfies $\|\dot{\varsigma}_s(t)\| \leqslant f_s(\|\varsigma_s(t)\|+\|e_x(t)\|+\|e_z(t)\|)$. Then, we have

$$\frac{dh_s(t)}{dt} \leqslant f_s(1+h_s(t))^2 + f_s(1+h_s(t))\frac{\|e_z(t)\|}{\|\varsigma_s(t)\|}$$

for all $\forall t\in[t_x^i, t_x^{i+1})$. Since the system (E4) is asymptotically stable by Theorem 2, there exists a constant $d_s > 0$ such that $d_s \geqslant f_s(1+h_s(t))\frac{\|e_z(t)\|}{\|\varsigma_s(t)\|}$. Hence, we have $\frac{dt}{dh_s} \geqslant \frac{1}{f_s(1+h_s)^2+d_s}$. It is shown in (F1) that $h_s(t_x^{i+1})$ evolves from 0 to a lower bound g_s between the two triggering instants, then the execution time $\mathscr{L}_x^i = t_x^{i+1} - t_x^i \geqslant \int_0^{g_s}\frac{1}{f_s(1+h_s)^2+d_s}dh > 0$. Similarly for the execution time $\mathscr{L}_z^i = t_z^{i+1} - t_z^i$, we thus omit its proof for saving of space.

Appendix G Applications

We carry out a benchmark example and a two-area power system to support our theoretical results.

Appendix G.1 Benchmark example

Consider a benchmark SPS (1) [8, 29] with the following structure

$$\mathcal{A}_{11} = \begin{bmatrix} 0 & 0.4 \\ 0 & 0 \end{bmatrix},\ \mathcal{A}_{12} = \begin{bmatrix} 0 & 0 \\ 0.345 & 0 \end{bmatrix},\ \mathcal{A}_{21} = \begin{bmatrix} 0 & -0.524 \\ 0 & 0 \end{bmatrix},\ \mathcal{A}_{22} = \begin{bmatrix} -0.465 & 0.262 \\ 0 & -1 \end{bmatrix}$$
$$\mathcal{B}_1 = \begin{bmatrix} 0 & 0 \end{bmatrix}^T,\ \mathcal{B}_2 = \begin{bmatrix} 0 & 1 \end{bmatrix}^T,\ \mathcal{C} = I_2$$

and $\mu = 0.01$. The weight matrices in (2) are respectively taken as $\mathcal{Q} = I_2$ and $\mathcal{R} = 1$.

We first test the feasibility of composite control designs. To this end, the real parts of eigenvalues of the resulting closed-loop system under (B14) for different α are shown in Table G1. We can see that the convergence rate of both slow and fast states is guaranteed to be no slower than $\mathrm{e}^{-\alpha t}$ in different situations. Besides, the ideal values of feedback gains and kernel matrices for slow and fast subsystems are as follows

$$P_s = \begin{bmatrix} 1231.5429 & 267.8236 \\ 267.8236 & 96.6106 \end{bmatrix},\ P_f = \begin{bmatrix} 9.9948 & 3.5628 \\ 3.5628 & 1.47 \end{bmatrix},\ K_0 = \begin{bmatrix} 52.0614 & 18.7798 \end{bmatrix},\ K_2 = \begin{bmatrix} 3.5628 & 1.47 \end{bmatrix}$$

with $\alpha = 1.1$. Then, we use Algorithm 1 to learn the associated parameters of (4) in a completely model-free manner. For this, the initial states are set as $x(0) = [1, 1]^T$ and $z(0) = [-1, -1]^T$. The data information is recorded at 300 sampled times, and the threshold is $\sigma_s = \sigma_f = 0.001$. As shown in Fig. G1, the associated parameters can be obtained from Algorithm 1. We further evaluate the event-triggered performance by executing composite optimal controller (4). By Theorem 2, we design the parameters $c = 0.01$, $\gamma_s = 207.3$, and $\gamma_f = 38.2$. From Fig. G2, we see that both slow and fast states of the closed-loop system under (4) converge asymptotically to the origin. Also, the transmission instants of both slow and fast states and the control input trajectories are shown in Fig. G3. It is clear that the sampling of the slow and fast states is asynchronous. Furthermore, we list Table G2 to compare the performance between the proposed Algorithm 1 and the existing algorithms [2], [36] for learning composite control, in which the error is defined as $||K_1 - K_1^*|| + ||K_2 - K_2^*||$.

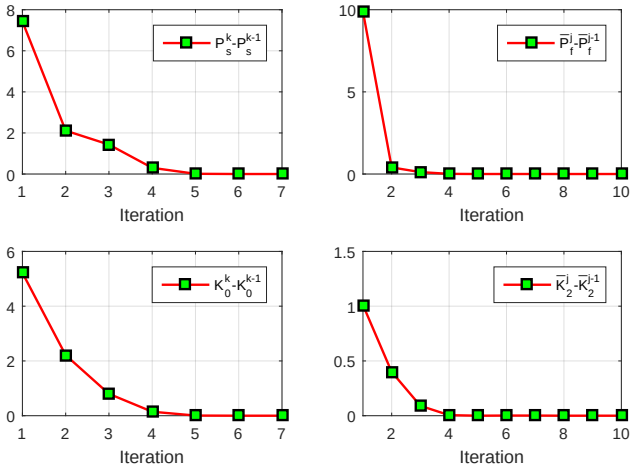

Figure G1 Convergence of P_s, K_0, $\bar{P}_f$, and $\bar{K}_2$ by Algorithm 1.

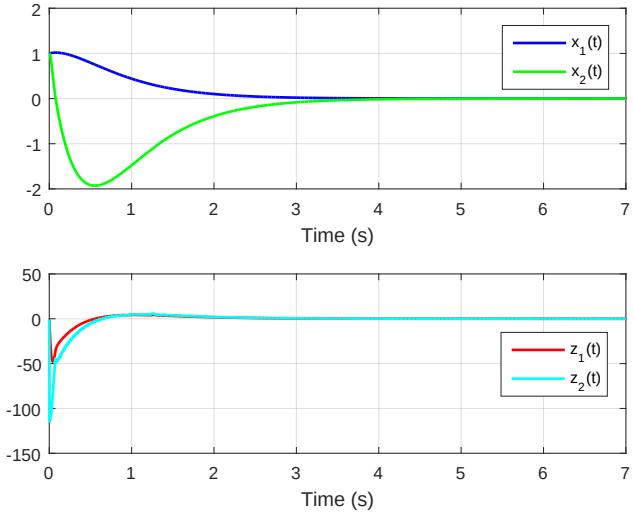

Figure G2 Evolution of the slow states (top) and the fast states (bottom) of the resulting closed-loop system.

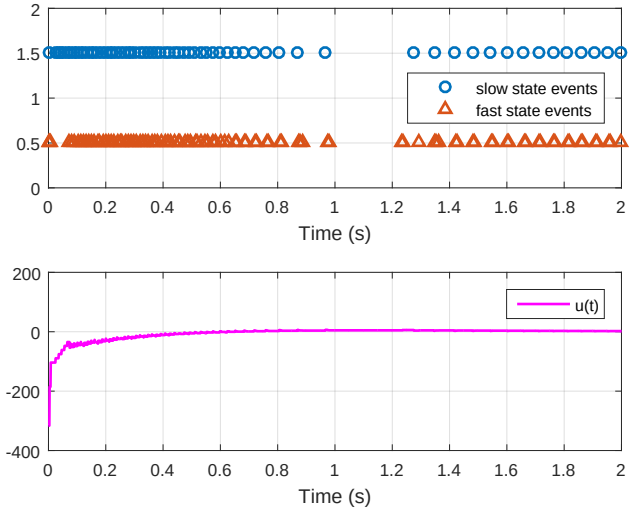

Figure G3 Transmission instants of slow and fast states (top), and evolution of the control input (bottom).

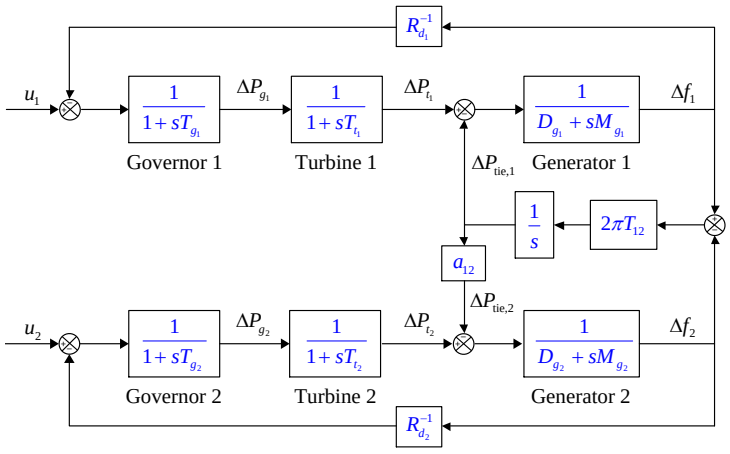

Figure G4 Structure of the two-area power system.

Table G3 Real Parts of Eigenvalues of the Two-Area Power System Under Composite Control

α	0.15	1	10	100
$\text{Re}(\lambda_1)$	-0.7120	-1.6942	-16.9786	-168.9831
$\text{Re}(\lambda_2)$	-0.7120	-2.0717	-16.9786	-168.9831
$\text{Re}(\lambda_3)$	-1.0173	-2.2692	-17.4612	-176.3966
$\text{Re}(\lambda_4)$	-1.2653	-2.2742	-23.3406	-235.9047
$\text{Re}(\lambda_5)$	-1.2653	-2.2742	-28.1932	-287.8315
$\text{Re}(\lambda_6)$	-11.6335	-153.1908	-1612.0252	-16166.3081
$\text{Re}(\lambda_7)$	-15.7804	-163.9115	-1688.6384	-16930.4912
$\text{Re}(\lambda_8)$	-27.8457	-217.3726	-2255.6098	-22642.2575
$\text{Re}(\lambda_9)$	-29.8685	-225.0417	-2310.8743	-23192.9442

Appendix G.2 Two-area power system

We take a two-area power system shown in Fig. G4 as an example to support our theoretical results. Following [35], the model of each area composed of generator, turbine, and governor is described by

$$\begin{aligned} M_{g_i}\frac{d}{dt}\Delta f_i &= -D_{g_i}\Delta f_i + \Delta P_{t_i} - \Delta P_{\text{tie},i} \\ T_{t_i}\frac{d}{dt}\Delta P_{t_i} &= -\Delta P_{t_i} + \Delta P_{g_i} \\ T_{g_i}\frac{d}{dt}\Delta P_{g_i} &= -\Delta P_{g_i} - \frac{1}{R_{d_i}}\Delta f_i + u_i \end{aligned}$$

where $i = 1, 2$, Δf_i is the frequency deviation, ΔP_{t_i} is the turbine output deviation, ΔP_{g_i} is the governor valve position deviation, u_i is the speed changer deviation, M_{g_i}, T_{t_i}, and T_{g_i} denote the generator inertia constant, turbine time constant, and governor time constant, respectively, R_{d_i} and D_{g_i} represent the speed drop and damping coefficient, a_{12} is coupling coefficient, respectively, $\Delta P_{\text{tie},i}$ is the tie-line power deviation

$$\frac{d}{dt}\Delta P_{\text{tie},1} = 2\pi T_{12}\left(\Delta f_1 - \Delta f_2\right)$$

where T_{12} represents the synchronizing power flow coefficient. The tie-line active power deviation exchange satisfies $\Delta P_{\text{tie},1} + \Delta P_{\text{tie},2} = 0$. Besides, the integral of area control error (ACE) is included into the state vector

$$\text{IACE}_i = \int \text{ACE}_i dt$$

where $\text{ACE}_i = \Delta P_{\text{tie},i} + \beta_i \Delta f_i$, and β_i is a frequency bias factor.

In this simulation, we suppose that these two areas are the same. The model parameters are specified as $T_g = 0.1$ s, $T_t = 0.2$ s, $M_g = 10$, $D_g = 0.5$, $R_d = 0.25$, $\beta = 4.5$, $a_{12} = -0.25$, and $T_{12} = 5.2$. Since the constants of the generator inertia and the integral control are much larger than the time constants of the governor and the turbine in this model, $x = (\Delta f_1, \Delta P_{\text{tie},1}, \text{IACE}_1, \Delta f_2, \text{IACE}_2)^T$ are taken as the slow states and $z = (\Delta P_{g_1}, \Delta P_{t_1}, \Delta P_{g_2}, \Delta P_{t_2})^T$ as the fast states. By (1), we set $\mu = T_t/(M_g/D_g) = 0.01$ and $\mathcal{C} = I_5$.

The weight matrices in (2) are respectively taken as $\mathcal{Q} = I_5$ and $\mathcal{R} = I_2$. We first test the feasibility of composite control designs. To this end, the real parts of eigenvalues of the resulting closed-loop power system under (B14) for different α are shown in Table G3. We can see that the convergence rate of both slow and fast states is guaranteed to be no slower than $\text{e}^{-\alpha t}$ in different situations. In case of $\alpha = 0.15$, we use Algorithm 1 to learn the associated parameters of (4) in a completely model-free manner. For this, the

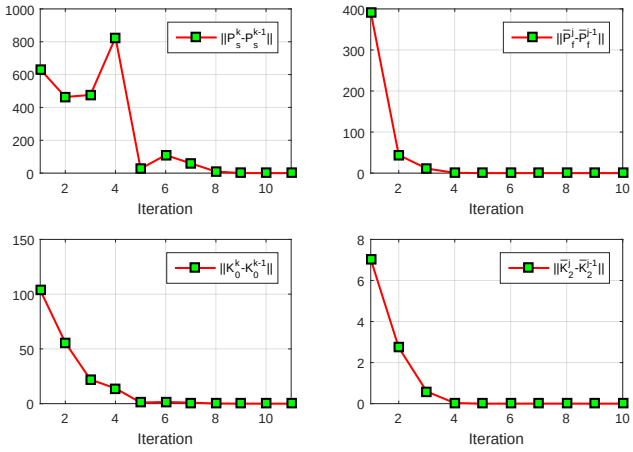

Figure G5 Convergence of P_s, K_0, $\bar{P}_f$, and $\bar{K}_2$ by Algorithm 1.

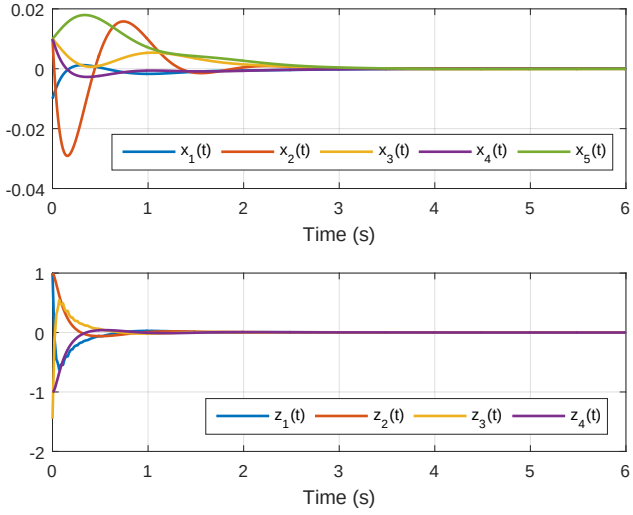

Figure G6 Evolution of the slow states (top) and the fast states (bottom).

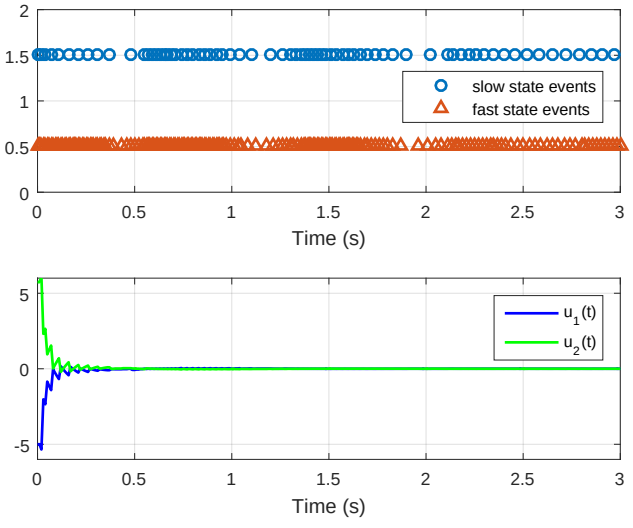

Figure G7 Transmission instants of slow and fast states (top), and evolution of the control input (bottom).

initial states are set as $x(0) = [-0.01, 0.01, 0.01, 0.01, 0.01]^T$ and $z(0) = [1, 1, -1.45, -1]^T$. The data information is recorded at 320 sampled times, and the threshold is $\sigma_s = \sigma_f = 0.001$. As shown in Fig. G5, the associated parameters can be obtained from Algorithm 1. We further evaluate the event-triggered performance by executing composite optimal controller (4). By Theorem 2, we design the

parameters $c = 0.03$, $\gamma_s = 2975$, and $\gamma_f = 17$. From Fig. G6, we see that both slow and fast states of the two-area power system under (4) converge asymptotically to the origin. Also, the transmission instants of both slow and fast states and the control input trajectories are shown in Fig. G7. It is clear that the sampling of the slow and fast states is asynchronous.

Remark 3. The eigenvalues of system (1) are clustered into two distinct groups, wherein the dynamics with larger eigenvalues are called fast state and the other is slow state. Due to the two-time-scale phenomenon of SPSs, we can also classify the fast and slow states through the rate of state change such as mechatronic systems [7] and industrial processes [1].

Remark 4. The proposed composite learning algorithm can be considered as a baseline framework to solve the optimal control problem associated with (1), (2). For system (1) with external disturbances, we can design a zero-sum game based control strategy to accommodate disturbance. In terms of the proposed composite learning framework, both optimal feedback control and worst-case disturbance can be learned simultaneously through single-loop policy iteration [46]. For system (1) with parameter uncertainties, we can guarantee the robustness of the control strategy by setting a larger weight matrix $\mathcal{Q}$ in (2). Our learning approach can be used to learn such robust control strategy by absorbing uncertainties into the control input [11].

References

1 Li J, Ni H, Chai T, et al. Reinforcement learning for optimal tracking of large-scale systems with multitime scales. Sci China Inf Sci, 2023, 66: 170201
2 Xue W, Fan J, Lopez V G, et al. Off-policy reinforcement learning for tracking in continuous-time systems on two time scales. IEEE Trans Neural Netw Learn Syst, 2021, 32: 4334–4346
3 Cheng J, Liu S, Yan H, et al. Sliding mode control for switching singularly perturbed systems: Adopting nonhomogeneous sojourn probabilities. IEEE Trans Autom Sci Eng, 2024, 21: 6575–6586
4 Mukherjee S, Chakrabortty A, Bai H, et al. Scalable designs for reinforcement learning-based wide-area damping control. IEEE Trans Smart Grid, 2021, 12: 2389–2401
5 Qi W, Li R, Park J H, et al. Observer-based stabilization for discrete nonlinear semi-Markov jump singularly perturbed models with mode-switching delay. Sci China Inf Sci, 2025, 68: 149201
6 Xie Y, Yu X, Shi Y, et al. SPT-based composite hierarchical antidisturbance control applied to a quadrotor UAV. IEEE Trans Ind Electron, 2023, 70: 635–645
7 Zhao J, Yang C, Gao W, et al. Reinforcement learning and optimal control of PMSM speed servo system. IEEE Trans Ind Electron, 2023, 70: 8305–8313
8 Kokotovic P V, Khalil H K, O'Reilly J. Singular Perturbation Methods in Control: Analysis and Design. Philadelphia: Society for Industrial and Applied Mathematics, 1999
9 Li Y, Wei L, Liu J, et al. Secure recursive state estimation for singularly perturbed discrete sequential systems under round-robin-like multichannel access policy. IEEE Trans Autom Sci Eng, 2025, 22: 3719–3730
10 Lei Y, Wang Y-W, Morărescu I-C, et al. Event-triggered fixed-time stabilization of two time scales linear systems. IEEE Trans Autom Control, 2023, 68: 1722–1729
11 Jiang Y, Jiang Z-P. Robust Adaptive Dynamic Programming. Hoboken: Wiley-IEEE Press, 2017
12 Wang D, Gao N, Liu D, et al. Recent progress in reinforcement learning and adaptive dynamic programming for advanced control applications. IEEE/CAA J Autom Sinica, 2024, 11: 18–36
13 Kiumarsi B, Vamvoudakis K G, Modares H, et al. Optimal and autonomous control using reinforcement learning: A survey. IEEE Trans Neural Netw Learn Syst, 2018, 29: 2042–2062
14 Cao W, Yang Q, Meng W, et al. Data-based robust adaptive dynamic programming for balancing control performance and energy consumption in wastewater treatment process, IEEE Trans Ind Informat, 2024, 20: 6622–6630
15 Han X, Zhao X, Xu X, et al. Trajectory tracking control for underactuated autonomous vehicles via adaptive dynamic programming, J Frankl Inst, 2024, 361: 474–488
16 Chakraborty S, Cui L, Ozbay K, et al. Automated lane changing control in mixed traffic: An adaptive dynamic programming approach, Transportation Research Part B: Methodological, 2024, 187: 103026.
17 Dong L, Tang Y, He H, et al. An event-triggered approach for load frequency control with supplementary ADP. IEEE Trans Power Syst, 2017, 32: 581–589
18 Yang X, He H, Zhong X. Adaptive dynamic programming for robust regulation and its application to power systems. IEEE Trans Ind Electron, 2018, 65: 5722–5732
19 Mu C, Tang Y, He H. Improved sliding mode design for load frequency control of power system integrated an adaptive learning strategy, 2017, 64: 6742–6751
20 Jiang Y, Jiang Z-P. Robust adaptive dynamic programming for large-scale systems with an application to multimachine power systems, IEEE Trans Circuits Syst II-Exp Briefs, 2012, 59: 693–697
21 Liu M, Liu Q, Zhang L, et al. Adaptive dynamic programming-based fault-tolerant attitude control for flexible spacecraft with limited wireless resources. Sci China Inf Sci, 2023, 66: 202201
22 Mu C, Wang K, Xu X, et al. Safe adaptive dynamic programming for multiplayer systems with static and moving no-entry regions. IEEE Trans Artif Intell 2024, 5: 2079–2092
23 Wang Z, Wang Y, Liang L. Optimal dynamic output-feedback controller design for linear output regulation problems with its applications. Sci China Inf Sci, 2025, 68: 182203
24 Wang D, Zhao M, Ha M, et al. Stability and admissibility analysis for zero-sum games under general value iteration formulation. IEEE Trans Neural Netw Learn Syst, 2023, 34: 8707–8718
25 Wei Q, Yang Z, Su H, et al. Online adaptive dynamic programming for optimal self-learning control of VTOL aircraft systems with disturbances. IEEE Trans Autom Sci Eng, 2024, 21: 343–352
26 Zhao J, Yang C, Gao W. Reinforcement learning based optimal control of linear singularly perturbed systems. IEEE Trans Circuits Syst II-Exp Briefs, 2022, 69: 1362–1366
27 Fu Z-J, Xie W-F, Rakheja S, et al. Adaptive optimal control of unknown nonlinear systems with different time scales. Neurocomputing, 2017, 238: 179–190
28 Shen H, Wang Y, Wang J, et al. A fuzzy-model-based approach to optimal control for nonlinear Markov jump singularly perturbed systems: A novel integral reinforcement learning scheme. IEEE Trans Fuzzy Syst, 2023, 31: 3734–3740
29 Mukherjee S, Bai H, Chakrabortty A. Reduced-dimensional reinforcement learning control using singular perturbation approximations. Automatica, 2021, 126: 109451
30 Liu X, Yang C, Zhou L, et al. Suboptimal reduced control of unknown nonlinear singularly perturbed systems via reinforcement learning. Int J Robust Nonlinear Control, 2021, 31: 6626–6645
31 Yang C, Zhong S, Liu X, et al. Adaptive composite suboptimal control for linear singularly perturbed systems with unknown slow dynamics. Int J Robust Nonlinear Control, 2020, 30: 2625–2643
32 Zhou L, Zhao J, Ma L, et al. Decentralized composite suboptimal control for a class of two-time-scale interconnected networks with unknown slow dynamics. Neurocomputing, 2020, 382: 71–79

33 Wang J, Peng C, Park J H, et al. Reinforcement learning-based near optimization for continuous-time Markov jump singularly perturbed systems. IEEE Trans Circuits Syst II-Exp Briefs, 2023, 70: 2026–2030
34 Liu X, Yang C, Luo B, et al. Suboptimal control for nonlinear slow-fast coupled systems using reinforcement learning and Takagi-Sugeno fuzzy methods. Int J Adapt Contr Signal Process, 2021, 35: 1017–1038
35 Zhao J, Yang C, Gao W, et al. Adaptive optimal output regulation of interconnected singularly perturbed systems with application to power systems. IEEE/CAA J Autom Sinica, 2024, 11: 1–13
36 Li J, Yang M, Lewis F L, et al. Compensator-based self-learning: Optimal operational control for two-time-scale systems with input constraints. IEEE Trans Ind Informat, 2024, 20: 9465–9475
37 Xue S, Luo B, Liu D. Event-triggered adaptive dynamic programming for unmatched uncertain nonlinear continuous-time systems. IEEE Trans Neural Netw Learn Syst, 2021, 32: 2939–2951
38 Yang X, Zhu Y, Dong N, et al. Decentralized event-driven constrained control using adaptive critic designs. IEEE Trans Neural Netw Learn Syst, 2022, 33: 5830–5844
39 Zhang Y, Zhao B, Liu D, et al. Adaptive dynamic programming-based event-triggered robust control for multiplayer nonzero-sum games with unknown dynamics. IEEE Trans Cybern, 2023, 53: 5151–5164
40 Zhao F, Gao W, Jiang Z-P, et al. Event-triggered adaptive optimal control with output feedback: An adaptive dynamic programming approach. IEEE Trans Neural Netw Learn Syst, 2021, 32: 5208–5221
41 Yan Y, Yang C, Ma L, et al. Asynchronous event-triggered H_∞ control problem for two-time-scale systems based on double-rate sampled-data. J Frankl Inst, 2022, 359: 8848–8868
42 Bhandari M, Fulwani D M, Gupta R. Event-triggered composite control of a two time scale system. IEEE Trans Circuits Syst II-Exp Briefs, 2018, 65: 471–475
43 Lewis F L, Vrabie D, Syrmos V L. Optimal Control. Hoboken: Wiley, 2012.
44 Zhou B, Duan G, Lin Z. A parametric Lyapunov equation approach to the design of low gain feedback. IEEE Trans Autom Control. 2008, 53: 1548–1554
45 Kleinman D. On an iterative technique for Riccati equation computations. IEEE Trans Autom Control, 1968, 13: 114–115
46 Zhao J, Yang C, Gao W, et al. Novel single-loop policy iteration for linear zero-sum games. Automatica, 2024, 163: 111551