SCIENCE CHINA
Information Sciences

• **Supplementary File** •

Policy iteration-based adaptive optimal control for Markov jump systems: a transition-probability-free asynchronous approach

Weidi CHENG[1,2], Chengcheng REN[1,2], Shuping HE[1,2,3*] & Changyin SUN[4]

[1]*Key Laboratory of Intelligent Computing and Signal Processing (Ministry of Education), School of Electrical Engineering and Automation, Anhui University, Hefei* 230601*, China;*
[2]*Anhui Engineering Laboratory of Human-Robot Integration System and Intelligent Equipment, School of Electrical Engineering and Automation, Anhui University, Hefei* 230601*, China;*
[3]*Information Materials and Intelligent Sensing Laboratory of Anhui Province, Hefei* 230601*, China;*
[4]*Engineering Research Center of Autonomous Unmanned System Technology (Ministry of Education), School of Artificial Intelligence, Anhui University, Hefei* 230601*, China*

Appendix A Abstract

This paper proposes a policy iteration-based adaptive optimal controller for a class of continuous-time Markov jump systems with hidden Markov model. Since the mismatched mode phenomenon between the system and the controller, the asynchronous control scheme is considered. The internal dynamics are not available in time, and therefore the coupled algebraic Riccati equations (CAREs) cannot be solved directly. To address the adaptive optimal control problem, we present a novel algorithm by utilizing off-policy adaptive dynamic programming techniques to iteratively tackle the CAREs. It leverages asynchronous information from the state and input, eliminating the need for a priori knowledge of the system matrix. By using the constructed discounted cost function, the information on the coupled transition probabilities is also not required. Also, the convergence of the designed algorithm has been demonstrated. Finally, an example via an RLC circuit system is provided to validate the effectiveness of the designed algorithm.

Appendix B Introduction and discussion

In general, minimizing a predefined cost function to ensure the performance is a key feature of the optimization algorithm. Such algorithms are usually implemented on the actor-critic mechanism [1], to acquire the optimal control policy of dynamic systems through the function approximator. In the ideal case, the internal dynamics of the system are available to achieve optimal control. However, in practice, the system dynamics might not be accurately identified. To handle this problem, Vrabie *et al*. [2] introduced an online method that relies on adaptive critics to get the optimal control solution of linear systems, in which only partial knowledge of the system dynamics is available. Then, Jiang and Jiang [3] designed a model-free adaptive dynamic programming (ADP) algorithm to iteratively solve the algebraic Riccati equation (ARE) by incorporating real-time data from both state and input. Luo *et al*. [4] derived a data-based policy gradient ADP algorithm relying on gathered data and the policy gradient framework. Yang *et al*. [5] developed a dynamic prioritized policy gradient ADP approach to address the optimal control challenges in nonaffine systems, with particular application to the management of wastewater treatment processes. Lian *et al*. [6] proposed a data-driven inverse reinforcement learning (RL) technique for redesigning an unknown index function, eliminating the need for prior knowledge of the control policy gain and the system dynamics. Within the domain of multi-robot systems, Huo *et al*. [7] devised an efficient game-theoretic decision-making framework to handle the task allocation challenge. Nevertheless, the optimal control problem for Markov jump systems (MJSs) has rarely been considered, especially for hidden MJSs with mismatched jumping mode phenomenon between the system and the controller.

In recent decades, MJSs have drawn extensive interest due to the outstanding modeling capabilities for network control systems [8], manufacturing systems [9], etc. For linear time-invariant systems, the optimal control objective design problem to be solved is to find a unique positive definite solution for the underlying ARE. Due to the presence of the transition probabilities in the Markovian process, the learning algorithms for handling the MJSs optimal control problem refers to address the coupled ARE (CARE) or the coupled Lyapunov matrix equation (CLME). In [10], a necessary and sufficient condition was provided to study the stochastic stability of MJSs in terms of the corresponding CLME. In [11], a parallel iterative algorithm for the solution of the CLME was developed. And an implicit sequential algorithm for handling CLME of MJS was given in [12]. By introducing a tuning parameter, an explicit iterative algorithm [13] was given to deal with the matrix equations under consideration. In [14], it presented an integral RL tracking algorithm for MJSs, which uses an offline tracking iteration algorithm to solve CARE and avoids the transition probability requirement. In [15], it deduced an offline parallel Kleinman algorithm and an online parallel integral RL algorithm to obtain the suboptimal controller. It's worth noting that the majority of current findings regarding the control of MJSs rely on a crucial assumption that the controller has constant access to complete plant mode information; this assumption can ensure the controller modes run synchronously with the system modes [16–18]. However, maintaining such an ideal condition exists challenging in practical applications.

Therefore, it is important to investigate the problem of imperfect mode detection for MJSs. In recent years, Wang *et al*. [19] derived a stochastic scheduled controller to solve the almost surely exponential stabilization problem in which the controller operates

* Corresponding author (email: shuping.he@ahu.edu.cn)

sequentially. Furthermore, an asynchronous controller based on a stochastic transfer matrix scheme was constructed to achieve exponential almost sure stability of MJSs in [20]. Moreover, the hidden Markov model (HMM) is one of the effective tools to handle this problem. In this way, the system mode is identified via a mode-detector. The detected mode is associated with the system mode through conditional probabilities, but it may not be equivalent to the real mode. Some works on HMM-based asynchronous control can be found in [21–25]. However, to the authors' knowledge, the problem of asynchronous optimal control of HMM-based MJSs has not been adequately addressed. There are two main challenges to be addressed during the optimal control algorithm process for hidden MJSs, that is:

1. By the HMM, a conditional probability matrix (CPM) is used to represent the jump probability of a detected mode. This implies the existence of a novel coupling term in the CAREs associated with the hidden MJSs. Then, how do we solve this novel CAREs?

2. Due to the asynchronous characteristic between the system mode and the controller mode, it is difficult for the controller to obtain the real system mode information. This requires considering how to use online measured state and input information to iteratively design optimal control laws.

In this paper, we study a policy iteration-based adaptive optimal control scheme for MJSs by transition-probability-free asynchronous approach. The main contributions are summarized in the following:

1. An asynchronous infinite horizon performance index is established, where the jumping of the weight matrix and the control policy are based on the detected mode under conditional probability. The traditional performance index for the synchronous Markov model in [12–15] can be considered as a special case of this work.

2. An asynchronous off-policy RL technique is presented to address the CAREs iteratively and obtain the optimal control gains by employing the online measured state and input information without requiring the system matrix and coupled transition probability information.

3. A model-free asynchronous adaptive optimal stabilization control approach for continuous-time hidden MJSs is presented to ensure that the system output achieves stability. This is the first attempt to design the optimal control strategy for hidden MJSs, to the best of our knowledge.

Notation: For the probability space $(\Omega, \Upsilon, \mathrm{H})$, Υ is the σ-algebra of events, H is the probability measure, and Ω is the sample space. $\mathbb{R}$ denotes the set of real numbers, $\mathbb{Z}_+$ denotes the set of non-negative integers. $\|\cdot\|$ represents the Euclidean norm of a vector. The Kronecker product operator is denoted by $\otimes$. $vec(\bar{\zeta}) = [\zeta_1^T, \zeta_2^T, \ldots, \zeta_n^T]^T$ is a vector-valued function of a matrix $\bar{\zeta}$. I_n represents n-by-n identity matrix.

Appendix C Implementation of asynchronous policy iteration algorithm

The online implementation of the policy iteration approach can be achieved as:

$$\begin{aligned} x(t+o(t))^T P_l^{(\beta)} x(t+o(t)) - x(t)^T P_l^{(\beta)} x(t) \\ &= \int_t^{t+ot} [x^T (\tilde{A}_i^{(\beta)T} P_l^{(k)} + P_l^{(\beta)} \tilde{A}_i^{(\beta)}) x + 2(u_l + K_l^{(\beta)} x)^T B_i^T P_l^{(\beta)} x] d\tau \\ &= -\int_t^{t+o(t)} x^T \bar{\mathcal{Q}}_l^{(\beta)} x d\tau + 2\int_t^{t+o(t)} (u_l + K_l^{(\beta)} x)^T R_l K_l^{(\beta+1)} x d\tau \end{aligned} \tag{C1}$$

where $\bar{\mathcal{Q}}_l^{(\beta)} = Q_l + K_l^{(\beta)T} R_l K_l^{(\beta)} + \gamma \ln \mu P_l^{(\beta-1)}$.

Remark 1. The term $B_i^T P_l^{(\beta)}$ involving B_i is replaced by $R_l K_l^{(\beta+1)}$, and the relevant algorithm can be implemented without requiring the system matrices A_i and B_i.

Furthermore, to find $P_l^{(\beta)}$ and $K_l^{(\beta+1)}$, for $l \in \Lambda$, the following two operators are given, that is, $x \in \mathbb{R}^n \to \bar{\chi} \in \mathbb{R}^{\frac{1}{2}n(n+1)}$, $P_l \in \mathbb{R}^{n\times n} \to \bar{\Gamma}_l \in \mathbb{R}^{\frac{1}{2}n(n+1)}$ with

$$\begin{cases} \bar{\Gamma}_l = [p_{l11}, 2p_{l12}, \ldots, 2p_{l1n}, p_{l22}, 2p_{l23}, \ldots, 2p_{l(n-1)n}, p_{lnn}]^T \\ \bar{\chi} = [x_1^2, x_1x_2, \ldots, x_1x_n, x_2^2, x_2x_3, \ldots, x_{n-1}x_n, x_n^2]^T \end{cases}$$

According to the Kronecker product representation, the terms on the right-hand side of (C1) satisfies

$$x^T \bar{\mathcal{Q}}_l^{(\beta)} x = (x^T \otimes x^T) vec(\bar{\mathcal{Q}}_l^{(\beta)}) \tag{C2}$$

$$(u_l + K_l^{(\beta)} x)^T R_l K_l^{(\beta+1)} x = \left[(x^T \otimes x^T)(I_n \otimes K_l^{(\beta)T} R_l) + (x^T \otimes u_l^T)(I_n \otimes R_l)\right] vec(K_l^{(\beta+1)}) \tag{C3}$$

Then, we define $\sigma_{xx} \in \mathbb{R}^{\varepsilon \times \frac{1}{2}n(n+1)}$, $\xi_{xx} \in \mathbb{R}^{\varepsilon \times n^2}$, and $\xi_{xu} \in \mathbb{R}^{\varepsilon \times mn}$ where ε is a positive integer, and get

$$\begin{cases} \sigma_{xx} = [(\bar{\chi}_{t_1} - \bar{\chi}_{t_0}, \bar{\chi}_{t_2} - \bar{\chi}_{t_1}, \ldots, \bar{\chi}_{t_\varepsilon} - \bar{\chi}_{t_{\varepsilon-1}}]^T \\ \xi_{xx} = [\int_{t_0}^{t_1} x \otimes x d\tau, \int_{t_1}^{t_2} x \otimes x d\tau, \ldots, \int_{t_{\varepsilon-1}}^{t_\varepsilon} x \otimes x d\tau]^T \\ \xi_{xu} = [\int_{t_0}^{t_1} x \otimes u_l d\tau, \int_{t_1}^{t_2} x \otimes u_l d\tau, \ldots, \int_{t_{\varepsilon-1}}^{t_\varepsilon} x \otimes u_l d\tau]^T \end{cases} \tag{C4}$$

where $0 \leqslant t_0 < t_1 < \cdots < t_\varepsilon$. By substituting (C2)-(C4) into (C1), we obtain the following compact linear form:

$$\Phi_l^{(\beta)} \begin{bmatrix} \bar{\Gamma}_l^{(\beta)} \\ vec(K_l^{(\beta+1)}) \end{bmatrix} = \Psi_l^{(\beta)} \tag{C5}$$

where $\Phi_l^{(\beta)} \in \mathbb{R}^{\varepsilon \times [\frac{1}{2}n(n+1)+mn]}$ and $\Psi_l^{(\beta)} \in \mathbb{R}^{\varepsilon}$ given by:

$$\begin{cases} \Phi_l^{(\beta)} = [\sigma_{xx}, -2\xi_{xx}(I_n \otimes K_l^{(\beta)T} R_l) - 2\xi_{xu}(I_n \otimes R_l)] \\ \Psi_l^{(\beta)} = -\xi_{xx} vec(\bar{\mathcal{Q}}_l^{(\beta)}) \end{cases} \tag{C6}$$

Inspired by [3], to maintain the persistence of excitation condition, the rank condition $rank([\xi_{xx}, \xi_{xu}]) = \frac{n(n+1)}{2} + mn$ needs to be fulfilled to guarantee the existence and uniqueness of solution to (C5). When the full column rank condition is met, the solution to equation (C5) is given by:

$$\begin{bmatrix} \bar{\Gamma}_l^{(\beta)} \\ vec(K_l^{(\beta+1)}) \end{bmatrix} = (\Phi_l^{(\beta)T} \Phi_l^{(\beta)})^{-1} \Phi_l^{(\beta)T} \Psi_l^{(\beta)} \tag{C7}$$

Then, we have the transition-probability-free adaptive asynchronous optimal control Algorithm 1.

Remark 2. Due to the asynchronous property between the system mode and the controller mode, the information of the system mode available to the controller is often imprecise. Therefore, in (4), the term $x^T(\tilde{A}_i^{(\beta)T} P_l^{(\beta)} + P_l^{(\beta)} \tilde{A}_i^{(\beta)})x$ is replaced by $x^T \bar{\mathcal{Q}}_l^{(\beta)} x$, the term $B_i^T P_l^{(\beta)}$ is replaced by $R_l K_l^{(\beta+1)}$. After these transformations, both A_i and B_i do not appear explicitly in the policy update. The necessity for the system matrices can be substituted by the online measurement of the state and the input data.

Remark 3. It should also be noted that in (4), the requirement of the coupled transition probability $\phi_{ij}\tau_{il}$ is avoided. By introducing the discount factor μ, the term $\gamma \ln \mu P_l^{(\beta-1)}$ in (4) contains information about the coupled transition probability $\sum_{j=1}^{\mathcal{N}} \sum_{l=1}^{\mathcal{M}} \phi_{ij}\tau_{il} P_j$, that is, P_l, for the previous step contains information about the coupled transition probability updated by $\bar{\mathcal{Q}}_l^{(\beta)}$ iterations.

Appendix D Simulation study

In this section, a practical example via a RLC circuit system [26,27] is shown in the following Fig. D1. We present three simulation cases to illustrate the effectiveness of the proposed algorithm, where Case 1 and Case 2 are used to exhibit the generality of the proposed algorithm in both asynchronous and synchronous situations, Case 3 further expresses the practicality and universality of the algorithm by considering the asynchronous case of a different transfer probability matrix.

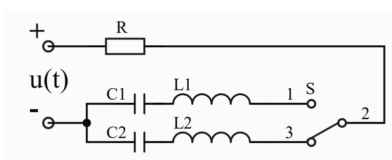

Figure D1 The RLC circuit.

The switch S has two positions with randomly transitions. Assume a Markov process with hidden system modes $\Lambda = \{1, 2\}$ is driving the state transitions between the two modes $\Xi = \{1, 2\}$.

Given that I_L represents the current of inductance L, u_C is the voltage of capacitance C, apply the Kirchhoff laws for $i = 1, 2$, and get

$$\frac{dI_L}{dt} = \frac{u_L}{L_i} = \frac{u - u_C - I_L}{L_i}; \quad \frac{du_C}{dt} = \frac{I_L}{C_i} \tag{D1}$$

Then, letting $x = [u_C, I_L]^T$, the state space model is

$$\dot{x} = \begin{bmatrix} 0 & \frac{1}{C_i} \\ -\frac{1}{L_i} & -\frac{R}{L_i} \end{bmatrix} x + \begin{bmatrix} 0 \\ \frac{1}{L_i} \end{bmatrix} u \tag{D2}$$

where the parameters of the RLC circuit system are given as $L_1 = 0.5H$, $C_1 = 5F$, $L_2 = 0.8H$, $C_2 = 8F$ and $R = 1.5\Omega$.

Regarding the parameters of the proposed asynchronous optimal algorithms, the discount factor μ and the parameter γ are given by $\mu = 0.5$, $\gamma = 0.1$, respectively. The state and the input are gathered at intervals of $0.01s$. The recursive iteration begins at $t = 2s$. Before that, the exploration noise θ is given by $\theta = \sum_{i=1}^{100} \sin(\varpi_i t)$ as the system input, where ϖ_i is randomly chosen from $[-50, 50]$. Moreover, the stopping criterion is established as $\|P_l^{(\beta)} - P_l^{(\beta-1)}\| \leqslant 1 \times 10^{-12}$. After $t = 2s$, the asynchronous control law obtained after iteration serves as the actual control signal to the system until the end of the simulation.

Moreover, the weighting matrices are given as $Q_1 = \text{diag}[1\ 1]$, $Q_2 = \text{diag}[1\ 0.1]$, $R_1 = 0.5$ and $R_2 = 1$. The initial control gain matrix is considered as $K_l^{(0)} = [1\ 1]$. To verify the applicability of the developed algorithm in both asynchronous and synchronous cases, two scenarios are considered in the simulation results.

Case 1 *(Asynchronous case):* The transition probability matrix Π is given by $\Pi = \begin{bmatrix} -4 & 4 \\ 5 & -5 \end{bmatrix}$, and the observation probability matrix Θ is given by $\Theta = \begin{bmatrix} 0.5 & 0.5 \\ 0.4 & 0.6 \end{bmatrix}$. By applying Algorithm 1, after 12th iterations, one can get

$$P_1^{(12)} = \begin{bmatrix} -16.8816 & -1.6630 \\ -1.6630 & 0.0838 \end{bmatrix}, \ K_1^{(12)} = \begin{bmatrix} -8.6052 & 0.2800 \end{bmatrix}, P_2^{(12)} = \begin{bmatrix} -9.1488 & -0.0272 \\ -0.0272 & 0.0038 \end{bmatrix}, \ K_2^{(12)} = \begin{bmatrix} -0.1436 & 0.0086 \end{bmatrix}.$$

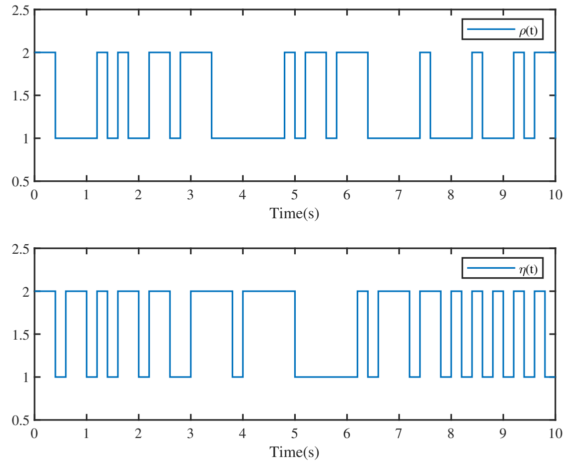

Figure D2 The modes of $\rho(t)$ and $\eta(t)$ modes in Case 1.

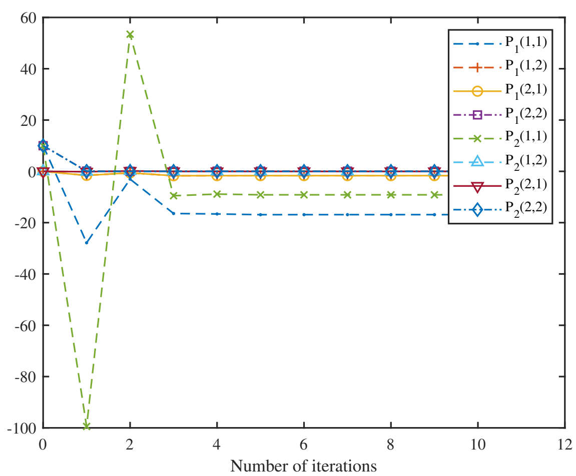

Figure D3 Convergence of P_1 and P_2 in Case 1.

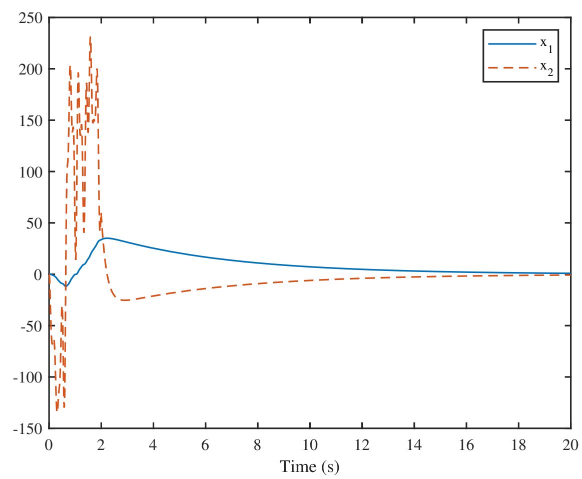

Figure D4 The system states in Case 1.

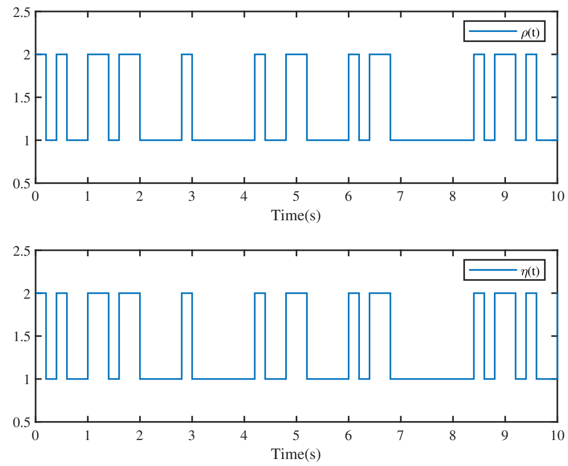

Figure D5 The modes of $\rho(t)$ and $\eta(t)$ modes in Case 2.

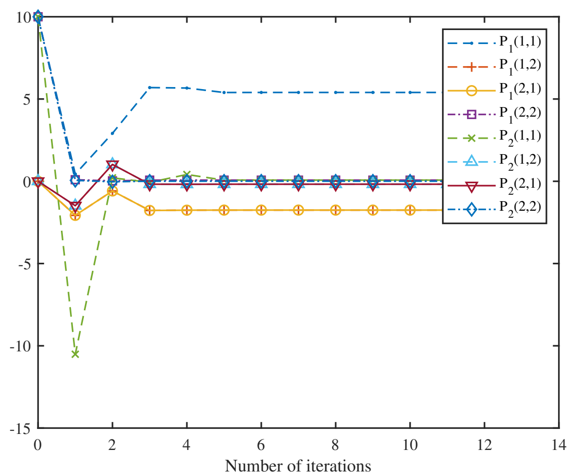

Figure D6 Convergence of P_1 and P_2 in Case 2.

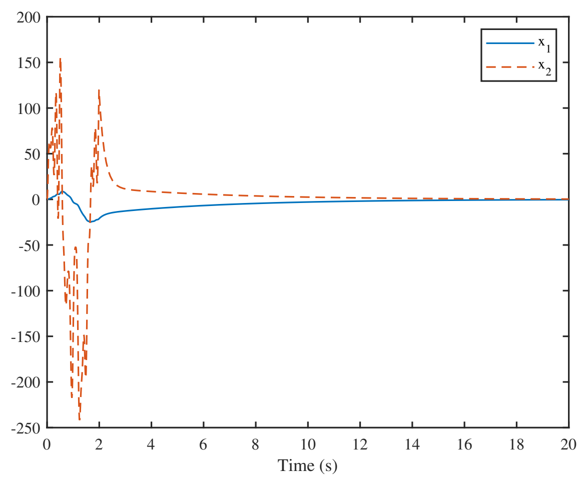

Figure D7 The system states in Case 2.

The asynchronous properties between the system mode $\rho(t)$ and the controller mode $\eta(t)$ are exhibited in Fig. D2. Further, Fig. D3 exhibits the iteration process of the matrix parameters P_1 and P_2 by CAREs, revealing evident convergence throughout the iterative learning process. This illustrates that Algorithm 1 successfully achieves the optimal asynchronous control solution. In Fig. D4, the system state trajectory is depicted with the initial state matrix $x_0 = [0.2\ 0.1]^T$, showing the ideal control performance and achieving system stabilization.

Case 2 *(Synchronous case):* In synchronization case, we give the relevant observation probability matrix Θ as $\Theta = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, and the transition probability matrix Π is still given by $\Pi = \begin{bmatrix} -4 & 4 \\ 5 & -5 \end{bmatrix}$. By applying Algorithm 1, after 14th iterations, one can get:

$$P_1^{(14)} = \begin{bmatrix} 5.3993 & -1.7516 \\ -1.7516 & 0.0702 \end{bmatrix},\ K_1^{(14)} = \begin{bmatrix} -4.1123 & 0.2217 \end{bmatrix}, P_2^{(14)} = \begin{bmatrix} 0.0699 & -0.1779 \\ -0.1779 & 0.0051 \end{bmatrix},\ K_2^{(14)} = \begin{bmatrix} -0.2737 & 0.0089 \end{bmatrix}.$$

From Fig. D5, it can be observed that the system mode $\rho(t)$ and the controller mode $\eta(t)$ are synchronous. From Fig. D6 and Fig. D7, the convergence of P_1 and P_2 are provided, and the developed optimal control scheme stabilize the system. It shows that the presented adaptive optimal Algorithm 1 can also be adapted in synchronous case.

Case 3 *(Asynchronous case of another transition probability matrix):* In this case, The transition probability matrix Π is given by $\Pi = \begin{bmatrix} -1 & 1 \\ 3 & -3 \end{bmatrix}$. The observation probability matrix Θ is given by $\Theta = \begin{bmatrix} 0.5 & 0.5 \\ 0.4 & 0.6 \end{bmatrix}$, which is equal to Case 1. By applying

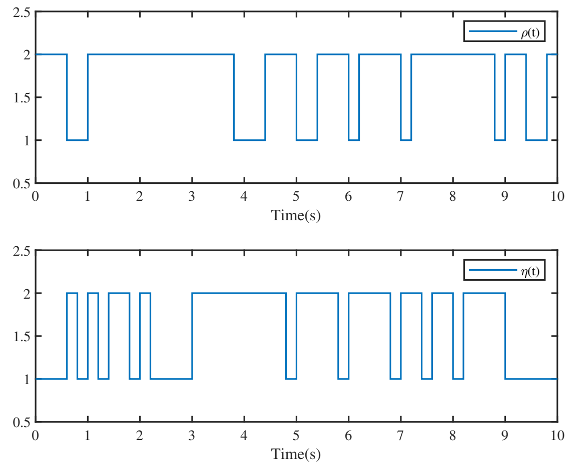

Figure D8 The modes of $\rho(t)$ and $\eta(t)$ modes in Case 3.

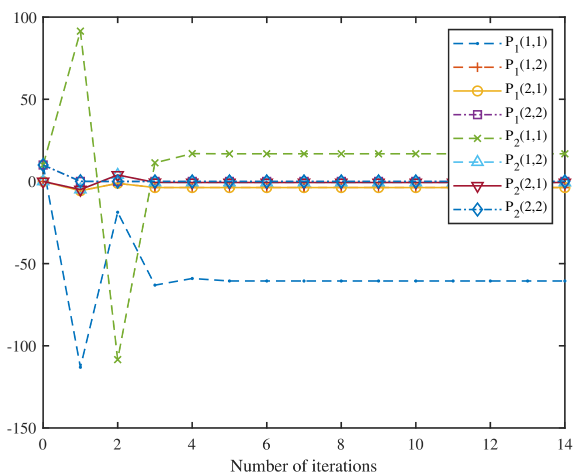

Figure D9 Convergence of P_1 and P_2 in Case 3.

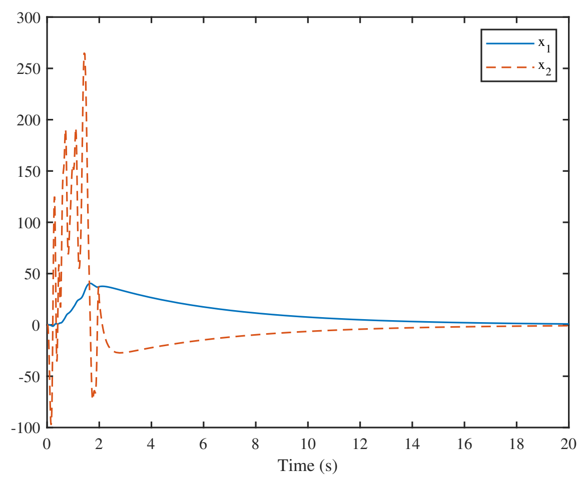

Figure D10 The system states in Case 3.

Algorithm 1, after 15th iterations, one can get:

$$P_1^{(15)} = \begin{bmatrix} -60.5695 & -3.7249 \\ -3.7249 & 0.0166 \end{bmatrix},\ K_1^{(15)} = \begin{bmatrix} -9.3594 & 0.0420 \end{bmatrix}, P_2^{(15)} = \begin{bmatrix} 16.8285 & -0.6444 \\ -0.6444 & 0.0069 \end{bmatrix},\ K_2^{(15)} = \begin{bmatrix} -1.7679 & 0.0133 \end{bmatrix}.$$

From Figs. D8-D10 it can be seen that different transfer probability matrices do not affect the effectiveness of the algorithm, the practicality and universality of the presented algorithm are verified.

Remark 4. Based on the results of the simulation study, it can be seen that the proposed asynchronous optimal control scheme achieves the desired control performance regardless of whether the system model and controller model are asynchronous or synchronous. In this study, an asynchronous infinite horizon performance index is established, where the jumping of the weight matrix and the control policy are based on the detected mode under conditional probability, which means that the traditional performance index for the synchronous Markov model can be considered as a special case of this work.

According to the observed simulation results, the developed algorithm achieves convergence and obtains the optimal matrix P_l and K_l. And there is not required a prior information about the system dynamics A_i, B_i, $i \in \Xi$ and the coupled transition probabilities $\phi_{ij}\tau_{il}$, $l \in \Lambda$. It also exhibits the effectiveness of our constructed adaptive asynchronous control scheme and policy policy iteration algorithm.

References

1 Miller W T, Thomas R S, Werbos P J. A Menu of designs for reinforcement learning over time. Neural Networks for Control, 1995, 67–95.

2 Vrabie D, Pastravanu O, Abu-Khalaf M, et al. Adaptive optimal control for continuous-time linear systems based on policy iteration. Automatica, 2009, 45: 477–484.

3 Jiang J, Jiang Z P. Computational adaptive optimal control for continuous-time linear systems with completely unknown dynamics. Automatica, 2012, 48: 2699–2704.

4 Luo B, Liu D, Wu H N, et al. Policy gradient adaptive dynamic programming for data-based optimal control. IEEE Trans Cybern, 2017, 47: 3341–3354.
5 Yang R, Wang D, Qiao J. Policy gradient adaptive critic design with dynamic prioritized experience replay for wastewater treatment process control. IEEE Trans Ind Inf, 2022, 18: 3150–3158.
6 Lian B, Xue W, Xie Y, et al. Off-policy inverse Q-learning for discrete-time antagonistic unknown systems. Automatica, 2023, 155: 111171.
7 Huo X, Zhang H, Wang Z, et al. Task allocation with minimum requirements for multiple mobile robot systems: a game-theoretical approach. IEEE Trans Netw Sci Eng, 2024, 11: 1202–1213.
8 Cheng P, He S, Luan X, et al. Finite-region asynchronous H_∞ control for 2D Markov jump systems. Automatica, 2021, 129: 109590.
9 Song J, Wang Y K, Niu Y, et al. Periodic event-triggered terminal sliding mode speed control for networked PMSM system: A ga-optimized extended state observer approach. IEEE/ASME Trans Mechatron, 2022, 27: 4153–4164.
10 Ji Y, Chizeck H J. Controllability, stabilizability, and continuous-time Markovian jump linear quadratic control. IEEE Trans Automat Contr, 1990, 35: 777–788.
11 Borno I. Parallel computation of the solutions of coupled algebraic Lyapunov equations. Automatica, 1995, 31: 1345–1347.
12 Qian Y Y, Pang W J. An implicit sequential algorithm for solving coupled Lyapunov equations of continuous-time Markovian jump systems. Automatica, 2015, 60: 245–250.
13 Sun H J, Zhang Y, Liu W. Explicit iterative algorithms for continuous coupled Lyapunov matrix equations. IEEE Trans Automat Contr, 2020, 65: 3631–3638.
14 Zhang K, Zhang H G, Cai Y, et al. Parallel optimal tracking control schemes for mode-dependent control of coupled Markov jump systems via integral RL method. IEEE Trans Autom Sci Eng, 2020, 17: 1332–1342.
15 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 Sys II, 2023, 70: 2026–2030.
16 Ren H, Wang Y, Liu M, et al. An optimal estimation framework of multi-agent systems with random transport protocol. IEEE Trans Signal Process, 2022, 70: 2548–2559.
17 Zhang X, He S, Stojanovic V, et al. Finite-time asynchronous dissipative filtering of conic-type nonlinear Markov jump systems. Sci China Inf Sci, 2021, 64: 152206.
18 Fan Z X, Li X, Liu R. ADP-based optimal control for systems with mismatched disturbances: A PMSM application. IEEE Trans Circuits Sys II, 2023, 70: 2057–2061.
19 Wang G, Sun Y. Almost sure stabilization of continuous-time jump linear systems via a stochastic scheduled controller. IEEE Trans Cybern, 2020, 52: 2712–2724.
20 Wang G, Xu L. Almost sure stability and stabilization of Markovian jump systems with stochastic switching. IEEE Trans Automat Contr, 2021, 67: 1529–1536.
21 Wu Z G, Shi P, Shu Z, et al. Passivity-based asynchronous control for Markov jump systems. IEEE Trans Automat Contr, 2017, 62: 2020–2025.
22 Ren H, Ma H, Li H, et al. Adaptive fixed-time control of nonlinear mass with actuator faults. IEEE/CAA J Autom Sinica, 2023, 10: 1252-1262.
23 Ren C, He S, Luan X, et al. Finite-time L2-gain asynchronous control for continuous-time positive hidden Markov jump systems via T-S fuzzy model approach. IEEE Trans Cybern, 2021, 51: 77–87.
24 Tao Y Y, Wu Z G, et al. Asynchronous stabilization for hidden Markov jump linear systems with complex transition probabilities. Automatica, 2023, 156: 111200.
25 Cheng P, Wu D, He S, et al. Asynchronous control for 2-D Markov jump cyber-physical systems against aperiodic denial-of-service attacks. Sci China Inf Sci, 2023, 66: 172204.
26 Wang G, Zhang Q. Robust control of uncertain singular stochastic systems with Markovian switching via proportional-derivative state feedback. IET Control Theory A, 2012, 6: 1089–1096.
27 Tian Y, Yan H, Dai W, et al. Observed-based asynchronous control of linear semi-Markov jump systems with time-varying mode emission probabilities. IEEE Trans Circuits Sys II, 2020, 67: 3147–3151.