



Event-triggered self-learning-based tracking control for nonlinear constrained-input systems with uncertain disturbances

Binbin Peng¹ · Xiaohong Cui^{1,2} · Kun Zhou^{1,2}

Received: 14 June 2023 / Accepted: 15 January 2024 / Published online: 25 February 2024
© The Author(s), under exclusive licence to Springer-Verlag London Ltd., part of Springer Nature 2024

Abstract

In this paper, an online event-triggered self-learning scheme based on adaptive dynamic programming (ADP) is developed to address tracking control design for nonlinear systems with constrained input and uncertain disturbance. Firstly, the value function with non-quadratic function is defined for the augmented nominal system, and the constrained robust tracking problem is equivalent to the optimal control for solving the tracking event-triggered Hamilton–Jacobi–Bellman (ETHJB) equation. Then, a single-critic network is developed to obtain the value function and control law related to the solution of the tracking ETHJB equation, greatly reducing approximation errors and computational costs. To alleviate the requirement for the entire state sampling, we propose a triggering rule that ensures system stability while limiting control updates. Theoretical proof demonstrates that the tracking state of the closed-loop system and the weight approximation error of the neural network are uniformly ultimately bounded (UUB). Finally, two examples are provided to validate the availability of the proposed scheme.

Keywords Adaptive dynamic programming (ADP) · Event-triggered control (ETC) · Integral reinforcement learning (IRL) · Tracking control · Neural network (NN)

1 Introduction

In practical engineering, the unknown environment and the uncertain model commonly affect nonlinear systems greatly. Therefore, it is necessary to design a suitable controller to fulfill control quality for the plant with uncertainty. Robust control not only guarantees the robust stability of the closed-loop nonlinear systems, but also optimizes some performance indicators, it has been widely recognized by scholars [1–4]. There are two methods available for solving robust control problems. The first strategy is to design a nominal system to convert the robust nonlinear control problem into the optimal control problem by constructing a nominal system. The solution of the

optimal control problem is equivalent to the solution of the original robust control problem [5, 6]. The other method is the H_∞ control method, which transforms the robust problem into an optimal control problem involving two non-cooperative players in a game. Then, the robust problem is equivalent to obtaining the solution of the Hamilton–Jacobi–Bellman (HJB) equation in optimal control [7–9]. We employ the first method to derive the robust control law without solving the H_∞ optimal control problem, avoiding the difficulties of determining whether the saddle point of two-player exists in H_∞ control problems. Additionally, [10] solves optimal control problems offline based on neural networks. In practice, solving the robust control problem online remains difficult.

Considering that the HJB equation suffers from the curse of dimensionality, direct solutions are nearly impossible. As a self-learning method, the introduction and application of adaptive dynamic programming (ADP) [11, 12] have been widely recognized in tracking problems [13], disturbance attenuation problems [14], and robust control problems. ADP, reinforcement learning (RL) [15], and adaptive critical learning (ACL) [16, 17] are

✉ Xiaohong Cui
18A0103159@cjlu.edu.cn

¹ College of Mechanical and Electrical Engineering, China Jiliang University, Hangzhou 310000, Zhejiang, China

² Key Laboratory of Intelligent Manufacturing Quality Big Data Tracing and Analysis of Zhejiang Province, China Jiliang University, Hangzhou 310018, Zhejiang, China

considered analogous characteristics. One of the advantages of ADP is that it uses function approximators, usually neural networks (NNs), to estimate the ideal value function related to the system performance. In [18], an ADP-based robust controller is designed to stabilize the system and efficiently reduce the effects of perturbations on system functionality. However, it should be noted that the majority of current ADP-based robust control systems merely address the regulation problem [19, 20]. In practice, not only does the designed controller need to satisfy the robustness of the plant, but it also requires the plant to track an anticipated trajectory, especially in noisy and uncertain environments. This results in a robust tracking problem, in which the goal is to track an expected trajectory with a given value function in the presence of modeling unknowns. [21] considers using policy iteration (PI) methods to solve nonlinear tracking control systems with uncertain dynamics. PI technique based on data-driven strategy is applied to deal with the tracking problem for nonlinear systems in [22]. However, the PI algorithm requires a more accurate model and a larger estimate error obtained by the neural network approximator. In [23], a PI-based integral reinforcement learning (IRL) is used to solve the tracking HJB equation of nonlinear systems. By adding integral operations and recording reinforcement signals from different time intervals, the weight values of NNs are updated, which fully utilizes the collected data and avoids dependence on known system information. By collecting the data of the integral interval, integral reinforcement learning technology is also applied in [24] to update the weight values of the critic-actor network and solve the optimal tracking control problem online. However, the presence of uncertain dynamics and external disturbances is not considered, and the designed multiple NNs approximator may increase the computational complexity of the system in previous studies. Therefore, the robust tracking control law is obtained by a single-critic network approximation in this paper.

When designing a tracking controller, we need to consider not only the impact of disturbances on system stability, but also the safety and physical characteristics of the controller (or actuator) [25]. One solution is to install constraints on the controller (or actuator). Currently, in theoretical research with input constraints, it is common to design a non-quadratic function for a system, where the control input is guaranteed to have a certain bound. Symmetrical input constraints are the most commonly considered constraints in [26–28], and numerous methods have been advocated to deal with them. However, many nonlinear plants in reality are subject to asymmetric input constraints, which adds challenges to the design of tracking controllers. Therefore, addressing asymmetric input constraints is also a consideration factor in this paper.

The ADP-based method typically relies on transferring data periodically with a fixed sampling period or time-triggered control (TTC). However, this may result in a large amount of transmitted data and increase the computational and storage burden of the control system [29–32]. To address this issue, event-triggered control (ETC) has been introduced as an alternative to TTC in [33, 34]. ETC executes non-periodically, which has better performance than TTC in terms of reducing the computational burden. In [35], an event-triggered ADP (ETADP) scheme is proposed to address HJB equations related to zero-sum game problems for nonlinear systems, which significantly reduces computational costs. In [36], an event-triggered network-based algorithm is utilized for input-constrained nonlinear systems with external disturbances to reduce unnecessary controller updates. In [37], for a class of nonlinear continuous-time systems with external disturbances, a network-based ETC control method is proposed, which can ensure the stability of the system and reduce the number of controller updates. [38] provides an ETC-based network learning method, which uses the optimal distributed control method to deal with approximately interconnected nonlinear systems, thereby reducing computational costs and communication waste. Similarly, a framework of the identified-critic network is constructed to address the ETC-based decentralized for interconnected nonlinear systems, which greatly saves communication resources for each subsystem in [39]. The commonality of the above studies is that ETC is well applied to optimal regulation problems, but the optimal robust tracking problem is not considered. To address this factor and reduce computational complexity, we propose an approximate optimal robust tracking control scheme for uncertain nonlinear systems with input constraints using the event-triggered self-learning method. Furthermore, a triggering condition is designed to ensure control system stability while updating the event-based robust control law. The main contributions of this paper are summarized as follows:

1. Different from [8, 11] to solve the H_∞ control problem, we overcome the difficulty of predetermining the existence of Nash equilibria points in non-cooperative games. Moreover, we propose the method realized by defining an augmented nominal system and only solving a constrained event-based optimal control laws.
2. Rather than considering nonlinear systems with symmetric input constraints, we construct cost functions with a discount factor and non-quadratic functions to solve robust optimal control problems with asymmetric input constraints.
3. An online neural network approximator is constructed based on an event-triggered self-learning framework to

avoid dependence on complete dynamics. Then, the weights of a single-critic network are updated using continuous integration interval information to obtain the value function and control policy related to the ETHJB equation. Based on an event-based mechanism, the designed triggering rule limits control updates, which reduces the computational cost and storage burden.

The remainder of this paper is structured as follows. The problem is described and transformed in Sect. 2. Section 3 is the introduction and implementation of the event-triggered adaptive dynamic programming scheme. Section 4 provides theoretical proof of system stability. Section 5 gives the simulation examples of the bionic joint model and the nonlinear system, and we provide the conclusion in Sect. 6.

Notation. Some parameters are strictly defined in this paper. $\mathbb{R}, \mathbb{R}^m,$ and $\mathbb{R}^{m \times n}$ are denoted the set of real matrices, and m and $m \times n$ are the correlative dimension matrix. $\underline{\lambda}(\cdot)$ is the minimum eigenvalue of the matrix. $\|\cdot\|$ is expressed as the 2-norm. $\nabla V \triangleq \partial V / \partial \xi$ is the derivative of $V(\xi)$ with respect to ξ .

2 Problem statement

The continuous-time dynamic system with uncertain disturbances is considered as

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

where $x \in \mathbb{R}^n$ is the state vector, $f(x) \in \mathbb{R}^n$ is the unknown internal dynamics, and $g(x) \in \mathbb{R}^{n \times m}$ is the input dynamics. $u(t) \in \mathcal{U} \subset \mathbb{R}^m$ is the input vector with asymmetric bounds, which is expressed as $\mathcal{U} = \{u_1, u_2, \dots, u_m\} : u_{min} \leq u_i \leq u_{max}, i = 1, 2, \dots, m, |u_{min}| \neq |u_{max}|\}$. $\bar{d}(x)$ is an uncertain disturbance with $\bar{d}(0) = 0$. Let $x(0) = x_0$ as the initial state of the system. Assuming that the uncertain item satisfies $\bar{d}(x)^T \bar{d}(x) \leq \bar{d}_m(x)^T \bar{d}_m(x)$ with $\bar{d}_m(0) = 0$, where $\bar{d}_m(x)$ is a known function.

For the tracking problem, a bounded reference signal $x_r(t)$ is given, and there exists a Lipschitz continuous command generator $\mathcal{G}(x_r(t))$ satisfying

$$\dot{x}_r(t) = \mathcal{G}(x_r(t)), \quad \mathcal{G}(0) = 0 \tag{2}$$

Then, the tracking error dynamic equation and its derivatives are expressed as

$$e_r(t) = x(t) - x_r(t) \tag{3}$$

$$\dot{e}_r(t) = \dot{x}(t) - \dot{x}_r(t) \tag{4}$$

Define the augmented matrix with tracking error and

expected trajectory as $\xi(t) = [e_r^T, x_r^T]^T \in \mathbb{R}^{2n}$. Then, the dynamic system after dimension is converted to

$$\dot{\xi} = F(\xi) + G(\xi)u + G(\xi)\bar{D}(\xi) \tag{5}$$

where $F(\xi) = \begin{bmatrix} f(e_r + x_r) - \mathcal{G}(x_r) \\ \mathcal{G}(x_r) \end{bmatrix}$ and $G(\xi) = \begin{bmatrix} g(e_r + x_r) \\ 0 \end{bmatrix}$. The uncertain disturbance after

augmentation satisfies with $\|\bar{D}(\xi)\| = \|\bar{D}(e_r + x_r)\| = \|\bar{D}(x)\| \leq \bar{d}_m(x) = \bar{d}_m(e_r + x_r) \triangleq \bar{d}_m(\xi)$.

Assumption 1 [16, 17] The unknown dynamic $F(\xi)$ is Lipschitz continuous with $f(0) = 0$. $G(\xi)$ and x_r are bounded, that is, $\|G(\xi)\| \leq \mathcal{L}_G$ and $\|x_r\| \leq b_m$ with $\mathcal{L}_G > 0$ and $b_m > 0$.

The existence of unknown disturbances may affect the stability of the tracking system (5). According to [5], we transform the robust tracking control problem into the optimal control problem for solving the HJB equation by defining the following nominal system

$$\dot{\xi} = F(\xi(t)) + G(\xi(t))u(t) \tag{6}$$

Assuming that the system (6) is controllable for any admissible input law u on the compact set Ω . Let $\mathcal{X}(\xi, u)$ be the basic utility function of optimal control and $\bar{d}_m(\xi) > 0$ be the additional utility function related to uncertain disturbances. To analyze the performance of the system (6), the following value function with the discount factor $\mu > 0$ is defined as

$$V(\xi) = \int_t^\infty e^{-\mu(\tau-t)} (\mathcal{X}(\xi(\tau), u(\tau)) + \|\bar{d}_m(\xi)\|^2) d\tau \tag{7}$$

where $\mathcal{X}(\xi, u) = \xi^T Q \xi + \mathcal{R}(u)$. $\mathcal{R}(u)$ is a non-quadratic function subject to asymmetrically constrained input, which is constructed as

$$\mathcal{R}(u) = 2\alpha \sum_{i=1}^m \int_{\beta_0}^u \Theta^{-1}\left(\frac{\eta - \beta_0}{\alpha}\right) R d\eta \tag{8}$$

where $\alpha = (u_{max} - u_{min})/2, \beta_0 = (u_{max} + u_{min})/2,$ and $\Theta^{-1}(\cdot)$ is an odd monotonic function with $\Theta^{-1}(0) = 0$. To get results, we use $\Theta(\cdot) = \tanh(\cdot) = (e^x + e^{-x})^{-1} (e^x - e^{-x})$. $Q = \begin{bmatrix} Q_0 & 0 \\ 0 & 1 \end{bmatrix}, Q_0 > 0$ and $R > 0$.

Remark 1 Due to the influence of reference signals and uncertain disturbances in the augmented state for the tracking system, the control term is usually nonzero when the state arrives at the equilibrium point $\xi = 0$. Thus, the involved decay term $e^{-\mu(\tau-t)}$ is introduced in the integral term in (7). If we let $\mu = 0$, the right side of (7) may be

unbounded, and $V(\xi)$ may be divergent. That is why we consider using the discount factor μ .

For any admissible input law $u(\xi) \in \Omega$, $V(\xi)$ is assumed to be a continuously differentiable value function, then the optimal value function can be rewritten as

$$V^*(\xi) = \min_{u \in \Omega} \int_t^\infty e^{-\mu(\tau-t)} \left(\mathcal{L}(\xi(\tau), u(\tau)) + \|\bar{d}_m(\xi)\|^2 \right) d\tau \tag{9}$$

The defined value function (9) includes the function $\bar{d}_m(\xi)$ related to disturbances. Thus, the lemmas of equivalence between the robust control of the system (5) and the optimal control of the system (6) are provided.

Lemma 1 [40] *Let $\xi = 0$ be the equilibrium point of the system (6). If the solution to the optimal control problem of nominal system with value function (7) exists, then it is the solution to the robust tracking system (5) as well.*

Then, the Hamiltonian function is defined as

$$\mathcal{H}(\xi, \nabla V^*, u) = \nabla V^{*\top} (F(\xi) + G(\xi)u) - \mu V^* + \xi^\top Q \xi + \mathcal{R}(u) + \|\bar{d}_m(\xi)\|^2 \tag{10}$$

The optimal value $V^*(\xi)$ satisfies $V^*(0) = 0$. By optimal Bellman’s principle, the tracking HJB equation is shown as

$$0 = \min_{u \in \Omega} H(\xi, \nabla V^*(\xi), u) \tag{11}$$

Then, the following optimal control law is obtained

$$0 = \frac{\partial H}{\partial u} \Rightarrow u^*(\xi) = -\alpha \Theta \left(\frac{1}{2\alpha} R^{-1} G^\top(\xi) \nabla V^*(\xi) \right) + \beta \tag{12}$$

where $\beta = [\beta_0, \beta_0, \dots, \beta_0]^\top \in \mathbb{R}^m$.

Lemma 2 [41] *For the nominal system with continuously differentiable optimal value functions. The optimal control $u^*(\xi)$ in (12) can make the system asymptotically stable if the condition $\mu \rightarrow 0$ holds.*

Remark 2 The proofs of Lemma 1 and Lemma 2 are similar in [40, 41], which not be presented here. Lemma 1 shows that robust control can be solved by obtaining the solution of the HJB equation of nominal control. Based on Lemma 2 in [41], the error dynamics can be locally asymptotically stable when the discount factor μ approaches zero infinitely. Significantly, since the provided reference signal may not be asymptotically stable, the discount factor needs to be introduced for the value function defined in the augmented system to ensure that (7) is bounded (i.e., Remark 1).

According to (9)–(12), the optimal tracking HJB equation based on time-triggered for the nominal system becomes

$$\begin{aligned} &\xi^T Q \xi + \mathcal{R} \left(-\alpha \Theta \left(\frac{1}{2\alpha} R^{-1} G^\top(\xi) \nabla V^* \right) + \beta \right) + \nabla V^{*\top} F(\xi) \\ &+ \|\bar{d}_m(\xi)\|^2 + \nabla V^{*\top} G(\xi) \beta - \alpha \nabla V^{*\top} G(\xi) \\ &\Theta \left(\frac{1}{2\alpha} R^{-1} G^\top(\xi) \nabla V^* \right) - \mu V^* = 0 \end{aligned} \tag{13}$$

3 Solve the event-based tracking HJB equation

3.1 Event-triggered self-learning scheme

It is observed that (13) is a time-triggered HJB equation with locally uncertain dynamics $F(\xi)$. Then, the analytical solution of the HJB equation is a challenge to solve the optimal control problems. Meanwhile, the controller needs a heavy computational burden and numerous storage spaces by the time-triggered control with periodic sampling time, which imposes higher communication requirements on the actuator and plant. Thus, we prefer to propose an event-triggered self-learning algorithm to address such a problem.

To achieve online iterative learning and avoid dependence on overall system dynamics, we use a self-learning scheme known as integral reinforcement learning (IRL) to relax the demand for drift dynamics $F(\xi)$. For the integral time $T > 0$, the Bellman equation in IRL form is shown as

$$\begin{aligned} V(\xi(t-T)) &= \int_{t-T}^t e^{-\mu(\tau-t+T)} \\ &\left(\xi^\top Q \xi + \mathcal{R}(u) + \|\bar{d}_m(\xi)\|^2 \right) d\tau + e^{-\mu T} V(\xi(t)) \end{aligned} \tag{14}$$

Noting that (14) does not contain the dynamics of the tracking system.

Before introducing the event-triggered scheme, we firstly define a forward time sequence related to the event and express it as $t_q \triangleq \{t_q < t_{q+1}, q = 0, 1, 2, \dots\}$. If an event appears at a certain instant, the instant is defined as the triggering instant t_q . The current state $\xi(t)$ is denoted as sampling state $\hat{\xi}_q$ at the triggering instant. Thus, define the following triggering error as

$$e_q = \hat{\xi}_q - \xi(t), \quad t \in [t_q, t_{q+1}) \tag{15}$$

Only at the triggering instant will $e_q = 0$ be satisfied. However, during the non-triggered interval, there is a

triggering error in the augmented nominal system (6), and it converts to

$$\dot{\xi} = \mathcal{F}(\xi(t)) + G(\xi(t))u(\xi(t) + e_q) \quad t \in [t_q, t_{q+1}) \quad (16)$$

Then, (9) and (12) are updated in a non-periodic sampling manner by introducing the event-triggered control. Thus, the IRL-based event-triggered adaptive dynamic programming (ETADP) technology is shown in Algorithm 1.

Algorithm 1 IRL-based ETADP algorithm

Step 1: Select the initial permissible control law u_0 .
Step 2: Policy Evaluation During the interval time $[t - T, t]$, the value function $V^{[i]}(\xi)$ is obtained by solving the following Bellman equation

$$V^{[i]}(\xi(t - T)) = \int_{t-T}^t e^{-\mu(\tau-t+T)} \left(\mathcal{X}(\xi, u^{[i]}(\hat{\xi}_q)) + \|\bar{d}_m(\xi)\|^2 \right) d\tau + e^{-\mu T} V^{[i]}(\xi) \quad (17)$$

Step 3: Policy Improvement Update the control law by

$$u^{[i+1]}(\hat{\xi}_q) = -\alpha\Theta\left(\frac{1}{2\alpha}R^{-1}G^T(\hat{\xi}_q)\nabla V^{[i]}(\hat{\xi}_q)\right) + \beta \quad (18)$$

Step 4: If a positive constant \mathcal{V}_0 satisfies with $\|V^{[i+1]}(\xi) - V^{[i]}(\xi)\| \leq \mathcal{V}_0$, and then the iteration ends. Instead, continue to $i + 1$ and return to Step 2.

Remark 3 In Algorithm 1, the selected integration interval T is different from the traditional IRL Bellman equation in (14). Equation (14) yields reinforcement signals in a continuous periodic interval time. However, the ETC-based triggering interval $T_q = t_{q+1} - t_q$ is variable, and the event obtains new jump dynamics at different triggering integral instants. The implementation of the controller without periodic update law is determined by the triggering condition (or rule) in the event generator, which is the key to achieving ETC.

If a triggering error violates a designed event-triggered condition, an event occurs and obtains a new sampling state. Transmit the collected signal to the controller to generate a new control law at the triggering instant t_q . In flow dynamics, that is, during the interval time $t \in (t_q, t_{q+1})$ is not triggered. The sampling signal of the system comes from the sampling state stored in the zero-order hold (ZOH) at the last triggering instant and remains within the interval time. Therefore, the control law is obtained in the form of continuous segmented signals as follows

$$u^*(t) = \begin{cases} -\alpha\Theta\left(\frac{1}{2\alpha}R^{-1}G^T(\hat{\xi}_q)\nabla V^*(\hat{\xi}_q)\right) + \beta, & t \in [t_q, t_{q+1}) \\ u^*(\hat{\xi}_{q+1}), & t = t_{q+1} \end{cases} \quad (19)$$

Converting $u(\xi)$ in (11) to $u^*(\hat{\xi}_q)$, the event-based tracking HJB equation becomes

$$\begin{aligned} & \xi^T Q \xi + \mathcal{R}\left(-\alpha\Theta\left(\frac{1}{2\alpha}R^{-1}G^T(\hat{\xi}_q)\nabla V^*(\hat{\xi}_q)\right) + \beta\right) \\ & + \nabla V^{*T} F(\xi) + \|\bar{d}_m(\xi)\|^2 \\ & - \alpha \nabla V^{*T} G(\xi) \Theta\left(\frac{1}{2\alpha}R^{-1}G^T(\hat{\xi}_q)\nabla V^*(\hat{\xi}_q)\right) \\ & + \nabla V^{*T} G(\xi) \beta - \mu V^* = 0 \end{aligned} \quad (20)$$

Then, we need to design a triggering rule to implement equation (19). The following assumptions need to be proposed as [36–38]

Assumption 2 Let the optimal value function be continuously differentiable. Then, $V^*(\xi)$ and $\nabla V^*(\xi)$ are bounded by \mathcal{V}_1 and \mathcal{V}_2 with $\mathcal{V}_1 > 0$ and $\mathcal{V}_2 > 0$.

Assumption 3

1. There has a constant $\mathcal{K}_1 > 0$ satisfying the following inequality

$$\|u^*(\xi) - u^*(\hat{\xi}_q)\| \leq \mathcal{K}_1 \|\xi - \hat{\xi}_q\| = \mathcal{K}_1 \|e_q(t)\| \quad (21)$$

2. Let $F(\xi) + G(\xi)u^*(\hat{\xi}_q)$ be Lipschitz continuous on $\Omega \in \mathbb{R}^n$, and there exists positive constants \mathcal{L}_F and \mathcal{L}_G that make the following inequality hold

$$\|F(\xi) + G(\xi)u^*(\hat{\xi}_q)\| \leq \mathcal{L}_F\|\xi\| + \mathcal{L}_G\|e_q\| \quad (22)$$

Theorem 1 *The value function (7) with a non-quadratic function is defined for the nominal system (6). Let Assumptions 2–3 hold. The event-based control law is derived by (19), which can ensure the augmented system (5) is uniformly ultimately bounded (UUB) if the following triggering rule is satisfied*

$$\|e_q(t)\| \leq \sqrt{\frac{(1-\varrho^2)\underline{\lambda}(Q)}{v\mathcal{K}_1^2}}\|\xi\| \triangleq T_e \quad (23)$$

where $v > 0$ in (28) and $\varrho \in (0, 1)$ in (30) are adjustable parameters, and T_e denotes triggering threshold.

Proof Construct the Lyapunov function as $V^*(\xi)$ in (9) and use the dynamic equation (5) with the event-based optimal law $u^*(\hat{\xi}_q)$. The derivatives of $V^*(\xi)$ respecting to the time t are expressed as

$$\begin{aligned} \dot{V}^*(\xi) &= \nabla V^{*\top}(F(\xi) + G(\xi)u^*(\hat{\xi}_q) + G(\xi)\bar{D}(\xi)) \\ &= \nabla V^{*\top}(F(\xi) + G(\xi)u^*(\xi) + G(\xi)\bar{D}(\xi)) \\ &\quad + \nabla V^{*\top}G(\xi)(u^*(\hat{\xi}_q) - u^*(\xi)) \end{aligned} \quad (24)$$

According to (9)–(10), we have

$$\begin{aligned} \nabla V^{*\top}(F(\xi) + G(\xi)u^*(\xi)) \\ = -\xi^\top Q\xi - \mathcal{R}(u^*(\xi)) - \|\bar{d}_m(\xi)\|^2 + \mu V^* \end{aligned} \quad (25)$$

Then, convert (12) to

$$\nabla V^{*\top}G(\xi) = -2\alpha R\Theta^{-\top}(\alpha^{-1}(u^*(\xi) - \beta)) \quad (26)$$

Substituting (25) and (26) into (24), we obtain

$$\begin{aligned} \dot{V}^*(\xi) &= -\xi^\top Q\xi - \mathcal{R}(u^*(\xi)) - \|\bar{d}_m(\xi)\|^2 \\ &\quad - \underbrace{2\alpha R\Theta^{-\top}(\alpha^{-1}(u^*(\xi) - \beta)\bar{D}(\xi))}_{\mathcal{E}_1} \\ &\quad + \underbrace{2\alpha R\Theta^{-\top}(\alpha^{-1}(u^*(\xi) - \beta))(u^*(\xi) - u^*(\hat{\xi}_q))}_{\mathcal{E}_2} \\ &\quad + \mu V^* \end{aligned} \quad (27)$$

Introducing Young's inequality (i.e., $2a^\top b \leq a^2/v + vb^2$ with $v > 0$) and Assumption 2, \mathcal{E}_1 and \mathcal{E}_2 can be transformed as

$$\begin{aligned} \mathcal{E}_1 &\leq \frac{\|\alpha R\Theta^{-\top}(\alpha^{-1}(u^*(\xi) - \beta))\|^2}{v} + v\|\bar{d}_m(\xi)\|^2 \\ \mathcal{E}_2 &\leq \frac{\|\alpha R\Theta^{-\top}(\alpha^{-1}(u^*(\xi) - \beta))\|^2}{v} + v\|u^*(\xi) - u^*(\hat{\xi}_q)\|^2 \\ &\leq \frac{1}{4v}\|\nabla V^{*\top}G(\xi)\|^2 + v\mathcal{K}_1^2\|e_q(t)\|^2 \end{aligned} \quad (28)$$

According to (28), (27) can be translated to

$$\dot{V}^*(\xi) \leq -\xi^\top Q\xi - \mathcal{R}(u^*(\xi)) + \mu V^* + v\mathcal{K}_1^2\|e_q(t)\|^2 \quad (29)$$

Assuming that the conditions $\xi^\top Q\xi \geq \underline{\lambda}(Q)\|\xi\|^2$ containing the minimum eigenvalue of Q and $-\mathcal{R}(u^*(\xi)) \leq 0$ defined in (8) with the optimal control law are satisfied. Based on Assumptions 1–2, we have

$$\begin{aligned} \dot{V}^*(\xi) &\leq -\varrho^2\underline{\lambda}(Q)\|\xi\|^2 - (1-\varrho^2)\underline{\lambda}(Q)\|\xi\|^2 \\ &\quad + v\mathcal{K}_1^2\|e_q(t)\|^2 + \mu\mathcal{V}_1 \end{aligned} \quad (30)$$

If the triggering rule is devised as (23), (30) converts to

$$\dot{V}^*(\xi) \leq -\varrho^2\underline{\lambda}(Q)\|\xi\|^2 + \mu\mathcal{V}_1 \quad (31)$$

so that $\dot{V}^*(\xi) < 0$ is valid when the following inequality is not within the range of the set Λ_ξ related to the augmented state as

$$\Lambda_\xi = \left\{ \xi \in \mathbb{R} : \|\xi\| \leq \sqrt{\frac{\mu\mathcal{V}_1}{\varrho^2\underline{\lambda}(Q)}} \right\} \quad (32)$$

Thus, we can conclude the augmented nominal system state is UUB by using the Lyapunov stability theorem. \square

Remark 4 After the introduction of ETC, the nominal system (6) includes both continuous sampling signals and discrete sampling signals, which is known as the hybrid system. The existence of Zeno behavior in hybrid systems may result in infinite discrete sampling signals at finite triggering intervals. The existence of the Zeno behavior cannot ensure the stability of the system.

Theorem 2 *Assuming that Assumption 3 holds, and the triggering rule is designed as (23). Then, the minimal triggering instant satisfies $(T_q)_{\min} > 0$, where $(T_q)_{\min} = (t_{q+1} - t_q)_{\min}$ is a lower bounded.*

Proof Considering that $\bar{D}(\xi)$ is bounded on Ω , then $\|G(\xi)\bar{D}(\xi)\| \leq \mathcal{D}_m$ for any $\xi \in \Omega$ is defined with a positive constant \mathcal{D}_m . According to Assumption 3, the tracking system (5) with the optimal event-based control law is converted to

$$\begin{aligned} \|\dot{\xi}\| &= \|F(\xi) + G(\xi)u^*(\hat{\xi}_q)\| + \|G(\xi)\bar{D}(\xi)\| \\ &\leq \mathcal{L}_F\|\xi\| + \mathcal{L}_G\|e_q\| + \mathcal{D}_m \end{aligned} \tag{33}$$

A sampling state $\hat{\xi}_q = \xi(t_q)$ occurs at the triggering instant $t = t_q$, and we obtain $\dot{e}_q = -\dot{\xi}(t)$. Thus, (33) is converted to

$$\begin{aligned} \|\dot{e}_q\| &= \|\dot{\xi}\| \leq \mathcal{L}_F\|\hat{\xi}_q - e_q\| + \mathcal{L}_G\|e_q\| + \mathcal{D}_m \\ &\leq \mathcal{L}_F\|\hat{\xi}_q\| + (\mathcal{L}_F + \mathcal{L}_G)\|e_q\| + \mathcal{D}_m \end{aligned} \tag{34}$$

From (15), we obtain the triggering error $e_q = 0$ at the triggering instant $t = t_q$. Accordingly, the solution of equation (34) is

$$\|\dot{e}_q\| \leq \frac{\mathcal{L}_F\|\hat{\xi}_q\| + \mathcal{D}_m}{\mathcal{L}_F + \mathcal{L}_G} \left(e^{(\mathcal{L}_F + \mathcal{L}_G)(t-t_q)} - 1 \right) \tag{35}$$

Based on Theorem 1, the tracking state is UUB. Hence, we obtain the state $\|\hat{\xi}_{q+1}\| \neq 0$ at the triggering instant $t = t_{q+1}$. Then, (35) is converted into

$$T_q = t_{q+1} - t_q \geq \frac{1}{\mathcal{L}_F + \mathcal{L}_G} \ln \left(1 + \mathcal{M}_0 \sqrt{\frac{(1 - \varrho^2)\underline{\lambda}(Q_0)}{v\mathcal{K}_1^2}} \right) \tag{36}$$

where

$$\mathcal{M}_0 = \frac{(\mathcal{L}_F + \mathcal{L}_G)\|e_{q+1}\|}{\mathcal{L}_F\|\hat{\xi}_q\| + \mathcal{D}_m} > 0 \tag{37}$$

According to (36) and (37), \mathcal{M}_0 is bounded by a minimum constant $\mathcal{M}_{min} > 0$. Then, we obtain

$$(T_q)_{min} = \frac{1}{\mathcal{L}_F + \mathcal{L}_G} \ln \left(1 + \mathcal{M}_{min} \sqrt{\frac{(1 - \varrho^2)\underline{\lambda}(Q_0)}{v\mathcal{K}_1^2}} \right) > 0 \tag{38}$$

Therefore, there exists the minimum interval in the control system that eliminates the Zeno behavior. \square

3.2 A single-critic network approximate solver

The optimal value function and optimal control law of the tracking nominal system can be approximated by a single-critic neural network. Due to the universal approximation principle of neural networks, the optimal value function (9) and its derivative are reconstructed by

$$V^*(\xi) = \mathcal{W}_c^T \psi(\xi) + \mathcal{E}_0(\xi) \tag{39}$$

$$\nabla V^*(\xi) = \nabla \psi^T(\xi) \mathcal{W}_c + \nabla \mathcal{E}_0(\xi) \tag{40}$$

where $\mathcal{W}_c \in \mathbb{R}^l$ is a constant vector, l indicates the number of hidden layer neurons, and $\psi(\xi)$ is suitable basis function vector. The approximate error $\mathcal{E}_0(\xi)$ and its gradient

$\nabla \mathcal{E}_0(\xi)$ are confined by positive constants and satisfied with $\|\mathcal{E}_0(\xi)\| \leq \mathcal{E}_m$ and $\|\nabla \mathcal{E}_0(\xi)\| \leq \mathcal{E}_M$.

The correlated value function for the tracking Bellman equation is approximated by the critic network. Then, the tracking Belman error generated by NNs approximation is obtained

$$\begin{aligned} \epsilon_B(t) \triangleq & \int_{t-T}^t e^{-\mu(\tau-t+T)} \left(2\alpha \sum_{i=1}^m \int_{\beta_0}^{u^*} \Theta^{-1} \left(\frac{\eta - \beta_0}{\alpha} \right) R d\eta \right. \\ & \left. + \xi^T Q \xi + \|\bar{d}_m(\xi)\|^2 \right) d\tau + \mathcal{W}_c^T \Delta \psi(\xi(t)) \end{aligned} \tag{41}$$

where $\Delta \psi(\xi(t)) = e^{-\mu T} \psi(\xi(t)) - \psi(\xi(t-T))$. ϵ_B is bounded by a constant and satisfies by $\|\epsilon_B\| \leq \epsilon_M$. Then, the reinforcement reward in the integral domain is denoted as

$$p(t) = \int_{t-T}^t e^{-\mu(\tau-t+T)} \left(\xi^T Q \xi + \mathcal{R}(u^*(\xi)) + \|\bar{d}_m(\xi)\|^2 \right) d\tau \tag{42}$$

Combining (12) and (40), the optimal control law based on neural network at triggering instant $t = t_q$ by the mean value theorem is represented as

$$u^*(\hat{\xi}_q) = -\alpha \Theta(\Psi(\hat{\xi}_q)) + \beta + \mathcal{E}_u(\hat{\xi}_q) \tag{43}$$

where $\Psi(\hat{\xi}_q) = (R^{-1}G^T(\hat{\xi}_q)\nabla \psi^T(\hat{\xi}_q)\mathcal{W}_c)/(2\alpha)$.

$\mathcal{E}_u(\hat{\xi}_q) = -\varepsilon_u(\hat{\xi}_q)(R^{-1}G^T(\hat{\xi}_q)\nabla \mathcal{E}_0(\hat{\xi}_q))/2$ with $\varepsilon_u(\hat{\xi}_q) = [1 - \Theta^2(\delta_1(\hat{\xi}_q)), \dots, 1 - \Theta^2(\delta_m(\hat{\xi}_q))]^T \in \mathbb{R}^m$, and we obtain the point $\delta_i(\hat{\xi}_q) \in \mathbb{R}, i = 1, 2, \dots, m$ within the integration interval of $\Psi(\hat{\xi}_q)$ and $\Psi(\hat{\xi}_q) + (R^{-1}G^T(\hat{\xi}_q)\nabla \mathcal{E}_0(\hat{\xi}_q))/(2\alpha)$.

We need to design a controller for the unknown value function (39) and obtain the optimal value function by adjusting the weight of the critic network to ensure convergence. Since the ideal critic NN weight \mathcal{W}_c obtained the optimal approximate solution is unknown, and the current weight $\hat{\mathcal{W}}_c$ is applied to approximate

$$\hat{V}(\xi) = \hat{\mathcal{W}}_c^T \psi(\xi) \tag{44}$$

$$\nabla \hat{V}(\xi) = \nabla \psi^T(\xi) \hat{\mathcal{W}}_c \tag{45}$$

Then, the approximate Bellman error equation is

$$\delta_B(t) = p(t) + \hat{\mathcal{W}}_c^T \Delta \psi(\xi(t)) \tag{46}$$

The error δ_B can be regarded as the corresponding error of the time difference (TD) error of a continuous-time system [42]. Therefore, we find the optimal value function by adjusting the critic NN weight to minimize the TD error δ_B . Define the objective function as $\mathcal{E}_B = (\delta_B^T \delta_B)/2$, and the adaptive turning law of the critic network is obtained

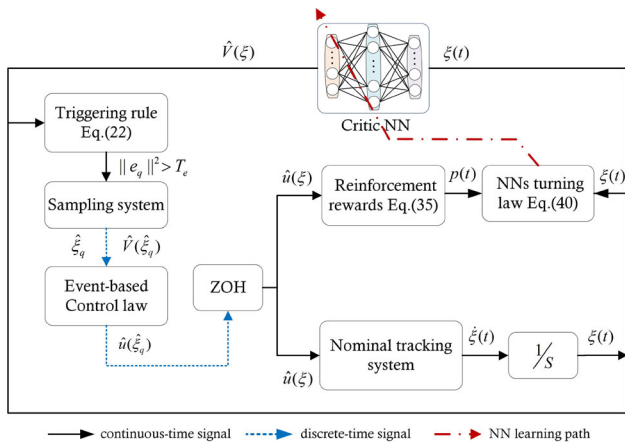


Fig. 1 Proposed algorithm structure diagram

$$\begin{aligned} \dot{\hat{W}}_c &= -a_c \frac{\partial \mathcal{E}_B}{\partial \hat{W}_c} = -a_c \frac{\Delta\psi(\xi)}{(1 + \Delta\psi^T(\xi)\Delta\psi(\xi))^2} \delta_B \\ &= -a_c \frac{\Delta\psi(\xi)}{(1 + \Delta\psi^T(\xi)\Delta\psi(\xi))^2} (p(t) + \hat{W}_c^T \Delta\psi(\xi)) \end{aligned} \tag{47}$$

Remark 5 The single-critic network is a three-level neural network that includes input layers, hidden layers, and output layers. The structure diagram of the principle is shown in Fig. 1. Assuming that the activation function $\Delta\psi$ is continuously exciting in the interval $[t - T, t]$. The approximate weight \hat{W}_c can be guaranteed to converge to ideal weight W_c based on the persistent excitation (PE) conditions [43]. In the simulation experiment, the PE condition is composed of detection signals with different amplitude function combinations.

The IRL-based event-triggered ADP is implemented in an online manner by tuning the critic network. The optimal control law is obtained by adjusting the weight of critic NN. According to (19), the event-based control law can be approximated by

$$\hat{u}(\hat{\xi}_q) = -\alpha \Theta \left(\frac{1}{2\alpha} R^{-1} G^T(\hat{\xi}_q) \nabla \psi^T(\hat{\xi}_q) \hat{W}_c \right) + \beta \tag{48}$$

Then, the critic network approximation error can be defined as $\tilde{W}_c = W_c - \hat{W}_c$. To make $\hat{W}_c \rightarrow W_c$ be satisfied, define the control input errors as

$$\begin{aligned} \delta_u = u^*(\hat{\xi}_q) - \hat{u}(\hat{\xi}_q) &= \alpha \Theta \left(\frac{1}{2\alpha} R^{-1} G^T(\hat{\xi}_q) \nabla \psi^T(\hat{\xi}_q) \hat{W}_c \right) \\ &\quad - \alpha \Theta(\Psi(\hat{\xi}_q)) + \mathcal{E}_u(\hat{\xi}_q) \end{aligned} \tag{49}$$

where $\mathcal{E}_u(\hat{\xi}_q)$ is from (43). It is the ideal result $\delta_u = 0$ to get the optimal laws.

The system obtains new sampling signals during the event-triggered integration interval depending on the triggering condition. Therefore, the signal acquisition for reinforcement rewards is obtained in a non-periodic manner. Based on (41) and (42), the Bellman error with event-based control laws is yielded by

$$\check{\delta}_B(t) = \check{p}(t) + \hat{W}_c^T \Delta\psi(\xi(t)) \tag{50}$$

where

$$\check{p}(t) = \int_{t-T}^t e^{-\mu(\tau-t+T)} \left(\xi^T Q \xi + \mathcal{R}(\hat{u}(\hat{\xi}_q)) + \|\bar{d}_m(\xi)\|^2 \right) d\tau$$

is the reinforcement rewards in the triggering interval.

From (47), denote $\bar{\Delta}\psi(\xi) = \Delta\psi/(1 + \Delta\psi^T \Delta\psi)$ and $m_\psi(\xi) = 1/(1 + \Delta\psi^T \Delta\psi)$. Then, the critic weight error dynamics can be given as

$$\dot{\tilde{W}}_c = -\dot{\hat{W}}_c = -a_c \bar{\Delta}\psi \bar{\Delta}\psi^T \tilde{W}_c + a_c \bar{\Delta}\psi m_\psi \epsilon_B \tag{51}$$

4 Stability analysis

To facilitate the stability analysis of the system, the following assumptions are given [27–29]

Assumption 4

- $\mathcal{E}_u(\hat{\xi}_q)$ is bounded by a positive constant with $\|\mathcal{E}_u(\hat{\xi}_q)\| \leq c_{\mathcal{E}_u}$.
- $\psi(\xi)$ and $\nabla\psi(\xi)$ are bounded by positive constants and satisfied by $\|\psi(\xi)\| \leq \psi_m$ and $\|\nabla\psi(x)\| \leq \psi_M$.

Theorem 3 The initial control for the dynamical system (6) is admissible. Let Assumptions 1–4 be valid, and the optimal event-based control law $u^*(\hat{\xi}_q)$ is implement by turning the weight dynamic (47) for the tracking nominal system (6). The tracking system state ξ and weight approximation errors \tilde{W}_c are UUB only when the triggering condition (22) holds and the following inequality satisfies

$$3a_c \underline{\lambda}(\bar{\Delta}\psi \bar{\Delta}\psi^T) - 16 \mathcal{L}_G^2 \psi_M^2 \|R^{-1}\|^2 > 0 \tag{52}$$

where $\underline{\lambda}(\bar{\Delta}\psi \bar{\Delta}\psi^T)$ is the minimum eigenvalue of $\bar{\Delta}\psi \bar{\Delta}\psi^T$ in (59), and a_c is the learning rate of NNs.

Proof For the nominal system (6) with the optimal value function (9), the following Lyapunov function is applied for

$$\mathcal{L}(t) = \mathcal{L}_1(t) + \mathcal{L}_2(t) + \mathcal{L}_3(t) \tag{53}$$

where $\mathcal{L}_1 = V^*(\xi)$, $\mathcal{L}_2 = V^*(\hat{\xi}_q)$, and $\mathcal{L}_3 = (1/2) \tilde{W}_c^T \tilde{W}_c$.

Based on Remark 4, the discrete sampling state $\xi(t_q)$ and the continuous sampling state $\xi(t)$ are included in hybrid systems due to the introduction of ETC. Thus, we describe the following two situations.

Case 1. For the triggering condition is not satisfied, that is, during the flow dynamic on $t \in [t_q, t_{q+1})$, we have

$$\begin{aligned} \dot{\mathcal{L}}_1(t) &= \dot{V}^*(\xi) = \nabla V^{*\top}(F(\xi) + G(\xi)\hat{u}(\hat{\xi}_q)) \\ &= \nabla V^{*\top}(F(\xi) + G(\xi)u^*(\xi)) \\ &\quad + \nabla V^{*\top}G(\xi)(\hat{u}(\hat{\xi}_q) - u^*(\xi)) \end{aligned} \tag{54}$$

Substituting equations (24) and (25) into (54), we obtain

$$\begin{aligned} \dot{\mathcal{L}}_1(t) &= -\xi^\top Q \xi - \mathcal{R}(u^*(\xi)) - \|\bar{d}_m(\xi)\|^2 + \mu V^* \\ &\quad + \underbrace{2\alpha R \Theta^{-\top}(\alpha^{-1}(u^*(\xi) - \beta))}_{\mathcal{E}_3}(u^*(\xi) - \hat{u}(\hat{\xi}_q)) \end{aligned} \tag{55}$$

Similar to (27) and using the inequality $2a^\top b \leq \|a\|^2 + \|b\|^2$, \mathcal{E}_3 can be converted to

$$\begin{aligned} \mathcal{E}_3 &\leq \|\alpha R \Theta^{-\top}(\alpha^{-1}(u^*(\xi) - \beta))\|^2 + \|u^*(\xi) - \hat{u}(\hat{\xi}_q)\|^2 \\ &\leq \frac{1}{4} \|\nabla V^{*\top}G(\xi)\|^2 + \underbrace{\|u^*(\xi) - \hat{u}(\hat{\xi}_q)\|^2}_{\mathcal{M}_1} \end{aligned} \tag{56}$$

By using the function property $|\tanh(y)| \leq 1$ for $\forall y \in \mathbb{R}$ and equation (49), we can get $\Theta(\Psi(\hat{\xi}_q)) \leq k_0$ with $k_0 \leq 1$. Based on Young's inequality (i.e., $\|a \pm b\|^2 \leq 2\|a\|^2 + 2\|b\|^2$) and Assumptions 3–4, \mathcal{M}_1 can be converted to

$$\begin{aligned} \mathcal{M}_1 &= \|u^*(\xi) - u^*(\hat{\xi}_q) + u^*(\hat{\xi}_q) - \hat{u}(\hat{\xi}_q)\|^2 \\ &\leq 2\|u^*(\xi) - u^*(\hat{\xi}_q)\|^2 + 2\|u^*(\hat{\xi}_q) - \hat{u}(\hat{\xi}_q)\|^2 \\ &\leq 2\|\alpha \Theta(\frac{1}{2\alpha} R^{-1} G^\top(\hat{\xi}_q) \nabla \psi^\top(\hat{\xi}_q) \hat{\mathcal{W}}_c) \\ &\quad - \alpha \Theta(\Psi(\hat{\xi}_q)) + \mathcal{E}_u(\hat{\xi}_q)\|^2 + 2\mathcal{K}_1^2 \|e_q\|^2 \\ &\leq 4\alpha^2 \|\Theta(\frac{1}{2\alpha} R^{-1} G^\top(\hat{\xi}_q) \nabla \psi^\top(\hat{\xi}_q) \hat{\mathcal{W}}_c) \\ &\quad - \Theta(\Psi(\hat{\xi}_q))\|^2 + 2\mathcal{K}_1^2 \|e_q\|^2 + 4\|\mathcal{E}_u(\hat{\xi}_q)\|^2 \\ &\leq 2\|R^{-1} G^\top(\hat{\xi}_q) \nabla \psi^\top(\hat{\xi}_q) \hat{\mathcal{W}}_c\|^2 + 8\alpha^2 k_0^2 \\ &\quad + 2\mathcal{K}_1^2 \|e_q\|^2 + 4c_{\mathcal{E}_u}^2 \\ &\leq 2\mathcal{K}_1^2 \|e_q\|^2 + 4c_{\mathcal{E}_u}^2 + 8\alpha^2 k_0^2 + 2\mathcal{L}_G^2 \psi_M^2 \|R^{-1}\|^2 \|\hat{\mathcal{W}}_c\|^2 \end{aligned} \tag{57}$$

According to (56) and (57), (55) is rewritten as

$$\begin{aligned} \dot{\mathcal{L}}_1(t) &\leq -\xi^\top Q \xi - \mathcal{R}(u^*(\xi)) + \frac{1}{4} \gamma_2^{-2} \mathcal{L}_G^2 + 2\mathcal{K}_1^2 \|e_q\|^2 \\ &\quad + 8\alpha^2 k_0^2 + \mu \mathcal{V}_1 \\ &\quad + 2\mathcal{L}_G^2 \psi_M^2 \|R^{-1}\|^2 \|\hat{\mathcal{W}}_c\|^2 + 4c_{\mathcal{E}_u}^2 - \|\bar{d}_m(\xi)\|^2 \end{aligned} \tag{58}$$

During the no-triggered interval $t \in [t_q, t_{q+1})$, we obtain $\dot{\mathcal{L}}_2(t) = V^*(\hat{\xi}_q) = 0$. According to (51) and $m_\psi \leq 1$, $\dot{\mathcal{L}}_3(t)$ is given as

$$\begin{aligned} \dot{\mathcal{L}}_3 &= \tilde{\mathcal{W}}_c^\top \dot{\tilde{\mathcal{W}}}_c = -a_c \tilde{\mathcal{W}}_c^\top \bar{\Delta} \psi \bar{\Delta} \psi^\top \tilde{\mathcal{W}}_c + a_c \tilde{\mathcal{W}}_c^\top \bar{\Delta} \psi m_\psi \epsilon_B \\ &\leq a_c \left(\frac{1}{4} \tilde{\mathcal{W}}_c^\top \bar{\Delta} \psi \bar{\Delta} \psi^\top \tilde{\mathcal{W}}_c + \epsilon_B^\top \epsilon_B \right) - a_c \tilde{\mathcal{W}}_c^\top \bar{\Delta} \psi \bar{\Delta} \psi^\top \tilde{\mathcal{W}}_c \\ &\leq -\frac{3}{4} a_c \underline{\lambda}(\bar{\Delta} \psi \bar{\Delta} \psi^\top) \|\tilde{\mathcal{W}}_c\|^2 + a_c \epsilon_M^2 \end{aligned} \tag{59}$$

According to $\xi^\top Q \xi \geq \underline{\lambda}(Q) \|\xi\|^2$ and combining (58) and (59), the derivative of $\mathcal{L}(t)$ becomes

$$\begin{aligned} \dot{\mathcal{L}}(t) &\leq -\varrho^2 \underline{\lambda}(Q) \|\xi\|^2 - (1 - \varrho^2) \underline{\lambda}(Q) \|\xi\|^2 \\ &\quad - \frac{3}{4} a_c \underline{\lambda}(\bar{\Delta} \psi \bar{\Delta} \psi^\top) \|\tilde{\mathcal{W}}_c\|^2 \\ &\quad + 2\mathcal{K}_1^2 \|e_q\|^2 + 2\mathcal{L}_G^2 \psi_M^2 \|R^{-1}\|^2 \|\mathcal{W}_c - \tilde{\mathcal{W}}_c\|^2 + \mu \mathcal{V}_1 \\ &\quad + \frac{1}{4} \gamma_2^{-2} \mathcal{L}_G^2 + 4c_{\mathcal{E}_u}^2 + 8k_0^2 \alpha^2 + a_c \epsilon_M^2 \end{aligned} \tag{60}$$

If the triggering condition is provided by (22), then we have

$$\begin{aligned} \dot{\mathcal{L}}(t) &\leq -\left(\frac{3}{4} a_c \underline{\lambda}(\bar{\Delta} \psi \bar{\Delta} \psi^\top) - 4\mathcal{L}_G^2 \psi_M^2 \|R^{-1}\|^2 \right) \|\tilde{\mathcal{W}}_c\|^2 \\ &\quad - \varrho^2 \underline{\lambda}(Q) \|\xi\|^2 + \Gamma_0 \end{aligned} \tag{61}$$

where

$$\begin{aligned} \Gamma_0 &= 4\mathcal{L}_G^2 \psi_M^2 \|R^{-1}\|^2 \mathcal{W}_c^2 + \frac{1}{4} \gamma_2^{-2} \mathcal{L}_G^2 + 4c_{\mathcal{E}_u}^2 + 8k_0^2 \alpha^2 \\ &\quad + a_c \epsilon_M^2 + \mu \mathcal{V}_1 \end{aligned}$$

Thus, we can conclude that $\dot{\mathcal{L}}(t)$ is negative definite if one of the following inequalities is not within the range of the sets Ω_ξ and $\Omega_{\tilde{\mathcal{W}}_c}$ as

$$\Omega_\xi = \left\{ \xi \in \mathbb{R} : \|\xi\| \leq \sqrt{\frac{\Gamma_0}{\varrho^2 \underline{\lambda}(Q)}} \right\}$$

$$\Omega_{\tilde{W}_c} = \left\{ \tilde{W}_c : \|\tilde{W}_c\| \leq \sqrt{\frac{4\Gamma_0}{3a_c \underline{\lambda}(\Delta\psi \Delta\psi^T) - 16\mathcal{L}_c^2 \varrho_M^2 \|R^{-1}\|^2}} \right\} \tag{62}$$

Case 2. The event obtains a jump dynamic at the triggering instant $t = t_{q+1}$. Then, the difference of Lyapunov function (53) is considered as

$$\Delta\mathcal{L}(t_{q+1}) = \Delta\mathcal{L}_1(t_{q+1}) + \Delta\mathcal{L}_2(t_{q+1}) + \Delta\mathcal{L}_3(t_{q+1}) \tag{63}$$

where $\Delta\mathcal{L}_1 = V^*(\xi(t_{q+1})) - V^*(\xi(t_{q+1}^-))$, $\Delta\mathcal{L}_2 = V^*(\hat{\xi}_{q+1}) - V^*(\hat{\xi}_q)$, and $\Delta\mathcal{L}_3 = \frac{1}{2}\tilde{W}_c^T(t_{q+1})\tilde{W}_c(t_{q+1}) - \frac{1}{2}\tilde{W}_c^T(t_{q+1}^-)\tilde{W}_c(t_{q+1}^-)$. There exists $\sigma \in (0, t_{q+1} - t_q)$ with $\xi(t_{q+1}^-) = \lim_{\sigma \rightarrow 0^+} \xi(t_{q+1} - \sigma)$ and $\tilde{W}_c(t_{q+1}^-) = \lim_{\sigma \rightarrow 0^+} \tilde{W}_c(t_{q+1} - \sigma)$.

According to Case 1, we can conclude that the nominal system state ξ is UUB during the no-triggered interval and $\dot{\mathcal{L}}(t) < 0$ hold. Then, we have $V^*(\hat{\xi}_{q+1}) \leq V^*(\hat{\xi}_q)$, that is, $\Delta\mathcal{L}_2(t_{q+1}) < 0$. Owing to the augmented state $\xi(t)$, the weight approximation error \tilde{W}_c is continuous and \mathcal{L}_1 and \mathcal{L}_3 from (53) are strictly incremental during the triggering interval $t \in [t_q, t_{q+1})$, and then we obtain

$$\begin{aligned} & V^*(\xi(t_{q+1}^-)) + \frac{1}{2}\tilde{W}_c^T(t_{q+1}^-)\tilde{W}_c(t_{q+1}^-) \\ &= \lim_{\sigma \rightarrow 0^+} V^*(\xi(t_{q+1} - \sigma)) \\ &+ \frac{1}{2}\lim_{\sigma \rightarrow 0^+} \tilde{W}_c^T(t_{q+1} - \sigma)\tilde{W}_c(t_{q+1} - \sigma) \\ &\geq V^*(\xi(t_{q+1})) + \frac{1}{2}\tilde{W}_c^T(t_{q+1})\tilde{W}_c(t_{q+1}) \end{aligned} \tag{64}$$

that is, $\Delta\mathcal{L}_1(t_{q+1}) + \Delta\mathcal{L}_3(t_{q+1}) \leq 0$. Thus, we can derive that $\Delta\mathcal{L}(t) < 0$ at the triggering instant $t = t_{q+1}$ if only one of the inequalities in (62) does not hold. Combining the discussion of the two cases, the tracking system state ξ and the weight approximation error \tilde{W}_c are UUB if $\xi \notin \Omega_\xi$ or $\tilde{W}_c \notin \Omega_{\tilde{W}_c}$ holds. \square

5 Simulation results

In this part, the actionability of the proposed algorithm is verified by applying the results of two examples. Firstly, the antagonistic bionic joint is taken as the controlled object to achieve optimal tracking results while reducing the number of controller updates. Then, the optimal tracking is obtained while ensuring the stability of a continuous-time nonlinear system in Sect. 5.2.

5.1 Example 1

Given that two pneumatic artificial muscles (PAMs) drive the end load in a confrontational manner. From Fig. 2, one PAM is used to release the air pressure, while the other PAM is used to input the air pressure, which causes the pulley to revolve by θ . The total force exerted by each PAM throughout an operation about the air pressure coefficient can be expressed as F_{P1} and F_{P2} . Assuming that two PAMs are increased or decreased with the same air pressure value p . The effective working length of the two PAMs is shown as $s_{1,2} = s_0 \pm \theta R$. The dynamic equation of torque generated by two PAMs is

$$J\ddot{\theta} + c\dot{\theta} + Mgl\sin\theta = (F_{P1} - F_{P2})R + \bar{d}(x) \tag{65}$$

where J is the equivalent rotational inertia of the pulley and joint. l is the length of joint, and M is the mass of pulley. $c\dot{\theta}$ means the joint damping term and R is the pulley radius. $\bar{d}(x)$ is unknown disturbance with a upper bound $\bar{d}_m(x)$.

Define $x_1 = \theta$, $x_2 = \dot{\theta}$, and (65) can be transformed as

$$\begin{cases} \dot{x}_1 = x_2 \\ \dot{x}_2 = f(x_1, x_2) + gu(t) + h\bar{d}(x) \end{cases} \tag{66}$$

where $f(x_1, x_2) = 2I^{-1}((b_1 + b_2 p_0)R^2 - 2I^{-1}c)x_2 - 2I^{-1}((k_1 + k_2 p_0)R^2 x_1 - (Mgl \sin x_1)/2)$, $g = RI^{-1}(f_2 - k_2 s_0)$, and $h = I^{-1}$, $p(t) = u(t)$. b_1 and b_2 are the damping coefficients, k_1 and k_2 are the spring coefficients, and f_1 and f_2 are the shrinkage force parameter of PAM. p_0 indicates the initial pressure. Other relevant parameters and reference values related to the model are given in Table 1. The constrained control satisfies by $u \in \{u \in \mathbb{R} : -25 \leq \|u\| \leq 30\}$. We analyze the system performance at the unknown disturbances $\bar{d}(x) = p_0 x_1 \cos x_2$ with $p_0 \in [-1, 1]$ and choose the function related to disturbance as $\bar{d}_m(x) = x$.

The desirable trajectories are considered as $x_{r1} = \sin(t)$ and $x_{r2} = \cos(t)$. Give the initial augmented state $\xi = [e_{r1}, e_{r2}, x_{r1}, x_{r2}]^T = [2, -2, 1, 0]^T$. The parameters

Table 1 Identified parameters of the model

Parameters	Value	Coefficients	Value
Load mass/ M	2 kg	f_1	14.89
Pulley mass/ m	0.5 kg	f_2	2.73
Pulley radius/ R	0.5 m	k_1	4.82
Arm length/ L	0.5 m	k_2	1.32
PAM initial length/ s_0	0.3 m	b_1	1.22
Damping coefficient/ c	0.2	b_2	1.35

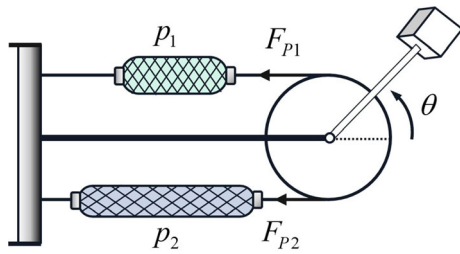


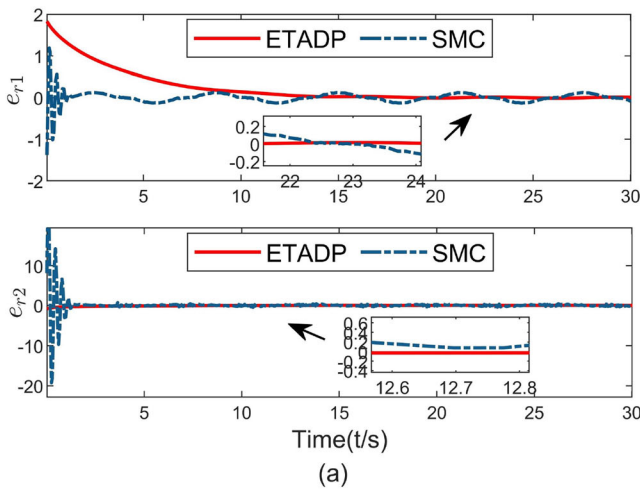
Fig. 2 Antagonistic bionic joint model

related to the value function (9) are $Q_0 = 100\text{diag}\{1, 1\}$, $R = 0.2$, and $\mu = 0.2$. The learning rate of the NNs is $a_c = 5$ and the detection signal is selected as $n(t) = 1.69 \sin(t)^2 \cos(2t) +$

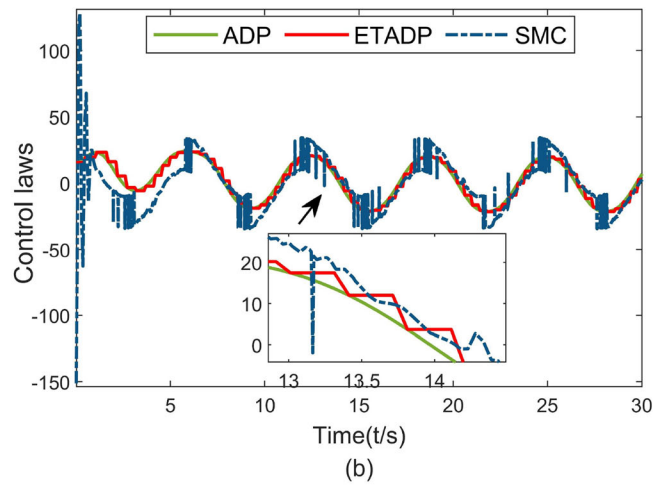
$1.69 \sin(2t)^2 \cos(0.1t) + 0.3 \sin(2.3t)^4 \cos(7t)$. The parameter of the triggering condition is determined in $\varrho \in (0, 1)$. The approximate NN weight vector is $\hat{W}_c = [\hat{w}_{c1}, \hat{w}_{c2}, \dots, \hat{w}_{c10}]^T$, and the activation function $\psi(\xi(t))$ is selected as

$$\psi(\xi) = [e_{r1}^2, e_{r1}e_{r2}, e_{r1}x_{r1}, e_{r1}x_{r2}, e_{r2}^2, e_{r2}x_{r2}, x_{r1}^2, x_{r1}x_{r2}, x_{r2}^2]^T \tag{67}$$

To demonstrate the performance of the proposed algorithm in tracking accuracy and optimal control, the application of sliding mode control (SMC) at biomimetic joints in [44] is utilized for comparison. The tracking error of the PAMs joint is shown in Fig. 3a and ultimately approaches zero within a certain range. The angle error e_{r1} and speed error

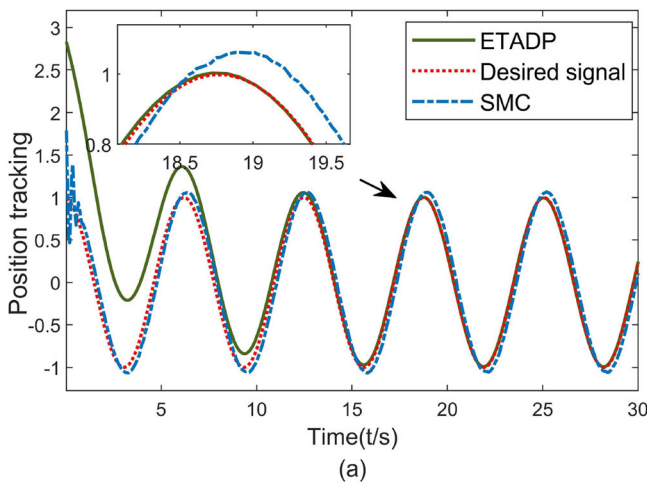


(a) Tracking errors of e_{r1} and e_{r2} .

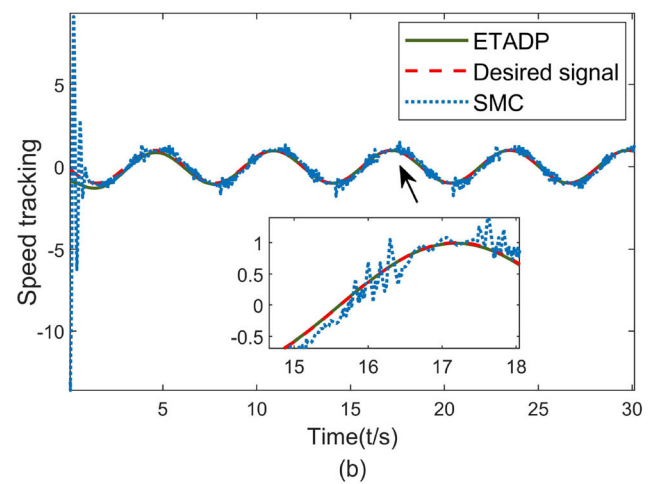


(b) Control inputs under different methods.

Fig. 3 Tracking errors and control laws



(a) Position tracking of PAMs model.



(b) Speed tracking of PAMs model.

Fig. 4 Position tracking and speed tracking under three control methods

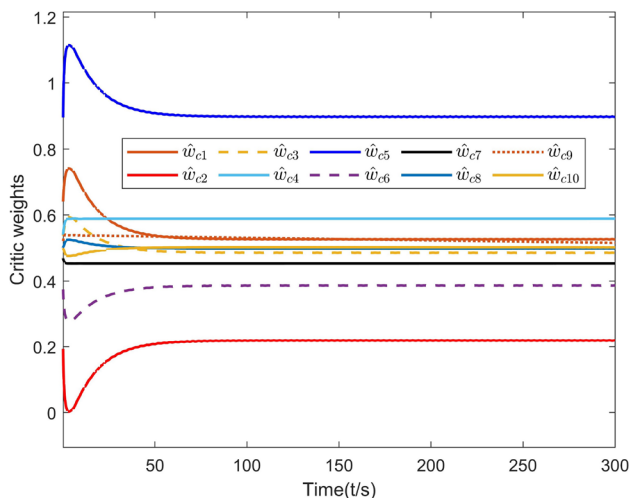


Fig. 5 The convergence trajectory of $\hat{\mathcal{W}}_c$

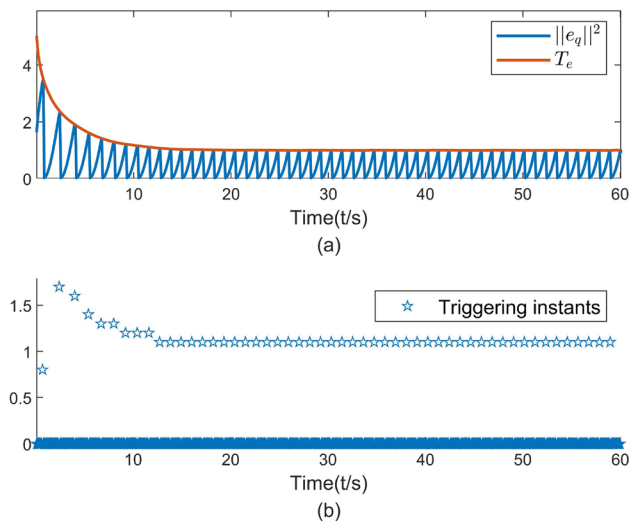


Fig. 6 a Event error $\|e_q\|^2$ and triggering threshold $\|T_e\|$. b Triggering intervals

e_{r2} of SMC are greater than ETADP. The constrained input laws based on TTC and ETC are shown in Fig. 3b. Compared with the TTC and SMC, the ETC consumes the least control decision. The tracking performances of position and speed for the PAMs model are displayed in Fig. 4a, b. After the online learning of 50 seconds under the PE condition, the single-critic NN weights converge to

$[0.5265, 0.2193, 0.4854, 0.5886, 0.8987, 0.3866, 0.4540, 0.4983, 0.5157, 0.5017]^T$ in Fig. 5. The triggering condition and sampling interval are demonstrated in Fig. 6. In addition, the horizontal coordinate is the sampling state obtained at this instant, and the vertical coordinate represents the difference between adjacent triggering intervals. According to the simulation results, the minimum sampling interval is $(T_q)_{\min} = 0.2s$. Thus, the Zeno behavior is successfully avoided. Experiments show that the sampling state under TTC and SMC is 1000, while the ETC only needs 247 samples. The ETC-based system greatly reduces the controller updates. Furthermore, the performance of the three control methods is shown in Table 2. From Table 2, it can be concluded that the advantages of the proposed ETADP method are mainly reflected in control accuracy, decreased controller updates, and reduced calculation costs.

5.2 Example 2

Consider the following nonlinear continuous-time system with uncertain disturbances

$$\dot{x} = f(x(t)) + gu(t) + g\bar{d}(x) \tag{68}$$

where $f(x) = \begin{bmatrix} x_2 \\ -0.5x_1 - 0.5x_2(1 - x_1^2) \end{bmatrix}$ and $g = [0, 1]^T$.

The two state vectors are denoted as $x = [x_1, x_2]^T$. The constrained control satisfies $u \in \{u \in \mathbb{R} : -2 \leq \|u\| \leq 4\}$. Unknown item related to the disturbances is defined as $\bar{d}(x) = p_1x_1 \sin(x_2 + p_2)$ with $p_1 \in (-2, 2)$ and $p_2 \in (-1, 3)$, which is satisfied with $\|\bar{d}(x)\| \leq \bar{d}_m(x)$. Define the augmented state vector as $\xi = [e_{r1}, e_{r2}, x_{r1}, x_{r2}]^T$. Then, the desired trajectory $x_r = [x_{r1}, x_{r2}]^T$ is provided by

$$\dot{x}_r = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} x_r \tag{69}$$

Give the initial tracking state as $x_r = [1, 0]^T$. The parameters are selected as $Q_0 = 100I_2$ and $R = 0.5I_1$, $\mu = 0.4$, $T = 0.1$, $\varrho \in (0, 1)$. The activation function of the single-critic NN and PE conditions are the same as in Example 1. Then, the approximated weight vectors are denoted as $\hat{\mathcal{W}}_c = [\hat{w}_{c1}, \hat{w}_{c2}, \dots, \hat{w}_{c10}]^T$. For the

Table 2 Performance comparison of three control methods

Methods	e_{r1}	e_{r2}	Sampling states	Minimum interval
SMC [44]	0.1217°	0.1972°/s	1000	0.1s
ADP	0.01°	0.0035°/s	1000	0.1s
ETADP	0.01°	0.0035°/s	247	0.2s

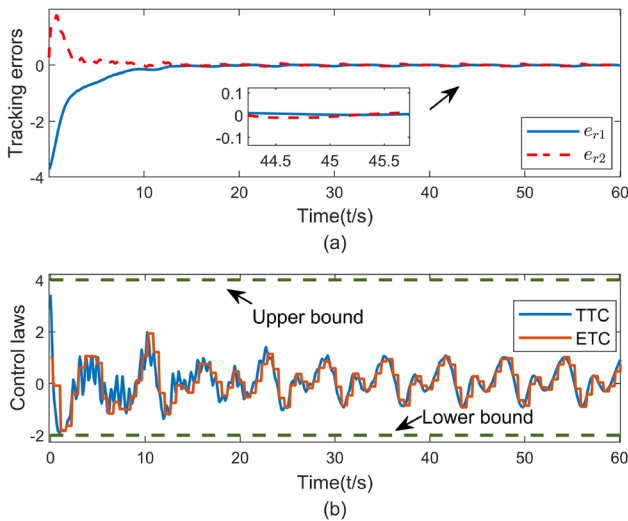


Fig. 7 **a** Tracking errors e_{r1} and e_{r2} . **b** Control laws with asymmetric constraints based on TTC and ETC ($p_1 = 1$ and $p_2 = 0$)

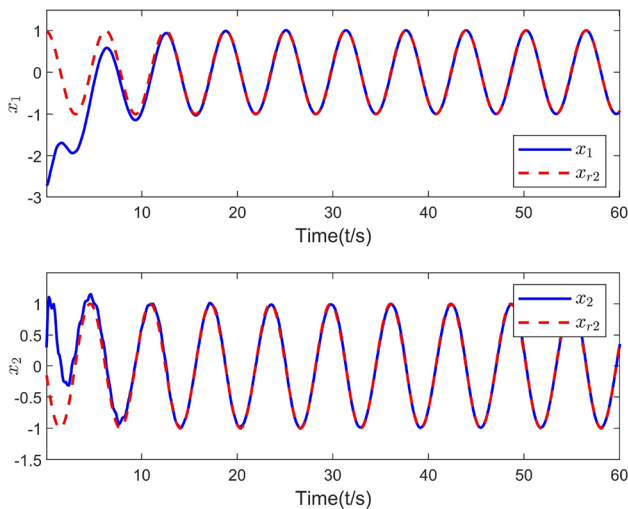


Fig. 8 Actual trajectories x_1, x_2 and desired trajectories x_{r1}, x_{r2} ($p_1 = 1$ and $p_2 = 0$)

determination of the unknown coefficients of the disturbance, we consider two cases as $p_1 = 1, p_2 = 0$ and $p_1 = -1, p_2 = 2$.

Case 1: $p_1 = 1, p_2 = 0$

Give the initial state as $x_0 = [-3, -1]^T$. The error trajectories of the tracking system are described in Fig. 7a. The tracking errors fluctuate around zero and tend to stabilize. The control laws $u(\hat{\xi}_q)$ and $u(\xi)$ are displayed in Fig. 7b. We represent a comparative result of input laws under TTC and ETC, and fewer controller updates under ETC. From Fig. 7b, $u(\hat{\xi}_q)$ and $u(\xi)$ remain within the asymmetric input ranges ($u_{min} = -2, u_{max} = 4$). The tracking performance of the augmented robust system is shown in Fig. 8. In Fig. 9, the single-critic NN weights can

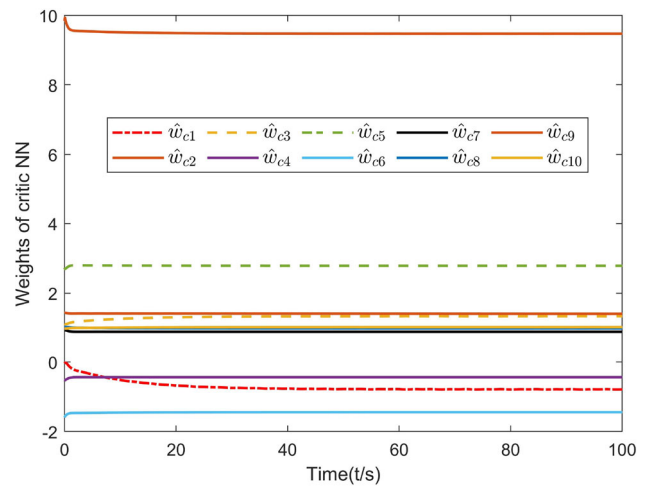


Fig. 9 Weight trajectories of critic network ($p_1 = 1$ and $p_2 = 0$)

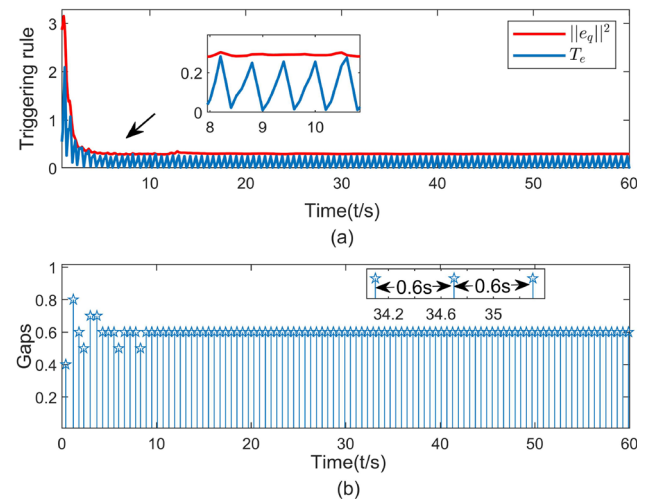


Fig. 10 **a** Triggering rule. **b** Triggering gaps ($p_1 = 1$ and $p_2 = 0$)

be ensured to converge under the action of PE condition after 50s, which maintain them at $[-0.7861, 9.4755, 1.3352, -0.4379, 2.7899, -1.4381, 0.8816, 0.9817, 1.4047, 1.0184]^T$. The trajectories of event error and the triggering threshold are shown in Fig. 10a. It can be seen intuitively that the event errors are always within the triggering thresholds. The interval time between each occurrence of a new event is shown in Fig. 10b. From Table 3, there are only 166 sampling states performed based on ETC. Instead, the system updates its states at every instant under TTC. The minimal triggering interval $(T_q)_{min} > 0$ is valid, and the Zeno behavior did not occur. The trajectory of actual signals and desired signals is displayed in Fig. 11.

Case 2: $p_1 = -1, p_2 = 2$

Give the initial state as $x_0 = [-2, -1.5]^T$. The trajectories of tracking errors are described in Fig. 12a. It is

Table 3 Comparative results of TTC and ETC

Performance	Time-triggered control (TTC)		Event-triggered control (ETC)	
	Case 1	Case 2	Case 1	Case 2
Sampling states	1000	1000	166	219
Minimal interval (s)	0.1	0.1	0.4	0.4
Error e_{r1}	0.0200	0.0154	0.0200	0.0154
Error e_{r2}	0.0325	0.0176	0.0325	0.0176
Average interval (s)	0.1	0.1	0.5977	0.4513

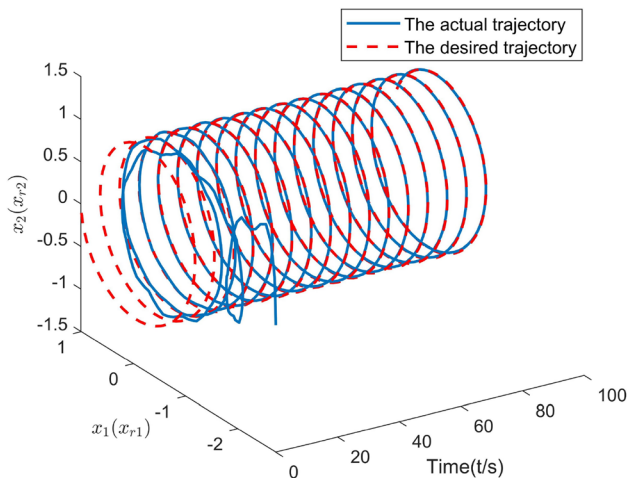


Fig. 11 Actual and desired tracking trajectory ($p_1 = 1$ and $p_2 = 0$)

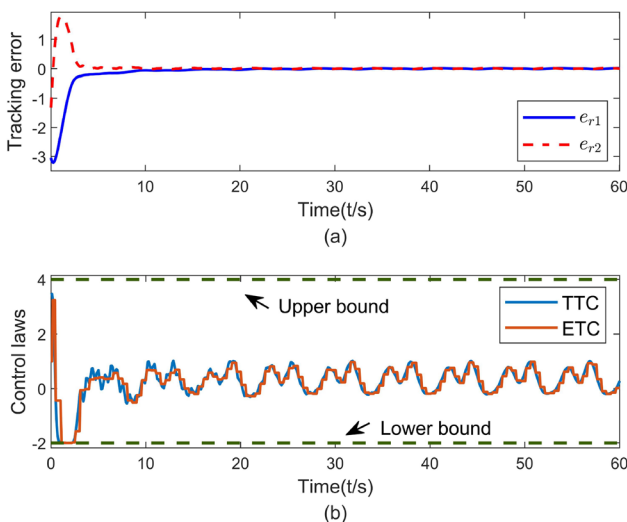


Fig. 12 **a** Tracking errors e_{r1}, e_{r2} . **b** Control laws with asymmetric constraints based on TTC and ETC ($p_1 = -1$ and $p_2 = 2$)

obvious that e_{r1} and e_{r2} gradually converge to asymptotically stable. From Fig. 12b, we can obtain the ETC-based input law $u(\hat{\xi}_q)$ with constrained conditions ($u_{min} = -2, u_{max} = 4$). Compared with TTC, the ETC-based control laws are updated in a non-periodic sampling manner. Figure 13 shows the tracking performance of the

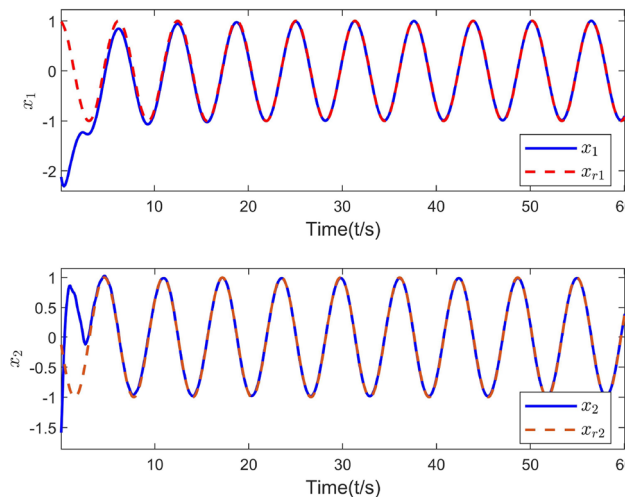


Fig. 13 The actual trajectories x_1, x_2 and desired trajectories x_{r1}, x_{r2} ($p_1 = -1$ and $p_2 = 2$)

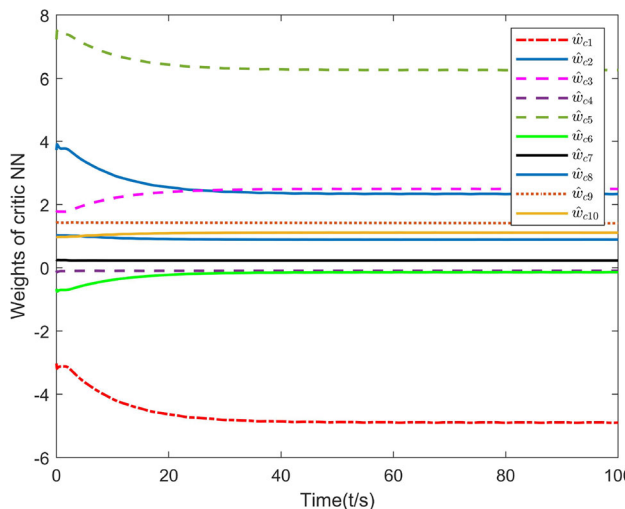


Fig. 14 Weight trajectories of critic NN ($p_1 = -1$ and $p_2 = 2$)

nonlinear system. The NN can be ensured to converge under the action of PE condition. Figure 14 demonstrates the approximal critic weight vectors are converged to $[-4.9850, 2.5005, 6.3542, 0.2073, 1.4214, 2.3553, -0.1263, -0.1712, 0.8905, 1.1092]^T$ after online learning. The relationship between event error and triggering

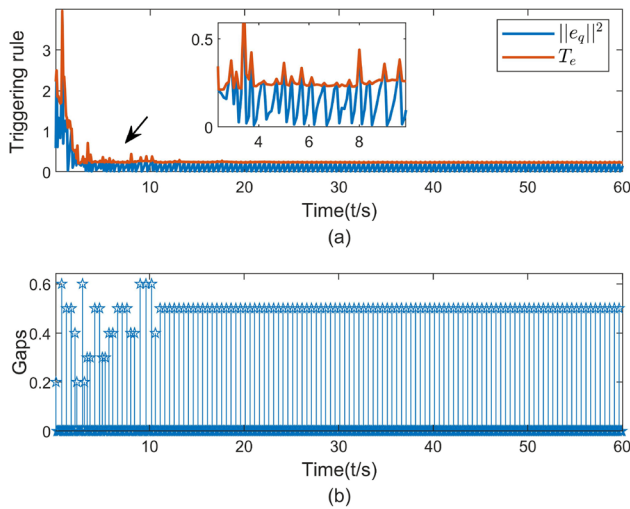


Fig. 15 a Triggering rule. b Triggering gaps ($p_1 = -1$ and $p_2 = 2$)

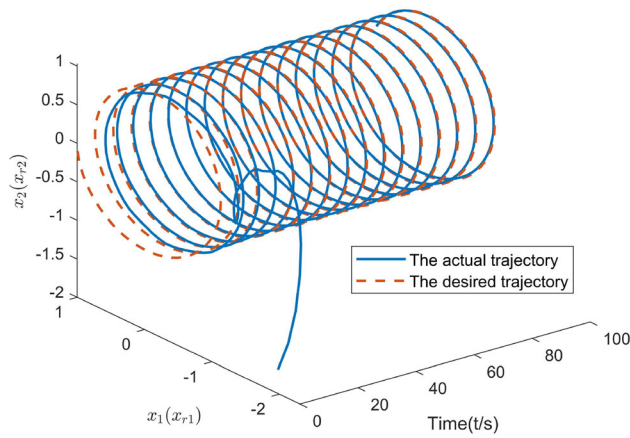


Fig. 16 Actual and desired tracking trajectory ($p_1 = -1$ and $p_2 = 2$)

threshold is shown in Fig. 15a, which demonstrates the effectiveness of the designed triggering condition (23). Figure 15b indicates the sampling gaps, and the number of sampling states is shown in Table 3. Compared with the sampling states of 1000 under TTC, only 219 sampling states are performed based on ETC. The results indicate that ETC can effectively reduce the controller updates, thereby relaxing communication load and reducing computational costs. From Table 3, considering ETC similarly, disturbance in Case 1 requires fewer sampling states, while the tracking accuracy and speed in Case 2 are better than in Case 1. Similarly, the minimum triggering gap exists, so we effectively avoid the Zeno behavior. In Fig. 16, the actual trajectory and the desired trajectory are displayed.

6 Conclusion

An online event-triggered ADP method is suggested to overcome the robust tracking control for nonlinear systems with unknown dynamics and constrained inputs. The value function with the non-quadratic function is constructed to equate the robust tracking problem with the optimal tracking control. Then, a single-critic network is applied to approximate the ideal value function and obtain the solution of the HJB equation for the optimal control law action. The designed ETC-based robust control law is updated in a continuous segmented manner, which significantly reduces the number of controller updates and relaxes computational costs. Two examples are provided to verify the robustness of the proposed algorithm for tracking control systems. However, in practice, complicated nonlinear systems are accompanied by model-free or uncertain mismatches, and solving such problems is our future work.

Author Contributions Peng Binbin performed the experiment and wrote the manuscript; Cui Xiaohong contributed significantly to analysis and manuscript preparation; Zhou Kun perform the analysis with constructive discussions and wrote the manuscript.

Funding This research was funded by the Natural Science Foundation of Zhejiang under Grant LY22F030009, the Funding Project of Basic Scientific Research Business Fee in Zhejiang Province (2022YW20,2022YW84), the National Natural Science Foundation of China under Grant 61903351.

Data availability All data generated or analyzed during this study are included in this article.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

Ethics approval This article does not contain any studies with animals performed by any of the authors.

Consent to participate Informed consent was obtained from all individual.

References

- Gao X, Li T, Yuan L, Bai W (2021) Robust fuzzy adaptive output feedback optimal tracking control for dynamic positioning of marine vessels with unknown disturbances and uncertain dynamics. *International Journal of Fuzzy Systems* 23(7):2283–2296
- Jafarlou F, Peimani M, Lotfivand N (2022) Fractional order adaptive sliding-mode finite time control for cable-suspended parallel robots with unknown dynamics. *International Journal of Dynamics and Control* 10(5):1674–1684
- Yang C, Xia J, Park JH, Shen H, Wang J (2022) Sliding mode control for uncertain active vehicle suspension systems: an event-

- triggered H_∞ control scheme. *Nonlinear Dynamics* 103(4):3209–3221
4. Feng J, Zhang J, Zhang G, Xie S, Ding Y, Liu Z (2021) Uav dynamic path planning based on obstacle position prediction in an unknown environment. *IEEE Access* 9:154679–154691
 5. Yang X, He H (2019) Adaptive critic designs for event-triggered robust control of nonlinear systems with unknown dynamics. *IEEE transactions on cybernetics* 49(6):2255–2267
 6. Cheng L, Wang Z, Jiang F, Li J (2021) Adaptive neural network control of nonlinear systems with unknown dynamics. *Advances in Space Research* 67(3):1114–1123
 7. Cui X, Peng B, Cui Y, Chen W (2023) Event-triggered neural experience replay learning for nonzero-sum tracking games of unknown continuous-time nonlinear systems. *International Journal of Robust and Nonlinear Control* 33(12):1–23
 8. Liu P, Zhang H, Sun J, Tan Z (2022) Event-triggered adaptive integral reinforcement learning method for zero-sum differential games of nonlinear systems with incomplete known dynamics. *Neural Comput & Applic* 34(13), 10775–10786
 9. Shen M, Wang X, Park JH, Yi Y, Che W-W (2023) Extended disturbance-observer-based data-driven control of networked nonlinear systems with event-triggered output. *IEEE Transactions on Systems, Man, and Cybernetics: Systems* 53(5):3129–3140
 10. Abu-Khalaf M, Lewis FL (2005) Nearly optimal control laws for nonlinear systems with saturating actuators using a neural network HJB approach. *Automatica* 41(5):779–791
 11. Yasini S, Sistani MBN, Karimpour A (2015) Approximate dynamic programming for two-player zero-sum game related to H_∞ control of unknown nonlinear continuous-time systems. *International Journal of Control, Automation and Systems* 13:99–109
 12. Mu C, Wang K (2019) Approximate-optimal control algorithm for constrained zero-sum differential games through event-triggering mechanism. *Nonlinear Dynamics* 95:2639–2657
 13. Liu Y, Zhu Q (2022) Adaptive neural network asymptotic control design for MIMO nonlinear systems based on event-triggered mechanism. *Inf. Sci.* 603:91–105
 14. Wang X, Karimi HR, Shen M, Liu D, Li L-W, Shi J (2022) Neural network-based event-triggered data-driven control of disturbed nonlinear systems with quantized input. *Neural Netw.* 156:152–159
 15. Yang D, Li T, Zhang H, Xie X (2019) Event-trigger-based robust control for nonlinear constrained-input systems using reinforcement learning method. *Neurocomputing* 340(7):158–170
 16. Yang X, He H (2019) Adaptive critic designs for event-triggered robust control of nonlinear systems with unknown dynamics. *IEEE Transactions on Cybernetics* 49(6):2255–2267
 17. Yang X, Zhu Y, Dong N, Wei Q (2022) Decentralized event-driven constrained control using adaptive critic designs. *IEEE Transactions on Neural Networks and Learning Systems* 33(10):5830–5844
 18. Liu P, Zhang H, Sun J, Tan Z (2022) Event-triggered adaptive integral reinforcement learning method for zero-sum differential games of nonlinear systems with incomplete known dynamics. *Neural Computing and Applications* 34:10775–10786
 19. Wang D, Mu C, Liu D, Ma H (2018) On mixed data and event driven design for adaptive-critic-based nonlinear H_∞ control. *IEEE transactions on neural networks and learning systems* 29(4):993–1005
 20. Zhu Y, Zhao D, Li X (2017) Iterative adaptive dynamic programming for solving unknown nonlinear zero-sum game based on online data. *IEEE Transactions on Neural Networks and Learning Systems* 28(3):714–725
 21. Mu C, Zhang Y (2020) Learning-based robust tracking control of quadrotor with time-varying and coupling uncertainties. *IEEE Transactions on Neural Networks and Learning Systems* 31(1):259–273
 22. Gao W, Jiang Z-P (2017) Learning-based adaptive optimal tracking control of strict-feedback nonlinear systems. *IEEE transactions on neural networks and learning systems* 29(6):2614–2624
 23. Modares H, Lewis FL (2014) Optimal tracking control of nonlinear partially-unknown constrained-input systems using integral reinforcement learning. *Automatica* 50(7):1780–1792
 24. Liu C, Zhang H, Xiao G, Sun S (2019) Integral reinforcement learning based decentralized optimal tracking control of unknown nonlinear large-scale interconnected systems with constrained-input. *Neurocomputing* 323:1–11
 25. Zhang H, Zhang K, Xiao G, Jiang H (2019) Robust optimal control scheme for unknown constrained-input nonlinear systems via a plug-n-play event-sampled critic-only algorithm. *IEEE Transactions on Systems, Man, and Cybernetics: Systems* 50(9):3169–3180
 26. Zhang S, Liu D, Zhang Y, Zhao B (2021) Observer-based event-triggered robust control for input constrained uncertain nonlinear systems. 2021 36th Youth Academic Annual Conference of Chinese Association of Automation (YAC), 865–870
 27. Song R, Liu L, Xia L, Lewis FL (2022) Online optimal event-triggered H_∞ control for nonlinear systems with constrained state and input. *IEEE Transactions on Systems, Man, and Cybernetics: Systems* 53(1):131–141
 28. Wu Q, Zhao B, Liu D, Polycarpou MM (2023) Event-triggered adaptive dynamic programming for decentralized tracking control of input constrained unknown nonlinear interconnected systems. *Neural Networks* 157:336–349
 29. Zhang H, Cui X, Luo Y, Jiang H (2018) Finite-horizon H_∞ tracking control for unknown nonlinear systems with saturating actuators. *IEEE transactions on neural networks and learning systems* 29(4):1200–1212
 30. Wang X, Karimi HR, Shen M, Liu D, Li L-W, Shi J (2022) Neural network-based event-triggered data-driven control of disturbed nonlinear systems with quantized input. *Neural Netw.* 156:152–159
 31. Wei Q, Song R, Yan P (2016) Data-driven zero-sum neuro-optimal control for a class of continuous-time unknown nonlinear systems with disturbance using ADP. *IEEE transactions on neural networks and learning systems* 27(2):444–458
 32. Cao L, Cheng Z, Liu Y, Li H (2022) Event-based adaptive NN fixed-time cooperative formation for multiagent systems. *IEEE Transactions on Neural Networks and Learning Systems*, 1–11
 33. Liu D, Huang Y, Wang D, Wei Q (2013) Neural-network-observer-based optimal control for unknown nonlinear systems using adaptive dynamic programming. *International Journal of Control* 86(9):1554–1566
 34. Fan B, Yang Q, Tang X, Sun Y (2018) Robust ADP design for continuous-time nonlinear systems with output constraints. *IEEE transactions on neural networks and learning systems* 29(6):2127–2138
 35. Xue S, Luo B, Liu D (2020) Event-triggered adaptive dynamic programming for zero-sum game of partially unknown continuous-time nonlinear systems. *IEEE Transactions on Systems, Man, and Cybernetics: Systems* 50(9):3189–3199
 36. Hu C, Zou Y, Li S (2023) Observer-based event-triggered ADP approach for input-constrained nonlinear systems with disturbances. *International Journal of Robust and Nonlinear Control* 10(9):5179–5197
 37. Mu C, Wang D, Sun C, Zong Q (2017) Robust adaptive critic control design with network-based event-triggered formulation. *Nonlinear Dynamics* 90:2023–2035
 38. Narayanan V, Jagannathan S (2018) Event-triggered distributed approximate optimal state and output control of affine nonlinear

- interconnected systems. *IEEE Transactions on Neural Networks and Learning Systems* 29(7):2846–2856
39. Wang T, Wang H, Xu N, Zhang L, Alharbi KH (2023) Sliding-mode surface-based decentralized event-triggered control of partially unknown interconnected nonlinear systems via reinforcement learning. *Inf. Sci.* 641:119070
 40. Zhao J, Na J, Gao G (2022) Robust tracking control of uncertain nonlinear systems with adaptive dynamic programming. *Neurocomputing* 471:21–30
 41. Modares H, Lewis FL (2014) Optimal tracking control of nonlinear partially-unknown constrained-input systems using integral reinforcement learning. *Automatica* 50(7):1780–1792
 42. Mu C, Wang K, Qiu T (2022) Dynamic event-triggering neural learning control for partially unknown nonlinear systems. *IEEE transactions on cybernetics* 52(4):2200–2213
 43. Vamvoudakis KG, Lewis FL (2010) Online actor-critic algorithm to solve the continuous-time infinite horizon optimal control problem. *Automatica* 46(5):878–888
 44. Dao Q-T, Mai D-H, Nguyen D-K, Ly N-T (2022) Adaptive parameter integral sliding mode control of pneumatic artificial muscles in antagonistic configuration. *Journal of Control, Automation and Electrical Systems* 33:1116–1124

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.