Neurocomputing 245 (2017) 46–54

Contents lists available at ScienceDirect

Neurocomputing

journal homepage: www.elsevier.com/locate/neucom

Neural-network-based adaptive guaranteed cost control of nonlinear dynamical systems with matched uncertainties☆

Chaoxu Mu[a], Ding Wang[b,*]

[a] Tianjin Key Laboratory of Process Measurement and Control, School of Electrical and Information Engineering, Tianjin University, Tianjin 300072, China
[b] The State Key Laboratory of Management and Control for Complex Systems, Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China

ARTICLE INFO

Article history:
Received 3 August 2016
Revised 2 November 2016
Accepted 13 March 2017
Available online 24 March 2017

Communicated by Bo Shen.

Keywords:
Adaptive dynamic programming (ADP)
Guaranteed cost control
Neural networks
Uncertain dynamics
Stability

ABSTRACT

In this paper, we investigate the neural-network-based adaptive guaranteed cost control for continuous-time affine nonlinear systems with dynamical uncertainties. Through theoretical analysis, the guaranteed cost control problem is transformed into designing an optimal controller of the associated nominal system with a newly defined cost function. The approach of adaptive dynamic programming (ADP) is involved to implement the guaranteed cost control strategy with the neural network approximation. The stability of the closed-loop system with the guaranteed cost control law, the convergence of the critic network weights and the approximate boundary of the guaranteed cost control law are all analyzed. Two simulation examples have been conducted and all simulation results have indicated the good performance of the developed guaranteed cost control strategy.

1. Introduction

During control practices, the design on nonlinear optimal regulators always involves to solve the Hamiltonian–Jacobi–Bellman (HJB) equation. Although the nonlinear optimal problem is well described by the HJB equation from the point of view of mathematics, it is well known that the analytical solution of nonlinear HJB equation is almost impossible to be obtained since the problem itself encounters the partial differential equation. The approach of dynamic programming was deemed as a basic strategy to handle optimal control problems of nonlinear systems, but there still exists a serious issue called the curse of dimensionality [1,2]. To overcome the difficulty in coping with optimal control problems of nonlinear systems, based on function approximation structures like neural networks and support vector machines, approximate or adaptive dynamic programming (ADP) was proposed as a kind of effective method to solve nonlinear optimal control problems in forward time [3–8]. There exists a fundamental idea in ADP which is similar as adaptive control systems with function approximation techniques [9–13].

ADP and relevant methods have gained much development in various control issues of nonlinear systems. In [14], the control-constrained approximate optimal control was proposed for discrete-time nonlinear systems by the finite-horizon nonquadratic cost function based on an adaptive critic network structure. In [15], the optimal control problem of unknown and nonaffine nonlinear systems was addressed by the ADP-based method and recurrent neural networks. A novel policy iteration method of global ADP was proposed in [16] for the adaptive optimal control of nonlinear polynomial systems. The H_∞ state feedback control problem for a class of affine nonlinear discrete-time systems was investigated in [17] with unknown system dynamics. More topics of nonlinear systems have also been conducted with the ADP-based approach, such as optimal control of discrete-time nonlinear systems [18–21], optimal control of continuous-time nonlinear systems [22–24], optimal tracking control [25–27], robust control [28–30], differential games [31–33], and so on.

It is worth mentioning that the researchers have paid great attention on system uncertainties since uncertain systems exist in a broad rage [34–37]. For example, in [34], the robust filtering problem of the stochastic systems with polytopic uncertainties was investigated. [35] presented the state estimation for complex networks with uncertain inner coupling. In [36], the robust stabilization of uncertain switched neutral systems was intensively developed based on Lyapunov stability theory and the dynamic

☆ This work was supported in part by the National Natural Science Foundation of China under Grants U1501251, 61533017, 61304018, 61304086, and 61411130160, in part by Beijing Natural Science Foundation under Grant 4162065, in part by Tianjin Natural Science Foundation under Grant 14JCQNJC05400, in part by Research Fund of Tianjin Key Laboratory of Process Measurement and Control under Grant TKLPMC-201612, and in part by the Early Career Development Award of SKLMCCS.

* Corresponding author.
E-mail addresses: cxmu@tju.edu.cn (C. Mu), ding.wang@ia.ac.cn (D. Wang).

http://dx.doi.org/10.1016/j.neucom.2017.03.047
0925-2312/

output feedback technique. In [37], a robust output observer-based control was proposed for the problem of uncertainty in switched neutral systems with interval time-varying mixed delays. Simultaneously, the optimal control design on uncertain systems is more and more frequently consulted. As is known to all, the direct optimal control design of uncertain nonlinear systems is pretty difficult, since coping with the cost function of the uncertain system is not an easy task. Therefore, some researchers have paid attention to study the boundedness of the cost function with respect to the uncertain system, not limited to directly optimize it. The guaranteed cost control strategy, proposed by Chang and Peng [38], possesses the advantage of providing an upper bound on a given cost function, and therefore the cost function can be limited within this boundary even if the degradation of control performance is incurred by system uncertainties.

When the boundedness of the cost function of the uncertain systems is considered, it results in the guaranteed cost control problem. There are some results in the field of the guaranteed cost control design with the ADP-based approach. In [39], the ADP-based guaranteed cost control was firstly proposed for continuous-time uncertain nonlinear systems, and a modified cost function was established as the guaranteed cost function of the studied uncertain system. In [40], for a class of time-varying uncertain nonlinear systems, a neural-network-based approximate optimal guaranteed cost controller was developed with respect to a finite-horizon cost function. In [41], an ADP-based guaranteed cost control algorithm was presented for the tracking control of continuous-time uncertain nonlinear systems. In [42], the decentralized guaranteed cost control design was constructed for large-scale uncertain nonlinear systems, where both dynamical uncertainties and interconnections were discussed for better control performance. Using the ADP-based approach, an available guaranteed cost control scheme can be obtained for nonlinear system by approximately solving a optimal control problem. Most of the existing results of this field are obtained in terms of optimal regulation or tracking problems [39–42], not the guaranteed cost control problem regarding to unknown system dynamics. This greatly motivates our research.

This paper mainly contributes to the neural-network-based guaranteed cost control for continuous-time uncertain nonlinear systems with unknown dynamics. First, the cost functions of the original uncertain system and its associated nominal system are both defined. It is proved that the optimal control of the associated nominal system can implement the guaranteed cost control for the original uncertain system. Second, a neural network identifier is involved to the guaranteed cost control scheme. The weights of neural network identifier are adaptively updated to ensure the asymptotical stability of the identification error. Third, with the ADP-based technology, a novel learning control framework is built to approximately solve the optimal control of the nominal system for implementing the guaranteed cost control law of the original uncertain system. Theoretical analysis is provided for the stability of the closed-loop system with the learning-based guaranteed cost control, as well as the approximate boundary of the guaranteed cost control law.

The rest of this paper is organized as follows. The problem description of the guaranteed cost control for a class of continuous-time affine nonlinear systems is provided in Section 2, and the problem transformation is also stated. The adaptive critic methodology within the ADP framework is developed for the guaranteed cost control design in Section 3, and the associated theoretical analysis is also presented in this section. The simulation studies on two nonlinear systems are shown in Section 4. Finally, the conclusion remarks are given in Section 5.

2. Problem statement of the guaranteed cost control

We study a class of continuous-time affine nonlinear systems with the formula

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

where $x(t) \in \Omega_x \subseteq \mathbb{R}^n$ is the state vector and $u(t) \in \Omega_u \subseteq \mathbb{R}^m$ is the control vector, $f(x)$ and $g(x)$ are differentiable in their arguments with $f(0) = 0$ and $\|g(x)\| \le g_M$, g_M is a positive constant. $\Delta f(x)$ denotes the uncertain dynamics. In this paper, we consider that the uncertain dynamics satisfies the matching condition, i.e., it is in the range space of $g(x)$ rendering $\Delta f(x) = g(x)d(x)$ with $d(x) \in \mathbb{R}^m$. Besides, assume that $d(x)$ is upper bounded by a known function $D(x)$, i.e., $\|d(x)\| \le D(x)$ with $D(0) = 0$. Here, we also assume that $d(0) = 0$ such that $x = 0$ is an equilibrium of system (1).

For the uncertain nonlinear system (1), consider the cost function given as

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

where $R(x, u) = Q(x) + u^T u$ is the utility function, and $Q(x) = x^T Qx$ with $Q = Q^T \ge 0$.

The purpose of designing the guaranteed cost controller is to find a feedback control function $u(x)$ and determine a finite upper bound function $\Phi(u)$ (where $\Phi(u) \le M < +\infty$ with M being a positive constant), such that the closed-loop system (1) is robustly stable and the cost function (2) is bounded as $J(x, u) \le \Phi(u)$. Therefore, the function $\Phi(u)$ can effectively bound the cost function of system (1), named as the guaranteed cost function. Furthermore, when $\Phi(u)$ is minimal, it is termed as the optimal guaranteed cost and is written as Φ^*, where $\Phi^* = \min_u \Phi(u)$. The associated control law u^* is called the optimal guaranteed cost control law with $u^* = \arg\min_u \Phi(u)$. In this paper, we will study how to obtain the optimal guaranteed cost control law u^* for system (1) with the cost function (2).

Without considering the uncertainty, the controlled system turns to its nominal version, which plays an important role in the control design. The nominal system corresponding to (1) is formulated as

$$\dot{x}(t) = f(x) + g(x)u(t). \tag{3}$$

Similar as the classical literature of nonlinear optimal control, assume that the right side of (3) is Lipschitz continuous on a set $\Omega_x \subseteq \mathbb{R}^n$ containing the origin such that system (3) is controllable.

The following conclusions present the achievement of nonlinear robust control and the existence of guaranteed cost function of system (1), as an improvement of the result of [39]. Note that $\nabla(\cdot) \triangleq \partial(\cdot)/\partial x$ is employed to denote the gradient operator.

Theorem 1. *There exists a continuously differentiable function $V(x)$ for system (1) with $V(x) \ge 0$ and $V(0) = 0$ only at $x = 0$, such that for a feedback control law $u(x)$ satisfying*

$$(\nabla V(x))^T g(x) = -2u^T(x), \tag{4a}$$

$$R(x, u) + (\nabla V(x))^T (f(x) + g(x)u(x)) + \theta D^2(x) = 0, \tag{4b}$$

where $\theta \ge 1$. Then system (1) is locally asymptotically stable within the neighbourhood of the origin under the control law $u(x)$.

Proof. Define $V(x)$ as the Lyapunov function and let $\dot{V}(x)$ be its derivative along system (1). By using (4), we can find that

$$\begin{aligned}\dot{V}(x) &= (\nabla V(x))^T (f(x) + g(x)u(x) + \Delta f(x)) \\ &= -\theta D^2(x) - R(x, u) - 2u^T(x)d(x).\end{aligned} \tag{5}$$

Adding and subtracting a quadratic term $d^T(x)d(x)$ and considering $\|d(x)\| \le D(x)$, it follows from (5) that

$$\begin{aligned}\dot{V}(x) =& -\theta D^2(x) - x^T Qx + d^T(x)d(x) \\ & - (u(x)+d(x))^T(u(x)+d(x)) \\ \le & - x^T Qx - (\theta - 1)D^2(x) \\ & - (u(x)+d(x))^T(u(x)+d(x)).\end{aligned}$$

Then, we derive that $\dot{V}(x) \le -x^T Qx < 0$ holds with $\theta \ge 1$ for $x \ne 0$. Hence, we obtain the asymptotic stability of system (1) with the feedback control law $u(x)$. □

Theorem 2. *The cost function of system (3) defined as*

$$\phi(x,u) = \int_t^\infty \Big\{R(x(\tau),u(x(\tau))) + \theta D^2(x(\tau))\Big\}\mathrm{d}\tau \tag{6}$$

can ensure that $J(x,u) \le \phi(x,u) = V(x)$ *holds.*

Proof. Recalling the definition of $J(x, u)$ and $\phi(x, u)$ in (2) and (6), it is obvious that $J(x, u) \le \phi(x, u)$. When $V(x)$ is taken as the Lyapunov function of system (3), we have its derivative as

$$\dot{V}(x) = (\nabla V(x))^T\big(f(x) + g(x)u(t)\big). \tag{7}$$

Based on (4b), it can be deduced that

$$(\nabla V(x))^T\big(f(x) + g(x)u(t)\big) = -R(x,u) - \theta D^2(x). \tag{8}$$

Combing (7) with (8), it yields

$$\dot{V}(x) = -R(x,u) - \theta D^2(x). \tag{9}$$

Integrating (9) in the interval $[0, \infty)$, it can be concluded that $\phi(x,u) = V(x)$. As a result, $J(x,u) \le \phi(x,u) = V(x)$ is proved, which guarantees the boundedness of the cost function $J(x, u)$. □

Theorem 1 provides the control law $u(x)$ that can robustly stabilize the original uncertain system (1). Through the conclusion of Theorem 2, $\phi(x, u)$ is seen as a guaranteed cost function of system (1). For the purpose of attaining the optimal guaranteed cost controller, i.e., dealing with the optimality of the guaranteed cost function, a control law $u(x)$ is pursued to minimize the bound function $\phi(x, u)$. Then, the original control problem described in this section is transformed into an optimal control problem. That is to say, for the nominal system (3), the optimal controller which minimizes the cost function (6) is the desired guaranteed cost solution of the original control problem.

Based on the above analysis, we aim to solve the optimal control problem of the nominal system. The designed optimal feedback control must be admissible. For system (3), observing

$$\begin{aligned}\phi(x,u) &= \int_t^\infty \Big\{R(x(\tau),u(x(\tau))) + \theta D^2(x(\tau))\Big\}\mathrm{d}\tau \\ &= \int_t^{t_1} \Big\{R(x,u) + \theta D^2(x)\Big\}\mathrm{d}\tau + \Phi\big(x(t_1),u\big).\end{aligned} \tag{10}$$

According to Theorem 2, it has $\phi(x,u) = V(x)$. Therefore,

$$V(x) = \int_t^{t_1} \Big\{R(x,u) + \theta D^2(x)\Big\}\mathrm{d}\tau + V\big(x(t_1)\big), \tag{11}$$

it follows that

$$\begin{aligned}&\lim_{t\to t_1}\int_t^{t_1}\frac{1}{t_1-t}\big(R(x,u)+\theta D^2(x)\big)\mathrm{d}\tau + \lim_{t\to t_1}\frac{1}{t_1-t}\big(V(x(t_1)) - V(x(t))\big) \\ &= R(x,u) + \theta D^2(x) + (\nabla V(x))^T\big(f(x)+g(x)u(t)\big) = 0.\end{aligned} \tag{12}$$

(12) is clearly equivalent to (4b). The above analysis demonstrates that (4b) is the nonlinear Lyapunov equation.

Note that in the following, $V(x)$ is employed to represent the cost function of system (3), instead of $\phi(x, u)$. Ω_x is a compact subset of $\mathbb{R}^n$ and Ω_u is the set of admissible controls on Ω_x. In light of the classical optimal control theory, we define the Hamiltonian function of transformed problem as

$$H(x,u,\nabla V(x)) = \theta D^2(x) + R(x,u) + (\nabla V(x))^T\big(f(x) + g(x)u(t)\big). \tag{13}$$

The optimal cost function of the nominal system (3) is defined as $V^*(x) = \min_{u\in\Omega_u}\phi(x,u)$, which satisfies the continuous-time HJB equation of the form

$$0 = \min_{u\in\Omega_u} H(x,u,\nabla V^*(x)). \tag{14}$$

Hence, the optimal control of system (3) is obtained by

$$u^*(x) = -\frac{1}{2}g^T(x)\nabla V^*(x). \tag{15}$$

Using the optimal control $u^*(x)$, the HJB equation becomes

$$0 = \theta D^2(x) + R(x,u^*) + (\nabla V^*(x))^T\big(f(x) + g(x)u^*(x)\big) \tag{16}$$

with $V^*(0) = 0$. Based on (15), (16) can also be written as

$$\begin{aligned}0 =& \theta D^2(x) + Q(x) + (\nabla V^*(x))^T f(x) \\ & - \frac{1}{4}(\nabla V^*(x))^T g(x)g^T(x)\nabla V^*(x)\end{aligned} \tag{17}$$

with $V^*(0) = 0$.

Suppose that the HJB equation (17) has a continuously differentiable solution $V^*(x)$. It can be found that the optimal guaranteed cost can be gained when letting $u = u^*(x)$. Then, $\Phi(u^*) = V^*(x)$, which implies that $\Phi^* = V^*(x)$ and $u^* = \arg\min_u \Phi(u)$. Therefore, once the solution of (17) related with (3) is gotten, we can develop the optimal guaranteed cost controller of system (1).

Since it is always difficult to analytically solve the nonlinear optimal control, kinds of ADP-based methods combining the idea of reinforcement learning have been proposed to get the approximate solution. In what follows, we turn to the adaptive guaranteed cost control design with the ADP-based method.

3. Neural adaptive design for guaranteed cost control

In this section, we develop the neural adaptive design for the guaranteed cost control problem, including a neural network identification of the unknown dynamics, the neural network implementation of adaptive guaranteed cost control, and some related theoretical analysis.

3.1. Neural network identification

In this paper, we assume that the internal dynamics of system (3) are unknown. A neural network identifier with three layers is introduced to reconstruct the dynamics by using the system data. Define the number of the hidden layer neurons as h_m, then system (3) can be expressed by

$$\dot{x}(t) = Ax(t) + w_m^T\sigma_m(\chi) + \zeta_m(x), \tag{18}$$

where $A \in \mathbb{R}^{n\times n}$ is a designed stable matrix, $w_m \in \mathbb{R}^{h_m\times n}$ is the ideal weight matrix between the hidden layer and the output layer, $\sigma_m(\cdot) \in \mathbb{R}^{h_m}$ is a monotonically increasing activation function, $\chi = \kappa_m^T z(t)$ with $\chi \in \mathbb{R}^{h_m}$, $\kappa_m \in \mathbb{R}^{(n+m)\times h_m}$ is the used weight matrix between the input layer and the hidden layer, $z(t) = [x^T(t), u^T(t)]^T \in \mathbb{R}^{n+m}$ is the augmented input vector, and $\zeta_m(x) \in \mathbb{R}^n$ is the reconstruction error.

For simplicity, let the input-to-hidden weight matrix κ_m be constant and only regulate the hidden-to-output weight matrix. Then, the output of neural network identifier is

$$\dot{\hat{x}}(t) = A\hat{x}(t) + \hat{w}_m^T(t)\sigma_m(\hat{\chi}), \tag{19}$$

where $\hat{w}_m(t)$ is the estimated weight matrix of w_m at time t, $\hat{x}(t)$ is the estimated system state vector, and $\hat{\chi} = \kappa_m^T[\hat{x}^T(t), u^T(t)]^T$.

Let $\tilde{w}_m(t) = \hat{w}_m(t) - w_m$ be the weight error matrix and $\tilde{x}(t) = \hat{x}(t) - x(t)$ be the identification error. Then, according to (18) and (19), the dynamical equation with respect to the identification error can be derived as

$$\begin{aligned}\dot{\tilde{x}}(t) =& A\tilde{x}(t) + \hat{w}_m^T(t)\sigma_m(\hat{\chi}) - w_m^T\sigma_m(\chi) - \zeta_m(x)\\ =& A\tilde{x}(t) + \tilde{w}_m^T(t)\sigma_m(\hat{\chi}) + w_m^T\big(\sigma_m(\hat{\chi}) - \sigma_m(\chi)\big) - \zeta_m(x).\end{aligned} \tag{20}$$

With regard to the weight matrices and the reconstruction error, one lemma and two assumptions that are commonly used in the community [15,33], are provided for the following theoretical analysis.

Lemma 1 (cf. [15]). *With the activation function $\sigma_m(\cdot)$, we can derive the following inequality*

$$\|\sigma_m(\chi_a) - \sigma_m(\chi_b)\| \le \lambda_{\sigma_m}\|\chi_a - \chi_b\| \tag{21}$$

for any two vectors χ_a and χ_b with same dimensions, $\lambda_{\sigma_m} \ge 0$ is a constant.

Assumption 1. The weight matrices w_m and κ_m are bounded such that $\|w_m\| \le w_M$ and $\|\kappa_m\| \le \kappa_M$, where w_M and κ_M are positive constants.

Assumption 2. The reconstruction error $\zeta_m(x)$ is upper bounded by a function of the identification error, such that $\zeta_m^T(x)\zeta_m(x) \le \lambda_{\zeta_m}\tilde{x}^T(t)\tilde{x}(t)$, where $\lambda_{\zeta_m} < 1$ is a positive constant.

Theorem 3. *Using the neural network identifier (19), if the weight matrix is tuned by*

$$\dot{\hat{w}}_m(t) = -\eta_m\sigma_m(\hat{\chi})\tilde{x}^T(t), \tag{22}$$

where $\eta_m > 0$ is the learning rate of the neural network identifier, then the identification error $\tilde{x}(t)$ is asymptotically stable.

Proof. Choose a Lyapunov function as

$$L_m(t) = L_{11}(t) + L_{12}(t),$$

where

$$L_{11}(t) = \tilde{x}^T(t)\tilde{x}(t), \tag{23a}$$

$$L_{12}(t) = \eta_m^{-1}\mathrm{tr}\{\tilde{w}_m^T(t)\tilde{w}_m(t)\}. \tag{23b}$$

Take the derivative of $L_{11}(t)$ along the trajectory of the error system (20) and obtain

$$\begin{aligned}\dot{L}_{11}(t) =& 2\tilde{x}^T(t)\Big(A\tilde{x}(t) + w_m^T\big(\sigma_m(\hat{\chi}) - \sigma_m(\chi)\big) - \zeta_m(x)\Big)\\ &+ 2\tilde{x}^T(t)\tilde{w}_m^T(t)\sigma_m(\hat{\chi}).\end{aligned} \tag{24}$$

For the term $\dot{L}_{12}(t)$, using the updating rule (22) and the property of trace operation, we find that

$$\dot{L}_{12}(t) = 2\eta_m^{-1}\mathrm{tr}\big(\tilde{w}_m^T(t)\dot{\hat{w}}_m(t)\big) = -2\tilde{x}^T(t)\tilde{w}_m^T(t)\sigma_m(\hat{\chi}). \tag{25}$$

According to (24) and (25), we can obtain

$$\begin{aligned}\dot{L}_m(t) =& 2\tilde{x}^T(t)A\tilde{x}(t) + 2\tilde{x}^T(t)w_m^T(t)\big(\sigma_m(\hat{\chi}) - \sigma_m(\chi)\big)\\ &- 2\tilde{x}^T(t)\zeta_m(x).\end{aligned} \tag{26}$$

Adopting (21) and using Assumption 1, we have

$$\|\hat{\chi} - \chi\| \le \|\kappa_m\|\|\hat{z}(t) - z(t)\| \le \|\kappa_m\|\|\hat{x}(t) - x(t)\| \le \kappa_M\|\tilde{x}(t)\|,$$

such that

$$\begin{aligned}&2\tilde{x}^T(t)w_m^T\big(\sigma_m(\hat{\chi}) - \sigma_m(\chi)\big)\\ &\le \tilde{x}^T(t)w_m^Tw_m\tilde{x}(t) + \big(\sigma_m(\hat{\chi}) - \sigma_m(\chi)\big)^T\big(\sigma_m(\hat{\chi}) - \sigma_m(\chi)\big)\\ &\le \tilde{x}^T(t)w_m^Tw_m\tilde{x}(t) + \lambda_{\sigma_m}^2\kappa_M^2\tilde{x}^T(t)\tilde{x}(t).\end{aligned} \tag{27}$$

Considering Assumption 2, we derive

$$\begin{aligned}-2\tilde{x}^T(t)\zeta_m(x) &\le \tilde{x}^T(t)\tilde{x}(t) + \zeta_m^T(x)\zeta_m(x)\\ &\le (1 + \lambda_{\zeta_m})\tilde{x}^T(t)\tilde{x}(t).\end{aligned} \tag{28}$$

Based on (27) and (28), we further obtain

$$\begin{aligned}\dot{L}_m(t) \le& \tilde{x}^T(t)\big(2A + w_m^Tw_m + (1 + \lambda_{\zeta_m} + \lambda_{\sigma_m}^2\kappa_M^2)I_n\big)\tilde{x}(t)\\ =& -\tilde{x}^T(t)\mathcal{M}\tilde{x}(t),\end{aligned}$$

where the matrix $\mathcal{M}$ stands for $\mathcal{M} = -2A - w_M^2I_n - (1 + \lambda_{\zeta m} + \lambda_{\sigma m}^2\kappa_M^2)I_n$, I_n represents the $n \times n$ identity matrix. If A is selected to ensure that $\mathcal{M} > 0$, then the time derivative of the Lyapunov function is $\dot{L}_m(t) < 0$ for any $\tilde{x}(t) \ne 0$. Thus, we find that the identification error can approach zero as time goes to infinity (i.e., $\tilde{x}(t) \to 0$ as $t \to \infty$), which ends the proof. □

According to Theorem 3, the neural network identifier is asymptotically stable. It means that the estimated system states would finally converge to the actual states. Hence, after a sufficient learning process, an available neural network identifier can be obtained with finally convergent weights as follows:

$$\dot{x}(t) = f(x) + g(x)u(x) = Ax(t) + \hat{w}_m^T\sigma_m(\chi), \tag{29}$$

where $x(t)$ can be actually understood in the sense of the final output derived from the model network, instead of $\hat{x}(t)$ for simplicity. Take the partial derivative of (29) with respect to $u(x)$ and derive that

$$g(x) = \hat{w}_m^T\left(\frac{\partial\sigma_m(\chi)}{\partial\chi}\right)\kappa_m^T\left[\begin{array}{c}0_{n\times m}\\ \hdashline I_m\end{array}\right], \tag{30}$$

where the term $\partial\sigma_m(\chi)/\partial\chi$ is in fact a h_m-dimensional square matrix. Eq. (30) reconstructs the information of the control matrix.

3.2. Neural adaptive guaranteed cost control design

The optimal cost $V^*(x)$ for system (3) can be expressed by a neural network on a compact set Ω_x, i.e.,

$$V^*(x) = w_c^T\sigma_c(x) + \zeta_c(x), \tag{31}$$

where $w_c \in \mathbb{R}^{h_c}$ is the ideal weight vector, $\sigma_c(x) \in \mathbb{R}^{h_c}$ is the activation function, h_c is the number of hidden layer neurons, and $\zeta_c(x)$ represents the approximation error by neural networks.

Take the derivative of (31) to get $\nabla V^*(x)$ as

$$\nabla V^*(x) = (\nabla\sigma_c(x))^Tw_c + \nabla\zeta_c(x). \tag{32}$$

Based on this expression, the Lyapunov equation (4b) becomes

$$\begin{aligned}&H(x, u, w_c)\\ &= \theta D^2(x) + R(x, u) + \big(w_c^T\nabla\sigma_c(x) + (\nabla\zeta_c(x))^T\big)\dot{x}(t) = 0.\end{aligned} \tag{33}$$

Substitute (32) into (15) to get the optimal feedback control under the ideal weight w_c, which is

$$u^*(x) = -\frac{1}{2}g^T(x)\big((\nabla\sigma_c(x))^Tw_c + \nabla\zeta_c(x)\big). \tag{34}$$

The above equation represents the neural network expression of the optimal feedback control if $V^*(x) = w_c^T\sigma_c(x) + \zeta_c(x)$ is regarded as a neural network expression of the optimal cost function. In this sense, it is that $\nabla V^*(x)$ and $u^*(x)$ can be written in terms of the ideal weight w_c in (32) and (34), respectively. Hence, the Hamiltonian equation $H(x, u^*, \nabla V^*) = 0$ implies $H(x, u, w_c) = 0$ as shown in (33), which is in fact expanded as

$$H(x, u, w_c) = \theta D^2(x) + R(x, u) + w_c^T\nabla\sigma_c(x)\dot{x}(t) - e_H = 0, \tag{35}$$

where

$$e_H = -(\nabla\zeta_c(x))^T\dot{x}(t) \tag{36}$$

represents the residual error of the Hamiltonian function, which is caused when conducting the approximation operation by using neural networks.

Under the framework of ADP-based approximate optimal control design, a critic neural network is built based on an estimated weight vector $\hat{w}_c(t)$, i.e.,

$$\hat{V}(x) = \hat{w}_c^T(t)\sigma_c(x) \tag{37}$$

to approximate the cost function. Obviously, we can get that

$$\nabla\hat{V}(x) = (\nabla\sigma_c(x))^T\hat{w}_c(t). \tag{38}$$

Similarly, the approximate optimal control function is derived as

$$\hat{u}(x) = -\frac{1}{2}g^T(x)(\nabla\sigma_c(x))^T\hat{w}_c(t). \tag{39}$$

With the estimation weight vector $\hat{w}_c(t)$, the approximate Hamiltonian function can be written as

$$\hat{H}(x, u, \hat{w}_c) = \theta D^2(x) + R(x, u) + \hat{w}_c^T(t)\nabla\sigma_c(x)\dot{x}(t). \tag{40}$$

Define $e_c(t) = \hat{H}(x, u, \hat{w}_c) - H(x, u, w_c)$ as the approximate error of the critic network, which means $e_c(t) = \hat{H}(x, u, \tilde{w}_c)$, where $\tilde{w}_c(t) = w_c - \hat{w}_c(t)$ denotes the estimation error of the critic network weight vector. Therefore, we derive that

$$e_c(t) = e_H - \tilde{w}_c^T(t)\nabla\sigma_c(x)\dot{x}(t). \tag{41}$$

For training the critic network, we design $\hat{w}_c(t)$ to minimize the network error function expressed as $E_c = (1/2)e_c^T(t)e_c(t)$. The weight vector is updated via a steepest descent algorithm, i.e.,

$$\dot{\hat{w}}_c(t) = -\eta_c\left(\frac{\partial E_c}{\partial\hat{w}_c(t)}\right), \tag{42}$$

where $\eta_c > 0$ is the learning rate of the training process for the critic network.

Next, we consider the weight error of the critic network. Considering (41), it can be found that

$$\frac{\partial e_c(t)}{\partial\hat{w}_c(t)} = \nabla\sigma_c(x)\dot{x}(t) \triangleq \varphi(x). \tag{43}$$

Hence, according to the fact that $\hat{w}_c(t) = w_c - \tilde{w}_c(t)$, the dynamics of the weight estimation error are

$$\dot{\tilde{w}}_c(t) = -\dot{\hat{w}}_c(t) = \eta_c e_c(t)\left(\frac{\partial e_c(t)}{\partial\hat{w}_c(t)}\right). \tag{44}$$

Then, we can further deduce that

$$\dot{\tilde{w}}_c(t) = \eta_c\big(e_H - \tilde{w}_c^T(t)\nabla\sigma_c(x)\dot{x}(t)\big)\nabla\sigma_c(x)\dot{x}(t). \tag{45}$$

3.3. Stability analysis

Before the related stability analysis, we first introduce two common assumptions.

Assumption 3. The activation function $\sigma_c(\cdot)$ and its derivative $\nabla\sigma_c(\cdot)$, the reconstruction error $\zeta_c(x)$ and its derivative $\nabla\zeta_c(x)$, the ideal weight vector w_c are all bounded, i.e., $\|\sigma_c(\cdot)\| \le \sigma_M$, $\|\nabla\sigma_c(\cdot)\| \le \sigma_D$, $\|\zeta_c(t)\| \le \zeta_M$, $\|\nabla\zeta_c(t)\| \le \zeta_D$ and $\|w_c\| \le w_{cM}$, where σ_M, σ_D, ζ_M, ζ_D and w_M are all positive constants.

Assumption 4. Considering system (3) is controllable, e_H and $\varphi(x)$ are assumed to be bounded by $\|e_H\| \le e_M$ and $\|\varphi(x)\| \le \varphi_M$, respectively, where e_M and φ_M are both positive constants.

The convergence of the critic network weights and the uniformly ultimately bounded stability of the closed-loop system are shown in the following theorem.

Theorem 4. *Consider the nonlinear system (3). Let the feedback control law be obtained by (39) and the weight vector of the critic network be trained by (42). Then, the weight estimation error $\tilde{w}_c(t)$ and the state vector $x(t)$ of the closed-loop system are uniformly ultimately bounded.*

Proof. Let the Lyapunov function be chosen as $L(t) = L_{21}(t) + L_{22}(t)$, where $L_{21}(t) = (1/2)\eta_c^{-1}\tilde{w}_c^T(t)\tilde{w}_c(t)$ and $L_{22}(t) = V^*(x)$. The derivative of $L(t)$ is formulated as $\dot{L}(t) = \dot{L}_{21}(t) + \dot{L}_{22}(t)$.

For deriving the term $\dot{L}_{21}(t)$, we consider (45) and obtain that

$$\begin{aligned}\dot{L}_{21}(t) =& \eta_c^{-1}\dot{\tilde{w}}_c^T(t)\tilde{w}_c(t)\\ =& \big(e_{cH} - \tilde{w}_c^T(t)\nabla\sigma_c(x)\dot{x}(t)\big)\tilde{w}_c^T(t)\nabla\sigma_c(x)\dot{x}(t).\end{aligned} \tag{46}$$

By applying the Cauchy–Schwarz inequality and Assumption 4, $\dot{L}_{21}(t)$ can be deduced as

$$\begin{aligned}\dot{L}_{21}(t) =& \eta_c^{-1}\left(\eta_c e_H\tilde{w}_c^T(t)\varphi(x) - \eta_c\big(\tilde{w}_c^T(t)\varphi(x)\big)^2\right)\\ \le& \frac{1}{2\eta_c}\left(e_H^2 + \big(\eta_c\tilde{w}_c^T(t)\varphi(x)\big)^2\right) - \big(\tilde{w}_c^T(t)\varphi(x)\big)^2\\ =& -\left(1 - \frac{\eta_c}{2}\right)\|\tilde{w}_c^T(t)\varphi(x)\|^2 + \frac{1}{2\eta_c}e_H^2.\end{aligned} \tag{47}$$

Observing (47), it can obtain $\dot{L}_{21}(t) \le 0$ with

$$0 < \eta_c < 2, \|\tilde{w}_c(t)\| \ge \sqrt{\frac{e_M^2}{\eta_c\varphi_M(2-\eta_c)}}. \tag{48}$$

Therefore, we can obtain that $\dot{L}_{21}(t) \le 0$ if $0 < \eta_c < 2$ and $\tilde{w}_c(t)$ lies outside the compact set

$$\Omega_{\tilde{w}_c} = \left\{\tilde{w}_c(t) : \|\tilde{w}_c(t)\| \le e_M\big(\eta_c\varphi_M(2-\eta_c)\big)^{-\frac{1}{2}}\right\}. \tag{49}$$

For the term $\dot{L}_{22}(t)$, we observe (17) and derive that

$$(\nabla V^*(x))^T f(x) = -\theta D^2(x) - Q(x) + \frac{1}{4}(\nabla V^*(x))^T gg^T\nabla V^*(x). \tag{50}$$

Then, applying (39) and (50), we can find that

$$\begin{aligned}\dot{L}_{22}(t) =& (\nabla V^*(x))^T\big(f(x) + g(x)\hat{u}(x)\big)\\ =& -\theta D^2(x) - Q(x) + \frac{1}{4}(\nabla V^*(x))^T g(x)g^T(x)\\ &\times\nabla V^*(x) + (\nabla V^*(x))^T g(x)\hat{u}(x)\\ =& -\theta D^2(x) - Q(x) + \frac{1}{4}(\nabla V^*(x))^T g(x)g^T(x)\\ &\times\nabla V^*(x) - \frac{1}{2}(\nabla V^*(x))^T g(x)g^T(x)(\nabla\sigma_c(x))^T\\ &\times(w_c - \tilde{w}_c(t))\\ =& -\theta D^2(x) - Q(x) - \frac{1}{4}(\nabla V^*(x))^T g(x)g^T(x)\\ &\times\nabla V^*(x) + \frac{1}{2}(\nabla V^*(x))^T g(x)g^T(x)\tilde{w}_c\\ &+\frac{1}{2}(\nabla V^*(x))^T g(x)g^T(x)\nabla\zeta_c(x).\end{aligned} \tag{51}$$

Since the quadratic bound of $d(x)$ can be determined in many cases. It is common to assume that $D(x) = \rho\|x\|$, where ρ is a positive constant. Let $\lambda_{\min}(Q)$ be the least eigenvalue of Q, and thus $Q(x) \ge \lambda_{\min}(Q)\|x\|^2$. Then, (51) can be derived that

$$\begin{aligned}\dot{L}_{22}(t) =& -\theta\rho^2\|x\|^2 - \lambda_{\min}(Q)\|x(t)\|^2 + \frac{1}{2}(\nabla V^*(x))^T\\ &\times g(x)g^T(x)(\tilde{w}_c(t) + \nabla\zeta_c(x)).\end{aligned} \tag{52}$$

By applying Assumption 3, the above inequality can be further deduced to

$$\dot{L}_{22}(t) = -\big(\theta\rho^2 + \lambda_{\min}(Q)\big)\|x(t)\|^2 + \psi^2, \tag{53}$$

where

$$\psi^2 = \frac{1}{2}(\nabla V^*(x))^T g(x) g^T(x)(\tilde{w}_c(t) + \nabla \zeta_c(x))$$
$$\leq \frac{1}{2}(\sigma_D w_{cM} + \zeta_D) g_M^2 \left(e_M \left(\eta_c \varphi_M (2 - \eta_c) \right)^{-\frac{1}{2}} + \zeta_D \right) \triangleq \psi_M^2 \quad (54)$$

Therefore, we can obtain that $\dot{L}_{22}(t) \leq 0$ if $x(t)$ lies outside the compact set

$$\Omega_x = \left\{ x(t) : \|x(t)\| \leq \psi_M \left(\theta \rho^2 + \lambda_{\min}(Q) \right)^{-\frac{1}{2}} \right\} \quad (55)$$

holds. According to (49) and (55), it can be concluded that that the weight estimation error $\tilde{w}_c(t)$ and the state vector $x(t)$ are uniformly ultimately bounded. Thus, the proof is completed. □

Corollary 1. *The approximate optimal control law $\hat{u}(x)$ is designed in the formula (39), which converges to the neighborhood of the optimal control $u^*(x)$, with a finite boundary given in (56).*

Proof. According to Theorem 4, it can derive that $\|\tilde{w}_c(t)\| \leq e_M \left(\eta_c \varphi_M (2 - \eta_c) \right)^{-\frac{1}{2}}$. Based on (34) and (39), we have that

$$\|u^*(x) - \hat{u}(x)\| = \frac{1}{2} \|g^T(x) \left((\nabla \sigma_c(x))^T \tilde{w}_c(t) + \nabla \zeta_c(x) \right)\|$$
$$\leq \frac{1}{2} g_M \left(\sigma_D e_M \left(\eta_c \varphi_M (2 - \eta_c) \right)^{-\frac{1}{2}} + \zeta_D \right) \triangleq \lambda_u, \quad (56)$$

where λ_u stands for the finite boundary with respect to the control signal. Thus the proof is completed. □

Via the developed approximate optimal control scheme, the guaranteed cost control strategy for the original uncertain system is attained. It is significant to note that the established method provides a combination of adaptivity and robustness, thus enlarges the application scope of approximate optimal strategy to the study of guaranteed cost control of nonlinear systems under uncertain environment.

Remark 1. The proposed method of this paper focuses on the ADP-based guaranteed cost control for uncertain nonlinear systems with unknown dynamics. Using the ADP-based approach, we can obtain a novel learning-based guaranteed cost control scheme. Compared with most of the existing results of this field in terms of optimal regulation or tracking problems [39–42], this proposed method further involves a neural network identifier to the guaranteed cost control framework for unknown system dynamics. The weights of the identifier are updated to guarantee the asymptotical stability of the identification error. Theoretical analysis is provided for the stability of the closed-loop system with the learning-based guaranteed cost control. With this kind of design, it means that the model network is introduced to implement the control scheme. Therefore, the proposed guaranteed cost control scheme is data-driven based on available system data without requiring the accurate model. It is a learning control framework with the ADP-based technology.

4. Simulation

In this section, two simulation cases are provided to demonstrate the control performance of the neural-network-based adaptive guaranteed cost control strategy.

Case 1. The studied nonlinear dynamic system with uncertainties is described in the following formula

$$\dot{x} = \begin{bmatrix} -x_1^3 - 2x_2 \\ x_1 + 0.5\cos(x_1^2)\sin(x_2^3) \end{bmatrix} + \begin{bmatrix} 1 \\ -1 \end{bmatrix} u(x) + \Delta f(t), \quad (57)$$

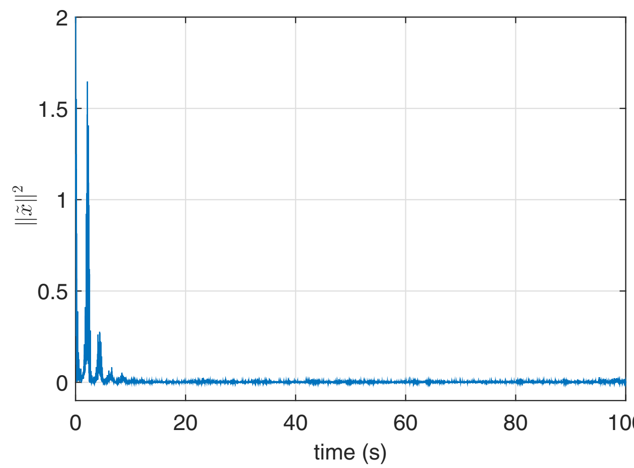


Fig. 1. The norm square of identification error.

where $x = [x_1, x_2]^T \in \mathbb{R}^2$, $u(x) \in \mathbb{R}$ and $\Delta f(t) = [d_1 x_2 \sin(d_2 x_1^2), -d_1 x_2 \sin(d_2 x_1^2)]^T$ with d_1 and d_2 belonging to the interval $[-1, 1]$. Considering $\Delta f(t)$ is matched, therefore $\Delta f(t) = g(x) d(x)$, where $d(x) = d_1 x_2 \sin(d_2 x_1^2)$ and $D(x) = \rho \|x\|$.

In this case, select $Q(x) = x^T x$, $d_1 = 0.8$, $d_2 = 0.5$. In order to solve the guaranteed cost control of system (57), the optimal control of its nominal system is considered with the associated cost function in the formula (6). According to Theorem 2, (6) is the guaranteed cost function of system (57). In order to obtain the optimal guaranteed cost control law $u(x)$ for system (57), it is transformed to minimize the guaranteed cost function $\phi(x, u)$. It means that the optimal guaranteed cost control law $u(x)$ can minimize the cost function of nominal system. The optimal guaranteed cost control law $u(x)$ is approximately solved by the ADP-based approach with the neural network implementation.

Since the internal dynamics are unknown, an identifier is conducted to capture the system dynamics, where a three-layer neural network is taken as the identifier with the structure of 3–6–2 (i.e., 3 input neurons, 6 hidden neurons, and 2 output neurons). The learning rate of the model network is $\eta_m = 0.1$ and the activation function is the tansig function. The selected negative definite matrix $A = -I_2$. We train the neural network identifier for 100 s to obtain the evolution result of the identification error, which is shown in Fig. 1.

The weight vector between the input layer and the hidden layer is initialized as

$$\kappa_m = \begin{bmatrix} -0.5895 & 0.0020 & 0.8314 & -0.5755 & -0.6453 & -0.9565 \\ -0.1547 & 0.2543 & 0.9095 & -0.8287 & 0.1488 & 0.2455 \\ -0.7355 & 0.1980 & -0.6288 & 0.6704 & -0.4947 & -0.3569 \end{bmatrix},$$

while the weight vector between the hidden layer and the output layer is finally convergent to

$$\hat{w}_m = \begin{bmatrix} 0.9441 & -1.3837 & -1.0364 & 0.2429 & -1.5269 & -1.5009 \\ 0.2808 & -0.1009 & 1.1530 & -0.3143 & 0.1585 & -0.2734 \end{bmatrix}^T.$$

Then, we finish training the neural identifier and keep all weights unchanged. According to (29) and (30), $g(x)$ of system (57) can be approximated as $g(x) = [0.9857 \ \ -1.0024]^T$ and the associated nominal dynamics are estimated as $\dot{x}(t) = Ax(t) + \hat{w}_m(t)\sigma(\kappa_m(\chi))$. Based on the the reconstructed neural network model, the ADP-based approach is used to approximately solve the optimal guaranteed cost control law. A critic neural network is constructed with the activation function $\sigma(x) = [x_1^2, x_1 x_2, x_2^2]^T$ and the learning rate $\eta_c = 0.01$. By taking the initial state vector of the nominal system

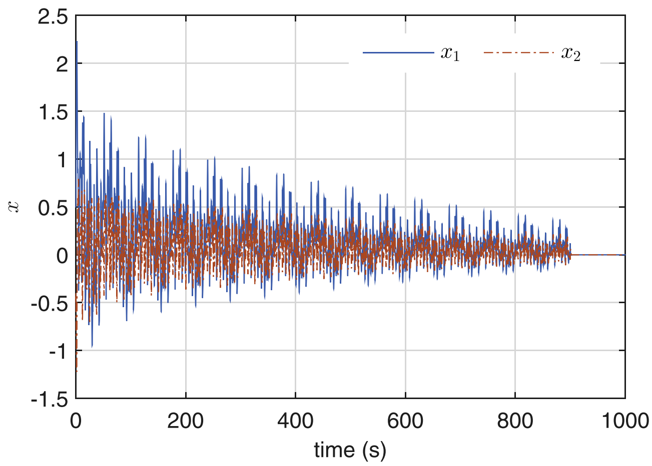


Fig. 2. The state curves with the exploration noises when training the critic network.

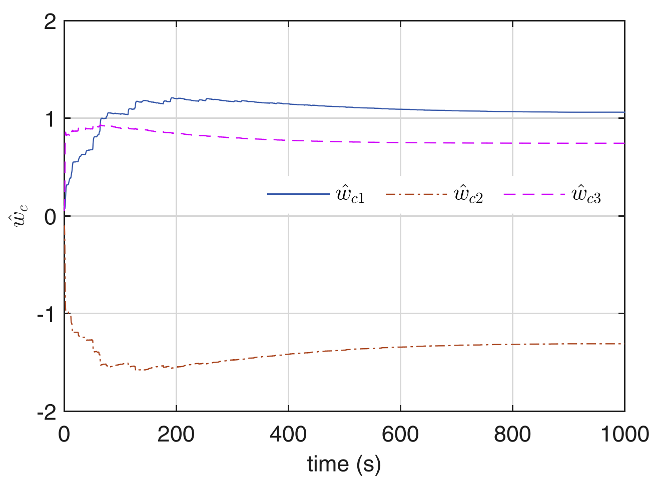


Fig. 3. The convergence curves of the weights when training the critic network.

as $[0.5,\ 0.5]^T$ and setting $\theta = 1.5$ and $\rho = 1$, the critic network is trained by the weight updating rule presented in (42). During the training process, for satisfying the persistency of excitation condition, an excitation noise $N(t) = \sin^2(t)\cos(t) + \sin^2(2t)\cos(0.1t) + \sin^2(-1.2t)\cos(0.5t) + \sin^5(t) + \sin^2(1.2t) + \cos(2.4t)\sin^3(2.4t)$ is added to the control input and thus further affects the system states. The training process lasts 1000 s and the exploration noise is turned off at 900 s when the weight vector of the critic network has converged to $\hat{w}_c = [1.0620 \quad -1.3086 \quad 0.7445]^T$, as shown in Figs. 2 and 3.

With the trained critic network weight vector $\hat{w}_c$, the approximated optimal guaranteed cost control law can be obtained according to (39). Then we apply the control law to system (57) for 15 s to evaluate the control performance. The state trajectories are presented in Fig. 4 and the used control variable is shown in Fig. 5. Fig. 6 depicts the cost function of system (57) and the guaranteed cost function, which shows that the cost function $J(x,\ u)$ of system (57) has been bounded by $\phi(x,\ u)$. These results have illustrated that system (57) can be well operated under the designed guaranteed cost control law.

Case 2. The studied three-order continuous-time uncertain nonlinear system is described as

$$\dot{x} = \begin{bmatrix} -x_1 + x_2 \\ -0.2x_2 - x_1x_3 \\ x_1x_2 - x_1^2x_3 \end{bmatrix} + \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix} u(x) + \Delta f(t), \tag{58}$$

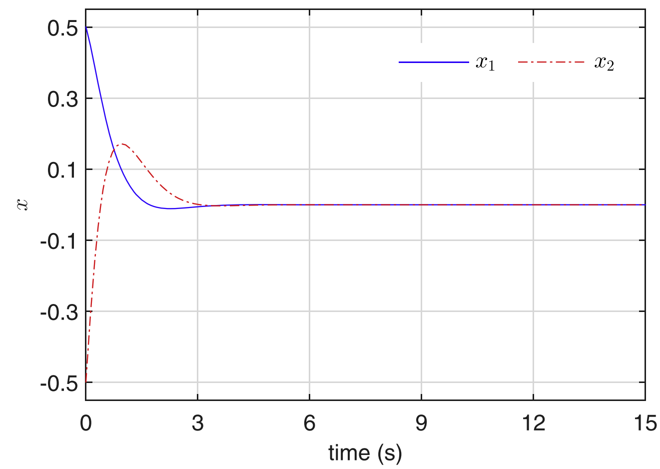


Fig. 4. The state trajectories of system (57).

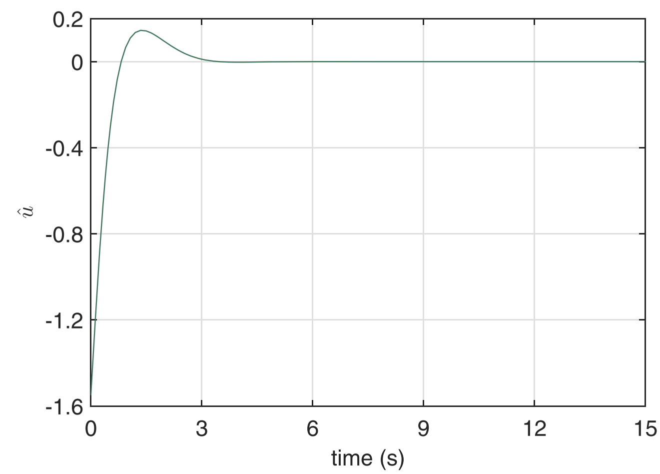


Fig. 5. The curve of guarantee cost control law.

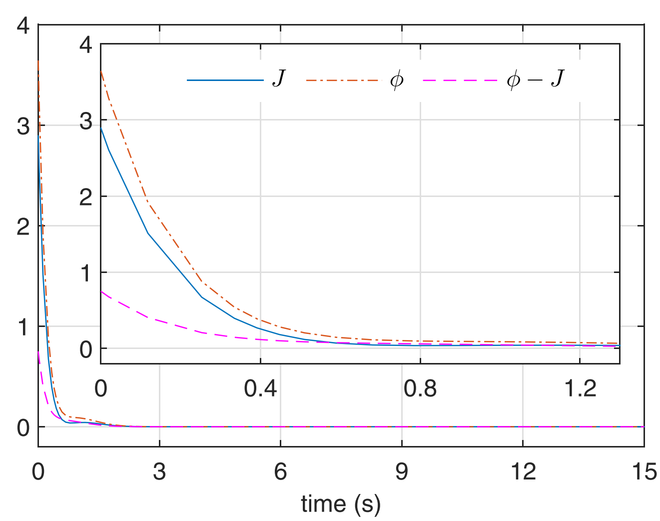


Fig. 6. The cost function of system (57) and its guarantee cost function.

where $x = [x_1, x_2, x_3]^T \in \mathbb{R}^3$, $u(x) \in \mathbb{R}^3$, and $\Delta f(t) = [-d_1x_3\sin(x_1x_2), -d_2x_1\sin(x_2), d_3x_2\cos(x_3)]^T$ with d_1, d_2 and d_3 belonging to the interval $[-1, 1]$. Since $\Delta f(t)$ is matched, then $d(x) = [-d_1x_3\sin(x_1x_2), d_2x_1\sin(x_2),\ d_3x_2\cos(x_3)]^T$ and its boundary $D(x)$ satisfies $D(x) = \rho\|x\|$.

In this case, $Q(x)$ is selected as $Q(x) = x^Tx$. The guaranteed cost control problem of system (58) is approximately solved by the ADP-based approach with the neural network implementation. The

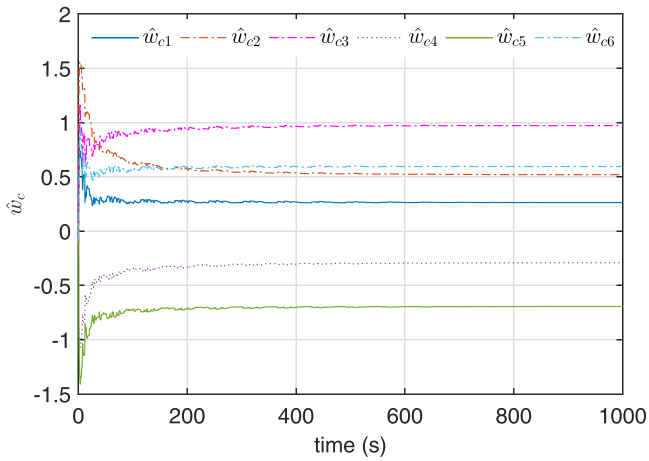


Fig. 7. The convergence curves of the weights when training the critic network.

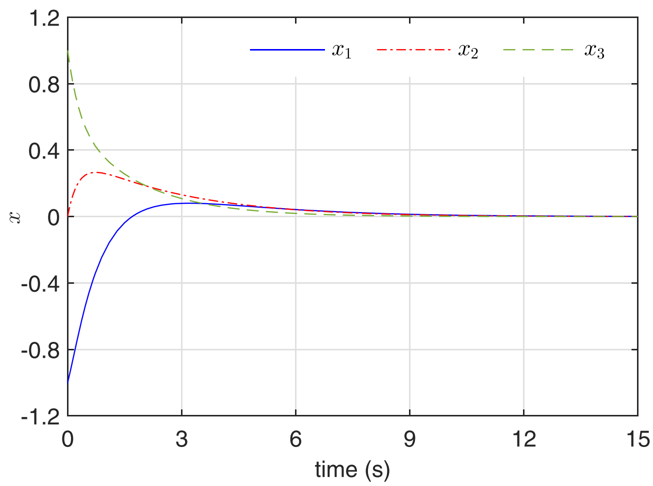


Fig. 8. The state trajectories of system (58).

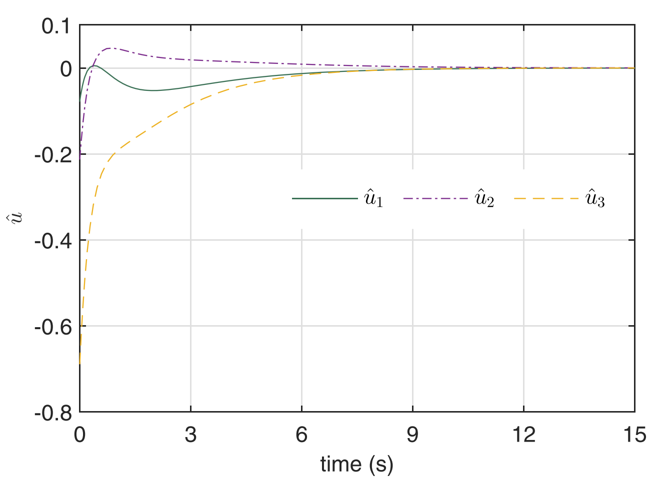


Fig. 9. The curves of control variables in the guarantee cost control law.

neural network identifier with the structure 6–8–3 is first built based on the data of the studied system, where eight hidden neurons all use the tansig function as the activation function. We set the negative definite matrix $A = -I_3$ and then train the neural network identifier for 100 s. The input-to-hidden weight vector is initialized as κ_m and the hidden-to-output weight vector is trained as $\hat{w}_m$. Thus, the nominal dynamics of system (58) are estimated as $\dot{x}(t) = Ax(t) + \hat{w}_m(t)\sigma(\kappa_m(\chi))$. By applying (30), $g(x)$ is obtained as

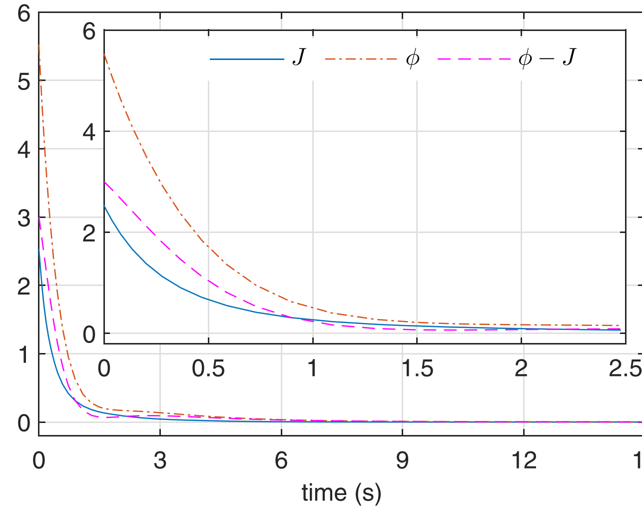


Fig. 10. The cost function of system (58) and its guarantee cost function.

$$g(x) = \begin{bmatrix} 1.1363 & 0.0565 & 0.0165 \\ 0.0055 & -1.0668 & -0.0092 \\ 0.0575 & -0.0055 & 1.0183 \end{bmatrix}.$$

After the identification of the nominal system dynamics, the ADP-based approach is used to solve the approximate optimal control law. The critic network is constructed to approximate the cost function of the nominal system with the structure 3–6–1, and the learning rate of the critic network is set as $\eta_c = 0.5$. The training process for the critic network lasts 1000 s. The exploration noise $N(t)$ is also added to the control variables for providing the persistent excitation. It is turned off at 900 s when all weights have converged with the exploration noise, which indicates a sufficient training process. Then, we can obtain that the weight vector of the critic network converges to $[0.2628 \quad 0.5198 \quad 0.9717 \quad -0.2912 \quad -0.6939 \quad 0.5951]^T$, as shown in Fig. 7. With the trained weights of the critic network, we can derive the approximated optimal guaranteed cost control law according to (39). In order to verify the performance of the guaranteed cost control, we apply the control law into system (58) to stabilize all states. Let the initial states be $x_0 = [1, 0, 1]^T$ and parameters in Δf be $d_1 = 1,\ d_2 = 0.5$ and $d_3 = 0.1$. Such that we select $\theta = 1.5$ and $\rho = 1$. Fig. 8 shows all system trajectories during the regulation, and Fig. 9 provides the values of all control variables. Fig. 10 presents the cost function $J(x, u)$ of system (58) and its guarantee cost function $\phi(x, u)$, which obviously shows that $J(x, u)$ is bounded by $\phi(x, u)$. These simulation results verify the effectiveness of the developed guaranteed cost control approach.

5. Conclusions

In this paper, the relationship between the guaranteed cost control strategy of original uncertain system and the optimal control of nominal system has been clarified. The neural network identifer with adaptive weights has been designed to obtain the unknown system dynamics. By using the ADP-based method to approximately solve the HJB equation of the optimal control problem of nominal system, a novel learning control framework of guaranteed cost control for original uncertain system has been established. The related theoretical analysis and simulation verification conducted on two examples have well demonstrated that the developed guaranteed control law is effective and robust. For further research work, this learning control framework of guaranteed cost control is also expected to be developed for non-affine uncertain systems including unknown dynamics and unmatched disturbances. Also, we will investigate the robust optimal control of uncertain nonlinear system with unknown dynamics and unmatched disturbances with the ADP-based method.

References

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

[2] F.L. Lewis, V.L. Syrmos, Optimal Control, John Wiley & Sons, New York, NY, 1995.

[3] D.V. Prokhorov, R.A. Santiago, D.C. Wunsch, Adaptive critic designs: a case study for neurocontrol, Neural Netw. 8 (9) (1995) 1367–1372.

[4] J. Si, Y.T. Wang, Online learning control by association and reinforcement, IEEE Trans. Neural Netw. 12 (2) (2001) 264–276.

[5] F.L. Lewis, D. Liu, Reinforcement Learning and Approximate Dynamic Programming for Feedback Control, Wiley, Hoboken, NJ, 2013.

[6] H. He, Z. Ni, J. Fu, A three-network architecture for on-line learning and optimization based on adaptive dynamic programming, Neurocomputing 78 (1) (2012) 3–13.

[7] H. Zhang, D. Liu, Y. Luo, D. Wang, Adaptive Dynamic Programming for Control: Algorithms and Stability, Springer, London, UK, 2013.

[8] D. Wang, C. Mu, D. Liu, Data-driven nonlinear near-optimal regulation based on iterative neural dynamic programming, Acta Automatica Sinica 43 (3) (2017) 354–363.

[9] L. Cheng, Z.G. Hou, M. Tan, Adaptive neural network tracking control for manipulators with uncertain kinematics, dynamics and actuator model, Automatica 45 (10) (2009) 2312–2318.

[10] B. Xu, Robust adaptive neural control of flexible hypersonic flight vehicle with dead-zone input nonlinearity, Nonlinear Dyn. 80 (3) (2015) 1509–1520.

[11] B. Gu, V.S. Sheng, K.Y. Tay, W. Romano, S. Li, Incremental support vector learning for ordinal regression, IEEE Trans. Neural Netw. Learn. Syst. 26 (7) (2015a) 1403–1416.

[12] B. Gu, V.S. Sheng, Z. Wang, D. Ho, S. Osman, S. Li, Incremental learning for ν-support vector regression, Neural Netw. 67 (2015b) 140–150.

[13] W. He, Y. Dong, C. Sun, Adaptive neural impedance control of a robotic manipulator with input saturation, IEEE Trans. Syst. Man Cybern. Syst. 46 (3) (2016) 334–344.

[14] 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.

[15] H. Zhang, L. Cui, X. Zhang, Y. Luo, Data-driven robust approximate optimal tracking control for unknown general nonlinear systems using adaptive dynamic programming method, IEEE Trans Neural Netw. 22 (12) (2011) 2226–2236.

[16] Y. Jiang, Z.P. Jiang, Global adaptive dynamic programming for continuous-time nonlinear systems, IEEE Trans. Automatic Control 60 (11) (2015) 2917–2929.

[17] H. Zhang, C. Qin, B. Jiang, Y. Luo, Online adaptive policy learning algorithm for h_∞ state feedback control of unknown affine nonlinear discrete-time systems, IEEE Trans. Cybern. 44 (12) (2014) 2706–2718.

[18] D. Wang, D. Liu, Q. Wei, D. Zhao, N. Jin, Optimal control of unknown nonaffine nonlinear discrete-time systems based on adaptive dynamic programming, Automatica 48 (8) (2012) 1825–1832.

[19] C. Mu, C. Sun, A. Song, H. Yu, Iterative GDHP-based approximate optimal tracking control for a class of discrete-time nonlinear systems, Neurocomputing 214 (2016) 775–784.

[20] J. Zhang, H. Zhang, Y. Luo, T. Feng, Model-free optimal control design for a class of linear discrete-time systems with multiple delays using adaptive dynamic programming, Neurocomputing 135 (2014) 163–170.

[21] Q. Zhao, H. Xu, S. Jagannathan, Near optimal output feedback control of nonlinear discrete-time systems based on reinforcement neural network learning, IEEE/CAA J. Automatica Sinica 1 (4) (2014) 372–384.

[22] K.G. Vamvoudakis, F.L. Lewis, Online actor-critic algorithm to solve the continuous-time infinite horizon optimal control problem, Automatica 46 (5) (2010) 878–888.

[23] Y. Jiang, Z.P. Jiang, Computational adaptive optimal control for continuous-time linear systems with completely unknown dynamics, Automatica 48 (10) (2012) 2699–2704.

[24] B. Luo, H.N. Wu, T. Huang, D. Liu, Reinforcement learning solution for HJB equation arising in constrained optimal control problem, Neural Netw. 71 (2015) 150–158.

[25] J. Na, G. Herrmann, Online adaptive approximate optimal tracking control with simplified dual approximation structure for continuoustime unknown nonlinear systems, IEEE/CAA J. Automatica Sinica 1 (4) (2014) 412–422.

[26] C. Mu, Z. Ni, C. Sun, H. He, Air-breathing hypersonic vehicle tracking control based on adaptive dynamic programming, IEEE Trans. Neural Netw. Learn. Syst. 28 (3) (2017) 584–598.

[27] C. Mu, Z. Ni, C. Sun, H. He, Data-driven tracking control with adaptive dynamic programming for a class of continuous-time nonlinear systems, IEEE Trans. Cybern. (2016), doi:10.1109/TCYB.2016.2548941.

[28] Y. Jiang, Z.P. Jiang, Robust adaptive dynamic programming and feedback stabilization of nonlinear systems, IEEE Trans. Neural Netw. Learn. Syst. 25 (5) (2014) 882–893.

[29] D. Wang, D. Liu, H. Li, Policy iteration algorithm for online design of robust control for a class of continuous-time nonlinear systems, IEEE Trans. Automation Sci. Eng. 11 (2) (2014) 627–632.

[30] D. Wang, C. Li, D. Liu, C. Mu, Data-based robust optimal control of continuous-time affine nonlinear systems with matched uncertainties, Inf. Sci. 366 (2016) 121–133.

[31] S. Mehraeen, T. Dierks, S. Jagannathan, M.L. Crow, Zero-sum two-player game theoretic formulation of affine nonlinear discrete-time systems using neural networks, IEEE Trans. Cybern. 43 (6) (2013) 1641–1655.

[32] C. Qin, H. Zhang, Y. Wang, Y. Luo, Neural network-based online h_∞ control for discrete-time affine nonlinear system using adaptive dynamic programming, Neurocomputing 198 (2016) 91–99.

[33] D. Zhao, Q. Zhang, D. Wang, Y. Zhu, Experience replay for optimal control of nonzero-sum game systems with unknown dynamics, IEEE Trans. Cybern. 46 (3) (2016) 854–865.

[34] B. Shen, Z. Wang, H. Shu, G. Wei, Robust h_∞ finite-horizon filtering with randomly occurred nonlinearities and quantization effects, Automatica 46 (11) (2010) 1743–1751.

[35] B. Shen, Z. Wang, D. Ding, H. Shu, h_∞ state estimation for complex networks with uncertain inner coupling and incomplete measurements, IEEE Trans. Neural Netw. Learn. Syst. 24 (12) (2013) 2027–2037.

[36] H. Ghadiri, M.R. Jahed-Motlagh, M.B. Yazdi, Robust stabilization for uncertain switched neutral systems with interval time-varying mixed delays, Nonlinear Anal. Hybrid Syst. 13 (2014a) 2–21.

[37] H. Ghadiri, M.R. Jahed-Motlagh, M.B. Yazdi, Robust output observer-based guaranteed cost control of a class of uncertain switched neutral systems with interval time-varying mixed delays, Int. J. Control Automation Syst. 12 (6) (2014b) 1167–1179.

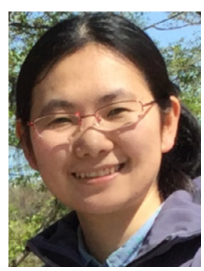
[38] S.S.L. Chang, T.K.C. Peng, Adaptive guaranteed cost control of systems with uncertain parameters, IEEE Trans. Automatic Control 17 (4) (1972) 474–483.

[39] D. Liu, D. Wang, F.Y. Wang, H. Li, X. Yang, Neural-network-based online HJB solution for optimal robust guaranteed cost control of continuous-time uncertain nonlinear systems, IEEE Trans. Cybern. 44 (12) (2014) 2834–2847.

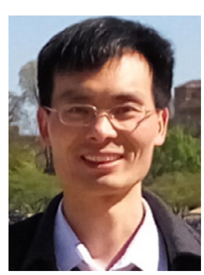
[40] H.N. Wu, M. Li, L. Guo, Finite-horizon approximate optimal guaranteed cost control of uncertain nonlinear systems with application to mars entry guidance, IEEE Trans. Neural Netw. Learn. Syst. 26 (7) (2015) 1456–1467.

[41] X. Yang, D. Liu, Q. Wei, D. Wang, Guaranteed cost neural tracking control for a class of uncertain nonlinear systems using adaptive dynamic programming, Neurocomputing 198 (2016) 80–90.

[42] D. Wang, D. Liu, C. Mu, H. Ma, Decentralized guaranteed cost control of interconnected systems with uncertainties: a learning-based optimal control strategy, Neurocomputing 214 (2016) 297–306.



Chaoxu Mu received the Ph.D. degree in control science and engineering from Southeast University, Nanjing, China, in 2012. She was a Visiting Ph.D. Student with the Royal Melbourne Institute of Technology University, Melbourne, VIC, Australia, from 2010 to 2011. She has been a Post-Doctoral Fellow with the Department of Electrical, Computer and Biomedical Engineering, The University of Rhode Island, Kingston, RI, USA, from 2014 to 2016. She is currently an Associate Professor with the School of Electrical and Information Engineering, Tianjin University, Tianjin, China. She has authored over 50 journal and conference papers, and co-authored one monograph. Her current research interests include nonlinear system control and optimization, adaptive and learning systems, and machine learning. She is the IEEE member and a member of the Asia-Pacic Neural Network Society.



Ding Wang received the B.S. degree in mathematics from Zhengzhou University of Light Industry, Zhengzhou, China, the M.S. degree in operations research and cybernetics from Northeastern University, Shenyang, China, and the Ph.D. degree in control theory and control engineering from Institute of Automation, Chinese Academy of Sciences, Beijing, China, in 2007, 2009, and 2012, respectively. He has been a Visiting Scholar with the Department of Electrical, Computer, and Biomedical Engineering, University of Rhode Island, Kingston, RI, USA, from 2015 to 2017. He is currently an Associate Professor with The State Key Laboratory of Management and Control for Complex Systems, Institute of Automation, Chinese Academy of Sciences. His research interests include adaptive and learning systems, intelligent control, and neural networks. He has published over 80 journal and conference papers, and coauthored two monographs. He was the Finance Chair of the 12th World Congress on Intelligent Control and Automation (WCICA 2016), the Secretariat of the 2014 IEEE World Congress on Computational Intelligence (IEEE WCCI 2014), and the Registration Chair of the 5th International Conference on Information Science and Technology (ICIST 2015) and the 4th International Conference on Intelligent Control and Information Processing (ICICIP 2013), and served as the program committee member of several international conferences. He was a recipient of the Excellent Doctoral Dissertation Award of Chinese Academy of Sciences in 2013, and a nomination of the Excellent Doctoral Dissertation Award of Chinese Association of Automation (CAA) in 2014. He serves as an Associate Editor of IEEE Transactions on *Neural Networks and Learning Systems* and *Neurocomputing*. He is a member of Institute of Electrical and Electronics Engineers (IEEE), Asia-Pacific Neural Network Society (APNNS), and CAA.