Nonlinear Dyn (2017) 90:2023–2035
DOI 10.1007/s11071-017-3778-5

ORIGINAL PAPER

Robust adaptive critic control design with network-based event-triggered formulation

Chaoxu Mu · Ding Wang · Changyin Sun · Qun Zong

Received: 11 February 2017 / Accepted: 29 August 2017 / Published online: 6 September 2017

Abstract In this paper, by incorporating the network-based event-triggered formulation, the robust adaptive critic control design for a class of nonlinear continuous-time systems is investigated to fulfill disturbance rejection. First, the designed problem with output information is formulated as a two-player zero-sum differential game and the adaptive critic mechanism is employed toward the event-based minimax optimization involving a suitable triggering condition. Then, the event-based optimal control law and the time-based worst-case disturbance law are learned by training the critic neural network. Besides, the closed-loop system is constructed with stability proof of the critic error dynamics and the sampled-data plant. The theoretical analysis has demonstrated that the infamous Zeno behavior of the proposed event-based adaptive critic design has been avoided. Finally, the developed method is applied toward the robot arm plant, as a mechanical component of the complex robot system, so as to substantiate the performance of disturbance rejection.

C. Mu · Q. Zong
Tianjin Key Laboratory of Process Measurement and Control, School of Electrical and Information Engineering, Tianjin University, Tianjin 300072, China

D. Wang (✉)
The State Key Laboratory of Management and Control for Complex Systems, Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China
e-mail: ding.wang@ia.ac.cn

C. Sun
School of Automation, Southeast University, Nanjing 210096, China

Keywords Adaptive critic control · Disturbance rejection · Event-triggered mechanism · Network-based systems · Robust H_∞ control

1 Introduction

With the framework of network-based systems, the control loops are often closed through a communication medium. As a hot topic of system and control community, it is significant to perform systematically theoretical researches and meaningfully industrial applications for network-based control design. The growing demands in reducing the computational load of networked control systems, or more extensively, the emerging cyber-physical systems bring a great attention to develop the mechanism of event-triggering control [1–4]. Dolk et al. [1] proposed the popular framework for output-based dynamic event-triggered control design under denial-of-service attacks. Wu et al. [3] dealt with the event-based optimal control of heating, ventilation and air-conditioning systems of buildings for the purpose of energy saving. Within these general event-based control approaches, the actuators are only updated under certain triggering conditions such that both control performance and system stability can be guaranteed toward the target objects.

Robustness is an important criterion to evaluate the performance of the designed controller with respect to uncertain disturbances and parameters of the controlled plant. In particular, the H_∞ method usually concerns to construct a control law for the worse-case uncertain plant. From the point of view of the minimax optimization, a H_∞ control problem can be considered as the two-player zero-sum differential game, where a controller is obtained that minimizes the cost function in the worst-case disturbance. In this case, it requires to get the Nash equilibrium solution via the Hamilton–Jacobi–Isaacs equation. As we all know, it is obviously difficult to get the analytic Nash equilibrium solution for nonlinear systems. Fortunately, the methodology of adaptive/approximate dynamic programming is developed to effectively solve the class of optimal control problems forward-in-time [5–7], with neural networks [8–11] as well as some new developed function approximation architectures like incremental support vector machines [12], and so on. The adaptive/approximate dynamic programming approach has made great progress in the aspects of optimal control for discrete-time nonlinear systems[13–17], continuous-time nonlinear systems [18–21] and some related applications [22–24]. Moreover, the problems of nonlinear H_∞ control and the nonzero-sum game have been revisited and studied with the approach of adaptive/approximate dynamic programming in [25–29].

Adaptive critic control, as one method of adaptive dynamic programming-based control, comes from the literature [30] where Prokhorov et al. proposed the adaptive critic design with neural networks. Since then, adaptive critic control has been developed as an important method of approximate optimal control approaches. In order to improve the robustness of adaptive critic control, the robust adaptive critic control methodology was proposed in references [31,32], which has recently achieved great development in [33–37]. However, these existing research results are obtained by the traditional design manner of time-based control, which would cause that actuators are frequently adjusted and thus energy consumption is enormous. Therefore, the time/event control structure has become an outlet to fulfill the event-based design and enhance the control efficiency [29,36,38–40]. In recent few years, the event-based adaptive critic design method has been developed as a new channel for the adaptive optimal stabilization of nonlinear systems [36,40–42]. With the new time/event control mechanism, the developed controller is updated once an event is triggered, which results in reducing the computational cost. Consider that most existing work conducted for the optimal regulation without involving output information, such as [18,19,21,26,40,41,43], motivates this extension work to nonlinear event-based zero-sum differential game problem with output information.

In this paper, the event-based robust H_∞ control with output information is investigated under the framework of adaptive critic designs. The contributions of this paper are listed as follows. For one thing, the framework of the event-based adaptive critic control with output information is established to study the nonlinear H_∞ feedback control. The two-player zero-sum differential game problem with output information is formulated, and the event-based minimax optimization involving a suitable triggering condition is designed within the event-based adaptive critic control framework. For another thing, by involving output information, both the event-based optimal control law and the time-based worst-case disturbance law are derived with stability proof, and the Zeno behavior in the event-based control is effectively avoided. This improves the results of traditional adaptive critic design such as [18,19,21,26] and event-based control design such as [40,41,43]. The rest of this paper is organized as follows: In Sect. 2, a succinct transformation of nonlinear H_∞ control with output information is described. The event-based adaptive critic design for the nonlinear H_∞ feedback control problem is intensively investigated in Sect. 3 with the analysis of closed-loop system stability and Zeno behavior exclusion. The application of a robot arm plant is provided in Sect. 4, and some concluding remarks are finally drawn in Sect. 5.

For the effective presentation, these notations are defined and used in the following sections. $\mathbb{R}$, $\mathbb{R}^n$ and $\mathbb{R}^{n\times m}$ define the set of all real numbers, the Euclidean space of all n-dimensional real vectors and the space of all $n \times m$ real matrices, respectively. $\mathbb{N} = \{0, 1, 2, \ldots\}$ defines the set of all nonnegative integers. I_n is the identity matrix in $\mathbb{R}^{n\times n}$. $\lambda_{\max}(\cdot)$ and $\lambda_{\min}(\cdot)$ represent the maximal and minimal eigenvalues of a matrix, while $\text{diag}\{\xi_1, \xi_2, \ldots, \xi_n\}$ expresses the $n\times n$ diagonal matrix with elements of $\xi_1, \xi_2, \ldots, \xi_n$. $\|\cdot\|$ denotes the 2-norm for a vector and the induced-norm for a matrix. Define Ω as a compact subset of $\mathbb{R}^n$, and $\mathscr{A}(\Omega)$ is the admissible control policy set on Ω. A superscript "T" and

$\nabla(\cdot) \triangleq \partial(\cdot)/\partial x$ mean the transpose operation and the gradient operation, respectively.

2 Problem description and preliminaries

The following class of affine nonlinear continuous-time systems is considered in this paper with external perturbations:

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

$$y(t) = Cx(t). \tag{1b}$$

In (1), $x(t)$ denotes the state vector belonging to $\Omega \subset \mathbb{R}^n$, $u(t) \in \mathbb{R}^m$ is the control input, $\nu(t) \in \mathbb{R}^q$ is the perturbation belonging to $L_2[0, \infty)$, $y(t) \in \mathbb{R}^p$ is the output vector, and $C \in \mathbb{R}^{p\times n}$ is a constant output matrix. $f(x)$, $g(x)$ and $h(x)$ are differentiable with $f(0) = 0$. $x(0)$ is the initial state vector, recorded as x_0, and $x = 0$ is the equilibrium point of the system.

Assumption 1 The nonlinear system (1) is controllable. The system function $f(x)$ defined on Ω is Lipschitz continuous and contains the origin.

With this assumption, considering the nonlinear H_∞ control design of system (1), a feedback control law $u(x)$ is expected to make the closed-loop system (1) asymptotically stable with a L_2-gain no larger than ι, which is

$$\int_0^\infty \big(y^{\mathsf{T}}(\tau)Py(\tau)+u^{\mathsf{T}}(\tau)u(\tau)\big)\mathrm{d}\tau \le \iota^2 \int_0^\infty \|\nu(\tau)\|^2\mathrm{d}\tau, \tag{2}$$

where P is a positive semidefinite matrix with appropriate dimension. Recalling (1b), it is obvious that (2) can be rewritten as

$$\int_0^\infty \big(x^{\mathsf{T}}(\tau)Qx(\tau)+u^{\mathsf{T}}(\tau)u(\tau)\big)\mathrm{d}\tau \le \iota^2 \int_0^\infty \|\nu(\tau)\|^2\mathrm{d}\tau, \tag{3}$$

where $Q = C^{\mathsf{T}}PC$ is nonnegative definite. If the closed-loop expression of system (1) satisfies the condition (3), then it has a L_2-gain no larger than ι. It is known to all that the nonlinear H_∞ control can be translated into the problem of two-player zero-sum differential game, where the minimizing player is considered as the control and the maximizing player is regarded as the disturbance [25,26]. Therefore, the solution of nonlinear H_∞ control is defined as a control pair with the form (u^*, ν^*). u^* and ν^* are the optimal control and the worst-case disturbance, respectively. Define the utility function $U\big(x(\tau), u(\tau), \nu(\tau)\big)$ as

$$U\big(x(\tau), u(\tau), \nu(\tau)\big) = x^{\mathsf{T}}(\tau)Qx(\tau) + u^{\mathsf{T}}(\tau)u(\tau) - \iota^2\nu^{\mathsf{T}}(\tau)\nu(\tau),$$

and the corresponding cost function is

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

where the cost function $J(x, u, \nu)$ can be simplified as $J(x)$ in the following text. The initial cost function at $t = 0$ is recorded as $J(x_0)$. In the two-player zero-sum game problem, the feedback control pair (u^*, ν^*) satisfies the Nash condition, i.e.,

$$J^*(x_0) = \min_u \max_\nu J(x_0, u, \nu) = \max_\nu \min_u J(x_0, u, \nu).$$

Considering that an admissible control policy $u \in \mathscr{A}(\Omega)$ is used, if the cost function in (4) is differentiable, then it derives the following nonlinear Lyapunov equation

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

with an initial condition $J(0) = 0$. Correspondingly, the Hamiltonian function of system (1) is defined as

$$H\big(x, u, \nu, \nabla J(x)\big) = U(x, u, \nu) + \big(\nabla J(x)\big)^{\mathsf{T}}\big(f(x) + g(x)u + h(x)\nu\big).$$

If the Bellman's optimality principle is used, then the optimal cost $J^*(x)$ can make sure that the Hamilton–Jacobi–Isaacs equation

$$\min_u \max_\nu H\big(x, u, \nu, \nabla J^*(x)\big) = 0$$

holds. That is to say, the control pair (u^*, ν^*) can be obtained by the following partial differential equations

$$\frac{\partial H\big(x, u, \nu, \nabla J^*(x)\big)}{\partial u} = 0,$$
$$\frac{\partial H\big(x, u, \nu, \nabla J^*(x)\big)}{\partial \nu} = 0.$$

Therefore, the optimal control and the worst-case disturbance are calculated by

$$u^*(x) = -\frac{1}{2} g^\mathsf{T}(x) \nabla J^*(x), \tag{5a}$$
$$\nu^*(x) = \frac{1}{2\iota^2} h^\mathsf{T}(x) \nabla J^*(x). \tag{5b}$$

By using (5), the Hamilton–Jacobi–Isaacs equation turns to the following expression

$$0 = x^\mathsf{T} Q x + \big(\nabla J^*(x)\big)^\mathsf{T} f(x) - \frac{1}{4}\big(\nabla J^*(x)\big)^\mathsf{T} g(x) g^\mathsf{T}(x) \times \nabla J^*(x) + \frac{1}{4\iota^2}\big(\nabla J^*(x)\big)^\mathsf{T} h(x) h^\mathsf{T}(x) \nabla J^*(x) \tag{6}$$

with $J^*(0) = 0$. It should be noted that (6) is the classical time-based Hamilton–Jacobi–Isaacs equation. An approximate solution is pursued to substitute for the analytic solution. The adaptive critic control-based method is taken as an effective approach to handle the problem.

3 Event-based robust adaptive critic control design and implementation

3.1 Event-based control design with Zeno behavior exclusion

In industrial practice, a sampling component is often incorporated into a networked system. With the event-triggering control method, a monotonically increasing sequence is usually defined as the triggering instants, i.e., $\{s_j\}_{j=0}^{\infty}$, where s_j expresses the jth consecutive sampling instant with $j \in \mathbb{N}$. The sampled state vector is denoted as $x(s_j) \triangleq \hat{x}_j$ for all $t \in [s_j, s_{j+1})$. The event-triggered error defines the gap between current and sampled states, which is represented as $\sigma_j(t) = \hat{x}_j - x(t),\ \forall t \in [s_j, s_{j+1})$.

In the event-based control, the triggering condition decides the triggering instants. That is to say, at the triggering instant $t = s_j$, when the triggering condition is activated, the system is sampled such that the event-triggered error $\sigma_j(t)$ is reset as zero. The control law $u(x(s_j)) = u(\hat{x}_j) \triangleq \mu(\hat{x}_j)$ is accordingly updated. By introducing a zero-order holder, the control sequence $\{\mu(\hat{x}_j)\}_{j=0}^{\infty}$ can be turned into a continuous-time signal in the form of a piecewise constant function with a constant value $\mu(\hat{x}_j)$ at the time interval $[s_j, s_{j+1})$, $j \in \mathbb{N}$. When the event-triggering mechanism is employed, the feedback control law in (5a) becomes

$$\mu^*(\hat{x}_j) = -\frac{1}{2} g^\mathsf{T}(\hat{x}_j) \nabla J^*(\hat{x}_j), \tag{7}$$

where $\nabla J^*(\hat{x}_j) = \big(\partial J^*(x)/\partial x\big)|_{x=\hat{x}_j}$. The disturbance law is unchanged during the time/event structure transformation. Additionally, we make the following assumptions which are reasonable and conventional in the event-based design.

Assumption 2 (*cf.* [41]) The control law $u(x)$ is Lipschitz continuous with regard to the event-triggered error $\sigma_j(t)$, which is formulated as $\|u(x(t)) - u(\hat{x}_j)\| \le M_u \|\sigma_j(t)\|$, where M_u is a positive constant.

Assumption 3 The control function matrix $g(x)$ is Lipschitz continuous associated with the event-triggered error $\sigma_j(t)$ and is also upper-bounded, which means $\|g(x) - g(\hat{x}_j)\| \le M_g \|\sigma_j(t)\|$ and $\|g(x)\| \le B_g$, where M_g and B_g are positive constants. The disturbance matrix $h(x)$ is bounded by a positive constant B_h, which is expressed as $\|h(x)\| \le B_h$.

The following theorem is provided to design a triggering condition.

Theorem 1 *Considering the nonlinear system* (1) *and its related cost function* (4), *for all* $t \in [s_j, s_{j+1})$ *with* $j \in \mathbb{N}$, *if the disturbance law and the event-based control law are given by* (5b) *and* (7), *respectively, and the triggering condition is given as*

$$\|\sigma_j(t)\|^2 \le \sigma_T = \frac{x^\mathsf{T} Q x + \|\mu^*(\hat{x}_j)\|^2 - \iota^2 \|\nu^*(x)\|^2}{M_u^2}, \tag{8}$$

where σ_T *is the threshold of the triggering condition, then the closed-loop system* (1) *is asymptotically stable.*

Proof Select $L_1(t) = J^*(x(t))$ as the Lyapunov function candidate. Using (5b) and (7), we take the time derivative of $L_1(t)$ along the trajectory of system (1a) to compute $\dot{L}_1(t) = \mathrm{d}J^*(x(t))/\mathrm{d}t$, which derives

$$\dot{L}_1(t) = \big(\nabla J^*(x)\big)^{\mathsf{T}}\big(f(x)+g(x)\mu^*(\hat{x}_j)+h(x)\nu^*(x)\big).$$

Note that formula (5) implies that

$$\big(\nabla J^*(x)\big)^{\mathsf{T}} g(x) = -2u^{*\mathsf{T}}(x), \tag{9a}$$
$$\big(\nabla J^*(x)\big)^{\mathsf{T}} h(x) = 2\iota^2\nu^{*\mathsf{T}}(x). \tag{9b}$$

Besides, Eq. (6) reveals

$$\big(\nabla J^*(x)\big)^{\mathsf{T}} f(x) = -x^{\mathsf{T}}Qx+u^{*\mathsf{T}}(x)u^*-\iota^2\nu^{*\mathsf{T}}(x)\nu^*(x). \tag{10}$$

By using (9) and (10), it can derive

$$\begin{aligned}\dot{L}_1(t) =& -x^{\mathsf{T}}Qx + u^{*\mathsf{T}}(x)u^* - 2u^{*\mathsf{T}}(x)\mu^*(\hat{x}_j)\\ &+ \iota^2\nu^{*\mathsf{T}}(x)\nu^*(x).\end{aligned} \tag{11}$$

By introducing Assumption 2, $\dot{L}_1(t)$ can be obtained as

$$\begin{aligned}\dot{L}_1(t) =& -x^{\mathsf{T}}Qx+\big(u^*(x)-\mu^*(\hat{x}_j)\big)^{\mathsf{T}}\big(u^*(x)-\mu^*(\hat{x}_j)\big)\\ &-\mu^{*\mathsf{T}}(\hat{x}_j)\mu^*(\hat{x}_j)+\iota^2\nu^{*\mathsf{T}}(x)\nu^*(x)\\ \le& -x^{\mathsf{T}}Qx + M_u^2\|\sigma_j(t)\|^2\\ &-\|\mu^*(\hat{x}_j)\|^2+\iota^2\|\nu^*(x)\|^2.\end{aligned}$$

It is obvious that $\dot{L}_1(t) < 0$ can be obtained for any $x \neq 0$ if the triggering condition (8) holds, which ends the proof. □

For the proposed network-based event-triggered H_∞ control problem, the jth inter-sample time is $s_{j+1} - s_j$. Denote the minimal inter-sample time as

$$\Delta s_{\min} = \min_{j\in\mathbb{N}}\{s_{j+1} - s_j\},$$

which might be zero and thus lead to the accumulation of the event times, i.e., the infamous Zeno behavior.

By using Assumptions 1 and 3, and considering the fact that the optimal control function and the worst-case disturbance function are upper-bounded, it can acquire two positive constants κ_1 and κ_2 such that

$$\|\dot{x}\| = \|f(x)+g(x)\mu^*+h(x)\nu^*\| \le \kappa_1\|x\|+\kappa_2 \tag{12}$$

holds, where κ_2 is a bounded term with respect to the control matrix, the optimal control, the disturbance matrix and the worst-case disturbance. Take the derivative of the triggering error $\sigma_j(t)$ and then yield $\dot{\sigma}_j(t) = -\dot{x}$ for $t \in [s_j, s_{j+1})$. Based on (12), it can be further found that

$$\begin{aligned}\|\dot{\sigma}_j\| &\le \kappa_1\|\hat{x}_j - \sigma_j\| + \kappa_2\\ &\le \kappa_1\|\sigma_j\| + \kappa_1\|\hat{x}_j\| + \kappa_2, \forall t \in [s_j, s_{j+1}).\end{aligned} \tag{13}$$

By using the initial condition $\sigma_j(s_j) = \hat{x}_j - x(s_j) = 0$ and the comparison lemma (seeing [44]), the following inequality can be derived based on the solution of (13), which is

$$\|\sigma_j\| \le \frac{\kappa_1\|\hat{x}_j\|+\kappa_2}{\kappa_1}(\sigma^{\kappa_1(t-s_j)}-1) \tag{14}$$

for any $t \in [s_j, s_{j+1})$. According to (14), we obtain that the jth inter-sample time satisfies

$$s_{j+1} - s_j \ge \frac{1}{\kappa_1}\ln(1+\bar{\kappa}_j) > 0, \tag{15}$$

where the term $\bar{\kappa}_j = \kappa_1\bar{\sigma}_T/(\kappa_1\|\hat{x}_j\|+\kappa_2)$ is positive with $\bar{\sigma}_T = \|\sigma_j(s_{j+1})\|$ and $\sigma_j(s_{j+1}) = \hat{x}_j - x(s_{j+1})$. The minimum of $\bar{\kappa}_j$ with regard to all $t \in [s_j, s_{j+1})$, $j \in \mathbb{N}$, is defined as $\kappa_{\min} = \min_{j\in\mathbb{N}}\bar{\kappa}_j > 0$. By minimizing both sides of (15), we can conclude the following remark.

Remark 1 Considering the nonlinear system (1) used the disturbance law (5b) and the event-based control law (7), the minimal inter-sample time $\Delta s_{\min}$ determined by (8) is lower-bounded such that

$$\Delta s_{\min} \ge \frac{1}{\kappa_1}\ln(1+\kappa_{\min}) > 0, \tag{16}$$

where κ_1 and $\kappa_{\min}$ are positive constants. Hence, the Zeno behavior in this event-based control design is avoided.

3.2 Neural network implementation with stability analysis

The adaptive critic control design with neural networks is a practical approach to obtain the approximate optimal control solution for nonlinear system control problems [5,18,21,27,40]. In the neural network implementation, l_c is denoted as the neuron number of the hidden layer. By adopting the universal approximation property of neural networks, the cost function $J(x)$ is reconstructed by a single-hidden-layer neural network as

$$J(x) = \omega_c^{\mathsf{T}} \varphi_c(x) + \epsilon_c(x),$$

where $\omega_c \in \mathbb{R}^{l_c}$ is the desired weight vector, $\varphi_c(x) \in \mathbb{R}^{l_c}$ denotes the activation function of the neural network, and $\epsilon_c(x) \in \mathbb{R}$ is the reconstruction error. The gradient of $J(x)$ is expressed as

$$\nabla J(x) = \big(\nabla \varphi_c(x)\big)^{\mathsf{T}} \omega_c + \nabla \epsilon_c(x).$$

It is obvious that the desired weight vector ω_c is unknown; thus, the critic neural network with an estimated weight vector $\hat{\omega}_c(t)$ is used to construct the cost function, which is

$$\hat{J}(x) = \hat{\omega}_c^{\mathsf{T}}(t) \varphi_c(x).$$

Similarly, the gradient of the estimated cost function $\hat{J}(x)$ can be formulated as

$$\nabla \hat{J}(x) = \big(\nabla \varphi_c(x)\big)^{\mathsf{T}} \hat{\omega}_c(t).$$

Therefore, the event-based optimal control and the time-based worst-case disturbance are formulated as

$$\mu(\hat{x}_j) = -\frac{1}{2} g^{\mathsf{T}}(\hat{x}_j)\Big(\big(\nabla \varphi_c(\hat{x}_j)\big)^{\mathsf{T}} \omega_c + \nabla \epsilon_c(\hat{x}_j)\Big),$$
$$\nu(x) = \frac{1}{2\iota^2} h^{\mathsf{T}}(x)\Big(\big(\nabla \varphi_c(x)\big)^{\mathsf{T}} \omega_c + \nabla \epsilon_c(x)\Big).$$

By introducing the critic neural network, the approximate values of the above control pair are

$$\hat{\mu}(\hat{x}_j) = -\frac{1}{2} g^{\mathsf{T}}(\hat{x}_j)\big(\nabla \varphi_c(\hat{x}_j)\big)^{\mathsf{T}} \hat{\omega}_c(t), \tag{17a}$$
$$\hat{\nu}(x) = \frac{1}{2\iota^2} h^{\mathsf{T}}(x)\big(\nabla \varphi_c(x)\big)^{\mathsf{T}} \hat{\omega}_c(t). \tag{17b}$$

In the sequel, we apply the neural network expression to the Hamiltonian function and derive that

$$\begin{aligned} H&\big(x, \mu(\hat{x}_j), \nu(x), \omega_c\big) \\ &= U\big(x, \mu(\hat{x}_j), \nu(x)\big) + \omega_c^{\mathsf{T}} \nabla \varphi_c(x)\big(f(x) + g(x)\mu(\hat{x}_j) \\ &\quad + h(x)\nu(x)\big) \\ &\triangleq e_{cH}, \end{aligned} \tag{18}$$

where the term

$$e_{cH} = -\big(\nabla \varepsilon_c(x)\big)^{\mathsf{T}}\big(f(x) + g(x)\mu(\hat{x}_j) + h(x)\nu(x)\big)$$

represents the residual error arising in the approximate operation. Meanwhile, the approximate Hamiltonian function is

$$\begin{aligned} \hat{H}&\big(x, \mu(\hat{x}_j), \nu(x), \hat{\omega}_c\big) \\ &= U\big(x, \mu(\hat{x}_j), \nu(x)\big) + \hat{\omega}_c^{\mathsf{T}}(t) \nabla \varphi_c(x)\big(f(x) \\ &\quad + g(x)\mu(\hat{x}_j) + h(x)\nu(x)\big) \\ &\triangleq e_c. \end{aligned} \tag{19}$$

Let us define the weight error vector as $\tilde{\omega}_c(t) = \omega_c - \hat{\omega}_c(t)$. Then, we combine (18) with (19) to yield

$$\begin{aligned} e_c = &- \tilde{\omega}_c^{\mathsf{T}}(t) \nabla \varphi_c(x)\big(f(x) + g(x)\mu(\hat{x}_j) + h(x)\nu(x)\big) \\ &+ e_{cH}. \end{aligned}$$

Next, we show how to train the critic neural network. Here, we aim at minimizing the objective function defined as $E_c = 0.5 e_c^2$ to get $\hat{\omega}_c(t)$. It should be pointed out that the control pair of (17) is often adopted during the learning process because the optimal control and the worst-case disturbance are unavailable to be obtained. Based on (19), the normalized steepest descent algorithm is employed to regulate the weight vector $\hat{\omega}_c(t)$:

$$\begin{aligned} \dot{\hat{\omega}}_c(t) &= -\alpha_c \frac{1}{(1 + \psi^{\mathsf{T}}\psi)^2}\left(\frac{\partial E_c}{\partial \hat{\omega}_c(t)}\right) \\ &= -\alpha_c \frac{\psi}{(1 + \psi^{\mathsf{T}}\psi)^2}\big(U(x, \hat{\mu}(\hat{x}_j), \hat{\nu}(x)) \\ &\quad + \psi^{\mathsf{T}} \hat{\omega}_c(t)\big), \end{aligned} \tag{20}$$

where $\alpha_c > 0.5$ is the learning rate of the critic neural network,

$\psi = \nabla\varphi_c(x)\big(f(x) + g(x)\hat{\mu}(\hat{x}_j) + h(x)\hat{\nu}(x)\big)$

is a l_c-dimensional column vector, and $(1 + \psi^{\mathsf{T}}\psi)^2$ is a regularized term [45].

For the sake of clarity, a simple diagram of the adaptive critic-based nonlinear H_∞ control design that integrated the event-based component is depicted in Fig. 1, where the solid blocks exhibit the network-based computation modules, while the dashed blocks reveal the time/event transformation components. The solid line denotes the signal flow path for H_∞ control design, while the dashed line represents the back-propagation path for neural network training.

By using $\dot{\tilde{\omega}}_c(t) = -\dot{\hat{\omega}}_c(t)$ and introducing the following notations

$$\psi_1 = \frac{\psi}{(1+\psi^{\mathsf{T}}\psi)}, \psi_2 = 1 + \psi^{\mathsf{T}}\psi,$$

the error dynamics of critic neural network are further investigated, which can be written as

$$\dot{\tilde{\omega}}_c(t) = -\alpha_c\psi_1\psi_1^{\mathsf{T}}\tilde{\omega}_c(t) + \alpha_c\frac{\psi_1}{\psi_2}e_{cH}. \tag{21}$$

It is well known that the persistence of excitation is necessary to execute the system identification [46]. Therefore, this assumption is also required in this paper since the parameters of critic neural network need to be identified such that the cost function can be approximated.

Assumption 4 (*cf.* [18]) The signal ψ_1 is with the property of persistent excitation in the time interval $[t, t+T]$, $T > 0$, i.e., there exist two constants $\varsigma_1 > 0$ and $\varsigma_2 > 0$ such that

$$\varsigma_1 I_{l_c} \le \int_t^{t+T} \psi_1(\tau)\psi_1^{\mathsf{T}}(\tau)\mathrm{d}\tau \le \varsigma_2 I_{l_c}$$

holds for all t.

Based on Assumption 4, the persistent excitation condition means that $\lambda_{\min}(\psi_1\psi_1^{\mathsf{T}}) > 0$, which is useful in the following stability analysis.

In the event-triggered control, the closed-loop sampled-data system contains a flow dynamics for all $t \in [s_j, s_{j+1})$ and a jump dynamics at all $t = s_{j+1}$ with $j \in \mathbb{N}$. Before proceeding the stability issue of the closed-loop system, Assumption 5 is required, which is similar in [27,36,42].

Assumption 5 The derivative of used activation function is Lipschitz continuous, i.e., $\|\nabla\varphi_c(x) - \nabla\varphi_c(\hat{x}_j)\| \le M_\varphi\|\sigma_j(t)\|$, where M_φ is a positive constant. $\nabla\varphi_c(x)$, $\nabla\epsilon_c(x)$ and e_{cH} are upper-bounded by $\|\nabla\varphi_c(x)\| \le B_\varphi$, $\|\nabla\epsilon_c(x)\| \le B_\epsilon$, and $\|e_{cH}\| \le B_e$, where B_φ, B_ϵ and B_e are positive constants.

Theorem 2 *With Assumptions 3 and 5, for the nonlinear system (1), the event-based approximate optimal control law is given by (17a), and the time-based approximate worst-case disturbance law is (17b), where the weight vector of critic neural network is updated according to (20). Then, the closed-loop system (1) is asymptotically stable, and the weight error vector is uniformly ultimately bounded with the following triggering condition*

$$\|\sigma_j(t)\|^2 \le \hat{\sigma}_T = \frac{x^{\mathsf{T}}Qx + \|\hat{\mu}(\hat{x}_j)\|^2 - \iota^2\|\hat{\nu}(x)\|^2}{2M_{\mathcal{L}}\|\hat{\omega}_c\|^2}, \tag{22}$$

where the inequality

$$\|\tilde{\omega}_c(t)\| > \sqrt{\frac{2B_g^2B_\epsilon^2 + \alpha_c^2B_e^2}{(2\alpha_c - 1)\lambda_{\min}(\psi_1\psi_1^{\mathsf{T}}) - 2B_g^2B_\varphi^2}} \tag{23}$$

is satisfied when $M_{\mathcal{L}} = M_g^2B_\varphi^2 + M_\varphi^2B_g^2$ *and* $\alpha_c > 0.5$.

Proof Construct a Lyapunov function candidate as the formula

$$L_2(t) = L_{21}(t) + L_{22}(t) + L_{23}(t),$$

where

$$L_{21}(t) = J^*(x), L_{22}(t) = J^*(\hat{x}_j),$$
$$L_{23}(t) = \frac{1}{2}\tilde{\omega}_c^{\mathsf{T}}(t)\tilde{\omega}_c(t).$$

When $t \in [s_j, s_{j+1})$, the events are not triggered. The time derivative of $L_2(t)$ is calculated as

$$\dot{L}_{21}(t) = \big(\nabla J^*(x)\big)^{\mathsf{T}}\big(f(x) + g(x)\hat{\mu}(\hat{x}_j) + h(x)\hat{\nu}(x)\big),$$

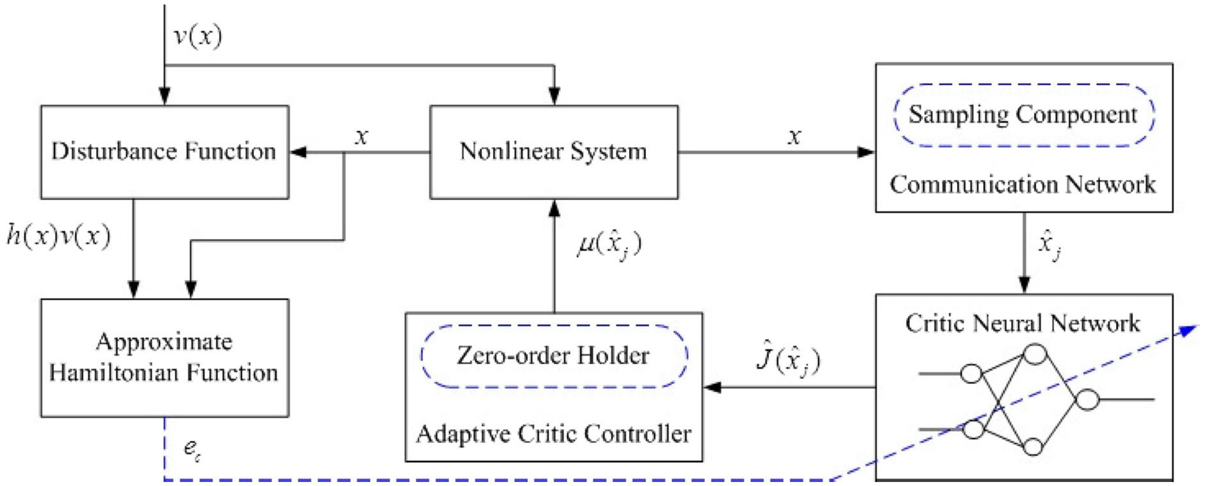

Fig. 1 Simple structure of adaptive critic H_∞ control design with network-based event-triggered mechanism

$\dot{L}_{22}(t) = 0$, and

$$\dot{L}_{23}(t) = -\alpha_c \tilde{\omega}_c^{\mathsf{T}}(t)\psi_1\psi_1^{\mathsf{T}}\tilde{\omega}_c(t) + \alpha_c \frac{\tilde{\omega}_c^{\mathsf{T}}(t)\psi_1}{\psi_2} e_{cH}.$$

For the term $\dot{L}_{21}(t)$, based on (5) and (6), by adding and subtracting $\hat{\mu}^{\mathsf{T}}(\hat{x}_j)\hat{\mu}(\hat{x}_j)$, $\dot{L}_{21}(t)$ can be obtained as

$$\begin{aligned}\dot{L}_{21}(t) =& -x^{\mathsf{T}}Qx + u^{*\mathsf{T}}(x)u^*(x) - 2u^{*\mathsf{T}}(x)\hat{\mu}(\hat{x}_j)\\ &- \iota^2 v^{*\mathsf{T}}(x)v^*(x) + 2\iota^2 v^{*\mathsf{T}}(x)\hat{v}(x)\\ \le& -x^{\mathsf{T}}Qx + \|u^*(x) - \hat{\mu}(\hat{x}_j)\|^2 - \|\hat{\mu}(\hat{x}_j)\|^2\\ &+ \iota^2\|\hat{v}(x)\|^2.\end{aligned}$$

Considering (5a) and using the neural network expression, the time-based optimal control can be reformulated as

$$u^*(x) = -\frac{1}{2}g^{\mathsf{T}}(x)\Big(\big(\nabla\varphi_c(x)\big)^{\mathsf{T}}\omega_c + \nabla\epsilon_c(x)\Big). \tag{24}$$

Using $\hat{\mu}(\hat{x}_j)$ in (17a) and $u^*(x)$ in (24), it follows from $\omega_c = \hat{\omega}_c(t) + \tilde{\omega}_c(t)$ that

$$\begin{aligned}&\|u^*(x) - \hat{\mu}(\hat{x}_j)\|^2\\ \le&\big\|\big[g^{\mathsf{T}}(\hat{x}_j)\big(\nabla\varphi_c(\hat{x}_j)\big)^{\mathsf{T}} - g^{\mathsf{T}}(x)\big(\nabla\varphi_c(x)\big)^{\mathsf{T}}\big]\hat{\omega}_c(t)\big\|^2\\ &+ \big\|g^{\mathsf{T}}(x)\Big(\big(\nabla\varphi_c(x)\big)^{\mathsf{T}}\tilde{\omega}_c(t) + \nabla\epsilon_c(x)\Big)\big\|^2.\end{aligned}$$

Recalling Assumptions 3 and 5, it yields

$$\begin{aligned}&\big\|g^{\mathsf{T}}(\hat{x}_j)\big(\nabla\varphi_c(\hat{x}_j)\big)^{\mathsf{T}} - g^{\mathsf{T}}(x)\big(\nabla\varphi_c(x)\big)^{\mathsf{T}}\big\|^2\\ =&\big\|\big(\nabla\varphi_c(\hat{x}_j) - \nabla\varphi_c(x)\big)g(\hat{x}_j) + \nabla\varphi_c(x)\big(g(\hat{x}_j) - g(x)\big)\big\|^2\\ \le& 2M_{\mathcal{L}}\|\sigma_j(t)\|^2.\end{aligned}$$

Thus, the following inequality can be obtained

$$\begin{aligned}\dot{L}_{21}(t) \le& -x^{\mathsf{T}}Qx - \|\hat{\mu}(\hat{x}_j)\|^2 + \iota^2\|\hat{v}(x)\|^2 + B_g^2B_\epsilon^2\\ &+ 2M_{\mathcal{L}}\|\hat{\omega}_c(t)\|^2\|\sigma_j(t)\|^2 + B_g^2B_\varphi^2\|\tilde{\omega}_c(t)\|^2.\end{aligned} \tag{25}$$

For the term $\dot{L}_{23}(t)$, by applying the Young's inequality into its second term, $\dot{L}_{23}(t)$ satisfies

$$\begin{aligned}\dot{L}_{23}(t) \le& -\Big(\alpha_c - \frac{1}{2}\Big)\lambda_{\min}(\psi_1\psi_1^{\mathsf{T}})\|\tilde{\omega}_c(t)\|^2\\ &+ \frac{1}{2}\alpha_c^2 B_e^2,\end{aligned} \tag{26}$$

where Assumption 5 and the fact $\psi_2 \ge 1$ are used. By combining (25) and (26), we can obtain that the overall time derivative of $L_2(t)$ is

$$\begin{aligned}\dot{L}_2(t) \le& -x^{\mathsf{T}}Qx + 2M_{\mathcal{L}}\|\hat{\omega}_c(t)\|^2\|\sigma_j(t)\|^2 - \|\hat{\mu}(\hat{x}_j)\|^2\\ &+ \iota^2\|\hat{v}(x)\|^2 + B_g^2B_\epsilon^2 + \frac{1}{2}\alpha_c^2B_e^2\\ &- \Big[\Big(\alpha_c - \frac{1}{2}\Big)\lambda_{\min}(\psi_1\psi_1^{\mathsf{T}}) - B_g^2B_\varphi^2\Big]\|\tilde{\omega}_c(t)\|^2.\end{aligned} \tag{27}$$

Therefore, it is clear that if (22) and (23) are satisfied, then $\dot{L}_2(t) < 0$ for any $x \neq 0$ can be obtained according to (27).

When $t = s_{j+1}$, the events are triggered. The difference of $L_2(t)$ is expressed as

$$\begin{aligned}\Delta L_2(t) &= L_2(\hat{x}_{j+1}) - L_2\big(x(s_{j+1}^-)\big)\\ &= \Delta L_{21}(t) + \Delta L_{22}(t) + \Delta L_{23}(t),\end{aligned}$$

where $x(s_{j+1}^-) = \lim_{\varepsilon\to 0} x(s_{j+1}-\varepsilon)$ and ε is a sufficiently small positive constant. For all $t \in [s_j, s_{j+1})$, $\dot{L}_2(t) < 0$ can be derived from (22), (23) and (27). Considering that the system states and the cost function are all continuous, it can acquire

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

and $\Delta L_{23}(t) \le 0$, where

$$\begin{aligned}\Delta L_{23}(t) =& \frac{1}{2}\big[\tilde{\omega}_c^{\mathsf{T}}(\hat{x}_{j+1})\tilde{\omega}_c(\hat{x}_{j+1}) - \tilde{\omega}_c^{\mathsf{T}}\big(x(s_{j+1}^-)\big)\\ &\times \tilde{\omega}_c\big(x(s_{j+1}^-)\big)\big].\end{aligned}$$

Hence, we obtain

$$\begin{aligned}\Delta L_2(t) &\le \Delta L_{22}(t) = J^*(\hat{x}_{j+1}) - J^*(\hat{x}_j)\\ &\le -\mathcal{K}(\|\sigma_{j+1}(s_j)\|),\end{aligned}$$

where $\mathcal{K}(\cdot)$ is a class-$\mathcal{K}$ function [44] and $\sigma_{j+1}(s_j) = \hat{x}_{j+1} - \hat{x}_j$. This leads to that $L_2(t)$ is decreasing for all $t = s_{j+1}$.

Based on these two cases, with the triggering condition (22) and the uniformly ultimately bounded weight error in (23), the closed-loop system (1) is asymptotically stable, which ends the proof. □

Remark 2 If we regard the first term of weight error dynamics (21) as a nominal system, which is written as $\dot{\tilde{\omega}}_{cn}(t) = -\alpha_c\psi_1\psi_1^{\mathsf{T}}\tilde{\omega}_{cn}(t)$, we can verify that it is exponentially stable. To this end, we choose a Lyapunov function as the form $L_{cn}(t) = 0.5\tilde{\omega}_{cn}^{\mathsf{T}}(t)\tilde{\omega}_{cn}(t)$ and differentiate it along the nominal part to yield $\dot{L}_{cn}(t) = -\alpha_c\tilde{\omega}_{cn}^{\mathsf{T}}(t)\psi_1\psi_1^{\mathsf{T}}\tilde{\omega}_{cn}(t)$, which clearly reveals that $\dot{L}_{cn}(t) \le 0$ and exhibits the stability of the nominal system. Moreover, the solution $\tilde{\omega}_{cn}(t)$ can be given by $\tilde{\omega}_{cn}(t) = \mathcal{T}(t,0)\tilde{\omega}_{cn}(0)$, where the state transition matrix is defined as $\dot{\mathcal{T}}(t,0) = -\alpha_c\psi_1\psi_1^{\mathsf{T}}\mathcal{T}(t,0)$. Hence, according to [44], there exist two constants ς_3 and ς_4 such that

$$\|\mathcal{T}(t,0)\| \le \varsigma_3 e^{-\varsigma_4 t}, \forall t \ge 0.$$

Under such circumstance, we can derive that

$$\|\tilde{\omega}_{cn}(t)\| \le \|\mathcal{T}(t,0)\|\|\tilde{\omega}_{cn}(0)\| \le \varsigma_3\|\tilde{\omega}_{cn}(0)\|e^{-\varsigma_4 t}.$$

Thus, it is shown that for the nominal part of the critic error dynamics (21), the equilibrium point is exponentially stable in case that ψ_1 satisfies the persistence of excitation condition. Note that this kind of stability with respect to the nominal system is stronger than the uniformly ultimately bounded stability of the whole error dynamics developed in Theorem 2. Nevertheless, the existence of the residual error-related term is indeed indispensable due to the neural network approximation, which eventually results in a weaker stability of the critic error dynamics.

It should be mentioned that although two triggering thresholds σ_T and $\hat{\sigma}_T$ are provided in Theorems 1 and 2, respectively, it is obvious that two thresholds work in different design stages. Overall, the event-based robust adaptive critic control algorithm can be summarized in Algorithm 1.

Algorithm 1 (Robust adaptive critic control method with network-based event-triggered formulation)

1: Set the parameter ι and the matrix P, calculate $Q = C^{\mathsf{T}}PC$ and the cost function. Choose an appropriate activation function $\varphi_c(x)$ and randomly initialize $\hat{\omega}_c(t)$.
2: Set the parameter values of α_c and $M_{\mathcal{L}}$. Then, conduct adaptive critic learning by employing the weight updating rule (20) and the triggering condition (22) with the threshold $\hat{\sigma}_T$.
3: After the online training process of step 2, retain the converged weight vector and then proceed the robust H_∞ control implementation.
4: Set the parameter value of M_u and carry out the robust adaptive critic control with the triggering condition (8) and the threshold σ_T.
5: Get the event-based optimal control law and the worst-case disturbance law, and then stop the algorithm.

4 Simulation analysis

In this section, a numerical example is conducted to demonstrate the effectiveness of the event-based non-

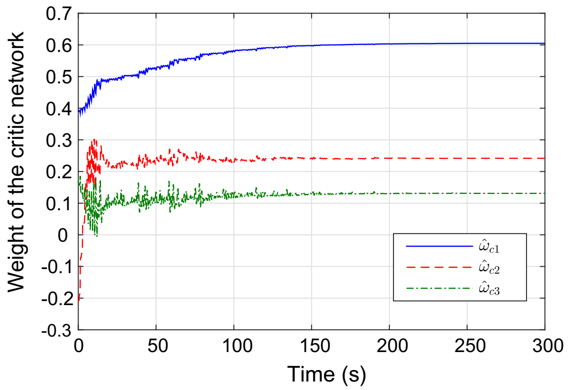

Fig. 2 Convergence curves of the critic network weights

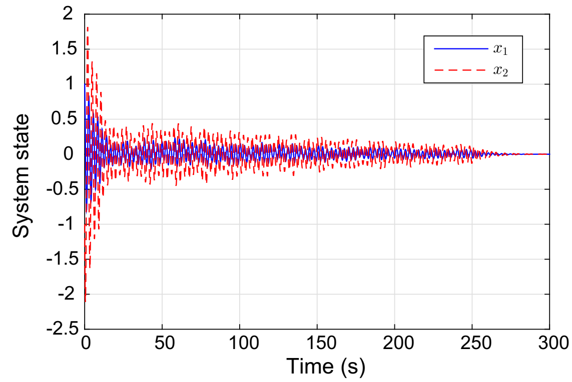

Fig. 3 Adaptive regulation of the state trajectories

linear H_∞ control. We consider a single-link robot arm with the description in [7,40,47,48], and the mechanical dynamics are derived by

$$\bar{G}\ddot{\theta}(t) = -M\bar{g}\bar{H}\sin(\theta(t)) - D\dot{\theta}(t) + u(t) + \nu(t), \tag{28a}$$

$$y(t) = \theta(t), \tag{28b}$$

where $\theta(t)$ denotes the angle position, $u(t)$ is the control, and $\nu(t)$ is the perturbation. $M = 10$ and $\bar{H} = 0.5$ are the mass and the length of the robot arm, respectively, $\bar{g} = 9.81$ is the gravity acceleration, $D = 2$ is the viscous friction, and $\bar{G} = 10$ is the inertia moment.

Define $x = [x_1, x_2]^\mathsf{T}$ with $x_1 = \theta$ and $x_2 = \dot{\theta}$ such that the dynamic equation of system (28) is rewritten as

$$\begin{bmatrix}\dot{x}_1\\ \dot{x}_2\end{bmatrix} = \begin{bmatrix} x_2 \\ -4.905\sin x_1 - 0.2x_2\end{bmatrix} + \begin{bmatrix}0\\0.1\end{bmatrix}(u+\nu), \tag{29a}$$

$$y = x_1. \tag{29b}$$

Obviously, the control and disturbance matrices are constants, which are both upper-bounded. For instance, we can choose $B_g = B_h = 0.1$. Then, the initial state vector of (29) is set as $x_0 = [1, -1]^\mathsf{T}$, and choose $P = 2$ so that $Q = \text{diag}\{2, 0\}$. The adaptive critic controller is designed for system (29) in the following.

In the simulation, the critic neural network is constructed as

$$\hat{J}(x) = \hat{\omega}_c^\mathsf{T}\varphi_c(x) = \hat{\omega}_{c1}x_1^2 + \hat{\omega}_{c2}x_1x_2 + \hat{\omega}_{c3}x_2^2,$$

where $\hat{\omega}_c = [\hat{\omega}_{c1}, \hat{\omega}_{c2}, \hat{\omega}_{c3}]^\mathsf{T}$ and $\varphi_c(x) = [x_1^2, x_1x_2, x_2^2]^\mathsf{T}$. Clearly, the derivative of the activation function is a 3×2 function matrix of the form

$$\nabla\varphi_c(x) = \begin{bmatrix} 2x_1 & x_2 & 0 \\ 0 & x_1 & 2x_2\end{bmatrix}^\mathsf{T}.$$

The neuron number of the hidden layer is often decided by computer experiment. We can certainly choose hidden neurons of any number, while we should also consider the complexity of the computation issue. In this case study, we find that selecting three hidden neurons can lead to satisfactory simulation results. In other words, it can be observed that the choice of the activation function is more of an art than science. For adjusting the critic network, we experimentally set $\alpha_c = 1.2$, $\iota = 2$, and $M_\mathcal{L} = 36$. The sampling time in the learning process is selected as 0.1 s. Note that we also employ a probing noise to ensure the persistency of excitation condition in the training process. The simulation results of the learning stage are shown in Figs. 2, 3 and 4. In Fig. 2, it can be observed that the critic network weight vector converges to $[0.6050, 0.2418, 0.1310]^\mathsf{T}$.

The adaptive regulation process of state trajectories and the triggering condition is displayed in Fig. 3, where the system is trained under the persistency of excitation condition and the states are regulated to zero once the excitation signals are stopped. Figure 4 provides the adjustment process of the triggering condition, in which the triggering condition with respect to $\|\sigma_j(t)\|^2$ and $\hat{\sigma}_T$ is shown. It can be observed that the time-based controller uses 3000 state samples, while the event-based controller only needs 1501 samples,

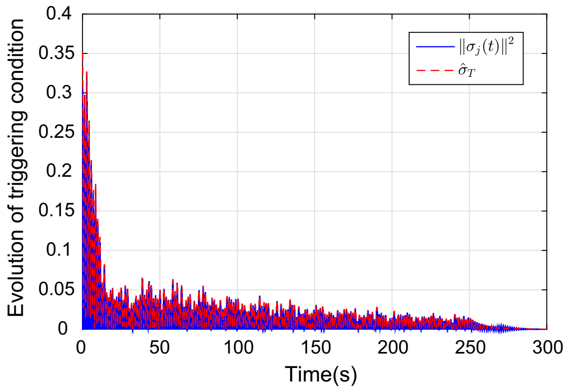

Fig. 4 Adjustment of the triggering condition with the relationship between $\|\sigma_j(t)\|^2$ and $\hat{\sigma}_T$

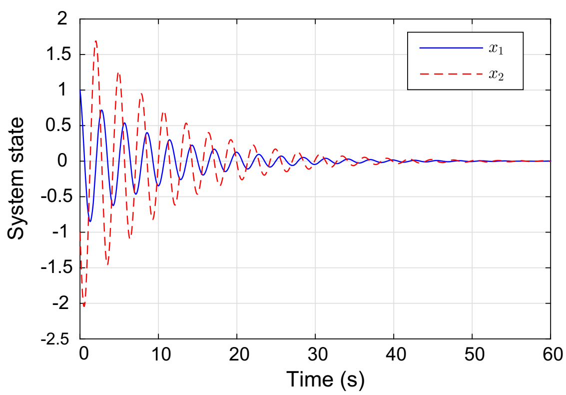

Fig. 5 System state trajectories under the robust H_∞ control

thereby resulting in an evident reduction in the data transmission.

For the controlled plant (29), the obtained control law is used for 60 s with the external perturbation $\nu(t) = 5e^{-(t)}\cos(t)$, $t > 0$, to evaluate the robust H_∞ control performance. Set $M_u = 5$, and the sampling time is 0.05 s. The simulation results with the H_∞ feedback control are exhibited in Figs. 5, 6, 7 and 8. Specifically, the system state trajectories and the control input trajectory are depicted in Figs. 5 and 6, respectively. Figure 7 shows the adjustment of the triggering condition under the robust H_∞ control.

Then, referring to the common definition in [27–29], a ratio function $\bar{\iota}(t)$ is defined as

$$\bar{\iota}(t) = \sqrt{\int_0^t \left(x^{\mathsf{T}}(\tau)Qx(\tau) + u^{\mathsf{T}}(\tau)u(\tau)\right)\mathrm{d}\tau \Big/ \int_0^t \|\nu(\tau)\|^2 \mathrm{d}\tau},$$

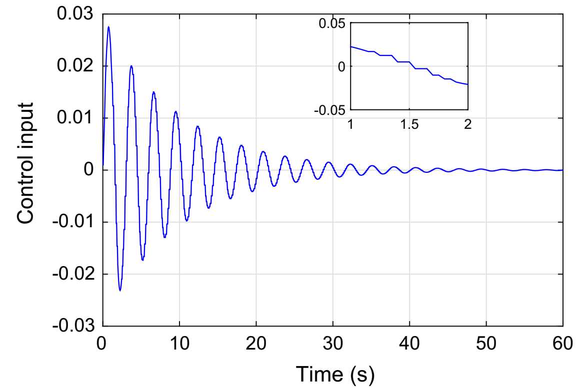

Fig. 6 Robust H_∞ control input

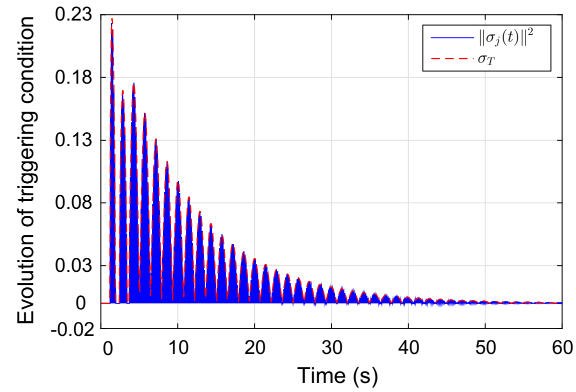

Fig. 7 Adjustment of the triggering condition under the robust H_∞ control

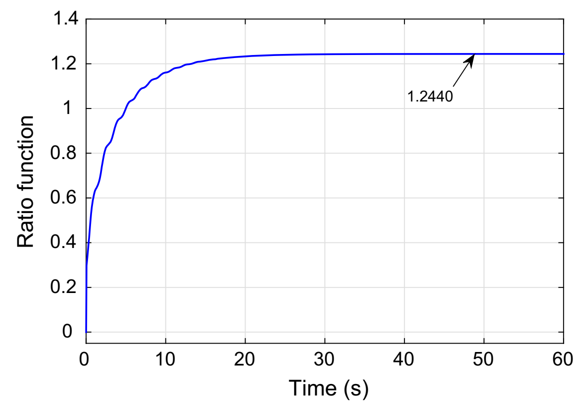

Fig. 8 Evolution curve of the ratio function

which is used to reflect the disturbance attenuation of the H_∞ control problem. In Fig. 8, the ratio $\bar{\iota}(t)$ gradually converges to 1.2440 over time. This implies that the designed H_∞ controller really works on attaining a prespecified L_2-gain performance level (i.e., $\bar{\iota}(t) < \iota = 2$).

These simulation results substantiate the effectiveness of the event-based robust adaptive critic control strategy with regard to the external disturbance, and consequently, it possesses the excellent ability of disturbance rejection.

5 Conclusion

In this paper, the event-based H_∞ feedback control of nonlinear dynamic systems involving output information has been intensively studied under the event-based adaptive critic design framework. It formulated the H_∞ control problem of disturbed nonlinear system as the problem of two-player zero-sum differential game. The event-based mechanism and the adaptive critic approach have been adopted to pursue the Nash equilibrium solution of the two-player zero-sum differential game such that the event-based approximate optimal control law and the time-based worst-case disturbance law were derived by the learning process of the critic network, where the triggering condition and its related threshold were provided. Simultaneously, this paper also presented the stability analysis of the closed-loop system and the weight estimation error of critic neural network. With the experimental verification of a single-link robot arm, the theoretical results have been well demonstrated and illustrated. Along this direction of the event-triggered adaptive critic control, some interesting research topics can be further studied in the future work, such as the event-triggered approximate optimal tracking control design for affine nonlinear systems with unmatched uncertainties, for nonaffine nonlinear systems including uncertainties and unknown dynamics.

Acknowledgements This work was supported by National Natural Science Foundation of China under Grants 61773284, 61773373, U1501251, 61533008 and 61520106009, Beijing Natural Science Foundation under Grant 4162065, Tianjin Natural Science Foundation under Grant 14JCQNJC05400, China Post-doctoral Science Foundation under Grant 2014M561559.

References

1. Dolk, V.S., Tesi, P., Persis, C.D., Heemels, W.P.M.H.: Event-triggered control systems under denial-of-service attacks. IEEE Trans. Control Netw. Syst. **41**(93), 93–105 (2017)
2. Liu, S., Xie, L., Quevedo, D.E.: Event-triggered quantized communication based distributed convex optimization. IEEE Trans. Control Netw. Syst (2016). doi:10.1109/TCNS.2016.2585305 **(in press)**
3. Wu, Z., Jia, Q.S., Guan, X.: Optimal control of multiroom HVAC system: an event-based approach. IEEE Trans. Control Syst. Technol. **24**(2), 662–669 (2016)
4. Wang, D., Mu, C., He, H., Liu, D.: Event-driven adaptive robust control of nonlinear systems with uncertainties through NDP strategy. IEEE Trans. Syst. Man Cybern. Syst. **47**(7), 1358–1370 (2017)
5. Werbos, P.J.: Approximate dynamic programming for real-time control and neural modeling. In: White, D.A., Sofge, D.A. (eds.) Handbook of Intelligent Control: Neural, Fuzzy, and Adaptive Approaches. Van Nostrand Reinhold, New York (1992)
6. Dong, N., Chen, Z.Q.: A novel ADP based model-free predictive control. Nonlinear Dyn. **69**(1–2), 89–97 (2012)
7. Mu, C., Wang, D., He, H.: Novel iterative neural dynamic programming for data-based approximate optimal control design. Automatica **81**, 240–252 (2017)
8. Hendzel, Z.: An adaptive critic neural network for motion control of a wheeled mobile robot. Nonlinear Dyn. **50**(4), 849–855 (2007)
9. He, W., Chen, Y., Yin, Z.: Adaptive neural network control of an uncertain robot with full-state constraints. IEEE Trans. Cybern. **46**(3), 620–629 (2016)
10. Wang, Y., Cheng, L., Hou, Z.G., Yu, J., Tan, M.: Optimal formation of multi-robot systems based on a recurrent neural network. IEEE Trans. Neural Netw. Learn. Syst. **27**(2), 322–333 (2016)
11. Xie, X., Yue, D., Zhang, H., Peng, C.: Control synthesis of discrete-time T-S fuzzy systems: reducing the conservatism whilst alleviating the computational burden. IEEE Trans. Cybern. **47**(9), 2480–2491 (2017)
12. Gu, B., Sheng, V.S., Wang, Z., Ho, D., Osman, S., Li, S.: Incremental learning for ν-support vector regression. Neural Netw. **67**, 140–150 (2015)
13. Dierks, T., Thumati, B.T., Jagannathan, S.: Optimal control of unknown affine nonlinear discrete-time systems using offline-trained neural networks with proof of convergence. Neural Netw. **22**(5–6), 851–860 (2009)
14. Wang, D., Liu, D., Wei, Q., Zhao, D., Jin, N.: Optimal control of unknown nonaffine nonlinear discrete-time systems based on adaptive dynamic programming. Automatica **48**(8), 1825–1832 (2012)
15. Heydari, A., Balakrishnan, S.N.: Finite-horizon control-constrained nonlinear optimal control using single network adaptive critics. IEEE Trans. Neural Netw. Learn. Syst. **24**(1), 145–157 (2013)
16. Zhao, Q., Xu, H., Jagannathan, S.: Near optimal output feedback control of nonlinear discrete-time systems based on reinforcement neural network learning. IEEE/CAA J. Autom. Sin. **1**(4), 372–384 (2014)
17. Mu, C., Ni, Z., Sun, C., He, H.: Air-breathing hypersonic vehicle tracking control based on adaptive dynamic programming. IEEE Trans. Neural Netw. Learn. Syst. **28**(3), 584–598 (2017)
18. Vamvoudakis, K.G., Lewis, F.L.: Online actor–critic algorithm to solve the continuous-time infinite horizon optimal control problem. Automatica **46**(5), 878–888 (2010)

19. Jiang, Y., Jiang, Z.P.: Global adaptive dynamic programming for continuous-time nonlinear systems. IEEE Trans. Autom. Control **60**(11), 2917–2929 (2015)
20. Gao, W., Jiang, Z.P.: Adaptive dynamic programming and adaptive optimal output regulation of linear systems. IEEE Trans. Autom. Control **61**(12), 4164–4169 (2016)
21. Wang, D., Liu, D., Mu, C., Ma, H.: Decentralized guaranteed cost control of interconnected systems with uncertainties: a learning-based optimal control strategy. Neurocomputing **214**, 297–306 (2016)
22. Bian, T., Jiang, Y., Jiang, Z.P.: Decentralized adaptive optimal control of large-scale systems with application to power systems. IEEE Trans. Industr. Electron. **62**(4), 2439–2447 (2015)
23. Zhang, H., Jiang, H., Luo, Y., Xiao, G.: Data-driven optimal consensus control for discrete-time multi-agent systems with unknown dynamics using reinforcement learning method. IEEE Trans. Industr. Electron. **64**(5), 4091–4100 (2017)
24. Zhang, H., Liu, D., Luo, Y., Wang, D.: Adaptive Dynamic Programming for Control: Algorithms and Stability. Springer, London (2013)
25. Abu-Khalaf, M., Lewis, F.L., Huang, J.: Policy iterations on the Hamilton–Jacobi–Isaacs equation for H_∞ state feedback control with input saturation. IEEE Trans. Autom. Control **51**(12), 1989–1995 (2006)
26. Liu, D., Li, H., Wang, D.: Neural-network-based zero-sum game for discrete-time nonlinear systems via iterative adaptive dynamic programming algorithm. Neurocomputing **110**, 92–100 (2013)
27. Zhang, H., Qin, C., Jiang, B., Luo, Y.: Online adaptive policy learning algorithm for H_∞ state feedback control of unknown affine nonlinear discrete-time systems. IEEE Trans. Cybern. **44**(12), 2706–2718 (2014)
28. Luo, B., Wu, H.N., Huang, T.: Off-policy reinforcement learning for H_∞ control design. IEEE Trans. Cybern. **45**(1), 65–76 (2015)
29. Wang, D., Mu, C., Liu, D., Ma, H.: On mixed data and event driven design for adaptive-critic-based nonlinear H_∞ control. IEEE Trans. Neural Netw. Learn. Syst. (2016). doi:10.1109/TNNLS.2016.2642128 **(in press)**
30. Prokhorov, D.V., Wunsch, D.C.: Adaptive critic designs. IEEE Trans. Neural Netw. **8**(5), 997–1007 (1997)
31. Wang, D., Liu, D., Li, H.: Policy iteration algorithm for online design of robust control for a class of continuous-time nonlinear systems. IEEE Trans. Autom. Sci. Eng. **11**(2), 627–632 (2014)
32. Jiang, Y., Jiang, Z.: Robust adaptive dynamic programming and feedback stabilization of nonlinear systems. IEEE Trans. Neural Netw. Learn. Syst. **25**(5), 882–893 (2014)
33. Liu, D., Yang, X., Wang, D., Wei, Q.: Reinforcement-learning-based robust controller design for continuous-time uncertain nonlinear systems subject to input constraints. IEEE Trans. Cybern. **45**(7), 1372–1385 (2015)
34. Bian, T., Jiang, Y., Jiang, Z.: Decentralized adaptive optimal control of large-scale systems with application to power systems. IEEE Trans. Industr. Electron. **62**(4), 2439–2447 (2015)
35. Wang, D., Li, C., Liu, D., Mu, C.: Data-based robust optimal control of continuous-time affine nonlinear systems with matched uncertainties. Inf. Sci. **366**, 121–133 (2016)
36. Zhang, Q., Zhao, D., Wang, D.: Event-based robust control for uncertain nonlinear systems using adaptive dynamic programming. IEEE Trans. Neural Netw. Learn. Syst. doi:10.1109/TNNLS.2016.2614002 **(in press)**
37. Mu, C., Sun, C., Wang, D., Song, A.: Adaptive tracking control for a class of continuous-time uncertain nonlinear systems using the approximate solution of HJB equation. Neurocomputing **260**, 432–442 (2017)
38. Liu, Y., Lee, S.M.: Improved results on sampled-data synchronization of complex dynamical networks with time-varying coupling delay. Nonlinear Dyn. **81**(1), 931–938 (2015)
39. Liu, Y., Guo, B.Z., Park, J.H., Lee, S.M.: Nonfragile exponential synchronization of delayed complex dynamical networks with memory sampled-data control. IEEE Trans. Neural Netw. Learn. Syst. (2016). doi:10.1109/TNNLS.2016.2614709 **(in press)**
40. Zhong, X., He, H.: An event-triggered ADP control approach for continuous-time system with unknown internal states. IEEE Trans. Cybern. **47**(3), 683–694 (2017)
41. Vamvoudakis, K.G.: Event-triggered optimal adaptive control algorithm for continuous-time nonlinear systems. IEEE/CAA J. Autom. Sin. **1**(3), 282–293 (2014)
42. Sahoo, A., Xu, H., Jagannathan, S.: Neural network-based event-triggered state feedback control of nonlinear continuous-time systems. IEEE Trans. Neural Netw. Learn. Syst. **27**(3), 497–509 (2016)
43. Zhang, Q., Zhao, D., Zhu, Y.: Event-triggered H_∞ control for continuous-time nonlinear system via concurrent learning. IEEE Trans. Syst. Man Cybern. Syst. **47**(7), 1071–1081 (2017)
44. Khalil, H.K.: Nonlinear Systems, 3rd edn. Prentice-Hall, Upper Saddle River (2002)
45. Beard, R.W., Saridis, G.N., Wen, J.T.: Galerkin approximations of the generalized Hamilton–Jacobi–Bellman equation. Automatica **33**, 2159–2177 (1997)
46. Krstic, M., Kanellakopoulos, I., Kokotovic, P.: Nonlinear and Adaptive Control Design. Wiley, New York (1995)
47. Kim, Y.H., Lewis, F.L., Abdallah, C.T.: A dynamic recurrent neural-network-based adaptive observer for a clas of nonlinear systems. Automatica **33**, 1539–1543 (1997)
48. Zhong, X., Ni, Z., He, H.: A theoretical foundation of goal representation heuristic dynamic programming. IEEE Trans. Neural Netw. Learn. Syst. **27**(12), 2513–2525 (2016)