Event-Triggered Adaptive Dynamic Programming for Uncertain Nonlinear Systems

Qichao Zhang[1,2], Dongbin Zhao*[1,2], and Ding Wang[1,2]

[1]The State Key Laboratory of Management and Control for Complex Systems,
Institute of Automation, Chinese Academy of Sciences, Beijing, 100190, China
[2] University of Chinese Academy of Sciences, Beijing, 100049, China
Email: zhangqichao2013@163.com, dongbin.zhao@ia.ac.cn, ding.wang@ia.ac.cn

Abstract. In this paper, the robust control for a class of continuous-time nonlinear system with unmatched uncertainties is investigated using an event-triggered adaptive dynamic programming method. First, the robust control problem is solved using the optimal control method. Under the event-triggered mechanism, the solution of the optimal control problem can asymptotically stabilize the uncertain system with an designed triggering condition. That is, the designed event-triggered controller is robust to the original uncertain system. Then, a single critic network structure with experience replay technique is constructed to approach the optimal control policies. Finally, a simulation example is provided to demonstrate the effectiveness of the proposed control scheme.

Keywords: Adaptive dynamic programming, event-triggered control, robust control, neural network

1   Introduction

As many practical control systems become more and more complex, the uncertainties arise in the system models frequently. These uncertainties may severely degrade the system performance, and even lead to system instability, so it is necessary to design the robust controller for uncertain nonlinear systems. Lin *et al.* [1] established a connection between the robust control problem and optimal control problem. They proposed an indirect approach for the robust stabilization by designing a corresponding optimal controller. In [2], the uncertainties was divided into matched and unmatched ones and proved that the optimal controller can stabilize the linear and nonlinear robust control systems in the same way. However, the detailed approach to solve the Hamilton-Jacobi-Bellman (HJB) equation to obtain the optimal control policy was not discussed.

As is known, it is intractable to give an analytic solution to the HJB equation for the nonlinear systems [3]. Recently, adaptive dynamic programming (ADP) which was proposed by Werbos [4] has been widely applied to approximate the solution of the HJB equation. For example, the H_∞ control approach based on ADP was investigated for the uncertain nonlinear systems in [5]. Jiang *et al.* proposed a robust adaptive dynamic programming (RADP) methodology in [6].

Liu *et al.* [7] investigated the optimal robust guaranteed cost control problem using corresponding optimal control method. In [8], the robust control problem of nonlinear systems with matched uncertainties was converted into an optimal control problem of an nominal system. For the nonlinear deterministic systems, Zhao *et al.* developed several ADP methods to solve the corresponding optimal control problems [9,10]. However, the aforementioned approaches are conducted predicated on the traditional time-triggered strategy.

In general, the amount of transmitted data is huge using the traditional time-triggered approach. To mitigate the unnecessary waste of communication resources, event-triggered control (ETC) method has received great interests among the control researchers.Recently, the ETC method has been integrated with the ADP approach to solve the optimal control problems [11]. In [12], Sahoo *et al.* proposed a neural network (NN)-based ETC scheme for nonlinear discrete-time systems using ADP approach. Vamvoudakis proposed an optimal adaptive ETC algorithm based on the actor-critic structure for CT nonlinear systems with guaranteed performance in [13]. In [14], the event-triggered reinforcement learning approach was developed for the nonlinear systems without requiring exact knowledge of system dynamics. However, the system uncertainties are not concerned in the existing work of event-triggered optimal control.

In this paper, we investigate the robust control problem of nonlinear systems with unmatched uncertainties using an optimal control approach.

Section II introduces the robust control problem of the nonlinear system with unmatched uncertainties and the traditional optimal control problem. In Section III, the connection between the robust stabilization and the optimal control problem is discussed. In section IV, the event-triggered ADP algorithm is proposed to approximate the optimal control policy. Simulation results and the conclusion are presented in the end.

2   Problem Statement

Consider the following CT uncertain nonlinear system

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

where $x = x(t) \subseteq \mathbb{R}^n$ is the state vector, $u = u(x) \in \mathbb{R}^m$ is the control input, $f(\cdot) \in \mathbb{R}^n$, $g(\cdot) \in \mathbb{R}^{n \times m}$ and $k(\cdot) \in \mathbb{R}^{n \times q}$ are differentiable nonlinear dynamics with $f(0) = 0$, and $W(\cdot) \in \mathbb{R}^q$ is the unknown nonlinear perturbation. Assume that $W(0) = 0$, so that $x = 0$ is an equilibrium of system (1).

The uncertainty $W(x)$ is known as an unmatched uncertainty for system (1), if $k(x) \neq g(x)$. In this paper, we aim to find a control policy so that the system (1) is globally asymptotically stable for all unmatched uncertainties $W(x)$ satisfying the following assumption.

Assumption 1 [2]:

1. The uncertainty $W(x)$ is bounded by a known non-negative function $W_M(x)$, i.e., $\|W(x)\| \leq W_M(x)$ with $W_M(0) = 0$.

2. There exists a non-negative function $g_M(x)$ such that

$$\|g^+(x)k(x)W(x)\|^2 \leq \frac{g_M^2(x)}{2} \qquad (2)$$

where $g^+(x)$ denotes the (Moore-Penrose) pseudoinverse of function $g(x)$.

Motivated by [2], the robust control problem of the uncertain nonlinear system will be converted into an optimal control problem of a corresponding auxiliary system with an appropriate cost function. First, the uncertainty term $k(x)W(x)$ is decomposed into a matched component and an unmatched one in the range space of $g(x)$

$$k(x)W(x) = g(x)g^+(x)k(x)W(x) + (I - g(x)g^+(x))k(x)W(x) \qquad (3)$$

Then, we can transform the robust control problem into an optimal control problem as follows.

Optimal Control Problem: For the corresponding auxiliary system

$$\dot{x} = f(x) + g(x)u(x) + \big(I - g(x)g^+(x)\big)k(x)w(x) \qquad (4)$$

where $w = w(x) \in \mathbb{R}^q$ is an augmented control to deal with the unmatched uncertainty component, and $[u^\mathrm{T}(x), w^\mathrm{T}(x)]^\mathrm{T}$ is a control policy pair of system (4).

Assume that the auxiliary system (4) is controllable. It is desired to find the optimal control policy pair $[u^{*\mathrm{T}}(x), w^{*\mathrm{T}}(x)]^\mathrm{T}$ that minimizes the cost function given by

$$V\big(x(0)\big) = \int_0^\infty \|r^\mathrm{T}\|^2 g_M^2(x) + \eta^2\|m^\mathrm{T}\|^2 W_M^2(x) + U(x,u,w)\mathrm{d}t \qquad (5)$$

where the utility $U(x,u,w) = x^\mathrm{T}Qx + u^\mathrm{T}(x)Ru(x) + \eta^2 w^\mathrm{T}(x)Mw(x)$, and $\eta > 0$ is a designed parameter. Here, Q, R and M are positive definite symmetric matrices. According to the principle of Cholesky decomposition, we have $R = rr^\mathrm{T}$ and $M = mm^\mathrm{T}$, where r and m are two appropriate lower triangular matrices.

Remark 1 *For the optimal control problem of the auxiliary system (4), the designed feedback control inputs should be admissible (see [5] for definition). In this paper, we use $\Phi(\Omega)$ to denote the set of admissible policies on a compact set Ω.*

For any admissible policies $u, w \in \Phi(\Omega)$, if the cost function (5) is continuously differentiable, the infinitesimal version of (5) is the so-called nonlinear Lyapunov equation

$$\begin{aligned} &\nabla V^\mathrm{T} \big(f(x) + g(x)u(t) + (I - g(x)g^+(x))k(x)w(t)\big) \\ &+ \|r^\mathrm{T}\|^2 g_M^2(x) + \eta^2\|m^\mathrm{T}\|^2 W_M^2(x) + U(x,u,w) = 0, \end{aligned} \qquad (6)$$

where $\nabla V = \partial V(x)/\partial x$ is the partial derivative of the cost function $V(x)$ with respect to the state x, and $V(0) = 0$.

Define the Hamiltonian function of system (4) as

$$
\begin{aligned}
H(x, \nabla V, u, w) =& \|r^{\mathrm{T}}\|^2 g_M^2(x) + \eta^2 \|m^{\mathrm{T}}\|^2 W_M^2(x) + U(x, u, w) \\
& + (\nabla V)^{\mathrm{T}} \left(f(x) + g(x)u(t) + (I - g(x)g^+(x))k(x)w(t) \right)
\end{aligned}
\tag{7}
$$

Then the optimal cost function of system (4)

$$
\begin{aligned}
V^*\big(x(0)\big) = \min_{u,w \in \Phi(\Omega)} \int_0^\infty & \big\{ \|r^{\mathrm{T}}\|^2 g_M^2(x(\tau)) + \eta^2 \|m^{\mathrm{T}}\|^2 \\
& \times W_M^2(x(\tau)) + U(x(\tau), u(\tau), w(\tau)) \big\} \,\mathrm{d}\tau
\end{aligned}
\tag{8}
$$

satisfies the associated HJB equation

$$
\min_{u,w \in \Phi(\Omega)} H(x, \nabla V^*, u, w) = 0
\tag{9}
$$

where $V^*(x)$ is a solution of the HJB equation.

Assume that the minimum policy pair on the left-hand side of (9) exists and is unique. According to the stationary conditions, the optimal control policies are given by

$$
u^*(x) = -\frac{1}{2} R^{-1} g^{\mathrm{T}}(x) \nabla V^*
\tag{10}
$$

$$
w^*(x) = -\frac{1}{2\eta^2} M^{-1} k^{\mathrm{T}}(x)(I - g(x)g^+(x))^{\mathrm{T}} \nabla V^*
\tag{11}
$$

Denote $d(x) = (I - g(x)g^+(x))k(x)$. Based on (10) and (11), the HJB equation (9) can be rewritten as

$$
\begin{aligned}
& H(x, \nabla V^*, u^*, w^*) \\
=& (\nabla V^*)^{\mathrm{T}} f(x) + x^{\mathrm{T}} Q x + \|r^{\mathrm{T}}\|^2 g_M^2(x) + \eta^2 \|m^{\mathrm{T}}\|^2 W_M^2(x) \\
& - \frac{1}{4} (\nabla V^*)^{\mathrm{T}} g(x) R^{-1} g^{\mathrm{T}}(x) \nabla V^* - \frac{1}{4\eta^2} (\nabla V^*)^{\mathrm{T}} d(x) M^{-1} d^{\mathrm{T}}(x) \nabla V^* = 0
\end{aligned}
\tag{12}
$$

So far, the robust control problem is transformed into a corresponding time-triggered optimal control problem. Then the traditional ADP technique can be employed to approximate the solution $V^*(x)$ of the HJB equation. In order to reduce the computational burden and save communication resources, the ETC mechanism is introduced in this paper. And an adaptive triggering condition will be designed to guarantee the stability of the uncertain system with an event-triggered optimal controller.

3   Event-Triggered Robust Optimal Controller

To propose the ETC mechanism, we first define a monotonically increasing sequence of triggering instants $\{\tau_j\}_{j=0}^\infty$, where τ_j is the jth consecutive sampling instant with $\tau_j < \tau_{j+1}, j \in \mathbb{N}$ with $\mathbb{N} = \{0, 1, 2, \cdots\}$. Then an sampled-data system characterized by the triggering instants is introduced, where the controller

is updated based on the sampled state $\hat{x}_j = x(\tau_j)$ for all $t \in [\tau_j, \tau_{j+1})$. Define the event-triggered error as

$$e_j(t) = \hat{x}_j - x(t), \forall t \in [\tau_j, \tau_{j+1}), j \in \mathbb{N} \tag{13}$$

where $x(t)$ and $\hat{x}_j$ denote the current state and the sampled state, respectively.

In the ETC method, the triggering condition is determined by the event-triggered error and a designed state-dependent threshold. When the event-triggered error exceeds the state-dependent threshold, an event is triggered. Then, the system states are sampled that resets the event-triggered error $e_j(t)$ to zero, and be held until the next triggering instant. Accordingly, the designed event-triggered controller $u(\hat{x}_j) \triangleq \mu(\hat{x}_j)$ is updated. Clearly, the control signal $\mu(\hat{x}_j)$ is a function of the event-based state vector, which is executed based on the latest sampled state $\hat{x}_j$ instead of the current value $x(t)$. That is, the event-triggered controller is only updated at the triggering instant sequence $\{\tau_j\}_{j=0}^\infty$ and remains unchanged in each time interval $t \in [\tau_j, \tau_{j+1})$. Hence, this control signal $\mu(\hat{x}_j)$ with $j \in \mathbb{N}$ is a piecewise constant function on each segment $[\tau_j, \tau_{j+1})$.

Under the event-triggering mechanism, the transformed optimal control problem in the previous section can be restated as follows.

With the event-triggered control input $\mu(\hat{x}_j)$, the sampled-data version of the auxiliary system (4) can be written as

$$\dot{x}(t) = f(x) + g(x)\mu\left(x(t) + e_j(t)\right) + d(x)w\left(x(t)\right) \tag{14}$$

Considering the event-triggered sampling rule, the optimal control policy (10) becomes

$$\mu^*(\hat{x}_j) = -\frac{1}{2}R^{-1}g^{\mathrm{T}}(\hat{x}_j)\nabla V^*(\hat{x}_j) \tag{15}$$

for all $t \in [\tau_j, \tau_{j+1})$, where $\nabla V^*(\hat{x}_j) = \partial V^*(x)/\partial x|_{x=\hat{x}_j}$.

By using the optimal cost function $V^*(x)$, the event-triggered controller (15) and the augmented controller (11), the Hamiltonian function (7) becomes

$$
\begin{aligned}
& H(x, \nabla V^*, \mu^*(\hat{x}_j), w^*(x)) \\
& = (\nabla V^*)^{\mathrm{T}}f(x) + x^{\mathrm{T}}Qx + \|r^{\mathrm{T}}\|^2 g_M^2(x) + \eta^2\|m^{\mathrm{T}}\|^2 W_M^2(x) \\
& \quad - \frac{1}{2}(\nabla V^*)^{\mathrm{T}}g(x)R^{-1}g^{\mathrm{T}}(\hat{x}_j)\nabla V^*(\hat{x}_j) \\
& \quad + \frac{1}{4}(\nabla V^*(\hat{x}_j))^{\mathrm{T}}g(\hat{x}_j)R^{-1}g^{\mathrm{T}}(\hat{x}_j)\nabla V^*(\hat{x}_j) \\
& \quad - \frac{1}{4\eta^2}(\nabla V^*)^{\mathrm{T}}d(x)M^{-1}d^{\mathrm{T}}(x)\nabla V^*
\end{aligned}
\tag{16}
$$

For convenience of analysis, results of this paper are based on the following assumptions.

Assumption 2 [8]: $f + gu + dw$ is Lipschitz continuous on $\Omega \in \mathbb{R}^n$ containing the origin.

Assumption 3 [13]: The controller $u(x)$ is Lipschitz continuous with respect to the event-triggered error,

$$\|u(x(t)) - u(\hat{x}_j)\| = \|u(x(t)) - u(x(t) + e_j(t))\| \leq L\|e_j(t)\| \tag{17}$$

where L is a positive real constant.

Remark 2 *This assumption is satisfied in many applications where the controller are affine with respect to e_j. Note that $w(t)$ is not the direct control policy of the robust control system (1), but it plays an important role in finding the event-triggered optimal control policy $\mu^*(\hat{x}_j)$ for the system (14).*

Remark 3 *Combined (12) and (16), we have*

$$
\begin{aligned}
&H(x, \nabla V^*, \mu^*(\hat{x}_j), w^*(x)) \\
=&(\nabla V^*)^T \left(f(x) + g(x)\mu^*(\hat{x}_j) + d(x)w^*(x)\right) + \|r^T\|^2 g_M^2(x) \\
&+ \eta^2 \|m^T\|^2 W_M^2(x) + U\left(x, \mu^*(\hat{x}_j), w^*(x)\right) \\
=& \left(r^T(u^*(x) - \mu^*(\hat{x}_j))\right)^T \left(r^T(u^*(x) - \mu^*(\hat{x}_j))\right)
\end{aligned}
\tag{18}
$$

It is called the event-triggered HJB equation. Different from the traditional HJB equation (12), the event-triggered HJB equation is only equal to zero at every triggering instant. In other words, a transformation error is introduced due to the event-triggered transformation from (10) to (15), which makes the function $H(x, \nabla V^*, \mu^*(\hat{x}_j), w^*(x))$ not equal to zero during each time interval $t \in (\tau_j, \tau_{j+1})$.

Theorem 1 *Suppose that $V^*(x)$ is the solution of the HJB equation (12). For all $t \in [\tau_j, \tau_{j+1}), j \in \mathbb{N}$, the control policies are given by (11) and (15), respectively. If the triggering condition is defined as follows*

$$\|e_j(t)\|^2 > \frac{(1 - \beta^2)}{2\|r^T\|^2\|L\|^2}\lambda_{\min}(Q)\|x\|^2 - \frac{\eta^2\|m^T w^*(x)\|^2}{\|r^T\|^2\|L\|^2} \triangleq \|e_T\|^2 \tag{19}$$

where e_T denotes the threshold, $\lambda_{\min}(Q)$ is the minimal eigenvalue of Q, and $\beta \in (0, 1)$ is a designed sample frequency parameter. Then the solution $\mu^*(\hat{x}_j)$ to the optimal control problem is also a solution to the robust control problem. That is, the system (1) can be globally asymptotically stable for all admissible uncertainties $W(x)$ under $\mu^*(\hat{x}_j)$.

Remark 4 *The corresponding proof will be given in a future work. Note that the control input $\mu^*(\hat{x}_j)$ is based on event-triggered mechanism while the augmented control input $w^*(x)$ is based on time-triggered mechanism in this paper.*

4   Approximate Optimal Controller Design

In this section, an online event-triggered ADP algorithm with a single NN structure is proposed to approximate the solution of the event-triggered HJB equation.

4.1   Event-Triggered ADP Algorithm via Critic Network

In the event-triggered ADP algorithm, only a single critic network with a three-layer network structure is required to approximate the optimal value function. The optimal value function based on NN can be formulated as

$$V^*(x) = W_c^{\mathrm{T}}\phi(x) + \varepsilon \tag{20}$$

where $W_c \in \mathbb{R}^N$ is the critic NN ideal weights, $\phi(x) \in \mathbb{R}^N$ is the activation function vector, N is the number of hidden neurons, and $\varepsilon \in \mathbb{R}$ is the critic NN approximation error. Then, we can obtain

$$\nabla V^*(x) = \nabla \phi^{\mathrm{T}}(x) W_c + \nabla \varepsilon \tag{21}$$

Since the ideal weight matrix is unknown, the actual output of critic NN can be presented as

$$\hat{V}(x) = \hat{W}_c^{\mathrm{T}}\phi(x) \tag{22}$$

where $\hat{W}_c$ represents the estimation of the unknown weight matrix W_c.

Accordingly, the augmented control policy (11) and the event-triggered control policy (15) can be approximated by

$$\hat{w}(x) = -\frac{1}{2\eta^2} M^{-1} d^{\mathrm{T}}(x) \nabla \phi^{\mathrm{T}}(x) \hat{W}_c \tag{23}$$

$$\hat{\mu}(\hat{x}_j) = -\frac{1}{2} R^{-1} g^{\mathrm{T}}(\hat{x}_j) \nabla \phi^{\mathrm{T}}(\hat{x}_j) \hat{W}_c \tag{24}$$

Using the neural network expression (20), the event-triggered HJB equation (18) becomes

$$\begin{aligned}
&H(x, W_c, \mu(\hat{x}_j), w) \\
=&\|r^{\mathrm{T}}\|^2 g_M^2(x) + \eta^2 \|m^{\mathrm{T}}\|^2 W_M^2(x) + U\big(x, \mu(\hat{x}_j), w\big) \\
&+ W_c^{\mathrm{T}}\nabla\phi(x)\big(f(x) + g(x)\mu(\hat{x}_j) + d(x)w\big) \\
=&\varepsilon_{cH} + \big(r^{\mathrm{T}}(u(x) - \mu(\hat{x}_j))\big)^{\mathrm{T}}\big(r^{\mathrm{T}}(u(x) - \mu(\hat{x}_j))\big)
\end{aligned} \tag{25}$$

where

$$\varepsilon_{cH} = -(\nabla\varepsilon)^{\mathrm{T}}\big(f(x) + g(x)\mu(\hat{x}_j) + d(x)w\big)$$

denotes the residual error. For fixed N, the NN approximation errors ε and $\nabla\varepsilon$ are bounded locally [5]. That is, $\forall \nabla\varepsilon_{\max} > 0, \exists N(\nabla\varepsilon_{\max}) : \sup\|\nabla\varepsilon\| < \nabla\varepsilon_{\max}$. Then, the residual error is bounded locally under the Lipschitz assumption on the system dynamics. That is, there exists $\varepsilon_{cH\,\max} > 0$ such that $|\varepsilon_{cH}| \leq \varepsilon_{cH\,\max}$.

Using (22) with the estimated weight vector, the approximate event-triggered HJB equation is

$$\begin{aligned}
&H(x, \hat{W}_c, \mu(\hat{x}_j), w) \\
=&\|r^{\mathrm{T}}\|^2 g_M^2(x) + \eta^2 \|m^{\mathrm{T}}\|^2 W_M^2(x) + U\big(x, \mu(\hat{x}_j), w\big) \\
&+ \hat{W}_c^{\mathrm{T}}\nabla\phi(x)\,(f(x) + g(x)\mu(\hat{x}_j) + d(x)w) \\
\triangleq& e_c
\end{aligned} \tag{26}$$

where e_c is a residual equation error.

Define $\varepsilon_u = \left(r^{\mathrm{T}}(u(x) - \mu(\hat{x}_j))\right)^{\mathrm{T}} \left(r^{\mathrm{T}}(u(x) - \mu(\hat{x}_j))\right)$ as the event-triggered transformation error. Letting the weight estimation error of the critic NN be $\tilde{W}_c = W_c - \hat{W}_c$ and by combining (25) with (26), we have

$$e_c = -\tilde{W}_c^{\mathrm{T}} \nabla \phi(x) \left(f(x) + g(x)\mu(\hat{x}_j) + d(x)w\right) + \varepsilon_{cH} + \varepsilon_u \qquad (27)$$

Based on experience replay technique [15], it is desired to choose $\hat{W}_c$ to minimize the corresponding squared residual error

$$E = \frac{1}{2} \left(e_c^{\mathrm{T}} e_c + \sum_{k=1}^{p} e^{\mathrm{T}}(t_k) e(t_k)\right)$$

where $e(t_k) = U\left(x(t_k), \hat{\mu}(\hat{x}_i), \hat{w}(t_k)\right) + \hat{W}_c^{\mathrm{T}}(t)\sigma_k$, $\sigma_k = \nabla \phi(x(t_k))(f(x(t_k)) + g(x(t_k))\hat{\mu}(\hat{x}_i) + k(x(t_k))\hat{w}(t_k))$ are stored data at time $t_k \in [\tau_i, \tau_{i+1})$, $i \in \mathbb{N}$, and p is the number of stored samples.

PE-Like Condition: The recorded data matrix $M = [\sigma_1, ..., \sigma_p]$ contains as many as linearly independent elements as the number of the critic NN's hidden neurons, such that $\mathrm{rank}(M) = N$.

The weights of the critic NN are tuned using the standard steepest descent algorithm, which is given by

$$\begin{aligned}
\dot{\hat{W}}_c = &-\alpha_c \frac{\partial E}{\partial \hat{W}_c} \\
= &-\alpha_c \sigma \left(\|r^{\mathrm{T}}\|^2 g_M^2(x) + \eta^2 \|m^{\mathrm{T}}\|^2 W_M^2(x)\right. \\
&\left. + \sigma^{\mathrm{T}} \hat{W}_c + U\left(x, \hat{\mu}(\hat{x}_j), \hat{w}(t)\right)\right) \\
&-\alpha_c \sum_{k=1}^{p} \sigma_k \left(\|r^{\mathrm{T}}\|^2 g_M^2(x(t_k)) + \eta^2 \|m^{\mathrm{T}}\|^2 W_M^2(x(t_k))\right. \\
&\left. + \sigma_k^{\mathrm{T}} \hat{W}_c + U\left(x(t_k), \hat{\mu}(\hat{x}_i), \hat{w}(t_k)\right)\right)
\end{aligned} \qquad (28)$$

where $\sigma = \nabla \phi(x) \left(f(x) + g(x)\hat{\mu}(\hat{x}_j) + d(x)\hat{w}(t)\right)$, and α_c denotes the learning rate.

Combining (25), (27) and (28), we have

$$\begin{aligned}
\dot{\tilde{W}}_c = &-\alpha_c \sigma(\sigma^{\mathrm{T}} \tilde{W}_c - \varepsilon_{cH} - \varepsilon_u) \\
&-\alpha_c \sum_{k=1}^{p} \sigma_k \left(\sigma_k^{\mathrm{T}} \tilde{W}_c - \varepsilon_{cH}(t_k) - \varepsilon_u(t_k)\right)
\end{aligned} \qquad (29)$$

where $\varepsilon_{cH}(t_k)$ and $\varepsilon_u(t_k)$ denote the residual error and the event-triggered transformation error at $t = t_k$, respectively.

Note that the closed-loop sampled-data system behaves as an impulsive dynamical system with the flow dynamics and jump dynamics. Define the augmented state $\Psi = [x^{\mathrm{T}}, \hat{x}_j^{\mathrm{T}}, \tilde{W}_c^{\mathrm{T}}]^{\mathrm{T}}$. From (13), (14) and (29), the dynamics of the

impulsive system during the flow $t \in [\tau_j, \tau_{j+1}), j \in \mathbb{N}$ can be described by

$$\dot{\Psi} = \left[\; F(\Psi)^{\mathrm{T}} \quad 0 \quad G(\Psi)^{\mathrm{T}} \; \right]^{\mathrm{T}} \tag{30}$$

where the nonlinear functions

$$\begin{aligned}
F(\Psi) =& f(x) + g(x)\mu^*(\hat{x}_j) + k(x)w^*(x) \\
&+ \frac{1}{2}g(x)R^{-1}g^{\mathrm{T}}(\hat{x}_j)\left(\nabla\phi^{\mathrm{T}}(\hat{x}_j)\tilde{W}_c + \nabla\varepsilon \right) \\
&+ \frac{1}{2\eta^2}d(x)M^{-1}d^{\mathrm{T}}(x)\left(\nabla\phi^{\mathrm{T}}(x)\tilde{W}_c + \nabla\varepsilon \right),
\end{aligned}$$

$$\begin{aligned}
G(\Psi) =& -\alpha_c\sigma(\sigma^{\mathrm{T}}\tilde{W}_c - \varepsilon_{cH} - \varepsilon_u) \\
&- \alpha_c\sum_{k=1}^{p}\sigma_k\left(\sigma_k^{\mathrm{T}}\tilde{W}_c - \varepsilon_{cH}(t_k) - \varepsilon_u(t_k) \right).
\end{aligned}$$

The jump dynamics at the triggering instant $t = \tau_{j+1}$ can be given by

$$\Psi(t) = \Psi\left(t^-\right) + \left[\; 0 \quad (x - \hat{x}_j)^{\mathrm{T}} \quad 0 \; \right]^{\mathrm{T}} \tag{31}$$

where $\Psi\left(t^-\right) = \lim_{\varrho\to0}\Psi\left(t - \varrho\right)$, and 0s are null vectors with appropriate dimensions.

4.2   Stability Analysis

In this subsection, the main theorem will be presented to guarantee the weight estimation error of the critic NN to be UUB. Meanwhile, the stability of the impulsive dynamical system based on the event-triggered optimal control and the augmented control will be guaranteed with a novel adaptive triggering condition. First, we give the following assumption.

Assumption 4:

1. $g(x)$ and $d(x)$ are upper bounded by positive constants such that $\|g(x)\| \leq g_{\max}$ and $\|d(x)\| \leq d_{\max}$.
2. The critic NN activation function and its gradient are bounded, i.e., $\| \phi(x) \| \leq \phi_{\max}$ and $\| \nabla\phi(x) \| \leq \nabla\phi_{\max}$, with $\phi_{\max}, \nabla\phi_{\max}$ being positive constants.
3. The critic NN ideal weight matrix is bounded by a positive constant, that is $\|W_c\| \leq W_{\max}$.

Theorem 2 *Suppose that Assumptions 1-4 hold. The tuning law for the CT critic neural network is given by (28). Then the closed-loop sampled-data system (14) is asymptotically stable and the critic weight estimation error is guaranteed to be UUB if the adaptive triggering condition*

$$\begin{aligned}
\|e_j(t)\|^2 >& \frac{1}{L^2\|r^T\|^2}\left((1 - \beta^2)\lambda_{\min}(Q)\|x\|^2 \right. \\
&\left. + \|r^T\hat{\mu}(\hat{x}_j)\|^2 - \eta^2\|m^T\hat{w}(x)\|^2 \right) \\
\triangleq& \left\| \hat{e}_T\left(x, \hat{x}_j, \hat{W}_c \right) \right\|^2
\end{aligned} \tag{32}$$

holds and the following inequality

$$\|\tilde{W}_c\| > \sqrt{\frac{\alpha_c^2 \eta^2 \lambda_{\min}(M) \sum_{k=1}^{p+1} (\varepsilon_{cH\max} + \varepsilon_{u\max})^2 + \Gamma}{4\eta^2 (\alpha_c - 1) \lambda_{\min}(H) \underline{\theta}(M)}} \tag{33}$$
$$\overset{\Delta}{=} \Pi_{\max}$$

is satisfied with $\alpha_c > 1$, where $\Gamma = 2 \left(d_{\max} \nabla \phi_{\max} \left(W_{\max} + \nabla \varepsilon_{\max}\right)\right)^2$.

Remark 5 *The proof of Theorem 2 will be presented in a future paper. Note that the triggering condition (32) is utilized to approximate the optimal control policy pair $[\hat{\mu}^{*T}(\hat{x}_j), \hat{w}^{*T}(x)]^T$ for the auxiliary sampled-date system while the triggering condition (19) in Theorem 1 is utilized to guarantee the robust stabilization of the original uncertain system with the obtained optimal control policy $\hat{\mu}^*(\hat{x}_j)$.*

5   Simulation

The example is considered as follows [2]:

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u + \begin{bmatrix} 0.2 \\ 0 \end{bmatrix} W(x) \tag{34}$$

where $W(x) = \lambda_1 x_1 \cos\left(\frac{1}{x_2 + \lambda_2}\right) + \lambda_3 x_2 \sin(\lambda_4 x_1 x_2)$, and $\lambda_1, \lambda_2, \lambda_3, \lambda_4$ are the unknown parameters. The last term reflects the unmatched uncertainty in the system. Assume that $\lambda_1 \in [-1, 1]$, $\lambda_2 \in [-100, 100]$, $\lambda_3 \in [-0.2, 1]$, and $\lambda_4 \in [-100, 0]$.

Clearly,

$$g^+(x) = \left(g^T(x)g(x)\right)^{-1} g^T(x) = g^T(x) = [0, 1],$$

$$\left(I - g(x)g^+(x)\right) k(x) = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 0.2 \\ 0 \end{bmatrix} = \begin{bmatrix} 0.2 \\ 0 \end{bmatrix},$$

$$\|W(x)\|^2 \leq x_1^2 + x_2^2 \overset{\Delta}{=} W_M^2(x),$$

$$2\|g^+(x)k(x)W(x)\|^2 = 0 \overset{\Delta}{=} g_M^2(x).$$

Set Q, R, r and m are the identity matrices with appropriate dimensions. We experimentally choose $\eta = 1$, $p = 10$, $\beta = 0.1$ and $L = 3$.

During the implementation process of the event-triggered ADP method, we choose a three-layer feedforward NN with structure 2-8-1 as the critic network. The critic NN activation function is chosen as $\phi(x) = [x_1^2 \ x_1 x_2 \ x_2^2 \ x_1^4 \ x_1^3 x_2 \ x_1^2 x_2^2 \ x_1 x_2^3 \ x_2^4]^T$. The initial state is selected as $x_0 = [1, -1]^T$, the learning rate is $\alpha_c = 0.1$, and the sampling time is chosen as 0.05s. The trajectories of the critic parameters are shown in Fig. 1. At the end of learning process, the parameters converge to $\hat{W}_c = [1.8594 \ 0.8845 \ 1.1560 \ 1.9860 \ 0.9272 \ 0.5403 \ 0.4344 \ 0.3737]^T$. From Fig. 2,

one can get the event-triggered error $e_j(t)$ and the threshold e_T converge to zero as the states converge to zero. In addition, the event-triggered error is forced to zero when the triggering condition is satisfied, that is the system states are sampled at the triggering instants. The sampling period during the event-triggered learning process for the control policy is provided in Fig. 3. Furthermore, the lower bound on the inter-sample times is found to be 0.15s. In particular, the event-triggered controller uses 47 samples of the state while the time-triggered controller uses 1000 samples, which means fewer transmissions are needed between the plant and the controller due to event-triggered sampling. That will reduce the number of controller updates during the learning process.
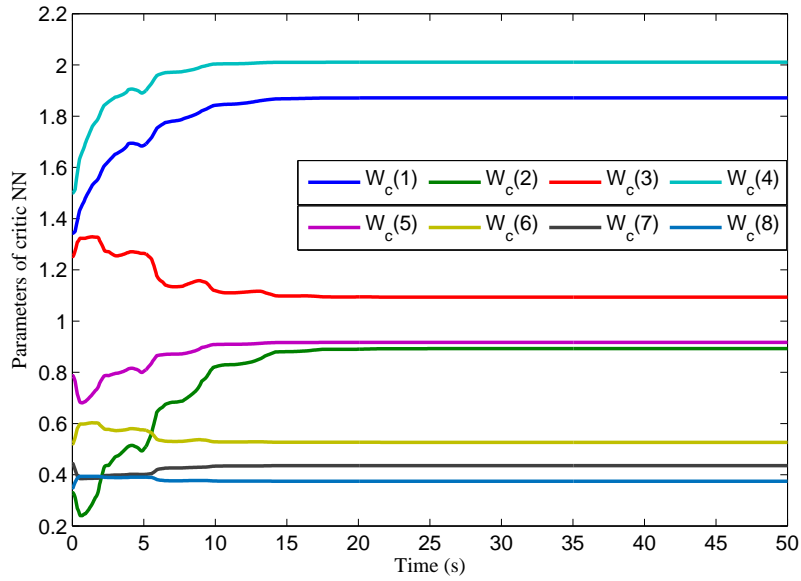
Fig. 1. Convergence of the critic parameters

Based on the converged weights $\hat{W}_c$, we can obtain the near-optimal control laws as

$$\begin{bmatrix} \hat{\mu}^*(\hat{x}_j) \\ \hat{w}^*(x) \end{bmatrix} = \begin{bmatrix} -\frac{1}{2}[0\ 1]\nabla\phi^{\mathrm{T}}(\hat{x}_j)\hat{W}_c \\ -\frac{1}{2}[1\ 0]\nabla\phi^{\mathrm{T}}(x)\hat{W}_c \end{bmatrix} \tag{35}$$

From [2], the optimal control laws are given as

$$\begin{bmatrix} u^*(x) \\ w^*(x) \end{bmatrix} = \begin{bmatrix} -1.2906x_1 - 2.1247x_2 \\ -0.5783x_1 - 0.2581x_2 \end{bmatrix} \tag{36}$$

Now, we apply the near-optimal control laws (35) with the triggering condition (19) and the optimal control laws (36) for the uncertain nonlinear system with
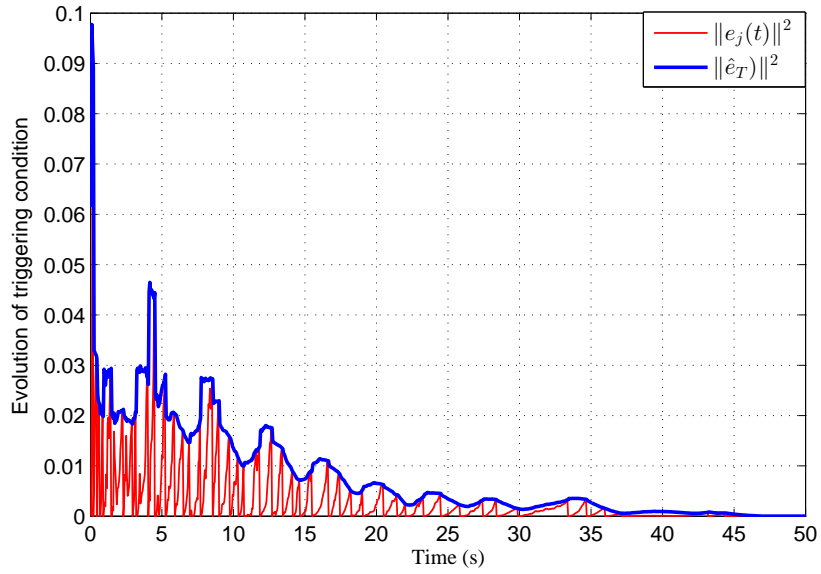
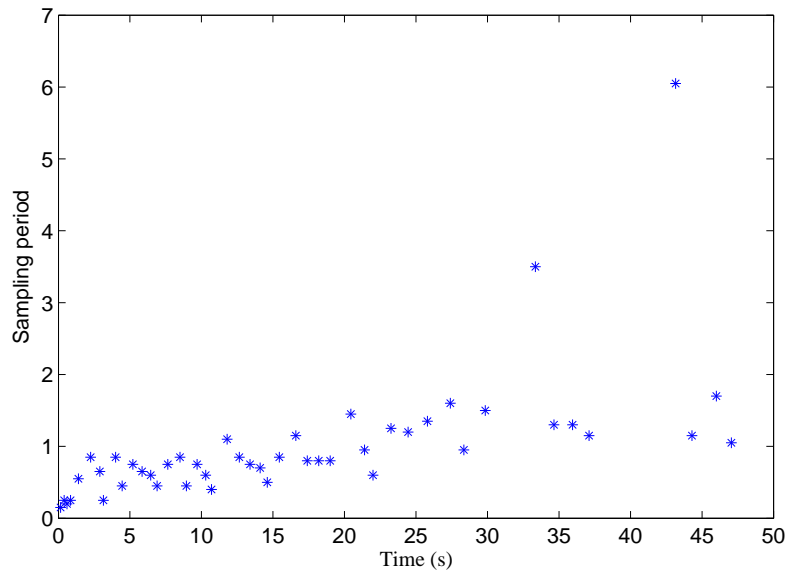**Fig. 2.** Response of $\|e_j(t)\|^2$ and $\|\hat{e}_T\|^2$

Fig. 3. Triggering instants during the learning process of the control input

$\lambda_1 = -1, \lambda_2 = -100, \lambda_3 = 0, \lambda_4 = -100$. Set the initial state be $x_0 = [1, -1]^\mathrm{T}$, and the sampling time be $0.05s$. The simulation results are given in Fig. 4.
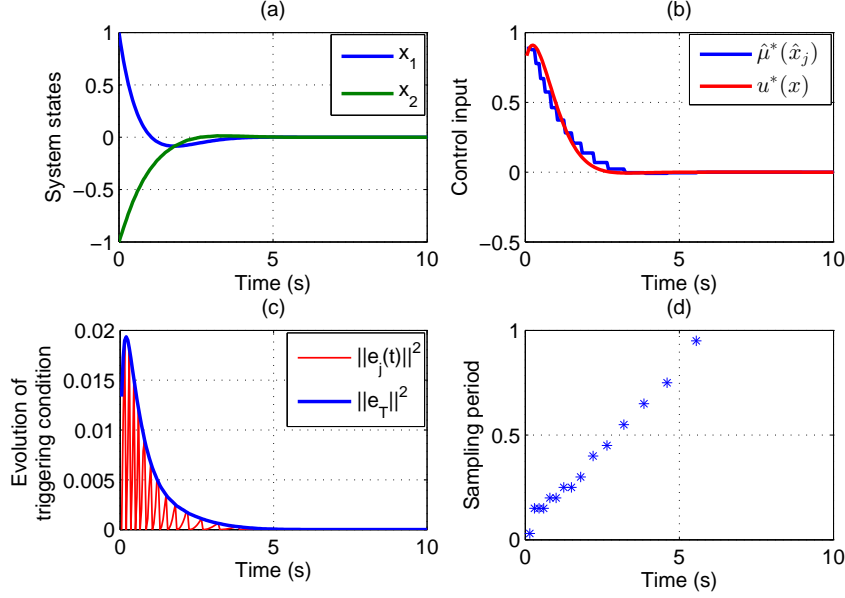

Fig. 4. Case 1: (a) State trajectory. (b) Near-optimal and optimal control inputs. (c) Response of $\|e_T\|^2$ and $\|e(t)\|^2$. (d) Sampling period.

We can observe the near-optimal controller is robust for the uncertain nonlinear system and adjusted with events.

6   Conclusion

In this paper, we propose an event-triggered ADP algorithm to solve the robust control problem of uncertain nonlinear systems. The robust control problem is described as an optimal control problem with an modified cost function. For implementation purpose, a critic NN is constructed to approximate the optimal value function. Finally, simulation results are given to demonstrate the effective of the event-triggered ADP scheme.

14 Qichao Zhang[1,2], Dongbin Zhao*[1,2], and Ding Wang[1,2]

7 The References Section

References

1. Lin, F., Brandt, R.D., Sun, J.: Robust control of nonlinear systems: Compensating for uncertainty. International Journal of Control, 56(6), 1453–1459 (1992)
2. Lin, F.: An optimal control approach to robust control design. International journal of control, 73(3), 177–186 (2000)
3. Lewis, F.L., Liu, D.: Reinforcement learning and approximate dynamic programming for feedback control. John Wiley & Sons, New Jersey (2013)
4. Werbos, P.J.: Advanced forecasting methods for global crisis warning and models of intelligence. General Systems Yearbook, 22(12), 25–38 (1977)
5. Abu-Khalaf, M., Lewis, F.L.: Nearly optimal control laws for nonlinear systems with saturating actuators using a neural network HJB approach. Automatica, 41(5), 779–791 (2005)
6. Jiang, Y., Jiang, Z.P.: Robust adaptive dynamic programming. Reinforcement Learning and Approximate Dynamic Programming for Feedback Control, 281–302 (2012)
7. Liu, D., Wang, D., Wang, F.Y., et al.: Neural-network-based online HJB solution for optimal robust guaranteed cost control of continuous-time uncertain nonlinear systems. IEEE Transactions on Cybernetics, 44(12), 2834–2847 (2014)
8. Wang, D., Liu, D., Li, H., et al.: Neural-network-based robust optimal control design for a class of uncertain nonlinear systems via adaptive dynamic programming. Information Sciences, 282: 167–179 (2014)
9. Zhao, D., Zhu, Y.: MECA near-optimal online reinforcement learning algorithm for continuous deterministic systems. IEEE Transactions on Neural Networks and Learning Systems, 26(2), 346–356 (2015)
10. Zhao, D., Zhang, Q., Wang, D., et al.: Experience replay for optimal control of nonzero-sum game systems with unknown dynamics. IEEE Transactions on Cybernetics, 46(3), 854–865 (2016)
11. Zhang, Q., Zhao, D., Zhu, Y.: Event-triggered H_∞ control for continuous-time nonlinear system via concurrent learning. IEEE Transactions on Systems, Man, and Cybernetics: Systems, DOI10.1109/TSMC.2016.2531680 (2016)
12. Sahoo, A., Xu, H., Jagannathan, S.: Near optimal event-triggered control of nonlinear discrete-time systems using neurodynamic programming. IEEE Transactions on Neural Networks and Learning Systems, 27(9), 1801–1815 (2016)
13. Vamvoudakis, K.G.: Event-triggered optimal adaptive control algorithm for continuous-time nonlinear systems. IEEE/CAA Journal of Automatica Sinica, 1(3), 282–293 (2014)
14. Zhong, X., Ni, Z., He, H., et al.: Event-triggered reinforcement learning approach for unknown nonlinear continuous-time system. In: 2014 IEEE International Joint Conference on Neural Networks, pp.3677–3684. IEEE Press, Beijing (2014)
15. Modares, H., Lewis, F.L., Naghibi-Sistani, M.B.: Adaptive optimal control of unknown constrained-input systems using policy iteration and neural networks. IEEE Transactions on Neural Networks and Learning Systems, 24(10), 1513–1525 (2013)