Learning-Based Event-Triggered Tracking Control for Nonlinear Networked Control Systems With Unmatched Disturbance

Jinliang Liu, Nan Zhang, Yan Li, Xiangpeng Xie, *Member, IEEE*, Engang Tian, and Jinde Cao, *Fellow, IEEE*

***Abstract*—This article concentrates on optimal tracking control for a class of nonlinear networked systems subjecting to limited network bandwidth and unmatched disturbance. Given the models of the control and reference systems, the considered optimal tracking control issue is initially formulated as a minimax optimization problem. Then, with the introduction of an event-triggered mechanism used for saving bandwidth, the formulated problem is transformed into solving an event-based Hamilton-Jacobi-Isaacs (HJI) equation by recurring to the Bellman optimality theory. Based on the HJI equation, we demonstrate that the stability of the concerned system in the sense of uniformly ultimately bounded (UUB) can be guaranteed with the envisioned optimal control and worst disturbance policies. Here, the disturbance policy can be varied periodically while the control policy can only be updated at event-triggering instants, which differs from the existed researches. Furthermore, we propose a reinforcement learning (RL)-based algorithm to handle the constructed HJI equation and thus settle the studied tracking control problem. The effectiveness of the algorithm is finally validated by both theoretical analysis and simulations.**

***Index Terms*—Event-triggered mechanism, networked control systems (NCSs), reinforcement learning (RL), tracking control, uniformly ultimately bounded (UUB).**

I. Introduction

NOWADAYS, networked control systems (NCSs) are acknowledged as the fundamental construction elements in modern control field. Through wired or wireless communication networks, NCSs can enable remote control in a flexible and cost-efficient manner, and thus have been extensively used in many practical applications. However, with the increasing complexity of internal structures and external factors, obvious nonlinearities are observed from NCSs, so it is greatly demanding to study nonlinear NCSs. To this end, lots of research issues, such as asynchronous control, consensus control, security filtering, and tracking control, have been conducted over nonlinear NCSs in the literatures. Among which, tracking control has attracted considerable attentions given its wide practicability [1], [2], [3], [4], [5]. Generally, the objective of tracking control is to design appropriate control strategies so as to synchronize the state trajectory of the controlled system with that of a preselected reference system. In the existed studies, many tracking control schemes for nonlinear NCSs have been proposed under diverse scenarios [6], [7], [8], [9], [10]. To mention a few, Zhang et al. [6] proposed three stable adaptive tracking schemes to deal with the tracking control issue for a class of uncertain MIMO nonlinear systems suffering from time-varying disturbance. In [7], a novel adaptive fuzzy control strategy was designed to guarantee the convergence of the tracking error and observation error for nonlinear systems with completely unknown dynamics. Besides, the fuzzy tracking control problem was settled in [8] for pure-feedback random nonlinear systems with input saturation. For second-order nonlinear multiagents systems, the expected tracking performance was achieved in [9] by a designed distributed sliding-mode control strategy. Yu et al. [10] investigated the adaptive tracking problem over uncertain switched random strict-feedback nonlinear systems.

The aforementioned researches provide many effective solutions to the tracking control problem over nonlinear NCSs, however, none of them explores the optimality of tracking control, where the optimality can be defined as the minimization of control cost or maximization of long-term control performance. With the rise of optimal control theory, optimal control in NCSs has extracted extensive research interests. A typical method for handling the optimal control issue is to transform the problem into solving the Hamilton-Jacobi-Isaacs (HJI) equation [11], [12], [13], [14], [15], [16], [17], [18], [19]. In view of this, the adaptive dynamic programming (ADP) method is initially adopted to obtain optimal control solution. For instance, Li et al. [11] proposed an ADP

Manuscript received 5 September 2022; revised 5 October 2022; accepted 14 November 2022. Date of publication 8 December 2022; date of current version 17 April 2023. This work was supported in part by the National Natural Science Foundation of China under Grant 61973152 and Grant 62273174; in part by the Natural Science Foundation of Jiangsu Province of China under Grant BK20211290; and in part by the Jiangsu Postgraduate Research Innovation Project under Grant KYCX21_1507. This article was recommended by Associate Editor Z. Liu. *(Corresponding author: Yan Li.)*

Jinliang Liu, Nan Zhang, and Yan Li are with the College of Information Engineering, Nanjing University of Finance and Economics, Nanjing 210023, Jiangsu, China (e-mail: liujinliang@vip.163.com; zhangnan_nufe@163.com; ylnjue@163.com).

Xiangpeng Xie is with the Institute of Advanced Technology, Nanjing University of Posts and Telecommunications, Nanjing 210023, Jiangsu, China (e-mail: xiexiangpeng1953@163.com).

Engang Tian is with the School of Optical-Electrical and Computer Engineering, University of Shanghai for Science and Technology, Shanghai 200093, China (e-mail: tianengang@163.com).

Jinde Cao is with the School of Mathematics and the Research Center for Complex Systems and Network Sciences, Southeast University, Nanjing 210096, Jiangsu, China (e-mail: jdcao@seu.edu.cn).

Color versions of one or more figures in this article are available at https://doi.org/10.1109/TSMC.2022.3224432.

Digital Object Identifier 10.1109/TSMC.2022.3224432

online algorithm to solve the optimal tracking control problem for nonlinear affine system with partial uncertain time-delay. The optimal control problems for discrete-time systems were settled in [12] by invariant ADP methods. Furthermore, reinforcement learning (RL) is viewed as an another promising technique for solving the optimal control problem, which differs from ADP in that the training model is avoided to be established [13]. RL-based strategies have been widely designed to solve the HJI equation in recent studies. For example, the optimal control problem for affine nonlinear discrete-time systems was investigated in [14] with a novel off-policy interleaved Q-learning algorithm. For a class of weakly coupled continuous-time nonlinear systems, a data-based online policy iteration learning algorithm was designed to solve the optimal robust problem in [15]. Besides, Pang and Jiang [16] proposed a novel infinite-horizon optimal control problem for continuous-time linear periodic (CTLP) systems. The optimal cooperative consensus problem for nonlinear multiagents systems was settled in [17] with the designed model-free policy iterative algorithm. In addition, a new data-driven performance-prescribed RL control (DPRLC) scheme was designed in [18] to address the optimal control issue for the unmanned surface vehicle.

With the comprehensive exploration on RL-based optimal control solutions, the researchers observed that the restriction of communication resources largely affects the design of corresponding strategies. Actually, the network bandwidth in NCSs is becoming limited given the expansion of the system scale and complexity of the controlled objectives. Then, event-triggered mechanisms, by which system signals are released into the communication network only while a specific triggering condition is satisfied, are now commonly utilized to alleviate the pressure on communication bandwidth in NCSs [20], [21], [22], [23], [24], [25]. Under event-triggered mechanisms, system controllers generally update their control policies at triggering instants. Thus, a number of event-triggered optimal control algorithms were proposed based on RL technique in [26], [27], [28], and [29]. To be specific, the event-triggered-based optimal tracking control problem over multiagents systems was settled based on multigradients distributed RL approach in [27]. An effective event-triggered data transmission strategy was designed in [28] to solve optimal tracking control issue occurred in nonlinear interconnected systems. Considering a class of nonlinear systems with actuator faults and limited communication resources, the event-driven control strategy and RL algorithm were presented in [29] to achieve the guaranteed control performance. For continuous-time nonlinear systems with disturbance, the event-driven H_∞ constrained optimal control problem was solved via single adaptive critic learning algorithm in [26].

In the above-mentioned researches, nevertheless, the triggering conditions are generally derived with the primary purpose of stabilizing the targeted systems rather than reducing the transmission of redundant system signals. Under the limited communication resources situation, it is more promising to adopt bandwidth-focused event-triggering schemes, i.e., the schemes dedicated to save network bandwidth, and then design effective RL-based algorithm to assure the stability of the considered systems. In addition, practical NCSs are inevitably affected by external disturbances, which thus is considered while designing optimal control strategies in some of the existed works [30], [31], [32], [33], [34]. For instance, the event-triggered zero-sum game problem was focused in [31] for partially unknown continuous-time nonlinear systems with unmatched disturbance. It is worth noting that the studies assumed that the disturbance policies are updated at the triggering instants. Actually, a more realistic scenario is to consider sampling-driven disturbance policies, i.e., the policies are updated at the sampling instants and then influence the occurrence of the "events".

Motivated by the above analysis, this article will study the event-triggered optimal tracking control problem over a nonlinear NCS with limited network bandwidth and sampling-driven unmatched disturbance, where the meaning of unmatched disturbance will be explained shortly. Specifically, we will adopt a bandwidth-focused event-triggered mechanism (BFETM) to ease the communication pressure, and then explore a RL-based optimal tracking control algorithm for the concerned system. The major contributions of the work can be summarized as follows.

1) To alleviate the communication constraint, a bandwidth-focused event-triggering scheme is adopted to effectively reduce the transmission of redundant data.
2) The studied event-triggering optimal tracking control issue is formulated as a minimax optimization problem, and then transferred into solving an event-triggered HJI equation based on the Bellman optimality theory.
3) A RL-based iterative algorithm is proposed to solve the established HJI equation, and thus the stability of the tracking error and approximated weight vector error can be guaranteed in the sense of uniformly ultimately bounded (UUB).
4) The effectiveness of the designed tracking control method is validated by conducting extensive simulations.

The remainder of this article is arranged as follows. In Section II, based on the depiction of the envisioned system and the adopted event-triggered mechanism, an event-based HJI equation is constructed to formulate the considered optimal tracking control problem. In Section III, the sufficient conditions for assuring the UUB of the concerned system are first analyzed, and then a RL-based iterative algorithm is proposed to obtain an efficient solution to the studied tracking control problem. The performance of the designed method is evaluated by conducting simulations in Section IV. Section V concludes this article.

II. System Model and Problem Formulation

In this section, the considered nonlinear NCS and reference model are described first, then the studied event-triggered optimal tracking control problem is formulated based on the Bellman optimality theory.

A. System Description

Considering a continuous-time NCS with unmatched disturbance as follows:

$$\dot{x}(t) = f(x(t)) + g(x(t))u(t) + h(x(t))w(t) \tag{1}$$

where $x(t) \in \mathbb{R}^n$ and $u(t) \in \mathbb{R}^m$ denote the system states and control input, respectively; $w(t) \in \mathbb{R}^p$ is the uncertainty disturbance and $w(t) \in L_2[0, \infty)$; $f(x(t)) \in \mathbb{R}^n$ is the known nonlinear function, $g(x(t)) \in \mathbb{R}^{n\times m}$ is the input gain function, and $h(x(t)) \in \mathbb{R}^{n\times p}$ is the disturbance gain function.

Remark 1: As mentioned in [35], for NCSs with disturbance, if $g(x(t)) = h(x(t))$, then the systems are said to suffer from the matched disturbance, otherwise, the systems are considered to work under the unmatched disturbance. The envisioned system in this article is assumed to be affected by unmatched disturbance.

The selected tracking reference model is given as

$$\dot{r}(t) = \phi(r(t)) \tag{2}$$

where $r(t) \in \mathbb{R}^n$ is the state of the reference model and $\phi(r(t))$ denotes the nonlinear continuous function. Based on (1) and (2), the tracking error is defined as $z(t) = x(t) - r(t)$, then the tracking error system can be obtained as

$$\dot{z}(t) = F(x(t), r(t)) + g(x(t))u(t) + h(x(t))w(t) \tag{3}$$

in which $F(x(t), r(t)) = f(x(t)) - \phi(r(t))$.

Definition 1 (UUB) [26]: The tracking error $z(t)$ is said to be stable in the sense of UUB, if there exist a compact set $\Omega \subset \mathbb{R}^n$, a constant θ and a time instant $t_f = \mathbb{T}(\theta, z_0)$ that $||z(t)|| \le \theta$ for all $t \ge t_0 + t_f$ with initial value $z(t_0) = z_0 \in \Omega$.

B. Problem Formulation

Focusing on the formulated tracking error system (3), the objective of the study is to design an optimal control policy to guarantee that the system is stable in the sense of UUB, and has L_2 gain no more than σ which can be formally described as

$$\int_0^\infty \left[z^T(t)Pz(t) + u^T(t)Ru(t)\right]dt \le \sigma^2 \int_0^\infty w^T(t)w(t)dt \tag{4}$$

where $P \in \mathbb{R}^{n\times n}$ and $Q \in \mathbb{R}^{m\times m}$ are the symmetric positive definite constant matrices; $\sigma > 0$ is the pregiven disturbance attenuation level. To achieve the objective, referring to [26], a value function is first designed as

$$V(z(t)) = \int_t^\infty e^{-a(\tau-t)}\Big[z^T(\tau)Pz(\tau) + u^T(\tau)Ru(\tau) - \sigma^2 w^T(\tau)w(\tau)\Big]d\tau \tag{5}$$

in which $a > 0$ denotes the discount factor and the decay term $e^{-a(\tau-t)}$ is introduced to ensure the convergence of $V(z(t))$. Then, by referring to [36], the studied H_∞ optimal tracking control problem can be modeled as

$$V^*(z(t)) = \min_{u\in\Phi_1(\Omega_1)} \max_{w\in\Phi_2(\Omega_2)} V(z(t)) \tag{6}$$

in which $\Phi_1(\Omega_1)$ and $\Phi_2(\Omega_2)$ represent the admissible sets of control input $u(t)$ and disturbance $w(t)$, respectively.

According to the Bellman optimality theory, the formulated minimax optimization problem in (6) can be transformed into solving the following HJI equation:

$$\min_{u\in\Phi_1(\Omega_1)} \max_{w\in\Phi_2(\Omega_2)} H\Big(z(t), V^*_{z(t)}, u(z(t)), w(z(t))\Big) = 0. \tag{7}$$

$H(\cdot)$ is called the Hamilton function and is defined as

$$\begin{aligned} &H\Big(z(t), V^*_{z(t)}, u(z(t)), w(z(t))\Big) \\ &= \Big(V^*_{z(t)}\Big)^T \dot{z}(t) + z^T(t)Pz(t) + u^T(z(t))Ru(z(t)) \\ &\quad - \sigma^2 w^T(z(t))w(z(t)) - aV^*(z(t)) \end{aligned} \tag{8}$$

where $V^*_{z(t)}$ denotes the partial derivative of $V^*(z(t))$ with respect to $z(t)$, i.e., $V^*_{z(t)} = ([\partial V^*(z(t))]/[\partial z(t)])$.

Following the method used in [37], we then set

$$\begin{cases} \dfrac{\partial H\Big(z(t), V^*_{z(t)}, u(z(t)), w(z(t))\Big)}{\partial u(z(t))} = 0 \\ \dfrac{\partial H\Big(z(t), V^*_{z(t)}, u(z(t)), w(z(t))\Big)}{\partial w(z(t))} = 0 \end{cases} \tag{9}$$

and obtain the optimal control input $u^*(z(t))$ and worst disturbance $w^*(z(t))$ as

$$\begin{cases} u^*(z(t)) = -\frac{1}{2}R^{-1}g^T(x(t))V^*_{z(t)} \\ w^*(z(t)) = \frac{1}{2\sigma^2}h^T(x(t))V^*_{z(t)}. \end{cases} \tag{10}$$

Based on $u^*(z(t))$ and $w^*(z(t))$, the HJI equation described by (7) can be rewritten as

$$\begin{aligned} &\min_{u\in\Phi_1(\Omega_1)} \max_{w\in\Phi_2(\Omega_2)} H\Big(z(t), V^*_{z(t)}, u(z(t)), w(z(t))\Big) \\ &= H\Big(z(t), V^*_{z(t)}, u^*(z(t)), w^*(z(t))\Big) \\ &= \Big(V^*_{z(t)}\Big)^T \big(F(x(t), r(t)) + g(x(t))u^*(z(t)) \\ &\qquad + h(x(t))w^*(z(t))\big) + \Xi^*\big(z(t), u^*(z(t)), w^*(z(t))\big) \\ &\quad - aV^*(z(t)) \\ &= 0 \end{aligned} \tag{11}$$

where

$$\begin{aligned} &\Xi^*\big(z(t), u^*(z(t)), w^*(z(t))\big) \\ &= z^T(t)Pz(t) + \big(u^*(z(t))\big)^T Ru^*(z(t)) \\ &\quad - \sigma^2\big(w^*(z(t))\big)^T w^*(z(t)). \end{aligned}$$

In order to effectively save network communication resources, a BFETM is adopted in this article. To be specific, the sampled data can be released only while the following event-triggering condition holds:

$$d_j^T(t)\Omega d_j(t) > z^T\big(t_j + lh\big)\Omega z\big(t_j + lh\big) \tag{12}$$

where $d_j(t) = z(t_j + lh) - z(t_j)$ and $l \in \{1, 2, 3, \ldots\}$; h denotes the sampling period; $z(t_j)$ is the transmitted data at the latest event-triggered instant t_j (t_j is a multiple of h and $t_0 = 0$); $z(t_j+lh)$ denotes the current sampled data; and Ω is a symmetric positive definite matrix with appropriate dimension. Based on (12), the next triggering instant t_{j+1} ($t_{j+1} > t_j$) can be expressed as

$$t_{j+1} = t_j + \min_{l\ge 1}\{lh|lh \text{ satisfying (12)}\}. \tag{13}$$

It is generally acknowledged that the control input $u(z(t))$ is unchanged during the two adjacent event-triggering instants, thus, we have

$$u(z(t)) = u\big(z(t_j)\big), \qquad t \in \big[t_j, t_{j+1}\big). \tag{14}$$

Under the above event-triggered mechanism, (10) can be rewritten as

$$\begin{cases} u^*(z(t)) = u^*\big(z(t_j)\big) = -\frac{1}{2}R^{-1}g^T\big(x(t_j)\big)V^*_{z(t_j)} \\ \qquad\qquad\qquad\qquad t \in \big[t_j, t_{j+1}\big) \\ w^*(z(t)) = \frac{1}{2\sigma^2}h^T(x(t))V^*_{z(t)} \end{cases} \tag{15}$$

and the event-triggered HJI equation can be modified as

$$\begin{aligned} &H\Big(z(t), V^*_{z(t)}, u^*\big(z(t_j)\big), w^*(z(t))\Big) \\ &= \Big(V^*_{z(t)}\Big)^T\big(F(x(t), r(t)) + g(x(t))u^*\big(z(t_j)\big) \\ &\quad + h(x(t))w^*(z(t))\big) + \Xi^*\big(z(t), u^*\big(z(t_j)\big), w^*(z(t))\big) \\ &\quad - aV^*(z(t)) \\ &= 0, \qquad t \in \big[t_j, t_{j+1}\big). \end{aligned} \tag{16}$$

Remark 2: In this article, we follow the general rule that the control policy $u(z(t))$ can only be updated at the event-triggering instants, meanwhile, we consider a more realistic scenario that the disturbance policy $w(z(t))$ will be updated at sampling instants, which may further affect the appearance of the specific event-triggering instants. Based on the event-driven control policy and sampling-driven disturbance policy, the objective event-triggered HJI equation is thus formulated by (16).

Remark 3: The adopted event-triggered scheme presented in (12) is devoted to reduce the redundant data transmission, i.e., the sampled data can be released only while the predesigned condition (12) is satisfied. Under such a BFETM, this article then dedicates to prove that the formulated tracking error system is stable in the sense of UUB. This differs from the relative researches [26], [27], [28], [29], which derive the event-triggering conditions during stabilizing the envisioned systems.

C. Stability Analysis

Before giving the specific stability analysis, the following three assumptions are first introduced (for easing description, $z(t)$, $z(t_j)$, $x(t)$, $x(t_j)$, $d_j(t)$, and $r(t)$ are referred to as z, z_j, x, x_j, d_j, and r, respectively, in the remainder of this article).

Assumption 1: Supposing that $u^*(z)$ is Lipschitz continuous on admissible set $\Phi_1(\Omega_1)$, which means that for all $u^*(z)$ and $u^*(z_j) \in \Phi_1(\Omega_1)$, there exists a positive constant K_{u^*} that makes the following condition holds:

$$||u^*(z) - u^*\big(z_j\big)|| \le K_{u^*}||z - z_j|| = K_{u^*}\big\|d_j\big\|. \tag{17}$$

It should be noted that the optimal value function $V^*(z)$ is continuously differentiable on its admissible set $\Phi_3(\Omega_3)$ [36], which means that both $V^*(z)$ and V^*_z are bounded on $\Phi_3(\Omega_3)$. Thus, we further give the following assumptions.

Assumption 2: Supposing that $V^*(z) \le \eta_m$ and $V^*_z \le \eta_{em}$, where η_m and η_{em} are positive constants.

Assumption 3: Supposing the system functions $g(x)$ and $h(x)$ are bounded by $||g(x)|| \le g_m$ and $||h(x)|| \le h_m$, respectively, in which g_m and h_m are positive constants.

Theorem 1: Given the optimal value function $V^*(z)$, the nonlinear tracking error system (3) is considered to be stable in the sense of UUB under the optimal control input $u^*(z)$ and worst disturbance $w^*(z)$ in (15), Assumptions 1 and 2, if the following condition holds:

$$\lambda_{\min}(P) - \lambda_{\max}(R)K^2_{u^*}\frac{\lambda_{\max}(\Omega)}{\lambda_{\min}(\Omega)} \ge 0. \tag{18}$$

Proof: For $V^*(z)$ which is the solution to the event-triggered HJI (16), it can be obtained that $V^*(z) \ge 0$ as illustrated in [38]. Therefore, we choose $V^*(z)$ as the Lyapunov function and take the derivative of it along with $\dot{z}$ defined in (3), then, we have

$$\begin{aligned} \dot{L} = \dot{V}^*(z) &= \big(V^*_z(z)\big)^T\dot{z} \\ &= \big(V^*_z(z)\big)^T\big(F(x, r) + g(x)u^*\big(z_j\big) + h(x)w^*(z)\big) \\ &= \big(V^*_z(z)\big)^T\big(F(x, r) + g(x)u^*(z) + h(x)w^*(z)\big) \\ &\quad + \big(V^*_z(z)\big)^T g(x)\big(u^*\big(z_j\big) - u^*(z)\big) \end{aligned} \tag{19}$$

in which the first item can be rewritten as

$$\begin{aligned} &\big(V^*_z\big)^T\big(F(x, r) + g(x)u^*(z) + h(x)w^*(z)\big) \\ &= aV^*(z) - \Xi^*\big(z, u^*(z), w^*(z)\big) \end{aligned} \tag{20}$$

according to (11). Moreover, based on (10), we can get

$$\begin{cases} \big(V^*_z\big)^T g(x) = -2(u^*(z))^T R \\ \big(V^*_z\big)^T h(x) = 2\sigma^2(w^*(z))^T. \end{cases} \tag{21}$$

Substituting (20), (21) into (19), we can further obtain

$$\begin{aligned} \dot{L} &= aV^*(z) - \Xi^*\big(z, u^*(z), w^*(z)\big) \\ &\quad - 2\big(u^*(z)\big)^T R\big(u^*\big(z_j\big) - u^*(z)\big) \\ &= aV^*(z) - z^TPz - \big(u^*(z)\big)^T Ru^*(z) \\ &\quad + \sigma^2\big(w^*(z)\big)^T w^*(z) - 2\big(u^*(z)\big)^T Ru^*\big(z_j\big) \\ &\quad + 2\big(u^*(z)\big)^T Ru^*(z)\big) \\ &= aV^*(z) - z^TPz + \sigma^2\big(w^*(z)\big)^T w^*(z) \\ &\quad + \big(u^*\big(z_j\big) - u^*(z)\big)^T R\big(u^*\big(z_j\big) - u^*(z)\big) \\ &\quad - \big(u^*\big(z_j\big)\big)^T Ru^*\big(z_j\big) \\ &\le a\eta_m - \lambda_{\min}(P)\|z\|^2 + \sigma^2\big(w^*(z)\big)^T w^*(z) \\ &\quad + \lambda_{\max}(R)||u^*\big(z_j\big) - u^*(z)||^2 \end{aligned} \tag{22}$$

in which

$$\begin{aligned} \sigma^2\big(w^*(z)\big)^T w^*(z) &\le \sigma^2\big\|w^*(z)\big\|^2 \\ &\le \sigma^2 * \frac{1}{4\sigma^4}\Big\|\big(V^*(z)\big)^T h(x)\Big\|^2 \\ &\le \frac{1}{4\sigma^2}h^2_m\eta^2_{em} \end{aligned} \tag{23}$$

based on (21) and Assumptions 2 and 3, furthermore, according to Assumption 1, it can be gotten that

$$||u^*\big(z_j\big) - u^*(z)||^2 \le K^2_{u^*}\big\|d_j\big\|^2 \tag{24}$$

and thus we have

$$\begin{aligned} \dot{L} &\le a\eta_m - \lambda_{\min}(P)\|z\|^2 + \frac{1}{4\sigma^2}h^2_m\eta^2_{em} \\ &\quad + \lambda_{\max}(R)K^2_{u^*}\big\|d_j\big\|^2. \end{aligned} \tag{25}$$

It is worth noting that when $t \in [t_j, t_{j+1})$, the event-triggering condition (12) will not be satisfied, which means that

$$d_j^T \Omega d_j \leq z^T \Omega z \tag{26}$$

and then, we get

$$\lambda_{\min}(\Omega)\left\|d_j\right\|^2 \leq \lambda_{\max}(\Omega)\|z\|^2. \tag{27}$$

Based on (22)–(27), it can be achieved that

$$\begin{aligned}\dot{L} &\leq a\eta_m - \lambda_{\min}(P)\|z\|^2 + \frac{1}{4\sigma^2}h_m^2\eta_{em}^2 \\ &\quad + \lambda_{\max}(R)K_{u^*}^2\frac{\lambda_{\max}(\Omega)}{\lambda_{\min}(\Omega)}\|z\|^2 \\ &= a\eta_m + \frac{1}{4\sigma^2}h_m^2\eta_{em}^2 \\ &\quad - \left(\lambda_{\min}(P) - \lambda_{\max}(R)K_{u^*}^2\frac{\lambda_{\max}(\Omega)}{\lambda_{\min}(\Omega)}\right)\|z\|^2. \end{aligned} \tag{28}$$

Apparently, if the condition (18) holds, then $\dot{L} \leq 0$ while

$$\|z\| \geq \sqrt{\frac{a\eta_m + \frac{1}{4\sigma^2}h_m^2\eta_{em}^2}{\lambda_{\min}(P) - \lambda_{\max}(R)K_{u^*}^2\frac{\lambda_{\max}(\Omega)}{\lambda_{\min}(\Omega)}}}. \tag{29}$$

Therefore, it can be summarized that under the condition (18), there exists a time instant t_f such that the tracking error z converges to a certain bound when $t > t_0 + t_f$. Then, the considered tracking error system (3) is stable in the sense of UUB according to Definition 1, and Theorem 1 is proved. ■

III. Solution to the Optimal Control Problem With RL-Based Method

Given that it is hard to obtain a closed-form solution of the event-triggered HJI equation presented by (16) [31], [32], [33], then a RL-based iterative algorithm is proposed to derive an approximate solution, which is followed by the effectiveness analysis of the algorithm. Considering that the control policy will not be updated between each of the two adjacent event-triggering instants under the event-triggered scheme (12), we then use $u^*(z_j)$ without emphasizing $t \in [t_j, t_{j+1})$ if there is no special needs in the remainder of this article.

A. RL-Based Iterative Algorithm

In this section, by recurring to the neural networks approximation method, a RL-based iteration algorithm is proposed to solve the formulated event-triggered HJI equation (16). Specifically, we first describe $V^*(z)$ using a critic network as follows:

$$V^*(z) = W_c^T\phi_c(z) + \varepsilon_c(z) \tag{30}$$

where $W_c \in \mathbb{R}^{n_c}$ denotes the optimal weight vector which needs to be designed, and $n_c \in \mathbb{R}$ is the number of hidden neurons; $\phi_c(z) = \text{col}\{\phi_{c_i}\}$ $(i = 1, 2, \ldots, n_c)$ is an undetermined activation function and ϕ_{c_i} are independent with each other; $\varepsilon_c(z)$ is the approximation error.

Then, the partial derivative of $V^*(z)$ with respect to the tracking error z is

$$V_z^* = \nabla\phi_c^T(z)W_c + \nabla\varepsilon_c(z). \tag{31}$$

Based on (31), the event-triggered-based optimal control input and worst disturbance in (15) can be rewritten as

$$\begin{cases} u^*(z) = u^*(z_j) = -\frac{1}{2}R^{-1}g^T(x_j)V_{z_j}^* \\ = -\frac{1}{2}R^{-1}g^T(x_j)\left(\nabla\phi_c^T(z_j)W_c + \nabla\varepsilon_c(z_j)\right) \\ w^*(z) = \frac{1}{2\sigma^2}h^T(x)V_z^* \\ \quad = \frac{1}{2\sigma^2}h^T(x)\left(\nabla\phi_c^T(z)W_c + \nabla\varepsilon_c(z)\right). \end{cases} \tag{32}$$

Actually, it is almost impossible to obtain the exact value of W_c, and thus $u^*(z)$ and $w^*(z)$ cannot be computed directly according to (32). In view of this, an approximate value of $V^*(z)$, denoted by $\hat{V}(z)$, is introduced, and formulated as

$$\hat{V}(z) = \hat{W}_c^T\phi_c(z) \tag{33}$$

where $\hat{W}_c \in \mathbb{R}^{n_c}$ is the approximation of W_c. Then, V_z^* can be approximated as

$$\hat{V}_z = \nabla\phi_c^T(z)\hat{W}_c. \tag{34}$$

Based on $\hat{V}_z$, the approximate value of the optimal control input and worst disturbance are denoted by $\hat{u}(z)$ and $\hat{w}(z)$, respectively, and represented as

$$\begin{cases} \hat{u}(z) = \hat{u}(z_j) = -\frac{1}{2}R^{-1}g^T(x_j)\hat{V}_{z_j} \\ = -\frac{1}{2}R^{-1}g^T(x_j)\nabla\phi_c^T(z_j)\hat{W}_c \\ \hat{w}(z) = \frac{1}{2\sigma^2}h^T(x)\hat{V}_z \\ = \frac{1}{2\sigma^2}h^T(x)\nabla\phi_c^T(z)\hat{W}_c. \end{cases} \tag{35}$$

The above method will result in an approximate value of the Hamiltonian function $H(z, V_z^*, u^*(z), w^*(z)))$, that is

$$\begin{aligned} &\hat{H}\left(z, \hat{V}_z, \hat{u}(z_j), \hat{w}(z)\right) \\ &= \hat{V}_z^T\left(F(x,r) + g(x)\hat{u}(z_j) + h(x(t))\hat{w}(z)\right) \\ &\quad + \hat{\Xi}\left(z, \hat{u}(z_j), \hat{w}(z)\right) - a\hat{V}(z) \\ &= W_c^T\left(\nabla\phi_c(z)\left(F(x,r) + g(x)\hat{u}(z_j)\right.\right. \\ &\qquad \left.\left. + h(x)\hat{w}(z)\right) - a\phi_c(z)\right) + \hat{\Xi}\left(z, \hat{u}(z_j), \hat{w}(z)\right) \end{aligned} \tag{36}$$

where $\hat{\Xi}(\cdot)$ is the approximate value of $\Xi^*(\cdot)$. We then use κ_c to denote the error between $\hat{H}(\cdot)$ and $H(\cdot)$, and thus have

$$\begin{aligned} \kappa_c &= \hat{H}\left(z, \hat{V}_z, \hat{u}(z_j), \hat{w}(z)\right) - H\left(z, V_z^*, u^*(z), w^*(z)\right) \\ &= \hat{W}_c^T\zeta + \hat{\Xi}\left(z, \hat{u}(z_j), \hat{w}(z)\right) \end{aligned} \tag{37}$$

in which $\zeta = \nabla\phi_c(z)(F(x,r) + g(x)\hat{u}(z_j) + h(x)\hat{w}(z)) - a\phi_c(z)$.

Then, we dedicate to guarantee that W_c can be effectively approximated by $\hat{W}_c$, i.e., let $\kappa_c \to 0$, and thus achieve the accurate approximation of $u^*(z)$ and $w^*(z)$. For this, we first design an objective function as follows:

$$E = \frac{1}{2\left(1 + \zeta^T\zeta\right)^2}\kappa_c^T\kappa_c. \tag{38}$$

Based on the gradient decent method, a tuning law of $\hat{W}_c$ is then constructed as

$$\dot{\hat{W}}_c = -l_c\frac{\partial E}{\partial \hat{W}_c} = -l_c\frac{\zeta}{\left(1 + \zeta^T\zeta\right)^2}\kappa_c \tag{39}$$

Algorithm 1: Proposed RL-Based Iterative Algorithm

```
Initializing û(0) and ŵ(0) according to Eq. (35) with the given Ŵ_c(0) and z(0);
for t = h; t ≤ T; t = t + h do
    Computing Ŵ_c(t) based on Eq. (41);
    Measuring the tracking error z(t) with û(t − h) and ŵ(t − h) according to Eq. (3);
    if The event-triggered condition Eq. (12) holds with z(t) then
        Calculating û(t) based on Eq. (35) with Ŵ_c(t);
    else
        û(t) = û(t − h);
    end
    Computing ŵ(t) according to Eq. (35) with Ŵ_c(t);
end
```

where l_c indicates the learning rate and will affect the convergence of the algorithm. Let $\tilde{W}_c = W_c - \hat{W}_c$ be the weight estimated error, and thus we can get

$$\dot{\tilde{W}}_c = -l_c \frac{\zeta\zeta^T}{\left(1+\zeta^T\zeta\right)^2}\tilde{W}_c + l_c \frac{\zeta}{\left(1+\zeta^T\zeta\right)^2}\epsilon_c \tag{40}$$

where ϵ_c is the residual error and can be represented as

$$\epsilon_c = -(\nabla\varepsilon_c(z))^T\big(F(x,r) + g(x)\hat{u}(z_j) + h(x)\hat{w}(z)\big) + a\varepsilon_c(z).$$

According to the tuning law, the value of $\hat{W}_c$ at time t, i.e., $\hat{W}_c(t)$, can be calculated as

$$\hat{W}_c(t) = \hat{W}_c(t-h) + h\left(-l_c \frac{\zeta(t-h)}{\left(1+\zeta^T(t-h)\zeta(t-h)\right)^2}\kappa_c(t-h)\right). \tag{41}$$

Then, $\hat{u}(z_j)$, and $\hat{w}(z)$ can be updated accordingly. To sum up, given the initial values of $\hat{W}_c$ and z, the proposed algorithm can update $\hat{u}(z_j)$ and $\hat{w}(z)$ over time. Without the loss of generality, the pseudo-code of the algorithm with a time window T is given in Algorithm 1.

B. Effectiveness Analysis for Proposed RL-Based Iterative Algorithm

In order to verify the effectiveness of the proposed algorithm, we will prove that the tracking error z and approximate error $\tilde{W}_c$ can be stabilized in the sense of UUB under Algorithm 1. The following assumption is first given to facilitate the further analysis.

Assumption 4: Supposing $||\nabla\phi_c|| \le \nabla\phi_{cm}$, $||\nabla\varepsilon_c|| \le \nabla\varepsilon_{cm}$, $||\varepsilon_c|| \le \varepsilon_{cm}$, and $||\epsilon_c|| \le \epsilon_{cm}$, where $\nabla\phi_{cm}$, $\nabla\varepsilon_{cm}$, ε_{cm}, and ϵ_{cm} are positive constants.

Theorem 2: Given the approximate optimal control input $\hat{u}(z_j)$ and worst disturbance $\hat{w}(z)$ presented in (35), the tuning law for $\hat{W}_c$ as described by (39), and Assumptions 1–4, the tracking error z in the system (3) and the estimated weight error $\tilde{W}_c$ are stable in the sense of UUB if

$$\begin{cases} \Pi_1 = \lambda_{\min}(P) - 2\lambda_{\max}(R)K_{u^*}^2 \frac{\lambda_{\max}(\Omega)}{\lambda_{\min}(\Omega)} > 0 \\ \Pi_2 = \frac{l_c}{2}\lambda_{\min}(\Gamma) - \frac{\lambda_{\max}(R)}{2}\lambda_{\min}^{-2}(R)g_m^2\nabla\phi_m^2 \\ \quad -\frac{1}{2\sigma^2}h_m^2\nabla\phi_{cm}^2 > 0 \end{cases} \tag{42}$$

where $\Gamma = ([\zeta\zeta^T]/[(1+\zeta^T\zeta)^2])$.

Proof: Constructing the following Lyapunov function:

$$\begin{aligned} L &= L_1 + L_2 + L_3 \\ &= V^*(z) + V^*(z_j) + \frac{1}{2}\tilde{W}_c^T\tilde{W}_c. \end{aligned} \tag{43}$$

With the adopted event-triggered scheme, we then consider the following two scenarios.

SI ($\boldsymbol{t \in [t_j, t_{j+1})}$): In such a scenario, it is obviously that $\dot{L}_2 = 0$, moreover, it can be calculated that

$$\begin{aligned} \dot{L}_1 &= \left(V_z^*\right)^T\dot{z} \\ &= \left(V_z^*\right)^T\big(F(x,r) + g(x)\hat{u}(z_j) + h(x)\hat{w}(z)\big) \\ &= \left(V_z^*\right)^T\big(F(x,r) + g(x)u^*(z) + h(x)w^*(z)\big) \\ &\quad + \left(V_z^*\right)^T g(x)\big(\hat{u}(z_j) - u^*(z)\big) \\ &\quad + \left(V_z^*\right)^T h(x)\big(\hat{w}(z) - w^*(z)\big). \end{aligned} \tag{44}$$

Considering Assumption 2, substituting (20) and (21) into (44), then it has

$$\begin{aligned} \dot{L}_1 &= aV^*(z) - z^TPz - \big(u^*(z)\big)^TRu^*(z) \\ &\quad + \sigma^2\big(w^*(z)\big)^Tw^*(z) - 2\big(u^*(z)\big)^TR\big(\hat{u}(z_j) - u^*(z)\big) \\ &\quad + 2\sigma^2\big(w^*(z)\big)^T\big(\hat{w}(z) - w^*(z)\big) \\ &= aV^*(z) - z^TPz + \big(u^*(z)\big)^TRu^*(z) \\ &\quad - 2\big(u^*(z)\big)^TR\hat{u}(z_j) + \sigma^2\big(w^*(z)\big)^Tw^*(z) \\ &\quad + 2\sigma^2\big(w^*(z)\big)^T\big(\hat{w}(z) - w^*(z)\big) \\ &= aV^*(z) - z^TPz - \big(\hat{u}(z_j)\big)^TR\hat{u}(z_j) \\ &\quad + \big(u^*(z) - \hat{u}(z_j)\big)^TR\big(u^*(z) - \hat{u}(z_j)\big) \\ &\quad + \sigma^2\big(w^*(z)\big)^Tw^*(z) + 2\sigma^2\big(w^*(z)\big)^T\big(\hat{w}(z) - w^*(z)\big) \\ &\le a\eta_m - \lambda_{\min}(P)\|z\|^2 + \lambda_{\max}(R)\big\|u^*(z) - \hat{u}(z_j)\big\|^2 \\ &\quad + 2\sigma^2\big\|w^*(z)\big\|^2 + \sigma^2\big\|\hat{w}(z) - w^*(z)\big\|^2. \end{aligned} \tag{45}$$

For the term $||u^*(z) - \hat{u}(z_j)||^2$ in (45), with Young's inequality, i.e., $||a+b||^2 \le 2||a||^2 + 2||b||^2$, and Assumption 1, we can get

$$\begin{aligned} &\big\|u^*(z) - \hat{u}(z_j)\big\|^2 \\ &\quad = \big\|\big(u^*(z) - u^*(z_j)\big) + \big(u^*(z_j) - \hat{u}(z_j)\big)\big\|^2 \\ &\quad \le 2\big\|u^*(z) - u^*(z_j)\big\|^2 + 2\big\|u^*(z_j) - \hat{u}(z_j)\big\|^2 \\ &\quad \le 2K_{u^*}^2\big\|d_j\big\|^2 + 2\big\|u^*(z_j) - \hat{u}(z_j)\big\|^2. \end{aligned} \tag{46}$$

According to Assumptions 3 and 4 and (32) and (35), it can be obtained that

$$\begin{aligned} &\big\|u^*(z_j) - \hat{u}(z_j)\big\|^2 \\ &= \Bigg\|-\frac{1}{2}R^{-1}g^T(x_j)\big(\nabla\phi_c^T(z_j)W_c + \nabla\varepsilon_c(z_j)\big) \\ &\qquad - \left(-\frac{1}{2}R^{-1}g^T(x_j)\nabla\phi_c^T(z_j)\hat{W}_c\right)\Bigg\|^2 \\ &= \left\|-\frac{1}{2}R^{-1}g^T(x_j)\nabla\phi_c^T(z_j)\tilde{W}_c - \frac{1}{2}R^{-1}g^T(x_j)\nabla\varepsilon_c(z_j)\right\|^2 \\ &\le \frac{1}{2}\lambda_{\min}^{-2}(R)g_m^2\nabla\phi_{cm}^2\big\|\tilde{W}_c\big\|^2 + \frac{1}{2}\lambda_{\min}^{-2}(R)g_m^2\nabla\varepsilon_{cm}^2. \end{aligned} \tag{47}$$

On the basis of (46), (47), it yields that

$$\left\|u^*(z)-\hat{u}(z_j)\right\|^2 \le 2K_{u^*}^2\left\|d_j\right\|^2+\lambda_{\min}^{-2}(R)g_m^2\nabla\phi_{cm}^2\left\|\tilde{W}_c\right\|^2 + \lambda_{\min}^{-2}(R)g_m^2\nabla\varepsilon_{cm}^2. \tag{48}$$

For the term $\sigma^2||\hat{w}(z)-w^*(z)||^2$ in (45), by means of (32) and (35), Assumptions 3 and 4, then it can be found that

$$\begin{aligned}\sigma^2&\left\|\hat{w}(z)-w^*(z)\right\|^2\\ &=\sigma^2\left\|\frac{1}{2\sigma^2}h^T(x)\nabla\phi_c^T(z)\hat{W}_c - \frac{1}{2\sigma^2}h^T(x)\left(\nabla\phi_c^T(z)W_c+\nabla\varepsilon_c(z)\right)\right\|^2\\ &\le\frac{1}{4\sigma^2}\|h(x)\|^2\left\|\nabla\phi_c^T(z)\tilde{W}_c+\nabla\varepsilon_c(z)\right\|^2\\ &\le\frac{1}{2\sigma^2}h_m^2\nabla\phi_{cm}^2\left\|\tilde{W}_c\right\|^2+\frac{1}{2\sigma^2}h_m^2\nabla\varepsilon_{cm}^2.\end{aligned} \tag{49}$$

In addition, according to Assumptions 2 and 3 and (10), the term $2\sigma^2||w^*(z)||^2$ can be handled as

$$2\sigma^2||w^*(z)||^2=2\sigma^2||\frac{1}{2\sigma^2}h^T(x)V_z^*||^2\le\frac{1}{2\sigma^2}h_m^2\eta_{em}^2. \tag{50}$$

So, based on (48)–(50), and (45) can be represented as

$$\begin{aligned}\dot{L}_1\le{}& a\eta_m-\lambda_{\min}(P)\|z\|^2+2\lambda_{\max}(R)K_{u^*}^2\left\|d_j\right\|^2\\ &+\lambda_{\max}(R)\lambda_{\min}^{-2}(R)g_m^2\nabla\phi_m^2\left\|\tilde{W}_c\right\|^2\\ &+\lambda_{\max}(R)\lambda_{\min}^{-2}(R)g_m^2\nabla\varepsilon_{cm}^2+\frac{1}{2\sigma^2}h_m^2\eta_{em}^2\\ &+\frac{1}{2\sigma^2}h_m^2\nabla\phi_{cm}^2\left\|\tilde{W}_c\right\|^2+\frac{1}{2\sigma^2}h_m^2\nabla\varepsilon_{cm}^2.\end{aligned} \tag{51}$$

We then focus on the derivative of L_3 with respect to t

$$\begin{aligned}\dot{L}_3&=\tilde{W}_c^T\left(-lc\frac{\zeta\zeta^T}{\left(1+\zeta^T\zeta\right)^2}\tilde{W}_c+lc\frac{\zeta}{\left(1+\zeta^T\zeta\right)^2}\epsilon_c\right)\\ &=-lc\tilde{W}_c^T\frac{\zeta\zeta^T}{\left(1+\zeta^T\zeta\right)^2}\tilde{W}_c+l_c\tilde{W}_c^T\frac{\zeta}{\left(1+\zeta^T\zeta\right)^2}\epsilon_c.\end{aligned} \tag{52}$$

It is obvious that $(1+\zeta^T\zeta)\ge 1$, then

$$\begin{aligned}l_c\tilde{W}_c^T\frac{\zeta}{\left(1+\zeta^T\zeta\right)^2}\epsilon_c&\le l_c\tilde{W}_c^T\frac{\zeta}{\left(1+\zeta^T\zeta\right)}\epsilon_c\\ &\le\frac{l_c}{2}\tilde{W}_c^T\frac{\zeta\zeta^T}{\left(1+\zeta^T\zeta\right)^2}\tilde{W}_c+\frac{l_c}{2}\epsilon_c^T\epsilon_c.\end{aligned} \tag{53}$$

Therefore, based on Assumption 4, it can be gotten that

$$\begin{aligned}\dot{L}_3&\le-\frac{l_c}{2}\tilde{W}_c^T\frac{\zeta\zeta^T}{\left(1+\zeta^T\zeta\right)^2}\tilde{W}_c+\frac{l_c}{2}\epsilon_c^T\epsilon_c\\ &\le-\frac{l_c}{2}\lambda_{\min}(\Gamma)\left\|\tilde{W}_c\right\|^2+\frac{l_c}{2}\epsilon_{cm}^2\end{aligned} \tag{54}$$

where $\Gamma=([\zeta\zeta^T]/[(1+\zeta^T\zeta)^2])$.

Combining (51) and (54), then by recurring to (27), we further get

$$\begin{aligned}\dot{L}\le{}& a\eta_m-\lambda_{\min}(P)\|z\|^2\\ &+2\lambda_{\max}(R)K_{u^*}^2\frac{\lambda_{\max}(\Omega)}{\lambda_{\min}(\Omega)}\|z\|^2\\ &+\lambda_{\max}(R)\lambda_{\min}^{-2}(R)g_m^2\nabla\phi_{cm}^2\left\|\tilde{W}_c\right\|^2\\ &+\lambda_{\max}(R)\lambda_{\min}^{-2}(R)g_m^2\nabla\varepsilon_{cm}^2+\frac{1}{2\sigma^2}h_m^2\eta_{em}^2\\ &+\frac{1}{2\sigma^2}h_m^2\nabla\phi_{cm}^2\left\|\tilde{W}_c\right\|^2+\frac{1}{2\sigma^2}h_m^2\nabla\varepsilon_{cm}^2\\ &-\frac{l_c}{2}\lambda_{\min}(\Gamma)\left\|\tilde{W}_c\right\|^2+\frac{l_c}{2}\epsilon_{cm}^2.\\ \triangleq{}& -\Pi_1\left\|\tilde{W}_c\right\|^2-\Pi_2\|z\|^2+\Pi_3\end{aligned} \tag{55}$$

where

$$\begin{aligned}\Pi_1&=\frac{l_c}{2}\lambda_{\min}(\Gamma)-\lambda_{\max}(R)\lambda_{\min}^{-2}(R)g_m^2\nabla\phi_{cm}^2-\frac{1}{2\sigma^2}h_m^2\nabla\phi_{cm}^2\\ \Pi_2&=\lambda_{\min}(P)-2\lambda_{\max}(R)K_{u^*}^2\frac{\lambda_{\max}(\Omega)}{\lambda_{\min}(\Omega)}\\ \Pi_3&=a\eta_m+\lambda_{\max}(R)\lambda_{\min}^{-2}(R)g_m^2\nabla\varepsilon_{cm}^2+\frac{1}{2\sigma^2}h_m^2\eta_{em}^2\\ &\quad+\frac{1}{2\sigma^2}h_m^2\nabla\varepsilon_{cm}^2+\frac{l_c}{2}\epsilon_{cm}^2.\end{aligned}$$

Thus, it can be concluded that $\dot{L}<0$ if (42) holds and one of the following inequalities is satisfied:

$$\begin{cases}\|z\|\ge\sqrt{\Pi_3/\Pi_2}\\ \left\|\tilde{W}_c\right\|\ge\sqrt{\Pi_3/\Pi_1}.\end{cases} \tag{56}$$

SII ($\boldsymbol{t=t_{j+1}}$): Under such a scenario, for the constructed Lyapunov function L, we have

$$\begin{aligned}\triangle L={}&V^*\left(z(t_{j+1})\right)-V^*\left(z(t_j)\right)+V^*\left(z(t_{j+1})\right)-V^*\left(z\left(t_{j+1}^-\right)\right)\\ &+\frac{1}{2}\tilde{W}_c^T\left(t_{j+1}\right)\tilde{W}_c\left(t_{j+1}\right)-\frac{1}{2}\tilde{W}_c^T\left(t_{j+1}^-\right)\tilde{W}_c\left(t_{j+1}^-\right)\end{aligned} \tag{57}$$

in which the $V^*(z(t_{j+1}^-))=\lim_{\rho\to0^+}V^*(z(t_{j+1}-\rho))$ with $\rho\in(0,t_{j+1}-t_j)$.

As discussed in *SI*, it is easy to obtain that $V^*(z(t_{j+1}))\le V^*(z(t_j))$ because $V^*(\cdot)$ is a continuous and monotonically decreasing function. Moreover, if either $||z||\ge\sqrt{\Pi_3/\Pi_2}$ or $||\tilde{W}_c||\ge\sqrt{\Pi_3/\Pi_1}$ for $t\in[t_j,t_{j+1})$, it can be derived that $\dot{L}=\dot{L}_1+\dot{L}_3<0$, which indicates that it is monotonically decreasing, and thus it has

$$L_1\left(t_{j+1}\right)+L_3\left(t_{j+1}\right)\le L_1\left(t_{j+1}-\rho\right)+L_3\left(t_{j+1}-\rho\right). \tag{58}$$

When $\rho\to0^+$

$$\begin{aligned}L_1\left(t_{j+1}\right)+L_3\left(t_{j+1}\right)&\le\lim_{\rho\to0^+}\left(L_1\left(t_{j+1}-\rho\right)+L_3\left(t_{j+1}-\rho\right)\right)\\ &=L_1\left(t_{j+1}^-\right)+L_3\left(t_{j+1}^-\right).\end{aligned} \tag{59}$$

So, we can achieve that $\Delta L<0$ at the event-triggering instant t_{j+1}.

Based on the aforementioned analysis under two scenarios, it can be concluded that the tracking error z and approximated weight error $\tilde{W}_c$ are stable in the sense of UUB if (42) holds based on Definition 1, thus, we can get the theorem. ■

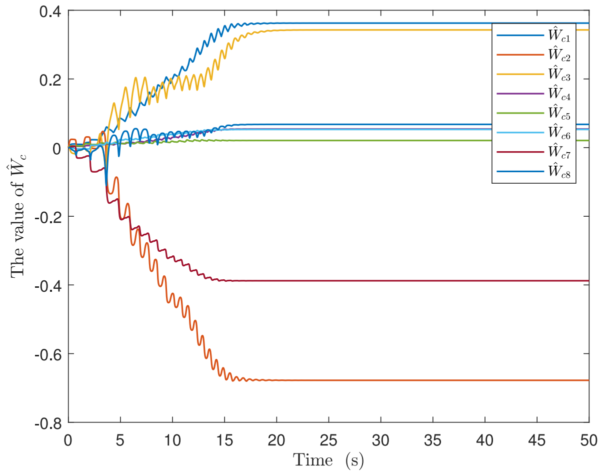


Fig. 1. Value of $\hat{W}_c$.

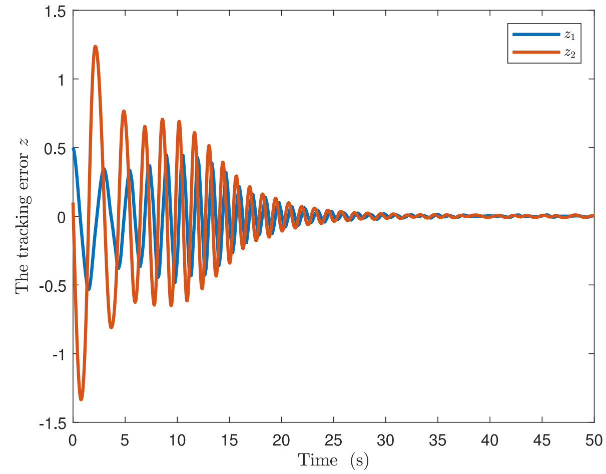


Fig. 2. Tracking error z.

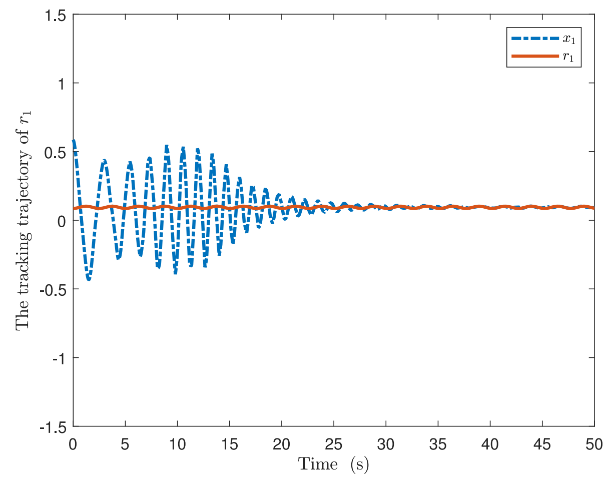


Fig. 3. Tracking trajectory of r_1.

IV. Numerical Example

In this section, a numerical example is designed to verify the effectiveness of the established method. The specific simulation settings are first described as follows. For the envisioned tracking error system (3), the corresponding nonlinear functions are given as

$$f(x) = \begin{bmatrix} \sin(x_2) - 0.065e^{3x_1} \\ -6\sin(x_1) + 0.8\cos(x_2) - 0.2e^{1.5x_1} \end{bmatrix}$$
$$g(x) = \begin{bmatrix} 2.5 \\ 0.2 \end{bmatrix}, h(x) = \begin{bmatrix} 10.2 \\ 8.5 \end{bmatrix}, \phi(r(t)) = \begin{bmatrix} 0.02\sin(2.5h) \\ 0.02\cos(2.5h) \end{bmatrix} \tag{60}$$

and the initial system states are designed to be $z(0) = [0.5\ 0.1]^T$ and $r(0) = [0.086\ 0.086]^T$.

For the defined value function $V(z)$ (5) and the adopted event-triggered mechanism (12), we set $P = 600I_2$, $R = 0.8$, $\Omega = 0.08I_2$, and $h = 0.04s$. For the proposed Algorithm 1, the parameters of the critic network are designed to be $l_c = 0.01$, $n_c = 8$, $\sigma = 8$, and $a = 0.6$. Besides, referring to [26], the basic function $\phi_c(z)$ is given as

$$\phi_c(z) = \begin{bmatrix} z_1^2 & z_1 z_2 & z_2^2 & z_1^4 & z_1^3 z_2 & z_1^2 z_2^2 & z_1 z_2^3 & z_2^4 \end{bmatrix}^T \tag{61}$$

moreover, the approximated weight vector is formulated as $\hat{W}_c = [\hat{W}_{c1}\ \hat{W}_{c2}\ \hat{W}_{c3}\ \hat{W}_{c4}\ \hat{W}_{c5}\ \hat{W}_{c6}\ \hat{W}_{c7}\ \hat{W}_{c8}]^T$, and the initial value of which is set to be $\hat{W}_c(0) = [0\ 0\ 0\ 0\ 0\ 0\ 0\ 0]^T$.

Based on the above settings, the simulation results are then presented by Figs. 1–7. The convergence process of $\hat{W}_c$ is depicted by Fig. 1, it can be seen that $\hat{W}_c$ finally converges to the value $\hat{W}_c = [0.3626\ -0.6776\ 0.3426\ 0.0542\ 0.0208\ 0.0528\ -0.3878\ 0.0677]^T$. Fig. 2 shows that the tracking error of the system can reach a stable state when the weight $\hat{W}_c$ converges to the certain value. The tracking trajectories of system states are given in Figs. 3 and 4, as shown, x_1 and x_2 can well track the objective trajectories r_1 and r_2 after about 40-s under the designed tracking control strategy. Moreover, the approximate optimal control input $\hat{u}(z_j)$, worst disturbance $\hat{w}(z)$, and the estimated Hamilton function $\hat{H}(\cdot)$ are demonstrated by Figs. 5–7, respectively. It can be seen from Figs. 5 and 6 that the approximated optimal control input always remains the same between two adjacent event-triggered instants while the disturbance updates along with the simulation time. Moreover, as shown in Fig. 7, the value of the estimated Hamilton function $\hat{H}(z, \hat{V}_z, \hat{u}(z_j), \hat{w}(z))$ fluctuates with the time and finally converges to zero, which verifies the effectiveness of the approximated optimal control input according to the optimality theory.

As illustrated in Remark 3, one of the differences of the work with the relative researches is that the tracking control strategy is designed based on a BFETM. To validate the effectiveness of such a design, we then compare some results obtained under the adopted BFETM and a stability-focused event-triggered mechanism (SFETM). The SFETM is constructed based on the approach proposed in [26], and thus can assist the stabilization of the considered tracking error system (3). Specifically, the event-triggering condition of the SFETM is derived as

$$\|d_j\|^2 \le \frac{(1-\eta^2)\lambda_{\min}(P)\|z\|^2}{2{K_{u^*}}^2\lambda_{\max}(R)} \tag{62}$$

where $\eta \in (0, 1)$ is a given constant and $K_{u^*} > 0$ is the parameter used in Assumption 1, which are then set to be

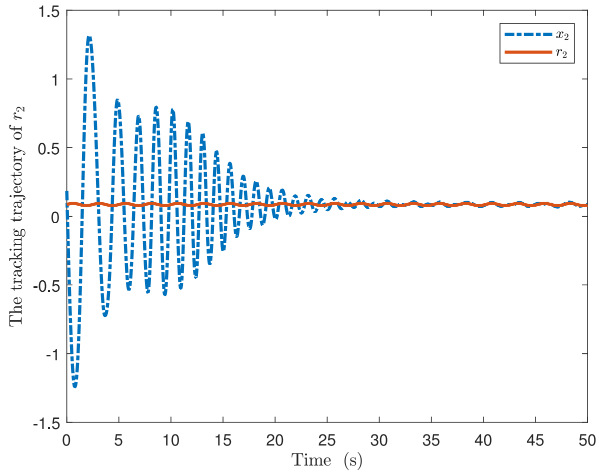


Fig. 4. Tracking trajectory of r_2.

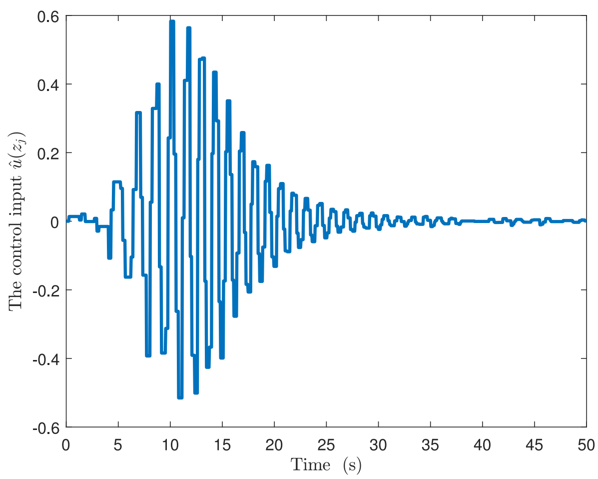


Fig. 5. Approximate optimal control input $\hat{u}(z_j)$.

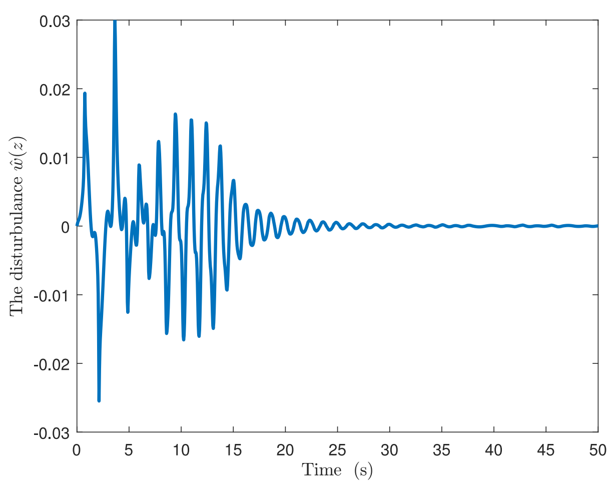


Fig. 6. Approximate worst disturbance $\hat{w}(z)$.

$\eta = 0.34$ and $K_{u^*} = 21$ in the simulation. With the SFETM, the ultimate bounds of the tracking error z and approximate weight error $\tilde{W}_c$ are

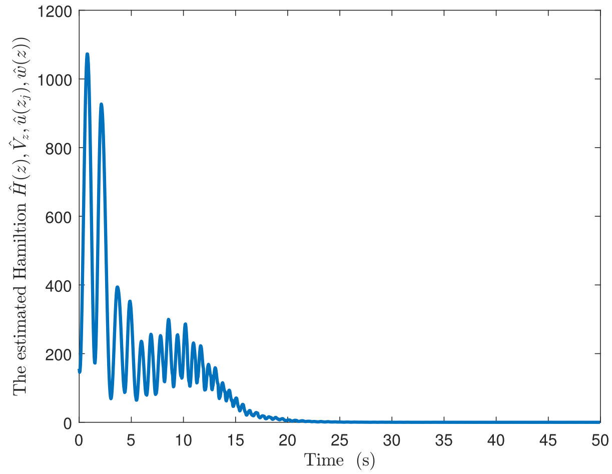


Fig. 7. Estimated Hamilton function $\hat{H}(z, \hat{V}_z, \hat{u}(z_j), \hat{w}(z))$.

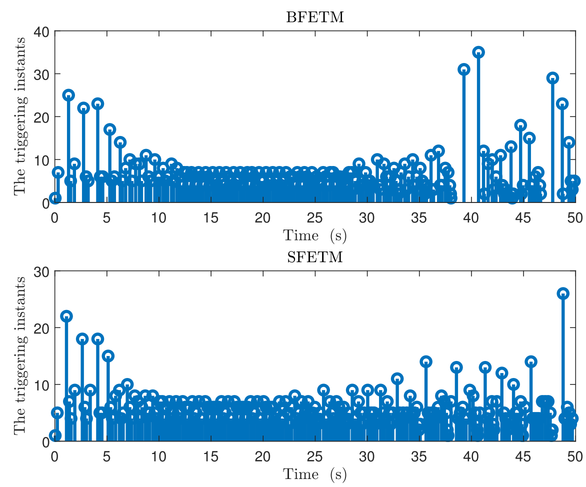


Fig. 8. Triggering instants with the two event-triggered mechanisms.

$$\begin{cases} \|z\| \geq \sqrt{\Pi_3/\eta^2\lambda_{\min}(P)} \\ \|\tilde{W}_c\| \geq \sqrt{\Pi_3/\Pi_1} \end{cases} \tag{63}$$

where Π_1 and Π_3 are the same as in (55).

In Fig. 8, the triggering instants under the BFETM and SFETM are presented. It can be found that less triggering instants are generated with the BFETM (actually, the numbers of the triggering instants under the BFETM and SFETM are, respectively, 187 and 219 within 50 s), and thus the bandwidth pressure can be effectively alleviated. Meanwhile, Fig. 9 shows the two-norm of the tracking error z obtained under the BFETM and SFETM, and we can see that the overall transient performance of the tracking error with the BFETM is not inferior to that with the SFETM, which further evidences the efficiency of using the BFETM.

V. Conclusion

In this article, for a bandwidth-limited nonlinear networked system with sampling-driven disturbance, we have developed a RL-based scheme to tackle the optimal tracking control problem. The tracking error system with a BFETM has been

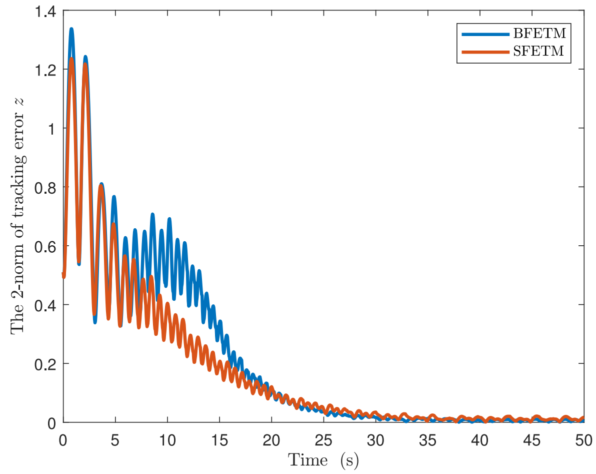


Fig. 9. Tracking errors obtained under the two event-triggered mechanisms.

established to model the considered problem. Then, by designing an appropriate value function and on account of the Bellman optimality theory, we have shown that the optimal tracking control strategy under the envisioned event-triggering scenario can be obtained by solving an event-based HJI equation. After proving the existence of the optimal control and worst disturbance policies that can stabilize the concerned system in the sense of UUB, a RL-based iterative algorithm with single critic network framework has been proposed to solve the constructed HJI equation. The effectiveness of the algorithm has been analytically evaluated and further confirmed by simulations. Based on the current research, the future work will focus on the optimal control problem for various complex industrial systems under limited communication resources and stochastic cyber attacks with the assistance of RL methods.

References

[1] J.-L. Liu, Y.-H. Dong, L.-J. Zha, E.-G. Tian, and X.-P. Xie, "Event-based security tracking control for networked control systems against stochastic cyber-attacks," *Inf. Sci.*, vol. 612, pp. 306–321, Oct. 2022.

[2] Y.-Y. Wang, B. Jiang, Z.-G. Wu, S.-R. Xie, and Y. Peng, "Adaptive sliding mode fault-tolerant fuzzy tracking control with application to unmanned marine vehicles," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 51, no. 11, pp. 6691–6700, Nov. 2021.

[3] Z.-Q. Yu et al., "Enhanced recurrent fuzzy neural fault-tolerant synchronization tracking control of multiple unmanned airships via fractional calculus and fixed-time prescribed performance function," *IEEE Trans. Fuzzy Syst.*, vol. 30, no. 10, pp. 4515–4529, Oct. 2022.

[4] J.-L. Liu, T.-T. Yin, J. Cao, D. Yue, and H. R. Karimi, "Security control for T–S fuzzy systems with adaptive event-triggered mechanism and multiple cyber-attacks," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 51, no. 10, pp. 6544–6554, Oct. 2021.

[5] Z.-H. Wu, E.-G. Tian, and H.-T. Chen, "Covert attack detection for LFC systems of electric vehicles: A dual time-varying coding method," *IEEE/ASME Trans. Mechatronics*, early access, Sep. 9, 2022, doi: 10.1109/TMECH.2022.3201875.

[6] Z.-Q. Zhang, X.-J. Xie, and S. S. Ge, "Adaptive tracking for uncertain MIMO nonlinear systems with time-varying parameters and bounded disturbance," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 51, no. 7, pp. 4479–4491, Jul. 2021.

[7] N. Wang, J.-C. Sun, and M.-J. Er, "Tracking-error-based universal adaptive fuzzy control for output tracking of nonlinear systems with completely unknown dynamics," *IEEE Trans. Fuzzy Syst.*, vol. 26, no. 2, pp. 869–883, Apr. 2018.

[8] T. Wang, M. Ma, J.-B. Qiu, and H.-J. Gao, "Event-triggered adaptive fuzzy tracking control for pure-feedback stochastic nonlinear systems with multiple constraints," *IEEE Trans. Fuzzy Syst.*, vol. 29, no. 6, pp. 1496–1506, Jun. 2021.

[9] D.-Y. Yao, H.-Y. Li, R.-Q. Lu, and Y. Shi, "Distributed sliding-mode tracking control of second-order nonlinear multiagent systems: An event-triggered approach," *IEEE Trans. Cybern.*, vol. 50, no. 9, pp. 3892–3902, Sep. 2020.

[10] Z.-X. Yu, H.-C. Yan, S.-G. Li, and Y. Dong, "Approximation-based adaptive tracking control for switched stochastic strict-feedback nonlinear time-delay systems with sector-bounded quantization input," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 48, no. 12, pp. 2145–2157, Dec. 2018.

[11] S. Li, L. Ding, H.-B. Gao, Y.-J. Liu, L. Huang, and Z.-Q. Deng, "ADP-based online tracking control of partially uncertain time-delayed nonlinear system and application to wheeled mobile robots," *IEEE Trans. Cybern.*, vol. 50, no. 7, pp. 3182–3194, Jul. 2020.

[12] Y.-H. Zhu, D.-B. Zhao, and H.-B. He, "Invariant adaptive dynamic programming for discrete-time optimal control," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 50, no. 11, pp. 3959–3971, Nov. 2020.

[13] Z.-H. Xu, T.-L. Shen, and D.-Z. Cheng, "Model-free reinforcement learning by embedding an auxiliary system for optimal control of nonlinear systems," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 33, no. 4, pp. 1520–1534, Apr. 2022.

[14] J.-N. Li, T.-Y. Chai, F. L. Lewis, Z.-T. Ding, and Y. Jiang, "Off-policy interleaved *Q*-learning: Optimal control for affine nonlinear discrete-time systems," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 30, no. 5, pp. 1308–1320, May 2019.

[15] C. Li, D.-R. Liu, and D. Wang, "Data-based optimal control for weakly coupled nonlinear systems using policy iteration," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 48, no. 4, pp. 511–521, Apr. 2018.

[16] B. Pang and Z.-P. Jiang, "Adaptive optimal control of linear periodic systems: An off-policy value iteration approach," *IEEE Trans. Autom. Control*, vol. 66, no. 2, pp. 888–894, Feb. 2021.

[17] H. Wang and M. Li, "Model-free reinforcement learning for fully cooperative consensus problem of nonlinear multiagent systems," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 33, no. 4, pp. 1482–1491, Apr. 2022.

[18] N. Wang, Y. Gao, and X.-F. Zhang, "Data-driven performance-prescribed reinforcement learning control of an unmanned surface vehicle," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 32, no. 12, pp. 5456–5467, Dec. 2021.

[19] F.-R. Qu, E.-G. Tian, and X. Zhao, "Chance-constrained H_∞ state estimation for recursive neural networks under deception attacks and energy constraints: The finite-horizon case," *IEEE Trans. Neural Netw. Learn. Syst.*, early access, Jan. 7, 2022, doi: 10.1109/TNNLS.2021.3137426.

[20] S.-H. Luo, F. L. Lewis, Y.-D. Song, and H. M. Ouakad, "Optimal synchronization of unidirectionally coupled FO chaotic electromechanical devices with the hierarchical neural network," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 33, no. 3, pp. 1192–1202, Mar. 2022.

[21] W.-L. Yang, D.-Z. Xu, L.-C. Jin, B. Jiang, and P. Shi, "Robust model predictive control for linear systems via self-triggered pseudo terminal ingredients," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 69, no. 3, pp. 1312–1322, Mar. 2022.

[22] J.-L. Liu, Z.-G. Wu, D. Yue, and J. H. Park, "Stabilization of networked control systems with hybrid-driven mechanism and probabilistic cyber attacks," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 51, no. 2, pp. 943–953, Feb. 2021.

[23] R. Nie, W.-L. He, W.-L. Du, Z.-Q. Lang, and S.-P. He, "Dynamic event-triggered SMC of multi-agent systems for consensus tracking," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 69, no. 3, pp. 1188–1192, Mar. 2022.

[24] Y. Li, F.-Y. Song, J.-L. Liu, X.-P. Xie, and E.-G. Tian, "Decentralized event-triggered synchronization control for complex networks with non-periodic DoS attacks," *Int. J. Robust Nonlinear Control*, vol. 32, no. 3, pp. 1633–1653, 2022.

[25] L.-J. Zha, R.-F. Liao, J.-L. Liu, X.-P. Xie, E.-G. Tian, and J.-D. Cao, "Dynamic event-triggered output feedback control for networked systems subject to multiple cyber attacks," *IEEE Trans. Cybern.*, vol. 52, no. 12, pp. 13800–13808, Dec. 2022.

[26] X. Yang and H.-B. He, "Event-driven H_∞ constrained control using adaptive critic learning," *IEEE Trans. Cybern.*, vol. 51, no. 10, pp. 4860–4872, Oct. 2021.

[27] W.-W. Bai, T.-S. Li, Y. Long, and C. L. P. Chen, "Event-triggered multi-gradient recursive reinforcement learning tracking control for multiagent systems," *IEEE Trans. Neural Netw. Learn. Syst.*, early access, Jul. 16, 2021, doi: 10.1109/TNNLS.2021.3094901.

[28] X. Yang and H.-B. He, "Decentralized event-triggered control for a class of nonlinear-interconnected systems using reinforcement learning," *IEEE Trans. Cybern.*, vol. 51, no. 2, pp. 635–648, Feb. 2021.

[29] H.-G. Zhang, Y.-L. Liang, H.-G. 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.

[30] X. Yang and H.-B. He, "Event-triggered robust stabilization of nonlinear input-constrained systems using single network adaptive critic designs," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 50, no. 9, pp. 3145–3157, Sep. 2020.

[31] S. Xue, B. Luo, and D.-R. Liu, "Event-triggered adaptive dynamic programming for zero-sum game of partially unknown continuous-time nonlinear systems," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 50, no. 9, pp. 3189–3199, Sep. 2020.

[32] X. Yang and Q.-L. Wei, "Adaptive critic designs for optimal event-driven control of a CSTR system," *IEEE Trans. Ind. Informat.*, vol. 17, no. 1, pp. 484–493, Jan. 2021.

[33] B. Zhao and D.-R. Liu, "Event-triggered decentralized tracking control of modular reconfigurable robots through adaptive dynamic programming," *IEEE Trans. Ind. Electron.*, vol. 67, no. 4, pp. 3054–3064, Apr. 2020.

[34] S.-X. Zhu, E.-G. Tian, D.-H. Xu, and J.-L. Liu, "An adaptive torus-event-based H_∞ controller design for networked T–S fuzzy systems under deception attacks," *Int. J. Robust Nonlinear Control*, vol. 32, no. 6, pp. 3425–3441, 2022.

[35] L. Feng and R. D. Brandt, "An optimal control approach to robust control of robot manipulators," *IEEE Trans. Robot. Autom.*, vol. 14, no. 1, pp. 69–77, Feb. 1998.

[36] M. Abu-Khalaf, J. Huang, and F. L. Lewis, *Nonlinear H_2/H_∞ Constrained Feedback Control: A Practical Design Approach Using Neural Networks*. London, U.K.: Springer, 2006.

[37] W. Rudin, *Principles of Mathematical Analysis*, vol. 3. New York, NY, USA: McGraw-Hill, 1976.

[38] A. van der Schaft, "L_2-gain analysis of nonlinear systems and nonlinear state-feedback H_∞ control," *IEEE Trans. Autom. Control*, vol. 37, no. 6, pp. 770–784, Jun. 1992.

Jinliang Liu received the Ph.D. degree in automatic control from Donghua University, Shanghai, China, in 2011.

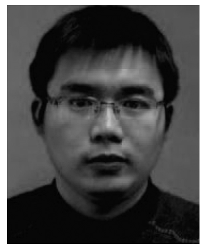
He was a Postdoctoral Research Associate with the School of Automation, Southeast University, Nanjing, China, from 2013 to 2016. He is currently a Professor with the College of Information Engineering, Nanjing University of Finance and Economics, Nanjing. His research interests include networked control systems, complex dynamical networks, and time-delay systems.

Nan Zhang received the B.S. degree in computer science from Nanjing University of Finance and Economics, Nanjing, China, in 2020, where he is currently pursuing the M.Sc. degree with the College of Information Engineering.

His research interests include networked control systems and learning-based optimal control methods.

Yan Li received the Ph.D. degree in computer science from the University of Science and Technology of China, Hefei, China, and the City University of Hong Kong, Hong Kong, in 2011.

She is currently an Associate Professor with the College of Information Engineering, Nanjing University of Finance and Economics, Nanjing, China. Her research interests include networked control systems, learning-based algorithms, and software-defined networking.



Xiangpeng Xie (Member, IEEE) received the B.S. and Ph.D. degrees in engineering from Northeastern University, Shenyang, China, in 2004 and 2010, respectively.

He is currently a Professor with the Institute of Advanced Technology, Nanjing University of Posts and Telecommunications, Nanjing, China. His research interests include fuzzy modeling and control synthesis, state estimations, optimization in process industries, and intelligent optimization algorithms.

Prof. Xie serves as an Associate Editor for the *International Journal of Fuzzy Systems* and *International Journal of Control, Automation, and Systems*.

Engang Tian received the B.S. degree in mathematics from Shandong Normal University, Jinan, China, in 2002, the M.Sc. degree in operations research and cybernetics from Nanjing Normal University, Nanjing, China, in 2005, and the Ph.D. degree in control theory and control engineering from Donghua University, Shanghai, China, in 2008.

He is currently a Professor with the School of Optical-Electrical and Computer Engineering, University of Shanghai for Science and Technology, Shanghai. His research interests include networked control systems, cyber attack, as well as nonlinear stochastic control and filtering.

Jinde Cao (Fellow, IEEE) received the B.S. degree in mathematics/applied mathematics from Anhui Normal University, Wuhu, China, in 1986, the M.S. degree in mathematics/applied mathematics from Yunnan University, Kunming, China, in 1989, and the Ph.D. degree in mathematics/applied mathematics from Sichuan University, Chengdu, China, in 1998.

He is currently an Endowed Chair Professor, the Dean of the School of Mathematics, and the Director of the Research Center for Complex Systems and Network Sciences with the School of Mathematics, Southeast University, Nanjing, China.