

Adaptive Event-Triggered Near-Optimal Tracking Control for Unknown Continuous-Time Nonlinear Systems

KUNFU WANG^{1,2}, QIJIA GU², BAIQIAO HUANG², QINGLAI WEI³, (Senior Member, IEEE),
AND TIANMIN ZHOU³

¹School of Mathematics and Statistics, Nanjing University of Information Science & Technology, Nanjing 210044, China

²Systems Engineering Research Institute, China State Shipbuilding Corporation Ltd., Beijing 100036, China

³State Key Laboratory for Management and Control of Complex Systems, Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China

Corresponding author: Tianmin Zhou (tianmin.zhou@ia.ac.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant 61973282 and Grant 62073321; in part by the National Key Research and Development Program of China under Grant 2021YFE0206100 and Grant 2018YFB1702300; and in part by the Science and Technology Development Fund, Macau SAR, under Grant 0015/2020/AMJ.

ABSTRACT This paper studies the event-triggered optimal tracking control (ETOTC) problem of continuous-time (CT) unknown nonlinear systems. In order to solve the ETOTC problem, an augmented system composed of the error system dynamics and the reference dynamics is used to introduce a new discounted performance index function (DPIF). A novel event-triggered (ET) adaptive dynamic programming (ADP) method is developed to solve the ET Hamilton-Jacobi-Bellman equation (HJBE). The presented method is implemented via an identifier-critic architecture, which consists of two neural networks (NNs): an identifier NN is applied to estimate the unknown system dynamics, and a critic NN is constructed to obtain the approximate solution of the ET HJBE. The augmented closed-loop system and the critic estimation error are proved to be ultimately uniformly bounded (UUB) by the Lyapunov direct method. Finally, two simulations illustrate the effectiveness of the developed method.

INDEX TERMS Adaptive dynamic programming, event-triggered mechanism, optimal tracking control, nonlinear system.

I. INTRODUCTION

Adaptive dynamic programming (ADP), a branch of the reinforcement learning, is widely concerned in solving optimal control problems in recent years [1]. On the basis of the neural network (NN) technology, the reinforcement learning and the dynamic programming (DP), ADP is an effective solution method in the Hamilton-Jacobi-Bellman equation (HJBE) of complicated nonlinear systems, and it can also overwhelm the “dimensionality curse” of the traditional DP [2]. Its essence is to use the online or offline data to estimate the system performance index function (PIF) and obtain the approximate optimal control law according to the Bellman’s optimality principle [3].

In the past decade, researches have proposed different ADP-based control methods to solve the problems of optimal control, trajectory tracking and robust stabilization in discrete-time (DT) or continuous-time (CT)

The associate editor coordinating the review of this manuscript and approving it for publication was Zheng H. Zhu ³.

systems [4]–[12]. As a human learning and decision-making process, ADP possesses several synonyms, such as approximate dynamic programming [13], adaptive critic designs [14], neuro-dynamic programming [15], relaxing dynamic programming [16], and reinforcement learning [17]. Based on different structures, ADP techniques can be classified into several categories, such as heuristic dynamic programming (HDP), dual HDP (DHP), globalized DHP (GDHP) and their action dependent form [18]–[22]. Based on different iteration methods, there are two major categories of ADP techniques, which are value iteration (VI) and policy iteration (PI) algorithms, respectively [23]–[29].

Most of the existing control methods are based on sampling with the period or triggering with the time. However, from the perspective of the resource allocation, the traditional methods have some disadvantages. For example, when the system works in a stable state, periodic sampling will cause the unnecessary waste of resources. For networked control systems, the periodic sampling can increase the computational cost as well as the communication burden [30]–[33].

In order to reduce the burden of computing and communication, researchers have proposed event-triggered control (ETC). The system signal sampling and the controller operation are driven by a specific event instead of the time instant. There are many types of the event which can be a variable of the system exceeding the limit value or a packet arriving at a node in the network control. Overall, the sampling period of the ETC system is time-varying, and the signals are sampled and transmitted according to the needs of the system [34]. Therefore, ETC is an effective control method for reducing computational costs and has a wide range of applications in areas such as the network system control and the complex system tracking [35]–[37]. ETC mainly discusses the introduction of the event-triggered (ET) mechanism, the stability analysis and the design method of ET condition, etc. Eqtami *et al.* [38] proposed an ETC method for DT systems, which uses the input-to-state stable technique to extend the ET conditions of CT systems to DT systems with guaranteed system stability. Wang and Lemmon [39] developed an ET mechanism for CT nonlinear systems and gave the minimum time interval between two adjacent triggered events. Li and Xu [40] designed a trigger condition that can guarantee the asymptotic stability of the CT nonlinear system according to Lyapunov's stability theorem.

Nowadays, more attention is paid to ETC methods based on ADP with the development of ADP theory. Zhong and He [41] presented an ET ADP control algorithm with a state observer based on the input and output data. For nonlinear DT systems, Dong *et al.* [42] designed a new ET condition based ADP and analyzed the stability of the system. Using a single NN approximation structure, Wang *et al.* [43] dealt with the nonlinear optimal regulation problem in the framework of the adaptive critic NN based on the ET mechanism. The optimal control law can be derived by training the critic NN. The decentralized ETC strategy was designed in [44] for a class of nonlinear systems with uncertain cross-linked terms. It is shown in this paper that the decentralized ETC policy for the whole system can be represented by optimal ETC policies of auxiliary subsystems. Luo *et al.* [45] studied the problem of event-triggered optimal control (ETOC) for CT systems and provide formal performance guarantees by proving a predetermined upper bound. In [46], a critic NN was used to obtain an approximate solution to the ET HJB equation, and the weights of the critic NN were updated using gradient descent and empirical replay techniques.

In this paper, a novel ETOTC ADP method is developed for unknown CT affine nonlinear systems. First, an augmented system composed of the error system dynamics and the reference dynamics is used to introduce a new discounted performance index function (DPIF) for ETOTC. Second, a NN-based identifier is designed to obtain the unknown system dynamic. Third, for the identified augmented system, a particular ET adaptive implement method is developed without initial stable control law, and the closed loop identified augmented system and the critic weight estimation error are proved to be ultimately uniformly bounded (UUB).

Different from the existing methods, the contributions are summarized as follows.

1) We extend the ADP-based ETC method to address the optimal tracking control problem. It makes the actual trajectory of the unknown system track its desired one. Compared with the related ADP-based ET tracking methods, the developed ETOTC method does not require a steady-state control law or a feedforward control law, which is easier to implement in practical engineering.

2) We establish a new ET condition of the error system, which ensures the signals in the augmented closed loop system and the critic weight estimation error to be UUB. Compared with the offline system identification, the developed ETOTC uses real-time data online to achieve system identification.

3) Under the ET mechanism, we propose a novel NN weights update law for optimal tracking of unknown systems without requiring an initial stable control law, which is an extension time-triggered optimal control methods.

In this paper, Section II formulates the problem of ETOTC of CT nonlinear systems without the steady-state control law. Section III develops the NN-based identifier and critic, and the ET learning rule of the critic weights is proposed with the stability analyses. Section IV shows the detailed simulations and analyses. Section V gives some conclusions.

Notations: We use \mathbb{R} to represent the set of real numbers. \mathbb{R}_+ is the set of nonnegative real numbers. \mathbb{R}^m is the Euclidean space of dimension m . $\mathbb{R}^{n \times m}$ is the set of $n \times m$ real matrices. \mathbb{N} is the nonnegative integer set. $\mathbf{1}_n$ is the column vector of n ones. $\|\cdot\|$ is the vector norm on the Euclidean space. $\lambda_{\max}(\cdot)$ and $\lambda_{\min}(\cdot)$ are the maximum and minimum eigenvalues of a matrix, separately. $\Omega_a = \{x \in \mathbb{R}^n : \|x\| \leq a, a \in \mathbb{R}_+\}$ is a compact subset of \mathbb{R}^n . $\mathcal{A}(\Omega_a)$ is the admissible control set on Ω_a . $C^1(\Omega_a : \mathbb{R}^+)$ is the set of once differentiable functions with respect to its argument on Ω_a . $diag(\cdot)$ is a diagonal matrix operator.

II. PROBLEM FORMULATION

Let us take the following CT time-invariant nonlinear system into account

$$\dot{x}(t) = f(x(t)) + g(x(t))u(t) \quad (1)$$

where $x(t) \in \mathbb{R}^n$ and $u(t) \in \mathbb{R}^m$ are n -dimensional state and m -dimensional control input. $f(x(t)) \in \mathbb{R}^n$ and $g(x(t)) \in \mathbb{R}^{n \times m}$ are local Lipschitz continuous and differentiable on a compact set Ω_a with $f(0) = 0$ and $g(0) \neq 0$.

The desired reference trajectory $x_d(t)$ satisfies

$$\dot{x}_d = f_d(x_d(t)) \quad (2)$$

where $f_d(x_d(t))$ is local Lipschitz continuous and differentiable on the compact set Ω_a with $f_d(0) = 0$. Then, the tracking error can be defined as

$$e_d = x(t) - x_d(t) \quad (3)$$

and according to (1), the tracking error system is

$$\dot{e}_d(t) = f(x(t)) + g(x(t))u(t) - f_d(x_d(t)). \quad (4)$$

Inspired by [50]–[52], we define the augmented system

$$\dot{x}_a(t) = f_a(x_a(t)) + g_a(x_a(t))u(t) \quad (5)$$

where $x_a(t) = [e_d^T(t), x_d^T(t)]^T \in \mathbb{R}^{2n}$ is the augmented state, $f_a(x_a(t)) = [f(e_d(t) + x_d(t)) - f_d(x_d(t)), f_d(x_d(t))]^T$ is the augmented internal dynamic, $g_a(x_a) = [g(e_d(t) + x_d(t)), 0]^T$ is the augmented input dynamic function.

Then, we define the infinite-horizon DPIF of the augmented system as

$$V(x_a(t)) = \int_t^\infty e^{-\gamma(\tau-t)} \left[x_a^T(\tau) Q_a x_a(\tau) + u^T(\tau) R u(\tau) \right] d\tau \quad (6)$$

where

$$Q_a = \begin{bmatrix} Q & 0_{n \times n} \\ 0_{n \times n} & 0_{n \times n} \end{bmatrix}, Q \in \mathbb{R}^{n \times n} > 0 \quad (7)$$

$R \in \mathbb{R}^{m \times m} > 0$, and $\gamma > 0$ is the discount factor.

Remark 1: For the standard solution to the optimal tracking control problem, the desired reference satisfies

$$\dot{x}_d(t) = f(x_d(t)) + g(x_d(t))u_d(t). \quad (8)$$

The premise of founding (8) is that the inverse of the input dynamic function $g(x_d(t))$ exists, then the steady-state control law is acquired as

$$u_d(t) = g^{-1}(x_d(t))(\dot{x}_d(t) - f(x_d(t))). \quad (9)$$

However, $g^{-1}(x_d(t))$ can not be solved exactly due to the rank deficient input dynamic function $g(x_d(t))$. Thus, the standard solution to the optimal tracking control problem is invalid.

With regard to any integral time $T > 0$, the related DPIF satisfies

$$\int_t^{t+T} e^{-\gamma(\tau-t)} \left[x_a^T(\tau) Q_a x_a(\tau) + u^T(\tau) R u(\tau) \right] d\tau + e^{-\gamma T} V(x_a(t+T)) - V(x_a(t)) = 0. \quad (10)$$

For all $x_a(t) = x_a$, if $V(x_a) \in C^1(\Omega_a : \mathbb{R}^+)$, we let $T \rightarrow 0$ in (10), and obtain the time-triggered nonlinear Lyapunov equation as follows:

$$0 = \nabla^T V(x_a)(f_a(x_a) + g_a(x_a)u(t)) + x_a^T Q_a x_a + u^T(t) R u(t) - \gamma V(x_a). \quad (11)$$

The Hamiltonian of the augmented system (5) is defined as

$$\begin{aligned} H(x_a, u(x_a), \nabla V(x_a)) &= \nabla^T V(x_a)(f_a(x_a) + g_a(x_a)u(x_a)) \\ &\quad + x_a^T Q_a x_a + u^T(x_a) R u(x_a) \\ &\quad - \gamma V(x_a). \end{aligned} \quad (12)$$

According to the well-known Bellman's principle of optimality, it follows from (6) that the optimal DPIF is

$$\begin{aligned} V^*(x_a) &= \min_{u \in \mathcal{A}(\Omega_a)} \int_t^\infty e^{-\gamma(\tau-t)} \left[x_a^T Q_a x_a + u^T(\tau) R u(\tau) \right] d\tau \end{aligned} \quad (13)$$

which satisfies the time-triggered HJBE

$$\min_{u \in \mathcal{A}(\Omega_a)} H(x_a, u(x_a), \nabla V^*(x_a)) = 0 \quad (14)$$

and the time-triggered optimal tracking control law is

$$u^*(x_a) = -\frac{1}{2} R^{-1} g_a^T(x_a) \nabla V^*(x_a). \quad (15)$$

On the basis of event-triggered mechanism, the triggered instants are a monotone increasing sequence $\{t_j\}_{j=0}^\infty$ determined by the event-triggering condition, where $t_0 = 0$ and $t_j < t_{j+1}, j \in \mathbb{N}$ are the sampling instants. Then, the sampled augmented system state is

$$x_{aj} = x_{aj}(t) = \begin{cases} x_a(t), & t = t_j \\ x_a(t_j), & t \in [t_j, t_{j+1}) \end{cases} \quad (16)$$

where $j \in \mathbb{N}$, and $t_{j+1} - t_j$ is called the execution interval. The corresponding ETOTC is derived as

$$\mu(x_{aj}) = u^*(x_{aj}) = -\frac{1}{2} R^{-1} g_a^T(x_{aj}) \nabla V^*(x_{aj}) \quad (17)$$

where $\nabla V^*(x_{aj}) = \partial V^*(x_a) / \partial x_a|_{x_a=x_{aj}}$ and by leading into the zero-order hold (ZOH), $\mu(x_{aj})$ becomes a continuous function with respect to x_{aj} . Notice that the ETOTC (17) depends on the solution of the HJBE (14).

Next, We define the error between the sampled augmented state $x_{aj}(t)$ and the real state $x_a(t)$ as

$$e_{aj}(t) = x_a(t) - x_{aj}(t). \quad (18)$$

According to the (5), (15) and (18), the closed augmented system is

$$\begin{aligned} \dot{x}_a &= f_a(x_a) + g_a(x_a)\mu(x_{aj}) \\ &= f_a(x_a) - \frac{1}{2} g_a(x_a) R^{-1} g_a^T(x_{aj}) \nabla V^*(x_{aj}) \\ &= f_a(x_a) - \frac{1}{2} g_a(x_a) R^{-1} g_a^T(x_a - e_{aj}) \nabla V^*(x_a - e_{aj}). \end{aligned} \quad (19)$$

Remark 2: With regard to the time-triggered optimal tracking control $u^(x_a)$, the sampling time of the augmented system (5) is always triggering. Different from $u^*(x_a)$, the ETOCT (15) remains unchanged in the execution interval, i.e. sampled time $x_a(t_j)$ at instant t_j is maintained until the new sampled augmented state $x_a(t_{j+1})$ is transmitted into the controller at instant t_{j+1} , which significantly reduces the computational and communication burden.*

III. ADAPTIVE-IDENTIFIER-CRITIC DESIGN BASED EVENT-TRIGGERED OPTIMAL TRACKING CONTROL

In this section, a NN-based identifier is first designed, and the identifier error and the identifier weight estimation error are UUB, which is proved in Theorem 1. Then, for the identified augmented system, a particular ET adaptive implement method is developed without initial stable control law, and the closed loop identified augmented system and critic weight estimation error are proved to be UUB in Theorem 2.

A. NN-BASED IDENTIFIER DESIGN

The rewritten system dynamic can be expressed as follows

$$\dot{x} = A_h x + W_f^T \sigma_f(x) + W_g^T \sigma_g(x) u + \varepsilon_x(x) \quad (20)$$

where $A_h \in \mathbb{R}^{n \times n}$ is a Hurwitz matrix, $W_f \in \mathbb{R}^{n \times n}$, $W_g \in \mathbb{R}^{n \times n}$, $\sigma_f(x) \in \mathbb{R}^n$, $\sigma_g(x) \in \mathbb{R}^{n \times m}$. The ideal NN weight matrices W_f and W_g satisfy $W_f^T W_f \leq A_f$ and $W_g^T W_g \leq A_g$ where A_f and A_g are prior known positive definite matrices. On the compact set $\Omega_{\tilde{x}}$, the activation functions $\sigma_f(x)$ and $\sigma_g(x)$ need to satisfy

$$\|\sigma_f(x_1) - \sigma_f(x_2)\| \leq K_f \|x_1 - x_2\| \quad (21)$$

$$\|\sigma_g(x_1) \mathbb{1}_m - \sigma_g(x_2) \mathbb{1}_m\| \leq K_g \|x_1 - x_2\| \quad (22)$$

where $\mathbb{1}_m = [1, 1, \dots, 1]^T \in \mathbb{R}^m$.

Then, the reconstruct system dynamic is

$$\dot{\hat{x}} = A_h \hat{x} + \hat{W}_f^T \sigma_f(\hat{x}) + \hat{W}_g^T \sigma_g(\hat{x}) u + B(x - \hat{x}) \quad (23)$$

where $B = \text{diag}(b_0, b_1, \dots, b_n) > 0$ is the designed matrix. The error dynamic of identifier is

$$\begin{aligned} \dot{\tilde{x}} = A_h \tilde{x} + \tilde{W}_f^T \sigma_f(\hat{x}) + \tilde{W}_g^T \sigma_g(\hat{x}) u + W_f^T \tilde{\sigma}_f(x, \hat{x}) \\ + W_g^T \tilde{\sigma}_g(x, \hat{x}) u - B \tilde{x} + \varepsilon_x(x) \end{aligned} \quad (24)$$

where $\tilde{x} = x - \hat{x}$, $\tilde{W}_f = W_f - \hat{W}_f$, $\tilde{W}_g = W_g - \hat{W}_g$, $\tilde{\sigma}_f(x, \hat{x}) = \sigma_f(x) - \sigma_f(\hat{x})$, $\tilde{\sigma}_g(x, \hat{x}) = \sigma_g(x) - \sigma_g(\hat{x})$.

Theorem 1: For affine unknown nonlinear system (19), if the weight of NN-based identifier (23) has the update rules

$$\dot{\hat{W}}_f = P_f \sigma_f(\hat{x}) \tilde{x}^T P_0 \quad (25)$$

$$\dot{\hat{W}}_g = P_g \sigma_g(\hat{x}) u \tilde{x}^T P_0 \quad (26)$$

where P_0 , P_f and P_g are the predefined positive definite matrices. Then, the identification error $\tilde{x}(t)$ converges to the following compact set

$$\Omega_{\tilde{x}} = \left\{ \tilde{x} : \|\tilde{x}\| \leq \frac{\varepsilon_{xM}}{\sqrt{\lambda_{\min}(Q_0) + 2\lambda_{\min}(BP_0B) - K_0}} \right\} \quad (27)$$

where $K_0 > 0$ is the constant parameter which can be determined in the following proof. Besides that, the weight estimation errors \tilde{W}_f and \tilde{W}_g of the NN-based identifier are assured of the UUB.

Proof: Choose the Lyapunov function candidate as follows:

$$L_1(t) = \frac{1}{2} \tilde{x}^T P_0 \tilde{x} + \frac{1}{2} \text{tr}(\tilde{W}_f^T P_f^{-1} \tilde{W}_f + \tilde{W}_g^T P_g^{-1} \tilde{W}_g) \quad (28)$$

Since A_h is a Hurwitz matrix, there exists a positive-definite symmetric matrix P_0 satisfy Lyapunov's equation

$$A_h^T P_0 + P_0 A_h = -Q_0 \quad (29)$$

The time derivative of (28) can be expressed as

$$\begin{aligned} \dot{L}_1(t) &= \frac{1}{2} \dot{\tilde{x}}^T P_0 \tilde{x} + \frac{1}{2} \tilde{x}^T P_0 \dot{\tilde{x}} + \text{tr}(\tilde{W}_f^T P_f^{-1} \dot{\tilde{W}}_f + \tilde{W}_g^T P_g^{-1} \dot{\tilde{W}}_g) \\ &= -\frac{1}{2} \tilde{x}^T (A_h^T P_0 + P_0 A_h) \tilde{x} + \tilde{x}^T P_0 \tilde{W}_f^T \sigma_f(\hat{x}) \\ &\quad + \tilde{x}^T P_0 \tilde{W}_g^T \sigma_g(\hat{x}) u + \tilde{x}^T P_0 W_f^T \tilde{\sigma}_f(x, \hat{x}) \\ &\quad + \tilde{x}^T P_0 W_g^T \tilde{\sigma}_g(x, \hat{x}) u - \tilde{x}^T P_0 B \tilde{x} + \tilde{x}^T P_0 \varepsilon_x(x) \\ &\quad - \text{tr}(\tilde{W}_f^T \sigma_f(\hat{x}) \tilde{x}^T P_0 + \tilde{W}_g^T \sigma_g(\hat{x}) u \tilde{x}^T P_0) \end{aligned} \quad (30)$$

Observing $\text{tr}(A_1 A_2) = \text{tr}(A_2 A_1) = A_2 A_1$ for $A_1 \in \mathbb{R}^{n \times n}$ and $A_2 \in \mathbb{R}^{1 \times n}$. Furthermore, according to Cauchy-Schwarz inequality, (30) becomes

$$\begin{aligned} \dot{L}_1(t) &\leq -\frac{1}{2} \tilde{x}^T Q_0 \tilde{x} + \frac{1}{2} \tilde{x}^T P_0 W_f^T W_f P_0 \tilde{x}^T \\ &\quad + \frac{1}{2} \tilde{\sigma}_f^T(x, \hat{x}) \tilde{\sigma}_f(x, \hat{x}) + \frac{1}{2} \tilde{x}^T P_0 W_g^T W_g P_0 \tilde{x}^T \\ &\quad + \frac{1}{2} u^T \tilde{\sigma}_g^T(x, \hat{x}) \tilde{\sigma}_g(x, \hat{x}) u - \tilde{x}^T B P_0 B \tilde{x} \\ &\quad + \frac{1}{2} \tilde{x}^T P_0^2 \tilde{x} + \frac{1}{2} \varepsilon_x^T(x) \varepsilon_x(x) \\ &\leq -\frac{1}{2} \lambda_{\min}(Q_0) \|\tilde{x}\|^2 + \frac{1}{2} \lambda_{\max}(P_0 A_f P_0) \|\tilde{x}\|^2 \\ &\quad + \frac{1}{2} K_f^2 \|\tilde{x}\|^2 + \frac{1}{2} \lambda_{\max}(P_0 A_g P_0) \|\tilde{x}\|^2 + \frac{1}{2} u_M^2 K_g^2 \|\tilde{x}\|^2 \\ &\quad - \lambda_{\min}(B P_0 B) \|\tilde{x}\|^2 + \frac{1}{2} \lambda_{\max}(P_0^2) \|\tilde{x}\|^2 + \frac{1}{2} \|\varepsilon_{xM}\|^2 \\ &= \frac{1}{2} (\lambda_{\max}(P_0^2) + \lambda_{\max}(P_0 A_f P_0) + \lambda_{\max}(P_0 A_g P_0) \\ &\quad - \lambda_{\min}(Q_0) - 2\lambda_{\min}(B P_0 B) + K_f^2 \\ &\quad + u_M^2 K_g^2) \|\tilde{x}\|^2 + \frac{1}{2} \|\varepsilon_{xM}\|^2 \end{aligned} \quad (31)$$

In order to guarantee that $\lambda_{\max}(P_0^2) + \lambda_{\max}(P_0 A_f P_0) + \lambda_{\max}(P_0 A_g P_0) - \lambda_{\min}(Q_0) - 2\lambda_{\min}(B P_0 B) + K_f^2 + u_M^2 K_g^2 < 0$, the parameter matrices B and Q_0 need to be selected properly. Then, $\dot{L}_1(t) < 0$ as long as the identifier error \tilde{x} satisfies the following inequality:

$$\|\tilde{x}\| > \frac{\varepsilon_{xM}}{\sqrt{\lambda_{\min}(Q_0) + 2\lambda_{\min}(B P_0 B) - K_0}} \quad (32)$$

where $K_0 = \lambda_{\max}(P_0^2) + \lambda_{\max}(P_0 A_f P_0) + \lambda_{\max}(P_0 A_g P_0) + K_f^2 + u_M^2 K_g^2$. According to the Lyapunov's direct methods, the identifier error \tilde{x} converges to $\Omega_{\tilde{x}}$ and the weight estimation errors \tilde{W}_f and \tilde{W}_g are UUB. This completes the proof. \square

B. ADAPTIVE-CRITIC DESIGN BASED EVENT-TRIGGERED OPTIMAL TRACKING CONTROL

In this part, we present an online implementation method based on NNs. We have the fact that the unknown nonlinear system (1) can be approximated by identifier NN with in

a sufficiently small compact set $\Omega_{\hat{x}}$. According to (23), the identified augmented system can be rewritten as follows:

$$\dot{\hat{x}}_a = \mathcal{F}_a(\hat{x}_a) + \mathcal{G}_a(\hat{x}_a)u(t) \quad (33)$$

where $\hat{x}_a = [\hat{e}_d^\top(t), x_d^\top(t)]^\top$, $\hat{e}_d = \hat{x}(t) - x_d(t)$,

$$\mathcal{F}_a(\hat{x}_a) = \begin{bmatrix} A_h(\hat{e}_d + x_d) \\ 0 \end{bmatrix} + \begin{bmatrix} B(e_d + x_d - \hat{e}_d - x_d) \\ 0 \end{bmatrix} + \begin{bmatrix} \hat{W}_f^\top \sigma_f(\hat{e}_d + x_d) - f_d(x_d) \\ f_d(x_d) \end{bmatrix} \quad (34)$$

and

$$\mathcal{G}_a(\hat{x}_a) = \begin{bmatrix} \hat{W}_g^\top \sigma_g(\hat{e}_d + x_d) \\ 0 \end{bmatrix}. \quad (35)$$

Then, we assume that the optimal value function $V^*(\hat{x}_a) \in C^1(\Omega_a; \mathbb{R}_+)$ exists. By the universal approximation property of feed-forward NNs [47], a critic NN is used to construct $V^*(\hat{x}_a)$ on the compact set Ω_a as

$$\begin{aligned} V_c^*(\hat{x}_a) &= \sum_{i=1}^l W_{ci} \sigma_{ci}(\hat{x}_a) + \varepsilon_c(\hat{x}_a) \\ &= (W_c)^\top \sigma_c(\hat{x}_a) + \varepsilon_c(\hat{x}_a) \end{aligned} \quad (36)$$

where $W_c = [W_{c1}, W_{c2}, \dots, W_{cl}]^\top$ is the ideal NN weight vector, $\sigma_c(\hat{x}_a) = [\sigma_{c1}(\hat{x}_a), \sigma_{c2}(\hat{x}_a), \dots, \sigma_{cl}(\hat{x}_a)]^\top$ is the activation function vector satisfying that $\sigma_{ci} \in C^1(\Omega_a; \mathbb{R})$ and $\sigma_{ci}(0) = 0$, and l is the quantity of activation functions with $\lim_{l \rightarrow \infty} \varepsilon_c(\hat{x}_a) = 0$. Generally, the activation functions are chosen to be linearly independent. However, the ideal critic NN weight vector W_c^* which offers the best approximation to the tracking HJBE is always unknown, then the optimal value function $\hat{V}(\hat{x}_a)$ can be approximated by

$$\hat{V}_c(\hat{x}_a) = \sum_{i=1}^l \hat{W}_{ci} \sigma_{ci}(\hat{x}_a) = \hat{W}_c^\top \sigma_c(\hat{x}_a). \quad (37)$$

According to the rewritten augmented system (33), (36) and (37), the optimal event-triggered control law and approximate event-triggered control law are derived separately.

$$\mu(\hat{x}_{aj}) = -\frac{1}{2}R^{-1}\mathcal{G}_a^\top(\hat{x}_{aj})(\nabla^\top \sigma_c(\hat{x}_{aj})W_c + \nabla \varepsilon_c(\hat{x}_{aj})) \quad (38)$$

and

$$\hat{\mu}(\hat{x}_{aj}) = -\frac{1}{2}R^{-1}\hat{\mathcal{G}}_a^\top(\hat{x}_{aj})\nabla^\top \sigma_c(\hat{x}_{aj})\hat{W}_c \quad (39)$$

where

$$\hat{x}_{aj}(t) = \begin{cases} \hat{x}_a(t), & t = t_j \\ \hat{x}_a(t_j), & t \in [t_j, t_{j+1}). \end{cases} \quad (40)$$

According to the approximate event-triggered control law (39), the identified closed-loop system is

$$\dot{\hat{x}}_a = \mathcal{F}_a(\hat{x}_a) + \mathcal{G}_a(\hat{x}_a)\hat{\mu}(\hat{x}_{aj}). \quad (41)$$

The triggering condition of the identified nonlinear system is

$$\|e_{aj}(t)\|^2 \leq \frac{2((1 - \xi^2)\lambda(Q)\|e_d\|^2 + \|R\hat{\mu}(\hat{x}_{aj})\|)}{\|\Lambda\|^2\|R^{-1}\|^2(\mathcal{G}_{aM}^2\mathcal{L}_{\sigma_{cM}}^2 + d_{\sigma_{cM}}^2\mathcal{L}_{\mathcal{G}_{aM}}^2)}\|\hat{W}_c\|^2. \quad (42)$$

Before implementing the classical adaptive-critic design approach, a special NN weight vector needs to be chosen in order to create an initial stable control law, and the critic NN weights can be trained by the classical approach. However, the special NN weight vector is difficult to choose, which may lead to the instability for the closed loop system. Inspired by [45], [48], [49], we adopt an auxiliary Lyapunov function to improve the learning criterion and employ it to tune the critic NN weights. Moreover, the following assumption is introduced.

Assumption 1: Consider that $L_a(\hat{x}_a) \in C^1(\Omega_a; \mathbb{R}_+)$ is an auxiliary differentiable Lyapunov function candidate, and $L_a(x)$ satisfies the following inequality.

$$\nabla^\top L_a(\hat{x}_a)(\mathcal{F}_a(\hat{x}_a) + \mathcal{G}_a(\hat{x}_a)\mu(\hat{x}_{aj})) \leq 0. \quad (43)$$

In addition, let $P_a \in \mathbb{R}^{n \times n}$ be a positive definite matrix, and the following condition is true on the compact set Ω_a .

$$\begin{aligned} \nabla^\top L_a(\hat{x}_a)(\mathcal{F}_a(\hat{x}_a) + \mathcal{G}_a(\hat{x}_a)\mu(\hat{x}_{aj})) \\ \leq -\lambda_{\min}(P_a)\|\nabla L_a(\hat{x}_a)\|^2. \end{aligned} \quad (44)$$

Remark 3: During the realization process, we can choose a polynomial, such as $\hat{x}_a^\top \hat{x}_a$, to determine the auxiliary Lyapunov function $L_a(\hat{x}_a)$, which guarantees the stability of the augmented state vector \hat{x}_a and critic NN weights \hat{W}_c .

Due to the fact that it is difficult to obtain the ideal critic NN weight vector W_c , we design an online learning approach to approximate W_c by employing the real-time data. Furthermore, $V_c^*(\hat{x}_a)$ and $\mu(\hat{x}_{aj})$ satisfy the event-triggered HJBE $H(\hat{x}_a, \mu(\hat{x}_{aj}), \nabla V^*(\hat{x}_a)) = 0$. Thus, we define the approximate error between $H(\hat{x}_a, \mu(\hat{x}_{aj}), \nabla V^*(\hat{x}_a))$ and $\hat{H}(\hat{x}_a, \hat{\mu}(\hat{x}_{aj}), \nabla \hat{V}(\hat{x}_a))$, that is

$$\begin{aligned} e_c(\hat{x}_a, \hat{W}_c) &= \hat{H}(\hat{x}_a, \hat{\mu}(\hat{x}_{aj}), \nabla \hat{V}_c(\hat{x}_a)) - H(\hat{x}_a, \mu(\hat{x}_{aj}), \nabla V_c^*(\hat{x}_a)) \\ &= \hat{H}(\hat{x}_a, \hat{\mu}(\hat{x}_{aj}), \nabla \hat{V}_c(\hat{x}_a)) \\ &= \hat{W}_c^\top \nabla \sigma_c(\hat{x}_a)(\mathcal{F}_a(\hat{x}_a) + \mathcal{G}_a(\hat{x}_a)\hat{\mu}(\hat{x}_{aj})) + \hat{x}_a^\top Q_a \hat{x}_a \\ &\quad + \hat{\mu}^\top(\hat{x}_{aj})R\hat{\mu}(\hat{x}_{aj}) - \gamma \hat{W}_c^\top \sigma_c(\hat{x}_a). \end{aligned} \quad (45)$$

The critic NN weight vector \hat{W}_c can be tuned by minimizing the square approximate errors $E_c(\hat{x}_a, \hat{W}_c) = (1/2)e_c^\top e_c$. Furthermore, in order to avoid $\dot{L}_a(\hat{x}_a) = \nabla^\top L_a(\hat{x}_a)(\mathcal{F}_a(\hat{x}_a) + \mathcal{G}_a(\hat{x}_a)\hat{\mu}(\hat{x}_{aj})) > 0$, the auxiliary term is introduced into the tuning law for \hat{W}_c , which can be derived as follow.

$$\begin{aligned} \dot{\hat{W}}_c &= -\eta_1 \frac{1}{(1 + (\delta \hat{\mu}(\hat{x}_a))^\top \delta \hat{\mu}(\hat{x}_a))^2} \frac{\partial E_c(\hat{x}_a, \hat{W}_c)}{\partial \hat{W}_c} \\ &\quad - \eta_2 \frac{\partial (\nabla^\top L_a(\hat{x}_a)(\mathcal{F}_a(\hat{x}_a) + \mathcal{G}_a(\hat{x}_a)\hat{\mu}(\hat{x}_{aj})))}{\partial \hat{W}_c} \end{aligned} \quad (46)$$

where $\eta_1 > 0$ is a learning rate parameter,

$$\delta^{\hat{\mu}}(\hat{x}_a) = \nabla \sigma_c(\hat{x}_a)(\mathcal{F}_a(\hat{x}_a) + \mathcal{G}_a(\hat{x}_a)\hat{\mu}(\hat{x}_{aj})) - \gamma \sigma_c(\hat{x}_a) \quad (47)$$

the term $(1 + \delta^{\hat{\mu}}(\hat{x}_a)\delta(\hat{x}_a))^2$ is the normalization of the tuning process, and η_2 is a constant parameter which is designed to avoid the instability of the closed-loop system.

Moreover, according to the approximate event-triggered control (39), the weights tuning law (46) becomes

$$\begin{aligned} \dot{\hat{W}}_c = & -\eta_1 \frac{\delta^{\hat{\mu}}(\hat{x}_a)}{(1 + (\delta^{\hat{\mu}}(\hat{x}_a))^{\top} \delta^{\hat{\mu}}(\hat{x}_a))^2} e_c(\hat{x}_a, \hat{W}_c) \\ & + \frac{1}{2} \eta_2 \nabla \sigma_c(\hat{x}_{aj}) \mathcal{G}_a(\hat{x}_{aj}) R^{-1} \mathcal{G}_a^{\top}(\hat{x}_a) \nabla L_a(\hat{x}_a). \end{aligned} \quad (48)$$

The critic NN weights tuning law (48) is contained with decreasing the HJBE approximate error and stabilizing the closed-loop system. Thus we can select initial critic NN weights $\hat{W}_c(0)$ arbitrarily without considering the initial stabilizing control law.

C. THEORETICAL ANALYSIS

In this part, we show the theoretical analysis of the ACD based ETOTC method. The critic error vector between W_c and \hat{W}_c is defined as $\tilde{W}_c = \hat{W}_c - W_c$, and we further have $\dot{\tilde{W}}_c = \dot{\hat{W}}_c$. By introducing $\delta_1 = \delta^{\hat{\mu}}(\hat{x}_a)/(1 + (\delta^{\hat{\mu}}(\hat{x}_a))^{\top} \delta^{\hat{\mu}}(\hat{x}_a))$, $\delta_2 = 1 + (\delta^{\hat{\mu}}(\hat{x}_a))^{\top} \delta^{\hat{\mu}}(\hat{x}_a)$, The critic error vector dynamic is

$$\begin{aligned} \dot{\tilde{W}}_c = & -\eta_1 \delta_1 \delta_1^{\top} \tilde{W}_c + \eta_1 \frac{\delta_1}{\delta_2} \nabla^{\top} \varepsilon_c(\hat{x}_a)(\mathcal{F}_a(\hat{x}_a) \\ & + \mathcal{G}_a(\hat{x}_a)\hat{\mu}(\hat{x}_{aj})) + \frac{1}{2} \eta_2 \nabla \sigma_c(\hat{x}_{aj}) \mathcal{G}_a(\hat{x}_{aj}) R^{-1} \\ & \times \mathcal{G}_a^{\top}(\hat{x}_a) \nabla L_a(\hat{x}_a). \end{aligned} \quad (49)$$

where $e_{cH} = \nabla^{\top} \varepsilon_c(\hat{x}_a)(\mathcal{F}_a(\hat{x}_a) + \mathcal{G}_a(\hat{x}_a)\hat{\mu}(\hat{x}_{aj}))$ is the residual error of reconstruct HJBE, and its norm bounded, i.e. $\|e_{cH}\| \leq e_M$ [43], [53]. During the training process, the persistence of excitation (PE) condition ensures $\lambda_{\min}(\delta_1 \delta_1^{\top}) > 0$, which is necessary in the stability analysis [54], [55].

Before proceeding the stability analysis, some essential assumptions are employed to support the analysis.

Assumption 2: The identified augmented input dynamic function $\mathcal{G}_a(\hat{x}_a)$ is local Lipschitz continuity on the compact set Ω_a such that $\|\mathcal{G}_a(\hat{x}_a) - \mathcal{G}_a(\hat{x}_{aj})\| \leq \mathcal{L}_{\mathcal{G}_{aM}} \|e_{aj}(t)\|$, where $\mathcal{L}_{\mathcal{G}_{aM}}$ is a positive constant, and $\mathcal{G}_a(\hat{x}_a)$ is norm bounded, i.e. $\|\mathcal{G}_a(\hat{x}_a)\| \leq \mathcal{G}_{aM}$, where \mathcal{G}_{aM} is a positive constant.

Assumption 3: The derivative of the activation function $\nabla \sigma_c(\hat{x}_a)$ is local Lipschitz continuity on the compact set Ω_a such that $\|\nabla \sigma_c(\hat{x}_a) - \nabla \sigma_c(\hat{x}_{aj})\| \leq \mathcal{L}_{\sigma_{cM}} \|e_{aj}(t)\|$, where $\mathcal{L}_{\sigma_{cM}}$ is a positive constant, and the terms $\nabla \sigma_c(\hat{x}_a)$, $\nabla \varepsilon_c(\hat{x}_a)$ and $V_c^*(\hat{x}_a)$ are norm bounded, i.e. $\|\nabla \sigma_c(\hat{x}_a)\| \leq d_{\sigma_{cM}}$, $\|\nabla \varepsilon_c(\hat{x}_a)\| \leq d_{\varepsilon_{cM}}$ and $\|V_c^*(\hat{x}_a)\| \leq V_{cM}$.

Theorem 2: For the identified system (33) associated with the approximate error of the event-triggered HJBE (45), let Assumptions 2–3 hold. The event-triggered approximate optimal control law (39), the critic weight tuning rule (48) and the event-triggered condition (42) guarantee that the closed loop system state and the weight estimation error are UUB.

Proof: Choose the Lyapunov function candidate as follows

$$L_2(t) = V_c^*(\hat{x}_a) + V_c^*(\hat{x}_{aj}) + 2\eta_2 L_a(\hat{x}_a) + \tilde{W}_c^{\top} \tilde{W}_c \quad (50)$$

Case 1: When $t \in [t_j, t_{j+1})$ for all j , there is no triggering in events. We can take the time derivative of (50) along with (19) and (49), the we have

$$\dot{L}_2(t) = \dot{V}_c^*(\hat{x}_a) + \dot{V}_c^*(\hat{x}_{aj}) + 2\eta_2 \dot{L}_a(\hat{x}_a) + \tilde{W}_c^{\top} \dot{\tilde{W}}_c. \quad (51)$$

For the first term of the right-hand side of (50), according to the Assumptions 2–3, we have

$$\begin{aligned} \dot{V}_c^*(\hat{x}_a) = & \nabla^{\top} V_c^*(\hat{x}_a)(\mathcal{F}_a(\hat{x}_a) + \mathcal{G}_a(\hat{x}_a)\hat{\mu}(\hat{x}_{aj})) \\ = & -\hat{x}_a^{\top} Q_a \hat{x}_a + (u^*(\hat{x}_a))^{\top} R u^*(\hat{x}_a) \\ & - 2(u^*(\hat{x}_a))^{\top} R \hat{\mu}(\hat{x}_{aj}) + \gamma V_c^*(\hat{x}_a) \\ = & -\hat{e}_d^{\top} Q \hat{e}_d + \gamma V_c^*(\hat{x}_a) + (\Lambda u^*(\hat{x}_a))^{\top} \Lambda u^*(\hat{x}_a) \\ & - 2(\Lambda u^*(\hat{x}_a))^{\top} \Lambda \hat{\mu}(\hat{x}_{aj}) + (\Lambda \hat{\mu}(\hat{x}_{aj}))^{\top} \Lambda \hat{\mu}(\hat{x}_{aj}) \\ & - (\hat{\mu}(\hat{x}_{aj}))^{\top} R \hat{\mu}(\hat{x}_{aj}) \\ \leq & -\hat{e}_d^{\top} Q \hat{e}_d + \gamma V_c^*(\hat{x}_a) - \lambda_{\min}(R) \|\hat{\mu}(\hat{x}_{aj})\|^2 \\ & + \|\Lambda\|^2 \|u^*(\hat{x}_a) - \hat{\mu}(\hat{x}_{aj})\|^2 \\ \leq & -\hat{e}_d^{\top} Q \hat{e}_d + \gamma V_c^*(\hat{x}_a) - \lambda_{\min}(R) \|\hat{\mu}(\hat{x}_{aj})\|^2 \\ & + \frac{1}{4} \|\Lambda\|^2 \|R^{-1}\|^2 (\|\mathcal{G}_a^{\top}(\hat{x}_a)(\nabla \varepsilon_c(\hat{x}_a) \\ & - \nabla^{\top} \sigma_c(\hat{x}_a) \tilde{W}_c)\|^2 + \|\mathcal{G}_a^{\top}(\hat{x}_a) \nabla^{\top} \sigma_c(\hat{x}_a) \\ & - \mathcal{G}_a^{\top}(\hat{x}_{aj}) \nabla^{\top} \sigma_c(\hat{x}_{aj})\|^2 \|\tilde{W}_c\|^2) \\ \leq & -\hat{e}_d^{\top} Q \hat{e}_d + \gamma V_c^*(\hat{x}_a) - \lambda_{\min}(R) \|\hat{\mu}(\hat{x}_{aj})\|^2 \\ & + \frac{1}{4} \|\Lambda\|^2 \|R^{-1}\|^2 (2\|\mathcal{G}_a(\hat{x}_a)\|^2 (\|\nabla \varepsilon_c(\hat{x}_a)\|^2 \\ & + \|\nabla \sigma_c(\hat{x}_a)\|^2 \|\tilde{W}_c\|^2) + 2(\|\nabla \sigma_c(\hat{x}_a) \\ & - \nabla \sigma_c(\hat{x}_{aj})\mathcal{G}_a(\hat{x}_a)\|^2 + \|\nabla \sigma_c(\hat{x}_{aj})(\mathcal{G}_a(\hat{x}_a) \\ & - \mathcal{G}_a(\hat{x}_{aj}))\|^2 \|\tilde{W}_c\|^2) \\ \leq & -\hat{e}_d^{\top} Q \hat{e}_d + \gamma V_c^*(\hat{x}_a) - \lambda_{\min}(R) \|\hat{\mu}(\hat{x}_{aj})\|^2 \\ & + \frac{1}{2} \|\Lambda\|^2 \|R^{-1}\|^2 (\mathcal{G}_{aM}^2 d_{\varepsilon_{cM}}^2 + \mathcal{G}_{aM}^2 d_{\sigma_{cM}}^2 \|\tilde{W}_c\|^2 \\ & + (\mathcal{G}_{aM}^2 \mathcal{L}_{\sigma_{cM}}^2 + d_{\sigma_{cM}}^2 \mathcal{L}_{\mathcal{G}_{aM}}^2) \|e_{aj}(t)\|^2 \|\tilde{W}_c\|^2). \end{aligned} \quad (52)$$

On the basis of the definition of $\hat{x}_{aj}(t)$ and (40), $\hat{x}_{aj}(t)$ is unchanged on $t \in [t_j, t_{j+1})$. Thus, the second term of the right-hand side of (50) becomes

$$\dot{V}_c^*(\hat{x}_{aj}) = 0. \quad (53)$$

For the third term of the right-hand side of (50), the derivative of $L_a(\hat{x}_a)$ is expressed as

$$\dot{L}_a(\hat{x}_a) = \nabla^{\top} L_a(\hat{x}_a)(\mathcal{F}_a(\hat{x}_a) + \mathcal{G}_a(\hat{x}_a)\hat{\mu}(\hat{x}_{aj})). \quad (54)$$

According to (48), (49) and Young's inequality, the fourth term of the right-hand side of (50) is

$$\begin{aligned} \tilde{W}_c^{\top} \dot{\tilde{W}}_c \leq & -2\eta_1 \tilde{W}_c^{\top} \delta_1 \delta_1^{\top} \tilde{W}_c + 2\eta_1 \frac{\delta_1}{\delta_2} \tilde{W}_c^{\top} \nabla^{\top} \varepsilon_c(\hat{x}_a)(\mathcal{F}_a(\hat{x}_a) \\ & + \mathcal{G}_a(\hat{x}_a)\hat{\mu}(\hat{x}_{aj})) + \eta_2 \tilde{W}_c^{\top} \nabla \sigma_c(\hat{x}_{aj}) \mathcal{G}_a(\hat{x}_{aj}) R^{-1} \\ & \times \mathcal{G}_a^{\top}(\hat{x}_a) \nabla L_a(\hat{x}_a) \end{aligned}$$

$$\begin{aligned}
 &\leq -(2\eta_1 - 1)\lambda_{\min}(\delta_1 \delta_1^T) \|\tilde{W}_c\|^2 + \eta_1^2 e_M^2 \\
 &\quad + \eta_2 \tilde{W}_c^T \nabla \sigma_c(\hat{x}_{aj}) \mathcal{G}_a(\hat{x}_{aj}) R^{-1} \mathcal{G}_a^T(\hat{x}_a) \nabla L_a(\hat{x}_a) \\
 &= -(2\eta_1 - 1)\lambda_{\min}(\delta_1 \delta_1^T) \|\tilde{W}_c\|^2 + \eta_1^2 e_M^2 \\
 &\quad + 2\eta_2 \nabla^T L_a(\hat{x}_a) \mathcal{G}_a(\hat{x}_a) \mu^*(\hat{x}_{aj}) \\
 &\quad + \eta_2 \nabla^T L_a(\hat{x}_a) \mathcal{G}_a(\hat{x}_a) R^{-1} \nabla \varepsilon_c(\hat{x}_{aj}) \\
 &\quad - 2\eta_2 \nabla^T L_a(\hat{x}_a) \mathcal{G}_a(\hat{x}_a) \hat{\mu}(\hat{x}_{aj}). \tag{55}
 \end{aligned}$$

Combining (52)–(55), and taking the optimal ET control law (38) into consideration, the overall time derivative of $L_2(t)$ is

$$\begin{aligned}
 \dot{L}_2(t) &\leq -\hat{e}_d^T Q \hat{e}_d + \gamma V_c^*(\hat{x}_a) - \lambda_{\min}(R) \|\hat{\mu}(\hat{x}_{aj})\|^2 \\
 &\quad + \frac{1}{2} \|\Lambda\|^2 \|R^{-1}\|^2 (\mathcal{G}_{aM}^2 d_{\varepsilon_{cM}}^2 + \mathcal{G}_{aM}^2 d_{\sigma_{cM}}^2 \|\tilde{W}_c\|^2 \\
 &\quad + (\mathcal{G}_{aM}^2 \mathcal{L}_{\sigma_{cM}}^2 + d_{\sigma_{cM}}^2 \mathcal{L}_{\mathcal{G}_{aM}}^2) \|e_{aj}(t)\|^2 \|\hat{W}_c\|^2) \\
 &\quad + 2\eta_2 \nabla^T L_a(\hat{x}_a) (\mathcal{F}_a(\hat{x}_a) + \mathcal{G}_a(\hat{x}_a) \mu^*(\hat{x}_{aj})) \\
 &\quad - (2\eta_1 - 1)\lambda_{\min}(\delta_1 \delta_1^T) \|\tilde{W}_c\|^2 + \eta_1^2 e_M^2 \\
 &\quad + \xi^2 \eta_2 \nabla^T L_a(\hat{x}_a) \mathcal{G}_a(\hat{x}_a) R^{-1} \nabla \varepsilon_c(\hat{x}_{aj}) \\
 &\leq -\lambda_{\min}(Q) \xi^2 \|\hat{e}_d\|^2 + (\xi^2 - 1)\lambda_{\min}(Q) \|\hat{e}_d\|^2 \\
 &\quad + \gamma \mathcal{V}_{cM} - \lambda_{\min}(R) \|\hat{\mu}(\hat{x}_{aj})\|^2 \\
 &\quad + \mathcal{M}_1 \|\tilde{W}_c\|^2 + \mathcal{M}_2 \|e_{aj}(t)\|^2 \|\hat{W}_c\|^2 \\
 &\quad - 2\eta_2 \lambda_{\min}(P_a) \|\nabla L_a(\hat{x}_a)\|^2 + \eta_1^2 e_M^2 \\
 &\quad + \frac{1}{2} \|\Lambda\|^2 \|R^{-1}\|^2 \mathcal{G}_{aM}^2 d_{\varepsilon_{cM}}^2 \\
 &\quad + \eta_2 \|R^{-1}\| \mathcal{G}_{aM}^2 d_{\varepsilon_{cM}} \|\nabla L_a(\hat{x}_a)\| \\
 &= -\xi^2 \lambda_{\min}(Q) \xi^2 \|\hat{e}_d\|^2 + (\xi^2 - 1)\lambda_{\min}(Q) \|\hat{e}_d\|^2 \\
 &\quad + \gamma \mathcal{V}_{cM} - \lambda_{\min}(R) \|\hat{\mu}(\hat{x}_{aj})\|^2 \\
 &\quad + \mathcal{M}_1 \|\tilde{W}_c\|^2 + \mathcal{M}_2 \|e_{aj}(t)\|^2 \|\hat{W}_c\|^2 \\
 &\quad - 2\eta_2 \lambda_{\min}(P_a) \left(\|\nabla L_a(\hat{x}_a)\| - \frac{\|R^{-1}\| \mathcal{G}_{aM}^2 d_{\varepsilon_{cM}}}{4\lambda_{\min}(P_a)} \right)^2 \\
 &\quad + \eta_1^2 e_M^2 + \frac{1}{2} \|\Lambda\|^2 \|R^{-1}\|^2 \mathcal{G}_{aM}^2 d_{\varepsilon_{cM}}^2 \\
 &\quad + \frac{\|R^{-1}\|^2 \mathcal{G}_{aM}^4 d_{\varepsilon_{cM}}^2}{16\lambda_{\min}^2(P_a)} \tag{56}
 \end{aligned}$$

where $\mathcal{M}_1 = \frac{1}{2} \|\Lambda\|^2 \|R^{-1}\|^2 \mathcal{G}_{aM}^2 d_{\sigma_{cM}}^2 - (2\eta_1 - 1)\lambda_{\min}(\delta_1 \delta_1^T)$, $\mathcal{M}_2 = \frac{1}{2} \|\Lambda\|^2 \|R^{-1}\|^2 (\mathcal{G}_{aM}^2 \mathcal{L}_{\sigma_{cM}}^2 + d_{\sigma_{cM}}^2 \mathcal{L}_{\mathcal{G}_{aM}}^2)$. When the ET condition (42) satisfies, (56) becomes

$$\begin{aligned}
 \dot{L}_2(t) &\leq -\xi^2 \lambda_{\min}(Q) \xi^2 \|\hat{e}_d\|^2 + \mathcal{M}_1 \|\tilde{W}_c\|^2 \\
 &\quad + \gamma \mathcal{V}_{cM} + \mathcal{M}_3 \tag{57}
 \end{aligned}$$

where \mathcal{M}_3 is formed as a positive constant

$$\begin{aligned}
 \mathcal{M}_3 &= \eta_1^2 e_M^2 + \frac{1}{2} \|\Lambda\|^2 \|R^{-1}\|^2 \mathcal{G}_{aM}^2 d_{\varepsilon_{cM}}^2 \\
 &\quad + \frac{\|R^{-1}\|^2 \mathcal{G}_{aM}^4 d_{\varepsilon_{cM}}^2}{16\lambda_{\min}^2(P_a)} \tag{58}
 \end{aligned}$$

According to (57), $\dot{L}_2(t) < 0$, provided \hat{e}_d lies out of the set

$$\Omega_{\hat{e}_d} = \left\{ \hat{e}_d : \|\hat{e}_d\| \leq \sqrt{\frac{\gamma \mathcal{V}_{cM} + \mathcal{M}_3}{\xi^2 \lambda_{\min}(Q)}} \right\} \tag{59}$$

or \tilde{W}_c lies out of the set

$$\Omega_{\tilde{W}_c} = \left\{ \tilde{W}_c : \|\tilde{W}_c\| \leq \sqrt{\frac{\gamma \mathcal{V}_{cM} + \mathcal{M}_3}{-\mathcal{M}_2}} \right\} \tag{60}$$

with the condition $\mathcal{M}_2 < 0$.

Case 2: If the state jumps at the ET time $t = t_{j+1}$, the time difference of the Lyapunov function candidate $L_2(t)$ is

$$\begin{aligned}
 \Delta L_2(t) &= L_2(\hat{x}_{aj+1}) - L_2(\hat{x}_a(t_{j+1}^-)) \\
 &= V_c^*(\hat{x}_{aj+1}) - V_c^*(\hat{x}_a(t_{j+1}^-)) + V_c^*(\hat{x}_{aj+1}) \\
 &\quad - V_c^*(\hat{x}_{aj}) + \eta_2 (L_a(\hat{x}_{aj+1}) - L_a(\hat{x}_a(t_{j+1}^-))) \\
 &\quad + \frac{1}{2} (\tilde{W}_c^T(\hat{x}_{aj+1}) \tilde{W}_c(\hat{x}_{aj+1}) \\
 &\quad - \tilde{W}_c^T(\hat{x}_a(t_{j+1}^-)) \tilde{W}_c(\hat{x}_a(t_{j+1}^-))) \tag{61}
 \end{aligned}$$

where $\hat{x}_a(t_{j+1}^-) = \lim_{\varepsilon \rightarrow 0^+} \hat{x}_a(t_{j+1} + \varepsilon)$. According to the fact that \hat{x}_a and $V^*(\hat{x}_a)$ are continuous in time interval $[t_j, t_{j+1})$, and $L_2(t) < 0$ when \hat{e}_d lies out of the set $\Omega_{\hat{e}_d}$ and \tilde{W}_c lies out of the set $\Omega_{\tilde{W}_c}$, we can obtain

$$\begin{aligned}
 \Delta L_2(t) &\leq V_c^*(\hat{x}_{aj+1}) - V_c^*(\hat{x}_{aj}) \\
 &\leq -\kappa (\|\hat{x}_{aj+1} - \hat{x}_{aj}\|) \tag{62}
 \end{aligned}$$

where $\kappa(\cdot)$ is a class κ -function. Therefore, the Lyapunov function $\Delta L_2(t) < 0$ at all the ET time t_{j+1} .

On the basis of the above two cases, $L_2(t) < 0$ in the time interval $[t_j, t_{j+1})$, and $\Delta L_2(t) < 0$ at the each ET time t_{j+1} for all $j \in \mathbb{N}$ as long as \hat{e}_d lies out of the compact set $\Omega_{\hat{e}_d}$ and \tilde{W}_c lies out of the set $\Omega_{\tilde{W}_c}$. Therefore, the triggering condition (42) and the inequalities (59) and (60) ensure that the identified closed-loop system (41) and the critic weight error dynamic are UUB. This completes the proof. \square

Remark 4: For the ETOTC, the minimal intersampling time

$$\Delta t_{\min} = \min_{j \in \mathbb{N}} \{t_{j+1} - t_j\} \tag{63}$$

may be zero, which lead to the notorious Zeno behaviour. On the basis of the result in [43] and [56], the minimal intersampling time Δt_{\min} of the ETOTC has the nonzero positive lower bound. Thus, the Zeno behaviour is avoided.

IV. SIMULATION STUDY

In this section, we provide two numerical examples to show the effectiveness of the adaptive-identifier-critic design based ETOTC method.

A. EXAMPLE 1

Consider a CT linear mass-spring-damper system [51] as

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} x_2 \\ -\frac{\kappa}{m} x_2 - \frac{l}{m} x_1 \end{bmatrix} + \begin{bmatrix} 0 \\ \frac{1}{m} \end{bmatrix} u \tag{64}$$

where x_1, x_2, κ, l and m are the position, velocity damping, stiffness of the spring and mass; The initial value of $[x_1, x_2]^T$ is set to $[-1, 1]^T$; u is the one dimensional control input; The

true parameters are set as $\kappa = 5N/m$, $l = 0.5N \cdot s/m$, and $m = 1kg$. Then, the desired reference trajectories are

$$x_{d1}(t) = \cos(\sqrt{5}t) - \frac{\sqrt{5}}{5} \sin(\sqrt{5}t) \quad (65)$$

$$x_{d2}(t) = -\cos(\sqrt{5}t) - \sqrt{5} \sin(\sqrt{5}t) \quad (66)$$

which are given by using the following command generator dynamics

$$\begin{bmatrix} \dot{x}_{d1} \\ \dot{x}_{d2} \end{bmatrix} = \begin{bmatrix} x_{d2} \\ -5x_{d1} \end{bmatrix}. \quad (67)$$

According to (64) and (67), we construct the augmented system indicated as below

$$\begin{bmatrix} \dot{e}_{d1} \\ \dot{e}_{d2} \\ \dot{x}_{d1} \\ \dot{x}_{d2} \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ -\frac{l}{m} & -\frac{\kappa}{m} & 5 - \frac{l}{m} & -\frac{\kappa}{m} \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -5 & 0 \end{bmatrix} \begin{bmatrix} e_{d1} \\ e_{d2} \\ x_{d1} \\ x_{d2} \end{bmatrix} + \begin{bmatrix} 0 \\ \frac{1}{m} \\ 0 \\ 0 \end{bmatrix} u \quad (68)$$

and the augmented system state vector $[e_{d1}, e_{d2}, x_{d1}, x_{d2}]^T$ is denoted as x_a .

Due to the unavailable knowledges of the system, the identifier NNs are employed and the identifier gains are

$$A_h = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}, B = Q_0 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}. \quad (69)$$

The activation functions $\sigma_f(x)$ and $\sigma_g(x)$ are

$$\sigma_f(x) = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \sigma_g(x) = \begin{bmatrix} 1 \\ 1 \end{bmatrix}. \quad (70)$$

The learning rate matrices of the identifier NNs are

$$P_f = \begin{bmatrix} 20 & 20 \\ 20 & 20 \end{bmatrix}, P_g = \begin{bmatrix} 10 & 10 \\ 10 & 10 \end{bmatrix}. \quad (71)$$

During the identifying process, an excitation signal $u = 70 \sin(10t) \cos(10t)$ is selected to guarantee that the identifier NNs weights converge to the real value. For the ETOTC problem, the parameters in the DPIF are chosen as $\gamma = 0.1$, $Q = \text{diag}(10, 10)$, and $R = 1$. The DPIF from $t = 0$ is

$$V(x_a(0)) = \int_0^\infty e^{-\gamma\tau} \left[x_a^T Q_a x_a + u^T R u \right] d\tau. \quad (72)$$

Then, the initial augmented system state vector is set as $x_a(0) = [-1, 1, -1, 1]^T$ and the local critic NN is constructed to approximate the optimal value function which is defined as

$$\begin{aligned} \hat{V}(x_a) = & \hat{W}_{c1}e_{d1}^2 + \hat{W}_{c2}e_{d2}^2 + \hat{W}_{c3}x_{d1}^2 + \hat{W}_{c4}x_{d2}^2 \\ & + \hat{W}_{c5}e_{d1}e_{d2} + \hat{W}_{c6}e_{d1}x_{d1} + \hat{W}_{c7}e_{d1}x_{d2} \\ & + \hat{W}_{c8}e_{d2}x_{d1} + \hat{W}_{c9}e_{d2}x_{d2} + \hat{W}_{c10}x_{d1}x_{d2}. \end{aligned} \quad (73)$$

For relaxing the condition of the initial stability controller, the auxiliary Lyapunov function is set as $L_a(x_a) = e_{d1}^2 + e_{d2}^2$. Moreover, for obtaining the satisfied simulation results, we set $\eta_1 = 12$, $\eta_2 = 3$, $\xi^2 = 0.5$, and

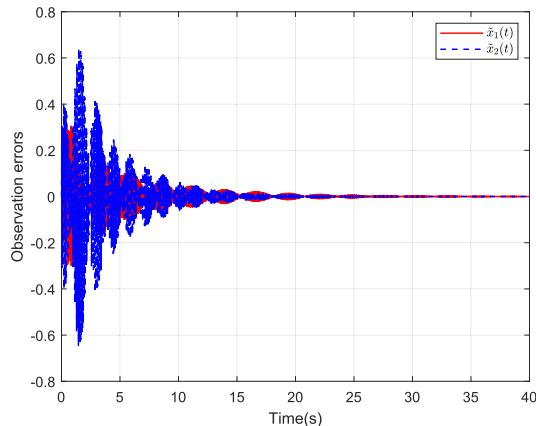


FIGURE 1. Identification errors of the system in example 1.

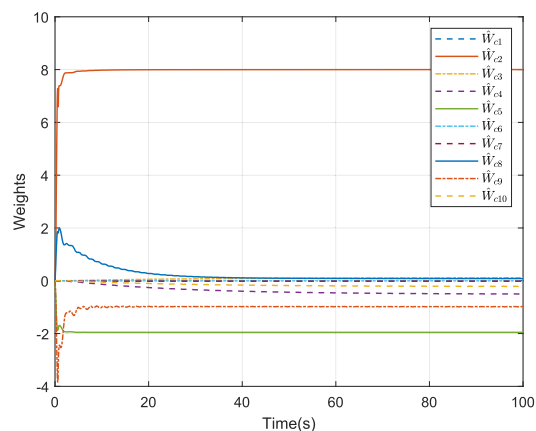


FIGURE 2. Convergence of the critic NN weights in example 1.

$\mathcal{G}_{aM}^2 \mathcal{L}_{\sigma_{cM}}^2 + d_{\sigma_{cM}}^2 \mathcal{L}_{\mathcal{G}_{aM}}^2 = 0.8$. During the training process, $0.3 \exp(-0.1t) \sin(10t) \cos(10t)$ is selected as exploring noise to satisfy the PE condition.

The simulation results are presented in Figs. 1–6. The system identification errors converge to a small region of the equilibrium point in Fig 1. The weight vector converges to $[0.0019, 8.0012, 0.1163, -0.5010, -1.9534, -0.0029, -0.0151, 0.08228, -0.9805, -0.2113]^T$, which is shown in Fig. 2. From Fig. 2, the values of the initial critic weights are all set to zeros, which means that the initial stable control law is unnecessary. Fig. 3 displays the sampling interval under the ET mechanism, which indicates that our developed algorithm avoids the Zeno behaviour.

Using the trained critic NN, we can obtain the approximate ET control law and the closed loop tracking system which are shown in Fig. 4 and 5, separately. On the basis of the ET mechanism, a stairstep graph of the controller is shown in Fig. 4. In Fig. 5, the desired trajectories can be tracked well and the triangles is the ET instants. Fig. 6 presents the evolution of samples. The ET controller only uses 675 samples of the state while the time-triggered controller needs 9047 samples, which greatly saves bandwidth and computational resources (or particularly, 92.54%).

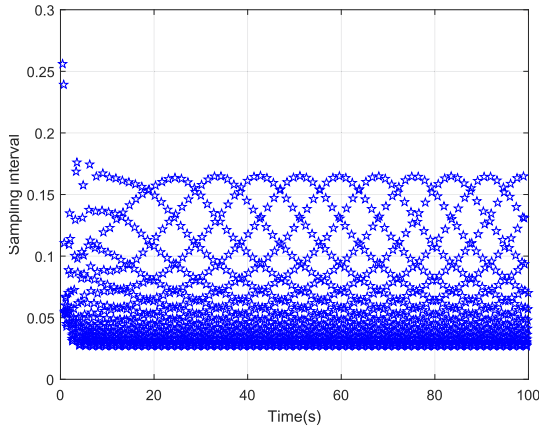


FIGURE 3. Sampling period in the learning process in example 1.

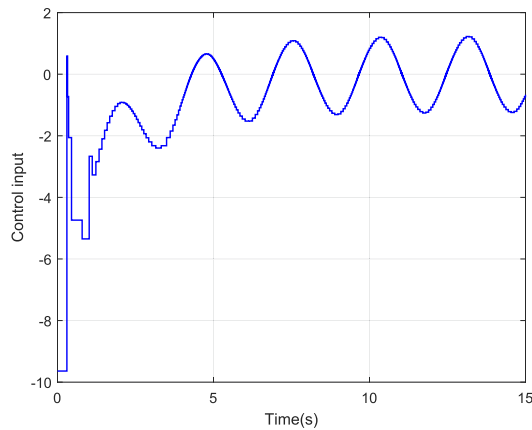


FIGURE 4. Control input for the proposed method in example 1.

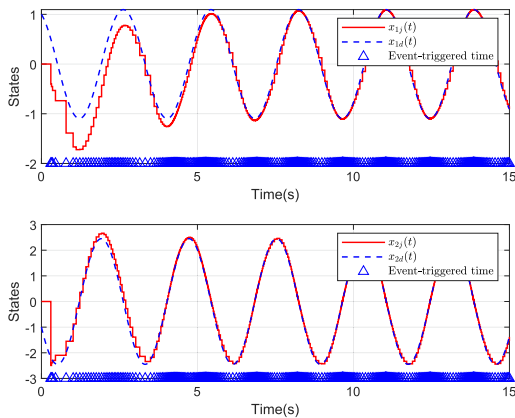


FIGURE 5. Evolution of tracking trajectories in example 1.

B. EXAMPLE 2

Consider a CT mass-spring-damper system with the nonlinearity $k(x) = -x^3$ in [51] as

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} x_2 \\ -x_2^3 - 0.5x_1 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u \quad (74)$$

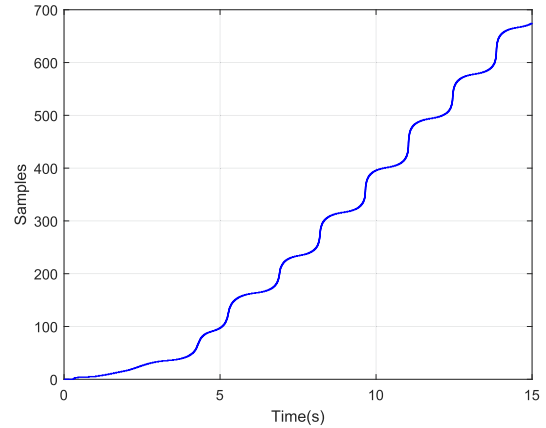


FIGURE 6. Evolution of sample times in example 1.

with the initial state $[-1, 1]^T$, then the Corresponding augmented system is

$$\begin{bmatrix} \dot{e}_{d1} \\ \dot{e}_{d2} \\ \dot{x}_{d1} \\ \dot{x}_{d2} \end{bmatrix} = \begin{bmatrix} e_{d2} \\ -5(e_{d1} + x_{d1})^3 - 0.5e_{d2} + 5x_{d1} - 0.5x_{d2} \\ x_{d2} \\ -5x_{d1} \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} u. \quad (75)$$

On account of the unavailable knowledges of the system, the identifier NNs are employed and the identifier gains are

$$A_h = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}, \quad B = Q_0 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}. \quad (76)$$

The activation functions $\sigma_f(x)$ and $\sigma_g(x)$ are

$$\sigma_f(x) = \begin{bmatrix} x_1^3 \\ x_1 \\ x_2 \end{bmatrix}, \quad \sigma_g(x) = \begin{bmatrix} 1 \\ 1 \end{bmatrix}. \quad (77)$$

The learning rate matrices of the identifier NNs are

$$P_f = \begin{bmatrix} 10 & 10 & 10 \\ 10 & 10 & 10 \\ 10 & 10 & 10 \end{bmatrix}, \quad P_g = \begin{bmatrix} 5 & 5 \\ 5 & 5 \end{bmatrix}. \quad (78)$$

During the identifying process, an excitation signal $u = 60 \sin(10t) \cos(10t)$ is selected to guarantee that the identifier NNs weights converge to the real value.

The parameters in the DPIF are chosen as $x_a = [e_{d1}, e_{d2}, x_{d1}, x_{d2}]^T$ with $x_a(0) = [-1, 1, -1, 1]^T$, $\gamma = 0.1$, $Q = \text{diag}(10, 10)$, and $R = 1$. The local critic NN is constructed to approximate the optimal value function which is defined as

$$\begin{aligned} \hat{V}(x_a) = & \hat{W}_{c1}e_{d1}^2 + \hat{W}_{c2}e_{d1}e_{d2} + \hat{W}_{c3}e_{d1}x_{d1} + \hat{W}_{c4}e_{d1}x_{d2} \\ & + \hat{W}_{c5}e_{d2}^2 + \hat{W}_{c6}e_{d2}x_{d1} + \hat{W}_{c7}e_{d2}x_{d2} + \hat{W}_{c8}x_{d1}^2 \\ & + \hat{W}_{c9}x_{d1}x_{d2} + \hat{W}_{c10}x_{d2}^2 + \hat{W}_{c11}e_{d1}^4 + \hat{W}_{c12}e_{d1}^3e_{d2} \\ & + \hat{W}_{c13}e_{d1}^3x_{d1} + \hat{W}_{c14}e_{d1}^3x_{d2} + \hat{W}_{c15}e_{d1}^2e_{d2}^2 \\ & + \hat{W}_{c16}e_{d1}^2e_{d2}x_{d1} + \hat{W}_{c17}e_{d1}^2e_{d2}x_{d2} + \hat{W}_{c18}e_{d1}^2x_{d1}^2 \\ & + \hat{W}_{c19}e_{d1}^2x_{d1}x_{d2} + \hat{W}_{c20}e_{d1}^2x_{d2}^2 + \hat{W}_{c21}e_{d1}e_{d2}^3 \end{aligned}$$

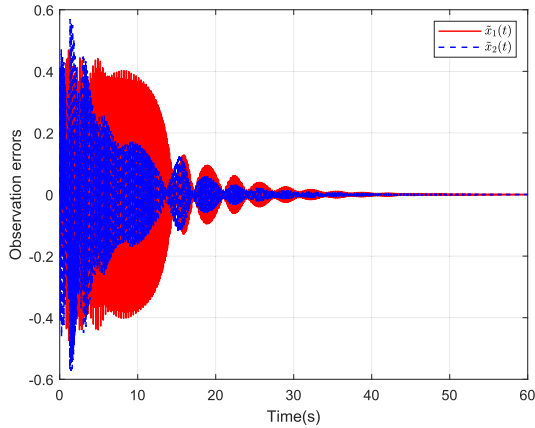


FIGURE 7. Identification errors of the system in example 2.

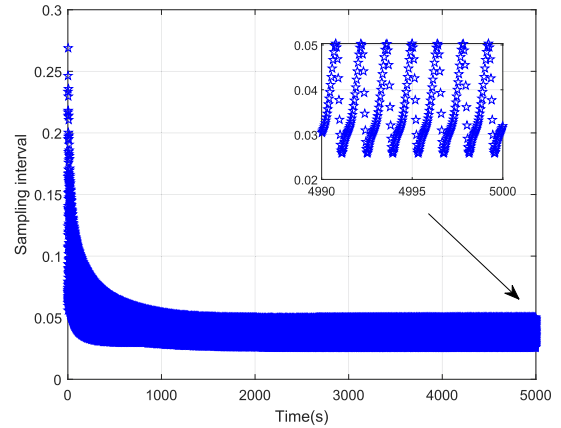


FIGURE 9. Sampling period in the learning process in example 2.

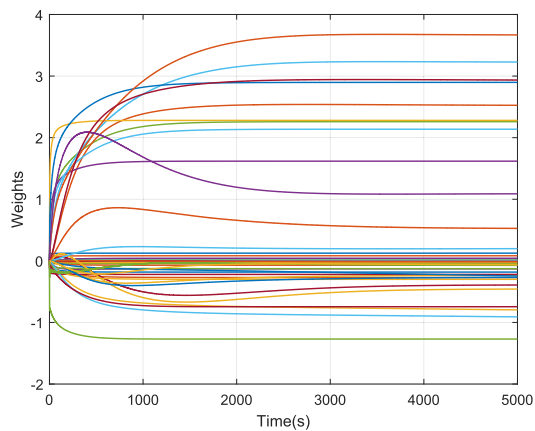


FIGURE 8. Convergence of the critic NN weights in example 2.

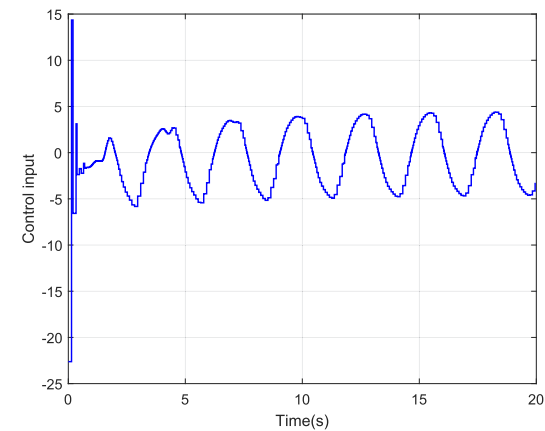


FIGURE 10. Control input for the proposed method in example 2.

$$\begin{aligned}
 & + \hat{W}_{c22}e_{d1}e_{d2}^2x_{d1} + \hat{W}_{c23}e_{d1}e_{d2}^2x_{d2} \\
 & + \hat{W}_{c24}e_{d1}e_{d2}x_{d1}^2 + \hat{W}_{c25}e_{d1}e_{d2}x_{d1}x_{d2} \\
 & + \hat{W}_{c26}e_{d1}e_{d2}x_{d2}^2 + \hat{W}_{c27}e_{d1}x_{d1}^3 + \hat{W}_{c28}e_{d1}x_{d1}^2x_{d2} \\
 & + \hat{W}_{c29}e_{d1}x_{d1}x_{d2}^2 + \hat{W}_{c30}e_{d1}x_{d2}^3 + \hat{W}_{c31}e_{d1}^4x_{d1} \\
 & + \hat{W}_{c32}e_{d2}^3x_{d1} + \hat{W}_{c33}e_{d2}^3x_{d2} + \hat{W}_{c34}e_{d2}^2x_{d1}^2 \\
 & + \hat{W}_{c35}e_{d2}^2x_{d1}x_{d2} + \hat{W}_{c36}e_{d2}^2x_{d2}^2 + \hat{W}_{c37}e_{d2}x_{d1}^3 \\
 & + \hat{W}_{c38}e_{d2}x_{d1}^2x_{d2} + \hat{W}_{c39}e_{d2}x_{d1}x_{d2}^2 + \hat{W}_{c40}e_{d2}x_{d1}^3 \\
 & + \hat{W}_{c41}x_{d1}^4 + \hat{W}_{c42}x_{d1}^3x_{d2} + \hat{W}_{c43}x_{d1}^2x_{d2}^2 \\
 & + \hat{W}_{c44}x_{d1}x_{d2}^3 + \hat{W}_{c45}x_{d2}^4. \tag{79}
 \end{aligned}$$

Similarly in order to relax the condition of the initial stability controller, the auxiliary Lyapunov function is set as $L_a(x_a) = e_{1d}^4 + 0.5e_{2d}^2 + x_{d1}^4 + x_{d2}^4$. Moreover, for obtaining the satisfied simulation results, we set $\eta_1 = 10$, $\eta_2 = 0.1$, $\xi^2 = 0.5$, and $\mathcal{G}_{aM}^2 \mathcal{L}_{\sigma_{cM}}^2 + d_{\sigma_{cM}}^2 \mathcal{L}_{\mathcal{G}_{aM}}^2 = 15$. During the training process, $0.3 \exp(-0.1t) \sin(10t) \cos(10t)$ is selected as exploring noise to satisfy the PE condition.

The simulation results are presented in Figs. 7–12. The system identification errors converge to a small region of the equilibrium point in Fig 7. Due to the nonlinearity of the mass-spring-damper system, we use the activation

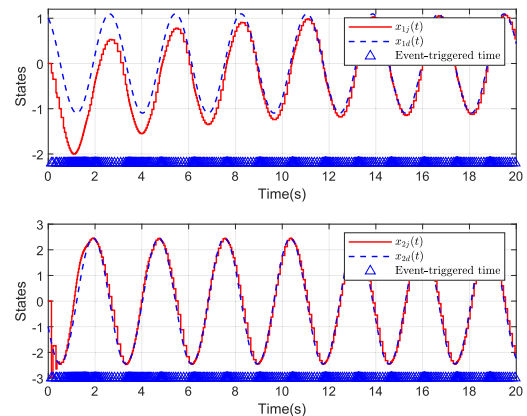


FIGURE 11. Evolution of tracking trajectories in example 2.

function of 45 neurons. The weight vector converges to $[0.0355, -0.2710, -0.7961, -0.0101, 2.2610, 3.2280, -0.3924, 0.0008, 2.5255, -0.2337, -6.7478 \times 10^{-6}, -0.0460, -0.0033, -0.0011, 0.1247, 0.0823, -0.0280, 0.0417, 0.0122, 0.0049, -0.2205, -0.1825, -0.0712, -0.3003, -0.1297, -0.1272, -0.9095, -0.0102, -0.2464, -0.0103, 2.2806, 1.6184, -1.2697, 2.1372, -0.7450,$

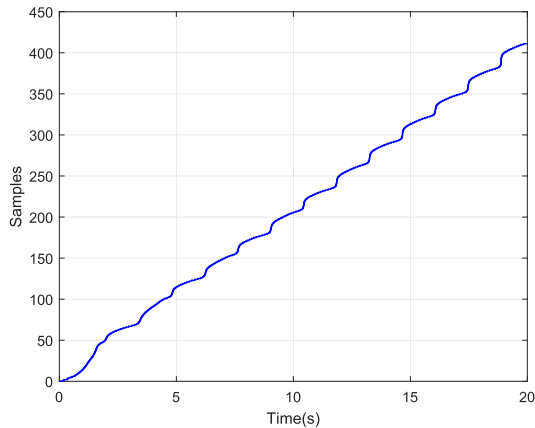


FIGURE 12. Evolution of sample times in example 2.

2.8976, 3.6685, -0.4582 , 1.0886, -0.0738 , 0.1980, 2.9358, -0.2684 , 0.5267, -0.0629 ^T, which is shown in Fig. 8. The sampling interval under the ET mechanism is displayed in Fig. 9, which indicates that our developed algorithm avoids the Zeno behaviour.

Using the trained critic NN, we can obtain the approximate ET control law and the closed loop tracking system which are shown in Fig. 10 and 11, separately. On the basis of the ET mechanism, a staircase graph of the controller is shown in Fig. 10. In Fig. 11, the desired trajectories can be tracked well and the triangles is the ET instants. Fig. 12 presents the evolution of samples. The ET controller only uses 412 samples of the state while the time-triggered controller needs 9679 samples, which greatly saves bandwidth and computational resources (or particularly, 95.74%).

V. CONCLUSION

In this paper, for the tracking control of the unknown continuous-time systems, an ET ADP-based method is developed. To handle the unknown systems, the identifier NNs are employed to learn the system dynamics. Then, the tracking control problem is transformed into the regulation problem of the identified augmented system. Only a critic NN is derived to reconstruct the DPIF and the actor is update by the ET condition. A novel online weight tuning law without requiring the initial admissible control is designed. Finally, the simulation results of the mass-spring-damper system show the effectiveness of the developed method. The deficiency of this method and the future works are discussed as follows.

1) The unknown systems studied in this paper are widely concerned. In these researches, the identifier process and the critic design are independent each other. In our future work, we will develop a new online learning method which the identification and the control are carried out simultaneously.

2) The actuator saturation is a common phenomenon in control systems, which reduces the system dynamic performance and induces the system instability. The research on the constrained systems will be one of our future work.

3) The critic NN structure used in this paper is simple and has limited expression ability. In our future work, the deep learning method will be employed, which can decrease the fitting error effectively.

REFERENCES

- [1] 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.
- [2] D. Liu, Q. Wei, D. Wang, X. Yang, and H. Li, *Adaptive Dynamic Programming with Applications in Optimal Control*. Cham, Switzerland: Springer, 2017.
- [3] R. Bellman, "Dynamic programming," *Science*, vol. 153, nos. 37–31, pp. 34–37, 1966.
- [4] Q. Wei, D. Liu, and H. Lin, "Value iteration adaptive dynamic programming for optimal control of discrete-time nonlinear systems," *IEEE Trans. Cybern.*, vol. 46, no. 3, pp. 840–853, Mar. 2016.
- [5] Q. Wei, B. Li, and R. Song, "Discrete-time stable generalized self-learning optimal control with approximation errors," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 29, no. 4, pp. 1226–1238, Apr. 2018.
- [6] Q. Wei, Z. Liao, Z. Yang, B. Li, and D. Liu, "Continuous-time time-varying policy iteration," *IEEE Trans. Cybern.*, vol. 50, no. 12, pp. 4958–4971, Dec. 2020.
- [7] Q. Wei, R. Song, Z. Liao, B. Li, and F. L. Lewis, "Discrete-time impulsive adaptive dynamic programming," *IEEE Trans. Cybern.*, vol. 50, no. 10, pp. 4293–4306, Oct. 2020.
- [8] X. Yang and H. He, "Event-driven H_∞ -constrained control using adaptive critic learning," *IEEE Trans. Cybern.*, vol. 51, no. 10, pp. 4860–4872, Oct. 2021.
- [9] H. Zhang, J. Duan, Y. Wang, and Z. Gao, "Bipartite fixed-time output consensus of heterogeneous linear multiagent systems," *IEEE Trans. Cybern.*, vol. 51, no. 2, pp. 548–557, Feb. 2021.
- [10] J. Lu, Q. Wei, and F.-Y. Wang, "Parallel control for optimal tracking via adaptive dynamic programming," *IEEE/CAA J. Automatica Sinica*, vol. 7, no. 6, pp. 1662–1674, Nov. 2020.
- [11] Q. Wei, H. Li, and F.-Y. Wang, "Parallel control for continuous-time linear systems: A case study," *IEEE/CAA J. Automatica Sinica*, vol. 7, no. 4, pp. 919–926, Jul. 2020.
- [12] Y. Zhang, Q. Li, W. Zhang, and Y. Yang, "Survey of multi-model adaptive control theory and its applications," *Chin. J. Eng.*, vol. 42, no. 2, pp. 135–143, Feb. 2020.
- [13] P. J. Werbos, "Approximate dynamic programming for real-time control and neural modeling," in *Handbook of Intelligent Control*, D. A. White and D. A. Sofge, Eds. New York, NY, USA: Van Nostrand Reinhold, 1992.
- [14] D. V. Prokhorov and D. C. Wunsch, "Adaptive critic designs," *IEEE Trans. Neural Netw.*, vol. 8, no. 5, pp. 997–1007, Sep. 1997.
- [15] D. P. Bertsekas and J. N. Tsitsiklis, *Neuro-Dynamic Programming*. Belmont, MA, USA: Athena Scientific, 1996.
- [16] B. Lincoln and A. Rantzer, "Relaxing dynamic programming," *IEEE Trans. Autom. Control*, vol. 51, no. 8, pp. 1249–1260, Aug. 2006.
- [17] L. P. Kaelbling, M. L. Littman, and A. W. Moore, "Reinforcement learning: A survey," *J. Artif. Intell. Res.*, vol. 41, no. 4, pp. 237–285, May 1996.
- [18] D. V. Prokhorov, R. A. Santiago, and D. C. Wunsch, "Adaptive critic designs: A case study for neurocontrol," *Neural Netw.*, vol. 8, no. 9, pp. 1367–1372, 1995.
- [19] D. Wang and D. Liu, "Neuro-optimal control for a class of unknown nonlinear dynamic systems using SN-DHP technique," *Neurocomputing*, vol. 121, pp. 218–225, Dec. 2013.
- [20] H. Zhang, J. Zhang, G.-H. Yang, and Y. Luo, "Leader-based optimal coordination control for the consensus problem of multiagent differential games via fuzzy adaptive dynamic programming," *IEEE Trans. Fuzzy Syst.*, vol. 23, no. 1, pp. 152–163, Feb. 2015.
- [21] P. Yan, D. Wang, H. Li, and D. Liu, "Error bound analysis of Q -function for discounted optimal control problems with policy iteration," *IEEE Trans. Syst., Man, Cybern. Syst.*, vol. 47, no. 7, pp. 1207–1216, Jul. 2017.
- [22] 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.
- [23] 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.

- [24] B. Luo, H.-N. Wu, T. Huang, and D. Liu, "Data-based approximate policy iteration for affine nonlinear continuous-time optimal control design," *Automatica*, vol. 50, no. 12, pp. 3281–3290, 2014.
- [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] D. P. Bertsekas, "Value and policy iterations in optimal control and adaptive dynamic programming," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 28, no. 3, pp. 500–509, Mar. 2017.
- [27] X. Yang, H. He, and X. Zhong, "Approximate dynamic programming for nonlinear-constrained optimizations," *IEEE Trans. Cybern.*, vol. 51, no. 5, pp. 2419–2432, May 2021.
- [28] H. Zhang, G. Xiao, Y. Liu, and L. Liu, "Value iteration-based H_∞ controller design for continuous-time nonlinear systems subject to input constraints," *IEEE Trans. Syst., Man, Cybern. Syst.*, vol. 50, no. 11, pp. 3986–3995, Nov. 2020.
- [29] Q. Wei, H. Li, X. Yang, and H. He, "Continuous-time distributed policy iteration for multicontroller nonlinear systems," *IEEE Trans. Cybern.*, vol. 51, no. 5, pp. 2372–2383, May 2021.
- [30] A. Anta and P. Tabuada, "To sample or not to sample: Self-triggered control for nonlinear systems," *IEEE Trans. Autom. Control*, vol. 55, no. 9, pp. 2030–2042, Sep. 2010.
- [31] E. Garcia and P. J. Antsaklis, "Model-based event-triggered control for systems with quantization and time-varying network delays," *IEEE Trans. Autom. Control*, vol. 58, no. 2, pp. 422–434, Feb. 2013.
- [32] J. Almeida, C. Silvestre, and A. Pascoal, "Synchronization of multiagent systems using event-triggered and self-triggered broadcasts," *IEEE Trans. Autom. Control*, vol. 62, no. 9, pp. 4741–4746, Sep. 2017.
- [33] K. Masako, "Event-triggered control with self-triggered sampling for discrete-time uncertain systems," *IEEE Trans. Autom. Control*, vol. 64, no. 3, pp. 1273–1279, Mar. 2019.
- [34] W. P. M. H. Heemels, M. C. F. Donkers, and A. R. Teel, "Periodic event-triggered control for linear systems," *IEEE Trans. Autom. Control*, vol. 58, no. 4, pp. 847–861, Apr. 2013.
- [35] M. Mazo, Jr., and P. Tabuada, "Decentralized event-triggered control over wireless sensor/actuator networks," *IEEE Trans. Autom. Control*, vol. 56, no. 10, pp. 2456–2461, Oct. 2011.
- [36] H. Li, Z. Chen, L. Wu, H.-K. Lam, and H. Du, "Event-triggered fault detection of nonlinear networked systems," *IEEE Trans. Cybern.*, vol. 47, no. 4, pp. 1041–1052, Apr. 2017.
- [37] A. Mustafa, N. K. Dhar, and N. K. Verma, "Event-triggered sliding mode control for trajectory tracking of nonlinear systems," *IEEE/CAA J. Automatica Sinica*, vol. 7, no. 1, pp. 307–314, Jan. 2020.
- [38] A. Eqtami, D. V. Dimarogonas, and K. J. Kyriakopoulos, "Event-triggered control for discrete-time systems," in *Proc. Amer. Control Conf.*, Baltimore, MD, USA, Jul. 2010, pp. 4719–4724.
- [39] X. Wang and M. D. Lemmon, "Event design in event-triggered feedback control systems," in *Proc. 47th IEEE Conf. Decis. Control*, Cancun, Mexico, Dec. 2008, pp. 2105–2110.
- [40] S. Li and B. Xu, "Co-design of event generator and controller for event-triggered control system," in *Proc. 30th Chin. Control Conf.*, Yantai, China, Jul. 2011, pp. 175–179.
- [41] X. Zhong and H. He, "An event-triggered ADP control approach for continuous-time system with unknown internal states," *IEEE Trans. Cybern.*, vol. 47, no. 3, pp. 683–694, Mar. 2017.
- [42] L. Dong, X. Zhong, C. Sun, and H. He, "Adaptive event-triggered control based on heuristic dynamic programming for nonlinear discrete-time systems," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 28, no. 7, pp. 1594–1605, Jul. 2017.
- [43] D. Wang, H. He, X. Zhong, and D. Liu, "Event-driven nonlinear discounted optimal regulation involving a power system application," *IEEE Trans. Ind. Electron.*, vol. 64, no. 10, pp. 8177–8186, Oct. 2017.
- [44] X. Yang and H. He, "Adaptive critic learning and experience replay for decentralized event-triggered control of nonlinear interconnected systems," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 50, no. 11, pp. 4043–4055, Nov. 2020.
- [45] B. Luo, Y. Yang, D. Liu, and H. Wu, "Event-triggered optimal control with performance guarantees using adaptive dynamic programming," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 31, no. 1, pp. 76–88, Jan. 2020.
- [46] X. Shan, L. Biao, and L. Derong, "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.
- [47] K. Hornik, M. Stinchcombe, and H. White, "Universal approximation of an unknown mapping and its derivatives using multilayer feedforward networks," *Neural Netw.*, vol. 3, no. 5, pp. 551–560, 1990.
- [48] T. Dierks and S. Jagannathan, "Optimal control of affine nonlinear continuous-time systems," in *Proc. Amer. Control Conf.*, Baltimore, MD, USA, Jun. 2010, pp. 1568–1573.
- [49] H. Zargazadeh, T. Dierks, and S. Jagannathan, "Optimal control of nonlinear continuous-time systems in strict-feedback form," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 26, no. 10, pp. 2535–2549, Oct. 2015.
- [50] J. Hou, D. Wang, D. Liu, and Y. Zhang, "Model-free H_∞ optimal tracking control of constrained nonlinear systems via an iterative adaptive learning algorithm," *IEEE Trans. Syst., Man, Cybern. Syst.*, vol. 50, no. 11, pp. 4097–4108, Nov. 2020.
- [51] H. Modares and F. L. Lewis, "Optimal tracking control of nonlinear partially-unknown constrained-input systems using integral reinforcement learning," *Automatica*, vol. 50, no. 7, pp. 1780–1792, Jul. 2014.
- [52] B. Kiumarsi, F. L. Lewis, H. Modares, A. Karimpour, and M. B. Naghibi-Sistani, "Reinforcement Q-learning for optimal tracking control of linear discrete-time systems with unknown dynamics," *Automatica*, vol. 50, no. 4, pp. 1167–1175, Apr. 2014.
- [53] D. Wang, C. Mu, H. He, and D. Liu, "Event-driven adaptive robust control of nonlinear systems with uncertainties through NDP strategy," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 47, no. 7, pp. 1358–1370, Jul. 2017.
- [54] K. G. Vamvoudakis and F. L. Lewis, "Online actor-critic algorithm to solve the continuous-time infinite horizon optimal control problem," *Automatica*, vol. 46, no. 5, pp. 878–888, 2010.
- [55] K. G. Vamvoudakis, A. Mojtodi, and H. Ferraz, "Event-triggered optimal tracking control of nonlinear systems," *Int. J. Robust Nonlinear Control*, vol. 27, no. 4, pp. 598–619, Mar. 2017.
- [56] B. Zhao and D. 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.



KUNFU WANG received the B.S. and M.S. degrees from Henan Polytechnic University, Jiaozuo, China, in 2011 and 2014, respectively. He is currently pursuing the Ph.D. degree in computer information technology with the Nanjing University of Information Science and Technology, Nanjing, China. He is currently a Senior Engineer with the System Engineering Research Institute, China State Shipbuilding Corporation, Beijing, China. His research interests include optimal control, deep reinforcement learning, and unmanned intelligence.



QIJIA GU received the B.S. degree in electronic information science and technology from Peking University, Beijing, China, in 2006. He is currently working at China State Shipbuilding Corporation, Beijing. His current research interests include integrated electronic information systems and systems engineering.



BAIQIAO HUANG received the Ph.D. degree in system engineering from Beihang University, Beijing, China. He is currently a Senior Engineer with the System Engineering Research Institute, China State Shipbuilding Corporation, Beijing. His research interests include artificial intelligence, unmanned intelligence, and system engineering.



QINGLAI WEI (Senior Member, IEEE) received the B.S. degree in automation and the Ph.D. degree in control theory and control engineering, from Northeastern University, Shenyang, China, in 2002 and 2009, respectively. From 2009 to 2011, he was a Postdoctoral Fellow with the State Key Laboratory for Management and Control of Complex Systems, Institute of Automation, Chinese Academy of Sciences, Beijing, China. He is currently a Professor with the Institute and the Associate Director of the Laboratory. He has authored four books, and published over 80 international journal articles. His research interests include adaptive dynamic programming, neural-networks-based control, optimal control, nonlinear systems, and their industrial applications.

Dr. Wei has been the Secretary of IEEE Computational Intelligence Society (CIS) Beijing Chapter, since 2015. He was a recipient of the IEEE/CAA JOURNAL OF AUTOMATICA SINICA Best Paper Award, the IEEE System, Man, and Cybernetics Society Andrew P. Sage Best Transactions Paper Award, the IEEE TRANSACTIONS ON NEURAL NETWORKS AND LEARNING SYSTEMS Outstanding Paper Award, the Outstanding Paper Award of *Acta Automatica Sinica*, the IEEE 6th Data Driven Control and Learning Systems Conference (DDCLS2017) Best Paper Award, and the Zhang Siying Outstanding Paper Award of Chinese Control and Decision Conference (CCDC). He was a recipient of Shuang-Chuang Talents in Jiangsu Province, China, Young Researcher Award of Asia Pacific Neural Network Society (APNNS), Young Scientist Award and Yang Jiachi Tech Award of Chinese Association of Automation (CAA). He is a Board of Governors (BOG) member of the International Neural Network Society (INNS) and a Council Member of CAA. He was the guest editor for several international journals.



TIANMIN ZHOU received the B.S. degree in management from Tongji University, Shanghai, China, in 2013, and the Ph.D. degree in systems science from Beijing Normal University, Beijing, China, in 2020. He is currently a Postdoctoral Fellow with the State Key Laboratory for Management and Control of Complex Systems, Institute of Automation, Chinese Academy of Sciences, Beijing. His research interests include adaptive dynamic programming, stochastic control, deep reinforcement learning, and artificial intelligence-based control.

Dr. Zhou is the Secretary-General of Technical Committee on Parallel Control and Management (TCPCM), Chinese Association of Automation (CAA). He serves as a Peer Reviewer for the IEEE/CAA JOURNAL OF AUTOMATION SINICA, IEEE TRANSACTIONS ON ARTIFICIAL INTELLIGENCE, and *Acta Automatica Sinica*, with a good reputation.

• • •