International Journal of Innovative Computing, Information and Control
Volume **20**, Number **5**, October **2024**
ICIC International ©2024 ISSN 1349-4198
pp. 1257–1275

EVENT-TRIGGERED ADAPTIVE CONTROL FOR NONLINEAR MULTI-PLAYER GAMES USING NEURAL CRITIC LEARNING

Ping Li[1,2], Huiyan Zhang[3], Wengang Ao[2,3] and Pengda Liu[3,*]
[1]School of Mechanical Engineering
[2]Chongqing Key Laboratory of Green Design and Manufacturing of Intelligent Equipment
[3]National Research Base of Intelligent Manufacturing Service
Chongqing Technology and Business University
No. 19, Xuefu Avenue, Nan'an District, Chongqing 400067, P. R. China
{ lpcq; huiyanzhang; aowg }@ctbu.edu.cn; *Corresponding author: liupengda@ctbu.edu.cn

Received November 2023; revised April 2024

Abstract. *In this paper, the event-triggered neural critic learning method is proposed to solve multi-player zero-sum games of nonlinear systems. First, the Hamilton-Jacobi-Isaacs equation is derived, which provides the foundation for solving zero-sum games with adaptive critic learning. Then, the integral reinforcement learning technique contributes to deriving the control and disturbance strategies without knowing the drift dynamics. To save computation/communication resources, dynamic event-triggered control mechanism is integrated into the critic learning method via introducing an internal dynamic variable. Furthermore, the stability of the closed-loop system and the uniform ultimately boundedness of critic weight error are demonstrated via Lyapunov theorem. Finally, the effectiveness of the proposed critic learning method is validated by two simulation experiments.*

Keywords: Adaptive dynamic programming, Adaptive critic learning, Reinforcement learning, Multi-player zero-sum games, Event-triggered control

1. **Introduction.** Optimal control, which aims to obtain a controller minimizing the user-defined performance index, has gained extensive research in control fields [1, 2, 3]. Dynamic programming has the ability to deal with the optimal control problems (OCPs) in some cases where the system is simple. Nonetheless, for the cases where the system becomes more complex, this technique may well present powerlessness due to the "curse of dimensionality" incurred by the growth calculations. Fortunately, adaptive critic learning (ACL), which is also known as adaptive dynamic programming, can approximate the optimal control and overcome this difficulty via the forward-in-time manner [4, 5, 6]. ACL methods generally utilize actor-critic framework to achieve the maximal cumulative rewards by interacting with the environment, such that the optimal solutions can be approximated [7, 8]. Due to the powerful approximating ability, in recent years, ACL methods have been reported to tackle different OCPs [9]. It is worth noting that most of the ACL approaches were proposed for OCPs of the systems with only one controller. However, there often exist the systems with more than one controller, of which the control issues can be described via multi-player games (MPGs) [10].

Game theory, in which nonzero-sum games (NZSGs) and zero-sum game (ZSG) are main branches, performs powerful ability to describe the complex decisions/behaviors [11]. Generally, the target for NZSGs is to achieve optimal controls while all the players reach their own minimal costs on the basis of the achievement of the system stability. While for the ZSGs, the players attempt to pursue their individual interests. NZSGs

---

DOI: 10.24507/ijicic.20.05.1257

or ZSG can be applied in multifarious control scenarios, such as autonomous vehicles changing lanes [12], load frequency control in power systems [4], and H_∞ control [13]. The key to tackling NZSGs or ZSG lies in solving the coupled Hamilton-Jacobi equations or the Hamilton-Jacobi-Isaacs equation (HJIE), with which are usually troublesome to deal [14]. Although the closed-form solutions for nonlinear game systems are so difficult to derive, one can utilize ACL method to approximate optimal strategies for these players. For example, an online ACL algorithm was developed in [15] to solve NZSGs of discrete-time unknown nonlinear systems. In [16], a new ACL method was proposed to approximate Nash equilibrium solution for achieving robust tracking control of NZSG systems. Furthermore, in [17], the nonlinear optimization problem of NZSGs with unknown drift dynamics was investigated with data-based integral reinforcement learning method.

With the progress of computer science, microelectronics, communication and power electronics, the networked control systems (NCSs) present the broad application prospect. There often exist a mass of data computation, exchange and transmission especially for the systems with multi-controllers, which may lead to transmission delay and eventually deteriorate the control performances [18, 19, 20, 21]. To improve the control efficiency, scholars developed the event-triggered mechanism (ETM), which can be fused with intelligent control ideas [22]. Different from traditional mechanism using all the data to reach the control object, this mechanism selectively filters the data and only utilizes the necessary data to compute and update the control strategy. In other words, ETM works with a nonperiodic manner where the filtered data drive the control to be updated aperiodically. Recently, ETM has been integrated into ACL frameworks to tackle different OCPs. In [23], the event-triggered robust control issues for NZSG systems was investigated with the aid of neural identifier and ACL method. In [24], a novel robust tracking control strategy for nonlinear unmatched uncertain systems was formulated using the event-based ACL approach. An event-triggered robust ACL algorithm was developed in [25] to solve a class of multiplayer Stackelberg-Nash games of uncertain nonlinear continuous-time systems. Through ACL approach, a dual event-triggered constrained control scheme was established in [26] for solving discrete-time nonlinear ZSG issue. An event-based ACL algorithm was proposed in [27] to study the OCPs of interconnected nonlinear systems subject to stochastic dynamics. In [28], the distributed formation control issue of multi-quadrotor unmanned aerial vehicle in the framework of event triggering was studied via ACL algorithm. On the foundation of traditional ETM, a dynamic event-triggered mechanism (DETM) was constructed in [29] to reduce more computation/communication costs. This dynamic triggering mechanism works via introducing a novel internal dynamic variable which can reasonably enlarge the triggering threshold, and the computation/communication burdens for the systems can be further alleviated.

Inspired by the above works, we investigate the MPGs for nonlinear systems by feat of ACL method under the framework of DETM. The contributions of this work are listed as follows. First, different from the works on NZSGs and two-player ZSG [30, 31, 32], the issue of multi-player games is considered which is of generality. The HJIE is constructed which forms the foundations for solving MPGs through adaptive critic networks. Then, the ACL approach is utilized to derive the approximated strategies for all the players. Furthermore, based on integral reinforcement learning (IRL) technique, the ACL approach of critic-only architecture is proposed to approximate the solutions of HJIE with drift dynamics unknown. Finally, in contrast to the works on [33, 34, 35], the dynamic triggering mechanism integrating dead-zone operation is fused with critic learning. The internal variable and the preliminary operation both contribute to the reduction of the triggering number. The dynamic triggering mechanism can be turned off by the preliminary dead-zone operation when the system state comes close enough to the origin, such that the unnecessary computation and communication costs can be further reduced.

The remainder of this work is listed as follows. Section 2 provides the descriptions of MPGs issue. The event-based optimal control strategies are derived in Section 3. Section 4 provides the dynamic event-triggered control structure and the stability demonstration. In Section 5, two examples are simulated to validate the proposed approach. Section 6 concludes the whole work.

Notations: $\mathbb{R}$, $\mathbb{R}^n$ and $\mathbb{R}^{n\times m}$ are the set including all real numbers, the n-dimensional Euclidean space and the space including all real matrices. The integer sets are defined that $\mathbb{N}_\mu = \{1, \ldots, N_\mu\}$ and $\mathbb{N}_\nu = \{1, \ldots, N_\nu\}$. $\bigtriangledown(\cdot) \triangleq \partial(\cdot)/\partial x$ and $\|\cdot\|$ are the gradient operator and Euclidean norm. $\lambda_{\min}(\cdot)$ denotes the minimum eigenvalue. $I_{n\times n}$ is the unix matrix with n dimension.

2. **Problem Formulation.** The nonlinear game system is presented as

$$\dot{x} = \mathcal{F}(x) + \sum_{\imath=1}^{N_\mu} \mathcal{G}_\imath(x)\mu_\imath + \sum_{\ell=1}^{N_\nu} \mathfrak{H}_\ell(x)\nu_\ell, \tag{1}$$

where $x \in \Theta \subset \mathbb{R}^n$, $\mu_\imath \in \mathfrak{U} \subset \mathbb{R}^{m\imath}$, and $\nu_\ell \in \mathfrak{V} \subset \mathbb{R}^{m\ell}$ are state vector, control strategy and disturbance, respectively. Herein, $\mathfrak{U}$ and $\mathfrak{V}$ are defined as $\mathfrak{U} = \{\mu_1, \ldots, \mu_{N_\mu}\}$ and $\mathfrak{V} = \{\nu_1, \ldots, \nu_{N_\ell}\}$. In addition, the system should satisfy the following assumptions [36].

Assumption 2.1. *The function $\mathcal{F}(x)$ is locally Lipschitz continuous.*

Assumption 2.2. $\forall\imath \in \mathbb{N}_\mu$, *$\mathcal{G}_\imath(x)$ is assumed to be bounded with $b_{\mathcal{G}\imath}$. Similarly, $\forall\ell \in \mathbb{N}_\nu$, $\mathfrak{H}_\ell(x)$ is bounded with positive constant $b_{\mathfrak{H}_\ell}$.*

Define the cost function as

$$\mathfrak{J}(x) = \int_t^\infty \psi(x, \mathfrak{U}, \mathfrak{V})d\upsilon, \tag{2}$$

where utility function $\psi(x, \mathfrak{U}, \mathfrak{V}) = x^T \mathcal{Q}x + \sum_{\imath=1}^{N_\mu} \mu_\imath^T \mathfrak{R}_\imath \mu_\imath - \gamma^2 \sum_{\ell=1}^{N_\nu} \nu_\ell^T \nu_\ell$. Here $\mathcal{Q}$ and $\mathfrak{R}_\imath$ are symmetric positive definite. To achieve the control objective of this work, the saddle points $\mathfrak{U}^*$ and $\mathfrak{V}^*$ should be derived which satisfies that

$$\mathfrak{J}(x, \mathfrak{U}^*, \mathfrak{V}) \leq \mathfrak{J}(x, \mathfrak{U}^*, \mathfrak{V}^*) \leq \mathfrak{J}(x, \mathfrak{U}, \mathfrak{V}^*), \tag{3}$$

where $\mathfrak{U}^* = \left\{\mu_1^*, \ldots, \mu_{N_\mu}^*\right\}$ and $\mathfrak{V}^* = \left\{\nu_1^*, \ldots, \nu_{N_\nu}^*\right\}$. Via differentiating the cost function (2), the Hamiltonian is derived

$$\mathcal{H}(x, \mathfrak{U}, \mathfrak{V}) = \psi(x, \mathfrak{U}, \mathfrak{V}) + \nabla\mathfrak{J}^T \left(\mathcal{F} + \sum_{\imath=1}^{N_\mu} \mathcal{G}_\imath\mu_\imath + \sum_{\ell=1}^{N_\nu} \mathfrak{H}_\ell\nu_\ell\right). \tag{4}$$

The stationary conditions are given as

$$\frac{\partial\mathcal{H}(\nabla\mathfrak{J}^*, x, \mathfrak{U}, \mathfrak{V})}{\partial\mu_\imath} = 0, \tag{5}$$

$$\frac{\partial\mathcal{H}(\nabla\mathfrak{J}^*, x, \mathfrak{U}, \mathfrak{V})}{\partial\nu_\ell} = 0, \tag{6}$$

where $\imath \in \mathbb{N}_\mu$ and $\ell \in \mathbb{N}_\nu$. Then the optimal control and worst disturbance are got as

$$\mu_\imath^* = -\frac{1}{2}\mathfrak{R}_\imath^{-1}\mathcal{G}_\imath^T\nabla\mathfrak{J}^*, \tag{7}$$

$$\nu_\ell^* = \frac{1}{2\gamma^2}\mathfrak{H}_\ell^T \nabla \mathfrak{J}^*. \tag{8}$$

In light of (4), (7) and (8), the HJIE can be formulated as

$$(\nabla \mathfrak{J}^*)^T \left(\mathcal{F}(x) - \frac{1}{4}\sum_{\imath=1}^{N_\mu} \mathcal{G}_\imath \mathfrak{R}_\imath^{-1} \mathcal{G}_\imath^T \nabla \mathfrak{J}^* + \frac{1}{4\gamma^2}\sum_{\ell=1}^{N_\nu} \mathfrak{H}_\ell \mathfrak{H}_\ell^T \nabla \mathfrak{J}^* \right) + x^T \mathcal{Q} x = 0 \tag{9}$$

with $\mathfrak{J}^*(0) = 0$.

From (9), it can be observed that the HJIE involves partial derivative parts and nonlinear terms, which increases the derivation difficulty of deriving the closed-form solution. Nonetheless, one can utilize ACL method to search the approximated solutions.

3. **Event-Triggered Mechanism Design.** Traditional control methods often achieve control objects with time-driven mechanism, which employs all the real-time data to compute/update control strategy and presents periodicity feature. However, for the systems like NCS where there often exist large amounts of data processing, using the gathered data indiscriminately is likely to cause communication delay, excessive consumption of system resources and even the decline of control performance [37, 38, 39]. To conquer this problem, ETM can be constructed which works as a data-filtration mechanism. The sequence $\{\delta_s\}_{s=1}^{\infty}$ is used for storing the selected data where s denotes the index. When the data are selected, we have the recorded state

$$\breve{x}_s(t) = x(t_s),\ t \in [\delta_s, \delta_{s+1}). \tag{10}$$

In general, ETM works on the event error given by

$$e_s = \breve{x}_s - x(t),\ t \in [\delta_s, \delta_{s+1}). \tag{11}$$

To judge whether an event occurs, the comparison is requisite between the error and pre-determined threshold. Additionally, from (11), one can deduce that at the recorded instants, the error becomes zero while in another triggering period, the error increases until reaching a new triggering instant. Based on the sequence $\{\delta_s\}_{s=1}^{\infty}$, we have

$$\breve{\mu}_\imath(t) = \mu_\imath(\breve{x}_s),\ t \in [\delta_s, \delta_{s+1}), \tag{12}$$

$$\breve{\nu}_\ell(t) = \nu_\ell(\breve{x}_s),\ t \in [\delta_s, \delta_{s+1}). \tag{13}$$

Based on (7) and (8), one can get

$$\breve{\mu}_\imath^*(t) = -\frac{1}{2}\mathfrak{R}_\imath^{-1}\mathcal{G}_\imath^T(\breve{x}_s)\nabla\breve{\mathfrak{J}}^*,\ t \in [\delta_s, \delta_{s+1}), \tag{14}$$

$$\breve{\nu}_\ell^*(t) = \frac{1}{2\gamma^2}\mathfrak{H}_\ell^T(\breve{x}_s)\nabla\breve{\mathfrak{J}}^*,\ t \in [\delta_s, \delta_{s+1}), \tag{15}$$

where $\nabla\breve{\mathfrak{J}}^* = \partial\mathfrak{J}^*/\partial x$ at the instant $t = \delta_s$. Let $\breve{\mathfrak{U}}^* = \left\{\breve{\mu}_1^*, \dots, \breve{\mu}_{N_\mu}^*\right\}$ and $\breve{\mathfrak{V}}^* = \left\{\breve{\nu}_1^*, \dots, \breve{\nu}_{N_\nu}^*\right\}$, and then the event-triggered version of HJIE can be rewritten as

$$\begin{aligned}
&\mathcal{H}\left(x, \breve{\mathfrak{U}}^*, \breve{\mathfrak{V}}^*, \mathfrak{J}^*\right) \\
&= (\nabla\mathfrak{J}^*)^T \left(\mathcal{F}(x) - \frac{1}{2}\sum_{\imath=1}^{N_\mu} \mathcal{G}_\imath \mathfrak{R}_\imath^{-1}\mathcal{G}_\imath^T \nabla\breve{\mathfrak{J}}^* + \frac{1}{2\gamma^2}\sum_{\ell=1}^{N_\nu}\mathfrak{H}_\ell\mathfrak{H}_\ell^T\nabla\breve{\mathfrak{J}}^* \right) + x^T\mathcal{Q}x \\
&\quad + \frac{1}{4}\sum_{\imath=1}^{N_\mu}\left(\nabla\breve{\mathfrak{J}}^*\right)^T \mathcal{G}_\imath\mathfrak{R}_\imath^{-1}\mathcal{G}_\imath^T\nabla\breve{\mathfrak{J}}^* - \frac{1}{4\gamma^2}\sum_{\ell=1}^{N_\nu}\left(\nabla\breve{\mathfrak{J}}^*\right)^T\mathfrak{H}_\ell\mathfrak{H}_\ell^T\nabla\breve{\mathfrak{J}}^*.
\end{aligned} \tag{16}$$

The following assumption is requisite for the triggering condition [40, 41].

Assumption 3.1. *For every* $x, \breve{x}_s \in \Theta$*, the following holds that* $\|\mu_\imath^* - \breve{\mu}_\imath^*\|^2 \leq c_\imath \|x - \breve{x}_s\|^2$ *with positive Lipschitz constant* $c_\imath$*.*

Theorem 3.1. *Consider system (1) and suppose Assumptions 2.1, 2.2 and 3.1 hold. Let* $\mathfrak{J}^*$ *be the solution of the HJIE (9). The control and disturbance strategies are updated by (14) and (15). Then system (1) is ensured to be stable with the triggering condition*

$$\|e_s\|^2 \leq \frac{1}{C_R}\left(\gamma_1 \lambda_{\min}(\mathcal{Q})\|x\|^2 + \Gamma_t\left(\breve{\mathfrak{U}}^*, \breve{\mathfrak{V}}^*\right)\right), \tag{17}$$

where $C_R = \sum_{\imath=1}^{N_\mu} c_\imath \|\mathfrak{R}_\imath\|$, $\gamma_1 \in (0,1)$ *is the regulated parameter and* $\Gamma_t\left(\breve{\mathfrak{U}}^*, \breve{\mathfrak{V}}^*\right) = \sum_{\imath=1}^{N_\mu} \breve{\mu}_\imath^{*T} \mathfrak{R}_\imath \breve{\mu}_\imath^* - \gamma^2 \sum_{\ell=1}^{N_\nu} \breve{\nu}_\ell^{*T} \breve{\nu}_\ell^*$*.*

Proof: Choose the Lyapunov function $\mathfrak{L}_1 = \mathfrak{J}^*(x)$, of which the derivative is given as

$$\dot{\mathfrak{L}}_1 = (\nabla \mathfrak{J}^*)^T \left(\mathcal{F}(x) + \sum_{\imath=1}^{N_\mu} \mathcal{G}_\imath \breve{\mu}_\imath^* + \sum_{\ell=1}^{N_\nu} \mathfrak{H}_\ell \breve{\nu}_\ell^*\right). \tag{18}$$

From (9), it can be derived

$$(\nabla \mathfrak{J}^*)^T \mathcal{F}(x) = \frac{1}{4}\sum_{\imath=1}^{N_\mu} \mathcal{G}_\imath \mathfrak{R}_\imath^{-1} \mathcal{G}_\imath^T \nabla \mathfrak{J}^* - \frac{1}{4\gamma^2}\sum_{\ell=1}^{N_\nu} \mathfrak{H}_\ell \mathfrak{H}_\ell^T \nabla \mathfrak{J}^* - x^T \mathcal{Q} x. \tag{19}$$

Based on (7), (8) and (19), one has

$$\begin{aligned}\dot{\mathfrak{L}}_1 =& \sum_{\imath=1}^{N_\mu} \left(\left(\mu_\imath^{*T} - \breve{\mu}_\imath^{*T}\right) \mathfrak{R}_\imath (\mu_\imath^* - \breve{\mu}_\imath^*) - \breve{\mu}_\imath^{*T} \mathfrak{R}_\imath \breve{\mu}_\imath^*\right) - x^T \mathcal{Q} x \\ &- \gamma^2 \sum_{\ell=1}^{N_\nu} \left(\left(\nu_\ell^{*T} - \breve{\nu}_\ell^{*T}\right)(\nu_\ell^* - \breve{\nu}_\ell^*) - \breve{\nu}_\ell^{*T} \breve{\nu}_\ell^*\right).\end{aligned} \tag{20}$$

Through using Young's inequality, one derives that

$$\dot{\mathfrak{L}}_1 \leq C_R \|e_s\|^2 - \Gamma_t\left(\breve{\mathfrak{U}}^*, \breve{\mathfrak{V}}^*\right) - \gamma_1 \lambda_{\min}(\mathcal{Q})\|x\|^2 - (1-\gamma_1)\lambda_{\min}(\mathcal{Q})\|x\|^2. \tag{21}$$

From (21), it can be deduced that triggering condition (17) can ensure the asymptotic stability of system (1). □

Remark 3.1. *Theorem 3.1 has demonstrated that with the utilization of condition (17), system states can be driven to be asymptotic stable. In (17), parameter* γ_1 *is used for regulating the frequency of event occurrence, that is, a smaller* γ_1 *will prompt the occurrence of event while a larger* γ_1 *will decline the frequency.*

4. **The Adaptive Critic Structure with the DETM.**

In this section, first, considering that the system dynamics is partially unknown, we utilize the IRL approach to tackle the multi-player game issue. Then, the DETM is integrated into the adaptive critic structure to approximate the solution of HJIE with less computation/communication costs. Finally, system states and critic weights are demonstrated to be UUB via using Lyapunov theory.

4.1. **Formulation of IRL approach.**

The IRL approach derived from RL algorithm, has the ability to search the solution where the system dynamics is unknown [42]. Thus, we can obtain the solutions of HJIE (9) approximately by virtue of IRL technique. For any $\mathcal{T}_I \in (0, t)$, $\mathfrak{J}^*$ can be presented in the form of integral Bellman equation [17].

$$\mathfrak{J}^*(x(t-\mathcal{T}_I)) = \mathfrak{J}^*(x(t)) + \int_{t-\mathcal{T}_I}^{t} \left(\psi(x, \mathfrak{U}^*, \mathfrak{V}^*)\right) d\varsigma. \tag{22}$$

When the admissible strategies are applied, then the following Lyapunov equation is derived.

$$\mathfrak{L}_E(\mathfrak{J}(x(t))) = \mathfrak{J}(x(t)) - \mathfrak{J}(x(t-\mathcal{T}_I)) + \int_{t-\mathcal{T}_I}^{t} \left(\psi(x, \mathfrak{U}, \mathfrak{V})\right) d\varsigma. \tag{23}$$

Remark 4.1. *Equation (22) reveals that when applying IRL technique to solving the game issues, the requirement for the information of system dynamics can be completely removed. Nevertheless, recalling the form of strategies in (7) and (8), we can deduce that the information associated with* $\mathcal{G}_\imath$ *and* $\mathfrak{H}_\ell$ *is necessary. Hence, in this work, the multi-player game issues of which the drift dynamics* $\mathcal{F}(x)$ *is unknown can be solved with the aid of IRL method.*

4.2. **The adaptive critic learning method.** Due to the universal approximation property of neural networks, the cost function can be given by

$$\mathfrak{J}^* = \omega^T \eta(x) + \tau, \tag{24}$$

where $\omega \in \mathbb{R}^{N_n}$, $\eta(x) \in \mathbb{R}^{N_n}$, $\tau \in \mathbb{R}$ and N_n are the ideal weight, activation function, reconstruction error and the neuro number for the critic network. Since the ideal weight is hard to acquire, the approximated one is presented

$$\hat{\mathfrak{J}} = \hat{\omega}^T \eta(x), \tag{25}$$

where the approximated weight $\hat{\omega} \in \mathbb{R}^{N_n}$.

Based on (7), (8) and (24), the event-based strategies are presented as

$$\breve{\mu}_\imath^* = -\frac{1}{2}\mathfrak{R}_\imath^{-1}\mathcal{G}_\imath^T(\breve{x}_s)\left((\nabla\eta(\breve{x}_s))^T\omega^* + \nabla\tau(\breve{x}_s)\right),\ t \in [\delta_s, \delta_{s+1}), \tag{26}$$

$$\breve{\nu}_\ell^* = \frac{1}{2\gamma^2}\mathfrak{H}_\ell^T(\breve{x}_s)\left((\nabla\eta(\breve{x}_s))^T\omega^* + \nabla\tau(\breve{x}_s)\right),\ t \in [\delta_s, \delta_{s+1}). \tag{27}$$

Similarly, applying the approximated cost (25), one has

$$\breve{\mu}_\imath = -\frac{1}{2}\mathfrak{R}_\imath^{-1}\mathcal{G}_\imath^T(\breve{x}_s)(\nabla\eta(\breve{x}_s))^T\hat{\omega}(\delta_s),\ t \in [\delta_s, \delta_{s+1}), \tag{28}$$

$$\breve{\nu}_\ell = \frac{1}{2\gamma^2}\mathfrak{H}_\ell^T(\breve{x}_s)(\nabla\eta(\breve{x}_s))^T\hat{\omega}(\delta_s),\ t \in [\delta_s, \delta_{s+1}). \tag{29}$$

Then, the approximated event-based Hamiltonian can be got

$$\begin{aligned}\mathfrak{L}_E\left(x, \breve{\mathfrak{U}}, \breve{\mathfrak{V}}, \hat{\mathfrak{J}}\right) &= \hat{\mathfrak{J}}(x(t)) - \hat{\mathfrak{J}}(x(t-\mathcal{T}_I)) + \int_{t-\mathcal{T}_I}^{t}\left(\psi\left(x, \breve{\mathfrak{U}}, \breve{\mathfrak{V}}\right)\right) d\varsigma \\ &= \hat{\omega}^T\theta(x) + \Upsilon\left(x, \breve{\mathfrak{U}}, \breve{\mathfrak{V}}\right) \triangleq \epsilon,\end{aligned} \tag{30}$$

where $\theta(x) = \eta(x(t)) - \eta(x(t-\mathcal{T}_I))$ and $\Upsilon\left(x, \breve{\mathfrak{U}}, \breve{\mathfrak{V}}\right) = \int_{t-\mathcal{T}_I}^{t}\left(\psi\left(x, \breve{\mathfrak{U}}, \breve{\mathfrak{V}}\right)\right) d\varsigma$. It is noted that the terms ϵ should be minimized via critic networks. Define the auxiliary item and weight error as $\varepsilon = \omega^{*T}\theta + \Upsilon$ and $\tilde{\omega} = \omega^* - \hat{\omega}$, then $\epsilon = \varepsilon - \tilde{\omega}^T\theta$. Based on gradient descent technique, we derive the critic update law

$$\dot{\hat{\omega}} = -\alpha\frac{\theta\varepsilon}{(\theta^T\theta + 1)^2} = -\alpha\breve{\theta}\varepsilon = -\alpha\breve{\theta}\varepsilon + \alpha\bar{\theta}\bar{\theta}^T\tilde{\omega}, \tag{31}$$

where $\alpha > 0$ is a critic learning coefficient, $\bar{\theta} = \theta/\left(\theta^T\theta + 1\right)$ and $\breve{\theta} = \theta/\left(\theta^T\theta + 1\right)^2$.

The derivation of $\tilde{\omega}$ is given as

$$\dot{\tilde{\omega}} = \alpha\breve{\theta}\varepsilon - \alpha\bar{\theta}\bar{\theta}^T\tilde{\omega}. \tag{32}$$

4.3. **Stability analysis.** Based on the traditional triggering mechanism, we construct the DETM with dead-zone operation which can drive control strategies to be updated in the manner more efficiently. For the sake of constructing DETM, an internal dynamic variable is indispensable defined as

$$\dot{\rho} = -\beta\rho + \Phi(x, e_s), \ \rho_0 = \rho(0) \geq 0, \tag{33}$$

where $\beta > 0$, $\Phi(x, e_s) = \gamma_2\lambda_{\min}(\mathcal{Q})\|x\|^2 + \Gamma_1\left(\breve{\mathfrak{U}}\right) - 2C_R\|e_s\|^2$ and $\Gamma_1\left(\breve{\mathfrak{U}}\right) = \sum_{\imath=1}^{N_\mu}\breve{\mu}_\imath^T\mathfrak{R}_\imath\breve{\mu}_\imath$. In fact, one can view (33) as a filter, and ρ is the filtered value for $\Phi(x, e_s)$. Additionally, $\Phi(x, e_s)$ includes all the static triggering information.

Via introducing variable ρ, we can construct the DETM as

$$\delta_{s+1} = \inf\left\{t > \delta_s : \Gamma(e_s) \geq \frac{1}{2C_R}\left(\gamma_2\lambda_{\min}(\mathcal{Q})\|x\|^2 + \Gamma_1\left(\breve{\mathfrak{U}}\right)\right) + T_f \triangleq \|T_d\|^2\right\}, \tag{34}$$

where $T_f = \rho/(2C_R\kappa)$ and

$$\Gamma(e_s) = \begin{cases} \|e_s\|^2, & \text{if } \|x\| > T_o; \\ 0, & \text{otherwise.} \end{cases} \tag{35}$$

For the clarity of illustration, the event-based critic learning approach is concluded in Algorithm 1.

Algorithm 1: Event-triggered adaptive control for nonlinear multi-player games using neural critic learning

```
Complete the algorithm initialization. Set the initial state $x_0$, the initial weight
vector $\hat{\omega}(0)$, the sampling period 0.01s, and neural learning time $t_{\max}$;
while $t < t_{\max}$ do
    Acquire the sampling state $\breve{x}_s(t)$ and the triggering error $e_s(t)$;
    Update the critic weight $\hat{\omega}$ with (31) and $\rho$ with (33);
    Verdict whether the triggering mechanism is activated in light of (35);
    if The control precision is not achieved then
        if Triggering condition is satisfied then
            Calculate the strategies with (28) and (29);
            Set $e_s = 0$;
        else
            Update $e_s(t) = \breve{x}_s(t) - x(t)$ and keep the strategies unchanged with ZOH;
        end
    else
        Unactivate the triggering mechanism and keep the strategies unchanged;
    end
    Update $x(t)$ with the strategies.
end
```

Lemma 4.1. *When triggering mechanism (34) is applied, variable ρ can be ensured to be positive.*

Proof: When $\kappa = 0$, we can deduce that $\rho \geq 0$. When $\kappa > 0$, based on (34) and (35) we can derive

$$\dot{\rho}(t) \geq -\left(\beta + \frac{1}{\kappa}\right)\rho(t),\ \ \rho_0 \geq 0. \tag{36}$$

Through applying comparison lemma, we have

$$\rho(t) \geq \rho_0 e^{-\left(\beta+\frac{1}{\kappa}\right)t}. \tag{37}$$

The inequation reveals that $\rho > 0$ if $\kappa > 0$. □

Before presenting the discussions for system stability, an assumption is necessary [43, 45].

Assumption 4.1. *The terms ω^*, $\nabla\eta$, $\nabla\epsilon$, $\nabla\tau$ and ε are bounded with positive constants. That is, $\|\omega^*\| \leq b_\omega$, $\|\nabla\eta\| \leq b_\eta$, $\|\nabla\epsilon\| \leq b_\epsilon$, $\|\nabla\tau\| \leq b_\tau$ and $\|\varepsilon\| \leq b_\varepsilon$ with positive constants b_ω, b_η, b_ϵ, b_τ and b_ε.*

Theorem 4.1. *Consider system (1) and let Assumptions 2.1, 2.2, 3.1 and 4.1 hold. The weights for critic networks are updated with the update laws (31). Then when the conditions (34) and (38) hold, the system state x and the critic weight errors $\tilde{\omega}$ are driven to be UUB with the control strategies (28)*

$$\frac{\alpha}{2}\lambda_{\min}\left(\Lambda\left(\bar{\theta}\right)\right) - C_{\omega 1} > 0. \tag{38}$$

Proof: Consider the Lyapunov candidate

$$\mathfrak{L}_y = \mathfrak{L}_{y1} + \mathfrak{L}_{y2} + \mathfrak{L}_{y3} + \mathfrak{L}_{y4}, \tag{39}$$

where $\mathfrak{L}_{y1} = \mathfrak{J}^*(\breve{x}_s)$, $\mathfrak{L}_{y2} = \mathfrak{J}^*(x)$, $\mathfrak{L}_{y3} = \frac{1}{2}\tilde{\omega}^T\tilde{\omega}$ and $\mathfrak{L}_{y4} = \rho$. Since there exists dynamics jump, the discussions are given in two cases.

Case i. No events occur when $t \in [\delta_s, \delta_{s+1})$. In this case, $\dot{\mathfrak{L}}_{y1} = 0$. Then according to (19), the derivative of $\mathfrak{L}_{y2}$ is given

$$\dot{\mathfrak{L}}_{y2} = (\nabla\mathfrak{J}^*)^T\left(\mathcal{F}(x) + \sum_{\imath=1}^{N_\mu}\mathcal{G}_\imath\breve{\mu}_\imath + \sum_{\ell=1}^{N_\nu}\mathfrak{H}_\ell\breve{\nu}_\ell\right). \tag{40}$$

Based on (9), (40) can be reformulated as

$$\begin{aligned}\dot{\mathfrak{L}}_{y2} =& -\gamma^2\sum_{\ell=1}^{N_\nu}\left(\left(\nu_\ell^{*T} - \breve{\nu}_\ell^T\right)\left(\nu_\ell^* - \breve{\nu}_\ell\right) - \breve{\nu}_\ell^T\breve{\nu}_\ell\right) - x^T\mathcal{Q}x\\ &+\sum_{\imath=1}^{N_\mu}\left(\left(\mu_\imath^{*T} - \breve{\mu}_\imath^T\right)\mathfrak{R}_\imath(\mu_\imath^* - \breve{\mu}_\imath) - \breve{\mu}_\imath^T\mathfrak{R}_\imath\breve{\mu}_\imath\right)\end{aligned} \tag{41}$$

Via using Young's inequation, one has

$$\begin{aligned}\dot{\mathfrak{L}}_{y2} \leq& \gamma^2\sum_{\ell=1}^{N_\nu}\breve{\nu}_\ell^T\breve{\nu}_\ell + \sum_{\imath=1}^{N_\mu}\left(\|\mathfrak{R}_\imath\|\|\mu_\imath^* - \breve{\mu}_\imath\|^2 - \breve{\mu}_\imath^T\mathfrak{R}_\imath\breve{\mu}_\imath\right) - x^T\mathcal{Q}x\\ \leq& \frac{1}{4\gamma^2}\sum_{\ell=1}^{N_\nu}\left\|\mathfrak{H}_\ell^T(\breve{x}_s)\nabla\eta^T(\breve{x}_s)\hat{\omega}\right\|^2 + \sum_{\imath=1}^{N_\mu}\left(\|\mathfrak{R}_\imath\|\|\mu_\imath^* - \breve{\mu}_\imath\|^2 - \breve{\mu}_\imath^T\mathfrak{R}_\imath\breve{\mu}_\imath\right) - x^T\mathcal{Q}x.\end{aligned} \tag{42}$$

According to (26) and (28), we gain

$$\begin{aligned}\|\mu_\imath^* - \breve{\mu}_\imath\|^2 \leq& 2c_\imath\|e_s\|^2 + \frac{1}{2}\left\|\mathfrak{R}_\imath^{-1}\right\|^2 b_{\mathcal{G}\imath}^2\left\|\nabla\eta^T(\breve{x}_s)\tilde{\omega} + \nabla\tau^T(\breve{x}_s)\right\|^2\\ \leq& 2c_\imath\|e_s\|^2 + \left\|\mathfrak{R}_\imath^{-1}\right\|^2 b_{\mathcal{G}\imath}^2\left(b_\eta^2\|\tilde{\omega}\|^2 + b_\tau^2\right).\end{aligned} \tag{43}$$

Substituting (43) into (42), one can derive

$$\begin{aligned}\dot{\mathfrak{L}}_{y2} &\leq \sum_{\imath=1}^{N_\mu}\left(\|\mathfrak{R}_\imath\|\|\mu_\imath^* - \breve{\mu}_\imath\|^2 - \breve{\mu}_\imath^T\mathfrak{R}_\imath\breve{\mu}_\imath\right) - x^T\mathcal{Q}x + \frac{b_{\mathfrak{H}}}{4\gamma^2}b_\eta^2\|\omega^* - \tilde{\omega}\| \\ &\leq 2C_R\|e_s\|^2 - \Gamma_1\left(\breve{\mathfrak{U}}\right) + C_{\omega1}\|\tilde{\omega}\|^2 - x^T\mathcal{Q}x + k_1, \qquad (44)\end{aligned}$$

where $b_{\mathfrak{H}} = \sum_{\ell=1}^{N_\nu} b_{\mathfrak{H}\ell}^2$, $C_{\omega1} = (1/(2\gamma^2))\,b_{\mathfrak{H}}b_\eta^2 + \sum_{\imath=1}^{N_\mu}\|\mathfrak{R}_\imath\|\left\|\mathfrak{R}_\imath^{-1}\right\|^2 b_{\mathcal{G}\imath}^2 b_\eta^2$, and the term $k_1 = (1/(2\gamma^2))\,b_{\mathfrak{H}}b_\omega^2 b_\eta^2 + \sum_{\imath=1}^{N_\mu}\|\mathfrak{R}_\imath\|\left\|\mathfrak{R}_\imath^{-1}\right\|^2 b_{\mathcal{G}\imath}^2 b_\tau^2$.

From (32), we get

$$\dot{\mathfrak{L}}_{y3} = -\alpha\tilde{\omega}^T\bar{\theta}\bar{\theta}^T\tilde{\omega} + \alpha\tilde{\omega}^T\breve{\theta}\varepsilon. \qquad (45)$$

Based on Young's inequality, it is gained that

$$\alpha\tilde{\omega}^T\breve{\theta}\varepsilon \leq \frac{\alpha}{2}\left(\tilde{\omega}^T\bar{\theta}\bar{\theta}^T\tilde{\omega} + \varepsilon^T\varepsilon\right). \qquad (46)$$

Substituting (46) into (45), we derive

$$\dot{\mathfrak{L}}_{y3} \leq -\frac{\alpha}{2}\lambda_{\min}\left(\Lambda\left(\bar{\theta}\right)\right)\|\tilde{\omega}\|^2 + \frac{\alpha}{2}b_\varepsilon^2. \qquad (47)$$

Integrating (33), (44) and (47), one has

$$\dot{\mathfrak{L}}_y \leq -(1-\gamma_2)\lambda_{\min}(\mathcal{Q})\|x\|^2 - \left(\frac{\alpha}{2}\lambda_{\min}\left(\Lambda\left(\bar{\theta}\right)\right) - C_{\omega1}\right)\|\tilde{\omega}\|^2 + b_0, \qquad (48)$$

with $b_0 = \alpha b_\varepsilon^2/2 + k_1$. Considering (34) and (38), one derives $\dot{\mathfrak{L}}_y < 0$ when either of the following conditions holds:

$$\|x\| > \sqrt{\frac{b_0}{(1-\gamma_2)\lambda_{\min}(\mathcal{Q})}} \triangleq b_1, \qquad (49)$$

and

$$\|\tilde{\omega}\| > \sqrt{\frac{b_0}{\frac{\alpha}{2}\lambda_{\min}\left(\Lambda\left(\bar{\theta}\right)\right) - C_{\omega1}}} \triangleq b_2. \qquad (50)$$

Case ii. An event is triggered when $t = \delta_{s+1}$. Then we present the difference of $\mathfrak{L}$

$$\begin{aligned}\Delta\mathfrak{L}_y = {}& \underbrace{\mathfrak{J}^*(x(\delta_{s+1})) - \mathfrak{J}^*(x(\delta_{s+1}^-))}_{\Delta\mathfrak{L}_{y1}} + \underbrace{\mathfrak{J}^*(\breve{x}_{s+1}) - \mathfrak{J}^*(\breve{x}_s)}_{\Delta\mathfrak{L}_{y2}} \\ & + \underbrace{\frac{1}{2\alpha}\left(\tilde{\omega}^T(\delta_{s+1})\tilde{\omega}(\delta_{s+1}) - \tilde{\omega}^T(\delta_{s+1}^-)\tilde{\omega}(\delta_{s+1}^-)\right)}_{\Delta\mathfrak{L}_{y3}} \\ & + \underbrace{\rho(\delta_{s+1}) - \rho(\delta_{s+1}^-)}_{\Delta\mathfrak{L}_{y4}}. \qquad (51)\end{aligned}$$

From the analysis above, we know that the system stability can be guaranteed. Recall that $\mathfrak{L}_y$ is decreasing, and then, based on the properties of limits, we have

$$\begin{aligned}0 < {}& \left(\mathfrak{J}^*(x(\delta_{s+1}^-)) - \mathfrak{J}^*(x(\delta_{s+1}))\right) + \rho(\delta_{s+1}^-) - \rho(\delta_{s+1}) \\ & + \left(\frac{1}{2\alpha}\tilde{\omega}^T(\delta_{s+1}^-)\tilde{\omega}(\delta_{s+1}^-) - \frac{1}{2\alpha}\tilde{\omega}^T(\delta_{s+1})\tilde{\omega}(\delta_{s+1})\right). \qquad (52)\end{aligned}$$

In addition, we gain that

$$\Delta\mathfrak{L}_{y2} = \mathfrak{J}^*(\breve{x}_{s+1}) - \mathfrak{J}^*(\breve{x}_s) \leq 0. \qquad (53)$$

This completes the proof. □

Remark 4.2. *With the introduction of* ρ*, the traditional ETM evolves the triggering mechanism which is endowed with dynamic characteristic. Meanwhile, the triggering threshold is enlarged to lower the frequency of event occurrence. Similar to the operations in (34), the static triggering threshold is presented to show the difference visibly*

$$T_s = \sqrt{\frac{1}{2C_R}\left(\gamma_2\lambda_{\min}(\mathcal{Q})\|x\|^2 + \Gamma_1\left(\breve{\mathfrak{U}}\right)\right)}. \tag{54}$$

Comparing (34) and (54), one can observe that there exists T_f *for the dynamic mechanism, which is the key to the triggering threshold. In an extreme case, when* T_f *becomes zero, the dynamic mechanism becomes the static one. Furthermore, the dead-zone operation in (35) works to reduce unnecessary computations with regard to ETM. More precisely, the DETM works only when the system state exceeds the pre-determined bound. Thus, we can ensure the achievement of control target with a lower computation cost.*

Remark 4.3. *In this section, we construct the ACL method integrating the dynamic triggering mechanism with dead-zone operation, which can ensure the system signals to be UUB. Generally, it is imperative to analyze whether the event-based control schemes can avoid the Zeno phenomenon. We can use comparison lemma to demonstrate that there exists a positive bound for the evolution of* $e_s(t)/x(t)$*. Furthermore, Zeno phenomenon can be demonstrated to be avoided with the technique similar to that in* [44, 45].

5. **Simulations.**

In this section, two examples are simulated to validate the proposed ACL method.

Example 5.1. *Consider the linear system*

$$\dot{x} = \begin{bmatrix} -0.2803 & -0.00065 \\ 0.35 & -0.25 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 0 \\ 0.12 \end{bmatrix}\mu_1 + \begin{bmatrix} 0.15 \\ 0.3 \end{bmatrix}\mu_2 + \begin{bmatrix} 0 \\ 0.25 \end{bmatrix}\nu, \tag{55}$$

where $x = [x_1, x_2]^T \in \mathbb{R}^2$ *is system state,* $\mu_1 \in \mathbb{R}$*,* $\mu_2 \in \mathbb{R}$ *and* $\nu \in \mathbb{R}$ *are strategies for the players.*

The experiment is designed with the following parameters. Let $\mathcal{Q} = 2I_{2\times 2}$*,* $\mathfrak{R}_1 = I_{2\times 2}$*,* $\mathfrak{R}_2 = 1.5I_{2\times 2}$*. The activation functions for critic learning are* $\eta(x) = \left[x_1^2, x_2^2, x_1x_2, x_1^4, x_2^4, x_1^3x_2, x_1^2x_2^2, x_1x_2^3\right]^T$*. The initial weights are set with the random parameters in* $[-0.5, 0.5]$*. The state is initialized by* $x_0^T = [-1, 1]$*. The other parameters are set as* $c_1 = c_2 = 6$*,* $\gamma_2 = 0.1$*,* $T_o = 0.00001$*,* $\rho_0 = 1$*,* $\beta = 0.1$*, and* $\kappa = 6$*. The simulation time is set as 120s and the probing noises are injected to facilitate the neural learning before* $t = 80s$*.*

The simulation results are shown in Figures 1-5. Figure 1 provides the weight evolutions of critic network, which shows the convergence of critic weights. Figure 2 indicates the stability of the addressed system. From Figures 3 and 4, one can observe the control and disturbance strategies are updated in nonperiodic form. To visually present the advantage of the constructed dynamic triggering mechanism, we designed the comparison experiments between the DETM-based method and ETM-based method in the context of the same parameters. In this comparison experiments, the triggering number for static ETM is 1631, while the number for DETM is only 973. Figure 5 provides the curves of the number that the sampled data employed by the critic learning method. From the evolution for the event-based method, one can observe that the triggering number no longer increases since the preliminary operation works.

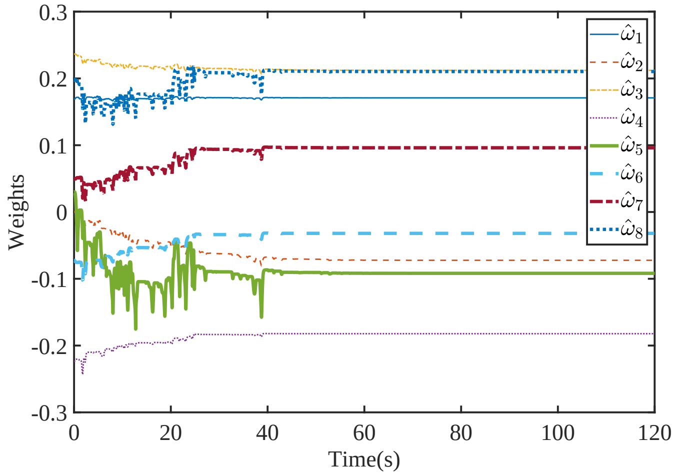


FIGURE 1. The evolutions of critic weights

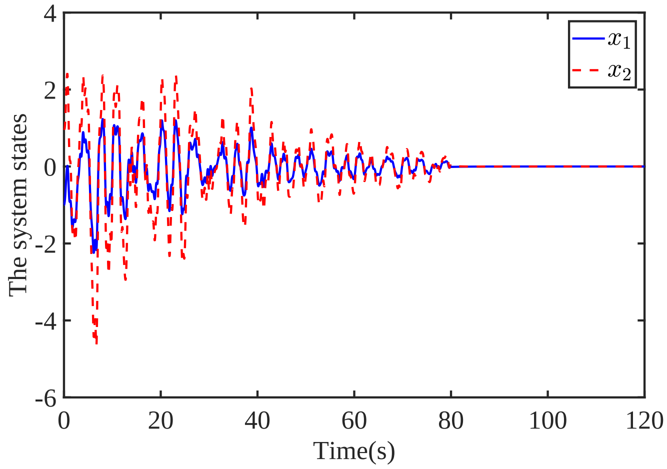


FIGURE 2. The system states

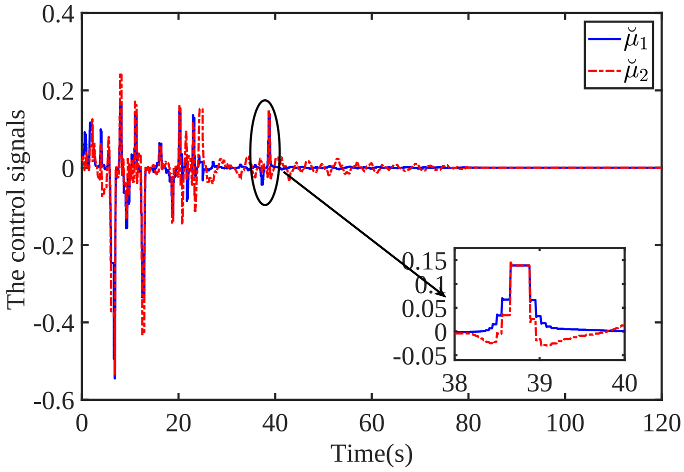


FIGURE 3. The control signals

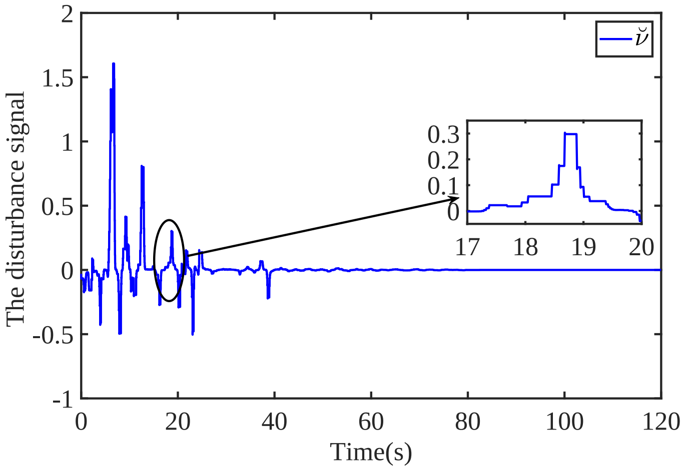


FIGURE 4. The disturbance signal

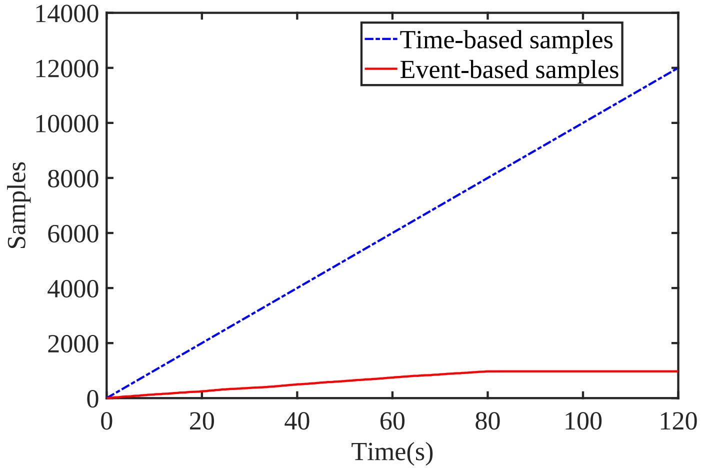


FIGURE 5. The evolutions of the triggering numbers

Example 5.2. *Consider the nonlinear system*

$$\dot{x} = \begin{bmatrix} 0.5x_2 - 0.8x_1 \\ -0.45x_1 - 0.6x_2 + 0.3x_1^2\cos(x_2) \end{bmatrix} \begin{bmatrix} 0.2 \\ 0.5 \end{bmatrix} \mu_1 + \begin{bmatrix} 0.24 \\ 0.3\sin(x_1) \end{bmatrix} \mu_2 + \begin{bmatrix} 0.1 \\ 0.6 \end{bmatrix} \mu_3 + \begin{bmatrix} \sin(x_1)\cos(x_1) \\ 1 \end{bmatrix} \nu_1 + \begin{bmatrix} 0 \\ x_1 + 1 \end{bmatrix} \nu_2, \qquad (56)$$

where $x = [x_1, x_2]^T \in \mathbb{R}^2$ *is system state,* $\mu_\imath$ *and* ν_ℓ *are the strategies. Let* $\mathcal{Q} = 3I_{2\times2}$, $\mathfrak{R}_1 = 2I_{2\times2}$, $\mathfrak{R}_2 = I_{2\times2}$, $\mathfrak{R}_3 = 2I_{2\times2}$. *The activation functions and the selection of initial weights for critic learning are the same as that in Example 5.1. The state is initialized with* $x_0^T = [0.5, 0.5]$. *The other parameters are set with* $c_1 = c_2 = 6$, $c_3 = 5$, $\gamma_2 = 0.1$, $T_o = 0.00001$, $\rho_0 = 1$, $\beta = 0.1$ *and* $\kappa = 6$.

The experiment results are given in Figures 6-12. Figure 6 presents the convergence evolutions of critic weights. Figure 7 reveals that the derived strategies can drive system state to reach the neighbourhood of the origin. The evolutions of control strategies in Figure 8 and the disturbances in Figure 9 present jump property caused by the proposed event-driven mechanism. Similar to the case in Example 5.1, we designed comparison experiments between the DETM-based critic learning method and the ETM-based one with the same parameters. Figure 10 shows that the dynamic judging thresholds are larger, and

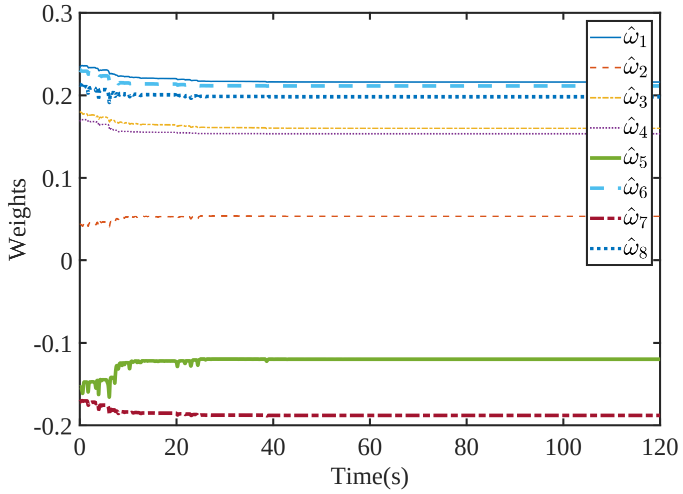


FIGURE 6. The evolutions of critic weights

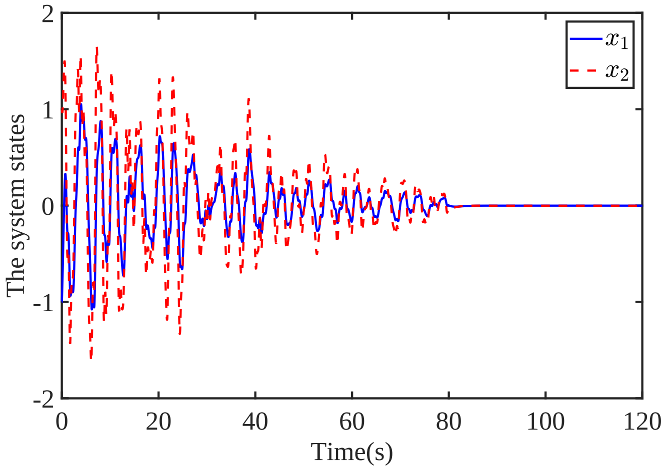


FIGURE 7. The system states

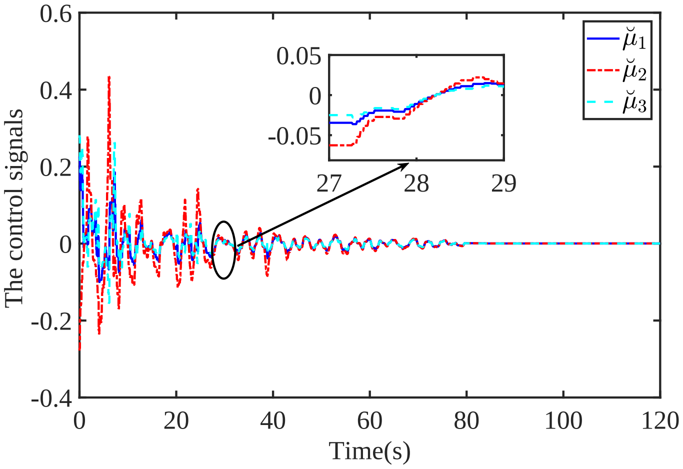


FIGURE 8. The control signals

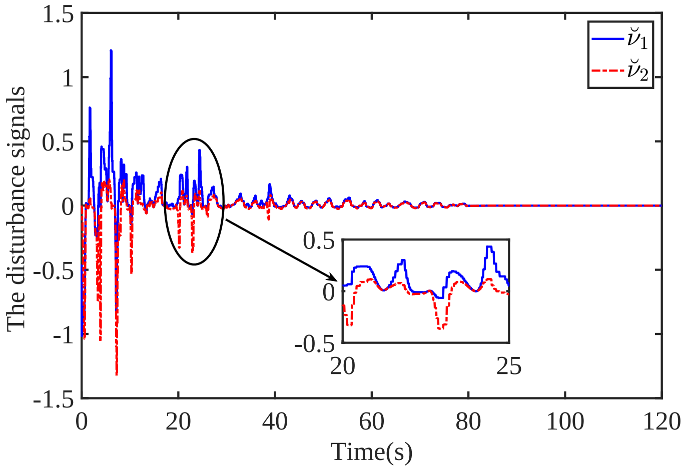
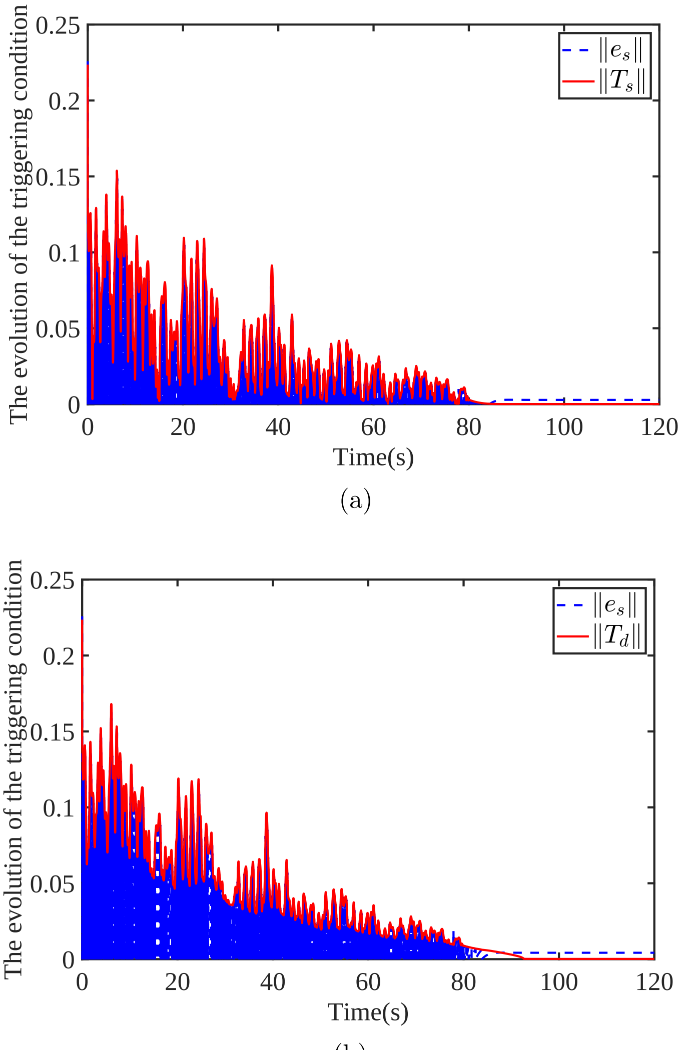


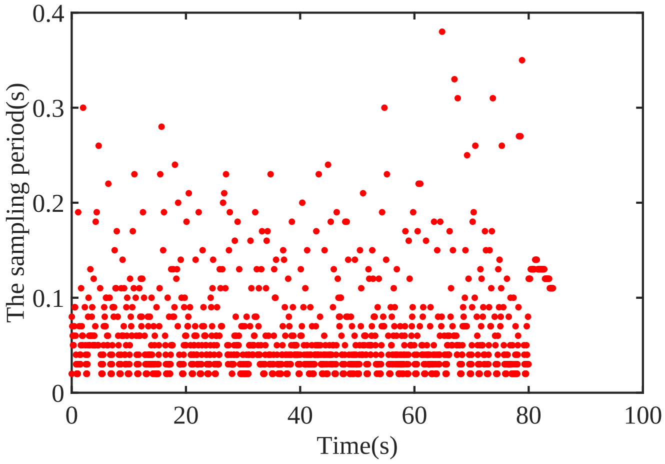
FIGURE 9. The disturbance signals



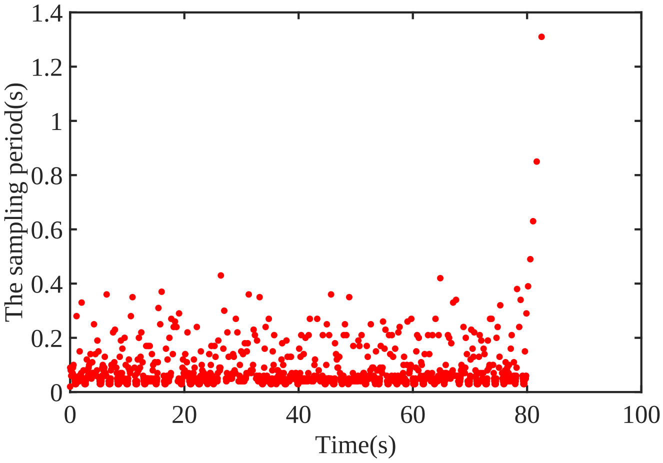
(a)

(b)

FIGURE 10. The static triggering condition and dynamic triggering condition



(a)



(b)

FIGURE 11. The sampling period for static triggering mechanism and dynamic triggering mechanism

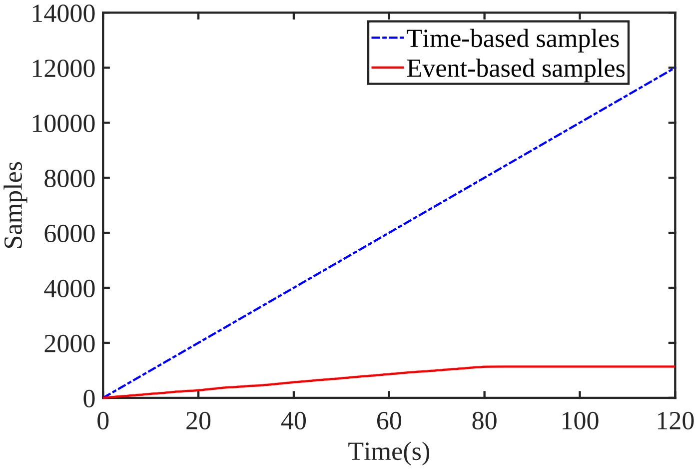


FIGURE 12. The evolutions of triggering numbers

from Figure 11 one can deduce that the dynamic triggering periods are also larger than the static ones. Besides, in Figure 10(b), after $t = 85s$ *the evolutions for* $\|e_s\|$ *pass through the triggering threshold evolutions. Then* $\|e_s\|$ *is always larger than* $\|T_d\|$ *or* $\|T_s\|$*, and does not return to zero anymore. This phenomenon is caused by activity results of the preliminary operation. In (35), when* x *comes close enough to the origin, dead-zone mechanism works and the ETM is turned off such that the unnecessary communication/computation can be avoided. Figure 12 indicates that during the whole critic learning process, the proposed method only utilizes 1137 data, while the time-based method requires for 12000 sampling data.*

To sum up, the experiments results provide the validations and supports for the proposed ACL method from multi-perspective.

6. **Conclusion.** In this paper, a dynamic event-triggered ACL approach was proposed to solve nonlinear multi-player games. The HJIE for the nonlinear game system is derived which provides the foundation of implementing adaptive critic learning. Then, integrating the IRL technique, the ACL method of critic-only structure is constructed to approximate the strategies for all the players without the drift dynamics information. Furthermore, dynamic triggering control mechanism with the preliminary operation is established to reduce the computation/communication costs of the proposed method. Finally, the effectiveness of the proposed approach is validated by Lyapunov theorem and two simulation experiments. In the future, we will further study game issues for stochastic nonlinear systems using critic learning with triggering mechanism.

Acknowledgment. This work was supported by the National Key R&D Program of China (2022YFE0107300), Science and Technology Research Project of Chongqing Municipal Education Commission (KJZD-K202300807, KJQN202300816), High-level Talents Research Project of CTBU (2356021), Key Platform Open Project of CTBU (KFJJ-2019062).

REFERENCES

[1] Y. Li, T. Yang and S. Tong, Adaptive neural networks finite-time optimal control for a class of nonlinear systems, *IEEE Trans. Neural Netw. Learn. Syst.*, vol.31, no.11, pp.4451-4460, 2020.

[2] Y. Zhang, B. Zhao, D. Liu and S. Zhang, Event-triggered control of discrete-time zero-sum games via deterministic policy gradient adaptive dynamic programming, *IEEE Trans. Syst., Man, Cybern.: Syst.*, vol.52, no.8, pp.4823-4835, 2022.

[3] Y. Ji, S. Liu, M. Zhou, Z. Zhao, X. Guo and L. Qi, A machine learning and genetic algorithm-based method for predicting width deviation of hot-rolled strip in steel production systems, *Inf. Sci.*, vol.589, pp.360-375, 2022.

[4] C. Mu, K. Wang, Z. Ni and C. Sun, Cooperative differential game-based optimal control and its application to power systems, *IEEE Trans. Ind. Inf.*, vol.16, no.8, pp.5169-5179, 2020.

[5] M. Liang, Y. Wang and D. Liu, An efficient impulsive adaptive dynamic programming algorithm for stochastic systems, *IEEE Trans. Cybern.*, DOI: 10.1109/TCYB.2022.3158898, 2022.

[6] M. Zhang, L. Li, J. Li, J. Duan, X. Yang and Q. Yuan, A knowledge-guided reinforcement learning model for news recommendation, *ICIC Express Letters, Part B: Applications*, vol.15, no.2, pp.117-128, 2024.

[7] X. Cui, B. Peng, B. Wang and L. Wang, Event-triggered neural experience replay learning for nonzero-sum tracking games of unknown continuous-time nonlinear systems, *Int. J. Robust Nonlinear Control*, vol.33, no.12, pp.6553-6575, 2023.

[8] C. Chen and D. Wang, Path planning of mobile robot based on the improved Q-learning algorithm, *International Journal of Innovative Computing, Information and Control*, vol.18, no.3, pp.687-702, 2022.

[9] B. Dong, T. An, X. Zhu, Y. Li and K. Liu, Zero-sum game-based neuro-optimal control of modular robot manipulators with uncertain disturbance using critic only policy iteration, *Neurocomputing*, vol.450, pp.183-196, 2021.

[10] A. Asgharnia, H. Schwartz and M. Atia, Learning multi-objective deception in a two-player differential game using reinforcement learning and multi-objective genetic algorithm, *International Journal of Innovative Computing, Information and Control*, vol.18, no.6, pp.1667-1688, 2022.

[11] A. W. Starr and Y. C. Ho, Nonzero-sum differential games, *J. Optim. Theory Appl.*, vol.3, no.3, pp.184-206, 1969.

[12] V. G. Lopez, F. L. Lewis, M. Liu, Y. Wan, S. Nageshrao and D. Filev, Game-theoretic lane-changing decision making and payoff learning for autonomous vehicles, *IEEE Trans. Veh. Techn.*, vol.71, no.4, pp.3609-3620, 2022.

[13] X. Yang, M. Xu and Q. Wei, Adaptive dynamic programming for nonlinear-constrained H_∞ control, *IEEE Trans. Syst., Man, Cybern.: Syst.*, vol.53, no.7, pp.4393-4403, 2023.

[14] C. Mu, K. Wang and Z. Ni, Adaptive learning and sampled-control for nonlinear game systems using dynamic event-triggering strategy, *IEEE Trans. Neural Netw. Learn. Syst.*, vol.33, no.9, pp.4437-4450, 2022.

[15] Q. Wei, L. Zhu, R. Song, P. Zhang, D. Liu and J. Xiao, Model-free adaptive optimal control for unknown nonlinear multiplayer nonzero-sum game, *IEEE Trans. Neural Netw. Learn. Syst.*, vol.33, no.2, pp.879-892, 2022.

[16] C. Qin, Z. Shang, Z. Zhang, D. Zhang and J. Zhang, Robust tracking control for non-zero-sum games of continuous-time uncertain nonlinear systems, *Mathematics*, vol.10, no.11, 1904, 2022.

[17] Q. Zhang and D. Zhao, Data-based reinforcement learning for nonzero-sum games with unknown drift dynamics, *IEEE Trans. Cybern.*, vol.49, no.8, pp.2874-2884, 2019.

[18] C. Deng, M. J. Er, G.-H. Yang and N. Wang, Event-triggered consensus of linear multiagent systems with time-varying communication delays, *IEEE Trans. Cybern.*, vol.50, no.7, pp.2916-2925, 2020.

[19] N. Zhao, X. Zhao, N. Xu and L. Zhang, Resilient event-triggered control of connected automated vehicles under cyber attacks, *IEEE/CAA J. Autom. Sinica*, DOI: 10.1109/JAS.2023.123483, 2023.

[20] H. Zhang, N. Zhao, S. Wang and R. K. Agarwal, Improved event-triggered dynamic output feedback control for networked T-S fuzzy systems with actuator failure and deception attacks, *IEEE Trans. Cybern.*, DOI: 10.1109/TCYB.2023.3264820, 2023.

[21] P. Chen, X. Luan, H. Wan and F. Liu, Event-triggered model reference adaptive sliding mode control with unknown direction control gain, *ICIC Express Letters*, vol.18, no.2, pp.159-165, 2024.

[22] C. Deng, C. Wen, W. Wang, X. Li and D. Yue, Distributed adaptive tracking control for high-order nonlinear multiagent systems over event-triggered communication, *IEEE Trans. Autom. Control*, vol.68, no.2, pp.1176-1183, 2023.

[23] Y. Zhang, B. Zhao, D. Liu and S. Zhang, Adaptive dynamic programming-based event-triggered robust control for multiplayer nonzero-sum games with unknown dynamics, *IEEE Trans. Cybern.*, vol.53, no.8, pp.5151-5164, 2023.

[24] R. Dahal and I. Kar, Robust tracking control of nonlinear unmatched uncertain systems via event-based adaptive dynamic programming, *Nonlinear Dyn.*, vol.109, pp.2831-2850, 2022.

[25] M. Lin, B. Zhao and D. Liu, Event-triggered robust adaptive dynamic programming for multiplayer Stackelberg-Nash games of uncertain nonlinear systems, *IEEE Trans. Cybern.*, DOI: 10.1109/TCYB.2023.3251653, 2024.

[26] D. Wang, L. Hu, M. Zhao and J. Qiao, Dual event-triggered constrained control through adaptive critic for discrete-time zero-sum games, *IEEE Trans. Syst., Man, Cybern.: Syst.*, vol.53, no.3, pp.1584-1595, 2023.

[27] Y. Zhao, B. Niu, G. Zong, N. Xu and A. M. Ahmad, Event-triggered optimal decentralized control for stochastic interconnected nonlinear systems via adaptive dynamic programming, *Neurocomputing*, vol.539, 126163, 2023.

[28] L. Dou, S. Cai, X. Zhang, X. Su and R. Zhang, Event-triggered-based adaptive dynamic programming for distributed formation control of multi-UAV, *J. Frankl. Inst.*, vol.359, no.8, pp.3671-3691, 2022.

[29] A. Girard, Dynamic triggering mechanisms for event-triggered control, *IEEE Trans. Autom. Control*, vol.60, no.7, pp.1992-1997, 2015.

[30] C. Qin, X. Qiao, J. Wang, D. Zhang, Y. Hou and S. Hu, Barrier-critic adaptive robust control of nonzero-sum differential games for uncertain nonlinear systems with state constraints, *IEEE Trans. Syst., Man, Cybern.: Syst.*, vol.54, no.1, pp.50-63, 2024.

[31] X. Cui, B. Wang, L. Wang and J. Chen, Online optimal learning algorithm for Stackelberg games with partially unknown dynamics and constrained inputs, *Neurocomputing*, vol.445, pp.1-11, 2021.

[32] R. Song, L. Liu and B. Hu, Aperiodic sampling artificial-actual H_∞ optimal control for interconnected constrained systems, *IEEE Trans. Autom. Sci. Engin.*, DOI: 10.1109/TASE.2023.3324643, 2023.

[33] X. Tong, D. Ma, Z. Wang, Z. Ming and X. Xie, Model-free adaptive dynamic event-triggered robust control for unknown nonlinear systems using iterative neural dynamic programming, *Inf. Sci.*, vol.655, 119866, 2024.

[34] P. Liu, W. Ao, Z. Ming, G. Huang and Z. Liu, Dynamic event-triggered optimal tracking control for constrained nonlinear stochastic systems, *J. Frankl. Instit.*, vol.360, no.2, pp.1145-1165, 2023.

[35] W. Li, H. Zhang, W. Wang and Z. Cao, Fully distributed event-triggered time-varying formation control of multi-agent systems subject to mode-switching denial-of-service attacks, *Appl. Mathem. Comput.*, vol.414, 126645, 2022.

[36] H. Zhang, H. Su, K. Zhang and Y. Luo, Event-triggered adaptive dynamic programming for non-zero-sum games of unknown nonlinear systems via generalized fuzzy hyperbolic models, *IEEE Trans. Fuzzy Syst.*, vol.27, no.11, pp.2202-2214, 2019.

[37] L. Li, R. Yang, Z. Feng and L. Wu, Event-triggered dissipative control for 2-D switched systems, *Inf. Sci.*, vol.589, pp.802-812, 2022.

[38] X. Sun and X. Song, Dissipative analysis and event-triggered exponential synchronization for fractional-order complex-valued reaction-diffusion neural networks, *International Journal of Innovative Computing, Information and Control*, vol.18, no.5, pp.1519-1536, 2022.

[39] R. Yang and S. Ding, Two-dimensional event-triggered sliding mode control for Roesser model with time delays, *ISA Trans.*, vol.124, pp.271-279, 2022.

[40] S. Zhang, B. Zhao, D. Liu and Y. Zhang, Observer-based event-triggered control for zero-sum games of input constrained multi-player nonlinear systems, *Neural Networks*, vol.144, pp.101-112, 2021.

[41] P. Liu, H. Zhang, Z. Ming, S. Wang and R. K. Agarwal, Dynamic event-triggered safe control for nonlinear game systems with asymmetric input saturation, *IEEE Trans. Cybern.*, DOI: 10.1109/TCYB.2024.3354945, 2024.

[42] J. Y. Lee, J. B. Park and Y. H. Choi, Integral reinforcement learning for continuous-time input-affine nonlinear systems with simultaneous invariant explorations, *IEEE Trans. Neural Netw. Learn. Syst.*, vol.26, no.5, pp.916-932, 2015.

[43] Z. Ming, H. Zhang, Y. Yan and J. Sun, Self-triggered adaptive dynamic programming for model-free nonlinear systems via generalized fuzzy hyperbolic model, *IEEE Trans. Syst., Man, Cybern.: Syst.*, vol.53, no.5, pp.2792-2801, 2023.

[44] P. Liu, H. Zhang, H. Ren and C. Liu, Online event-triggered adaptive critic design for multi-player zero-sum games of partially unknown nonlinear systems with input constraints, *Neurocomputing*, vol.462, pp.309-319, 2021.

[45] X. Yang and Q. Wei, Adaptive critic learning for constrained optimal event-triggered control with discounted cost, *IEEE Trans. Neural Netw. Learn. Syst.*, vol.32, no.1, pp.91-104, 2021.

Author Biography

Ping Li received the B.S. degree and the M.S. degree in Mechatronic Engineering and the Ph.D. degree in Control Theory and Control Engineering from Harbin Engineering University, China, in 2000, 2003 and 2006, respectively. He is currently a Professor with Chongqing Technology and Business University, China. His current research interests include data-based control methods for the complex mechatronic systems, cooperative control for multi-robot systems, and intelligent control.

Huiyan Zhang received the M.Sc. degree in Control Engineering and the Ph.D. degree in Control Theory and Control Engineering from Harbin Institute of Technology, Harbin, in September 2014 and April 2019, respectively. From September 2015 to September 2017, she was a Joint Training Ph.D. Student with the School of Electrical and Electronic Engineering, the University of Adelaide. She is currently an Associate Professor with Chongqing Technology and Business University. Her research interests include stochastic switched systems, event-triggered scheme, model reduction, balanced truncation, robust control, and filtering design.

Wengang Ao received the B.S. degree in Engineering Mechanics from Sichuan University, Chengdu, in Jun. 2000, and M.E. degree in Solid Mechanics from Chongqing University, Chongqing, in Jun. 2007. He is currently a Professor at Chongqing Technology and Business University. His research interests include strength theory, kinematics and dynamics, robust control, and intelligent robots.

Pengda Liu received the B.S. degree in Automation Control and the Ph.D. degree in Control Theory and Control Engineering from Northeastern University, Shenyang, China, in 2012 and 2022, respectively. He is currently a Lecturer with Chongqing Technology and Business University, Chongqing, China. His current research interests include adaptive dynamic programming, reinforcement learning, neural-networks-based control, optimal control, and their applications.