Optimal $\mathcal{H}_\infty$ Event-Triggered Control for Markov Jump Linear Systems Based on Actor-Critic Q-Learning Approach

Zhaowen Xu, *Member, IEEE*, and Zheng-Guang Wu

***Abstract*—In this paper, the optimal $\mathcal{H}_\infty$ control problem is investigated for a class of Markov jump linear systems, under event-triggering mechanism. To this end, a two-player zero-sum differential game-theoretic problem is formulated and a mode-dependent event-triggered version of Hamiltonian function is established and compared with its time-triggered counterpart. Based on this, the corresponding triggering event is given to ensure the stochastic stability of the considered system with a predefined $\mathcal{H}_\infty$ noise attenuation level index and the existence of a saddle point for the defined cost function is proved. The difference in the optimal cost between the time-triggered case and the event-triggered one is also shown. Further to derive an online solution without the full knowledge of the system model, a Q-learning framework is adopted to approximate the optimal cost and the optimal event-triggered control law combined with an actor-critic network. The error dynamics of the learning algorithm are analyzed and the convergence is proved. The stochastic stability of the resulting closed-loop system is also confirmed. Finally, a numerical example is given to show the effectiveness of the proposed results.**

***Note to Practitioners*—The wide applications of Markov jump system have gained its reputation in modeling systems with random changes from internal structure or external parameters. This work is motivated by the optimal $\mathcal{H}_\infty$ event-triggered control problem for Markov jump system. Existing results are mostly based on linear matrix inequality approach which is strictly model-based and offline thus vulnerable to modeling uncertainties. The optimality of the event-triggered controller has rarely been discussed. In this paper we adopt a Q-learning based actor-critic approach to perform the optimal $\mathcal{H}_\infty$ event-triggered controller design with respect to a user-defined triggering condition. The stability of the closed-loop system is proved as well as the convergence of the learning algorithm. It is data-based and easy to be implemented online without the full knowledge of the system parameters. Examples are given in the end to demonstrate the performance of the obtained results. In future research, network-induced disturbances such as time-delays, malicious attacks will be considered and discussed under the proposed method.**

Received 10 February 2025; revised 7 May 2025 and 22 May 2025; accepted 23 May 2025. Date of publication 26 May 2025; date of current version 10 June 2025. This article was recommended for publication by Associate Editor C. M. Abdissa and Editor S. Takai upon evaluation of the reviewers' comments. *(Corresponding author: Zheng-Guang Wu.)*

Zhaowen Xu is with the School of Information Science and Technology, Hangzhou Normal University, Hangzhou, Zhejiang 311121, China (e-mail: zhaowen_xu@outlook.com).

Zheng-Guang Wu is with the Institute of Intelligence Science and Engineering, Shenzhen Polytechnic University, Shenzhen 518055, China, and also with the National Laboratory of Industrial Control Technology, Institute of Cyber-Systems and Control, Zhejiang University, Hangzhou, Zhejiang 310027, China (e-mail: nashwzhg@zju.edu.cn).

Digital Object Identifier 10.1109/TASE.2025.3573908

***Index Terms*—Markov jump system, optimal event-triggered control, actor-critic network.**

I. Introduction

MARKOV jump system (MJS) is a class of hybrid systems capable of modeling system with abrupt changes either in system internal structure or external parameters. It has been widely acknowledged in areas such as power system, manufacturing process, network system [1], [2], [3], [4], [5], [6]. Up to now much progress in control area has been achieved, from fundamental control and filtering theory for nominal MJS [7], [8], to stability analysis and control synthesis when external disturbances or intrinsic modeling uncertainties are considered [9], [10]. Also the asynchronous control framework for the case of only partial mode information [11], [12], [13] has been constantly discussed. However, the results obtained are mainly based on the full knowledge of system model. They need to be computed offline and are vulnerable to uncertainties from system modeling error, which makes them not applicable in complex and fast change scenarios with high real-time demand like motion planning for autonomous driving or for robots in manufacturing. Recently, the reinforcement learning (RL) has been an intensively explored area [14], [15], [16], [17], [18]. From the perspective of control community, the word RL is interchangeable with adaptive/approximate dynamic programming (ADP), as is put in [19] and [20]. It maps a relationship between optimal control policy and system data with a given value function, which differs from traditional model-based control methods. Motivated by the efficiency of RL control design for unknown system, as well as its optimal performance, RL has attracted considerable attention in the field of MJSs in recent years. For example, in [21] a transition-probability-free optimal tracking control scheme for coupled MJS has been proposed by utilizing the integral RL. The stability and convergence has also been analyzed. In [22] the optimal regulation problem of the nonlinear semi-MJSs is addressed with state and input delays in an actor-critic neural network (NN) framework. The RL-based resilient control algorithm is investigated for a class of MJSs with completely unknown transition probability information in [23]. The secure control problem for discrete-time Markov jump cyber-physical systems with malicious attacks is studied in [24] based on a proposed hybrid learning scheme. In [25], the nonlinear MJSs are approximated by an identifier and the optimal control problem is discussed and solved based on

the adaptive dynamic programming technique. It can be seen that the RL method shows a potential application for solving control problems of MJSs. Unlike deterministic systems, the jumping parameter sampled from an underlying Markov process may present additional challenges over a traditional linear quadratic regulator control problem, which motivates our present work.

On the other hand, in modern control systems the communication channels are often shared with multiple sensors, actuators and controllers. Limited resources of bandwidth and energy will delay or fail the transmission of data. Event-triggered control strategy provides a wise way to determine the transmission of data based on a user designed triggering condition, making it possible to achieve the decongestion of network channels and energy saving for devices [26], [27], [28], [29]. Although event control problem is often discussed in the framework of reinforcement learning for nonlinear system so as to make its applications wide [30], [31], [32], there has been few reports on the optimal event-triggered control analysis and synthesis of MJSs. Usually this problem is investigated under various scenarios with the system model being as known a priori [33], [34], [35], [36], [37] where a series of linear matrix inequalities are involved. Yet the optimality of event triggered controller with respect to user defined triggering condition is still an open problem in the field of MJSs and has not fully been discussed, especially in the absence of the full knowledge of the system.

In this paper, we aim to design a model free algorithm for optimal event-triggered control of MJSs that can be implemented online. The main contributions are summarized as follows:

1) The optimal $\mathcal{H}_\infty$ event-triggered control problem is formulated for MJSs and tackled in this paper. A model-based parallel iteration algorithm and a model-free RL algorithm have been presented. Numerical examples are given to illustrate the effectiveness of the obtained results.
2) Compared with the works in [23], [38], and [39] where partial system matrix is still needed for designing controllers, here we utilizing a critic-actor network structure to implement a Q-learning algorithm, such that the control laws can be learned online and in a model-free fashion.
3) Unlike the commonly used methods where the performance index is optimized by solving a series of constraints expressed in terms of linear matrix inequalities (LMIs) [33], [34], [35], [36], [37], here the optimality of the mode dependent event-triggered controller with respect to the designed triggering event has been discussed under a Hamiltonian-driven ADP framework [26], and the difference of the optimal cost between the time triggered case and the event triggered one is also shown and proved.

Notations: The notations used throughout this paper are standard unless otherwise stated. Some of the symbols that easily cause confusion are listed in the following.

NOTATION AND CONVENTIONS

$\overline{\lambda}(A)$, $\underline{\lambda}(A)$	maximum, minimum eigenvalue of A
$\mathrm{tr}(A)$	trace of matrix A
$\mathbb{R}^+$	set of positive real numbers
$\|\cdot\|$	Euclidean norm
$\|x\|_R^2$	interchangeable with x^TRx for symmetric matrix $R > 0$
$\otimes$	Kronecker product
x	state of the considered system
$\hat{x}$	sampled state feedback signal
u	control input in general
u_c	control input in the time-triggered case
u_d	control input in the event-triggered case
$(\cdot)^*$	optimal value for $(\cdot)$
$(\cdot)^+$	right limit of $(\cdot)$
$\mathrm{vech}(A)$	half-vectorization of symmetric $n \times n$ matrix A obtained by vectorizing the lower or upper triangular part of A
$\mathcal{D}$	initial system state distribution
ρ	initial mode distribution
$\mathcal{L}_2[0,\infty)$	space of square-integrable vector function over $[0,\infty)$
J	infinite horizon cost functional
V_i^*	mode-dependent optimal value function
$\mathscr{H}_i$	mode-dependent Hamiltonian function
$\mathbb{E}$	mathematical expectation operator
$\mathscr{L}$	infinitesimal operator

II. PRELIMINARIES AND PROBLEM FORMULATION

Consider the following continuous-time Markov jump linear system:

$$\begin{cases} \dot{x}(t) &= A(r(t))x(t) + B(r(t))u(t) + D(r(t))w(t) \\ & x_0 \sim \mathcal{D}, r_0 \sim \rho \end{cases} \tag{1}$$

where $x(t) \in \mathbb{R}^{n_x}$ denote the system state. $u(t) \in \mathbb{R}^{n_u}$ is the control input, which shall be designed in this paper. $w(t)$ is the external disturbances belonging to $\mathcal{L}_2[0,\infty)$. $\{r(t), t \geq 0\}$ refers to a Markov process indicating the system operation mode. It takes values in a finite set $\mathcal{N} = \{1, 2, \ldots, N\}$. The transition probability is defined by

$$\mathbf{Pr}\{r(t+h) = j | r(t) = i\} = \begin{cases} \pi_{ij}h + o(h) & i \neq j \\ 1 - \pi_{ii}h + o(h) & i = j \end{cases} \tag{2}$$

where π_{ij} is the transition rate and satisfies $\sum_{j=1, j\neq i}^{N} \pi_{ij} = -\pi_{ii}$ and $\pi_{ij} > 0$. For notational convenience, the matrices associated with $r(t) = i$ will be written with a subscript i, say $A(r(t))$ will be written as A_i. Also the time tag t is omitted where needed if no ambiguity arises.

The initial distribution of r_0 is given by $\rho = [\rho_1, \rho_2, \ldots, \rho_N]$. It is assumed that the states and the system modes are fully accessible at any time. The expected signal to be evaluated is set to be $z(t)$ with

$$\mathbb{E}\{\|z(t)\|^2\} = \mathbb{E}\{x^T H_i x + u^T R_i u\} \tag{3}$$

where $H_i > 0$ and $R_i > 0$. Further we assume that $\mathbb{E}_{x_0\sim\mathcal{D}}\{x_0^T x_0\} > 0$ and $\rho_i > 0$. It shows that the covariance

of the initial state is full rank and the initial operation mode is arbitrary. As is put in [40], these assumptions are quite standard for learning-based control and can be informally interpreted as the persistently excitation condition. Before further presentation, the following definitions are introduced from [41] for later analysis.

Definition 1: For a given Lyapunov functional candidate $V_{r(t)}(x(t))$, its infinitesimal operator $\mathscr{L}$ is defined as

$$\mathscr{L}V_{r(t)}(x(t)) = \lim_{\delta\to 0}\frac{\mathbb{E}\{\hat{V}(t+\delta|x(t),r(t))\} - V_i(x(t))}{\delta} \tag{4}$$

where $\hat{V}(t+\delta|x(t),r(t)) = V_{r(t+\delta)}(x(t+\delta)|x(t),r(t)=i)$.

Definition 2: The considered continuous-time Markov jump linear system (1) with $w(t) = 0$ is said to be stochastically stable if

$$\mathbb{E}_{x_0\sim\mathcal{D},r_0\sim\rho}\left\{\int_0^\infty \|x(t)\|^2 dt\right\} \le \infty. \tag{5}$$

Definition 3: The considered continuous-time Markov jump linear system (1) is said to have an $\mathcal{H}_\infty$ noise attenuation level index less or equal than γ, if it satisfies Definition 1 and the following holds under zero initial condition with nonzero $w(t)$,

$$\mathbb{E}\left\{\int_t^\infty \|z(\tau)\|^2 \mathrm{d}\tau\right\} \le \gamma^2\mathbb{E}\left\{\int_t^\infty \|w(\tau)\|^2 \mathrm{d}\tau\right\}. \tag{6}$$

For the purpose of addressing the $\mathcal{H}_\infty$ control problem for system (1), we consider the following infinite horizon cost functional

$$J(u,w;x_0,r_0) = \mathbb{E}_{x_0\sim\mathcal{D},r_0\sim\rho}\left\{\frac{1}{2}\int_0^\infty (x^T H_i x + u^T R_i u - \gamma^2 w^T w)dt\right\}. \tag{7}$$

According to the zero-sum differential game theory [42], [43], [44], we formulate the control input u and the disturbance w as the two players. Then the $\mathcal{H}_\infty$ control problem can be solved if the feedback saddle point solution (u^*, w^*), also called the Nash equilibrium can be found such that

$$J(u^*,w^*;x_0,r_0) = \min_u \max_w J(u,w;x_0,r_0) \tag{8}$$

$$= \max_w \min_u J(u,w;x_0,r_0). \tag{9}$$

Now define the optimal value function starting from (x,i) as

$$V_i^*(x) = J(u^*,w^*;x,i). \tag{10}$$

Then the associated mode dependent Hamiltonian function can be established as

$$\begin{aligned}&\mathscr{H}_i(x,u,w,V_i^*)\\ &= \mathscr{L}V_i^*(x) + \frac{1}{2}x^T H_i x + \frac{1}{2}u^T R_i u - \frac{1}{2}\gamma^2 w^T w\\ &= \frac{\partial V_i^{*T}(x)}{\partial x}(A_i x + B_i u + D_i w) + \sum_{j=1}^N \pi_{ij}V_j^*(x)\\ &\quad + \frac{1}{2}x^T H_i x + \frac{1}{2}u^T R_i u - \frac{1}{2}\gamma^2 w^T w.\end{aligned} \tag{11}$$

Since the considered system is linear, the optimal value function V_i^* can be represented as quadratic in the state, which is

$$V_i^*(x) = \frac{1}{2}x^T P_i x \tag{12}$$

where $P_i > 0$ is symmetric. By employing the stationary condition

$$\frac{\partial\mathscr{H}_i(x,u,w,V_i^*)}{\partial u} = 0, \quad \frac{\partial\mathscr{H}_i(x,u,w,V_i^*)}{\partial w} = 0 \tag{13}$$

the optimal control law u_c^* in the time triggered case and the worst-case disturbance input w^* can be obtained as

$$u_c^* := u^* = -R_i^{-1}B_i^T P_i x, \quad w^* = \frac{1}{\gamma^2}D_i^T P_i x \tag{14}$$

where P_i is the unique solution that solves $\mathscr{H}_i(x,u_c^*,w^*,V_i^*) = 0$, i.e., the following coupled game algebraic Riccati equations (CGAREs)

$$\begin{aligned}&P_i A_i + A_i^T P_i + \sum_{i=1}^N \pi_{ij}P_j + H_i\\ &\quad - P_i B_i R_i^{-1}B_i^T + \gamma^{-2}P_i D_i D_i^T P_i = 0,\ i\in\mathcal{N}.\end{aligned} \tag{15}$$

We are now in a position to introduce the event-triggering mechanism, which has the potential to reduce the necessary data transmission between the controller and the considered plant. To this end, we first introduce a monotonically increasing time sequence $\{t_k\}_{k=0}^\infty$, denoting the event triggering time instant. The state feedback signal can be written as

$$\hat{x}(t) := \begin{cases}x(t_k), & \forall t\in[t_k,t_{k+1})\\ x(t), & t = t_{k+1}.\end{cases} \tag{16}$$

Let $e(t) = x(t) - \hat{x}(t)$ be the error vector and e_T be the triggering event threshold which will be designed later. The control signal is only updated when the triggering condition $\|e\| \le e_T$ is violated, which shows that

$$t_{k+1} = \inf\{t|t > t_k \wedge \|e\| > e_T\}. \tag{17}$$

Here we utilize an event-triggered version of the aforementioned game algebraic Riccati equation and present the following event-triggered controller:

$$u_d^* = -R_i^{-1}B_i^T P_i\hat{x} \tag{18}$$

then by comparing u_c^* and u_d^*, for system (1), we have $\|u_c^*(x) - u_d^*(\hat{x})\| = \|R_i^{-1}B_i^T P_i(x-\hat{x})\| \le \|R_i^{-1}B_i^T P_i\|\|(x-\hat{x})\| \le L\|e\|$ where $L = \max_{i\in\mathcal{N}}\|R_i^{-1}B_i^T P_i\|$.

Remark 1: The Hamiltonian of the event-triggered version $\mathscr{H}_i(x,u_d^*,w^*,V_i^*)$ can be expressed in terms of the time triggered case as

$$\begin{aligned}&\mathscr{H}_i(x,u_d^*,w^*,V^*)\\ &= \mathscr{H}_i(x,u_c^*,w^*,V^*) + \frac{1}{2}(u_c^* - u_d^*)^T R_i(u_c^* - u_d^*)\end{aligned} \tag{19}$$

by taking (14) and (18) into (11) and completing squares. Bearing in mind that $\mathscr{H}_i(x,u^*,w^*,V^*) = 0$, it follows from the above that

$$\mathscr{H}_i(x,u_d^*,w^*,V^*) \le \frac{1}{2}L^2\max_{i\in\mathcal{N}}\bar{\lambda}(R_i)\|e\|^2. \tag{20}$$

This reveals that the difference brought by the event triggering mechanism can be measured through the state error e.

III. Main Results

In this section, the results on event triggering design are presented and analyzed. The corresponding model-based algorithm and model-free Q-learning based algorithm are both given to solve the desired parameters. The convergence has also been discussed in the end.

A. Event Triggering Mechanism Design

Theorem 1: Given a constant $\gamma > 0$ and a positive definite and radially unbounded function defined in (12) with $P_i, i \in \mathcal{N}$ satisfying the coupled game algebra Riccati equation (15), the closed loop system (1) with event-triggered controller given by (18) is stochastically stable with a predefined noise attenuation level index γ as long as the event-triggered condition

$$\|e\|^2 \le e_T^2 := \frac{(1-\beta_x)\min_{i\in\mathcal{N}} \underline{\lambda}(H_i)}{L^2 \max_{i\in\mathcal{N}} \bar{\lambda}(R_i)}\|x\|^2 + \frac{(1-\beta_u)\min_{i\in\mathcal{N}} \underline{\lambda}(R_i)}{L^2 \max_{i\in\mathcal{N}} \bar{\lambda}(R_i)}\|u_d^*\|^2 \tag{21}$$

is satisfied where $\beta_x, \beta_u \in (0,1)$ are user defined constants. Moreover, the control policy (18) is optimal in the event-triggered sense and the optimal cost is given by

$$J(x_0, r_0; u_d^*, w^*) = J(x_0, r_0; u_c^*, w^*) + \frac{1}{2}\mathbb{E}\left\{\int_0^\infty (u_d^* - u_c^*)^T R_i (u_d^* - u_c^*) dt\right\} \tag{22}$$

where $J(x_0, r_0; u_c^*, w^*) = \mathbb{E}_{x_0\sim\mathcal{D}, r_0\sim\rho}\{x_0^T P_{r_0} x_0\}$ is the optimal cost for the time-triggered case.

Proof: Recall the time-triggered HJB equation and we have

$$\mathcal{H}_i(x, u_c^*, w^*, V_i^*) = 0$$

i.e.,

$$\frac{\partial V_i^{*T}(x)}{\partial x}(A_i x + B_i u_c^* + D_i w^*) + \sum_{j=1}^N \pi_{ij} V_j^*(x) + \frac{1}{2}x^T H_i x + \frac{1}{2}u_c^{*T} R_i u_c^* - \frac{1}{2}\gamma^2 w^{*T} w^* = 0. \tag{23}$$

Hence for the value function when event-triggered controller is used, we obtain its infinitesimal as

$$\begin{aligned}
&\mathscr{L}V_i^*(x)\\
&= \frac{\partial V_i^{*T}(x)}{\partial x}(A_i x + B_i u_d^* + D_i w^*) + \sum_{j=1}^N \pi_{ij} V_j^*(x)\\
&= \frac{\partial V_i^{*T}(x)}{\partial x} B_i (u_d^* - u_c^*) - \frac{1}{2}x^T H_i x - \frac{1}{2}u_c^{*T} R_i u_c^* + \frac{1}{2}\gamma^2 w^{*T} w^*\\
&\le \frac{1}{2}\|u_c^* - u_d^*\|_{R_i}^2 - \frac{1}{2}\|u_d^*\|_{R_i}^2 - \frac{1}{2}\|x\|_{H_i}^2 + \frac{1}{2}\gamma^2\|w^*\|^2\\
&\le \frac{1}{2}\|u_c^* - u_d^*\|_{R_i}^2 - \frac{1}{2}\|u_d^*\|_{R_i}^2 - \frac{1}{2}(1-\beta_u)\|u_d^*\|_{R_i}^2\\
&\quad - \frac{1}{2}\|x\|_{H_i}^2 - \frac{1}{2}(1-\beta_x)\|x\|_{H_i}^2 + \frac{1}{2}\gamma^2\|w^*\|^2
\end{aligned}$$

using the fact $x^T P_i B_i = -u_c^* R_i$ and $x^T P_i D_i = w^{*T}\gamma^2$. $\beta_x, \beta_u \in (0,1)$ are user defined constants.

Now let us examine the stochastic stability under the case of $w^* = 0$. If the condition (21) holds, $\mathscr{L}V_i^*(x) \le -\frac{1}{2}\beta_x\|x\|_{H_i}^2 - \frac{1}{2}\beta_u\|u_d^*\|_{R_i}^2$, which further guarantees the stochastic stability of the considered system by Definition 1.

Next we will show the optimality of the event-triggered control policy corresponding to the designed event condition (21) and the relevant cost compared with the time-triggered one.

Since the value function V_i^* can serve as a Lyapunov function, we can rewrite the cost functional as

$$J(u, w; x_0, r_0) = \frac{1}{2}\mathbb{E}_{x_0\sim\mathcal{D}, r_0\sim\rho}\left\{\int_0^\infty (x^T H_i x + u^T R_i u - \gamma^2 w^T w + \mathscr{L}V_i^*(x))dt + x_0 P_{r_0} x_0\right\}. \tag{24}$$

By utilizing the discussions in Remark 1 and completing the squares, we have

$$J(u_d, w; x_0, r_0) = \frac{1}{2}\mathbb{E}_{x_0\sim\mathcal{D}, r_0\sim\rho}\left\{\int_0^\infty (-\gamma^2\|w - w^*\|^2 \|u_d - u_d^*\|_{R_i}^2 + \|u_c^* - u_d^*\|_{R_i}^2)dt + x_0 P_{r_0} x_0\right\} \tag{25}$$

which shows that

$$J(u_d^*, w; x_0, r_0) \le J(u_d^*, w^*; x_0, r_0) \le J(u_d, w^*; x_0, r_0) \tag{26}$$

i.e., a saddle point exists and the optimal cost equals (22).

Algorithm 1 Optimal $\mathcal{H}_\infty$ Event-Triggered Control by Using Offline Parallel Iteration Algorithm

Step 1: Initialize a symmetric positive definite matrix sequence $\{P_i(0), i = 1, 2, \ldots, N\}$. Set event triggering parameters L, β_u, β_x according to Theorem 1.
Step 2: Let *converged* be a stopping flag and ε be the threshold. Do the following loop:
$k = 0$
converged $\leftarrow$ *false*
while not *converged* **do**
  for $i = 1 : N$ **do**

$$\bar{H}_i(k) = H_i + \sum_{j=1, j\ne i}^N \pi_{ij} P_j(k) \tag{27}$$

Solve the following equation for $P_i(k+1)$:

$$P_i(k+1)(\mathbf{A}_i - S_i P_i(k)) + (\mathbf{A}_i - S_i P_i(k))^T P_i(k+1) = -P_i(k) S_i P_i(k) - \bar{H}_i(k) \tag{28}$$

with $\mathbf{A}_i = A_i + \frac{\pi_{ii}}{2} I$ and $S_i = B_i R_i^{-1} B_i^T - \gamma^{-2} D_i D_i^T$
  end for
  if $\max_{i\in\mathcal{N}} \|P_i(k+1) - P_i(k)\| \ge \epsilon$ **then**
    $k = k + 1$
  else
    converged $\leftarrow$ *true*
  end if
end while
return: $P_i^* = P_i(k+1), i = 1, 2, \ldots, N$
Step 3: Substitute P_i^* in (18) to implement the optimal $\mathcal{H}_\infty$ event-triggered controller u_d^* with the event triggering condition (21).

Remark 2: The triggering event follows (21) and the controller gain is used as the same as in the time triggered case. We can use conventional parallel iteration algorithm [45] to solve the CGAREs (15), which is shown in Algorithm 1, to complete the event control design. It should be noted that there are other optimal options for chosen controller gains corresponding to different event triggering mechanism.

B. Q-Learning Based Actor-Critic Formulation and Learning Algorithm

Since solving the CGAREs is tedous and the result is vulnerable to system parameters, here a model free algorithm is preferred. Under Q-learning framework, we define the following mode dependent Q-function $\mathcal{Q}_i(x,u_d,w):\mathbb{R}^{n+m}\to\mathbb{R}^+$:

$$\mathcal{Q}^i(x,u_d,w) := V_i^*(x) + \mathscr{H}_i(x,u_d,w,V_i^*) - \mathscr{H}_i(x,u_c^*,w^*,V_i^*), \quad i\in\mathcal{N}. \tag{29}$$

Keep in mind that $\mathscr{H}_i(x,u_c^*,w^*,V^*)=0$, and the optimal time-triggered cost can be written in the form $V_i^*(x)=\frac{1}{2}x^TP_ix$. It follows from the definition of the mode dependent Hamiltonian function that this Q-function can be written in a compact quadratic formulation regarding the state x, control u_d and the disturbance w as follows:

$$\mathcal{Q}^i(x,u_d,w) := \frac{1}{2}U^T\bar{Q}^iU = \frac{1}{2}U^T\begin{bmatrix} Q_x^i & Q_{xu_d}^i & Q_{xw}^i \\ Q_{u_dx}^i & Q_{u_d}^i & Q_{u_dw}^i \\ Q_{wx}^i & Q_{wu_d}^i & Q_w^i \end{bmatrix}U$$

where $U=\begin{bmatrix} x^T & u_d^T & w^T\end{bmatrix}^T$ and

$$Q_x^i = P_i + H_i + P_iA_i + A^TP_i + \sum_{j=1}^N \pi_{ij}P_j$$
$$Q_{xu_d}^i = {Q_{xu_d}^i}^T = P_iB_i, Q_{xw}^i = {Q_{wx}^i}^T = P_iD_i$$
$$Q_{u_dw}^i = {Q_{wu_d}^i}^T = 0, Q_{u_d}^i = R_i, Q_w^i = -\gamma^2\, I.$$

Applying the stationary conditions, we obtain an equivalent expression for the optimal policies as

$$u_d^* = -(Q_{u_d}^i)^{-1}Q_{u_dx}^i\hat{x}, \quad w^* = -(Q_w^i)^{-1}Q_{wx}^ix. \tag{30}$$

Therefore by learning the action-dependent value function instead of solving the CGARE, we can develop an online learning algorithm of the controller gains where the knowledge of the system model is not necessary.

Denote $W_c^i := \frac{1}{2}\text{vech}(\bar{Q}^i)$ and the mode dependent Q-function can be rewritten as

$$\mathcal{Q}^i(x,u_d^*,w^*) = W_c^{iT}(U\otimes U), \quad i\in\mathcal{N} \tag{31}$$

with $W_c^i\in\mathbb{R}^{\frac{1}{2}(n_x+n_u+n_w)(n_x+n_u+n_w+1)}$ is the ideal weight. Denote $W_c^i=[n:m]$ as the vector with elements abstracted from the nth to the mth elements of W_c^i and also define the following indices

$$idx_1 = \frac{n_x(n_x+1)}{2}, \quad idx_2 = \frac{n_x(n_x+1)}{2} + n_xn_u$$
$$idx_3 = \frac{n_x(n_x+1)}{2} + n_x(n_u+n_w)$$
$$idx_4 = \frac{n_x(n_x+1)}{2} + n_x(n_u+n_w) + \frac{n_u(n_u+1)}{2}$$
$$idx_5 = \frac{n_x(n_x+1)}{2} + n_x(n_u+n_w) + \frac{n_u(n_u+1)}{2} + n_un_w$$
$$idx_6 = \frac{1}{2}(n_x+n_u+n_w)(n_x+n_u+n_w+1)$$

then we will have

$$\text{vech}(Q_x^i) := W_c^i[1:idx_1]$$
$$\text{vech}(Q_{xu_d}^i) := W_c^i[idx_1+1:idx_2]$$
$$\text{vech}(Q_{xw}^i) := W_c^i[idx_2+1:idx_3]$$
$$\text{vech}(Q_{u_d}^i) := W_c^i[idx_3+1:idx_4]$$
$$\text{vech}(Q_{u_dw}^i) := W_c^i[idx_4+1:idx_5]$$
$$\text{vech}(Q_w^i) := W_c^i[idx_5+1:idx_6].$$

Note that the above weights W_c^i are ideal but unknown for computing $\mathcal{Q}^i(x,u_d^*,w^*)$. We use the weight estimates $\hat{W}_c^i$ to establish the approximator $\hat{\mathcal{Q}}^i(x,u_d^*,w^*)$ as follows:

$$\hat{\mathcal{Q}}^i(x,u_d^*,w^*) = \hat{W}_c^{iT}(U\otimes U), \quad \forall i\in\mathcal{N} \tag{32}$$

where $\hat{W}_c^i\in\mathbb{R}^{\frac{1}{2}(n_x+n_u+n_w)(n_x+n_u+n_w+1)}$. By using a similar line, an approximate solution $\hat{u}_d$ of u_d^* and the disturbance $\hat{w}$ for w^* can be approximated by:

$$\hat{u}_d = \hat{W}_a^{iT}\hat{x}, \quad i\in\mathcal{N} \tag{33}$$
$$\hat{w} = \hat{W}_w^{iT}x, \quad i\in\mathcal{N} \tag{34}$$

where $\hat{W}_a^i\in\mathbb{R}^{n_x\times n_u}$ and $\hat{W}_w^i\in\mathbb{R}^{n_x\times n_w}$ are the weight estimates.

Following the results on integral RL method shown in [43] and [46], we can write a mode dependent version of integral Bellman equation for MJS in terms of $\mathcal{Q}^i$ as

$$\mathcal{Q}^i(x(t),u_d^*(t),w^*(t)) - \mathcal{Q}^i(x(t-T),u_d^*(t-T),w^*(t-T)) = -\frac{1}{2}\int_{t-T}^t\left(\|x\|_{H_i}^2 + \|u_d^*\|_{R_i}^2 - \gamma^2\|w^*\|^2 + \sum_{j=1}^N\pi_{ij}V_j^*(x)\right)\mathrm{d}\tau$$

where $T\in\mathbb{R}^+$ is a very small fixed time interval during which the system data is collected for learning. It is assumed in this paper that during this fixed time interval the system mode remains unchanged. In order to find the update law for the critic, we shall define the following error $e_c\in\mathbb{R}$

$$e_c = \hat{W}_c^{iT}(U(t)\otimes U(t)) - \hat{W}_c^{iT}(U(t-T)\otimes U(t-T)) + \frac{1}{2}\int_{t-T}^t\left(\|x\|_{H_i}^2 + \|\hat{u}_d\|_{R_i}^2 - \gamma^2\|w^*\|^2 + \sum_{j=1}^N\pi_{ij}V_j^*(x)\right)\mathrm{d}\tau. \tag{35}$$

From (30), the error $e_a\in\mathbb{R}^m$ for actor approximator and error $e_w\in\mathbb{R}^w$ for the worst case disturbance approximator can be given by

$$e_a = \hat{W}_a^{iT}\hat{x} + (Q_{u_d}^i)^{-1}Q_{u_dx}^i\hat{x} \tag{36}$$
$$e_w = \hat{W}_w^{iT}x + (Q_w^i)^{-1}Q_{xw}^ix \tag{37}$$

where the values of $Q^i_{u_d}$, $Q^i_{u_d x}$ Q^i_{xw} and Q^i_w can be extracted from the vector $\hat{W}^i_c$.

Now define the following squared-norms:

$$E_1 = \frac{1}{2}\|e_c\|^2$$
$$E_2 = \frac{1}{2}\|e_a\|^2$$
$$E_3 = \frac{1}{2}\|e_w\|^2.$$

By applying gradient descent, the updating laws for $\hat{W}^i_c$,$\hat{W}^i_a$ and $\hat{W}^i_w$ with which the errors e_c, e_a and e_w approach zero can be given by

$$\dot{\hat{W}}^i_c = -\alpha_c \frac{1}{\left(1+\sigma^T\sigma\right)^2}\frac{\partial E_1}{\partial \hat{W}^i_c} = -\alpha_c \frac{\sigma}{\left(1+\sigma^T\sigma\right)^2} e^T_c \tag{38}$$

$$\begin{cases} \dot{\hat{W}}^i_a & = 0 \forall t \in [t_k, t_{k+1}) \\ \hat{W}^{i+}_a & = \hat{W}^i_a - \alpha_a \frac{1}{(1+x^T x)}\frac{\partial K_2}{\partial \hat{W}^i_a} \\ & = \hat{W}^i_a - \alpha_a \frac{x}{(1+x^T x)} e^T_a t = t_{k+1} \end{cases} \tag{39}$$

$$\dot{\hat{W}}^i_w = -\alpha_w \frac{1}{\left(1+x^T x\right)^2}\frac{\partial E_3}{\partial \hat{W}^i_c} = -\alpha_c \frac{x}{\left(1+x^T x\right)^2} e^T_w \tag{40}$$

where $\sigma := U(t)\otimes U(t) - U(t-T)\otimes U(t-T)$, and $\alpha_c, \alpha_a, \alpha_w \in \mathbb{R}^+$ are user defined constant gains that determine the speed of convergence of the approximations for the critic and the two actors, respectively.

C. Stability and Convergence Analysis

In this subsection, the convergence of the updating laws for the weight parameters and the stochastic stability of the system state shall be analyzed. To this end, we first define for the weight estimation errors:

$$\begin{aligned} \tilde{W}^i_c &:= W^i_c - \hat{W}^i_c \\ \tilde{W}^i_a &:= -Q^i_{xu_d}(Q^i_{u_d})^{-1} - \hat{W}^i_a \\ \tilde{W}^i_w &= -Q^i_{xw}(Q^i_w)^{-1} - \hat{W}^i_w. \end{aligned} \tag{41}$$

Then the error dynamics can be written respectively as

$$\dot{\tilde{W}}^i_c = -\alpha_c \frac{\sigma\sigma^T}{\left(1+\sigma^T\sigma\right)^2}\tilde{W}^i_c \tag{42}$$

$$\begin{cases} \dot{\tilde{W}}^i_a = 0, & \forall t \in [t_k, t_{k+1}) \\ \tilde{W}^{i+}_a = \tilde{W}^i_a - \alpha_a \frac{xx^T}{(1+x^T x)}\tilde{W}^i_a & \\ -\alpha_a \frac{xx^T}{(1+x^T x)}\tilde{Q}^i_{xu_d}(Q^i_{u_d u_d})^{-1}, & t = t_{k+1} \end{cases} \tag{43}$$

$$\dot{\tilde{W}}^i_w = -\alpha_w \frac{xx^T}{(1+x^T x)^2}\tilde{W}^i_w - \alpha_w \frac{xx^T}{(1+x^T x)^2}\tilde{Q}^i_{xw}(Q^i_{ww})^{-1}. \tag{44}$$

Hence the closed-loop system dynamic becomes

$$\begin{aligned} \dot{x} = A_i x + B_i(-(Q^i_{u_d})^{-1}Q^i_{u_d x} - \tilde{W}^{iT}_a)\hat{x} \\ + D_i(-(Q^i_w)^{-1}Q^i_{wx} - \tilde{W}^{iT}_w)x, \quad t \geq 0. \end{aligned} \tag{45}$$

Now we are in the position to examine the convergence of the weight estimation errors $\tilde{W}^i_c, \tilde{W}^i_a, \tilde{W}^i_w$ and the stochastic stability of the system. The results are shown in the following theorem.

Theorem 2: Consider the closed loop system (45) and a predefined noise attenuation level index $\gamma > 0$. The critic, the actor and the disturbance approximators are given by (32), (33) and (34) respectively. The weight parameters $\hat{W}^i_c$,$\hat{W}^i_a$ and $\hat{W}^i_w$ in the critic network and actor network are updated according to (38), (39) and (40), respectively. Then, the dynamic of the closed-loop system (45) is stochastically stable and the weight estimation errors defined in (41) converge as long as the following event-triggered conditions are satisfied

$$\|e\|^2 \leq e^2_T := \frac{(1-\beta^2_x)\min_{i\in\mathcal{N}}\underline{\lambda}(H_i)}{4\left(L^2+L^2_1\right)\max_{i\in\mathcal{N}}\bar{\lambda}(R_i)}\|x\|^2 + \frac{(1-\beta^2_u)\min_{i\in\mathcal{N}}\underline{\lambda}(R_i)}{4\left(L^2+L^2_1\right)\max_{i\in\mathcal{N}}\bar{\lambda}(R_i)}\|\hat{u}_d\|^2 \tag{46}$$

where L_1 is a positive constant of unity order and the following inequalities hold for the rest of the parameters:

$$\frac{\max_{i\in\mathcal{N}}\bar{\lambda}(R_i)L^2_1 + \gamma^2\tilde{Q}^i_{xw}(Q^i_w)^{-1}}{\beta^2_x\min_{i\in\mathcal{N}}\underline{\lambda}(H_i)} < \frac{1}{2} \tag{47}$$

$$\frac{\alpha_w}{4\gamma^2} - \alpha_c\underline{\lambda}(MM^T) < 0 \tag{48}$$

$$\frac{\alpha_w}{4\gamma^2} - \alpha_w + \gamma^2 < 0 \tag{49}$$

$$0 < \alpha_a < \min_{i\in\mathcal{N}}\frac{8\underline{\lambda}(R_i)-2}{\underline{\lambda}(R_i)+2} \tag{50}$$

$$\alpha_c \gg \alpha_a. \tag{51}$$

Proof: First we consider the dynamics in $[t_k, t_{k+1})$, i.e., the case where event is not triggered. During this time interval, $\hat{x}$ and $\tilde{W}_a$ remains unchanged. For dynamics of the rest of variables, we define the following Lyapunov function: $\mathbf{V} : \mathbb{R}^n \times \mathbb{R}^{\frac{1}{2}(n+m)(n+m+1)} \times \mathbb{R}^{n\times m} \to \mathbb{R}$,

$$\mathbf{V} = V^*_i(x) + V_c(\tilde{W}^i_c) + V_w(\tilde{W}^i_w), \quad t \geq 0 \tag{52}$$

where $V^*_i(x)$ is the optimal value function, $V_c(\tilde{W}^i_c) = \frac{1}{2}\|\tilde{W}^i_c\|^2$, $V_w(\tilde{W}^i_w) = \frac{1}{2}\mathrm{tr}\left\{\tilde{W}^{iT}_w\tilde{W}^i_w\right\}$.

Along the dynamics of the closed loop real-time system (45) in the time interval $[t_k, t_{k+1})$, it follows from the analysis in Theorem 1 that

$$\mathscr{L}\mathbf{V} = \frac{1}{2}\|u^*_c - u^*_d\|_{R_i} - \frac{1}{2}\|u^*_d\|^2_{R_i} - \frac{1}{2}\gamma^2\|w^*\|^2 + \gamma^2(w^*)^T\hat{w} - \frac{1}{2}\|x\|^2_{H_i} + \tilde{W}^{iT}_c\dot{\tilde{W}}^i_c + \mathrm{tr}\left\{\tilde{W}^{iT}_w\dot{\tilde{W}}^i_w\right\}. \tag{53}$$

Since the actor weights remain unchanged during $[t_k, t_{k+1})$, there exist a positive constant of unity order L_1 such that $\|\tilde{W}^i_a\| \leq L_1$. Therefore it can be calculated by using Young's inequality for the cross terms that

$$\begin{aligned} &\frac{1}{2}\|u^*_c - u^*_d\|_{R_i} - \frac{1}{2}\|u^*_d\|^2_{R_i} \\ &\leq \frac{1}{2}\bar{\lambda}(R_i)\|R^{-1}_i B^T_i P_i x - \hat{W}^{iT}_a\hat{x}\|^2 - \frac{1}{2}\underline{\lambda}(R_i)\|\hat{u}_d\|^2 \\ &= \frac{1}{2}\bar{\lambda}(R_i)\left\|R^{-1}_i B^T_i P_i x - \hat{W}^{iT}_a(x+e)\right\|^2 - \frac{1}{2}\underline{\lambda}(R_i)\|\hat{u}_d\|^2 \\ &= \frac{1}{2}\bar{\lambda}(R_i)\left\|\tilde{W}^{iT}_a x - \hat{W}^{iT}_a e\right\|^2 - \frac{1}{2}\underline{\lambda}(R_i)\|\hat{u}_d\|^2 \\ &\leq 2\max_{i\in\mathcal{N}}\bar{\lambda}(R_i)(L^2+L^2_1)\|e\|^2 + \max_{i\in\mathcal{N}}\bar{\lambda}(R)L^2_1\|x\|^2 - \frac{1}{2}\min_{i\in\mathcal{N}}\underline{\lambda}(R_i)\|\hat{u}_d\|^2 \end{aligned}$$

and

$$\begin{aligned}&-\frac{1}{2}\gamma^2\|w^*\|^2+\gamma^2(w^*)^T\hat{w}\\&=-\frac{1}{2}\gamma^2\|w^*-\hat{w}\|^2+\frac{1}{2}\gamma^2\|\hat{w}\|^2\leq\frac{1}{2}\gamma^2\|w^*-\tilde{w}\|^2\\&\leq\gamma^2(\|w^*\|^2+\|\tilde{w}\|^2)=\gamma^2((Q_w^i)^{-1}Q_{xw}^i)^2\|x\|^2+\gamma^2\|\tilde{W}_w^{iT}x\|^2.\end{aligned}$$

Next, denote the signal $M=\frac{\sigma}{(\sigma^T\sigma+1)}$ and assume that M is persistently exciting over the interval $[t,t+T]$ with $\int_t^{t+T}MM^Td\tau\geqslant\varrho I$, where $\varrho\in\mathbb{R}^+$and I is an identity matrix of appropriate dimensions and also that there exists $M_B\in\mathbb{R}^+$such that for all $t\geqslant t_0,\max\{|M|,|\dot{M}|\}\leqslant M_B$. Then we can obtain

$$\begin{aligned}\tilde{W}_c^{iT}\dot{\tilde{W}}_c^i=&-\alpha_c\tilde{W}_c^{iT}\frac{\sigma\sigma^T}{\left(1+\sigma^T\sigma\right)^2}\tilde{W}_c^i\\\leq&-\alpha_c\underline{\lambda}(MM^T)\|\tilde{W}_c^i\|^2.\end{aligned}$$

Finally, using Young's inequality and some algebra gives that

$$\begin{aligned}\mathrm{tr}\{\tilde{W}_w^{iT}\dot{\tilde{W}}_w^i\}=&-\alpha_w\mathrm{tr}\left\{\tilde{W}_w^{iT}\frac{xx^T}{1+x^Tx}\tilde{W}_w^i\right.\\&\left.+\tilde{W}_w^{iT}\frac{xx^T}{1+x^Tx}\tilde{Q}_{xw}^i(Q_{ww}^i)^{-1}\right\}\\\leq&-\left(\alpha_w-\frac{\alpha_w}{4\gamma^2}\right)\|\tilde{W}_w^{iT}x\|^2+\frac{\alpha_w}{4\gamma^2}\|\tilde{W}_c^i\|^2\end{aligned}$$

Combing all the above calculations, it yields that

$$\begin{aligned}\mathscr{L}\mathbf{V}\leq&\left(\frac{\alpha_w}{4\gamma^2}-\alpha_c\underline{\lambda}(MM^T)\right)\|\tilde{W}_c^i\|^2-\frac{\beta_u^2}{2}\min_{i\in\mathcal{N}}\underline{\lambda}(R_i)\|\hat{u}_d\|^2\\&+\left(\frac{\alpha_w}{4\gamma^2}-\alpha_w+\gamma^2\right)\|\tilde{W}_w^{iT}x\|^2+\boldsymbol{\Theta_1}+\boldsymbol{\Theta_2}\end{aligned}$$

where

$$\begin{aligned}\boldsymbol{\Theta_1}=&2\max_{i\in\mathcal{N}}\bar{\lambda}(R_i)(L^2+L_1^2)\|e\|^2\\&-\frac{1-\beta_x^2}{2}\min_{i\in\mathcal{N}}\underline{\lambda}(H_i)\|x\|^2-\frac{1-\beta_u^2}{2}\min_{i\in\mathcal{N}}\underline{\lambda}(R_i)\|\hat{u}_d\|^2\\\boldsymbol{\Theta_2}=&\left(\max_{i\in\mathcal{N}}\bar{\lambda}(R_i)L_1^2+\gamma^2\tilde{Q}_{xw}^i(Q_w^i)^{-1}-\frac{\beta_x^2}{2}\min_{i\in\mathcal{N}}\underline{\lambda}(H_i)\right)\|x\|^2.\end{aligned}$$

Taking into account the event triggered condition, we conclude that $\mathscr{L}\mathbf{V}\leq 0$, which shows the trajectories of $x,\tilde{W}_c^i$ and $\tilde{W}_w^i$ converges to 0.

Now we consider the jump dynamics in the triggering instant t_{k+1}. Note that $x,\tilde{W}_c^i$ and $\tilde{W}_w^i$ are continuous without jump dynamics. Here we only consider $\hat{x}$ and $\tilde{W}_a^i$. Based on the fact that $\hat{x}^+=x(t_{k+1})$ and the convergence of x, $x(t_k)$ converges to the origin asymptotically.

For the error dynamic of $\tilde{W}_a^i$, let $V_a(\tilde{W}_a^i)=\frac{1}{2\alpha_a}\mathrm{tr}\{\tilde{W}_a^{iT}\tilde{W}_a^i\}$. Then it can be calculated that

$$\begin{aligned}&V_a(\tilde{W}_a^{i+})-V_a(\tilde{W}_a^i)\\&=\frac{1}{\alpha_a}\mathrm{tr}\left\{\tilde{W}_a^{iT}\left(-\alpha_a\frac{xx^T}{1+x^Tx}\tilde{W}_a^i\right)\right\}\\&\quad+\frac{1}{\alpha_a}\mathrm{tr}\left\{\tilde{W}_a^{iT}\left(-\alpha_a\frac{xx^T}{1+x^Tx}\tilde{Q}_{xu_d}^i(Q_{u_d}^i)^{-1}\right)\right\}\\&\quad+\frac{1}{2\alpha_a}\mathrm{tr}\left\{\alpha_a^2\left(\frac{xx^T}{1+x^Tx}\tilde{W}_a^i\right)^T\left(\frac{xx^T}{1+x^Tx}\tilde{W}_a^i\right)\right\}\\&\quad+\frac{1}{\alpha_a}\mathrm{tr}\left\{\alpha_a^2\left(\frac{xx^T}{1+x^Tx}\tilde{W}_a^i\right)^T\left(\frac{xx^T}{1+x^Tx}\tilde{Q}_{xu_d}(Q_{u_d}^i)^{-1}\right)\right\}\\&\quad+\frac{1}{2\alpha_a}\mathrm{tr}\left\{\alpha_a^2\left(\frac{xx^T}{1+x^Tx}\tilde{Q}_{xu_d}(Q_{u_d}^i)^{-1}\right)^T\right.\\&\quad\left.\times\left(\frac{xx^T}{1+x^Tx}\tilde{Q}_{xu_d}(Q_{u_d}^i)^{-1}\right)\right\}.\end{aligned}$$

Keep in mind the following fact

$$\begin{aligned}&\left\|\frac{1}{1+x^Tx}\right\|\leq 1,\quad\left\|\frac{xx^T}{1+x^Tx}\right\|\leq 1\\&\left\|\frac{x}{1+x^Tx}\right\|\leq\frac{1}{2},\quad\left\|\frac{xx^T}{(1+x^Tx)^2}\right\|\leq\frac{1}{4}.\end{aligned}$$

After taking norms and using Young's inequality for the cross-terms and Cauchy-Schwarz inequality, we have

$$\begin{aligned}&V_a(\tilde{W}_a^{i+})-V_a(\tilde{W}_a^i)\\&\leq-\left\|\tilde{W}_a^{iT}x\right\|^2+\left\|\tilde{W}_a^{iT}x\right\|\left\|\frac{x}{1+x^Tx}\right\|\left\|\tilde{Q}_{xu_d}^i\right\|\bar{\lambda}((Q_{u_d}^i)^{-1})\\&\quad+\frac{\alpha_a}{2}\left\|\tilde{W}_a^{iT}x\right\|^2\left\|\frac{xx^T}{1+x^Tx}\right\|\\&\quad+\alpha_a\left\|\tilde{W}_a^{iT}x\right\|\left\|\frac{x}{1+x^Tx}\right\|\left\|\frac{xx^T}{1+x^Tx}\right\|\left\|\tilde{Q}_{xu_d}^i\right\|\bar{\lambda}((Q_{u_d}^i)^{-1})\\&\quad+\frac{\alpha_a}{2}\left\|\tilde{Q}_{xu_d}^i\right\|^2\left\|\frac{xx^T}{1+x^Tx}\right\|^2[\bar{\lambda}((Q_{u_d}^i)^{-1})]^2\\&\leq-\left\|\tilde{W}_a^{iT}x\right\|^2\left(1-\frac{\bar{\lambda}((Q_{u_d}^i)^{-1})}{4}-\frac{\alpha_a}{8}-\frac{\alpha_a\bar{\lambda}((Q_{u_d}^i)^{-1})}{4}\right)\\&\quad+\frac{[\bar{\lambda}((Q_{u_d}^i)^{-1})]^2}{4}\left\|\tilde{Q}_{xu_d}^i\right\|^2+\frac{\alpha_a\bar{\lambda}((Q_{u_d}^i)^{-1})}{4}\left\|\tilde{Q}_{xu_d}\right\|^2\\&\quad+\frac{\alpha_a[\bar{\lambda}((Q_{u_d}^i)^{-1})]^2}{2}\left\|\tilde{Q}_{xu_d}^i\right\|^2.\end{aligned}\tag{54}$$

Now define a compact set

$$\begin{aligned}\mathscr{R}=&\left\{\tilde{W}_a^i:\|\tilde{W}_a^i\|\leq\right.\\&\left.\sqrt{\frac{\frac{[\bar{\lambda}((Q_{u_d}^i)^{-1})]^2}{4}+\frac{\alpha_a\bar{\lambda}((Q_{u_d}^i)^{-1})}{4}+\frac{\alpha_a[\bar{\lambda}((Q_{u_d}^i)^{-1})]^2}{2}}{1-\frac{\bar{\lambda}((Q_{u_d}^i)^{-1})}{4}-\frac{\alpha_a}{8}-\frac{\alpha_a\bar{\lambda}((Q_{u_d}^i)^{-1})}{4}}}\|\tilde{Q}_{xu_d}^i\|\right\}.\end{aligned}$$

It can be guaranteed that $V_a(\tilde{W}_a^{i+})-V_a(\tilde{W}_a^i)<0$ whenever $\tilde{W}_a^i$ lies outside the compact set $\mathscr{R}$. Here we have to select $\alpha_a<\frac{8-2\bar{\lambda}((Q_{u_d}^i)^{-1})}{1+2\bar{\lambda}((Q_{u_d}^i)^{-1})}$, which can be ensured by (50) with $\bar{\lambda}((Q_{u_d}^i)^{-1})=1/\underline{\lambda}(R_i)$. Consequently, as the convergence of $\|\tilde{Q}_{xu_d}^i\|$, the set $\mathscr{R}$ shrinks and will eventually turn into a single point. This means that $\|\tilde{W}_a^i\|$ converges to 0.

Now based on the above discussions, the actor-critic Q-learning based online model-free learning scheme can be concluded in Algorithm 2.

Remark 3: According to Theorem 2, the tuning gain α_c and α_a satisfies $\alpha_c\gg\alpha_a$, which means the critic network has a way faster convergence than the actor approximator. This is essential for Algorithm 2 since it guarrantees the asymptotic stability of the set $\mathscr{R}$.

Algorithm 2 Optimal $\mathcal{H}_\infty$ Event-Triggered Control by Using Proposed Online Model-Free Algorithm

Step 1: Randomly initialize weights $\hat{W}_c^i$, $\hat{W}_a^i$ and $\hat{W}_w^i$. Select the user-defined parameters $\alpha_c, \alpha_a, \alpha_w$ indicating convergence speed and event triggering parameters L, L_1, β_u, β_x according to Theorem 2. Set the sampling time interval t_s and a samll time interval $T = nt_s$ where n is a user-defined number for collecting system data.
Step 2: Let *converged* be a stopping flag and ε be the threshold
Step 3: For a given noise attenuation level index γ, do the following loop:
$converged \leftarrow false$
while not *converged* **do**
run the system and store system data $x, \hat{u}_d, \hat{w}$ during every time interval T, then update:
$\hat{W}_c^i \leftarrow \hat{W}_c^i + \dot{\hat{W}}_c^i * T$
$\hat{W}_w^i \leftarrow \hat{W}_w^i + \dot{\hat{W}}_w^i * t_s$
if not (46) **then**
$\hat{W}_a^i \leftarrow$ (39)
else
$\hat{W}_a^i \leftarrow \hat{W}_a^i$
end if
if $E_2 < \varepsilon$ **then**
$converged \leftarrow true$
end if
end while
return: control network weights $\hat{W}_a^i$
Step 4: Obtain the optimal $\mathcal{H}_\infty$ event-triggered controller (33) with respect to the event triggering condition (46).

Remark 4: It can be found that in the literature the event-triggered control problem for MJSs are mostly discussed through LMI-based methods [33], [34], [35], [36], [37]. They rely heavily on accurate system models and solving complex equations or inequalities, which makes them vulnerable to system parameters. By contrast, the proposed approach presented in Algorithm 2 determines the controller gain through an iterative process using collected system data. This model-free and data-driven design enhances robustness against uncertainties in system parameters and structural variations. However, these methods are not absolutely superior or inferior to each other. They should be chosen with care according to real application environments.

Remark 5: Regarding the computational complexity, the number of variables that we need to solve is $\frac{1}{2}(n_x + n_u + n_w)(n_x + n_u + n_w + 1)$ *and the weight updates is through a simple gradient decent, which means that our approach has second order polynomial complexity in each iteration step.*

IV. Numerical Examples

In this section, a fourth-order Markov jump system is taken from [47] to illustrate the proposed results. This simulation was executed via MATLAB R2023a on a computer with a 3.60-GHz Intel Core i7-12700KF CPU. The system parameters

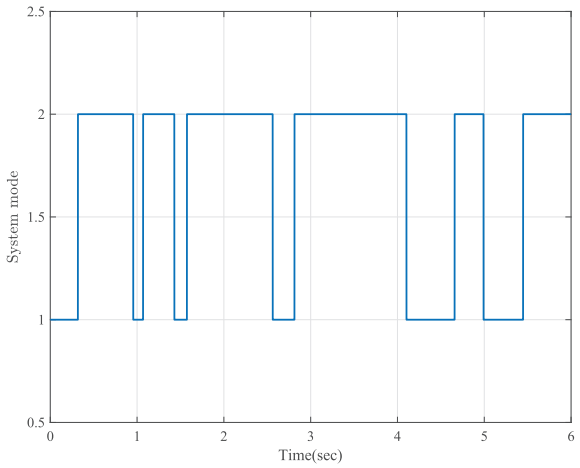

Fig. 1. System mode $r(t)$.

are as follows:

$$A_1 = \begin{bmatrix} -2.1051 & -11648 & 0.9347 & 0.5194 \\ -0.0807 & -2.8949 & 0.3835 & 0.8301 \\ 0.6914 & 10.5940 & -36.8199 & 3.8560 \\ 1.0692 & 13.4230 & 22.1185 & -13.1810 \end{bmatrix}$$

$$A_2 = \begin{bmatrix} -2.6430 & -1.2497 & 0.5269 & 0.6539 \\ -0.7910 & -2.8570 & 0.0920 & 0.4160 \\ 21.0357 & 22.8659 & -26.4655 & -1.7214 \\ 27.3096 & 7.8736 & -3.8604 & -29.5345 \end{bmatrix}$$

$$B_1 = \begin{bmatrix} 0.7564 & 0.9910 & 9.8255 & 7.2266 \end{bmatrix}^T$$

$$B_2 = \begin{bmatrix} 0.3653 & 0.2470 & 7.5336 & 6.5152 \end{bmatrix}^T$$

$$D_1 = D_2 = \begin{bmatrix} 1 & 0 & 0 & 0 \end{bmatrix}^T, R_1 = R_2 = 1$$

$$H_1 = H_2 = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}^T, [\pi_{ij}] = \begin{bmatrix} -2 & 2 \\ 1.5 & -1.5 \end{bmatrix}.$$

Fig. 1 shows a possible mode evolution.

To evaluate the $\mathcal{H}_\infty$ control performance, here we set the noise attenuation level index $\gamma = 2.5$. The rest of the parameters relating our model-free algorithm are $\alpha_a = 0.05$,$\alpha_c = 20$,$\alpha_w = 0.008$, $L = 10$, $L_1 = 2$ respectively. The Q-learning interval is chosen as $T = 0.01$ and the converge threshold is $\varepsilon = 0.001$. Then by online simulation, after 267 iteration steps, 0.1879s in runtime, the actor weights are converged as shown in Fig. 2, which is

$$W_a^1 = \begin{bmatrix} 0.1384 & 0.4800 & -0.5664 & -0.3969 \end{bmatrix}^T$$
$$W_a^2 = \begin{bmatrix} 0.0148 & 0.7691 & 0.3770 & 0.0636 \end{bmatrix}^T.$$

The system trajectory and the control action are shown in Fig. 3, which verifies the stability of the closed loop system and the effectiveness of the proposed method. In a 6s simulation with a sampling interval 0.001s, under our designed event condition, only 2064/6000 = 34.4% data is necessary to be transmitted, reducing a huge load of communication burden. The event-triggering instant and the event release interval are shown in Fig. 4. The comparison of the norm of

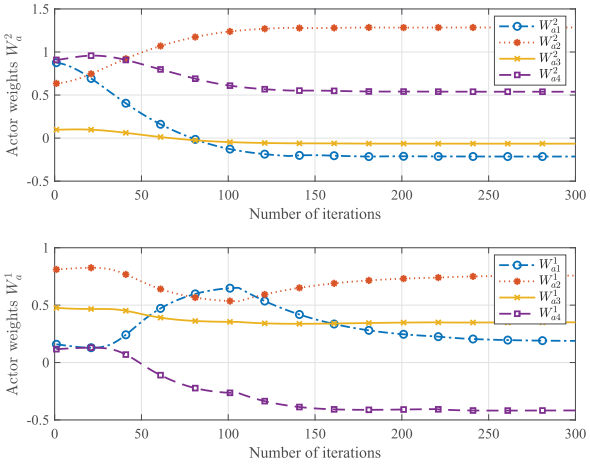

Fig. 2. The approximated actor weights W_a^i.

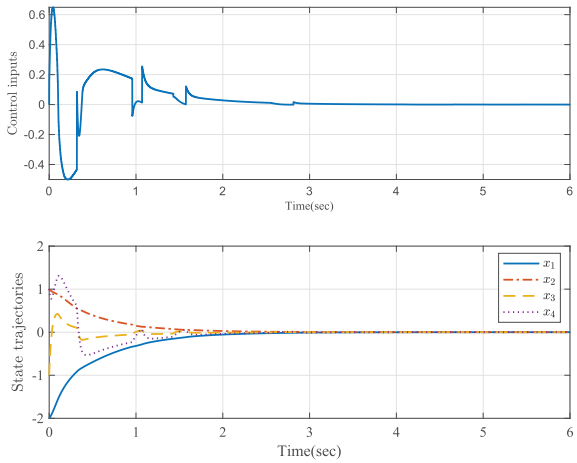

Fig. 3. State trajectories x_1, x_2, x_3, x_4 and the event-triggered control input $\hat{u}_d$ for the closed loop system.

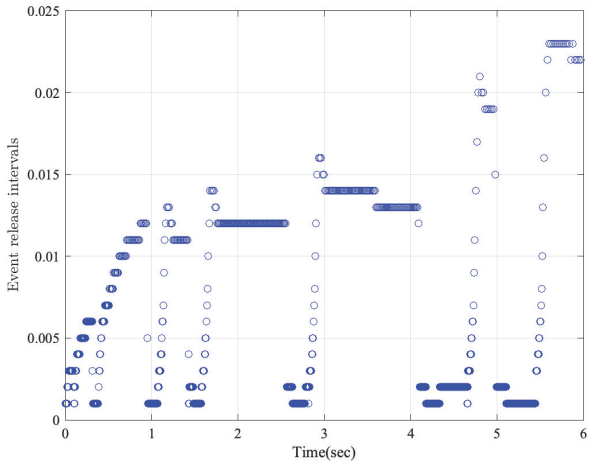

Fig. 4. Event release instant t_k and release intervals $(t_{k+1} - t_k)$.

error vector e with the event triggering threshold e_T defined in (46) is presented in Fig. 5.

The $\mathcal{H}_\infty$ attenuation level index is calculated to be $\gamma = 1.1944 \leq 2.5$. With this learned controller, we conduct a Monte

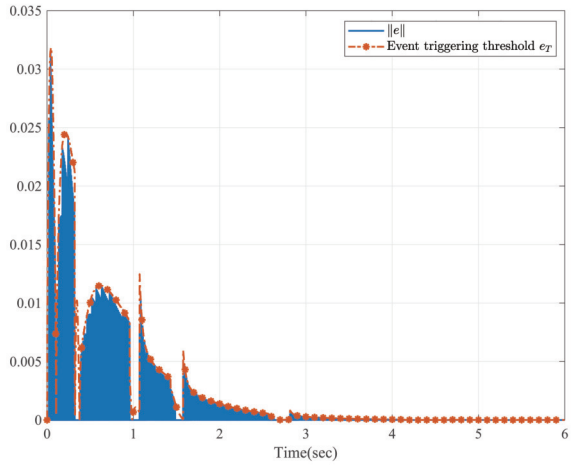

Fig. 5. The norm of state error vector $\|e\|$ and the event triggering threshold e_T.

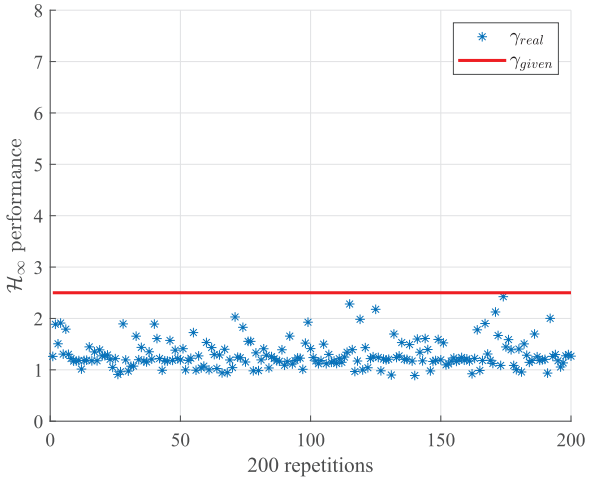

Fig. 6. $\mathcal{H}_\infty$ performance index γ_{real} vs Monte Carol simulation with 200 repetition experiments for arbitrarily obtained controller gain under given index $\gamma_{given} = 2.5$.

Carol simulation with 200 repetitions and the $\mathcal{H}_\infty$ performance is shown in Fig. 6.

To father illustrate the validity of our method, let the initial distribution of r_0 to be $\begin{bmatrix}0.2 & 0.8\end{bmatrix}$, and we conduct the experiment with the proposed learning control algorithm for 200 times.

The state trajectories are present in Fig. 7. Therefore the stochastic stability can be ensured.

To further illustrate the merits of the proposed results, here we make a brief comparison with the commonly used LMI-based methods under different cases in system matrix uncertainties to show how they perform in stability guarantee and $\mathcal{H}_\infty$ noise attenuation level index. The details of the used LMI-based methods for this comparison can be found in [36]and [37]. Let Δ be a random matrix with its values being uniformly distributed in $[0, 1]$. The system matrix uncertainties are set as $0 \cdot \Delta$ in case I, i.e., no uncertainties, 1.5Δ in case II and 2Δ in case III. The results are shown in Table I. As can be

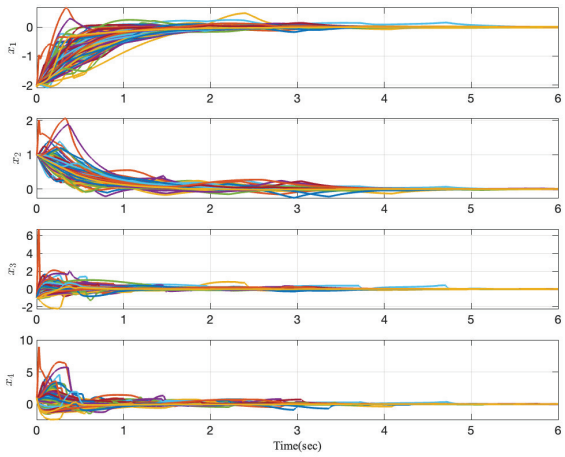

Fig. 7. State trajectories x_1, x_2, x_3, x_4 for a Monte Carol simulation with 200 repetition experiments executing the proposed algorithm.

TABLE I

COMPARISON BETWEEN LMI-BASED METHODS USED IN [36] AND [37] AND ALGORITHM 2 UNDER DIFFERENT CASES

		LMI-based methods	Algorithm 2
Case I	stability guarantee	Yes	Yes
	$\mathcal{H}_\infty$ index γ	1.1227	1.3242
Case II	stability guarantee	Yes	Yes
	$\mathcal{H}_\infty$ index γ	2.7371	1.8787
Case III	stability guarantee	No	Yes
	$\mathcal{H}_\infty$ index γ	-	2.0743

seen, when the system matrix is perfectly known and have no uncertainties, the LMI-based methods and Algorithm 2 have a similar performance index. However as the uncertainties goes higher, the LMI-based methods will result in an index larger than the predefined value, or even worse, losing its stability guarantee. On the contrary, Algorithm 2 works well in three cases, showing its robustness to system parameter uncertainties. This corresponds to what we have discussed in Remark 4.

V. CONCLUSION

This paper investigates the optimal event-triggered control problem for Markov jump linear system based on a model free Q-learning method. By recalling the zero-sum game-theoretic problem, the time-triggered optimal $\mathcal{H}_\infty$ controller gains can be computed by solving coupled game algebraic Riccati equations (CGAREs) through a mode-based parallel policy iteration algorithm. Then the event triggering mechanism is introduced where the feedback is updated in a sampled way. The corresponding triggering event is then designed by utilizing a Hamiltonian-driven adaptive dynamic programming. The difference of the optimal cost between the time triggered case and the event triggered one is also discussed. In light of the objective of developing an online model-free solution, a framework based on Q-learning and incorporating an actor-critic network has been adopted in order to approximate both the optimal cost and the event-triggered control law. The stability of the closed loop system and the convergence of the parameters associated with the network are analyzed and proved. A numerical example is presented in the end to validate the effectiveness of the proposed results. Future research directions include extending this method to address network-induced disturbances such as time delays and malicious attacks [24], as well as practical implementations on platforms of UAV [48], [49], [50], UGV [51] and mobile robots [52].

REFERENCES

[1] C. Wang, R. Li, X. Su, and P. Shi, "Output Feedback sliding mode control of Markovian jump systems and its application to switched boost converter," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 68, no. 12, pp. 5134–5144, Dec. 2021.

[2] H. Wan, X. Luan, H. R. Karimi, and F. Liu, "Dynamic self-triggered controller codesign for Markov jump systems," *IEEE Trans. Autom. Control*, vol. 66, no. 3, pp. 1353–1360, Mar. 2021.

[3] O. L. V. Costa and A. M. de Oliveira, "Feedback linear quadratic Nash equilibrium for discrete-time Markov jump linear systems," *Syst. Control Lett.*, vol. 192, Oct. 2024, Art. no. 105893.

[4] Z. Ni, F. Li, Y. Wang, and H. Shen, "Sliding-mode control for 2-D hidden Markov jump Roesser systems with partial information and its application in metal rolling process," *IEEE Trans. Autom. Sci. Eng.*, vol. 22, pp. 6851–6859, 2025.

[5] J. N. A. D. Bueno, L. B. Marcos, K. D. T. Rocha, and M. H. Terra, "Regulation of uncertain Markov jump linear systems with application on automotive powertrain control," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 53, no. 8, pp. 5019–5031, Aug. 2023.

[6] Y. Zacchia Lun, F. Smarra, and A. D'Innocenzo, "Optimal control over Markovian wireless communication channels under generalized packet dropout compensation," 2025, *arXiv:2501.17105*.

[7] O. L. V. Costa and A. M. de Oliveira, "Filtering for nonlinear and linear Markov jump systems," *IEEE Trans. Autom. Control*, vol. 69, no. 5, pp. 3309–3316, May 2024.

[8] C. Xiao, T. Hou, W. Zhang, and F. Deng, "Detectability, Riccati equations, and the game-based control of discrete-time MJLSs with the Markov chain on a borel space," 2025, *arXiv:2502.13404*.

[9] P. Cheng, H. Chen, S. He, and W. Zhang, "Asynchronous deconvolution filtering for 2-D Markov jump systems with packet loss compensation," *IEEE Trans. Autom. Sci. Eng.*, vol. 21, no. 3, pp. 4165–4176, Jul. 2024.

[10] Y. Zhao, L. Wang, X. Xie, H.-K. Lam, and J. Gu, "Asynchronous switching control for fuzzy Markov jump systems with periodically varying delay and its application to electronic circuits," *IEEE Trans. Autom. Sci. Eng.*, vol. 22, pp. 1100–1111, 2025.

[11] S. Wang and Z.-G. Wu, "Asynchronous Energy-to-Peak filtering for restricted Markov jump systems," *IEEE Trans. Autom. Control*, vol. 69, no. 5, pp. 2823–2838, May 2024.

[12] F. Stadtmann and O. L. V. Costa, "H_2-control of continuous-time hidden Markov jump linear systems," *IEEE Trans. Autom. Control*, vol. 62, no. 8, pp. 4031–4037, Aug. 2017.

[13] R. Weng et al., "Bumpless transfer MPC for hidden Markovian jump nonlinear systems with saturated inputs," *IEEE Trans. Autom. Sci. Eng.*, vol. 22, pp. 916–925, 2025.

[14] K. Zhang and Y. Peng, "Model-free tracking control for linear stochastic systems via integral reinforcement learning," *IEEE Trans. Autom. Sci. Eng.*, vol. 22, pp. 10835–10844, 2025.

[15] J. Tan, S. Xue, H. Li, Z. Guo, H. Cao, and D. Li, "Prescribed performance robust approximate optimal tracking control via Stackelberg game," *IEEE Trans. Autom. Sci. Eng.*, vol. 22, pp. 12871–12883, 2025.

[16] H. Zhang, Y. Liang, H. Su, and C. Liu, "Event-driven guaranteed cost control design for nonlinear systems with actuator faults via reinforcement learning algorithm," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 50, no. 11, pp. 4135–4150, Nov. 2020.

[17] L. Zhang, H. Zhang, J. Sun, and X. Yue, "ADP-based fault-tolerant control for multiagent systems with semi-Markovian jump parameters," *IEEE Trans. Cybern.*, vol. 54, no. 10, pp. 5952–5962, Oct. 2024.

[18] X. Shi, Y. Li, C. Du, C. Chen, G. Zong, and W. Gui, "Reinforcement learning-based optimal control for Markov jump systems with completely unknown dynamics," *Automatica*, vol. 171, Jan. 2025, Art. no. 111886.

[19] F. L. Lewis and D. Vrabie, "Reinforcement learning and adaptive dynamic programming for Feedback control," *IEEE Circuits Syst. Mag.*, vol. 9, no. 3, pp. 32–50, Aug. 2009.

[20] D. Liu, S. Xue, B. Zhao, B. Luo, and Q. Wei, "Adaptive dynamic programming for control: A survey and recent advances," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 51, no. 1, pp. 142–160, Jan. 2021.
[21] K. Zhang, H.-g. Zhang, Y. Cai, and R. Su, "Parallel optimal tracking control schemes for mode-dependent control of coupled Markov jump systems via integral RL method," *IEEE Trans. Autom. Sci. Eng.*, vol. 17, no. 3, pp. 1332–1342, Jul. 2020.
[22] L. Zhang, H. Zhang, X. Yue, and T. Wang, "Actor-critic optimal control for semi-Markovian jump systems with time delay," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 71, no. 4, pp. 2164–2168, Apr. 2024.
[23] K. Zhang, R. Su, and H. Zhang, "A novel resilient control scheme for a class of Markovian jump systems with partially unknown information," *IEEE Trans. Cybern.*, vol. 52, no. 8, pp. 8191–8200, Aug. 2022.
[24] H. Shen, Y. Wang, J. Wu, J. H. Park, and J. Wang, "Secure control for Markov jump cyber-physical systems subject to malicious attacks: A resilient hybrid learning scheme," *IEEE Trans. Cybern.*, vol. 54, no. 11, pp. 7068–7079, Nov. 2024.
[25] X. Zhong, H. He, H. Zhang, and Z. Wang, "Optimal control for unknown discrete-time nonlinear Markov jump systems using adaptive dynamic programming," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 25, no. 12, pp. 2141–2155, Dec. 2014.
[26] Y. Yang, K. G. Vamvoudakis, H. Modares, Y. Yin, and D. C. Wunsch, "Hamiltonian-driven hybrid adaptive dynamic programming," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 51, no. 10, pp. 6423–6434, Oct. 2021.
[27] J. Zhang, D. Yang, H. Zhang, Y. Wang, and B. Zhou, "Dynamic event-based tracking control of boiler turbine systems with guaranteed performance," *IEEE Trans. Autom. Sci. Eng.*, vol. 21, no. 3, pp. 4272–4282, Jul. 2024.
[28] F. Wang, Z. Hu, C. Zhou, and X. Lin, "Dynamic event-triggered consensus control of multiagent systems under Markov switching topology," *IEEE Trans. Autom. Sci. Eng.*, vol. 22, pp. 11163–11175, 2025.
[29] A. T. Demora and C. M. Abdissa, "Neural network-based lower limb prostheses control using super twisting sliding mode control," *IEEE Access*, vol. 13, pp. 24929–24953, 2025.
[30] X. Yang and Q. Wei, "Adaptive dynamic programming for robust event-driven tracking control of nonlinear systems with asymmetric input constraints," *IEEE Trans. Cybern.*, vol. 54, no. 11, pp. 6333–6344, Nov. 2024.
[31] D. Yang, Y. Sun, G. Zong, X. Zhao, and J. Xie, "Adaptive dynamic event-triggered tracking control for uncertain switched nonlinear systems under state-dependent switching," *IEEE Trans. Cybern.*, vol. 54, no. 12, pp. 7742–7753, Dec. 2024.
[32] Y. Hu, H. Yan, H. Zhang, M. Wang, and C. Chen, "Adaptive neural network output-Feedback control for uncertain nonlinear systems via event-triggered output," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 54, no. 10, pp. 5864–5875, Oct. 2024.
[33] Y.-Y. Tao, Z.-G. Wu, T. Huang, P. Chakrabarti, and C. K. Ahn, "Asynchronous event-triggered output-Feedback control of singular Markov jump systems," *IEEE Trans. Cybern.*, vol. 53, no. 11, pp. 7342–7352, Nov. 2023.
[34] Y. Gu, M. Shen, J. H. Park, Q.-G. Wang, Y. Yi, and Y. Sun, "Dynamic event-triggered H_∞ filtering for fuzzy Markov jump systems subject to mismatched quantization," *IEEE Trans. Autom. Sci. Eng.*, vol. 22, pp. 10639–10649, 2025.
[35] L. Yao, X. Huang, Z. Wang, and H. Shen, "Memory-based event-triggered control of Markov jump systems under hybrid cyber attacks: A switching-like adaptive law," *IEEE Trans. Autom. Sci. Eng.*, vol. 21, no. 4, pp. 6347–6357, Oct. 2024.
[36] Z. Xu, Y. Yu, Z.-G. Wu, and Y. Shen, "Asynchronous event-triggered H_∞ control for continuous-time Markov jump systems," *IEEE Control Syst. Lett.*, vol. 8, pp. 3386–3391, 2024.
[37] P. Zeng, F. Deng, Z.-H. Wu, T. Zhang, and X. Gao, "Event-triggered multiasynchronous H_∞ control for Markov jump systems with transmission delay," *IEEE Trans. Cybern.*, vol. 54, no. 1, pp. 476–485, Jan. 2024.
[38] H. Fang et al., "Adaptive optimization algorithm for nonlinear Markov jump systems with partial unknown dynamics," *Int. J. Robust Nonlinear Control*, vol. 31, no. 6, pp. 2126–2140, Apr. 2021.
[39] J. Wang, J. Wu, J. Cao, M. Chadli, and H. Shen, "Nonfragile output Feedback tracking control for Markov jump fuzzy systems based on integral reinforcement learning scheme," *IEEE Trans. Cybern.*, vol. 53, no. 7, pp. 4521–4530, Jul. 2023.
[40] J. P. Jansch-Porto, B. Hu, and G. E. Dullerud, "Policy optimization for Markovian jump linear quadratic control: Gradient method and global convergence," *IEEE Trans. Autom. Control*, vol. 68, no. 4, pp. 2475–2482, Apr. 2023.
[41] Z. Xu, Z.-G. Wu, H. Que, and P. Jiang, "Asynchronous control of nonlinear Markov jump systems with uncertainties using interval Type-2 polynomial fuzzy approach," *IEEE Trans. Fuzzy Syst.*, vol. 32, no. 3, pp. 883–892, Mar. 2024.
[42] L. Ye, Z. Zhao, and F. Liu, "Reinforcement learning for finite-horizon H_∞ tracking control of unknown discrete linear time-varying system," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 54, no. 10, pp. 6385–6396, Oct. 2024.
[43] K. G. Vamvoudakis and F. R. Pour Safaei, "Stochastic zero-sum Nash games for uncertain nonlinear Markovian jump systems," in *Proc. IEEE 56th Annu. Conf. Decis. Control (CDC)*, Dec. 2017, pp. 5582–5589.
[44] A. Al-Tamimi, F. L. Lewis, and M. Abu-Khalaf, "Model-free Q-learning designs for linear discrete-time zero-sum games with application to H-infinity control," *Automatica*, vol. 43, no. 3, pp. 473–481, Mar. 2007.
[45] J. Song, S. He, Z. Ding, and F. Liu, "A new iterative algorithm for solving H_∞ control problem of continuous-time Markovian jumping linear systems based on online implementation," *Int. J. Robust Nonlinear Control*, vol. 26, no. 17, pp. 3737–3754, Nov. 2016.
[46] D. Vrabie, O. Pastravanu, M. Abu-Khalaf, and F. L. Lewis, "Adaptive optimal control for continuous-time linear systems based on policy iteration," *Automatica*, vol. 45, no. 2, pp. 477–484, Feb. 2009.
[47] Z. Gajic and I. Borno, "Lyapunov iterations for optimal control of jump linear systems at steady state," *IEEE Trans. Autom. Control*, vol. 40, no. 11, pp. 1971–1975, Nov. 1995.
[48] M. M. Madebo, C. M. Abdissa, L. N. Lemma, and D. S. Negash, "Robust tracking control for quadrotor UAV with external disturbances and uncertainties using neural network based MRAC," *IEEE Access*, vol. 12, pp. 36183–36201, 2024.
[49] F. T. Yareshe, N. W. Madebo, C. M. Abdissa, and L. N. Lemma, "Trajectory tracking of fixed-wing UAV using ANFIS-based sliding mode controller," *IEEE Access*, vol. 13, pp. 61986–62003, 2025.
[50] E. A. Gedefaw, C. M. Abdissa, and L. N. Lemma, "An improved trajectory tracking control of quadcopter using a novel sliding mode control with fuzzy PID surface," *PLoS ONE*, vol. 19, no. 11, Nov. 2024, Art. no. e0308997.
[51] J. Tan, S. Xue, Q. Guan, T. Niu, H. Cao, and B. Chen, "Unmanned aerial-ground vehicle finite-time docking control via pursuit-evasion games," *Nonlinear Dyn.*, vol. 113, no. 13, pp. 16757–16777, Jul. 2025.
[52] Y. A. Wendemagegn, W. Ayalew Asfaw, C. M. Abdissa, and L. N. Lemma, "Enhancing trajectory tracking accuracy in three-wheeled mobile robots using backstepping fuzzy sliding mode control," *Eng. Res. Exp.*, vol. 6, no. 4, Oct. 2024, Art. no. 045204.

Zhaowen Xu (Member, IEEE) was born in 1991. He received the B.S. degree from Wuhan University, Wuhan, China, in 2013, and the Ph.D. degree in control science and engineering from Zhejiang University, Hangzhou, China, in 2018.

He is currently an Associate Professor with the School of Information Science and Technology, Hangzhou Normal University, Hangzhou. His research interests include control and filtering for Markov jump systems and applications.

Zheng-Guang Wu was born in 1982. He received the B.S. and M.S. degrees in mathematics from Zhejiang Normal University, Jinhua, China, in 2004 and 2007, respectively, and the Ph.D. degree in control science and engineering from Zhejiang University, Hangzhou, China, in 2011.

He is currently with the Institute of Cyber-Systems and Control, Zhejiang University. His current research interests include hybrid systems, networked systems, and computational intelligence.