KYBERNETIKA — VOLUME *59* (2023), NUMBER 3, PAGES 365–391

EVENT-TRIGGERED OPTIMAL CONTROL OF COMPLETELY UNKNOWN NONLINEAR SYSTEMS VIA IDENTIFIER-CRITIC LEARNING

Zhinan Peng, Zhiquan Zhang, Rui Luo, Yiqun Kuang, Jiangping Hu, Hong Cheng and Bijoy Kumar Ghosh

This paper proposes an online identifier-critic learning framework for event-triggered optimal control of completely unknown nonlinear systems. Unlike classical adaptive dynamic programming (ADP) methods with actor-critic neural networks (NNs), a filter-regression-based approach is developed to reconstruct the unknown system dynamics, and thus avoid the dependence on an accurate system model in the control design loop. Meanwhile, NN adaptive laws are designed for the parameter estimation by using only the measured system state and input data, and facilitate the identifier-critic NN design. The convergence of the adaptive laws is analyzed. Furthermore, in order to reduce state sampling frequency, two kinds of aperiodic sampling schemes, namely static and dynamic event triggers, are embedded into the proposed optimal control design. Finally, simulation results are presented to demonstrate the effectiveness of the proposed event-triggered optimal control strategy.

Keywords: optimal control, unknown nonlinear system, adaptive dynamic programming, identifier-critic neural networks, event-triggered mechanism

Classification: 93C10, 68T07

1. INTRODUCTION

With the rapid progress of control theory, numerous engineering applications have recognized that developing an admissible controller to ensure the system stability is a basic requirement, and that the resulting control cost should also be minimized. As a result, researchers have dedicated decades of effort to optimal control of uncertain nonlinear systems [12, 15].

From the principle of optimality, an optimal control problem can be generally solved with the help of a Hamilton-Jacobi-Bellman (HJB) equation, which is a partial differential equation (PDE). Werbos proposed an adaptive dynamic programming (ADP) method to compute approximate solutions to HJB equations [27], which combined neural networks (NNs) and dynamic programming into an actor-critic adaptive control architecture. Till now, the ADP framework has been developed to address a variety of control

DOI: 10.14736/kyb-2023-3-0365

issues, including stabilization control, tracking control, H_∞ control for single-agent systems [3, 9, 22] or multi-agent systems [19, 20].

It is important to note that the ADP framework often relies on the system dynamics information, but accurate system modeling is generally challenging in many real-world scenarios. For instance, autonomous underwater vehicles and exoskeleton systems have complex and uncertain dynamics and thus difficult to model accurately [2]. Alternatively, the structure of the system dynamics may be known in some application scenarios, but the model parameters may be unknown or uncertain, as in the case of hypersonic vehicles [16]. When faced with the lack of system information, identifier/observer-based reconstruction methods have been employed to facilitate the ADP framework. Particularly, NNs have been applied for system identification in recent years due to their exceptional performance in nonlinear function fitting. For example, for a partially unknown system (drift dynamic), S. Bhasin *et al.* combined the prior actor-critic ADP architecture with a NN identifier and proposed a triple-NNs structure, namely, actor-critic-identifier NNs in [1]. Under this framework, Lv et al. proposed NN weight tuning laws for online system identification based on an auxiliary filtering operation [14]. Luo et al. studied an adaptive optimal control of unknown nonlinear systems by using an identifier-critic learning framework with relaxed persistence of excitation [13]. Different from the NN identifier methods, Jiang et al. [7] developed a model-free control method based on policy iteration (PI) to solve the algebraic Riccati equation (ARE), which is another way that does not rely on complete system information. However, it is noted that, in the above mentioned literature, data sampling and control updating are both implemented in a periodic manner, which results in frequent data updates and heavy computational burden.

In recent years, researchers have been attempting to reduce the update frequency of controllers due to limited resources and computational bandwidth in practical control systems. To achieve this goal, event-triggered communication mechanisms have been proposed by designing an event generator that updates control inputs *aperiodically* [5, 6, 23, 26, 31]. Tabuada proposed a *static triggering mechanism* (STM) by designing a triggering rule based on the system state in [23], which promoted subsequent studies. For instance, Yang [31] proposed a STM-based event-sampled robust control method for unknown nonlinear systems by using adaptive critic learning. An event-triggered data-driven iterative learning control method was proposed for disturbed nonlinear discrete-time systems in [21]. To further reduce the triggering frequency, Girard developed a *dynamic triggering mechanism* (DTM) by incorporating an internal filter variable into the design of triggering rules [4]. It has been shown that the closed-loop system's stability can be guaranteed under STM and DTM, respectively. Recently, event-triggered mechanisms have been applied to the ADP framework and improve learning and control efficiency [18, 29]. For example, Wang et al. embedded STM into critic learning framework to solve optimal H_∞ control problems [25]. Xu et al. proposed an event-triggered ADP control method to solve a tracking control problem under STM in [28].

To the best of our knowledge, there are still some open issues in the existing studies: 1) Most of the literatures did not study optimal control with dynamic event-triggered schemes. Meanwhile, although some event-triggered mechanisms were adopted in the field of learning control of nonlinear systems [25], they need complete or partial knowledge of system dynamics; 2) The reference in [24] tried to relax the dependence of partial system dynamics $g(x)$ by utilizing an NN-based identification; however, drift dynamics were still required and the identification process was only implemented offline, which is not practical in many real-time cases. As a result, we try to combine system identification methods, online learning frameworks, and event-triggered mechanisms to cope with these issues, which becomes the motivation of this paper.

Based on the previous discussions, this paper presents a novel online identifier-critic learning framework for event-triggered optimal control of completely unknown nonlinear systems. The main contributions of the paper are summarized as follows:

1. Unlike existing ADP methods that used classical actor-critic NNs, this paper proposes a simplified online identifier-critic learning framework to address event-triggered optimal control of completely unknown nonlinear systems. The proposed method uses a filter regressor based parameter estimation method to reconstruct the unknown system dynamics, thereby reducing the dependence on accurate system models in control designs.

2. In contrast to previous methods that use gradient-descent based NN adaptive laws [24, 25], which only converge to the bounded neighborhood of their actual values, this paper proposes novel NN weight adaptive laws for the identifier NN and the critic NN based on parameter estimation. The convergence performance of the estimated weight is ensured, and the learning process of the identifier and critic NNs is simultaneous and online, using only the measured system state and input data.

3. To reduce the state sampling frequency, two aperiodic sampling schemes, namely static and dynamic event- triggers, are embedded into the proposed learning control framework. Additionally, the stability of the closed-loop system is proven with the help of a Lyapunov function method.

The rest of this paper is organized as follows. Section II describes the system model and optimal control problem. Section III proposes the identifier-critic learning framework and the NNs' adaptive laws. Section IV introduces two event-triggered mechanisms and the main theoretical results. Section V demonstrates the effectiveness of the proposed learning control methods via numerical simulations. Finally, Section VI concludes this paper.

2. PRELIMINARIES AND PROBLEM FORMULATION

In this paper, we consider a class of nonlinear continuous-time systems

$$\dot{x} = f(x) + g(x)\,u, \tag{1}$$

where $x \in \mathbb{R}^n$ and $u \in \mathbb{R}^m$ denote the system state and control input, respectively. $f(x) \in \mathbb{R}^n$ is the drift dynamics and $g(x) \in \mathbb{R}^{n\times m}$ is the input dynamics. It is assumed that $f(x)$ and $g(x)$ are unknown. $f(x) + g(x)u$ is Lipschitz continuous, and can be stabilized on a compact set $\Omega \subseteq \mathbb{R}^n$.

In order to analyze the stability of the closed-loop system of 1, we need the following notation:

Definition 1. (Uniform Ultimate Boundedness, Lewis et al. [11]) The solution of the system (1) is said to be uniform ultimate boundedness (UUB) if there exists a compact set $C \subset \mathbb{R}^n$ containing the initial value $x(t_0) = x_0$, a constant $B > 0$ and a time $T = T(B, x_0)$ such that $||x(t) - x_0|| < B$ for all $t \geq t_0 + T$.

In order to evaluate the system performance, an infinite-horizon cost function $J(x(t))$ is defined as follows:

$$J(x(t)) = \int_t^\infty r(x(\tau), u(\tau))\,\mathrm{d}\tau, \tag{2}$$

where $r(x(t)) = x(t)^\top Qx(t) + u(t)^\top Ru(t)$ denotes the utility function, $Q \in \mathbb{R}^{n\times n}$ and $R \in \mathbb{R}^{m\times m}$ are symmetric positive definite matrices, respectively.

The objective of this paper is to design a controller to stabilize the system (1) and, simultaneously, minimize the cost function $J(x(t))$. According to the optimal principle, the Bellman equation is obtained by taking time derivative of (2):

$$r\left(x, u\right) + \nabla J^\top\left(x\right)\left(f\left(x\right) + g\left(x\right)u\right) = 0, \tag{3}$$

where $\nabla J\left(x\right) = \partial J\left(x\right)/\partial x$. Based on (3), we define the Hamilton function as

$$H\left(x, u, \nabla J\right) = r\left(x, u\right) + \nabla J^\top\left(f\left(x\right) + g\left(x\right)u\right). \tag{4}$$

The optimal cost function J^* satisfies the following Hamilton-Jacobi-Bellman equation:

$$\min_u\left[H\left(x, u, \nabla J^*\right)\right] = 0. \tag{5}$$

Then, the optimal controller can be obtained as follows:

$$0 = \frac{\partial H\left(x, u, d, \nabla J^*\left(x\right)\right)}{\partial u}, \tag{6}$$

$$u^*\left(x\right) = -\frac{1}{2}R^{-1}g^\top\left(x\right)\left(\nabla J^*\left(x\right)\right). \tag{7}$$

By applying (7) to (5), the HJB equation can be expressed as

$$0 = x^\top Qx + \nabla J^{*\top} f\left(x\right) - \frac{1}{4}\nabla J^{*\top} g\left(x\right)R^{-1}g^\top\left(x\right)\nabla J^*. \tag{8}$$

As we know, it is difficult to compute the solution to the HJB equation because of the existence of the partial differential part in (8). In addition, since the information of $f(x)$ and $g(x)$ is unknown, thus the optimal controller u^* still cannot be applied directly to stabilize the nonlinear system.

In order to address these difficulties, in Section 3, a filter-regressor-based system identification will be proposed for reconstructing the unknown dynamics ($f(x)$ and $g(x)$), and then a critic NN is established to solve the approximate solution to the HJB equation, where an online NN weight adaption laws are proposed based on filtering operations.

3. ONLINE IDENTIFIER-CRITIC NNS LEARNING FRAMEWORK DESIGNS

This section presents an online NN-based identifier-critic learning control structure. It contains two NNs, i. e., identifier and critic neural networks, and is used to estimate the unknown dynamics and approximate the cost function.

3.1. Filter-regressor based system identification

In this subsection, we will describe the design process of the NN-based identifier with a filter-regressor method. Two NN-based identifiers are designed as follows:

$$f(x) = \alpha\delta(x) + e_f, \tag{9}$$

$$g(x) = \beta\gamma(x) + e_g, \tag{10}$$

where $\alpha \in \mathbb{R}^{n\times l_\alpha}$ and $\beta \in \mathbb{R}^{n\times l_\beta}$ are the ideal and unknown bounded weights of drift and input dynamics, respectively. $\delta(x) : \mathbb{R}^n \to \mathbb{R}^{l_\alpha}$ and $\gamma(x) : \mathbb{R}^n \to \mathbb{R}^{l_\beta\times m}$ are the activation functions. l_α and l_β respectively represent the dimensions of the NNs' hidden layers. $e_f \in \mathbb{R}^n$ and $e_g \in \mathbb{R}^{n\times m}$ are reconstruction errors, respectively.

According to (9) and (10), the system dynamics can be reformulated by a compact form:

$$\dot{x} = W_i^\top \theta_i(x, u) + \varepsilon_i, \tag{11}$$

where $W_i^\top = [\alpha, \beta]^\top \in \mathbb{R}^{(l_\alpha+l_\beta)\times n}$ denotes the unknown identifier weight. $\theta_i = [\delta^\top, u^\top\gamma^\top]$ is the *regressor vector*, and $\varepsilon_i = e_f + e_g u$ is the integrated reconstruction error. Before continuing the discussion, the following assumption is made.

Assumption 1. The signal θ_i is assumed to be persistently excited (PE) if there exist two positive constants λ_1 and λ_2 such that the following condition holds during the time interval $[t, t+T]$,

$$\lambda_1 I \le \int_t^{t+T} \theta_i(\tau)\theta_i(\tau)^\top \,\mathrm{d}\tau \le \lambda_2 I. \tag{12}$$

To estimate the unknown weight of the NN identifier, inspired by [10], a filtering operation is employed to transfer the system identification to a linear regression problem. First, x and θ_i are filtered by a low-pass filter operator $1/(l_i s+1)$, that is,

$$l_i\dot{x}_f + x_f = x, \tag{13}$$

$$l_i\dot{\theta}_{if} + \theta_{if} = \theta_i, \tag{14}$$

where l_i is the filter constant. x_f and θ_{if} are the filtered variables of x and θ_i. The reconstruction error ε_i is filtered by $l_i\dot{\varepsilon}_{if} + \varepsilon_{if} = \varepsilon_i$.

By the filtering operation, according to (13) and (14), the NN-based system (11) can be constructed by the following *linear regression* form, i. e.,

$$\dot{x}_f = \frac{x - x_f}{l_i} = W_i^\top \theta_{if} + \varepsilon_{if}. \tag{15}$$

To facilitate matrix processing, we multiply the transpose of the last two terms of equation (15) by θ_{if}. This yields:

$$\theta_{if}\left[\frac{x - x_f}{l_i}\right]^\top = \theta_{if}\theta_{if}^\top W_i + \theta_{if}\varepsilon_{if}^\top. \tag{16}$$

Similar to the first filtering process, $\theta_{if}\left[\frac{x-x_f}{l_i}\right]^\top$ and $\theta_{if}\theta_{if}{}^\top$ are filtered by a low-pass filter $1/(s+\xi)$, respectively, as

$$\dot{S}_i = -\xi S_i + \theta_{if}\theta_{if}^\top, \tag{17}$$

$$\dot{T}_i = -\xi T_i + \theta_{if}\left[\frac{x-x_f}{l_i}\right]^\top, \tag{18}$$

where $S_i \in \mathbb{R}^{(l_\alpha+l_\beta)\times(l_\alpha+l_\beta)}$ and $T_i \in \mathbb{R}^{(l_\alpha+l_\beta)\times n}$ denote two intermediate filtered regressor matrices of $\theta_{if}\left[\frac{x-x_f}{l_i}\right]^\top$ and $\theta_{if}\theta_{if}{}^\top$, respectively.

Assuming zero initial conditions, that is, $S_i(0) = 0$ and $T_i(0) = 0$, we can obtain analytical solutions to the differential equations (17) and (18) in the following form:

$$S_i(t) = \int_0^t e^{-\xi(t-\tau)}\theta_{if}(\tau)\,\theta_{if}(\tau)^\top\,\mathrm{d}\tau, \tag{19}$$

$$T_i(t) = \int_0^t e^{-\xi(t-\tau)}\theta_{if}(\tau)\left[\frac{x(\tau)-x_{if}(\tau)}{l_i}\right]^\top\,\mathrm{d}\tau. \tag{20}$$

Based on S_i and T_i, we define an another intermediate variable $P_i \in \mathbb{R}^{(l_\alpha+l_\beta)\times n}$ as

$$P_i = S_i\hat{W}_i - T_i. \tag{21}$$

Then, the weight tuning law for identifier NN is designed as

$$\dot{\hat{W}}_i = -\Pi_i P_i, \tag{22}$$

where $\hat{W} = \left[\hat{\alpha},\hat{\beta}\right]^\top$ denotes the estimation of W. Π_i denotes the learning gain matrix.

Remark 1. By using (15), (19), and (20), one can obtain $T_i = S_iW_i - e_i$, where $e_i = \int_0^t e^{-\xi(t-\tau)}\theta if(\tau)\,\varepsilon if(\tau)^\top\,\mathrm{d}\tau$. It should be noted that if the number of neuron nodes is large enough, the reconstruction error ε_i will converge to zero, which will result in $e_i \to 0$.

Remark 2. According to the core filtering steps (17), (18), and their solutions $S_i(t) = \int_0^t e^{-\xi(t-\tau)}\theta_{if}(\tau)\,\theta_{if}(\tau)^\top\,\mathrm{d}\tau$ and $T_i(t) = \int_0^t e^{-\xi(t-\tau)}\theta_{if}(\tau)\left[\frac{x(\tau)-x_{if}(\tau)}{l_i}\right]^\top\,\mathrm{d}\tau$, these filtering steps allow for the accumulation of previous and current data, weighted by an exponential term in the time domain. By applying this filtering process, we can improve the excitation and adequacy of the original regressor vector θ_i in the NN-based approximate system (11), by using only measured data.

3.2. Critic learning based optimal control

Based on the identification given in the previous subsection, the reconstructed system dynamics of (1) can be re-expressed as

$$\dot{x} = \hat{\alpha}\delta(x) + \hat{\beta}\gamma(x) + e_{Ni} + \varepsilon_i, \tag{23}$$

where $\hat{\alpha}$ and $\hat{\beta}$ denote the online estimated weights of identifier NNs, respectively. e_{Ni} is an estimation error introduced by the NN, i. e., $e_{Ni} = (W_i - \hat{W}_i)^\top \theta_i(x)$. Since the accurate dynamics $f(x)$ and $g(x)$ are unknown, the estimated system model (23) is employed in the critic learning design below.

According to (4), the approximate Hamilton function can be written as

$$H\left(x, u, \nabla J\right) = r\left(x, u\right) + \nabla J^\top (\hat{\alpha}\delta\left(x\right) + \hat{\beta}\gamma\left(x\right) + e_{Ni} + \varepsilon_i). \tag{24}$$

Therefore, with the approximate Hamiltonian function and HJB equation, the optimal control policy u^* can be re-written as:

$$u^* = -\frac{1}{2}R^{-1}\left[\hat{\beta}\gamma\left(x\right)\right]^\top \nabla J^*\left(x\right). \tag{25}$$

As is well known, the HJB equation and optimal control (25) are difficult to be computed. Thus, we introduce an approximation method to estimate the optimal cost function by using a three-layer critic NNs, which is represented as

$$J^*\left(x\right) = {W_c}^\top \theta_c\left(x\right) + \varepsilon_c, \tag{26}$$

and its derivative is given by

$$\nabla J^*\left(x\right) = \nabla {\theta_c}^\top\left(x\right) W_c + \nabla\varepsilon_c, \tag{27}$$

where $W_c \in \mathbb{R}^k$ and $\theta_c \in \mathbb{R}^k$ denote the ideal weight and activation function of the critic NNs, respectively. k denotes the number of neurons of the critic NNs. $\nabla\theta_c$ and $\nabla\varepsilon_c$ are the partial derivatives of the activation function $\theta_c(x)$ and approximation error ε_c, respectively. Since the ideal weights are unknown in the learning process, the true optimal cost function $\hat{J}(x)$ is approximated as

$$\hat{J}(x) = \hat{W}_c^\top \theta_c(x), \tag{28}$$

where $\hat{W}_c$ is the estimated weights of critic NN. Thus, the estimated control policy $\hat{u}$ can be expressed as

$$\hat{u} = -\frac{1}{2}R^{-1}\left[\hat{\beta}\gamma\left(x\right)\right]^\top \nabla{\theta_c}^\top \hat{W}_c. \tag{29}$$

Based on critic and identifier NNs, the estimated HJB equation can be represented as

$${W_c}^\top \nabla\theta_c\left(\hat{\alpha}\delta + \hat{\beta}\gamma u\right) + r\left(x, u\right) + \varepsilon_T = 0, \tag{30}$$

where $\varepsilon_T = {W_c}^\top \nabla\theta_c\left(e_{Ni} + \varepsilon_i\right) + \nabla\varepsilon_c\left(\hat{\alpha}\delta + \hat{\beta}\gamma u + e_{Ni} + \varepsilon_i\right)$ is the reconstruction error. According to Weierstrass approximation theorem, if l_α, l_β and k tend to infinity, ε_T will be vanished.

Then, the approximated HJB equation can be formulated as a *linear regression equation*, that is,

$$\Psi = -{W_c}^\top \Gamma - \varepsilon_T, \tag{31}$$

where $\Gamma = \nabla\theta_c(\hat{\alpha}\delta + \hat{\beta}\gamma u)$ and $\Psi = r\left(x, u\right)$.

Similar to the filtering procedure proposed in the identifier NN designs, two intermediate filtered regressor matrix $S_c \in \mathbb{R}^{(k\times k)}$ and $T_c \in \mathbb{R}^k$ are introduced for filtering $\Gamma\Gamma^\top$ and $\Gamma\Psi$ in (31), respectively. By using a low-pass filter operator $1/(s+l_c)$ with zero initial conditions, we have

$$\dot{S}_c = -l_c S_c + \Gamma\Gamma^\top, \tag{32}$$

$$\dot{T}_c = -l_c S_c + \Gamma\Psi, \tag{33}$$

where l_c is a positive parameter. Then, we define an another vector $P_c \in \mathbb{R}^k$ based on S_c and T_c as

$$P_c = S_c\hat{W}_c + T_c. \tag{34}$$

Then, the weight tuning law of critic NN is designed as

$$\dot{\hat{W}}_c = -\Pi_c P_c, \tag{35}$$

where Π_c is a learning rate matrix of the critic NN.

Remark 3. This paper proposes a new weight adaptation method based on a filter-regressor approach, which diverges from existing system identification techniques [30, 31] that aim to minimize the identifier error between the system state x and the estimation state $\hat{x}$. Although these methods show that the identifier state $\hat{x}$ converges to its true value x, they do not prove the convergence of the identifier weights to ideal model parameters. In contrast, the proposed method is a direct weight adaptation method based on the parameter estimation error. The advantage of this method is that the weights can be 'directly' estimated rather than updated to minimize the identifier error, which guarantees the convergence of the NN weights to ideal model parameters.

Till now, the derivations of the identifier and critic NNs with weight estimation laws have been completed. In the above controller designs, the system state must be periodically sampled to compute the controller, which results in high energy consumption and communication bandwidth. In the following, to improve learning and control efficiency, an event-triggered sampling mechanism is introduced into the proposed identifier-critic learning control designs. The block diagram of the proposed event-triggered optimal control method is shown in Figure 1.

4. EVENT-BASED IDENTIFIER-CRITIC LEARNING CONTROL

In this section, two triggering mechanisms are incorporated into the proposed learning-based optimal control. By using the event-based scheme, the controller is updated only when the triggering event occurs at τ_j, and remains unchanged in the time interval $[\tau_j, \tau_{j+1})$. The system state vector at the triggering instant is denoted by

$$x_j = x(\tau_j) = x(t)|_{t=\tau_j}. \tag{36}$$

Then, we define an error between the actual system state $x(t)$ and the system state at triggering instant τ_j x_j as $e_j(t) = x_j - x(t)$ in the time interval $[\tau_j, \tau_{j+1})$. The event-based control policy $\hat{u}(x_j)$ at triggering time τ_j is given by

$$\hat{u}(x_j) = -\frac{1}{2}R^{-1}\left[\hat{\beta}\gamma\left(x_j\right)\right]^\top \nabla{\theta_c}^\top(x_j)\hat{W}_c, t \in [\tau_j, \tau_{j+1}). \tag{37}$$

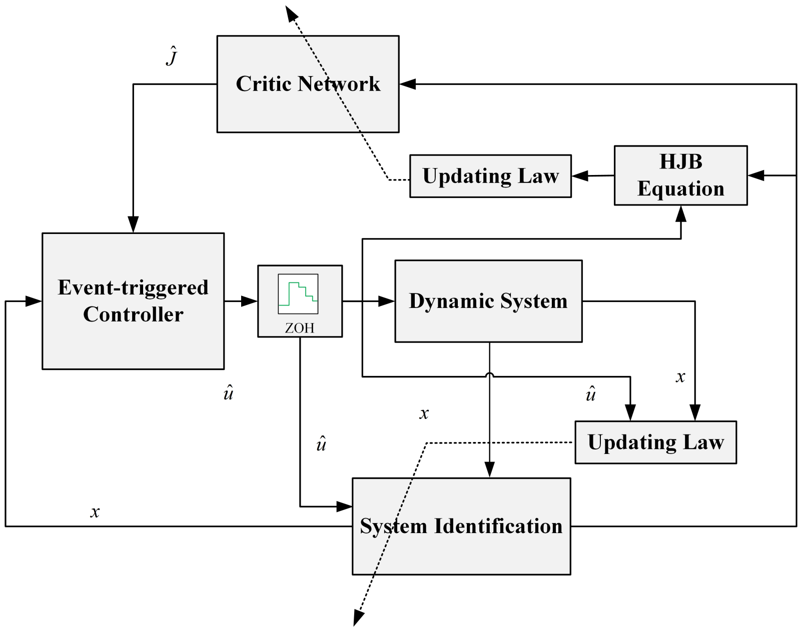


Fig. 1. Block diagram of the proposed identifier-critic-based event-triggered optimal control methods.

Then, the weight tuning law of the critic NN is designed as

$$\dot{\hat{W}}_c = -\Pi_c\big(S_c(\hat{u}(x_j))\hat{W}_c + T_c(\hat{u}(x_j))\big), \tag{38}$$

where $S_c(\hat{u}(x_j))$ and $T_c(\hat{u}(x_j))$ are computed by using (32) and (33) with $\Gamma(\hat{u}(x_j)) = \nabla\theta_c(\hat{\alpha}\delta + \hat{\beta}\gamma\hat{u}(x_j))$ and $\Psi(\hat{u}(x_j)) = r\,(x, \hat{u}(x_j))$.

Remark 4. In critic NNs, most available weight adaptation laws are designed to minimize the Hamiltonian errors, i. e., $\hat{H} - H$, by utilizing gradient descent methods [29, 31]. In contrast to these approaches, we propose a filter-regressor-based weight adaptation method based on the parameter estimation error. By introducing some new filtered variables, the problem of system identification can be transformed into a problem of parameter estimation, such that the convergence of NN weights can be guaranteed.

Before presenting the triggering condition designs, we need to make the following assumption, as proposed in [25].

Assumption 2. The control policy $u(t)$ satisfies the Lipschitz condition, that is, there exists a constant $\mathcal{L}$ such that, for all $x, x_j \in \Omega$, $\|u(x) - \hat{u}\,(x_j)\| \le \mathcal{L}\,\|e_j\,(t)\|$.

4.1. Static and dynamic triggers mechanism design

We first present the design of the static trigger mechanism (STM), the triggering time τ_{j+1} is determined by

$$\begin{aligned}\tau_{j+1} = \arg\min_{t} \Big\{ t \in \mathbb{R}_0^+ | t > \tau_j \cap \Big[\rho x^T Q x \\ + \|\mu\|^2 \|\hat{u}(x_j)\|^2 - \mathcal{L}^2 \|\mu\|^2 \|e_j(t)\|^2 \leq 0 \Big] \Big\}.\end{aligned} \tag{39}$$

Based on (39), the static trigger condition is designed as

$$\|e_j(t)\|^2 \leq \frac{\rho x^\top Q x + \|\mu\|^2 \|\hat{u}(x_j)\|^2}{\mathcal{L}^2 \|\mu\|^2} = E_S, \tag{40}$$

where μ is decomposed from R, i. e., $R = \mu\mu^\top$, and $\rho \in (0, 1)$ is a positive constant. E_S is the trigger threshold of STM.

Inspired by [4], a dynamic trigger mechanism (DTM) is proposed based on the STM. First, we introduce a filtered variable $y_d(t)$, which has the following equation

$$\dot{y}_d(t) = -\phi y_d + \Big\{ \rho x^\top Q x + \|\mu\|^2 \|\hat{u}(x_j)\|^2 - \mathcal{L}^2 \|\mu\|^2 \|e_j\|^2 \Big\}, \tag{41}$$

where $\phi > 0$ is a positive filtering constant. The initial value of the filtered variable $y_d(0) \geq 0$. The triggering time τ_{j+1} is determined by

$$\begin{aligned}\tau_{j+1} = \arg\min_{t} \Big\{ t \in \mathbb{R}_0^+ | t > \tau_j \cap \Big[y_d(t) + \rho_d \Big(\rho x^T Q x \\ + \|\mu\|^2 \|\hat{u}(x_j)\|^2 - \mathcal{L}^2 \|\mu\|^2 \|e_j(t)\|^2 \Big) \leq 0 \Big] \Big\}.\end{aligned} \tag{42}$$

Hence, the dynamic trigger condition is designed as

$$\|e_j(t)\|^2 \leq \frac{\rho_d \left(\rho x^\top Q x + \|\mu\|^2 \|\hat{u}(x_j)\|^2 \right) + y_d}{\rho_d \mathcal{L}^2 \|\mu\|^2} = E_D, \tag{43}$$

where ρ_d is a positive constant and E_D is the trigger threshold of DTM.

Remark 5. It is worth noting that when y_d equals zero, the dynamic triggering mechanism becomes equivalent to the static triggering mechanism. Furthermore, it can be observed that the dynamic triggering condition has a higher threshold compared to that of the static scheme, which means that the dynamic triggering condition is relatively easier to satisfy, which results in fewer trigger events.

The convergence of the weight adaptation laws and the stability of the closed-loop system under static and dynamic triggers will be proved in the following subsection.

4.2. Stability and convergency analysis

Before giving the main theoretical results, the following assumptions are first made.

Assumption 3. The unknown dynamics $f(x)$ and $g(x)$ are bounded, such that $\exists\, t_f > 0,\ t_g > 0,\ \|f(x)\| \le t_f \|x\|, \|g(x)\| \le t_g$.

Assumption 4. The critic NN weight $\|W_c\|$, the derivative of critic NN activate function $\|\nabla\theta_c\|$, the derivative of approximate error $\|\nabla\varepsilon_c\|$ are all bounded, that is, $\|W_c\| \le \|W_{cm}\|$, $\|\nabla\theta_c\| \le \|\theta_{cm}\|$, $\|\nabla\varepsilon_c\| \le \|\theta_{cdm}\|$. $t_{\beta\gamma} = \|\hat{\beta}\gamma\|$, $t_{W_c} = \|\hat{W}_c\|$ are also bounded variables.

Theorem 1. Let Assumptions 1-4 hold. Consider system (1) with weight tuning laws of identifier NN and critic NN updated by (22) and (35), respectively. Then, if the event is generated by the static triggering scheme (40), we can conclude: 1) Without regard for reconstruction errors, the system is asymptotically stable and the weight errors $\tilde{W}_i = W_i - \hat{W}_i$, $\tilde{W}_c = W_c - \hat{W}_c$ all converge to zero; 2) If reconstruction errors exist, the system stability and the convergence of the weight estimation errors $\tilde{W}_i$ and $\tilde{W}_c$ are all ensured in the sense of uniformly ultimately bounded (UUB).

Proof. The Lyapunov function candidate L_1 is constructed as

$$\begin{aligned} L_1 =& L_{11} + L_{12} + L_{13} + L_{14} + L_{15} + L_{16} \\ =& \frac{1}{2} tr\left(\tilde{W}_i^\top \Pi_i^{-1} \tilde{W}_i\right) + \frac{1}{2}\tilde{W}_c^\top \Pi_c^{-1} \tilde{W}_c \\ &+ J^*(x) + J^*(x_k) + \sigma_i {e_i}^\top e_i + \sigma_c {e_c}^\top e_c. \end{aligned} \tag{44}$$

The whole proof process is divided into two cases: 1) the event is not triggered and 2) the event is triggered.

Case 1. In event holding interval, we have $\dot{L}_{14} = 0$. The derivative of Lyapunov function L_1 is given by

$$\dot{L}_1 = \dot{L}_{11} + \dot{L}_{12} + \dot{L}_{13} + \dot{L}_{15} + \dot{L}_{16}. \tag{45}$$

For $\dot{L}_{11}$, a equivalent deformation related to variables S_i, T_i and W_i are made based on equations (15), (19) and (20):

$$T_i = S_i W_i + \int_0^t e^{-\xi(t-\tau)} \theta_{if}(\tau)\, \varepsilon_{if}^\top \, \mathrm{d}\tau. \tag{46}$$

We represent the integral term as $e_i = -\int_0^t e^{-\xi(t-\tau)} \theta_{if}(\tau)\, \varepsilon_{if}^\top \, \mathrm{d}\tau$. The variable P_i can thus be re-expressed as

$$P_i = S_i \hat{W}_i - T_i = S_i \hat{W}_i - (S_i W_i - e_i) = -S_i \tilde{W}_i + e_i. \tag{47}$$

Hence, $\dot{L}_{11}$ can be decomposed as

$$\dot{L}_{11} = tr(\tilde{W}_i^\top \Pi_i^{-1} \dot{\tilde{W}}_i) = -tr(\tilde{W}_i^\top S_i \tilde{W}_i) + tr(\tilde{W}_i^\top e_i). \tag{48}$$

According to Assumption 1, the intermediate variable S_i satisfies $\lambda_{\min}(S_i) > \omega_i > 0$, where $\lambda_{\min}(S_i)$ denotes the minimal eigenvalue of matrix S_i. Then, based on the the Young's inequality: $\exists c > 0$, $ab \leq a^2c/2 + b^2/2c$, the following inequity can be further derived as

$$\begin{aligned} \dot{L}_{11} &\leq -\omega_i \left\| \tilde{W}_i \right\|^2 + \left\| \tilde{W}_i^{\top} e_i \right\|^2 \\ &\leq -\left(\omega_i + \frac{1}{2c} \right) \left\| \tilde{W}_i \right\|^2 + \frac{c \| e_i \|^2}{2}. \end{aligned} \tag{49}$$

The derivative of L_{12} can be deduced by a similar procedure. From (32), (33) and (34), a linear equation related to S_c, T_c and W_c can be expressed as

$$\begin{aligned} T_c &= -S_c W_c - \int_0^t e^{-l_c(t-\tau)} \varepsilon_T(\tau) \Gamma(\tau) \, \mathrm{d}\tau \\ &= -S_c W_c + e_c, \end{aligned} \tag{50}$$

where e_c denotes the integral term $e_c = -\int_0^t e^{-l_c(t-\tau)} \Gamma(\tau) \varepsilon_T(\tau) \, \mathrm{d}\tau$. Further, plugging (50) into (34) yields

$$P_c = -S_c \tilde{W}_c + e_c. \tag{51}$$

Therefore, if S_c satisfies PE condition in Assumption 1, we have $\lambda_{\min}(S_c) \geq \omega_c \geq 0$, $\dot{L}_{12}$ can be further derived as

$$\begin{aligned} \dot{L}_{12} &= \tilde{W}_c^{\top} \Pi_c^{-1} \dot{\tilde{W}}_c \\ &= -\tilde{W}_c^{\top} S_c \tilde{W}_c + \tilde{W}_c^{\top} e_c \\ &\leq -\omega_c \left\| \tilde{W}_c \right\|^2 + \left\| \tilde{W}_c^{\top} e_c \right\|^2 \\ &\leq -\left(\omega_c + \frac{1}{2c} \right) \left\| \tilde{W}_c \right\|^2 + \frac{c \| e_c \|^2}{2}. \end{aligned} \tag{52}$$

By using (7) and (8), the following equations can be obtained

$$\nabla J^{*\top} g(x) = -2u^{*\top}(x) R, \tag{53}$$

$$\nabla J^{*\top} f(x) = \frac{1}{4} \nabla J^{*\top} g(x) R^{-1} g^{\top}(x) \nabla J^* - x^{\top} Q x. \tag{54}$$

For $\dot{L}_{13}$, by applying the chain rule, we can get

$$\dot{J}^*(x) = \frac{\partial J^*(x)}{\partial x} \frac{\partial x}{\partial t} = \nabla J^{*\top} \left(f(x) + g(x) \hat{u}(x_j) \right). \tag{55}$$

By applying (53) and (54) to (55), we can make the following deformation

$$\begin{aligned}\dot{J}^*(x) =&\frac{1}{4}\nabla J^{*\top}(x)\, g^\top(x)\, R^{-1} g(x)\, \nabla J^*(x)\\&- x^\top Qx - 2u^{*\top}(x)\, R\hat{u}(x_j)\\=&u^{*\top}(x)\, Ru^*(x) - x^\top Qx - 2u^{*\top}(x)\, R\hat{u}(x_j)\\=&\left[u^*(x) - \hat{u}(x_j) + \hat{u}(x_j)\right] R\left[u^*(x) - \hat{u}(x_j) - \hat{u}(x_j)\right]\\&- x^\top Qx\\=&\left\|\mu^\top\left(u^*(x) - \hat{u}(x_j)\right)\right\|^2 - \left\|\mu^\top \hat{u}(x_j)\right\|^2\\&- \rho x^\top Qx - (1-\rho)\, x^\top Qx.\end{aligned} \tag{56}$$

According to $\int_0^t e^{-\xi(t-\tau)}\theta_{if}(\tau)\,\varepsilon_{if}^\top\,\mathrm{d}\tau$ in Remark 1, we have

$$\dot{e}_i = -\xi e_i + \theta_{if}\varepsilon_{if}^\top. \tag{57}$$

Using Young's inequity, $\dot{L}_{15}$ can be derived as

$$\begin{aligned}\dot{L}_{15} &= 2\sigma_i {e_i}^\top(-\xi e_i + \theta_{if}\varepsilon_{if}^\top)\\&\le \frac{1}{\mathrm{c}}\left\|\sigma_i\theta_{if}\varepsilon_{if}\right\|^2 + (-2\sigma_i\xi + c)\left\|e_i\right\|^2.\end{aligned} \tag{58}$$

By using the fact $\dot{e}_c = -l_c e_c + \Gamma\varepsilon_T$ and Young's inequity, $\dot{L}_{16}$ is derived as

$$\begin{aligned}\dot{L}_{16} =&2\sigma e_c^\top \dot{e}_c\\=&2\sigma_c e_c^\top\big(-l_c\|e_c\| + \Gamma\left[{W_c}^\top\nabla\theta_c(e_{Ni}+\varepsilon_i) + \nabla\varepsilon_c(f+gu)\right]\big)\\=&\frac{1}{c}\sigma_c^2 W_{cm}^2\theta_{cm}^2\|\Gamma\|^2\|e_{Ni}\|^2 + \frac{1}{c}\sigma_c^2 W_{cm}^2\theta_{cm}^2\|\Gamma\|^2\|\varepsilon_i\|^2\\&+\frac{1}{c}\sigma_c^2 t_f^2\theta_{cdm}^2\|\Gamma\|^2\|x\|^2\\&+\frac{1}{4c}\lambda_{\max}^2\left(R^{-1}\right)\sigma_c^2 t_g^2 t_{Wc}^2 t_{\beta\gamma}^2\theta_{cm}^2\|\Gamma\|^2\|\nabla\varepsilon_c\|^2\\&+(4c - 2\sigma_c l_c)\left\|e_c\right\|^2.\end{aligned} \tag{59}$$

Based on the aforementioned processing (49)-(59), we have

$$\begin{aligned}\dot{L}_1 =&\dot{L}_{11} + \dot{L}_{12} + \dot{L}_{13} + \dot{L}_{15} + \dot{L}_{16}\\ \leq&\left(-\left(1-\rho\right)\lambda_{\min}\left(Q\right) + \frac{1}{c}\sigma_c^2 t_f^2 \theta_{cdm}^2 \|\Gamma\|^2\right)\|x\|^2\\ &-\left(\omega_i - \frac{1}{2c} - \frac{1}{c}\sigma_c^2 W_{cm}^2 \theta_{cm}^2 \|\Gamma\|^2 \|\theta_i\|^2\right)\|\tilde{W}_i\|^2\\ &-\left(\omega_c - \frac{1}{2c}\right)\|\tilde{W}_c\|^2\\ &+\left(-2\sigma_i\xi + \frac{3c}{2}\right)\|e_i\|^2 + \left(\frac{9c}{2} - 2\sigma_c l_c\right)\|e_c\|^2\\ &+\frac{1}{\mathrm{c}}\|\sigma_i\theta_{if}\varepsilon_{if}\|^2 + \frac{1}{c}\sigma_c^2 W_{cm}^2\theta_{cm}^2\|\Gamma\|^2\|e_{Ni}\|^2\\ &+\frac{1}{c}\sigma_c^2 W_{cm}^2\theta_{cm}^2\|\Gamma\|^2\|\varepsilon_i\|^2\\ &+\frac{1}{4c}\lambda_{\max}^2\left(R^{-1}\right)\sigma_c^2 t_g^2 t_{Wc}^2 t_{\beta\gamma}^2 \theta_{cm}^2\|\Gamma\|^2\|\nabla\varepsilon_c\|^2\\ &+\|\mu^\top\left(u^*\left(x\right) - \hat{u}\left(x_j\right)\right)\|^2 - \|\mu^\top\hat{u}\left(x_j\right)\|^2 - \rho x^\top Q x.\end{aligned} \tag{60}$$

To facilitate analysis, we define the following notations as

$$\begin{aligned}p_1 =&-\left(1-\rho\right)\lambda_{\min}\left(Q\right) + \frac{1}{c}\sigma_c^2 t_f^2\theta_{cdm}^2\|\Gamma\|^2,\\ p_2 =&-\left(\omega_i + \frac{1}{2c} - \frac{1}{c}\sigma_c^2 W_{cm}^2\theta_{cm}^2\|\Gamma\|^2\|\theta_i\|^2\right),\\ p_3 =&-(\omega_c + \frac{1}{2c}), p_4 = -2\sigma_i\xi + \frac{3c}{2},\\ p_5 =&\frac{9c}{2} - 2\sigma_c l_c,\\ E =&\frac{1}{\mathrm{c}}\|\sigma_i\theta_{if}\varepsilon_{if}\|^2 + \frac{1}{c}\sigma_c^2 W_{cm}^2\theta_{cm}^2\|\Gamma\|^2\|\varepsilon_i\|^2\\ &+\frac{1}{4c}\lambda_{\max}^2\left(R^{-1}\right)\sigma_c^2 t_g^2 t_{Wc}^2 t_{\beta\gamma}^2\theta_{cm}^2\|\Gamma\|^2\|\nabla\varepsilon_c\|^2.\end{aligned} \tag{61}$$

If parameters σ_i, σ_c and c, respectively, satisfy the following inequities:

$$\begin{aligned}c >&\max\Big\{(\sigma_{\mathrm{c}}^2 \mathrm{t}_{\mathrm{f}}^2\theta_{\mathrm{cdm}}^2\|\Gamma\|^2/\left(1-\rho\right)\lambda_{\min}\left(\mathrm{Q}\right)), (1/2\omega_{\mathrm{c}}),\\ &((\frac{1}{2} + \sigma_c^2 W_{cm}^2\theta_{cm}^2\|\Gamma\|^2\|\theta_i\|^2)\Big/\omega_i)\Big\},\\ \sigma_i >&\frac{3c}{4\xi}, \sigma_c > \frac{9c}{4l_c}.\end{aligned} \tag{62}$$

Then, inequality (60) can be further expressed as

$$\begin{aligned}\dot{L}_1 \leq & p_1\|x\|^2 + p_2\|\tilde{W}_i\|^2 + p_3\|\tilde{W}_c\|^2 + p_4\|e_i\|^2 \\ & + p_5\|e_c\|^2 + \|\mu^\top \left(u^*(x) - \hat{u}(x_j)\right)\|^2 \\ & - \|\mu^\top \hat{u}(x_j)\|^2 - \rho x^\top Q x + E \\ \leq & p_1\|x\|^2 + p_2\|\tilde{W}_i\|^2 + p_3\|\tilde{W}_c\|^2 + p_4\|e_i\|^2 \\ & + p_5\|e_c\|^2 + \mathcal{L}^2\|\mu\|^2\|e_j(t)\|^2 - \|\mu\|^2\|\hat{u}(x_j)\|^2 \\ & - \rho x^\top Q x + E.\end{aligned} \tag{63}$$

(i) If the reconstruction errors of both identifier and critic NNs are zero, i. e., $\|\varepsilon_i\| = \|\nabla\varepsilon_c\| = \|\varepsilon_{if}\| = 0$, then $\dot{L}_1$ can be re-expressed as

$$\begin{aligned}\dot{L}_1 \leq\ & p_1\|x\|^2 + p_2\|\tilde{W}_i\|^2 + p_3\|\tilde{W}_c\|^2 + p_4\|e_i\|^2 \\ & + p_5\|e_c\|^2 + \mathcal{L}^2\|\mu\|^2\|e_j(t)\|^2 - \|\mu\|^2\|\hat{u}(x_j)\|^2 \\ & - \rho x^\top Q x.\end{aligned} \tag{64}$$

If the triggering error $\|e_j(t)\|^2$ satisfies the triggering condition (40), according to Lyapunov theorem, the NNs weight errors $\|\tilde{W}_i\|$ and $\|\tilde{W}_c\|$ will decay to zero when $t \to \infty$, the closed-loop system is therefore asymptotically stable.

(ii) If the reconstruction errors of both identifier and critic NNs exist, i.e, $E \neq 0$. In this sense, we can conclude that if the following inequalities

$$\begin{aligned}\|x\| >& \sqrt{-E/p_1}, \\ \|\tilde{W}_i\| >& \sqrt{-E/p_2}, \|\tilde{W}_c\| > \sqrt{-E/p_3}, \\ \|e_i\| >& \sqrt{-E/p_4}, \|e_c\| > \sqrt{-E/p_5},\end{aligned} \tag{65}$$

hold, then the convergence of weight estimation errors for identifier NNs and critic NNs, i. e., $\|\tilde{W}_i\|$ and $\|\tilde{W}_c\|$, and the stability of closed-loop system are all guaranteed in the sense of uniformly ultimately bounded (UUB).

Case 2. We consider the situation when an event is triggered. The difference of Lyapunov function L_1 at triggering time $t = \tau_{j+1}$ is defined as

$$\Delta L_1 = \Delta L_{13} + \Delta L_{14} + \underbrace{\Delta L_{11} + \Delta L_{12} + \Delta L_{15} + \Delta L_{16}}_{\Delta L_\varepsilon}, \tag{66}$$

where

$$\Delta L_{13} = \lim_{\Delta t \to 0} \left\{J^*(x_{j+1}) - J^*\left(x\left(\tau_{j+1} - \Delta t\right)\right)\right\}, \tag{67}$$

$$\Delta L_{14} = J^*(x_{j+1}) - J^*(x_j), \tag{68}$$

$$\Delta L_\varepsilon = \lim_{\Delta t \to 0} \left\{L_\varepsilon|_{t=\tau_{j+1}} - L_\varepsilon|_{t=\tau_{j+1}-\Delta t}\right\}. \tag{69}$$

Since we have proved in Case 1 that $\dot{L}_1 < 0$, and since $\tilde{W}_i$, $\tilde{W}_c$, e_i, and e_c are continuous at $t \in [\tau_j, \tau_{j+1})$, we can further obtain

$$\Delta L_1 \leq J^* (x_{j+1}) - J^* (x_j) \leq -K \|x_{j+1} - x_j\| , \tag{70}$$

where $K(\cdot)$ is a class-K function [8]. Based on the above two cases, we can conclude that identifier NNs weight errors $\tilde{W}_i$, critic NNs weight errors $\tilde{W}_c$ and the stability of closed-loop system are all UUB. This completes the proof. □

Remark 6. It is noted from Theorem 1 that the PE condition is necessary for ensuring the convergence performance of NN weights. To the best of our knowledge, adding exploration noise to the controller during the learning process is an effective way of meeting this condition. Additionally, some studies have explored relaxing the PE condition by introducing concurrent learning techniques [30], such that the PE condition can be checked via the rank condition of historical data. Inspired by this idea, we plan to investigate the potential of obtaining relaxed PE conditions for the proposed parameter estimation method.

Theorem 2. Let Assumptions 1-4 hold. Consider system (1) with weight tuning laws of identifier NN and critic NN updated by (22) and (35), respectively. If the event is triggered by the dynamic triggering mechanism (43), then the system stability and the convergence of the weight estimation errors $\tilde{W}_i$ and $\tilde{W}_c$ are all UUB.

Proof. We select the following Lyapunov function candidate L_2 as

$$L_2 = L_1 + y_d. \tag{71}$$

Based on results of (63), the derivation of L_2 can be deduced as

$$\begin{aligned} \dot{L}_2 =& \dot{L}_1 + \dot{y}_d \\ \leq& p_1\|x\|^2 + p_2\|\tilde{W}_i\|^2 + p_3\|\tilde{W}_c\|^2 + p_4\|e_i\|^2 \\ &+ p_5\|e_c\|^2 + E - \phi y_d. \end{aligned} \tag{72}$$

By using comparison lemma [8] and (41), we can obtain $y_d \geq y_d(0) e^{-(\phi + \frac{1}{\rho})t}$, which means that y_d is non-negative. Then, if the conditions in (65) hold, one can conclude that $\|x\|$, $\|\tilde{W}_i\|$ and $\|\tilde{W}_c\|$ are all UUB. This finishes the proof. □

5. SIMULATION RESULTS

In this section, two simulation examples are given to verify the effectiveness of our main results.

Example 1. Consider a nonlinear continuous system as follows:

$$\begin{aligned} \dot{x} =& \begin{bmatrix} -x_1 + x_2 \\ -0.5x_1 - 0.5x_2 \left(1 - (\cos(2x_1) + 2)^2\right) \end{bmatrix} \\ &+ \begin{bmatrix} 0 \\ \cos(2x_1) + 2 \end{bmatrix} u. \end{aligned} \tag{73}$$

We first construct an identifier NN to estimate system parameters of both drift dynamics $f(x)$ and input dynamics $g(x)$, simultaneously. The activation function θ_c is chosen as

$$\theta_c = \left[x_1, x_2, x_2\left(1 - (\cos(2x_1) + 2)^2\right), \cos(2x_1)\, u, u\right]^\top . \tag{74}$$

According to the system model (73), the ideal weight matrix W_i is

$$W_i = \begin{bmatrix} W_i^{11}, W_i^{12}, W_i^{13}, W_i^{14}, W_i^{15} \\ W_i^{21}, W_i^{22}, W_i^{23}, W_i^{24}, W_i^{25} \end{bmatrix} = \begin{bmatrix} -1, 1, 0, 0, 0 \\ -0.5, 0, -0.5, 1, 2 \end{bmatrix}.$$

Some parameters of system identification are chosen as: $l_i = 0.001$, $\xi = 1$, $\Pi_i = 200$. The initial system state is $x(0) = [3, -1]^\top$, and the initial identifier NN weights are

$$W_i(0) = \begin{bmatrix} 0.4854, 0.1419, 0.9157, 0.9595, 0.0357 \\ 0.8003, 0.4218, 0.7922, 0.6557, 0.8491 \end{bmatrix}.$$

For critic NNs, the activation function θ_c is selected as

$$\theta_c = \begin{bmatrix} x_1^2 & x_1 x_2 & x_2^2 \end{bmatrix}^\top . \tag{75}$$

Some parameters of critic NNs are chosen as: $l_c = 150$, $\Pi_c = 200diag(0.3, 1, 1)$. Q and R are all identity matrices. The initial critic NN weights are set to be $W_c(0) = \left[W_c^1(0), W_c^2(0), W_c^3(0)\right]^\top = [0.9509, 0.7223, 0.4001]^\top$. In the proposed event-triggered designs, some parameters are set as: $\phi = 1$, $\rho = 0.84$, $\mathcal{L} = 30$, $\mu = 1$ and $\rho_d = 10$, and $y_d(0) = 1$. To satisfy PE condition in Assumption 1 in the training process, a probing noise δ_p is added to control input u for 1 second, i. e.,

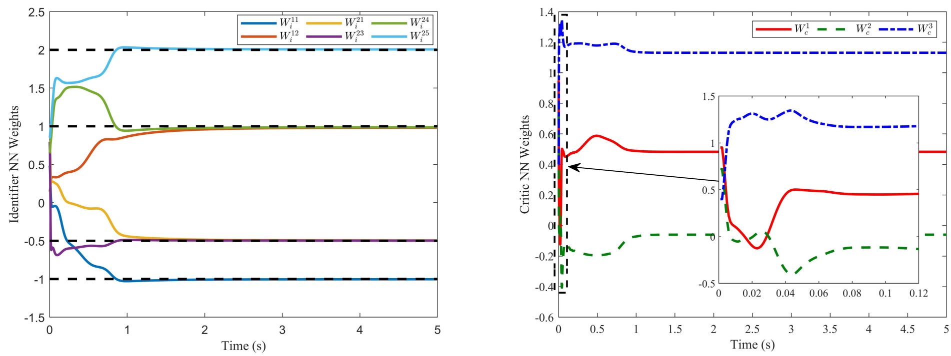
$$\delta_p = 0.25e^{-t}(\sin^2(t)\cos(t) + \sin^2(2t)\cos(0.1t) + \sin^2(-1.2t)\cos(0.5t) + \sin^5(t)).$$

Figure 2(a) shows the training process of the identifier weights W_i^{11}, W_i^{12}, W_i^{21}, W_i^{23}, W_i^{24}, W_i^{25} under dynamic trigger mechanisms. It shows that the weight tuning law can drive the estimated weights approach the true values, that is,

$$W_i = \begin{bmatrix} -1.0029, 0.9818, -0.0015, 0.0061, -0.0030 \\ -0.4945, 0.0341, -0.4972, 0.9886, 2.0057 \end{bmatrix}.$$

Figure 2(b) illustrates the training process of the critic NN weights. Figure 3 compares the trajectories of system states $x = [x_1, x_2]^\top$ under the proposed STM and DTM-based critic learning control approaches, as well as the event-triggered scheme, $\|e_j(t)\|^2 \leq \frac{(1-\rho^2)\lambda_{\min}(Q)}{(1+\epsilon^2)\mathcal{L}^2}\|x\|^2$ with $\epsilon = 1$, proposed in [30]. It is observed that the evolution of the system state is similar across all three methods. Figure 4 shows the comparison of control inputs under the proposed STM and DTM, as well as the event-triggered scheme proposed in [30].

The performance of the event-triggered control is presented in Figure 5 and Figure 6. Figure 5(a) and Figure 5(b) illustrate the relationship between triggering error and triggering threshold for the static and dynamic triggering methods, respectively. With the static and dynamic triggering schemes, the triggering times are reduced to



(a) (b)

Fig. 2. Training process of (a) the identifier NN weight W_i and (b) the critic NN weight W_c.

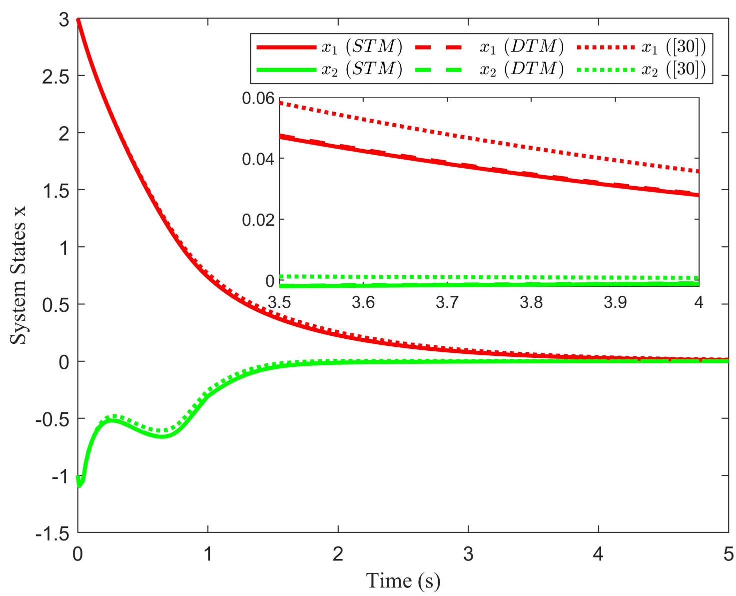


Fig. 3. The evolution of the state $x = [x_1, x_2]^\top$ under STM, DTM and the event-triggered mechanism in [30].

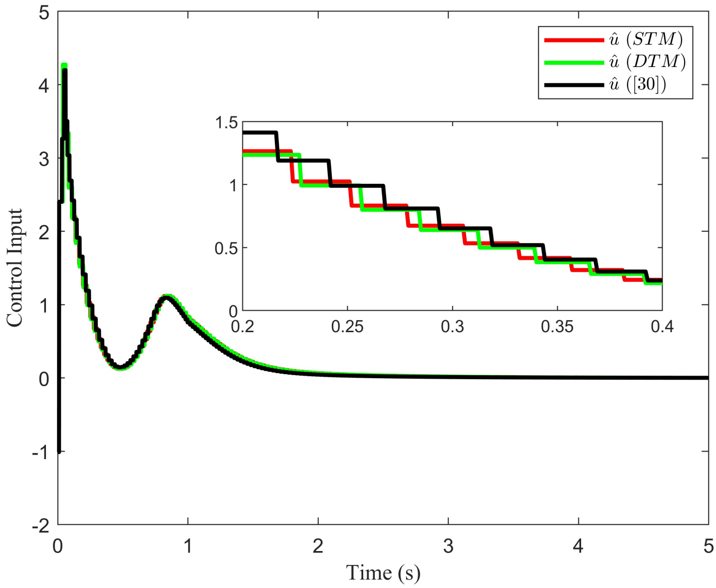


Fig. 4. The trajectories of control inputs $\hat{u}(x_k)$ under STM, DTM and in [30].

175 and 92, respectively, compared to 5000 times for the time-based triggering scheme. In other words, the static and dynamic triggering mechanisms reduce sampling times by approximately 96.50% and 98.16%, respectively. Figure 6 depicts the trigger interval $T_{trig} = \tau_j - \tau_{j-1}$ of the two triggering schemes. These results demonstrate that there are two advantages of the dynamic triggering mechanism when compared with the static triggering mechanism: 1) it triggers fewer state samples and 2) the triggering interval gradually increases over time. Therefore, the simulation results demonstrate that the proposed event-triggered mechanisms can effectively reduce system's computational resources.

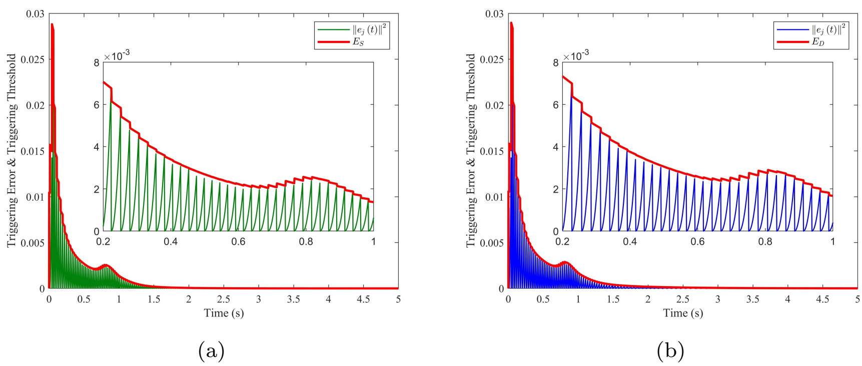


Fig. 5. Triggering error and threshold under (a) STM and (b) DTM.

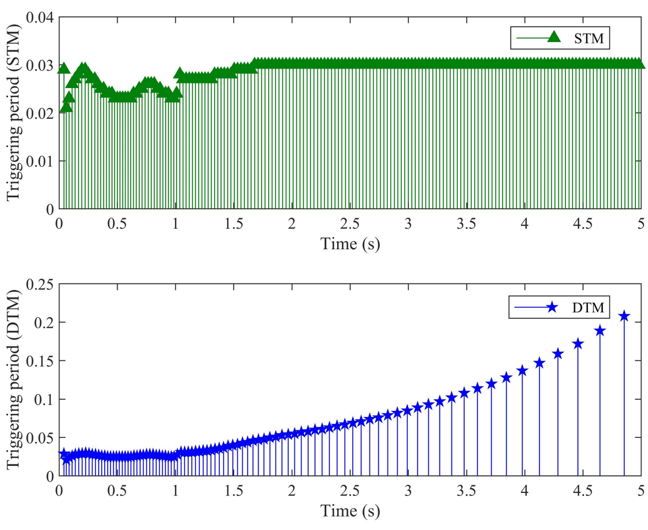


Fig. 6. The trigger instants under STM and DTM.

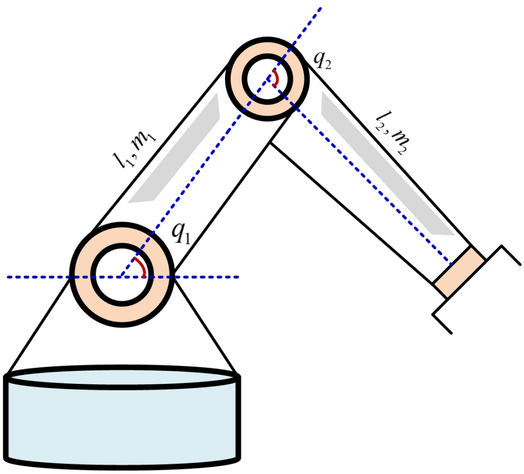


Fig. 7. Diagram of a 2-DOF robotic manipulator.

Example 2. Consider a 2-degree-of-freedom (DOF) robotic manipulator (RM) system [17], which is shown in Figure 7 to test the effectiveness of the proposed event-triggered optimal control method. The 2-DOF RM system model and system parameters are represented as

$$H(q)\ddot{q} + D(q,\dot{q})\dot{q} + G(q) = \tau, \tag{76}$$

$$H(q) = \begin{bmatrix} s_1 + s_2 + 2s_3\cos(q_2) & s_2 + s_3\cos(q_2) \\ s_2 + s_3\cos(q_2) & s_2 \end{bmatrix},$$

$$D(q,\dot{q}) = \begin{bmatrix} -s_3\dot{q}_2\sin(q_2) & -s_3(\dot{q}_1 + \dot{q}_2)\sin(q_2) \\ -s_3\dot{q}_1\sin(q_2) & 0 \end{bmatrix},$$

$$G(q) = \begin{bmatrix} 0 \\ 0 \end{bmatrix},$$

where $s_1 = m_1{l_{c1}}^2 + m_2{l_1}^2 + S_1,\ s_2 = m_2 l_{c2}^2 + S_2, s_3 = m_2 l_1 l_2,\ s_4 = m_1 l_{c2} + m_2 l_1,\ s_5 = m_2 l_{c2}$ with $l_{c1} = l_1/2,\ l_{c2} = l_2/2$. The joint state vector is $q = [q_1, q_2]^\top$ and velocity vector is $\dot{q} = [q_3, q_4]^\top$. $\tau = [\tau_1, \tau_2]^\top$ is the control input of the system. The system model parameters are given in the Table 1.

Length of link 1	l_1	1.01m
Length of link 2	l_2	0.82m
Mass of link 1	m_1	4kg
Mass of link 2	m_2	3kg
Inertia of link1	S_1	1.021kgm^2
Inertia of link2	S_2	0.5043kgm^2

Tab. 1. Parameters of the 2-DOF robotic system.

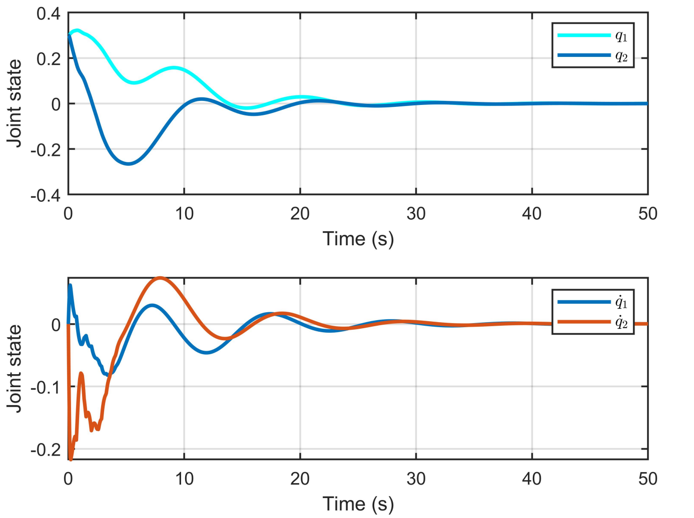


Fig. 8. The trajectories of robot system states $x = [q_1, q_2, \dot{q}_1, \dot{q}_2]^\top$ under DTM-based control methods.

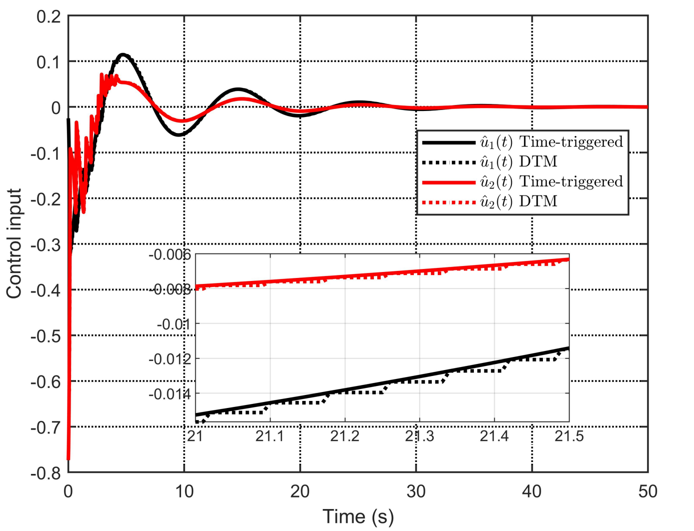


Fig. 9. The comparisons of control inputs $\hat{u}(x_k)$ under time-triggered method and DTM.

In the critic NN, the activation function is chosen as

$$\theta_c = [q_1^2, q_1q_2, q_1q_3, q_1q_4, q_2^2, q_2q_3, q_2q_4, q_3^2, q_3q_4, q_4^2]^\top.$$

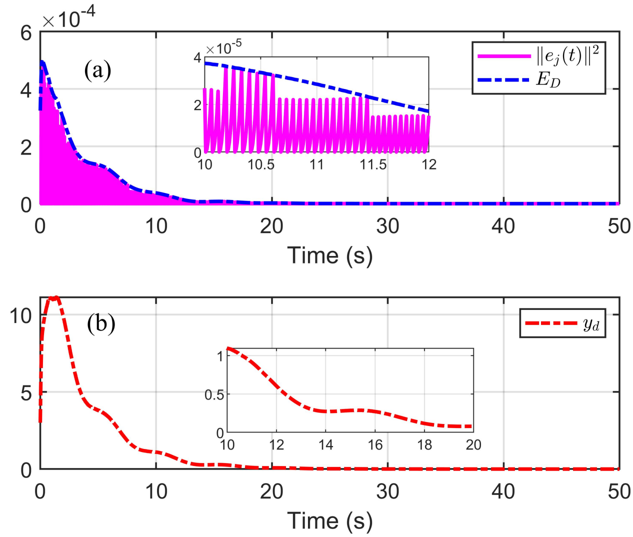


Fig. 10. (a) Triggering error and threshold under DTM; (b) The evolution of filtered variable $y_d(t)$.

The initial joint states of the 2-DOF robotic system are $q(0) = [0.3, 0.3]^\top$ and $\dot{q}(0) = [0, 0]^\top$. The initial critic NN weights are set to be

$$W_c(0) = [0.4795, 0.6393, 0.5447, 0.6473, 0.5439, 0.7210, 0.5225, 0.9937, 0.2187, 0.1058]^\top.$$

$Q = 10I_4$ and $R = I_2$. In our event-triggered scheme, some parameters are set to be: $\phi = 1$, $\rho = 0.84$, $\mathcal{L} = 27$, $\mu = 3$ and $\rho_d = 3$. The initial value of filtered variable is $y_d(0) = 3$. Figure 8 illustrates the trajectory of the system states. Figure 9 shows the comparisons of control inputs under the proposed DTM method and the traditional time-based control method. Figure 10 (a) reflects the relationship between the triggering error $\|e_j(t)\|^2$ and dynamic triggering threshold E_D. The evolution of filtered variable $y_d(t)$ is shown in Figure 10 (b), which confirms that the filtered variable $y_d(t)$ is non-negative. Therefore, the simulation results demonstrate the effectiveness of the proposed event-triggered identifier-critic learning control method.

Remark 7. Note that in order to test the proposed control methods on the RM system (Example 2), we should first transfer the RM system (76) to a general nonlinear continuous-time system (1). In this sense, as shown in [17], only $G = 0$ can satisfy $f(0) = 0$, which is a common assumption to obtain the main results in the ADP field.

6. CONCLUSIONS

This paper proposed an online identifier-critic learning control framework for unknown nonlinear systems based on event-triggered methods. Different from traditional ADP methods, a filter-regressor-based parameter estimation approach was proposed to reconstruct unknown system dynamics, thereby removing the need for an accurate system model in the control design loop. Meanwhile, NN weight adaptation rules of the identifier NN and the critic NN were designed for the parameter estimation by utilizing only measured system state and input data. Furthermore, two kinds of aperiodic sampling schemes were included in the proposed learning control framework to reduce state sampling frequency. The stability of the closed-loop system and the convergence of adaptive laws were analyzed. Finally, the simulation results were presented to demonstrate the performance of the proposed event-triggered learning control strategy.

ACKNOWLEDGEMENTS

This work was supported partially by the National Key Research and Development Program of China under Grant 2022YFE0133100, Grant 2018AAA0102504, partially by the National Natural Science Foundation of China under Grant 62203089, Grant 71503206, and Grant 61473061, partially by the China Postdoctoral Science Foundation under Grant 2021M700695, and Grant 2020M683274, and partially by the Sichuan Science and Technology Program under Grants 2022NSFSC0890, 2021YFG0184 and 2018GZDZX0037, and partially by the Guangdong Basic and Applied Basic Research Foundation under Grant 2022A1515110135.

(Received October 20, 2022)

REFERENCES

[1] S. Bhasin, R. Kamalapurkar, M. Johnson, K. G. Vamvoudakis, F. L. Lewis, and W. E. Dixon: A novel actor-critic-identifier architecture for approximate optimal control of uncertain nonlinear systems. Automatica *49* (2013), 82–92. DOI:10.1016/j.automatica.2012.09.019

[2] B. Chen, J. Hu, Y. Zhao, and B. K. Ghosh: Finite-time observer based tracking control of uncertain heterogeneous underwater vehicles using adaptive sliding mode approach. Neurocomputing *481* (2022), 322–332. DOI:10.1016/j.neucom.2022.01.038

[3] X. Fu and Z. Li: Neural network optimal control for nonlinear system based on zero-sum differential game. Kybernetika *57* (2021), 546–566. DOI:10.14736/kyb-2021-3-0546

[4] A. Girard: Dynamic triggering mechanisms for event-triggered control. IEEE Trans. Autom. Control *60* (2015), 1992–1997. DOI:10.1109/TAC.2014.2366855

[5] J. Hu, G. Chen, and H.-X. Li: Distributed event-triggered tracking control of leader-follower multi-agent systems with communication delays. Kybernetika *47* (2011), 630–643.

[6] J. Hu, J. Geng, and H. Zhu: An observer-based consensus tracking control and application to event-triggered tracking. Commun. Nonlinear Sci. Numer. Simul. *20* (2015), 559–570. DOI:10.1016/j.cnsns.2014.06.002

[7] Y. Jiang and Z. P. Jiang: Computational adaptive optimal control for continuous-time linear systems with completely unknown dynamics. Automatica *48* (2012), 2699–2704. DOI:10.1016/j.automatica.2012.06.096

[8] H. K. Khalil: Nonlinear Systems. Third Edition. Prentice-Hallm Upper Saddle River, NJ 2002.

[9] B. Kiumarsi and F. L. Lewis: Actor-critic-based optimal tracking for partially unknown nonlinear discrete-time systems. IEEE Trans. Neural Netw. Learn. Syst. *26* (2015), 140–151. DOI:10.1109/TNNLS.2014.2358227

[10] G. Kreisselmeier: Adaptive observers with exponential rate of convergence. IEEE Trans. Autom. Control *AC-22* (1977), 2–8.

[11] F. Lewis, S. Jagannathan, and A. Yesildirak: Neural Network Control of Robot Manipulators and Nonlinear Systems. Taylor and Francis, London 1999.

[12] F. L. Lewis, D. L. Vrabie, and V. L. Syrmos: Optimal Control. Third Edition. Wiley, New York 2012. DOI:10.1002/9781118122631

[13] R. Luo, Z. Peng, J. Hu, and B.K. Bijoy: Adaptive optimal control of completely unknown systems with relaxed PE conditions. In: Proc. IEEE 11th Data Driven Control and Learning Systems Conference (DDCLS), Chengdu 2022, pp. 836–841. DOI:10.1109/DDCLS55054.2022.9858418

[14] Y. Lv, J. Na, Q. Yang, X. Wu, and Y. Guo: Online adaptive optimal control for continuous-time nonlinear systems with completely unknown dynamics. Int. J. Control *89* (2016), 99–112. DOI:10.1080/00207179.2015.1060362

[15] R. Luo, Z. Peng, and J. Hu: On model identification based optimal control and it's applications to multi-agent learning and control. Mathematics *11* (2023), 906. DOI:10.3390/math11040906

[16] W. Makumi, M. L. Greene, Z. Bell, B. Bialy, R. Kamalapurkar, and W. Dixon: Hierarchical reinforcement learning and gains cheduling-based control of a hypersonic vehicle. AIAA SCITECH 2023 Forum,National Harbor, MD and Online, 2023, 1–11. DOI:10.2514/6.2023-2505

[17] Y. Ouyang, L. Dong, and C. Sun: Critic learning-based control for robotic manipulators with prescribed constraints. IEEE Trans. Cybern. *52* (2022), 2274–2283. DOI:10.1109/TCYB.2020.3003550

[18] Z. Peng, R. Luo, J. Hu, K. Shi, and B. K. Ghosh: Distributed optimal tracking control of discrete-time multiagent systems via event-triggered reinforcement learning. IEEE Trans. Circuits Syst. I-Regul. Pap. *69* (2022), 3689–3700. DOI:10.1109/TCSI.2022.3177407

[19] Z. Peng, R. Luo, J. Hu, K. Shi, S. K. Nguang, and B. K. Ghosh: Optimal tracking control of nonlinear multiagent systems using internal reinforce Q-learning. IEEE Trans. Neural Netw. Learn. Syst. *33* (2022), 4043–4055. DOI:10.1109/TNNLS.2021.3055761

[20] Z. Peng, Y. Zhao, J. Hu, R. Luo, B. K. Ghosh, and S. K. Nguang: Input-output data-based output antisynchronization control of multiagent systems using reinforcement learning approach. IEEE Trans. Ind. Inform. *17* (2021), 7359–7367. DOI:10.1109/TII.2021.3050768

[21] M. Shen, X. Wang, J. H. Park, Y. Yi, and W.-W. Che: Extended disturbance-observer-based data-driven control of networked nonlinear systems with event-triggered output. IEEE Trans. Syst. Man Cybern. Syst. to be published. DOI:10.1109/TSMC.2022.3222491

[22] R. Song, F. Lewis, Q. Wei, H. G. Zhang, Z. P. Jiang, and D. Levine: Multiple actor-critic structures for continuous-time optimal control using input-output data. IEEE Trans. Neural Netw. Learn. Syst. *26* (2015), 851–865. DOI:10.1109/TNNLS.2015.2399020

[23] P. Tabuada: Event-triggered real-time scheduling of stabilizing control tasks. IEEE Trans. Autom. Control *52* (2007), 1680–1685. DOI:10.1109/TAC.2007.904277

[24] K. Wang and C. Mu: Event-sampled learning for unknown nonlinear systems related to dynamic triggering method. In: Proc. IEEE Conference on Decision and Control (CDC), Jeju 2020, pp. 5200–5205. DOI:10.1109/CDC42340.2020.9303929

[25] D. Wang, C. Mu, and D. Liu: Adaptive critic designs for solving event-based H_∞ control problems. In: Proc. American Control Conference (ACC), Seattle 2017, pp. 2435–2400. DOI:10.23919/ACC.2017.7963318

[26] X. Wang, W. Qin, J. H. Park, and M. Shen: Event-triggered data-driven control of discrete-time nonlinear systems with unknown disturbance. ISA Trans. *128* (2022), 256–264. DOI:10.1016/j.isatra.2021.11.026

[27] P. J. Werbos: Approximate dynamic programming for real-time control and neural modeling. In: Handbook of Intelligent Control: Neural, Fuzzy, and Adaptive Approaches (D. A. White and D. A. Sofge, Eds.), Van Nostrand Reinhold, New York 1992, ch. 13.

[28] N. Xu, B. Niu, H. Wang, X. Huo, and X. Zhao: Single-network ADP for solving optimal event-triggered tracking control problem of completely unknown nonlinear systems. Int. J. Intell. Syst. *36* (2021), 4795–4815. DOI:10.1002/int.22491

[29] S. Xue, B. Luo, D. Liu, and Y. Gao: Adaptive dynamic programming-based event-triggered optimal tracking control. Int. J. Robust Nonlinear Control *31* (2021), 7480–7497. DOI:10.1002/rnc.5687

[30] X. Yang and H. He: Adaptive critic designs for event-triggered robust control of nonlinear systems with unknown dynamics. IEEE Trans. Cybern. *49* (2019), 2255–2267. DOI:10.1109/TCYB.2018.2823199

[31] X. Yang, H. He, and D. Liu: Event-triggered optimal neuro-controller design with reinforcement learning for unknown nonlinear systems. IEEE Trans. Syst. Man Cybern. Syst. *49* (2019), 1866–1878. DOI:10.1109/TSMC.2017.2774602

Zhinan Peng, School of Automation Engineering, University of Electronic Science and Technology of China, Chengdu 611731, P. R. China; Institute of Electronic and Information Engineering of University of Electronic Science and Technology of China, Dongguan 523808. P. R. China.
e-mail: zhinanpeng@uestc.edu.cn

Zhiquan Zhang, School of Engineering and Applied Science, Department of Electrical and System Engineering, University of Pennsylvania, Philadelphia, PA, 19104. USA.
e-mail: zzq2000@seas.upenn.edu

Rui Luo, School of Automation Engineering, University of Electronic Science and Technology of China, Chengdu 611731. P. R. China.
e-mail: nicole9922@163.com

Yiqun Kuang, School of Automation Engineering, University of Electronic Science and Technology of China, Chengdu 611731. P. R. China.
e-mail: yqkuang@uestc.edu.cn

Jiangping Hu, Corresponding author, School of Automation Engineering, University of Electronic Science and Technology of China, Chengdu 611731, P. R. China; Yangtze Delta Region Institute (Huzhou), University of Electronic Science and Technology of China, Huzhou 313001. P. R. China.
e-mail: hujp@uestc.edu.cn

Hong Cheng, School of Automation Engineering, University of Electronic Science and Technology of China, Chengdu 611731. P. R. China.
e-mail: hcheng@uestc.edu.cn

Bijoy Kumar Ghosh, Department of Mathematics and Statistics, Texas Tech University, Lubbock, TX, 79409-1042, USA; School of Automation Engineering, University of Electronic Science and Technology of China, Chengdu 611731. P. R. China.
e-mail: bijoy.ghosh@ttu.edu