Neurocomputing 214 (2016) 848–856

Contents lists available at ScienceDirect

Neurocomputing

journal homepage: www.elsevier.com/locate/neucom

Event-based input-constrained nonlinear H∞ state feedback with adaptive critic and neural implementation ☆

Ding Wang [a,b], Chaoxu Mu [b,*], Qichao Zhang [a], Derong Liu [c]

[a] The State Key Laboratory of Management and Control for Complex Systems, Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China
[b] Tianjin Key Laboratory of Process Measurement and Control, School of Electrical Engineering and Automation, Tianjin University, Tianjin 300072, China
[c] School of Automation and Electrical Engineering, University of Science and Technology Beijing, Beijing 100083, China

ARTICLE INFO

Article history:
Received 11 April 2016
Received in revised form
16 June 2016
Accepted 13 July 2016
Communicated by Bo Shen
Available online 20 July 2016

Keywords:
Adaptive critic learning (ACL)
Adaptive dynamic programming (ADP)
Event-based control
Hamilton–Jacobi–Isaacs (HJI) equation
Input constraints
Neural networks
Nonlinear H_∞ control
State feedback

ABSTRACT

In this paper, the continuous-time input-constrained nonlinear H_∞ state feedback control under event-based environment is investigated with adaptive critic designs and neural network implementation. The nonlinear H_∞ control issue is regarded as a two-player zero-sum game that requires solving the Hamilton–Jacobi–Isaacs equation and the adaptive critic learning (ACL) method is adopted toward the event-based constrained optimal regulation. The novelty lies in that the event-based design framework is combined with the ACL technique, thereby carrying out the input-constrained nonlinear H_∞ state feedback via adopting a non-quadratic utility function. The event-based optimal control law and the time-based worst-case disturbance law are derived approximately, by training an artificial neural network called a critic and eventually learning the optimal weight vector. Under the action of the event-based state feedback controller, the closed-loop system is constructed with uniformly ultimately bounded stability analysis. Simulation studies are included to verify the theoretical results as well as to illustrate the event-based H_∞ control performance.

1. Introduction

When coping with nonlinear optimal regulation designs during various control practices, we always encounter the difficulty of solving the Hamilton–Jacobi–Bellman (HJB) equation [1]. In particular, the input-constrained control design is a more complicated problem than the traditional unconstrained design task [2,3]. Though dynamic programming is deemed as a basic strategy to handle optimal control problems, there still exists a serious issue called the “curse of dimensionality”. Similarly, from the point of minimax optimization, the H_∞ control problem can be formulated as a two-player zero-sum differential game. In order to obtain a controller that minimizes the cost function in the worst-case disturbance, it is required to find the Nash equilibrium solution by dealing with the Hamilton–Jacobi–Isaacs (HJI) equation. Nevertheless, it is also intractable to gain an analytic solution of the HJI equation in case of nonlinear systems. Therefore, various approximate methods have been proposed to overcome the difficulty in handling the HJB and HJI equations. Among that, by involving neural networks for function approximation, the adaptive or approximate dynamic programming (ADP) was founded by Werbos [4] to solve optimal control problems forward-in-time. As Lewis and Liu [5] stated, there exists a fundamental idea in ADP which is similar as designing advanced adaptive systems with neural network technique (see, e.g., [6–9]). Note that therein, various nonlinearities, such as uncertain dynamics, input saturation, and dead-zone input, were considered and handled by constructing powerfully adaptive and neural systems. Hence, it is greatly important to understand and construct more intelligent adaptive systems, especially optimal adaptive systems, with the help of ADP methodology. Actually, it is observed that ADP and related fields have gained much development in various topics, such as adaptive optimal regulation [10–13], optimal tracking control [14–17], robust optimal control [18–20] and so on. Recently, the nonlinear H_∞ control and the non-zero-sum game have also been paid special attention under ADP framework [21–26]. For example, Abu-Khalaf

☆This work was supported in part by the National Natural Science Foundation of China under Grants 61233001, 61273136, 61273140, 61304018, 61304086, 61533017, 61573353, U1501251, and 61411130160, in part by Beijing Natural Science Foundation under Grant 4162065, in part by Tianjin Natural Science Foundation under Grant 14JCQNJC05400, in part by Research Fund of Tianjin Key Laboratory of Process Measurement and Control (TKLPMC-201612), and in part by the Early Career Development Award of SKLMCCS.

* Corresponding author.
E-mail addresses: ding.wang@ia.ac.cn (D. Wang), cxmu@tju.edu.cn (C. Mu), zhangqichao2014@ia.ac.cn (Q. Zhang), derong@ustb.edu.cn (D. Liu).

http://dx.doi.org/10.1016/j.neucom.2016.07.002
0925-2312/

et al. [21] performed policy iterations on both players (constrained control law and disturbance law), so as to solve the HJI equation of H_∞ state feedback control problem with input saturation. Liu et al. [25] studied the neural-network-based zero-sum game for a class of nonlinear discrete-time systems via the iterative ADP algorithm. Note that the above results are all derived with the traditional time-based design manner.

When designing automatic control systems, especially distributed and networked systems, the event-based approach has gained much attention in the last decade, since people can benefit greatly from it in terms of decreasing the computation complexity and enhancing the control efficiency [27–34]. Eqtami et al. [28] proposed an event-based strategy for state feedback control of discrete-time nonlinear systems. Tallapragada and Chopra [29,30] developed an event-based control algorithm for trajectory tracking of continuous-time nonlinear systems. Li et al. [31,32] and Dong et al. [33] studied the event-triggered state estimation and synchronization control for complex networks with the involvement of time-varying delays, uncertain inner coupling, and state-dependent noises. Ma et al. [34] constructed both centralized and decentralized event-triggered control protocols for group consensus to cope with energy consumption and communication constraint that may be encountered in physical implementations. In particular, the combination of event-based mechanism and ADP method provides a novel channel for achieving advanced nonlinear optimal control with adaptivity [35–38]. Among them, Vamvoudakis [35] originally proposed an event-based adaptive optimal control strategy for continuous-time affine nonlinear systems. Note that under the new framework, the ADP-based controller is updated only when an event is triggered, thereby reducing the computational burden of learning and control. However, the existing research is conducted either for nonlinear regulation problem, or without considering the input constraints, which calls for an extension to input-constrained zero-sum game design under event-based formulation. This motivates our research. Note that the main difficulty and challenge of introducing the event-based framework is how to conduct the critic learning and analyze the closed-loop stability in case that the event-based state vector is taken into consideration.

This paper focuses on the event-based input-constrained nonlinear H_∞ state feedback control with the idea of ADP. In order to emphasize the ability of adaptivity and self-learning, we call the ADP architecture established here as adaptive critic learning (ACL). The main contribution of this paper lies in that the event-based design framework is combined with the ACL technique, so as to accomplish the input-constrained nonlinear H_∞ state feedback. The rest of this paper is organized as follows. A brief description of the input-constrained nonlinear H_∞ control problem is provided in Section 2. The ACL methodology for the event-based input-constrained nonlinear H_∞ control is developed in Section 3 with closed-loop stability analysis. The simulation studies and the concluding remarks are presented in Sections 4 and 5, respectively.

For convenience, the following notations will be utilized throughout the paper. $\mathbb{R}$ represents the set of all real numbers. $\mathbb{R}^n$ is the Euclidean space of all n-dimensional real vectors. $\mathbb{R}^{n\times m}$ is the space of all $n\times m$ real matrices. $\|\cdot\|$ denotes the vector norm of a vector in $\mathbb{R}^n$ or the matrix norm of a matrix in $\mathbb{R}^{n\times m}$. I_n represents the $n\times n$ identity matrix. $\lambda_{\max}(\cdot)$ and $\lambda_{\min}(\cdot)$ calculate the maximal and minimal eigenvalues of a matrix, respectively. C^κ denotes the class of functions having continuous κ-th derivative. Let Ω be a compact subset of $\mathbb{R}^n$ and $\Psi(\Omega)$ be the set of admissible controls on Ω. $\mathbb{N}=\{0,1,2,\ldots\}$ denotes the set of all non-negative integers. In addition, the superscript "T" is taken for representing the transpose operation and $\nabla(\cdot)\triangleq\partial(\cdot)/\partial x$ is employed to denote the gradient operator.

2. Problem description of the input-constrained nonlinear H_∞ control problem

Let us consider a class of continuous-time nonlinear systems with external perturbations described by

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

$$z(t)=Q(x(t)), \tag{1b}$$

where $x(t)\in\Omega\subset\mathbb{R}^n$ is the state vector, $u(t)\in\Omega_u\subset\mathbb{R}^m$ is the control vector, $v(t)\in\mathbb{R}^q$ is the perturbation vector with $v(t)\in L_2[0,\infty)$, $z(t)\in\mathbb{R}^p$ is the objective output, and $f(\cdot)$, $g(\cdot)$, and $h(\cdot)$ are differentiable in their arguments with $f(0)=0$. The constrained control set is defined as $\Omega_u=\{u\in\mathbb{R}^m: |u_i|<\bar{u}_i, i=1,2,\ldots,m\}$. We let $x(0)=x_0$ be the initial state and $x=0$ be the equilibrium point of the controlled plant.

Assumption 1. The system function $f(x)$ is Lipschitz continuous on a set Ω in $\mathbb{R}^n$ containing the origin and the system (1a) is controllable.

With Assumption 1, for nonlinear H_∞ control, it is aimed at obtaining a state feedback control law $u=u(x)$ such that the closed-loop form of system (1a) is asymptotically stable, and simultaneously, has L_2-gain no larger than ϱ, that is

$$\int_0^\infty\left(\|Q(x)\|^2+Y(u)\right)\mathrm{d}\tau\le\varrho^2\int_0^\infty v^{\mathrm{T}}(\tau)Pv(\tau)\mathrm{d}\tau, \tag{2}$$

where $\|Q(x)\|^2=x^{\mathrm{T}}(t)Qx(t)$ and Q and P are symmetric positive definite matrices with appropriate dimensions. If the condition (2) is satisfied, the closed-loop system is said to have L_2-gain no larger than ϱ. For unconstrained control problem, we often select a quadratic utility regarding to u as $Y(u)=u^{\mathrm{T}}(t)Ru(t)$ with R being a symmetric positive definite matrix. However, for input-constrained control problem, inspired by [2,21], a non-quadratic utility is adopted by choosing

$$Y(u)=2\int_0^u\varphi^{-\mathrm{T}}(\varsigma)R\mathrm{d}\varsigma, \tag{3}$$

where $\varphi(\cdot)\in\mathbb{R}^m$ is a m-dimensional function, $\varphi^{-\mathrm{T}}$ denotes $(\varphi^{-1})^{\mathrm{T}}$, and $\varphi^{-1}(\varsigma)=(\varphi_1^{-1}(\varsigma_1),\varphi_2^{-1}(\varsigma_2),\ldots,\varphi_m^{-1}(\varsigma_m))^{\mathrm{T}}$. Meanwhile, $\varphi_i(\cdot)$ is a strictly monotonic odd function satisfying $|\varphi_i(\cdot)|<1$ $(i=1,2,\ldots,m)$ and belonging to C^κ $(\kappa\ge1)$ and $L_2(\Omega)$.

Remark 1. It is important to point out that this kind of non-quadratic utility is a nominal choice in light of the literature, such as [2,21]. Clearly, $Y(u)$ is positive definite since $\varphi_i(\cdot)$ is a monotonic odd function, for instance, $\varphi_i(\cdot)=\tanh(\cdot)$.

As is known, the H_∞ control problem can be formulated as a two-player zero-sum differential game, where the control input is a minimizing player while the disturbance is a maximizing one [21,23,24,38]. Note that the solution of H_∞ control problem is the saddle point of zero-sum game theory, denoted as (u^*,v^*), where u^* and v^* are the optimal control and the worst-case disturbance, respectively.

Define the infinite horizon cost function as

$$J(x(t),u,v)=\int_t^\infty U(x(\tau),u(\tau),v(\tau))\mathrm{d}\tau, \tag{4}$$

where

$$U(x,u,v)=x^{\mathrm{T}}Qx+Y(u)-\varrho^2v^{\mathrm{T}}Pv$$

represents the utility function. For the two-player zero-sum differential game, our goal is to find the feedback saddle point solution (u^*,v^*), such that the following Nash condition holds:

$$J^*(x_0)=\min_u\max_v J(x_0,u,v)=\max_v\min_u J(x_0,u,v).$$

Note that the proposed two-player optimal control problem has a unique solution if the above condition is satisfied [22].

For any admissible control law $u \in \Psi(\Omega)$, if the associated cost function (4) is continuously differentiable, then its infinitesimal version is the nonlinear Lyapunov equation

$$0 = U(x, u, v) + (\nabla J(x))^{\mathrm{T}}(f(x) + g(x)u + h(x)v)$$

with $J(0) = 0$. Define the Hamiltonian of system (1a) as

$$H(x, u, v, \nabla J(x)) = U(x, u, v) + (\nabla J(x))^{\mathrm{T}}(f(x) + g(x)u + h(x)v). \quad (5)$$

Based on optimal control theory, the optimal cost function $J^*(x)$ satisfies the HJI equation

$$0 = \min_u \max_v H(x, u, v, \nabla J^*(x)). \quad (6)$$

The saddle point solution (u^*, v^*) satisfies the following two stationary conditions:

$$\frac{\partial H(x, u, v, \nabla J^*(x))}{\partial u} = 0; \quad (7a)$$

$$\frac{\partial H(x, u, v, \nabla J^*(x))}{\partial v} = 0. \quad (7b)$$

Combining (6) with (7) yields the optimal control law and the worst-case disturbance law as follows:

$$u^*(x) = -\varphi\left(\frac{1}{2}R^{-1}g^{\mathrm{T}}(x)\nabla J^*(x)\right); \quad (8a)$$

$$v^*(x) = \frac{1}{2\varrho^2}P^{-1}h^{\mathrm{T}}(x)\nabla J^*(x). \quad (8b)$$

Clearly, the formula (8) reveals that

$$(\nabla J^*(x))^{\mathrm{T}}g(x) = -2\varphi^{-\mathrm{T}}(u^*(x))R; \quad (9a)$$

$$(\nabla J^*(x))^{\mathrm{T}}h(x) = 2\varrho^2 v^{*\mathrm{T}}(x)P. \quad (9b)$$

Using the expressions of (8) and (9b), the HJI equation (6), i.e., $H(x, u^*, v^*, \nabla J^*(x)) = 0$, turns to be the following form:

$$\begin{aligned}
0 &= x^{\mathrm{T}}Qx + Y(u^*) - \varrho^2 v^{*\mathrm{T}}(x)Pv^*(x) \\
&\quad + (\nabla J^*(x))^{\mathrm{T}}(f(x) + g(x)u^*(x) + h(x)v^*(x)) \\
&= x^{\mathrm{T}}Qx + Y(u^*) + (\nabla J^*(x))^{\mathrm{T}}(f(x) + g(x)u^*(x)) + \varrho^2 v^{*\mathrm{T}}(x)Pv^*(x) \\
&= x^{\mathrm{T}}Qx + Y(u^*) + (\nabla J^*(x))^{\mathrm{T}}(f(x) + g(x)u^*(x)) \\
&\quad + \frac{1}{4\varrho^2}(\nabla J^*(x))^{\mathrm{T}}h(x)P^{-1}h^{\mathrm{T}}(x)\nabla J^*(x), J^*(0) = 0.
\end{aligned} \quad (10)$$

Note that (10) is the traditional time-based HJI equation, which is difficult to solve in theory. This motivates us to pursue an alternate avenue to overcome the difficulty. Fortunately, the ADP method is an effective channel to solve the nonlinear optimization and optimal control problems. Hence, in the sequel, we construct a neural-network-based optimal control method with event-based design fashion to handle the input-constrained nonlinear H_∞ control problem.

3. Event-based input-constrained nonlinear H_∞ state feedback control design with ACL

In this section, we develop the event-based input-constrained nonlinear H_∞ control law with ACL, which includes three parts: event-based formulation of the optimal control design, neural network implementation of the ACL controller, and stability analysis of the closed-loop system.

3.1. Event-based formulation of the optimal control design

In the general framework of the event-based control design, we define a monotonically increasing sequence of triggering instants $\{s_j\}_{j=0}^{\infty}$, where s_j represents the jth consecutive sampling instant satisfying $s_j < s_{j+1}$ with $j \in \mathbb{N}$. Then, the output of the sampled-data component is a sequence of sampled state denoted as $x(s_j) \triangleq \hat{x}_j$ for all $t \in [s_j, s_{j+1})$. The event-triggered error function between the current state and the sampled state is defined as

$$e_j(t) = \hat{x}_j - x(t), \quad \forall\, t \in [s_j, s_{j+1}). \quad (11)$$

Usually, under the event-based formulation, the triggering instants are determined by a certain triggering condition. We say an event is triggered if it is not satisfied at $t = s_j$. The triggering condition is often established in terms of the event-triggered error and a state-dependent threshold. At every triggering instant, the system state is sampled that resets the event-triggered error $e_j(t)$ to zero, and accordingly, the state feedback controller $u(x(s_j)) = u(\hat{x}_j) \triangleq \mu(\hat{x}_j)$ is updated. Note that the control $\mu(\hat{x}_j)$ is a function of the event-based state vector rather than a time-based manner and the control sequence $\{\mu(\hat{x}_j)\}_{j=0}^{\infty}$ becomes a continuous-time signal via the function of a zero-order hold. Hence, this control signal can actually be regarded as a piecewise constant function and during any time interval $[s_j, s_{j+1})$, it is $\mu(\hat{x}_j)$ with $j \in \mathbb{N}$.

Using the fact that $\hat{x}_j = x(t) + e_j(t)$ and applying the control signal $\mu(\hat{x}_j)$, the closed-loop form of system (1a) turns to be the following form:

$$\begin{aligned}
\dot{x}(t) &= f(x(t)) + g(x(t))\mu(\hat{x}_j) + h(x(t))v(x(t)) \\
&= f(x(t)) + g(x(t))\mu(x(t) + e_j(t)) + h(x(t))v(x(t)), \\
&\quad \forall\, t \in [s_j, s_{j+1}).
\end{aligned} \quad (12)$$

With the event-based formulation, the state feedback control law (8a) becomes

$$\mu^*(\hat{x}_j) = -\varphi\left(\frac{1}{2}R^{-1}g^{\mathrm{T}}(\hat{x}_j)\nabla J^*(\hat{x}_j)\right). \quad (13)$$

Note that in this paper, we keep the disturbance law unchanged between the time/event transformation. Then, applying the time-based worst-case disturbance law (8b) and the event-based optimal control law (13), the event-based HJI equation can be written as

$$\begin{aligned}
&H(x, \mu^*(\hat{x}_j), v^*(x), \nabla J^*(x)) \\
&\quad = x^{\mathrm{T}}Qx + Y(\mu^*(\hat{x}_j)) + (\nabla J^*(x))^{\mathrm{T}}(f(x) + g(x)\mu^*(\hat{x}_j)) \\
&\quad + \frac{1}{4\varrho^2}(\nabla J^*(x))^{\mathrm{T}}h(x)P^{-1}h^{\mathrm{T}}(x)\nabla J^*(x), \quad J^*(0) = 0,
\end{aligned}$$

which, generally, is not equal to zero and thus is not same as the time-based HJI equation (10).

For the event-based control design, we make the following assumption and then develop the main theorem for stating the asymptotic stabilization.

Assumption 2 (*cf.* [35]). The feedback controller $u(x)$ is Lipschitz continuous with respect to the event-triggered error $e_j(t)$ such that $\| u(x(t)) - u(\hat{x}_j)\| \le \mathcal{L}_1 \| e_j(t)\|$, where $\mathcal{L}_1$ is a positive real constant.

Theorem 1. *For the nonlinear system (1a) with cost function (4), the sampled-data closed-loop system is formed as (12). The disturbance law is given by (8b) while the sampled-data control law is developed by (13) for all $t \in [s_j, s_{j+1})$ with $j \in \mathbb{N}$. In case that the triggering* *condition is defined as*

$$\| e_j(t)\|^2 \le \frac{(1-\eta_1^2)\lambda_{\min}(Q)\| x\|^2 + Y(u^*) - \lambda_1 - \varrho^2 v^{*\mathrm{T}}(x)Pv^*(x)}{\lambda_{\max}(R^{\mathrm{T}}R)\mathcal{L}_1^2} \triangleq e_T, \tag{14}$$

where e_T is the positive threshold and $\eta_1 \in (0, 1)$ is a designed parameter of the sample frequency, then the closed-loop system (12) is asymptotically stable.

Proof. Choose $L_1(t) = J^*(x(t))$ as the Lyapunov function. With the disturbance law (8b) and the sampled-data control law (13), we take the derivative of the Lyapunov function $L_1(t)$ along the trajectory of system (12) and obtain that $\dot{L}_1(t) = \mathrm{d}J^*(x(t))/\mathrm{d}t$ equals to

$$\dot{L}_1(t) = (\nabla J^*(x))^{\mathrm{T}}\Big(f(x) + g(x)\mu^*(\hat{x}_j) + h(x)v^*(x)\Big). \tag{15}$$

According to the time-based HJI equation (10), we have

$$\begin{aligned}(\nabla J^*(x))^{\mathrm{T}}(f(x) + h(x)v^*(x)) = {} & -x^{\mathrm{T}}Qx - Y(u^*) - (\nabla J^*(x))^{\mathrm{T}}g(x)u^*(x) \\ & + \varrho^2 v^{*\mathrm{T}}(x)Pv^*(x)\end{aligned}$$

which, combined with (9a), yields the reduction of (15) to

$$\begin{aligned}\dot{L}_1(t) = {} & -x^{\mathrm{T}}Qx - Y(u^*) + 2\varphi^{-\mathrm{T}}(u^*(x))R\Big(u^*(x) - \mu^*(\hat{x}_j)\Big) \\ & + \varrho^2 v^{*\mathrm{T}}(x)Pv^*(x)\end{aligned}$$

By applying the Cauchy–Schwarz inequality to the term $2\varphi^{-\mathrm{T}}(u^*(x))R\Big(u^*(x) - \mu^*(\hat{x}_j)\Big)$ and using Assumption 2, we further obtain

$$\begin{aligned}\dot{L}_1(t) \le {} & -x^{\mathrm{T}}Qx - Y(u^*) + \|\varphi^{-1}(u^*(x))\|^2 + \lambda_{\max}(R^{\mathrm{T}}R)\| u^*(x) \\ & - \mu^*(\hat{x}_j)\|^2 + \varrho^2 v^{*\mathrm{T}}(x)Pv^*(x) \\ \le {} & -\eta_1^2\lambda_{\min}(Q)\| x\|^2 + (\eta_1^2 - 1)\lambda_{\min}(Q)\| x\|^2 - Y(u^*) + \lambda_1 \\ & + \lambda_{\max}(R^{\mathrm{T}}R)\mathcal{L}_1^2\| e_j(t)\|^2 + \varrho^2 v^{*\mathrm{T}}(x)Pv^*(x),\end{aligned} \tag{16}$$

where the bounded term λ_1 is introduced when considering the inequality

$$\|\varphi^{-1}(u^*(x))\|^2 = \sum_{i=1}^{m}\Big(\varphi_i^{-1}(u_i^*(x))\Big)^2 \le \sum_{i=1}^{m}\Big(\varphi_i^{-1}(\bar{u}_i)\Big)^2 \triangleq \lambda_1. \tag{17}$$

In case the triggering condition (14) holds, (16) implies that $\dot{L}_1(t) \le -\eta_1^2\lambda_{\min}(Q)\| x\|^2 < 0$ for any $x \ne 0$. As a result, the conditions for Lyapunov local stability theory are satisfied, which ends the proof. □

3.2. Neural network implementation of the ACL controller

In neural network implementation, let us denote l_c as the number of neurons in the hidden layer. According to the universal approximation property, the cost function $J(x)$ can be reconstructed by a neural network with a single hidden layer on a compact set Ω as $J(x) = \omega_c^{\mathrm{T}}\sigma_c(x) + \varepsilon_c(x)$, where $\omega_c \in \mathbb{R}^{l_c}$ is the ideal weight vector, $\sigma_c(x) \in \mathbb{R}^{l_c}$ is the activation function, and $\varepsilon_c(x) \in \mathbb{R}$ is the reconstruction error. Then, the gradient vector is

$$\nabla J(x) = (\nabla\sigma_c(x))^{\mathrm{T}}\omega_c + \nabla\varepsilon_c(x). \tag{18}$$

Since the ideal weight is unknown, a critic neural network is built to approximate the cost function as $\hat{J}(x) = \hat{\omega}_c^{\mathrm{T}}\sigma_c(x)$, where $\hat{\omega}_c \in \mathbb{R}^{l_c}$ denotes the estimated weight vector. Similarly, we have the gradient vector

$$\nabla\hat{J}(x) = (\nabla\sigma_c(x))^{\mathrm{T}}\hat{\omega}_c. \tag{19}$$

Under the circumstance of neural network expression, according to (8b), (13), and (18), the event-based optimal control law and the time-based worst-case disturbance law are written as

$$\mu(\hat{x}_j) = -\varphi\bigg(\frac{1}{2}R^{-1}g^{\mathrm{T}}(\hat{x}_j)\Big((\nabla\sigma_c(\hat{x}_j))^{\mathrm{T}}\omega_c + \nabla\varepsilon_c(\hat{x}_j)\Big)\bigg); \tag{20a}$$

$$v(x) = \frac{1}{2\varrho^2}h^{\mathrm{T}}(x)\Big((\nabla\sigma_c(x))^{\mathrm{T}}\omega_c + \nabla\varepsilon_c(x)\Big). \tag{20b}$$

By combining (8b) and (13) with (19), the approximate values of the event-based optimal control law (20a) and time-based worst-case disturbance law (20b) are

$$\hat{\mu}(\hat{x}_j) = -\varphi\bigg(\frac{1}{2}R^{-1}g^{\mathrm{T}}(\hat{x}_j)(\nabla\sigma_c(\hat{x}_j))^{\mathrm{T}}\hat{\omega}_c\bigg); \tag{21a}$$

$$\hat{v}(x) = \frac{1}{2\varrho^2}h^{\mathrm{T}}(x)(\nabla\sigma_c(x))^{\mathrm{T}}\hat{\omega}_c. \tag{21b}$$

As for the Hamiltonian (5), when taking the neural network expression (18) into account, it becomes

$$\begin{aligned}H\Big(x, \mu(\hat{x}_j), v(x), \omega_c\Big) = {} & U(x, \mu(\hat{x}_j), v(x)) \\ & + \omega_c^{\mathrm{T}}\nabla\sigma_c(x)\Big(f(x) + g(x)\mu(\hat{x}_j) + h(x)v(x)\Big) \triangleq e_{cH},\end{aligned} \tag{22}$$

where

$$e_{cH} = -(\nabla\varepsilon_c(x))^{\mathrm{T}}\Big(f(x) + g(x)\mu(\hat{x}_j) + h(x)v(x)\Big)$$

represents the residual error due to the neural network approximation. Using (19), the approximate Hamiltonian is

$$\begin{aligned}\hat{H}\Big(x, \mu(\hat{x}_j), v(x), \hat{\omega}_c\Big) = {} & U(x, \mu(\hat{x}_j), v(x)) \\ & + \hat{\omega}_c^{\mathrm{T}}\nabla\sigma_c(x)\Big(f(x) + g(x)\mu(\hat{x}_j) + h(x)v(x)\Big) \triangleq e_c.\end{aligned} \tag{23}$$

Let the error vector between the ideal weight and the estimated value be $\tilde{\omega}_c = \omega_c - \hat{\omega}_c$. Combining (22) with (23) yields

$$e_c = -\tilde{\omega}_c^{\mathrm{T}}\nabla\sigma_c(x)\Big(f(x) + g(x)\mu(\hat{x}_j) + h(x)v(x)\Big) + e_{cH}. \tag{24}$$

For training the critic network, we design $\hat{\omega}_c$ to minimize the objective function $E_c = (1/2)e_c^{\mathrm{T}}e_c$. Note that the approximated control law (21a) and disturbance law (21b) are often used during the learning process because of the unavailability of the optimal control law $\mu(\hat{x}_j)$ and the worst-case disturbance law $v(x)$. Based on (23), we adopt the normalized steepest descent algorithm to tune the weight vector as follows:

$$\begin{aligned}\dot{\hat{\omega}}_c = {} & -\alpha_c\frac{1}{(1+\phi^{\mathrm{T}}\phi)^2}\left(\frac{\partial E_c}{\partial\hat{\omega}_c}\right) \\ = {} & -\alpha_c\frac{\phi}{(1+\phi^{\mathrm{T}}\phi)^2}\Big(U(x, \hat{\mu}(\hat{x}_j), \hat{v}(x)) + \phi^{\mathrm{T}}\hat{\omega}_c\Big),\end{aligned} \tag{25}$$

where $\alpha_c > 1/2$ is the learning rate, the vector ϕ is used to denote $\nabla\sigma_c(x)\Big(f(x) + g(x)\hat{\mu}(\hat{x}_j) + h(x)\hat{v}(x)\Big)$, and the term $(1+\phi^{\mathrm{T}}\phi)^2$ is employed for normalization. According to [39], if the set $\{\sigma_{ci}\}_{i=1}^{l_c}$ is linearly independent and the control law $\hat{\mu}$ can stabilize system (1a), then the set $\{\phi_i\}_{i=1}^{l_c}$ is also linearly independent with i being the i-th element of the vectors involved. Then, recalling $\dot{\tilde{\omega}}_c = -\dot{\hat{\omega}}_c$ and (24), we further derive that the error dynamics is

$$\dot{\tilde{\omega}}_c = -\alpha_c\frac{\phi}{(1+\phi^{\mathrm{T}}\phi)^2}(\phi^{\mathrm{T}}\tilde{\omega}_c - e_{cH}). \tag{26}$$

Under the event-based control mechanism, the closed-loop sampled-data system includes a flow dynamics for all $t \in [s_j, s_{j+1})$ and a jump dynamics for all $t = s_{j+1}$ with $j \in \mathbb{N}$. When defining an augmented state vector as $\mathcal{Z} = [x^{\mathrm{T}}, \hat{x}_j^{\mathrm{T}}, \tilde{\omega}_c^{\mathrm{T}}]^{\mathrm{T}} \in \mathbb{R}^{2n+l_c}$ and based on (11), (12), and (26), the dynamics of the impulsive system can be described by

$$\begin{cases} \dot{\mathcal{Z}} = \begin{bmatrix} f(x)+g(x)\hat{\mu}(\hat{x}_j)+h(x)\hat{v}(x) \\ \hdashline 0_{n\times 1} \\ \hdashline -\alpha_c \dfrac{\phi}{(1+\phi^{\mathsf{T}}\phi)^2}(\phi^{\mathsf{T}}\tilde{\omega}_c - e_{cH}) \end{bmatrix}, t \in [s_j, s_{j+1}); \\ \mathcal{Z}(t) = \mathcal{Z}(t^-) + \begin{bmatrix} 0_{n\times 1} \\ \hdashline x - \hat{x}_j \\ \hdashline 0_{l_c\times 1} \end{bmatrix}, \qquad t = s_{j+1}, \end{cases} \tag{27}$$

where $\mathcal{Z}(t^-) = \lim_{\epsilon\to 0}. \mathcal{Z}(t-\epsilon)$. Note that as an impulsive system, the related stability issue should be directed against two cases for a complete discussion.

3.3. Stability analysis of the closed-loop system

In what follows, the uniformly ultimately bounded (UUB) stability of the closed-loop system is analyzed. Before proceeding, we present a fact as follows.

Fact 1. *With the description of the non-quadratic utility (3), it is clear that $\varphi(\cdot)$ is Lipschitz continuous such that for any m-dimensional vectors ξ_1 and ξ_2, we have $\|\varphi(\xi_1) - \varphi(\xi_2)\| \le \mathcal{L}_2 \|\xi_1 - \xi_2\|$, where $\mathcal{L}_2$ is a positive real constant.*

In addition, the following assumptions are required for UUB stability, as usually stated in ADP literature like [36,38].

Assumption 3. For the control matrix $g(x)$, it is Lipschitz continuous such that $\|g(x) - g(\hat{x}_j)\| \le \mathcal{L}_3 \|e_j(t)\|$, where $\mathcal{L}_3$ is a positive constant and it is also upper bounded such that $\|g(x)\| \le g_{\max}$, where $g_{\max}$ is a positive constant.

Assumption 4. The derivative of the activation function is Lipschitz continuous such that $\|\nabla\sigma_c(x) - \nabla\sigma_c(\hat{x}_j)\| \le \mathcal{L}_4 \|e_j(t)\|$, where $\mathcal{L}_4$ is a positive constant. The derivative term $\nabla\sigma_c(x)$ is upper bounded such that $\|\nabla\sigma_c(x)\| \le \nabla\sigma_{c\max}$, where $\nabla\sigma_{c\max}$ is a positive constant. The derivative of the reconstruction error is upper bounded such that $\|\nabla\varepsilon_c(x)\| \le \nabla\varepsilon_{c\max}$, where $\nabla\varepsilon_{c\max}$ is a positive constant. The residual error term is upper bounded by a positive constant, i.e., $|e_{cH}| \le e_{cH\max}$.

Theorem 2. *For nonlinear system (1a), we suppose that Assumptions 3 and 4 hold. The event-based approximate optimal control law and time-based approximate worst-case disturbance law are given by (21a) and (21b), respectively, where the constructed critic network is tuned by (25). Then, under their conjunct actions, the closed-loop system (12) is asymptotically stable and the weight estimation error is UUB if the triggering condition*

$$\|e_j(t)\|^2 \le \frac{(1-\eta_2^2)\lambda_{\min}(Q)\|x\|^2 - \lambda_1 - \varrho^2\hat{v}^{\mathsf{T}}(x)P\hat{v}(x)}{\lambda_{\max}(R^{\mathsf{T}}R)\mathcal{L}_2^2\|R^{-1}\|^2\|\hat{\omega}_c\|^2\lambda_2} \triangleq \hat{e}_T, \tag{28}$$

and the inequality

$$\|\tilde{\omega}_c\| > \sqrt{\frac{\lambda_4}{\lambda_3}} \tag{29}$$

are satisfied, where $\hat{e}_T$ is the positive threshold, $\eta_2 \in (0,1)$ is the parameter to be designed for reflecting the sample frequency, and the terms λ_2, λ_3, and λ_4 are given in (35) and (42) with the learning rate being chosen as $\alpha_c > 1/2$.

Proof. Considering the impulsive dynamical system (27), we choose a Lyapunov function candidate composed of three terms as $L_2(t) = L_{21}(t) + L_{22}(t) + L_{23}(t)$, where $L_{21}(t) = J^*(x)$, $L_{22}(t) = J^*(\hat{x}_j)$, and $L_{23}(t) = (1/2)\tilde{\omega}_c^{\mathsf{T}}\tilde{\omega}_c$.

Case 1: *The events are not triggered, i.e.,* $\forall\, t \in [s_j, s_{j+1})$. Taking the time derivative of the Lyapunov function along the trajectory of the system (27), it obtains

$$\dot{L}_{21}(t) = (\nabla J^*(x))^{\mathsf{T}}\Big(f(x) + g(x)\hat{\mu}(\hat{x}_j) + h(x)\hat{v}(x)\Big); \tag{30a}$$

$$\dot{L}_{22}(t) = 0; \tag{30b}$$

$$\dot{L}_{23}(t) = -\alpha_c\tilde{\omega}_c^{\mathsf{T}}\frac{\phi}{(1+\phi^{\mathsf{T}}\phi)^2}\big(\phi^{\mathsf{T}}\tilde{\omega}_c - e_{cH}\big). \tag{30c}$$

For the first term $\dot{L}_{21}(t)$, based on (9) and (10), we can obtain the following formula:

$$\begin{aligned} \dot{L}_{21}(t) = &-x^{\mathsf{T}}Qx - Y(u^*) - \varrho^2 v^{*\mathsf{T}}(x)Pv^*(x) \\ &- (\nabla J^*(x))^{\mathsf{T}}g(x)(u^*(x) - \hat{\mu}(\hat{x}_j)) + (\nabla J^*(x))^{\mathsf{T}}h(x)\hat{v}(x) \\ = &-x^{\mathsf{T}}Qx - Y(u^*) - \varrho^2 v^{*\mathsf{T}}(x)Pv^*(x) \\ &+ 2\varphi^{-\mathsf{T}}(u^*(x))R(u^*(x) - \hat{\mu}(\hat{x}_j)) + 2\varrho^2 v^{*\mathsf{T}}(x)P\hat{v}(x). \end{aligned} \tag{31}$$

Letting $P = \mathcal{P}^{\mathsf{T}}\mathcal{P}$, where $\mathcal{P}$ is also a symmetric positive definite matrix, considering the inequality

$$2\varrho^2 v^{*\mathsf{T}}(x)P\hat{v}(x) \le \varrho^2\Big(v^{*\mathsf{T}}(x)Pv^*(x) + \hat{v}^{\mathsf{T}}(x)P\hat{v}(x)\Big),$$

and eliminating the effect of the term $-Y(u^*)$, it follows from (31) that

$$\begin{aligned} \dot{L}_{21}(t) \le &-x^{\mathsf{T}}Qx + \varrho^2\hat{v}^{\mathsf{T}}(x)P\hat{v}(x) + 2\varphi^{-\mathsf{T}}(u^*(x))R(u^*(x) - \hat{\mu}(\hat{x}_j)) \\ \le &-x^{\mathsf{T}}Qx + \varrho^2\hat{v}^{\mathsf{T}}(x)P\hat{v}(x) + \lambda_1 + \lambda_{\max}(R^{\mathsf{T}}R)\|u^*(x) - \hat{\mu}(\hat{x}_j)\|^2, \end{aligned} \tag{32}$$

where λ_1 is given in (17). Considering (8a) and (18), the time-based optimal control can be rewritten as

$$u^*(x) = -\varphi\left(\frac{1}{2}R^{-1}g^{\mathsf{T}}(x)\Big((\nabla\sigma_c(x))^{\mathsf{T}}\omega_c + \nabla\varepsilon_c(x)\Big)\right). \tag{33}$$

Adopting the neural network expression of $\hat{\mu}(\hat{x}_j)$ and $u^*(x)$, i.e., (21a) and (33), and observing Fact 1, we have

$$\|u^*(x) - \hat{\mu}(\hat{x}_j)\| \le \frac{1}{2}\mathcal{L}_2\Big\|R^{-1}g^{\mathsf{T}}(\hat{x}_j)(\nabla\sigma_c(\hat{x}_j))^{\mathsf{T}}\hat{\omega}_c - R^{-1}g^{\mathsf{T}}(x)\Big((\nabla\sigma_c(x))^{\mathsf{T}}\omega_c + \nabla\varepsilon_c(x)\Big)\Big\|,$$

which can be combined with the relationship $\omega_c = \hat{\omega}_c + \tilde{\omega}_c$ to further yield

$$\begin{aligned} \|u^*(x) - \hat{\mu}(\hat{x}_j)\|^2 &\le \frac{1}{4}\mathcal{L}_2^2\Big\|R^{-1}g^{\mathsf{T}}(\hat{x}_j)(\nabla\sigma_c(\hat{x}_j))^{\mathsf{T}}\hat{\omega}_c - R^{-1}g^{\mathsf{T}}(x)(\nabla\sigma_c(x))^{\mathsf{T}}\hat{\omega}_c - \\ &\le \frac{1}{2}\mathcal{L}_2^2\bigg(\Big\|R^{-1}\Big(g^{\mathsf{T}}(\hat{x}_j)(\nabla\sigma_c(\hat{x}_j))^{\mathsf{T}} - g^{\mathsf{T}}(x)(\nabla\sigma_c(x))^{\mathsf{T}}\Big)\hat{\omega}_c\Big\|^2 \\ &\quad + \Big\|R^{-1}g^{\mathsf{T}}(x)\Big((\nabla\sigma_c(x))^{\mathsf{T}}\tilde{\omega}_c + \nabla\varepsilon_c(x)\Big)\Big\|^2\bigg). \end{aligned} \tag{34}$$

By using the bounded conditions in Assumptions 3 and 4, we can find that

$$\begin{aligned} &\Big\|g^{\mathsf{T}}(\hat{x}_j)(\nabla\sigma_c(\hat{x}_j))^{\mathsf{T}} - g^{\mathsf{T}}(x)(\nabla\sigma_c(x))^{\mathsf{T}}\Big\|^2 \\ &= \Big\|(\nabla\sigma_c(\hat{x}_j) - \nabla\sigma_c(x))g(\hat{x}_j) + \nabla\sigma_c(x)(g(\hat{x}_j) - g(x))\Big\|^2 \\ &\le 2\Big(\|(\nabla\sigma_c(\hat{x}_j) - \nabla\sigma_c(x))g(\hat{x}_j)\|^2 + \|\nabla\sigma_c(x)(g(\hat{x}_j) - g(x))\|^2\Big) \\ &\le 2\Big(\mathcal{L}_3^2\nabla\sigma_{c\max}^2 + \mathcal{L}_4^2 g_{\max}^2\Big)\|e_j(t)\|^2 \triangleq 2\lambda_2\|e_j(t)\|^2, \end{aligned} \tag{35}$$

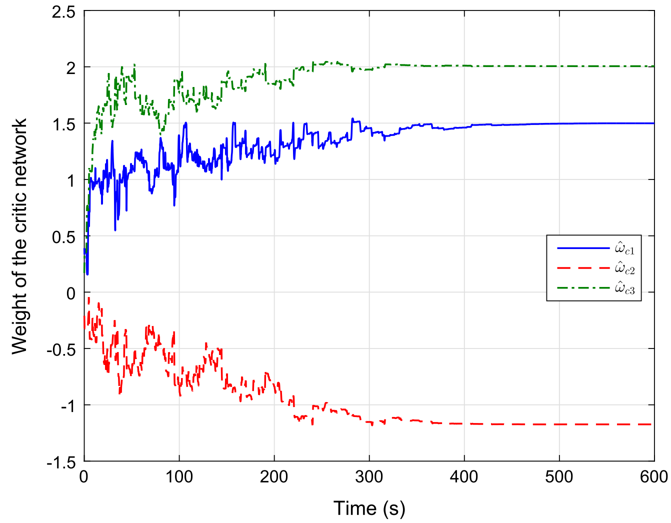

Fig. 1. Convergence process of the weight vector.

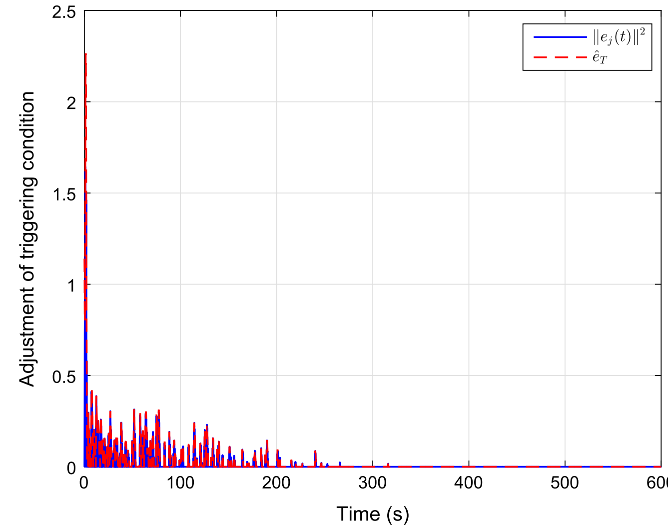

Fig. 3. Adjustment of the triggering condition.

where the constant λ_2 is introduced to denote the bounded term $\mathcal{L}_3^2\nabla\sigma_{c\,\max}^2 + \mathcal{L}_4^2 g_{\max}^2$. Besides, it can be observed that

$$\|(\nabla\sigma_c(x))^{\mathsf{T}}\tilde{\omega}_c + \nabla\varepsilon_c(x)\|^2 \le 2(\nabla\sigma_{c\,\max}^2 \,\|\tilde{\omega}_c\|^2 + \nabla\varepsilon_{c\,\max}^2). \tag{36}$$

Then, according to (34)–(36), we further obtain the reduction of (32) as follows:

$$\begin{aligned}\dot{L}_{21}(t) \le\; & -x^{\mathsf{T}}Qx + \varrho^2\hat{v}^{\mathsf{T}}(x)P\hat{v}(x) + \lambda_1 \\ & + \lambda_{\max}(R^{\mathsf{T}}R)\mathcal{L}_2^2\,\|R^{-1}\|^2\Big(\|\hat{\omega}_c\|^2\lambda_2\,\|e_j(t)\|^2 \\ & + g_{\max}^2\,\nabla\sigma_{c\,\max}^2\,\|\tilde{\omega}_c\|^2 + g_{\max}^2\,\nabla\varepsilon_{c\,\max}^2\Big).\end{aligned} \tag{37}$$

For the term $\dot{L}_{23}(t)$, we expand the time derivative (30c), introduce a l_c-dimensional column vector $\phi_1 = \phi/(1+\phi^{\mathsf{T}}\phi)$ and a scalar $\phi_2 = 1+\phi^{\mathsf{T}}\phi$, and then find that

$$\dot{L}_{23}(t) = -\alpha_c\tilde{\omega}_c^{\mathsf{T}}\phi_1\phi_1^{\mathsf{T}}\tilde{\omega}_c + \alpha_c\frac{\tilde{\omega}_c^{\mathsf{T}}\phi_1}{\phi_2}e_{cH}. \tag{38}$$

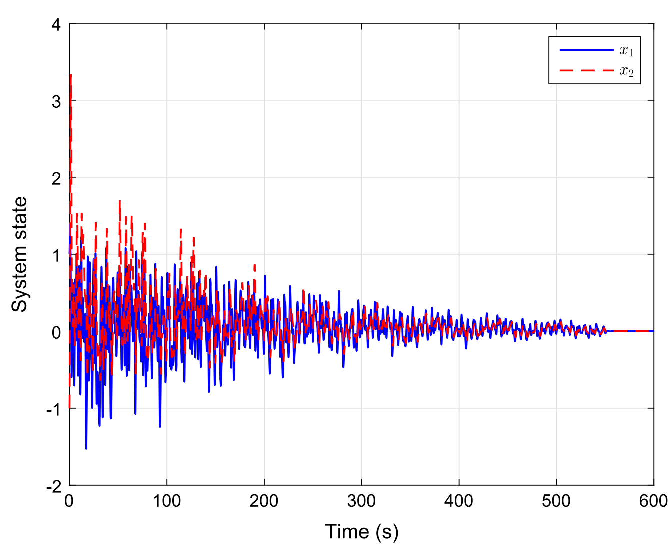

Fig. 2. The state trajectory during the learning phase.

When applying Young's inequality to the second term of (38), we can derive that

$$\alpha_c\frac{\tilde{\omega}_c^{\mathsf{T}}\phi_1}{\phi_2}e_{cH} \le \frac{1}{2}\left(\tilde{\omega}_c^{\mathsf{T}}\phi_1\phi_1^{\mathsf{T}}\tilde{\omega}_c + \alpha_c^2\frac{e_{cH}^2}{\phi_2^2}\right). \tag{39}$$

By recalling Assumption 4 and taking account of the fact that $\phi_2 = 1+\phi^{\mathsf{T}}\phi \ge 1$, it follows from (38) and (39) that

$$\begin{aligned}\dot{L}_{23}(t) &\le -\left(\alpha_c - \frac{1}{2}\right)\tilde{\omega}_c^{\mathsf{T}}\phi_1\phi_1^{\mathsf{T}}\tilde{\omega}_c + \alpha_c^2\frac{e_{cH\,\max}^2}{2\phi_2^2} \\ &\le -\left(\alpha_c - \frac{1}{2}\right)\lambda_{\min}(\phi_1\phi_1^{\mathsf{T}})\,\|\tilde{\omega}_c\|^2 + \frac{1}{2}\alpha_c^2 e_{cH\,\max}^2.\end{aligned} \tag{40}$$

Note that if the persistence of excitation like condition is satisfied, we have $\lambda_{\min}(\phi_1\phi_1^{\mathsf{T}}) > 0$.

By combining (30), (37), and (40), we can obtain the overall time derivative as

$$\begin{aligned}\dot{L}_2(t) \le\; & -x^{\mathsf{T}}Qx + \lambda_1 + \varrho^2\hat{v}^{\mathsf{T}}(x)P\hat{v}(x) - \lambda_3\,\|\tilde{\omega}_c\|^2 + \lambda_4 \\ & + \lambda_{\max}(R^{\mathsf{T}}R)\mathcal{L}_2^2\,\|R^{-1}\|^2\|\hat{\omega}_c\|^2\lambda_2\,\|e_j(t)\|^2 \\ \le\; & -\eta_2^2\lambda_{\min}(Q)\|x\|^2 + \left(\eta_2^2 - 1\right)\lambda_{\min}(Q)\|x\|^2 + \lambda_1 \\ & + \lambda_{\max}(R^{\mathsf{T}}R)\mathcal{L}_2^2\,\|R^{-1}\|^2\|\hat{\omega}_c\|^2\lambda_2\,\|e_j(t)\|^2 \\ & + \varrho^2\hat{v}^{\mathsf{T}}(x)P\hat{v}(x) - \lambda_3\,\|\tilde{\omega}_c\|^2 + \lambda_4,\end{aligned} \tag{41}$$

where the positive terms λ_3 and λ_4 are

$$\begin{aligned}\lambda_3 = & \left(\alpha_c - \frac{1}{2}\right)\lambda_{\min}(\phi_1\phi_1^{\mathsf{T}}) \\ & - \lambda_{\max}(R^{\mathsf{T}}R)\mathcal{L}_2^2\,\|R^{-1}\|^2\,g_{\max}^2\,\nabla\sigma_{c\,\max}^2;\end{aligned} \tag{42a}$$

$$\lambda_4 = \lambda_{\max}(R^{\mathsf{T}}R)\mathcal{L}_2^2\,\|R^{-1}\|^2\,g_{\max}^2\,\nabla\varepsilon_{c\,\max}^2 + \frac{1}{2}\alpha_c^2 e_{cH\,\max}^2. \tag{42b}$$

Then, we can find that if the triggering condition (28) and the inequality (29) are satisfied, the time derivative inequality (41) becomes $\dot{L}_2(t) \le -\eta_2^2\lambda_{\min}(Q)\|x\|^2 < 0$ for any $x \ne 0$. In other words, the derivative of the Lyapunov function candidate is negative during the flow for all $t \in [s_j, s_{j+1})$.

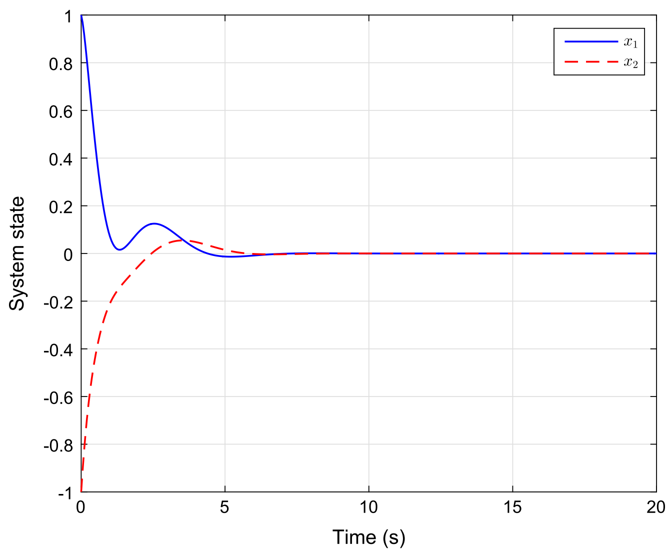

Fig. 4. The state trajectory.

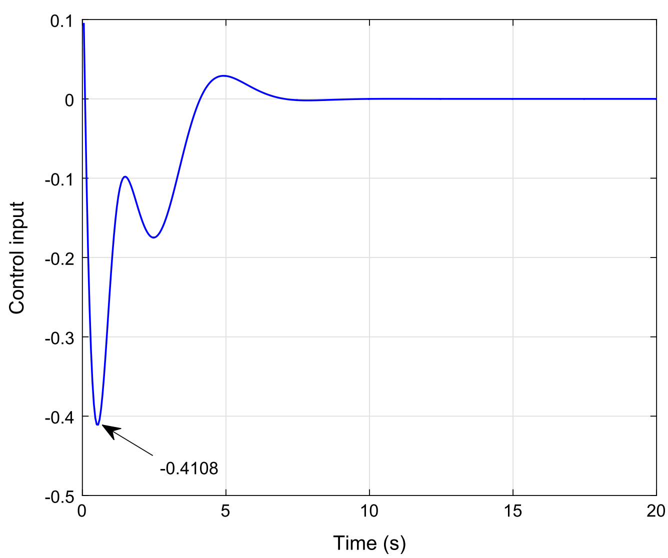

Fig. 5. The control input.

Case 2: *The events are triggered, i.e.*, $\forall\, t = s_{j+1}$. The difference of the chosen Lyapunov function candidate is

$$\Delta L_2(t) = L_2(\hat{x}_{j+1}) - L_2(x(s_{j+1}^-)) = \Delta L_{21}(t) + \Delta L_{22}(t) + \Delta L_{23}(t).$$

According to (28), (29), and (41), we know that $\dot{L}_2(t) < 0$ for all $t \in [s_j, s_{j+1})$. Since the system state and cost function are continuous, we can obtain the time differences

$$\Delta L_{21}(t) = J^*(\hat{x}_{j+1}) - J^*(x(s_{j+1}^-)) \le 0; \tag{43a}$$

$$\Delta L_{22}(t) = J^*(\hat{x}_{j+1}) - J^*(\hat{x}_j); \tag{43b}$$

$$\Delta L_{23}(t) = \frac{1}{2}\Big(\tilde{\omega}_c^{\mathrm{T}}(\hat{x}_{j+1})\tilde{\omega}_c(\hat{x}_{j+1}) - \tilde{\omega}_c^{\mathrm{T}}(x(s_{j+1}^-))\tilde{\omega}_c(x(s_{j+1}^-))\Big) \le 0. \tag{43c}$$

Based on (43) we derive $\Delta L_2(t) \le -\mathcal{K}(\| e_{j+1}(s_j)\|)$, where $\mathcal{K}(\cdot)$ is a class-$\mathcal{K}$ function and $e_{j+1}(s_j) = \hat{x}_{j+1} - \hat{x}_j$. This implies that the Lyapunov function candidate $L_2(t)$ is also decreasing at the triggering instants $\forall\, t = s_{j+1}$.

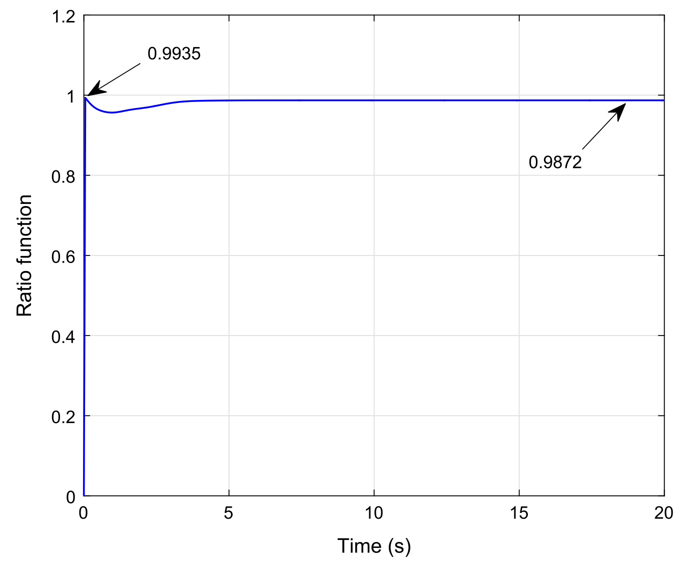

Fig. 6. Adjustment of the ratio function.

Considering Case 1 and Case 2, the triggering condition (28) and the inequality (29) ensure that the closed-loop impulsive system is asymptotically stable and the weight estimation error is UUB. At this point, the proof is finished.□

Remark 2. Observing (42a), we find that the learning rate should be chosen satisfying $\alpha_c > 1/2$, so as to ensure that the term λ_3 is positive. According to (29), we find that a larger learning rate can lead to a greater value of λ_3 and meanwhile, a smaller bound of $\tilde{\omega}_c$. In practical, we really want to decrease the bound of UUB stability as much as possible. In this sense, we can increase the learning rate as much as possible. However, we cannot do that due to the inherent weakness of neural network technique, such as the local minimum point. Consequently, it is always an experimental choice with engineering experience and intuition after considering a tradeoff between control accuracy and computation complexity.

Remark 3. Note that the two triggering thresholds e_T and $\hat{e}_T$ designed in Theorems 1 and 2 are different from each other. In general, $\hat{e}_T$ is employed to help to learn the critic weight and then facilitate conducting the input-constrained H_∞ control implementation with the involvement of e_T.

4. Simulation studies

In this section, an example is provided to verify the effectiveness of the ACL strategy for event-based nonlinear H_∞ control with input constraints. We consider an input-affine nonlinear system with an external perturbation as follows:

$$\dot{x} = \begin{bmatrix} -x_1^3 - 2x_2 \\ x_1 - x_2 + 0.5\cos x_1^2 \sin x_2^3 \end{bmatrix} + \begin{bmatrix} 1 \\ \sin x_1 \end{bmatrix} u(x) + \begin{bmatrix} -\cos x_2 \\ 0 \end{bmatrix} v(x), \tag{44}$$

where $x = [x_1, x_2]^{\mathrm{T}} \in \mathbb{R}^2$, $u(x) \in \mathbb{R}$, and $v(x) \in \mathbb{R}$ are the state, control, and perturbation variables, respectively. The control input is constrained to bound as $|u| \le \bar{u} = 0.5$. As for the infinite horizon cost function (4), the utility function is chosen as

$$U(x, u, v) = x^{\mathrm{T}}Qx + 2\int_0^u \tanh^{-\mathrm{T}}(\varsigma)R\mathrm{d}\varsigma - \varrho^2 v^{\mathrm{T}}Pv,$$

where $\varrho = 1.5$, $Q = 2I_2$, $R=I$, and $P = 0.5I$ with I being the identity matrix with suitable dimension.

We adopt the idea of ACL to design the event-based optimal control. Let the number of neurons in the hidden layer be $l_c=3$. We denote the weight vector of the critic network as the form $\hat{\omega}_c = [\hat{\omega}_{c1}, \hat{\omega}_{c2}, \hat{\omega}_{c3}]^{\mathsf{T}}$. The activation function of the critic network is experimentally selected as $\sigma_c(x) = [x_1^2, x_1x_2, x_2^2]^{\mathsf{T}}$. Besides, set the learning rate of the critic network as $\alpha_c = 0.9$ and let the initial state of the controlled plant be $x_0 = [1, -1]^{\mathsf{T}}$. As for other prespecified parameters of the triggering condition (28), we choose $\mathcal{L}_2 = 1$, $\eta_2 = 0.5$ and $\lambda_2 = 9$. In addition, the sampling time of the learning process is set as 0.1 s experimentally. During simulation, we add a probing noise to guarantee the persistency of excitation condition and observe that the weight vector of the critic network finally converges to $\hat{\omega}_c = [1.4985, -1.1734, 2.0055]^{\mathsf{T}}$, as shown in Fig. 1. In fact, we can observe that the convergence of the weight vector has occurred after 550 s. Then, the probing signal is turned off. The adjustment of the state trajectory during the learning phase is presented in Fig. 2. We see that the state vector converges to zero after the probing noise is turned off. Additionally, the adjustment of the triggering condition is depicted in Fig. 3.

Next, we turn to evaluate the H_∞ control performance with the obtained control law $\mu^*(\hat{x}_j)$ and the triggering condition (14). We set the prespecified parameters as $\mathcal{L}_1 = 1$ and $\eta_1 = 0.5$ and select the sampling time as 0.05 s. We apply the obtained control law to the controlled plant (44) for 20 s with the following external perturbation being introduced:

$$v(t) = \begin{cases} 3e^{-(t-t_0)}\cos(t-t_0), & \text{if } t > t_0; \\ 0, & \text{else.} \end{cases}$$

The corresponding simulation results are displayed in Figs. 4 and 5 when selecting $t_0 = 0$ s. In detail, Fig. 4 shows the system state trajectory while Fig. 5 displays the curve of event-based state feedback controller. Here, we can observe that the control signal does not reach the constrained bound all the time (i.e., $|u| \le 0.4108 < \bar{u} = 0.5$).

Finally, let us define a ratio function $\bar{\varrho}(t)$, which is used to reflect the disturbance attenuation and formed as

$$\bar{\varrho}^2(t) = \frac{\int_{t_0}^{t}\left(x^{\mathsf{T}}(\tau)Qx(\tau) + 2\int_0^{u(\tau)} \tanh^{-\mathsf{T}}(\varsigma)R\mathrm{d}\varsigma\right)\mathrm{d}\tau}{\int_{t_0}^{t} v^{\mathsf{T}}(\tau)Pv(\tau)\mathrm{d}\tau}.$$

Via simulation, it is found that the ratio $\bar{\varrho}(t)$ converges to 0.9872 as illustrated in Fig. 6. This implies that the H_∞ control law designed in this paper can attain a prespecified L_2-gain performance level when regarding the closed-loop system (i.e., $\bar{\varrho}(t) \le 0.9935 < \varrho = 1.5$).

The above experimental results demonstrate that the obtained event-based optimal controller possesses an excellent ability of disturbance rejection.

5. Concluding remarks

An event-based input-constrained H_∞ state feedback control approach of nonlinear dynamical systems is developed with ACL. The nonlinear H_∞ control problem is transformed into a two-player zero-sum differential game, which is solved by introducing the event-based control mechanism and adopting the ACL-based optimization methodology. The approximated event-based optimal control law and time-based worst-case disturbance law are derived via training a single critic network with closed-loop stability proof.

Note that in the current study, the control function derived by (8a) is usually limited to a range $|u_i^*| \le 1$ when selecting $\varphi_i(\cdot) = \tanh(\cdot)$, $i = 1, 2, \ldots, m$. In fact, we can introduce a diagonal matrix $\bar{U} = \mathrm{diag}\{\bar{u}_1, \bar{u}_2, \ldots, \bar{u}_m\}$ and adopt a non-quadratic utility function as the form

$$Y(u) = 2\int_0^u \varphi^{-\mathsf{T}}(\bar{U}^{-1}\varsigma)R\mathrm{d}\varsigma, \tag{45}$$

such that the control function can be obtained by a more general formulation

$$u^*(x) = -\bar{U}\varphi\left(\frac{1}{2}R^{-1}g^{\mathsf{T}}(x)\nabla J^*(x)\right), \tag{46}$$

where $g(x)$ is the known control matrix. Then, how to obtain the triggering condition in case that (45) and (46) are considered deserves further study in the future work. Additionally, how to reduce the requirement of relying on the system dynamics is also a potential research direction.

References

[1] F.L. Lewis, D.L. Vrabie, V.L. Syrmos, Optimal Control, third edition, John Wiley & Sons, Inc., Hoboken, NJ, 2012.
[2] S.E. Lyshevski, Nonlinear discrete-time systems: constrained optimization and application of nonquadratic costs, in Proceedings of the American Control Conference, Philadelphia, PA, USA, June 1998, pp. 3699–3703.
[3] S. Wang, X. Ren, J. Na, X. Gao, Robust tracking and vibration suppression for nonlinear two-inertia system via modified dynamic surface control with error constraint, Neurocomputing 203 (2016) 73–85.
[4] P.J. Werbos, Approximate dynamic programming for real-time control and neural modeling, in: D.A. White, D.A. Sofge (Eds.), Handbook of Intelligent Control: Neural, Fuzzy, and Adaptive Approaches, Van Nostrand Reinhold, New York, NY, 1992 (Chapter 13).
[5] F.L. Lewis, D. Liu, Reinforcement Learning and Approximate Dynamic Programming for Feedback Control, Wiley, Hoboken, NJ, 2013.
[6] L. Cheng, Z.G. Hou, M. Tan, Adaptive neural network tracking control for manipulators with uncertain kinematics, dynamics and actuator model, Automatica 45 (10) (2009) 2312–2318.
[7] L. Cheng, W. Liu, Z.G. Hou, J. Yu, M. Tan, Neural-network-based nonlinear model predictive control for piezoelectric actuators, IEEE Trans. Ind. Electron. 62 (12) (2015) 7717–7727.
[8] W. He, Y. Dong, C. Sun, Adaptive neural impedance control of a robotic manipulator with input saturation, IEEE Trans. Syst. Man Cybern. Syst. 46 (3) (2016) 334–344.
[9] B. Xu, C. Yang, Y. Pan, Global neural dynamic surface tracking control of strict-feedback systems with application to hypersonic flight vehicle, IEEE Trans. Neural Netw. Learn. Syst. 26 (10) (2015) 2563–2575.
[10] D. Wang, D. Liu, Neuro-optimal control for a class of unknown nonlinear dynamic systems using SN-DHP technique, Neurocomputing 121 (2013) 218–225.
[11] A. Heydari, S.N. Balakrishnan, Finite-horizon control-constrained nonlinear optimal control using single network adaptive critics, IEEE Trans. Neural Netw. Learn. Syst. 24 (1) (2013) 145–157.
[12] H. Zargarzadeh, T. Dierks, S. Jagannathan, Optimal control of nonlinear continuous-time systems in strict-feedback form, IEEE Trans. Neural Netw. Learn. Syst. 26 (10) (2015) 2535–2549.
[13] D. Wang, C. Mu, D. Liu, Data-driven nonlinear near-optimal regulation based on iterative neural dynamic programming, Acta Autom. Sin., accepted for publication.
[14] J. Na, G. Herrmann, Online adaptive approximate optimal tracking control with simplified dual approximation structure for continuous time unknown nonlinear systems, IEEE/CAA J. Autom. Sin. 1 (4) (2014) 412–422.
[15] C. Mu, Z. Ni, C. Sun, H. He, Air-breathing hypersonic vehicle tracking control based on adaptive dynamic programming, IEEE Trans. Neural Netw. Learn. Syst., in press, http://dx.doi.org/10.1109/TNNLS.2016.2516948.
[16] C. Mu, Z. Ni, C. Sun, H. He, Data-driven tracking control with adaptive dynamic programming for a class of continuous-time nonlinear systems, IEEE Trans. Cybern., (in press, http://dx.doi.org/10.1109/TCYB.2016.2548941).
[17] X. Yang, D. Liu, Q. Wei, D. Wang, Guaranteed cost neural tracking control for a class of uncertain nonlinear systems using adaptive dynamic programming, Neurocomputing 198 (2016) 80–90.
[18] Y. Jiang, Z.P. Jiang, Robust adaptive dynamic programming and feedback stabilization of nonlinear systems, IEEE Trans. Neural Netw. Learn. Syst. 25 (5) (2014) 882–893.
[19] D. Wang, D. Liu, Q. Zhang, D. Zhao, Data-based adaptive critic designs for nonlinear robust optimal control with uncertain dynamics, IEEE Trans. Syst. Man Cybern. Syst., (in press, http://dx.doi.org/10.1109/TSMC.2015.2492941).
[20] D. Wang, C. Li, D. Liu, C. Mu, Data-based robust optimal control of continuous-time affine nonlinear systems with matched uncertainties, Inf. Sci. 366 (2016) 121–133.
[21] M. Abu-Khalaf, F.L. Lewis, J. Huang, Policy iterations on the Hamilton-Jacobi-Isaacs equation for H_∞ state feedback control with input saturation, IEEE

Trans. Autom. Control 51 (12) (2006) 1989–1995.
[22] K.G. Vamvoudakis, F.L. Lewis, Online solution of nonlinear two-player zero-sum games using synchronous policy iteration, Int. J. Robust Nonlinear Control 22 (2012) 1460–1483.
[23] H. Zhang, C. Qin, B. Jiang, Y. Luo, Online adaptive policy learning algorithm for H_∞ state feedback control of unknown affine nonlinear discrete-time systems, IEEE Trans. Cybern. 44 (12) (2014) 2706–2718.
[24] B. Luo, H.N. Wu, T. Huang, Off-policy reinforcement learning for H_∞ control design, IEEE Trans. Cybern. 45 (1) (2015) 65–76.
[25] D. Liu, H. Li, D. Wang, Neural-network-based zero-sum game for discrete-time nonlinear systems via iterative adaptive dynamic programming algorithm, Neurocomputing 110 (2013) 92–100.
[26] D. Zhao, Q. Zhang, D. Wang, Y. Zhu, Experience replay for optimal control of nonzero-sum game systems with unknown dynamics, IEEE Trans. Cybern. 46 (3) (2016) 854–865.
[27] P. Tabuada, Event-triggered real-time scheduling of stabilizing control tasks, IEEE Trans. Autom. Control 52 (9) (2007) 1680–1685.
[28] A. Eqtami, D.V. Dimarogonas, K.J. Kyriakopoulos, Event-triggered control for discrete-time systems, in: Proceedings of the American Control Conference, Baltimore, MD, USA, June 2010, pp. 4719–4724.
[29] P. Tallapragada, N. Chopra, On event triggered trajectory tracking for control affine nonlinear systems, in: Proceedings of the 50th IEEE Conference on Decision and Control and European Control Conference, Orlando, FL, USA, December 2011, pp. 5377–5382.
[30] P. Tallapragada, N. Chopra, On event triggered tracking for nonlinear systems, IEEE Trans. Autom. Control 58 (9) (2013) 2343–2348.
[31] Q. Li, B. Shen, Y. Liu, F.E. Alsaadi, Event-triggered H_∞ state estimation for discrete-time stochastic genetic regulatory networks with Markovian jumping parameters and time-varying delays, Neurocomputing 174 (B) (2016) 912–920.
[32] Q. Li, B. Shen, J. Liang, H. Shu, Event-triggered synchronization control for complex networks with uncertain inner coupling, Int. J. General Syst. 44 (2) (2015) 212–225.
[33] H. Dong, Z. Wang, F.E. Alsaadi, B. Ahmad, Event-triggered robust distributed state estimation for sensor networks with state-dependent noises, Int. J. General Syst. 44 (2) (2015) 254–266.
[34] H. Ma, D. Liu, D. Wang, F. Tan, C. Li, Centralized and decentralized event-triggered control for group consensus with fixed topology in continuous time, Neurocomputing 161 (2015) 267–276.
[35] K.G. Vamvoudakis, Event-triggered optimal adaptive control algorithm for continuous-time nonlinear systems, IEEE/CAA J. Autom. Sin. 1 (3) (2014) 282–293.
[36] A. Sahoo, H. Xu, S. Jagannathan, Neural network-based event-triggered state feedback control of nonlinear continuous-time systems, IEEE Trans. Neural Netw. Learn. Syst. 27 (3) (2016) 497–509.
[37] X. Zhong, H. He, An event-triggered ADP control approach for continuous-time system with unknown internal states, IEEE Trans. Cybern., (in press, http://dx.doi.org/10.1109/TCYB.2016.2523878).
[38] Q. Zhang, D. Zhao, Y. Zhu, Event-triggered H_∞ control for continuous-time nonlinear system via concurrent learning, IEEE Trans. Syst. Man Cybern. Syst., (in press, http://dx.doi.org/10.1109/TSMC.2016.2531680).
[39] R.W. Beard, G.N. Saridis, J.T. Wen, Galerkin approximations of the generalized Hamilton–Jacobi–Bellman equation, Automatica 33 (1997) 2159–2177.

Ding Wang received the B.S. degree in mathematics from Zhengzhou University of Light Industry, Zhengzhou, China, the M.S. degree in operations research and cybernetics from Northeastern University, Shenyang, China, and the Ph.D. degree in control theory and control engineering from Institute of Automation, Chinese Academy of Sciences, Beijing, China, in 2007, 2009, and 2012, respectively. He has been a Visiting Scholar with the Department of Electrical, Computer, and Biomedical Engineering, University of Rhode Island, Kingston, RI, USA, since 2015. He is currently an Associate Professor with The State Key Laboratory of Management and Control for Complex Systems, Institute of Automation, Chinese Academy of Sciences. His research interests include adaptive and learning systems, intelligent control, and neural networks. He has published over 60 journal and conference papers, and coauthored two monographs.

He was the Secretariat of the 2014 IEEE World Congress on Computational Intelligence (IEEE WCCI 2014), and the Registration Chair of the 5th International Conference on Information Science and Technology (ICIST 2015) and the 4th International Conference on Intelligent Control and Information Processing (ICICIP 2013), and served as the program committee member of several international conferences. He is the Finance Chair of the 12th World Congress on Intelligent Control and Automation (WCICA 2016). He was a recipient of the Excellent Doctoral Dissertation Award of Chinese Academy of Sciences in 2013, and a nomination of the Excellent Doctoral Dissertation Award of Chinese Association of Automation (CAA) in 2014. He serves as an Associate Editor of IEEE Transactions on Neural Networks and Learning Systems and Neurocomputing. He is a member of Institute of Electrical and Electronics Engineers (IEEE), Asia-Pacific Neural Network Society (APNNS), and CAA.

Chaoxu Mu received the B.S. degree from the Harbin Institute of Technology, Harbin, China, in 2006, the M.S. degree from Hohai University, Nanjing, China, in 2009, and the Ph.D. degree in control science and engineering from Southeast University, Nanjing, China, in 2012. She was a Visiting Ph.D. Student with the Royal Melbourne Institute of Technology University, Melbourne, VIC, Australia, from 2010 to 2011. She has been a Post-Doctoral Fellow with the Department of Electrical, Computer and Biomedical Engineering, The University of Rhode Island, Kingston, RI, USA, since 2014. She is currently a Lecturer with the School of Electrical and Automation Engineering, Tianjin University, Tianjin, China. She has authored over 40 journal and conference papers, and co-authored one monograph. Her current research interests include computational intelligence, nonlinear system control and optimization, adaptive and learning systems, smart grid, and machine learning. Dr. Mu is a member of the Asia-Pacific Neural Network Society and the Youth Committee of the Chinese Association of Automation.

Qichao Zhang received the B.S. degree in automation from Northeastern Electric Power University, Jilin, China and the M.S. degree in control theory and control engineering from Northeast University in 2012 and 2014, respectively. He is currently working toward the Ph.D. degree in control theory and control engineering at The State Key Laboratory of Management and Control for Complex Systems, Institute of Automation, Chinese Academy of Sciences, Beijing, China. His research interests include reinforcement learning, game theory and multi-agent systems.

Derong Liu received the Ph.D. degree in electrical engineering from the University of Notre Dame in 1994. He was a Staff Fellow with General Motors Research and Development Center, from 1993 to 1995. He was an Assistant Professor with the Department of Electrical and Computer Engineering, Stevens Institute of Technology, from 1995 to 1999. He joined the University of Illinois at Chicago in 1999, and became a Full Professor of Electrical and Computer Engineering and of Computer Science in 2006. He was selected for the "100 Talents Program" by the Chinese Academy of Sciences in 2008, and he served as the Associate Director of The State Key Laboratory of Management and Control for Complex Systems at the Institute of Automation, from 2010 to 2015. He is now a Full Professor with the School of Automation and Electrical Engineering, University of Science and Technology Beijing. He has published 15 books (six research monographs and nine edited volumes). He is an elected AdCom member of the IEEE Computational Intelligence Society and he is the Editor-in-Chief of the Artificial Intelligence Review. He was the General Chair of 2014 IEEE World Congress on Computational Intelligence and is the General Chair of 2016 World Congress on Intelligent Control and Automation. He received the Faculty Early Career Development Award from the National Science Foundation in 1999, the University Scholar Award from University of Illinois from 2006 to 2009, the Overseas Outstanding Young Scholar Award from the National Natural Science Foundation of China in 2008, and the Outstanding Achievement Award from Asia Pacific Neural Network Assembly in 2014. He is a Fellow of the IEEE and a Fellow of the International Neural Network Society.