

Event-triggered intelligent critic control with input constraints applied to a nonlinear aeroelastic system

Sun, Bo; Wang, Xuerui; van Kampen, Erik Jan

DOI

[10.1016/j.ast.2021.107279](https://doi.org/10.1016/j.ast.2021.107279)

Publication date

2022

Document Version

Final published version

Published in

Aerospace Science and Technology

Citation (APA)

Sun, B., Wang, X., & van Kampen, E. J. (2022). Event-triggered intelligent critic control with input constraints applied to a nonlinear aeroelastic system. *Aerospace Science and Technology*, 120, Article 107279. <https://doi.org/10.1016/j.ast.2021.107279>

Important note

To cite this publication, please use the final published version (if applicable). Please check the document version above.

Copyright

Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

Takedown policy

Please contact us and provide details if you believe this document breaches copyrights. We will remove access to the work immediately and investigate your claim.



Event-triggered intelligent critic control with input constraints applied to a nonlinear aeroelastic system

Bo Sun^{a,*}, Xuerui Wang^{a,b}, Erik-Jan van Kampen^a

^a Department of Control and Operations, Delft University of Technology, 2629 HS Delft, the Netherlands

^b Department of Aerospace Structures and Materials, Delft University of Technology, 2629 HS Delft, the Netherlands

ARTICLE INFO

Article history:

Received 4 January 2021

Received in revised form 6 September 2021

Accepted 7 December 2021

Available online 11 December 2021

Communicated by Z. Weiwei

Keywords:

Adaptive dynamic programming

Event-triggered control

Artificial neural network

Input constraints

Nonlinear aeroelastic system

ABSTRACT

In this paper, we establish an event-triggered intelligent control scheme with a single critic network, to cope with the optimal stabilization problem of nonlinear aeroelastic systems. The main contribution lies in the design of a novel triggering condition with input constraints, avoiding the Lipschitz assumption on the inverse hyperbolic tangent function. Based on an improved weight updating criterion that eliminates the requirement of initial admissible control, the control law is obtained approximately by online training of a single critic network. The Lyapunov stability and the Zeno phenomenon of the closed-loop system are analysed. The feasibility of the established algorithm is verified by applying it to an optimal stabilization task of a nonlinear aeroelastic system. The results reveal that the developed approach can handle input-constrained optimal control problems, with performance comparable to the time-based method that updates control inputs at each instant, while reducing the computational and communication's load.

© 2021 The Author(s). Published by Elsevier Masson SAS. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Aeroelastic systems exhibit a variety of unstable phenomena, such as flutter and limit-cycle oscillations (LCOs), which can significantly degrade the flight performance of an aircraft [1–3]. For this reason, stable controller design for aeroelastic systems has been receiving considerable attention for decades in aerospace engineering research groups [2,4–6]. Most current controllers are designed based on feedback linearization approaches [4]. However, with the rapid development of aviation technologies, these traditional methods show their limitations in dealing with stronger nonlinearities. Input nonlinearities such as saturation constraints commonly exist in real systems [7,8], but they are rarely investigated for aeroelastic systems in existing literature. Furthermore, complex systems usually involve multiple control loops closed through some communication mediums, which brings a growing interest in enhancing resource utilization [9]. Motivated by the demands for tackling these challenges, this paper aims at developing a constrained-input optimal control approach with reduced computational and communication's cost for nonlinear aeroelastic systems.

Optimal control problems pursue optimal control policies for dynamical systems by maximizing/minimizing a pre-defined performance function that captures desired objectives [10]. When

dealing with optimal control problems, it is common to solve the Hamilton-Jacobi-Bellman (HJB) equation, but there are few effective approaches to obtain its analytical solutions for nonlinear systems [10,11]. Adaptive dynamic programming (ADP) provides a promising method to acquire numerical solutions of general HJB equations. By incorporating artificial neural networks (ANNs), ADP acquires a more powerful generalization capability and has been successfully applied to a variety of aerospace systems [12–17]. As a branch of reinforcement learning (RL), the principle of ADP lies in the effective iterations between policy improvement and policy evaluation [10,18], which are sometimes approximated by an actor network and a critic network, respectively [12,13,15]. However, for continuous-time (CT) systems, by solving the HJB equation, the single critic network (SCN) architecture is able to perform ADP with lower computational cost and eliminating the approximation error introduced by the actor network [8,19]. Different from the actor-critic architecture, where the input saturation constraints are addressed by the bounded output neurons of the actor network [13], the SCN structure ordinarily utilizes a non-quadratic cost function, such that the control inputs derived from the solution of HJB equation can be bounded by a hyperbolic tangent function [8,20,21].

Although time-based ADP approaches provide a mature and normative solution to nonlinear optimal control problems, the need for reducing transmitted data is not fully satisfied. Arising from networked control systems [22,23], event-triggered control (ETC) has attained a lot of attention in recent years because of its ability to reduce computational and communication's load [10]. A

* Corresponding author.

E-mail addresses: b.sun-1@tudelft.nl (B. Sun), X.Wang-6@tudelft.nl (X. Wang), E.vanKampen@tudelft.nl (E.-J. van Kampen).

cross fertilization of ETC and ADP produces event-triggered ADP, which has successfully been implemented for optimal stabilization of both discrete-time systems [24,25] and CT systems [26,27]. The key attribute of the event-triggered mechanism lies in that the control signals are updated only when a certain condition is triggered [24]. Therefore, designing a sound triggering condition is the principal task of ETC. For CT systems adopting ETC methods, the inter-execution time can be zero, resulting in the accumulation of event times. This is the infamous Zeno phenomenon that must be avoided in the controller design. The related analysis is conducted in [28,29] without incorporating ANNs and in [9] without taking the input constraints into account. Based on these studies, the closed-loop analysis is carried out to ensure that the Zeno phenomenon is inapplicable to the proposed method.

Developing from the time-based ADP, event-triggered ADP methods inherit and continue most of the properties and techniques of the time-based ADP, including the technique for handling input constraints with a non-quadratic cost function [29,30]. However, in most existing literature, the triggering-condition is derived by involving a Lipschitz constant of the inverse hyperbolic tangent function without effectively narrowing its domain. Although satisfying experimental results can be obtained in certain circumstances, this derivation is not mathematically rigorous. According to [31], in which the actor-critic structure is adopted, this paper replaces this Lipschitz constant using meticulously mathematical transformations with the SCN architecture.

In addition, the initial admissible control is a requirement for both time-based and event-triggered ADP methods, which weakens their application, especially for closed-loop online learning control. Inspired by [9,32,29], an improved weight updating rule is designed by adding a stabilizing term based on the Lyapunov stability theory, such that the requirement of initial admissible control is eliminated.

The contributions of this paper are summarized as follows:

1. It is the first time that an ADP-based controller is developed for a nonlinear aeroelastic system. This paper develops a general control method that can be applied directly without making coordinate transformations.
2. A novel triggering-condition incorporating input constraints is derived without requiring the Lipschitz assumption on the inverse hyperbolic tangent function.
3. The demand for the initial admissible control is relaxed by an improved critic weight updating criterion.
4. The Zeno phenomenon is analysed and avoided regarding the closed-loop system with the event-triggered control strategy.

The remainder of this paper is organized as follows: Section 2 states the constrained-input optimal control problem for a CT nonlinear aeroelastic system under the event-triggered framework. Section 3 provides the implementation of the event-triggered controller using ANN, and analyses the closed-loop stability as well as avoids the Zeno phenomenon. The simulation verification is presented in Section 4, and Section 5 summarizes this paper and states further research.

The main notations used in what follows are listed. \mathbb{N} is the set of all natural numbers. \mathbb{R} denotes the set of all real numbers. \mathbb{R}^n indicates the Euclidean space of all n -dimensional real vectors. $\mathbb{R}^{n \times m}$ is the space of all $n \times m$ real matrices. $|\cdot|$ is the scalar absolute value and $\|\cdot\|$ is the norm of the corresponding vector or matrix. $(\cdot)^-$ denotes the left continuity and $(\cdot)^T$ represents the transpose operation. I_n denotes the $n \times n$ identity matrix and $\mathbf{1}$ is a column vector with all elements equal to one. $\bar{\lambda}(\cdot)$ and $\underline{\lambda}(\cdot)$ respectively represents the maximal and minimal eigenvalues of a matrix. Denote Ω as a compact subset of \mathbb{R}^n , Ω_u as a compact

subset of \mathbb{R}^m , and $\mathcal{A}(\Omega)$ as the set of admissible controllers on Ω . The symbol $\nabla(\cdot) \triangleq \partial(\cdot)/\partial x$ stands for the gradient operator.

2. Problem description

A typical aeroelastic wing section plant with two degrees of freedom is modeled in this section. Then, we describe the constrained-input optimal control problem of general nonlinear systems, and present the event-triggered control mechanism.

2.1. Aeroelastic wing section model

With the wide usage of composite materials, high aspect-ratio aircraft wing can suffer from aeroelastic instability phenomena, including the LCOs [3,33]. If not suppressed by active control, LCOs can lead to structural failure and even flight accidents [2]. The schematic of an aeroelastic wing section controlled by a single trailing-edge flap is illustrated in Fig. 1 [4], where c.m. is the abbreviation of center of mass. It has two degrees of freedom: the plunge displacement h and the pitch angle θ . In this problem, it is assumed in the undisturbed case, that the freestream is along the airfoil chord, and thus pitch angle θ is equal to angle of attack α . Consequently, the governing expressions of motion are presented as [4,6]:

$$\begin{bmatrix} m_t & m_w x_\alpha b \\ m_w x_\alpha b & I_\alpha \end{bmatrix} \begin{bmatrix} \ddot{h} \\ \ddot{\alpha} \end{bmatrix} + \begin{bmatrix} c_h & 0 \\ 0 & c_\alpha \end{bmatrix} \begin{bmatrix} \dot{h} \\ \dot{\alpha} \end{bmatrix} + \begin{bmatrix} k_h(h) & 0 \\ 0 & k_\alpha(\alpha) \end{bmatrix} \begin{bmatrix} h \\ \alpha \end{bmatrix} = \begin{bmatrix} -L \\ M \end{bmatrix}, \quad (1)$$

where $k_h(h)$ and $k_\alpha(\alpha)$ respectively represents the plunge and pitch stiffness, which can be formulated by nonlinear polynomials as [2]:

$$\begin{aligned} k_h &= 2844(1 + 0.9h^2) \\ k_\alpha &= 2.82(1 - 22.1\alpha + 1315.5\alpha^2 - 8580\alpha^3 + 17289.7\alpha^4), \end{aligned} \quad (2)$$

and the remaining constant parameters are listed in Table 1; L and M respectively denotes the aerodynamic force and moment, which are formulated in a quasi-steady form as [4]:

$$\begin{aligned} L &= \rho U^2 b c_{L\alpha} \left(\alpha + \frac{\dot{h}}{U} + \bar{a} b \frac{\dot{\alpha}}{U} \right) + \rho U^2 b c_{L\beta} \beta, \\ M &= \rho U^2 b^2 c_{m\alpha} \left(\alpha + \frac{\dot{h}}{U} + \bar{a} b \frac{\dot{\alpha}}{U} \right) + \rho U^2 b^2 c_{m\beta} \beta, \end{aligned} \quad (3)$$

where $\bar{a} = 0.5 - a$, and β is the control surface deflection. As presented by (1) - (3), the motion dynamics of the aeroelastic system are nonlinear. To describe the system more profoundly, the properties of a simplified linear system are provided. By neglecting the nonlinear terms in (2), the flutter speed of the resulting linear aeroelastic system is 12.41 m/s. The natural frequencies of the corresponding linear undamped aeroelastic system are 9.11 rad/s and 13.28 rad/s.

The complete flight control system often involves the actuator, which can be described as a first-order component [13]:

$$\dot{\beta} - k_\beta \beta = k_\beta \beta_c, \quad (4)$$

where β_c is the deflection command directly generated by the controller. In this case, the complete state vector is $x = [x_1, x_2, x_3, x_4, x_5]^T = [h, \alpha, \dot{h}, \dot{\alpha}, \beta]^T$, and the control input is $u = \beta_c$. Besides, due to mechanical limitations, the control surface deflection always has constraints, which should be taken into consideration in the controller design process.

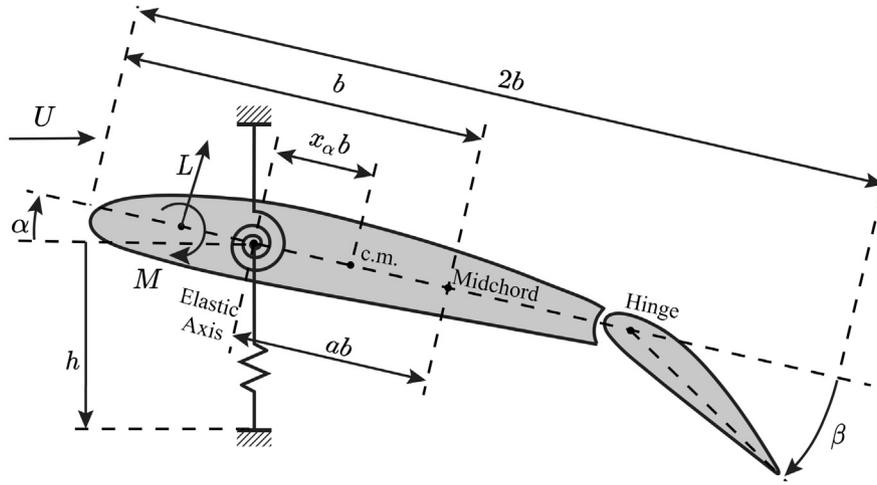


Fig. 1. An two-degree-of-freedom aeroelastic system with one control surface.

Table 1
Constant parameters for the aeroelastic system.

Symbol	Meaning	Value
a	nondimensional distance from the midchord to the elastic axis	-0.6847
b	wing semichord	0.135 m
m_t	total mass of the wing and the flap	12.387 kg
m_w	mass of the wing	2.049 kg
x_α	nondimensional distance from the elastic axis to the c.m.	$(0.0873 - (b + ab))/b$
I_α	moment of inertia	$m_w x_\alpha^2 b^2 + 0.0517$ kg m
c_h	plunge damping coefficient	27.43 kg/s
c_α	pitch damping coefficient	0.036 kg m ² /s
$c_{L\alpha}$	lift coefficient per α	6.28
$c_{m\alpha}$	moment coefficient per α	$(0.5 + a)c_{L\alpha}$
$c_{L\beta}$	lift coefficient per β	3.358
$c_{m\beta}$	moment coefficient per β	-1.94
ρ	density of air	1.225 kg/m ³
U	freestream velocity	15 m/s

2.2. Optimal control design with input constraints

To provide a general description, we consider a class of nonlinear CT systems formulated by:

$$\dot{x}(t) = f(x(t)) + g(x(t))u(x(t)), \quad (5)$$

where $x(t) \in \Omega \subset \mathbb{R}^n$ is the state vector and $u(x(t)) \in \Omega_u$ is the control signal vector, and $\Omega_u = \{u | u \in \mathbb{R}^m, |u_i| < u_b, i = 1, \dots, m\}$, in which u_b is the control saturating bound. $f(\cdot)$ is Lipschitz continuous in Ω satisfying $f(0) = 0$. The initial state at $t = 0$ is $x(0) = x_0$, and $x = 0$ is the equilibrium point of the system. System (5) is generally assumed to be controllable. For simplicity, we denote $x(t)$ by x hereafter.

For system (5), an infinite-horizon cost function can be defined as:

$$J(x) = \int_t^\infty x^T Q x + Y(u) d\tau, \quad (6)$$

where $Q \in \mathbb{R}^{n \times n}$ is positive semi-definite and is set to be a diagonal matrix in this paper, and $Y(u)$ is a positive semi-definite integrand function utilized to handle control input constraints. We denote $U(x, u(x)) = x^T Q x + Y(u)$ as the utility function, and $U(x, u(x))$ satisfies $U(x, u) \geq 0$ and $U(0, 0) = 0$. Inspired by [8,21], we define $Y(u)$ as:

$$\begin{aligned} Y(u) &= 2u_b \int_0^u \tanh^{-T}(v/u_b) R dv \\ &= 2u_b \sum_{i=1}^m \int_0^{u_i} \tanh^{-T}(v_i/u_b) r_i dv_i, \end{aligned} \quad (7)$$

where $\tanh^{-T}(\cdot)$ stands for $(\tanh^{-1}(\cdot))^T$, and $\tanh^{-1}(\cdot)$ is the inverse hyperbolic tangent function, which is a monotonic odd function; $R = \text{diag}([r_1, \dots, r_m]) \in \mathbb{R}^{m \times m}$ is a positive definite weight matrix, where $\text{diag}(\cdot)$ reshapes the vector to a diagonal matrix. $Y(u)$ satisfies $Y(u) \geq 0$, and only when $u = 0$, $Y(u) = 0$.

Admissible control is a prerequisite of optimal feedback stabilization, such that the cost function $J(x)$ is guaranteed to be finite. Choosing an admissible control law $u(x) \in \mathcal{A}(\Omega)$, and accordingly the Hamiltonian is defined as:

$$H(x, u(x), \nabla J(x)) = U(x, u(x)) + \nabla J^T(x)[f(x) + g(x)u(x)]. \quad (8)$$

The optimal value of the cost function given in (6) is:

$$J^*(x) = \min_{u(x) \in \mathcal{A}(\Omega)} \int_t^\infty U(x(\tau), u(x(\tau))) d\tau, \quad (9)$$

and it satisfies the HJB equation:

$$0 = \min_{u(x) \in \mathcal{A}(\Omega)} H(x, u(x), \nabla J^*(x)). \quad (10)$$

By making $\partial H(x, u(x), \nabla J(x))/\partial u(x) = 0$, the corresponding optimal feedback control solution is derived by:

$$\begin{aligned} u^*(x) &= \arg \min_{u(x) \in \mathcal{A}(\Omega)} H(x, u(x), \nabla J^*(x)) \\ &= -u_b \tanh(D^*), \end{aligned} \quad (11)$$

where $\tanh(\cdot)$ denotes the hyperbolic tangent function, and D^* is given by:

$$D^* = \frac{1}{2u_b} R^{-1} g^T(x) \nabla J^*(x). \quad (12)$$

The control input u^* is bounded by u_b , and the nonquadratic cost (7) regarding u^* is:

$$Y(u^*(x)) = u_b \nabla J^{*T}(x) g(x) \tanh(D^*) + u_b^2 \underline{R} \ln(\mathbf{1} - \tanh^2(D^*)), \quad (13)$$

where $\nabla J^{*\top}(x)$ denotes $(\nabla J^*(x))^\top$ and $\underline{R} = [r_1, \dots, r_m]^\top$. Substituting (11) and (12) into the HJB equation produces:

$$0 = x^\top Q x + u_b^2 \underline{R} \ln(\underline{1} - \tanh^2(D^*)) + \nabla J^{*\top}(x) f(x), \quad (14)$$

with $J^*(0) = 0$ that leads to $H(x, u^*(x), \nabla J^*(x)) = 0$.

2.3. Event-triggered scheme design

Considering the event-triggered scheme, we define a sequence of triggering instants $\{s_k\}_{k=0}^\infty$, where s_k satisfies $s_k < s_{k+1}$ with $k \in \mathbb{N}$. The output of the sampled-data module is $x(s_k) \triangleq x_k$ for all $t \in [s_k, s_{k+1})$. Subsequently, we define the gap function using the event error:

$$e_k(t) = x_k - x, \forall t \in [s_k, s_{k+1}). \quad (15)$$

We denote $e_k(t)$ briefly by e_k hereafter. Everytime when a certain triggering condition is satisfied, the event-triggered state vector is updated and the event error e_k is reset to zero. At every triggering instant (instead of time instant), the state feedback control law $u(x(s_k)) = u(x_k)$ is accordingly updated. By introducing a zero-order holder (ZOH), the control sequence $\{u(x_k)\}_{k=0}^\infty$ actually turns to be a piecewise signal that remains constant during the time interval $[s_k, s_{k+1})$, $\forall k \in \mathbb{N}$. Based on the control signal $u(x_k)$, system (5) takes the form:

$$\dot{x} = f(x) + g(x)u(x + e_k), \forall t \in [s_k, s_{k+1}). \quad (16)$$

Considering the event-triggered framework, combined with (12), the feedback control function (11) becomes:

$$u^*(x_k) = -u_b \tanh(D_k^*), \quad (17)$$

where D_k^* is given as:

$$D_k^* = \frac{1}{2u_b} R^{-1} g^\top(x_k) \nabla J^*(x_k). \quad (18)$$

For system (5), with the infinite-horizon cost function represented by (6), we define a triggering condition as follows:

$$\|e_k\|^2 > \|e_T\|^2, \quad (19)$$

where e_T is the threshold to be determined. We say the event is triggered if (19) is satisfied, and in the following section, we present the details of how to determine the threshold.

3. Intelligent critic control implementation

Since (14) is a nonlinear partial differential equation intractable to be solved analytically, in this section an ANN with an improved updating rule is used to approximate the optimal control policy. Then, the system stability is analysed and the Zeno phenomenon is avoided regarding the closed-loop system.

3.1. Improved neural control implementation

In light of the powerful generalization property of ANNs, the optimal cost function can be reconstructed as follows:

$$J^*(x) = w_c^\top \sigma_c(x) + \varepsilon_c(x), \quad (20)$$

where $w_c \in \mathbb{R}^{l_c}$ stands for the ideal weight, l_c is the number of neurons, $\sigma_c(x) \in \mathbb{R}^{l_c}$ denotes the activation function, and $\varepsilon_c(x) \in \mathbb{R}$ represents the neural approximation error. Accordingly, the gradient vector of the optimal cost is:

$$\nabla J^*(x) = \nabla \sigma_c^\top(x) w_c + \nabla \varepsilon_c(x). \quad (21)$$

Since the ideal weight vector is unavailable in advance, a critic network is constructed to approximate the cost function with an estimated weight vector $\hat{w}_c \in \mathbb{R}^{l_c}$ such that:

$$\hat{J}^*(x) = \hat{w}_c^\top \sigma_c(x). \quad (22)$$

Similarly, we determine:

$$\nabla \hat{J}^*(x) = \nabla \sigma_c^\top(x) \hat{w}_c. \quad (23)$$

Considering the ANN formulation (21), (18) can be rewritten as:

$$D_k^* = \frac{1}{2u_b} [R^{-1} g^\top(x_k) (\nabla \sigma_c^\top(x_k) w_c + \nabla \varepsilon_c(x))]. \quad (24)$$

Based on the mean-value theorem [21], we can accordingly rewrite (17) as:

$$u^*(x_k) = -u_b \tanh(D_k) + \varepsilon_{u_k^*}, \quad (25)$$

where $D_k = 1/(2u_b)(R^{-1} g^\top(x_k) \nabla \sigma_c^\top(x_k) w_c)$, and $\varepsilon_{u_k^*} = -1/2(\underline{1} - \tanh^2(\xi)) R^{-1} g^\top(x_k) \nabla \varepsilon$, where $\xi \in \mathbb{R}^m$ is selected between D_k and D_k^* .

Hence, according to (17) and (23), the event-triggered approximate optimal policy can be formulated as:

$$\hat{u}(x_k) = -u_b \tanh(\hat{D}_k), \quad (26)$$

and \hat{D}_k is modulated as:

$$\hat{D}_k = \frac{1}{2u_b} (R^{-1} g^\top(x_k) \nabla \sigma_c^\top(x_k) \hat{w}_c). \quad (27)$$

Substituting (25) into the Hamiltonian (8) yields that:

$$H(x, u^*(x_k), w_c) = w_c^\top \nabla \sigma_c(x) [f(x) + g(x)u^*(x_k)] + x^\top Q x + Y(u^*(x_k)) \triangleq e_{cH}, \quad (28)$$

where $u^*(x_k) = u_b \tanh(D_k^*)$, and $e_{cH} = -\nabla \varepsilon_c^\top(x) [f(x) + g(x)u^*(x_k)]$ is the residual error brought by the ANN. Utilizing (23), the approximate Hamiltonian is presented as:

$$\hat{H}(x, u^*(x_k), \hat{w}_c) = \hat{w}_c^\top \nabla \sigma_c(x) [f(x) + g(x)u^*(x_k)] + x^\top Q x + Y(u^*(x_k)) \triangleq e_c. \quad (29)$$

Defining the critic error vector as $\tilde{w}_c = w_c - \hat{w}_c$ and combining (28) with (29), we obtain an equivalent expression of e_c :

$$e_c = e_{cH} - \tilde{w}_c^\top \nabla \sigma_c(x) [f(x) + g(x)u^*(x_k)]. \quad (30)$$

Hence, the aim of training the critic network is to obtain an appropriate weight vector \hat{w}_c such that the objective function $E_c = (1/2)e_c^\top e_c$ is minimum. It is worth mentioning that the actual control law utilized during the learning process is the approximated control (26). In [30], a direct gradient-descent method is applied to adjust critic weight vector:

$$\dot{\hat{w}}_{c, \text{trad}} = -\eta_c \frac{\partial e_c}{\partial \hat{w}_c} e_c = -\eta_c \phi e_c, \quad (31)$$

where $\eta_c > 0$ is the learning rate parameter, and $\phi = \nabla \sigma_c(x) [f(x) + g(x)\hat{u}(x_k)]$.

The admissible control is essential for the general ADP-based optimal control design but is intractable to obtain in advance. To overcome this challenge, inspired by [9,29,32], we bring in an extra stabilizing term to improve the direct gradient-descent method

and adopt it to enhance the ANN weight updating. Similar to [9,21,32], we make the following assumption:

Assumption 1. Consider system (5) with the cost function (6) and its closed-loop form governed by the event-triggered optimal controller (17) and (24). Let $J_s(x)$ be a continuously differentiable Lyapunov function candidate satisfying:

$$\dot{J}_s^*(x) = \nabla J_s^T(x)[f(x) + g(x)u^*(x_k)] < 0.$$

Then, there exists a positive definite matrix $M \in \mathbb{R}^{n \times n}$ such that the following inequality holds:

$$\dot{J}_s^*(x) = \nabla J_s^T(x)M\nabla J_s(x) \leq -\underline{\lambda}(M)\|\nabla J_s(x)\|^2. \quad (32)$$

When adopting the event-triggered approximate optimal control (26), we should exclude the following case to guarantee the system stability:

$$\dot{J}_s(x) = \nabla J_s^T(x)[f(x) + g(x)\hat{u}(x_k)] > 0.$$

Hence, the learning performance is reinforced by adjusting the time derivative of $J_s(x)$ along the direction of the negative gradient, which is modulated as follows:

$$\begin{aligned} \dot{\hat{w}}_{c,stab} &= -\eta_s \frac{\partial \nabla J_s^T(x)[f(x) + g(x)\hat{u}(x_k)]}{\partial \hat{w}_c} \\ &= \frac{1}{2}\eta_s \nabla \sigma(x_k)g(x_k)R^{-1}(\mathbf{1} - \tanh^2(\hat{D}_k))g^T(x)\nabla J_s(x), \end{aligned} \quad (33)$$

where $\eta_s > 0$ is the designed learning rate. By combining the stabilizing term (33) and the traditional rule (31), we established the improved ANN learning criterion as follows:

$$\dot{\hat{w}}_c = \dot{\hat{w}}_{c,trad} + \Xi(x, \hat{u}(x_k))\dot{\hat{w}}_{c,stab}, \quad (34)$$

where $\Xi(x, \hat{u}(x_k))$ is a sign function utilized to eliminate the effect of the reinforced term when the system is already stable, which is defined as:

$$\Xi(x, \hat{u}(x_k)) = \begin{cases} 0, & \text{when } \dot{J}_s(x) < 0, \\ 1, & \text{elsewhere.} \end{cases} \quad (35)$$

Remark 1. The improved updating rule (34) with the reinforced term relaxes the demand for initial admissible control, which implies that the critic weight vector can initially be set as any random vector.

3.2. Closed-loop stability analysis

We firstly construct the error dynamics of the critic network by defining $\tilde{w}_c = w_c - \hat{w}_c$ and finding that $\dot{\tilde{w}}_c = -\dot{\hat{w}}_c$. Consequently, the critic error dynamics is presented as:

$$\begin{aligned} \dot{\tilde{w}}_c &= -\eta_c \phi(\phi^T \tilde{w}_c - e_{cH}) - \frac{1}{2}\eta_s \Xi(x, \hat{u}(x_k))\nabla \sigma(x_k)g(x_k) \\ &\quad \times R^{-1}(\mathbf{1} - \tanh^2(\hat{D}_k))g^T(x)\nabla J_s(x). \end{aligned} \quad (36)$$

Remark 2. The persistent excitation (PE) assumption is required. If the PE condition holds, we easily derive $\underline{\lambda}(\phi\phi^T) > 0$ [26], which is of great significance for stability analysis. A common approach to achieve PE is introducing a probing noise to excite the system [29,13,32].

Subsequently, we study the closed-loop stability based on the approximate event-triggered feedback control incorporating the weight estimation dynamics. Before proceeding, the following assumptions are required, which are commonly employed in ADP literature, such as [21,23,9,29].

Assumption 2. $g(x)$ is Lipschitz continuous rendering $\|g(x) - g(x_k)\| \leq L_g\|e_k\|$, and is upper bounded as $\|g(x)\| \leq b_g$, where L_g and b_g are positive real constants.

Assumption 3. Denote $L_{\nabla\sigma_c}$, $b_{\nabla\sigma_c}$, $b_{\nabla\varepsilon_c}$, $b_{\varepsilon_{u^*}}$, and $b_{e_{cH}}$ as positive real constants. $\nabla\sigma_c(x)$ is Lipschitz continuous guaranteeing $\|\nabla\sigma_c(x) - \nabla\sigma_c(x_k)\| \leq L_{\nabla\sigma_c}\|e_k\|$. $\nabla\sigma_c(x)$, $\nabla\varepsilon(x)$, ε_{u^*} , and e_{cH} are all upper bounded, such that $\|\nabla\sigma_c(x)\| \leq b_{\nabla\sigma_c}$, $\|\nabla\varepsilon(x)\| \leq b_{\nabla\varepsilon_c}$, $\|\varepsilon_{u^*}\| \leq b_{\varepsilon_{u^*}}$, and $|e_{cH}| \leq b_{e_{cH}}$.

Theorem 1. Considering Assumptions 1-3, utilizing the formula (34) to update the critic network, and with the event-triggered approximate optimal control policy (26), the closed-loop system is asymptotically stable while the weight error dynamics is ultimately uniformly bounded (UUB) if

$$\|e_k\|^2 \leq \frac{(1-\eta)\underline{\lambda}(Q)\|x\|^2 + Y(u^*(x_k))}{C_1\|\hat{w}_c\|^2} \triangleq \|\hat{e}_T\|^2, \quad (37)$$

and $\|\tilde{w}_c\|^2 \geq \mathcal{W}$, where $\eta \in (0, 1)$, C_1 and \mathcal{W} are given in the proof.

Proof. We construct a Lyapunov function candidate as:

$$\mathcal{L} = L_x + L_{x_k} + L_{\tilde{w}_c} + L_{J_s}, \quad (38)$$

where $L_x = J^*(x)$, $L_{x_k} = J^*(x_k)$, $L_{\tilde{w}_c} = (1/2)\tilde{w}_c^T\tilde{w}_c$, and $L_{J_s} = \eta_s J_s(x)$. The proof consists of two situations conforming to whether the event is triggered or not.

Situation 1: the events are not triggered, i.e., $\forall t \in [s_k, s_{k+1})$. Computing the time derivative of the Lyapunov function, the second term is $\dot{L}_{x_k} = 0$.

Considering the closed loop system using the approximate feedback control (26), and the optimal HJB equation (14), the first term can be derived as:

$$\begin{aligned} \dot{L}_x &= \dot{J}^*(x) = \nabla J^{*T}(x)[f(x) + g(x)\hat{u}(x_k)] \\ &= -x^T Q x - u_b^2 R \ln(\mathbf{1} - \tanh^2(D^*)) + \nabla J^{*T}(x)g(x)\hat{u}(x_k). \end{aligned} \quad (39)$$

According to the definition of utility function, the second term of (39) is converted into:

$$\begin{aligned} u_b^2 \bar{R} \ln(\mathbf{1} - \tanh^2(D^*)) &= \int_{\hat{u}(x_k)}^{u^*(x)} 2u_b \tanh^{-1}(v/u_b) R dv + Y(\hat{u}(x_k)) \\ &\quad - u_b \nabla J^{*T}(x)g(x) \tanh(D^*). \end{aligned} \quad (40)$$

Besides, (11) and (12) imply that:

$$\nabla J^{*T}(x_k)g(x) = -2u_b \tanh^{-1}(u^*(x_k)/u_b)R. \quad (41)$$

Therefore, the last term in (39) can be rewritten as:

$$\begin{aligned} \nabla J^{*T}(x)g(x)\hat{u}(x_k) &= \int_{u^*(x)}^{\hat{u}(x_k)} 2u_b D^{*T}(x) R dv \\ &\quad - u_b \nabla J^{*T}(x)g(x) \tanh(D^*(x)). \end{aligned} \quad (42)$$

Substituting (40) and (42) into (39) yields:

$$\dot{\mathcal{L}}_x = -x^T Q x - Y(\hat{u}(x_k)) + \int_{u^*(x)}^{\hat{u}(x_k)} 2u_b[\tanh^{-1}(v/u_b) + D^*]^T R dv. \quad (43)$$

Letting $v = -u_b \tanh(\omega)$, the last term in (43) is written as:

$$\begin{aligned} & \int_{u^*(x)}^{\hat{u}(x_k)} 2u_b[\tanh^{-1}(v/u_b) + D^*]^T R dv \\ &= \int_{D^*}^{\hat{D}_k} 2u_b^2(\omega - D^*)^T R [1 - \tanh^2(\omega)] d\omega \\ &\leq u_b^2(\hat{D}_k - D^*)^T R (\hat{D}_k - D^*). \end{aligned} \quad (44)$$

Therefore, (43) satisfies:

$$\begin{aligned} \dot{\mathcal{L}}_x &\leq -x^T Q x - Y(\hat{u}(x_k)) \\ &\quad + \frac{1}{4}\underline{\lambda}(R) \|\mathbf{g}^T(x_k) \nabla \hat{J}(x_k) - \mathbf{g}^T(x) \nabla J^*(x)\|^2, \end{aligned} \quad (45)$$

where $\nabla \hat{J}(x_k) = \nabla \sigma^T(x_k) \hat{w}_c$. According to Assumptions 2 and 3, we obtain:

$$\begin{aligned} & \|\mathbf{g}^T(x_k) \nabla \hat{J}(x_k) - \mathbf{g}^T(x) \nabla J^*(x)\|^2 \\ &= \|(\mathbf{g}^T(x) \nabla \sigma_c^T(x) - \mathbf{g}^T(x_k) \nabla \sigma_c^T(x_k)) \hat{w}_c + \mathbf{g}^T(x) (\nabla \sigma_c^T(x) \hat{w}_c \\ &\quad + \nabla \varepsilon_c(x))\|^2 \\ &\leq 2\|\nabla \sigma_c(x) \mathbf{g}(x) - \nabla \sigma_c(x_k) \mathbf{g}(x_k)\|^2 \|\hat{w}_c\|^2 \\ &\quad + 2b_g^2 (b_{\nabla \sigma_c}^2 \|\hat{w}_c\|^2 + b_{\varepsilon_c}^2), \end{aligned} \quad (46)$$

in which

$$\begin{aligned} & \|\nabla \sigma_c(x) \mathbf{g}(x) - \nabla \sigma_c(x_k) \mathbf{g}(x_k)\|^2 \\ &\leq 2\|\nabla \sigma_c(x) (\mathbf{g}(x) - \mathbf{g}(x_k))\|^2 + 2\|(\nabla \sigma_c(x) - \nabla \sigma_c(x_k)) \mathbf{g}(x_k)\|^2 \\ &\leq 2(b_{\nabla \sigma_c}^2 L_g^2 + L_{\nabla \sigma_c}^2 b_g^2) \|e_k\|^2. \end{aligned} \quad (47)$$

Therefore (45) continues as:

$$\begin{aligned} \dot{\mathcal{L}}_x &\leq C_1 \|\hat{w}_c\|^2 \|e_k\|^2 + C_2 \|\hat{w}_c\|^2 + C_3 - Y(\hat{u}(x_k)) - \eta \underline{\lambda}(Q) \|x\|^2 \\ &\quad - (1 - \eta) \underline{\lambda}(Q) \|x\|^2, \end{aligned} \quad (48)$$

where $C_1 = \underline{\lambda}(R) (b_{\nabla \sigma_c}^2 L_g^2 + L_{\nabla \sigma_c}^2 b_g^2)$, $C_2 = \frac{1}{2} \underline{\lambda}(R) b_g^2 b_{\nabla \sigma_c}^2$, and $C_3 = \frac{1}{2} \underline{\lambda}(R) b_g^2 b_{\varepsilon_c}^2$.

Then, we investigate the last two terms in (38). By taking the definition of $\Xi(x, \hat{u}(x_k))$ into consideration, two scenarios are examined separately.

I: $\Xi(x, \hat{u}(x_k)) = 0$. We have $\dot{L}_s < 0$ and

$$\dot{L}_{\tilde{w}_c} = -\eta_c \tilde{w}_c^T \phi (\phi^T \tilde{w}_c + e_{cH}) \leq -C_4 \|\tilde{w}_c\|^2 + C_5, \quad (49)$$

where $C_4 = \frac{1}{2} \eta_c \underline{\lambda}(\phi \phi^T)$ and $C_5 = \frac{1}{2} \eta_c e_{cH}^2$. Combining (48) and (49), we observe:

$$\begin{aligned} \dot{\mathcal{L}} &= \dot{L}_x + \dot{L}_{\tilde{w}_c} \\ &\leq C_1 \|\hat{w}_c\|^2 \|e_k\|^2 + C_2 \|\hat{w}_c\|^2 + C_3 - C_4 \|\tilde{w}_c\|^2 + C_5 \\ &\quad - \eta \underline{\lambda}(Q) \|x\|^2 - (1 - \eta) \underline{\lambda}(Q) \|x\|^2 - Y(\hat{u}(x_k)). \end{aligned} \quad (50)$$

By choosing an appropriate η_c , we can achieve $C_4 > C_2$. Therefore, when $\|\tilde{w}_c\|^2 \geq (C_3 + C_5)/(C_4 - C_2) \triangleq \mathcal{W}_1$ and $\|e_k\|^2 < \|\hat{e}_T\|^2$, we have $\dot{\mathcal{L}} < 0$, $\forall x \neq 0$.

II: $\Xi(x, \hat{u}(x_k)) = 1$. We combine \dot{L}_{J_s} with the stabilization term of $\dot{L}_{\tilde{w}_c}$ as:

$$\begin{aligned} \dot{L}' &= \dot{L}_{J_s} + \dot{L}'_{\tilde{w}_c} \\ &= \eta_s \nabla J_s^T(x) [f(x) + g(x) \hat{u}(x_k)] \\ &\quad - \frac{1}{2} \eta_s \tilde{w}_c^T \nabla \sigma(x_k) g(x_k) R^{-1} (\mathbf{1} - \tanh^2(\hat{D}_k)) \mathbf{g}^T(x) \nabla J_s(x) \\ &= \eta_s \nabla J_s^T(x) f(x) - \eta_s \nabla J_s^T(x) g(x) [u_b \tanh(\hat{D}_k) \\ &\quad + \frac{1}{2} (\mathbf{1} - \tanh^2(\hat{D}_k)) R^{-1} \mathbf{g}^T(x_k) \sigma^T(x_k) \tilde{w}_c]. \end{aligned} \quad (51)$$

By taking the first-order Taylor series expansion of $\tanh(D_k)$, we obtain:

$$\begin{aligned} \tanh(D_k) &= \frac{1}{2u_b} (\mathbf{1} - \tanh^2(\hat{D}_k)) R^{-1} \mathbf{g}^T(x_k) \sigma^T(x_k) \tilde{w}_c \\ &\quad + \tanh(\hat{D}_k) + o[(D_k - \hat{D}_k)^2], \end{aligned} \quad (52)$$

where $o[(D_k - \hat{D}_k)^2]$ has a bound, which is denoted by b_{oD} [21]. Substituting (52) into (25) and then (51) yields:

$$\begin{aligned} \dot{L}' &= \eta_s \dot{J}_s^*(x) - \eta_s \nabla J_s^T(x) g(x) \varepsilon_{u^*} + \eta_s \nabla J_s^T(x) g(x) u_b o[(D_k - \hat{D}_k)^2] \\ &\leq -\eta_s \underline{\lambda}(M) \|\nabla J_s(x)\|^2 + \eta_s C_6 \|\nabla J_s(x)\| \\ &= -\eta_s \underline{\lambda}(M) \left(\|\nabla J_s(x)\| - \frac{C_6}{2\underline{\lambda}(M)} \right)^2 + C_7, \end{aligned} \quad (53)$$

where $C_6 = b_g (b_{\varepsilon_{u^*}} + u_b b_{oD})$, and $C_7 = \eta_s C_6^2 / (4\underline{\lambda}(M))$. Accordingly, we obtain:

$$\begin{aligned} \dot{\mathcal{L}} &= \dot{L}_x + \dot{L}_{\tilde{w}_c} + \dot{L}_{J_s} \\ &\leq -\eta \underline{\lambda}(Q) \|x\|^2 - (1 - \eta) \underline{\lambda}(Q) \|x\|^2 - Y(\hat{u}(x_k)) \\ &\quad + C_1 \|\hat{w}_c\|^2 \|e_k\|^2 + C_2 \|\hat{w}_c\|^2 + C_3 - C_4 \|\tilde{w}_c\|^2 \\ &\quad + C_5 - \eta_s \underline{\lambda}(M) \left(\|\nabla J_s(x)\| - \frac{C_6}{2\underline{\lambda}(M)} \right)^2 + C_7. \end{aligned} \quad (54)$$

Therefore, when $\|\tilde{w}_c\|^2 \geq (C_3 + C_5 + C_7)/(C_4 - C_2) \triangleq \mathcal{W}_2$ and $\|e_k\|^2 < \|\hat{e}_T\|^2$, we have $\dot{\mathcal{L}} < 0$, $\forall x \neq 0$.

Consequently, combining I and II, we can conclude that, if the event is not triggered, i.e., $\|e_k\|^2 < \|\hat{e}_T\|^2$, if $\|\tilde{w}_c\|^2 \geq \max\{\mathcal{W}_1, \mathcal{W}_2\} \triangleq \mathcal{W}$ is satisfied, then we have $\dot{\mathcal{L}} < 0$, $\forall x \neq 0$.

Situation 2: the events are triggered, i.e., $\forall t = s_{k+1}$. Utilizing (38), the difference terms are derived as:

$$\begin{aligned} \Delta \mathcal{L} &= \Delta L_x + \Delta L_{x_k} + \Delta L_{\tilde{w}_c} + \Delta L_{J_s} \\ &= J^*(x_{k+1}) - J^*(x(s_{k+1}^-)) + J^*(x_{k+1}) - J^*(x_k) \\ &\quad + \frac{1}{2} \tilde{w}_c^T(x_{k+1}) \tilde{w}_c(x_{k+1}) - \frac{1}{2} \tilde{w}_c^T(x(s_{k+1}^-)) \tilde{w}_c(x(s_{k+1}^-)) \\ &\quad + \eta_s J_s(x_{k+1}) - \eta_s J_s(x(s_{k+1}^-)). \end{aligned} \quad (55)$$

From Situation 1, we find that if $\|e_k\|^2 \leq \|\hat{e}_T\|^2$ and $\|\tilde{w}_c\|^2 \geq \mathcal{W}$, $\dot{\mathcal{L}} < 0$, $\forall t \in [s_k, s_{k+1})$. Since the states and cost function are both continuous, according to the property of the limit, we have $\Delta L_x + \Delta L_{\tilde{w}_c} + \Delta L_{J_s} < 0$ and then obtain $\Delta \mathcal{L} < \Delta L_{x_k} \leq -\kappa(\|e_{k+1} - e_k\|)$, in which $\kappa(\cdot)$ is a class- κ function [34]. Hence, (38) is still decreasing when $\forall t = s_{k+1}$.

In summary, if $\|e_k\|^2 \leq \|\hat{e}_T\|^2$ and $\|\tilde{w}_c\|^2 \geq \mathcal{W}$ hold, we can reach the conclusion that the closed-loop state is asymptotically

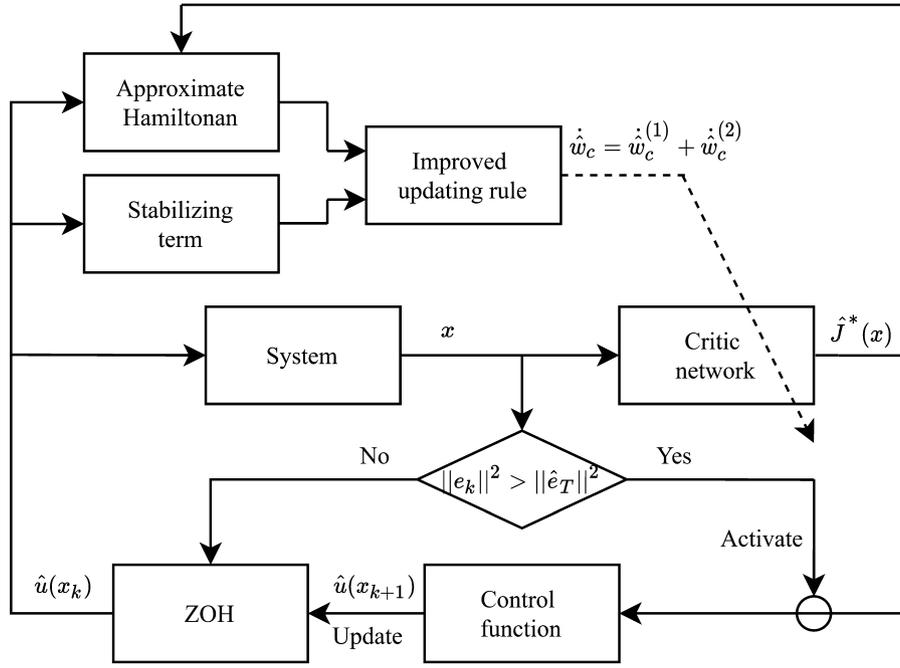


Fig. 2. Structural diagram of the developed event-triggered ADP control method, where solid lines represent the feed-forward flow of signals and the dashed line is the back-propagation path.

stable while the critic error dynamics is UUB, which ends the proof. \square

3.3. Analysis of Zeno phenomenon in the closed-loop system

For nonlinear CT systems with event-triggered control inputs, the inter-execution time is denoted as $\Delta s = s_{k+1} - s_k$, and the minimal inter-execution time $\Delta s_{\min} = \min_{k \in \mathbb{N}} \{s_{k+1} - s_k\}$ might be zero, which can lead to the accumulation of event times, a.k.a., the Zeno phenomenon. Hence, the condition of $\Delta s > 0$ should be guaranteed such that the undesired Zeno phenomenon is avoided.

Theorem 2. Considering the closed-loop form of the nonlinear system (5) governed by the event-triggered approximate optimal control (26), the k -th inter-execution time Δs_k determined by (37) has a lower bound as:

$$\Delta s_k \geq \frac{1}{k_f} \ln \left(\frac{k_f \|\hat{e}_T\|}{k_f \|e_k\| + b_g u_b} + 1 \right) > 0, \quad k \in \mathbb{N}, \quad (56)$$

where k_f is a positive constant.

Proof. We apply the approximate optimal control (26) to formulate the closed-loop dynamics as follows:

$$\dot{x} = f(x) - g(x)\hat{u}(x_k). \quad (57)$$

By noticing the fact that $\hat{u}(x_k)$ is upper bounded by u_b , and according to Assumption 2, we can derive that:

$$\|\dot{x}\| = \|f(x) - g(x)\hat{u}(x_k)\| \leq k_f \|x\| + b_g u_b. \quad (58)$$

Considering (15), we can further derive that:

$$\|\dot{e}_k\| \leq k_f \|x_k\| + k_f \|e_k\| + b_g u_b, \quad \forall t \in [s_k, s_{k+1}). \quad (59)$$

Since $e_k(s_k) = x_k - x(s_k) = 0$, by employing the comparison lemma [34,9] to solve (59), for any $t \in [s_k, s_{k+1})$, we have:

$$\|e_k\| \leq \frac{k_f \|e_k\| + b_g u_b}{k_f} \left(e^{k_f(t-s_k)} - 1 \right). \quad (60)$$

Therefore, the k -th inter-execution time Δs_k satisfies:

$$\Delta s_k = s_{k+1} - s_k \geq \frac{1}{k_f} \ln \left(\frac{k_f \|\hat{e}_T\|}{k_f \|e_k\| + b_g u_b} + 1 \right). \quad (61)$$

According to (37), $\|\hat{e}_T\| > 0$. In summary, $\Delta s_k > 0$ for any $x_k \neq 0$, i.e., $\Delta s_{\min} > 0$, which ends the proof. \square

Overall, the structural diagram of the present control implementation is depicted in Fig. 2 to clarify the design procedure.

4. Simulation study

Finally, we verify the effectiveness of the proposed control approach through the numerical simulation experiments based on the nonlinear aeroelastic system demonstrated above. Considering the cost function (6) from $t = 0$, we choose $Q = I_5$ and $R = 1$ as a trade-off between fast stabilizing and avoiding aggressive control, and set the deflection constraint as $u_b = 10$ deg. For purpose of simulation, we set the simulation frequency as 1 kHz whereas the sensing frequency as 100 Hz. Let the initial state vector be $x_0 = [-0.01 \text{ m}, 10 \text{ deg}, 0, 0, 0]^T$. In what follows, we implement the event-triggered intelligent critic control with the facilitation of an ANN.

A critic network is constructed to approximate the optimal cost function. The number of neurons and the nonlinearity of the activation function positively correlate with the approximation precision. However, more neurons with higher nonlinearities can also increase computational load and cause the overfitting that harms the control robustness [13]. For balancing control accuracy and computational complexity, we choose the activation function as $\sigma_c(x) = [x_1 x_2, x_1 x_3, x_1 x_4, x_1 x_5, x_2 x_3, x_2 x_4, x_2 x_5, x_3 x_4, x_3 x_5, x_4 x_5]^T$. We choose $J_s(x) = 0.5x^T x$ to enhance stability and experimentally set $\eta_c = 0.05$, $\eta_s = 0.001$, $\eta = 0.1$, and $C_1 = 250$. As claimed in Remark 2, an exploration noise u_e is introduced to satisfy the PE condition. The probing noise is designed as a composition of decaying sinusoidal functions, whose formula is $u_e = -0.05e^{-20t}(\sin^2(100t)\cos(100t) + \sin^2(2t)\cos(0.1t) + \sin^2(1.2t) \times \cos(0.5t) + \sin^5(t))$ deg. Only at the instant when it is triggered,

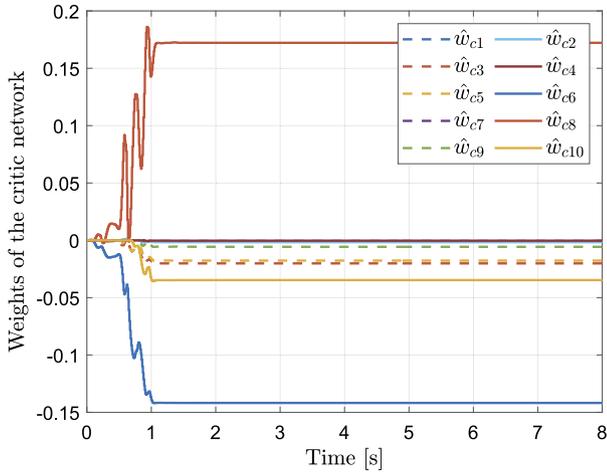


Fig. 3. Convergence process of the critic weights. (For interpretation of the colours in the figure(s), the reader is referred to the web version of this article.)

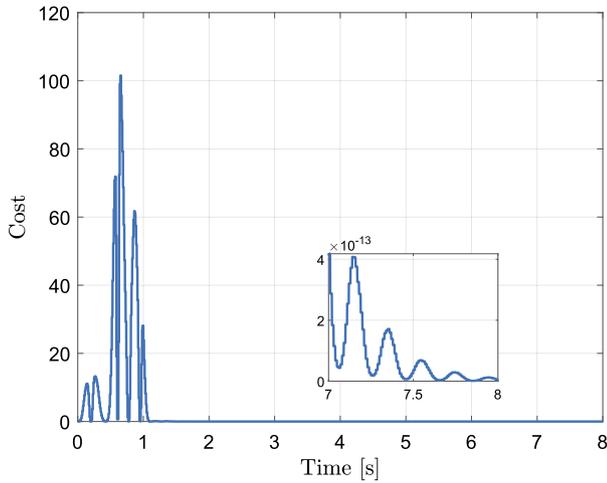


Fig. 4. Evolution of the approximate cost function.

will u_e really be added to the control command. Since the critic network has different hidden neurons, the weights can initially be set as zero. Recalling the triggering condition in (37), we find that $\|\hat{w}_c\|$ appears in the denominator, which can cause a large time interval of control at the beginning. Therefore, we manually set an upper bound for the inter-execution time as $\Delta s_{\max} = 0.1$ s. This configuration is set in the engineering sense for safety guarantee, and does not affect the theoretical completeness.

The simulation is conducted in an online manner, which means that the control policy improves in a closed-loop way. For presenting the advantage of the ETC scheme, a time-based approach is adopted for comparison, whose settings are exactly same as the proposed intelligent critic control approach except for the event-triggered scheme, i.e., the time-based control approach updates the control input at each time instant. We can observe from Fig. 3 that the convergence of the weight vector occurs around 1 s. Subsequently, we display the trajectory of the approximated cost function in Fig. 4, which presents the direct performance of the controller. Due to the initial zero values of the weight vector, the initial approximate cost is zero, and subsequently grows as the learning continues. Then because of the convergence of the weight vector, the approximate cost function swiftly decreases to a low level. Furthermore, the triggering threshold trajectory is displayed in Fig. 5, which presents a trend to zero along with the event error. The inter-execution time is depicted in Fig. 6. It is worth mentioning that 800 samples are utilized by the time-based controller,

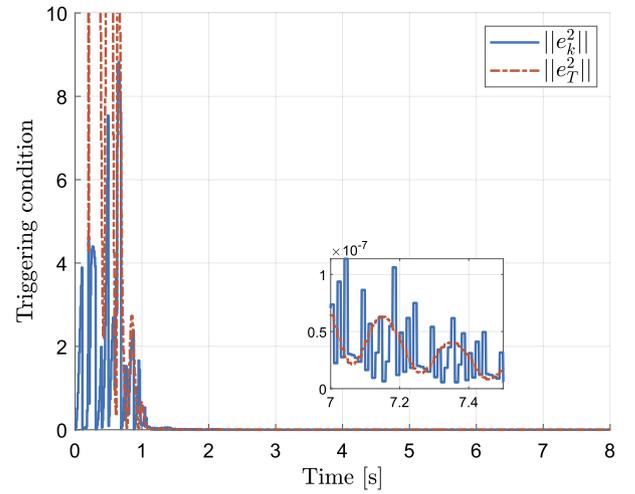


Fig. 5. Evolution of the triggering condition.

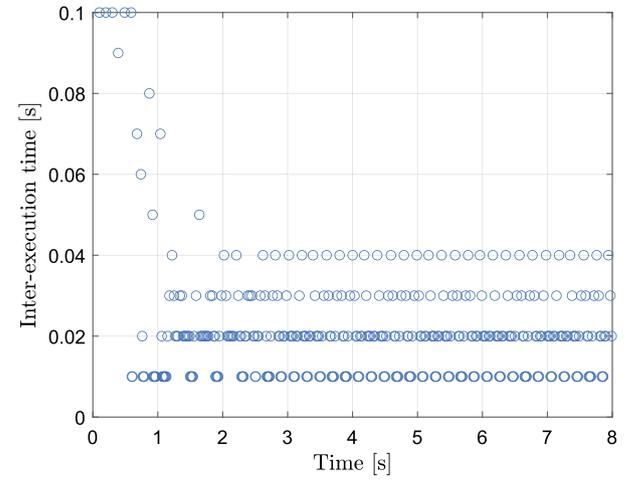


Fig. 6. Evolution of the inter-execution time.

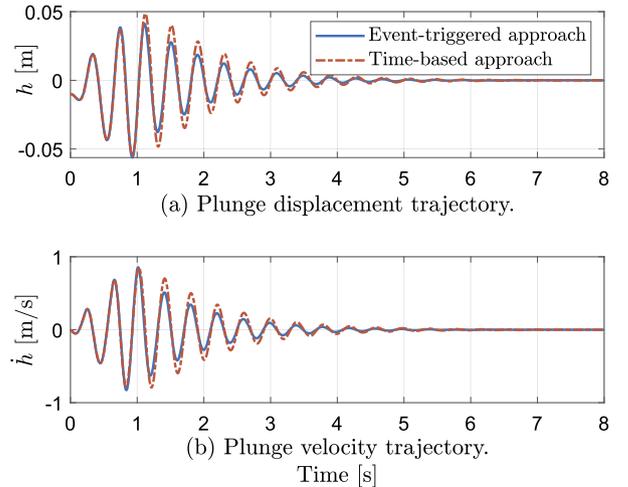


Fig. 7. Evolution of the plunge motion states.

whereas the proposed event-triggered approach only requires 366 samples. Therefore, the event-triggered method reduces the control updates in the learning process up to 54.25%, and thus improves the resource utilization.

Figs. 7 and 8 present the aeroelastic system states trajectory divided into plunge and pitch motion, respectively. We compare the

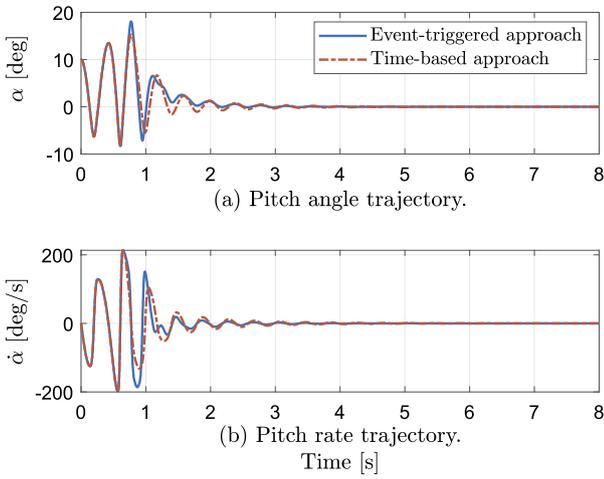


Fig. 8. Evolution of the pitch motion states.

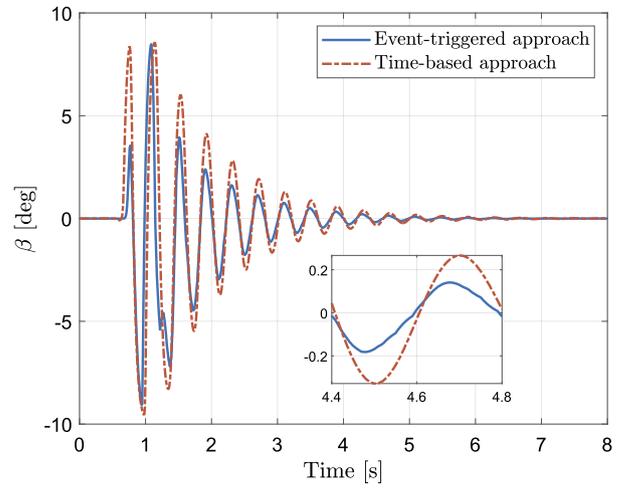


Fig. 10. Real deflection of the control surface.

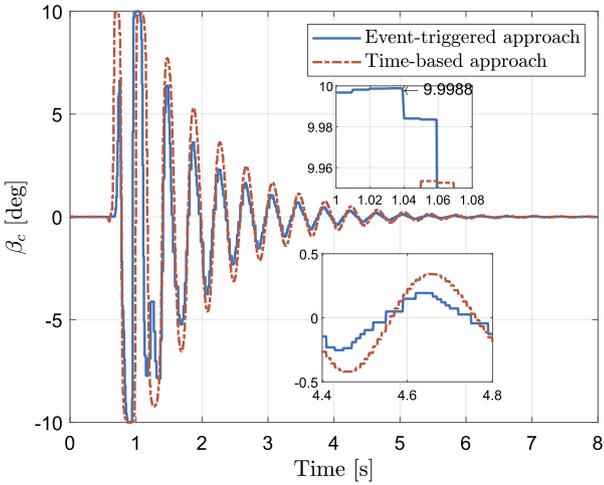


Fig. 9. Control command generated by the controller.

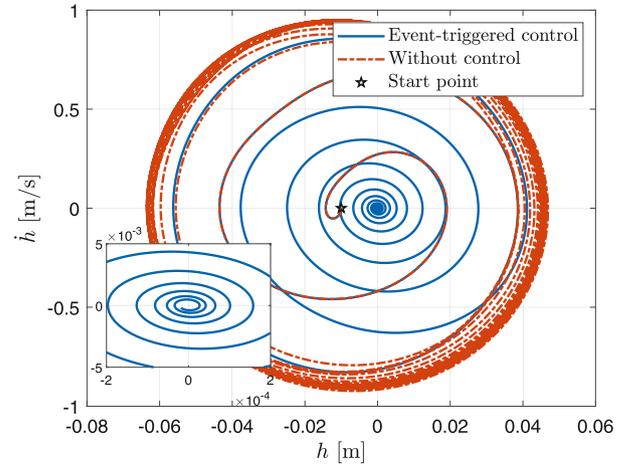


Fig. 11. Phase portrait of the plunge motion states.

results between the event-triggered and time-based approaches, and observe that, although the event-triggered controller utilizes fewer data samples, the state variables eventually converge to a small vicinity of zero without deteriorating the converge rate. Figs. 9 and 10 respectively presents the control command directly generated by the controller, $u(\beta_c)$, and the real deflection of the control surface, $x_5(\beta)$. The developed event-triggered approach has overall comparable curve to the time-based approach. Due to the event-triggered mechanism, the control command signal is stepwise. Nevertheless, the control command signal has to go through an actuator and the real deflection is adequately smooth for the wing surface control. Furthermore, we observe that the control command (incorporating exploration noise) is bounded by the pre-designed saturation constraints, i.e., $|u| < u_b$. Therefore, we conclude that the control input constraints problem has been overcome.

The phase portraits of plunge and pitch motions are illustrated in Figs. 11 and 12, respectively. As can be observed, the trajectories of the proposed method and the open-loop simulation almost coincide at the beginning. This phenomenon is due to the collective effect caused by LCOs and the initial unlearned policy, and disappears quickly as the weight vector updates. Then all states are stabilized to a small vicinity of the equilibrium point.

To further verify its performance, robustness tests are carried out with different freestream velocities using the proposed event-triggered intelligent optimal control strategy. In addition to nomi-

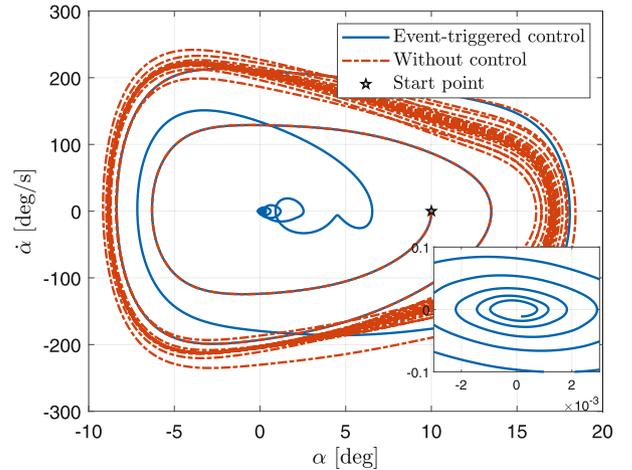


Fig. 12. Phase portrait of the pitch motion states.

nal velocities, the freestream is also assumed to be disturbed. The uncertain freestream is modeled as a composition of a white noise with unit variance and a sinusoidal gust $0.5 \cos(4\pi t + 0.3\pi)$ m/s. With other settings unchanged, the results are depicted in Figs. 13 - 15, in which the nominal freestream velocity U respectively equals to 12 m/s, 15 m/s and 27 m/s. For simplicity, the plunge displacement is selected as a representative to show the state

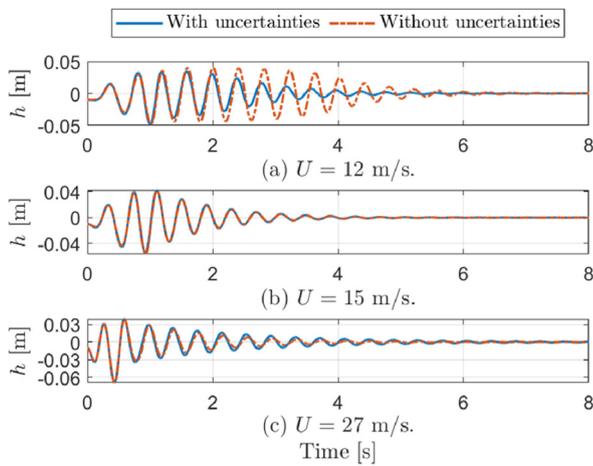


Fig. 13. Evolution of plunge displacements with different freestream velocities.

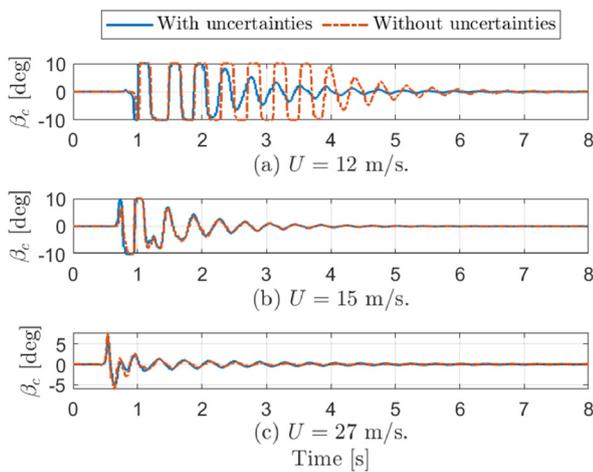


Fig. 14. Control commands generated by the controller with different freestream velocities.

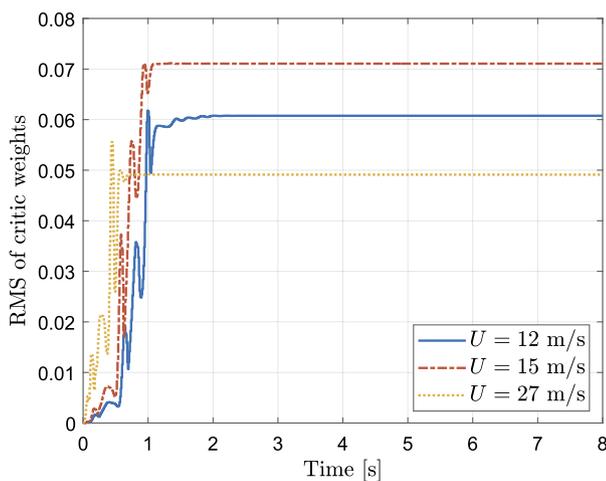


Fig. 15. Convergence process of the RMS of critic weights with different freestream velocities.

evolution, as illustrated in Fig. 13. It can be observed that in all conditions, the controller manages to stabilize the plunge displacement within 8 s. The situations with and without uncertainties demonstrate similar performance for $U = 15$ m/s and $U = 27$ m/s, whereas for $U = 12$ m/s, the control performance is even better with uncertainties involved. The reason behind this phenomenon lies in that when $U = 12$ m/s the flutter frequency is low and

it is difficult to fully excite the system. The velocity uncertainties disturb the system but meanwhile provide stronger excitation and thus speed up the learning process. In fact, this acceleration phenomenon also takes place in the other two situations though it is more obvious with lower freestream velocity, which illustrates the robustness of the proposed control method to uncertainties. With lower incoming flow speed ($U = 12$ m/s), the control effectiveness is also lower, and therefore it requires larger deflection of control surface to generate sufficient control torque, leading to the aggressive but saturated control command shown in Fig. 14 (a). When the freestream velocity is higher ($U = 27$ m/s), the flutter frequency is higher, which provides more excitations at the initial stage. Therefore, it can be seen from Fig. 14 (c) that the control command becomes effective earlier than for the other two conditions. However, it is remarkable that if $U < 12$ m/s or $U > 27$ m/s, the controller is capable for stabilizing the system within 8 s with current settings due to the insufficient control effectiveness or the excessive flutter frequency, respectively. Nevertheless, this can be improved by adapting hyperparameters. Fig. 15 compares the root means square (RMS) of critic weights with 3 different freestream velocities in the presence of uncertainties. Consistent with the above, the convergence speed is faster when the freestream velocity is higher because of the stronger excitation. These curves are different in that the control policy is learned online, and therefore the controller adapts to different conditions in real time, which validates the adaptability of the proposed control approach. The simulation results collectively verify the feasibility and the effectiveness of the event-triggered intelligent optimal control approach developed in this paper.

5. Conclusion

In this paper, we develop an event-triggered intelligent optimal control scheme, and apply it to an aeroelastic system control problem. Taking the input constraints into account, we derive a novel triggering condition without making the Lipschitz assumption on the inverse hyperbolic tangent function. The controller is conducted by adopting the adaptive dynamic programming (ADP) technique with a single critic network.

The theoretical analysis of the closed-loop system shows that, with the derived event-triggered controller, the system states can be guaranteed asymptotically stable, while the Zeno phenomenon is also avoided during the learning phase. The simulation results demonstrate that the nonlinear aeroelastic system is successfully stabilized with input saturation constraints handled. Besides, compared to the conventional time-based ADP method, the present event-triggered ADP method can achieve comparable performance with reduced control updates, which presents the advantages of the developed method in saving the computational and communication's load. Furthermore, the robustness tests demonstrate that the designed controller is able to adapt online to different situations and has the robustness to uncertainties to some extent.

At the current stage, we concentrate on the control algorithm development based on the known system dynamics and perfect measurements. Due to the fact that uncertainties generally exist in the real world, further investigation into robust control methods is recommended. Besides, due to the limitation of actuator power, the deflection rate of the control surface should also be constrained, which can be studied in the future.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

- [1] J. Yuan, N. Qi, Z. Qiu, F. Wang, Adaptive rbf observer-sliding mode controller design for a two dimensional aeroelastic system with unsteady aerodynamics, *Aerosp. Sci. Technol.* 80 (2018) 482–495.
- [2] K.W. Lee, S.N. Singh, Robust finite-time continuous control of an unsteady aeroelastic system, *J. Guid. Control Dyn.* 41 (4) (2018) 978–986.
- [3] D. Li, J. Xiang, S. Guo, Adaptive control of a nonlinear aeroelastic system, *Aerosp. Sci. Technol.* 15 (5) (2011) 343–352.
- [4] H. Lhachemi, Y. Chu, D. Saussié, G. Zhu, Flutter suppression for underactuated aeroelastic wing section: nonlinear gain-scheduling approach, *J. Guid. Control Dyn.* 40 (8) (2017) 2102–2109.
- [5] W. Su, W. Song, A real-time hybrid aeroelastic simulation platform for flexible wings, *Aerosp. Sci. Technol.* 95 (2019) 105513.
- [6] T.W. Strganac, J. Ko, D.E. Thompson, A.J. Kurdila, Identification and control of limit cycle oscillations in aeroelastic systems, *J. Guid. Control Dyn.* 23 (6) (2000) 1127–1133.
- [7] B. Sun, E.-J. van Kampen, Reinforcement-learning-based adaptive optimal flight control with output feedback and input constraints, *J. Guid. Control Dyn.* 44 (9) (2021) 1685–1691.
- [8] A. Heydari, S.N. Balakrishnan, Finite-horizon control-constrained nonlinear optimal control using single network adaptive critics, *IEEE Trans. Neural Netw. Learn. Syst.* 24 (1) (2013) 145–157.
- [9] D. Wang, H. He, D. Liu, Improving the critic learning for event-based nonlinear H_∞ control design, *IEEE Trans. Cybern.* 47 (10) (2017) 3417–3428.
- [10] B. Kiumarsi, K.G. Vamvoudakis, H. Modares, F.L. Lewis, Optimal and autonomous control using reinforcement learning: a survey, *IEEE Trans. Neural Netw. Learn. Syst.* 29 (6) (2018) 2042–2062.
- [11] C. Peng, J. Ma, Online integral reinforcement learning control for an uncertain highly flexible aircraft using state and output feedback, *Aerosp. Sci. Technol.* (2020) 106442.
- [12] Y. Zhou, E.-J. van Kampen, Q. Chu, Incremental model based online heuristic dynamic programming for nonlinear adaptive tracking control with partial observability, *Aerosp. Sci. Technol.* 105 (2020) 106013.
- [13] B. Sun, E.-J. van Kampen, Incremental model-based global dual heuristic programming with explicit analytical calculations applied to flight control, *Eng. Appl. Artif. Intell.* 89 (2020) 103425.
- [14] X. Han, Z. Zheng, L. Liu, B. Wang, Z. Cheng, H. Fan, Y. Wang, Online policy iteration adp-based attitude-tracking control for hypersonic vehicles, *Aerosp. Sci. Technol.* 106 (2020) 106233.
- [15] Y. Zhou, E.-J. van Kampen, Q.P. Chu, Incremental model based online dual heuristic programming for nonlinear adaptive control, *Control Eng. Pract.* 73 (2018) 13–25.
- [16] C. Liu, C. Dong, Z. Zhou, Z. Wang, Barrier Lyapunov function based reinforcement learning control for air-breathing hypersonic vehicle with variable geometry inlet, *Aerosp. Sci. Technol.* 96 (2020) 105537.
- [17] B. Sun, E.-J. van Kampen, Intelligent adaptive optimal control using incremental model-based global dual heuristic programming subject to partial observability, *Appl. Soft Comput.* 103 (2021) 107153.
- [18] R.S. Sutton, A.G. Barto, *Reinforcement Learning: An Introduction*, MIT Press, 2018.
- [19] H. Dong, X. Zhao, B. Luo, Optimal tracking control for uncertain nonlinear systems with prescribed performance via critic-only adp, *IEEE Trans. Syst. Man Cybern. Syst.* (2020), <https://doi.org/10.1109/TSMC.2020.3003797> (early access).
- [20] J. Na, B. Wang, G. Li, S. Zhan, W. He, Nonlinear constrained optimal control of wave energy converters with adaptive dynamic programming, *IEEE Trans. Ind. Electron.* 66 (10) (2018) 7904–7915.
- [21] D. Liu, X. Yang, D. Wang, Q. Wei, Reinforcement-learning-based robust controller design for continuous-time uncertain nonlinear systems subject to input constraints, *IEEE Trans. Cybern.* 45 (7) (2015) 1372–1385.
- [22] D.V. Dimarogonas, E. Frazzoli, K.H. Johansson, Distributed event-triggered control for multi-agent systems, *IEEE Trans. Autom. Control* 57 (5) (2012) 1291–1297.
- [23] V. Narayanan, H. Modares, S. Jagannathan, F.L. Lewis, Event-driven off-policy reinforcement learning for control of interconnected systems, *IEEE Trans. Cybern.* (2020), <https://doi.org/10.1109/TCYB.2020.2991166> (early access).
- [24] B. Sun, E.-J. van Kampen, Event-triggered constrained control using explainable global dual heuristic programming for nonlinear discrete-time systems, *Neurocomputing* 468 (2022) 452–463.
- [25] A. Sahoo, H. Xu, S. Jagannathan, Near optimal event-triggered control of nonlinear discrete-time systems using neurodynamic programming, *IEEE Trans. Neural Netw. Learn. Syst.* 27 (9) (2016) 1801–1815.
- [26] K.G. Vamvoudakis, Event-triggered optimal adaptive control algorithm for continuous-time nonlinear systems, *IEEE/CAA J. Autom. Sin.* 1 (3) (2014) 282–293.
- [27] A. Sahoo, H. Xu, S. Jagannathan, Approximate optimal control of affine nonlinear continuous-time systems using event-sampled neurodynamic programming, *IEEE Trans. Neural Netw. Learn. Syst.* 28 (3) (2016) 639–652.
- [28] P. Tabuada, Event-triggered real-time scheduling of stabilizing control tasks, *IEEE Trans. Autom. Control* 52 (9) (2007) 1680–1685.
- [29] S. Xue, B. Luo, D. Liu, Y. Li, Adaptive dynamic programming based event-triggered control for unknown continuous-time nonlinear systems with input constraints, *Neurocomputing* 396 (2020) 191–200.
- [30] D. Wang, C. Mu, X. Yang, D. Liu, Event-based constrained robust control of affine systems incorporating an adaptive critic mechanism, *IEEE Trans. Syst. Man Cybern. Syst.* 47 (7) (2017) 1602–1612.
- [31] Y. Zhu, D. Zhao, H. He, J. Ji, Event-triggered optimal control for partially unknown constrained-input systems via adaptive dynamic programming, *IEEE Trans. Ind. Electron.* 64 (5) (2017) 4101–4109.
- [32] D. Wang, H. He, D. Liu, Intelligent optimal control with critic learning for a nonlinear overhead crane system, *IEEE Trans. Ind. Inform.* 14 (7) (2018) 2932–2940.
- [33] L. Sanches, T.A. Guimarães, E.D. Marques, Aeroelastic tailoring of nonlinear typical section using the method of multiple scales to predict post-flutter stable lcos, *Aerosp. Sci. Technol.* 90 (2019) 157–168.
- [34] H.K. Khalil, J.W. Grizzle, *Nonlinear Systems*, 3rd edition, Prentice Hall, Upper Saddle River, NJ, 2002.