

An Approximate Optimal Control Approach for Robust Stabilization of a Class of Discrete-Time Nonlinear Systems With Uncertainties

Ding Wang, *Member, IEEE*, Derong Liu, *Fellow, IEEE*, Hongliang Li, Biao Luo, and Hongwen Ma

Abstract—In this correspondence paper, the robust stabilization of a class of discrete-time nonlinear systems with uncertainties is investigated by using an approximate optimal control approach. The robust control problem is transformed into an optimal control problem under some proper restrictions on the bound of the uncertainties. For the purpose of dealing with the transformed optimal control, the discrete-time generalized Hamilton–Jacobi–Bellman equation is introduced and then solved using the successive approximation method with neural network implementation. In addition, a numerical simulation is included to illustrate the effectiveness of the robust control strategy.

Index Terms—Adaptive dynamic programming (ADP), generalized Hamilton–Jacobi–Bellman (GHJB) equation, neural networks, optimal control, robust control, successive approximation method, uncertainties.

I. INTRODUCTION

It is known that the model uncertainties must be considered during the controller design process since they may cause deterioration of the control systems. In general, we say a controller is robust if it works even if the actual system deviates from its nominal model based on which the controller is designed. In fact, the robustness of control systems has been attended and studied by control scientists for many years. Robust control has become an important topic of modern control theory [1]–[3]. Lin *et al.* [3] pointed out that under proper restricted conditions, the robust control problem can be converted into an optimal control problem. Though the detailed operation procedure was not given, it provided an alternative method to deal with the robust stabilization problem. Thus, optimal control methods can be employed to design robust controllers. In fact, the research on linear optimal control has matured during the last several decades. However, when studying the nonlinear optimal control problem, we have to solve the Hamilton–Jacobi–Bellman (HJB) equation, which is often a difficult task. Therefore, some indirect methods have been proposed in order to overcome the difficulty in solving the nonlinear HJB equation. In [4], a recursive method was employed to deal with the optimal control problem for continuous-time nonlinear systems by successively solving the generalized HJB (GHJB) equation. The GHJB equation, which gives the cost of an arbitrary control

law, can be used to improve the performance of this control and to approximate the HJB equation successively as well. In [5], adaptive (or approximate) dynamic programming (ADP) was presented to solve the optimal control problem, mainly for discrete-time nonlinear systems, based on function approximation structures such as neural networks. In recent years, the research on optimal control based on GHJB formulation and the ADP approach has acquired much attention from scholars [6]–[10]. Specifically, ADP has become one of the key directions for future research in intelligent control and understanding intelligence.

Recently, by taking Taylor series expansion of the cost function, Chen and Jagannathan [11] and Jagannathan [12] applied the GHJB formulation to study the optimal control of discrete-time affine nonlinear systems, while the system uncertainties were not taken into consideration. Even so, it is meaningful that the discrete-time GHJB equation and the related discrete-time HJB equation are well-defined. Therefore, the application scope of the GHJB-formulation-based method is greatly extended. In addition, it presents another effective way to solve the constrained optimal control problem of discrete-time nonlinear systems [13]. After that, Adhyaru *et al.* [2] studied the bounded robust control of continuous-time constrained nonlinear systems with uncertainties by deriving the neural-network-based HJB solution, but the proposed approach was not suitable for discrete-time systems. Therefore, in this correspondence paper, we will investigate the robust stabilization of a class of discrete-time nonlinear systems with uncertainties using the discrete-time GHJB-formulation-based optimal control approach. Remarkably, this paper extends [11] and [12], which focuses on the GHJB-based optimal control for discrete-time affine nonlinear system, to robust controller design of uncertain nonlinear system. Additionally, this paper also develops a new robust control method for discrete-time nonlinear systems with uncertainties under the framework of the idea of ADP.

II. PROBLEM STATEMENT

In this paper, we study a class of discrete-time nonlinear systems described by

$$x_{k+1} = f(x_k) + g(x_k)(u(x_k) + d(x_k)) \quad (1)$$

where $x_k \in \mathbb{R}^n$ is the state vector, $u(x_k) \in \mathbb{R}^m$ is the control vector, and $f(\cdot)$ and $g(\cdot)$ are differentiable in their arguments with $f(0) = 0$. In (1), the term $g(x_k)d(x_k)$ with $d(x_k) \in \mathbb{R}^m$ is the unknown perturbation that represents the matched uncertainty of system dynamics.

Note that in this paper, stability is always with respect to $x = 0$. In addition, it is assumed that $d(x_k)$ is bounded by a known function $d_M(x_k)$, that is

$$\|d(x_k)\| \leq d_M(x_k), \forall k. \quad (2)$$

Moreover, we assume that $d(0) = 0$, so that $x = 0$ is an equilibrium of system (1). It is also assumed that the bounded function $d_M^2(x)$ is differentiable and $d_M(0) = 0$.

Here, we aim at investigating the robust control problem for uncertain nonlinear systems (1). In other words, we should develop a

Manuscript received July 20, 2014; revised December 15, 2014 and March 31, 2015; accepted May 18, 2015. This work was supported in part by the National Natural Science Foundation of China under Grant 61034002, Grant 61233001, Grant 61273140, and Grant 61304086, and in part by the Early Career Development Award of the State Key Laboratory of Management and Control for Complex Systems. This paper was recommended by Associate Editor A. H. Tan.

D. Wang, B. Luo, and H. Ma are with the State Key Laboratory of Management and Control for Complex Systems, Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China (e-mail: ding.wang@ia.ac.cn; biao.luo@ia.ac.cn; mahongwen2012@ia.ac.cn).

D. Liu is with the School of Automation and Electrical Engineering, University of Science and Technology Beijing, Beijing 100083, China (e-mail: derong@ustb.edu.cn).

H. Li is with IBM Research—China, Beijing 100193, China (e-mail: hongliang.li625@foxmail.com).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TSMC.2015.2466191

feedback control law $u(x)$, such that the closed-loop system is globally asymptotically stable for all uncertainties $d(x_k)$. In the following, we will display the result that this problem is closely related to the optimal controller design of the corresponding nominal system, with an appropriate choice of the cost function.

Consider the nominal system with respect to system (1) given as follows:

$$x_{k+1} = f(x_k) + g(x_k)u(x_k). \quad (3)$$

For system (3), we assume that $f + gu$ is Lipschitz continuous on a set Ω in \mathbb{R}^n containing the origin. In addition, we assume the system is controllable, i.e., there exists a continuous control law on Ω that can stabilize the system asymptotically.

When dealing with the optimal control of system (3), we desire to find the control law $u(x)$ which minimizes the infinite horizon cost function

$$J(x_k, u) = \sum_{q=k}^{\infty} \left\{ \rho d_M^2(x_q) + U(x_q, u(x_q)) \right\}. \quad (4)$$

In (4), ρ is a positive number and $U(\cdot, \cdot)$ is the utility function with $U(0, 0) = 0$ and $U(x_q, u(x_q)) \geq 0$ for any x_q and $u(x_q)$. In this paper, the utility function is chosen as the quadratic form $U(x_q, u(x_q)) = x_q^T Q x_q + u^T(x_q) R u(x_q)$, where Q is a positive definite matrix and R is a symmetric positive definite matrix, all with suitable dimensions. Note that for the optimal control problem, the designed feedback control law must be admissible [7], [8], [11], [12].

Now, we define the optimal cost function as follows:

$$J^*(x_k) = \min_{\{u(\cdot)\}} \sum_{q=k}^{\infty} \left\{ \rho d_M^2(x_q) + U(x_q, u(x_q)) \right\} \quad (5)$$

where $\{u(\cdot)\}$ denotes the sequence of control input, i.e., $u(x_k), u(x_{k+1}), \dots$. According to Bellman's optimality principle, we can obtain the discrete-time HJB equation

$$J^*(x_k) = \min_{u(x_k)} \left\{ \rho d_M^2(x_k) + U(x_k, u(x_k)) + J^*(x_{k+1}) \right\}. \quad (6)$$

The corresponding optimal control u^* is

$$u^*(x_k) = -\frac{1}{2} R^{-1} g^T(x_k) \frac{\partial J^*(x_{k+1})}{\partial x_{k+1}}. \quad (7)$$

Then, using the optimal control u^* , the discrete-time HJB equation (6) becomes

$$J^*(x_k) = \rho d_M^2(x_k) + U(x_k, u^*(x_k)) + J^*(f(x_k) + g(x_k)u^*(x_k)). \quad (8)$$

When studying the linear quadratic regulator problem, the discrete-time HJB equation reduces to the Riccati equation that can be solved efficiently. However, for the general nonlinear problem, it is not the case. Furthermore, the optimal control $u^*(x_k)$ is related to x_{k+1} and $J^*(x_{k+1})$, which cannot be determined at present time step k . Hence, in the following, we will employ the GHJB formulation to deal with the optimal control design problem. Moreover, the robust controller for system (1) can be established based on the optimal controller.

III. ROBUST CONTROLLER DESIGN BASED ON OPTIMAL CONTROL APPROACH USING DISCRETE-TIME GHJB FORMULATION

In this section, the discrete-time GHJB equation considering the modified cost function is defined first. Then, the successive approximation method is developed to solve the discrete-time GHJB equation and a neural network is constructed for facilitating the implementation, which results in an approximate optimal control. At last, the approximate optimal controller of system (3) is proved to be a robust stabilizer of system (1).

A. Discrete-Time GHJB Equation

In this part, motivated by the results of [4] and [7] for continuous-time systems and [11]–[13] for discrete-time systems, the discrete-time GHJB equation with the modified cost function is considered.

Lemma 1: Given an admissible control $\mu(x) \in \Omega_u$ (Ω_u is the set of admissible controls) for system (3), there exists a positive definite and continuously differentiable function $V(x)$ satisfying $V(x_0) = J(x_0, \mu)$ if the following equation holds:

$$\begin{aligned} & \rho d_M^2(x) + U(x, \mu(x)) + (\nabla V(x))^T (f(x) + g(x)\mu(x) - x) \\ & + \frac{1}{2} (f(x) + g(x)\mu(x) - x)^T \nabla^2 V(x) \\ & \times (f(x) + g(x)\mu(x) - x) = 0, V(0) = 0 \end{aligned} \quad (9)$$

where $\nabla V(x) = \partial V(x)/\partial x$ and $\nabla^2 V(x) = \partial^2 V(x)/\partial x^2$ is the Hessian matrix given as

$$\nabla^2 V(x) = \begin{bmatrix} \frac{\partial^2 V(x)}{\partial x_1^2} & \frac{\partial^2 V(x)}{\partial x_1 \partial x_2} & \dots & \frac{\partial^2 V(x)}{\partial x_1 \partial x_n} \\ \frac{\partial^2 V(x)}{\partial x_2 \partial x_1} & \frac{\partial^2 V(x)}{\partial x_2^2} & \dots & \frac{\partial^2 V(x)}{\partial x_2 \partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 V(x)}{\partial x_n \partial x_1} & \frac{\partial^2 V(x)}{\partial x_n \partial x_2} & \dots & \frac{\partial^2 V(x)}{\partial x_n^2} \end{bmatrix}. \quad (10)$$

Proof: Assume that a positive definite $V(x)$ exists and is continuously differentiable. Besides, assume that the high-order terms of the Taylor series expansion of $V(x)$ are small and can be neglected. Since μ is admissible, $x_k = 0$ and $V(x_k) = 0$ when $k \rightarrow \infty$. Denoting

$$\begin{aligned} \Delta V(x_k) &= V(x_{k+1}) - V(x_k) \\ &\approx (\nabla V(x_k))^T (f(x_k) + g(x_k)\mu(x_k) - x_k) \\ &+ \frac{1}{2} (f(x_k) + g(x_k)\mu(x_k) - x_k)^T \nabla^2 V(x_k) \\ &\times (f(x_k) + g(x_k)\mu(x_k) - x_k) \end{aligned} \quad (11)$$

where $\nabla V(x_k) = \nabla V(x)|_{x=x_k}$, $\nabla^2 V(x_k) = \nabla^2 V(x)|_{x=x_k}$, and considering (4) and (9), we can easily find that $V(x_k) = J(x_k, \mu)$ based on the proof of [11] and [12]. ■

In view of Lemma 1, the positive definite function $V(x)$ is also the value function of the optimal control problem, with admissible control being introduced.

Now, we define the discrete-time GHJB equation for system (3) as follows:

$$\begin{aligned} \text{GHJB}(V(x), \mu(x)) &\triangleq \rho d_M^2(x) + U(x, \mu(x)) + (\nabla V(x))^T \\ &\times (f(x) + g(x)\mu(x) - x) \\ &+ \frac{1}{2} (f(x) + g(x)\mu(x) - x)^T \nabla^2 V(x) \\ &\times (f(x) + g(x)\mu(x) - x) = 0 \\ V(0) &= 0. \end{aligned} \quad (12)$$

The Hamiltonian function for system (3) is

$$\begin{aligned} H(x, \mu(x), V(x)) &= \rho d_M^2(x) + U(x, \mu(x)) + (\nabla V(x))^T \\ &\times (f(x) + g(x)\mu(x) - x) \\ &+ \frac{1}{2} (f(x) + g(x)\mu(x) - x)^T \nabla^2 V(x) \\ &\times (f(x) + g(x)\mu(x) - x). \end{aligned} \quad (13)$$

The optimal value $V^*(x)$ associated with the discrete-time GHJB equation (12) satisfies

$$0 = \min_{\mu \in \Omega_u} H(x, \mu(x), V^*(x)). \quad (14)$$

Similar as [11] and [12], we observe that $g^T(x)\nabla^2 V^*(x)g(x) + 2R$ is positive definite. In this sense, the optimal control is

$$\begin{aligned}\mu^*(x) &= \arg \min_{\mu \in \Omega_u} H(x, \mu(x), V^*(x)) \\ &= -\left(g^T(x)\nabla^2 V^*(x)g(x) + 2R\right)^{-1} g^T(x) \\ &\quad \times \left(\nabla V^*(x) + \nabla^2 V^*(x)(f(x) - x)\right).\end{aligned}\quad (15)$$

As a result, by substituting the optimal control (15) into the discrete-time GHJB equation (12), we can obtain the approximate version of the discrete-time HJB equation as follows:

$$\begin{aligned}\rho d_M^2(x) + U(x, \mu^*(x)) + (\nabla V^*(x))^T(f(x) + g(x)\mu^*(x) - x) \\ + \frac{1}{2}(f(x) + g(x)\mu^*(x) - x)^T \nabla^2 V^*(x) \\ \times (f(x) + g(x)\mu^*(x) - x) = 0, V^*(0) = 0.\end{aligned}\quad (16)$$

Remark 1: It is worth emphasizing that the discrete-time HJB equation (16) is attained under the framework of discrete-time GHJB equation. Actually, it can be regarded as an approximate version of the ideal HJB equation (8) by carrying out the Taylor series expansion. In addition, the optimal control (15) is also an approximation of the ideal optimal control (7) to a certain degree. That is to say, V^* and μ^* are approximate results of J^* and u^* , respectively. It is in this sense that $\mu^*(x_k)$ can be expressed in terms of x_k , not x_{k+1} . For this reason, the optimal control is available.

Remark 2: Incidentally, though errors inevitably exist during the approximation operations, it is applicable when the high-order terms of Taylor series expansion of the cost function are small. Therefore, the next focal point is designing V^* and μ^* on the basis of discrete-time GHJB formulation.

B. Approximation Method and Neural Network Implementation

In this part, the successive approximation method based on the discrete-time GHJB equation is introduced. The main idea is to construct two sequences, i.e., $\{\mu^{(i)}\}$ and $\{V^{(i)}\}$, where $i = 0, 1, 2, \dots$, such that $\mu^{(i)} \rightarrow \mu^*$ and $V^{(i)} \rightarrow V^*$ as $i \rightarrow \infty$.

Generally speaking, if a control function $\mu^{(i)} (\mu^{(i)} \in \Omega_u)$ and a cost function $V^{(i)}$ satisfy relationship that $\text{GHJB}(V^{(i)}, \mu^{(i)}) = 0$, an updated control can be derived by differentiating $H(x, \mu^{(i+1)}(x), V^{(i)}(x)) = 0$ with respect to $\mu^{(i+1)}$. Then, we have

$$\begin{aligned}\mu^{(i+1)}(x) &= -\left(g^T(x)\nabla^2 V^{(i)}(x)g(x) + 2R\right)^{-1} g^T(x) \\ &\quad \times \left(\nabla V^{(i)}(x) + \nabla^2 V^{(i)}(x)(f(x) - x)\right).\end{aligned}\quad (17)$$

In the following, two lemmas are given to show that the updated control function $\mu^{(i+1)}$ is admissible and the related cost function is reduced under its action.

Lemma 2: If $\mu^{(i)}(x) \in \Omega_u$, $x_0 \in \Omega$, and the positive definite and continuously differentiable cost function $V^{(i)}$ satisfies $\text{GHJB}(V^{(i)}, \mu^{(i)}) = 0$ with $V^{(i)}(0) = 0$, the updated control obtained by (17) is an admissible control for system (3) on Ω . In addition, if $V^{(i+1)}$ is the unique positive definite function satisfying the relationship that $\text{GHJB}(V^{(i+1)}, \mu^{(i+1)}) = 0$, then we have $V^{(i+1)}(x_0) \leq V^{(i)}(x_0)$.

Lemma 3: Given an initial control $\mu^{(0)} \in \Omega_u$, by solving $\text{GHJB}(V^{(i)}, \mu^{(i)}) = 0$ and updating the control via (17) with $i = 0, 1, 2, \dots$, the sequence $\{\mu^{(i)}\}$ is convergent as $i \rightarrow \infty$, i.e., $V^{(i)} \rightarrow V^*$ as $i \rightarrow \infty$. Furthermore, $\mu^{(i)} \rightarrow \mu^*$ as $i \rightarrow \infty$.

Remark 3: In this paper, Lemmas 2 and 3 can be easily obtained according to the convergence proof of [11] and [12]. However, it should be noted that the modified utility function $\rho d_M^2(x_k) + U(x_k, \mu(x_k))$ is employed instead of the traditional one $U(x_k, \mu(x_k))$. This reflects the distinguishing feature of the transformed optimal control problem when dealing with the robust control of uncertain nonlinear system.

Note that the discrete-time GHJB equation is easier to deal with than the discrete-time HJB equation in theory, but the closed-form solution still cannot be obtained. Hence, in this part, a neural network is constructed to solve the discrete-time GHJB equation, so that a control function in feedback form can be developed.

In view of the property that neural networks can be employed to approximate smooth functions on a prescribed compact set, the cost function $V(x)$ can be approximated by a neural network as follows:

$$\hat{V}(x) = \hat{\omega}_c^T \sigma_c(x) = \sum_{j=1}^l \hat{\omega}_{cj} \sigma_{cj}(x) \quad (18)$$

where $\hat{\omega}_c = [\hat{\omega}_{c1}, \hat{\omega}_{c2}, \dots, \hat{\omega}_{cl}]^T$ is the weight vector, $\sigma_c(x) = [\sigma_{c1}(x), \sigma_{c2}(x), \dots, \sigma_{cl}(x)]^T$ is the activation function, which is assumed to be second-order differentiable, and l is the number of neurons in the hidden layer, respectively. Note that for any $j = 1, 2, \dots, l$, the activation function $\sigma_{cj}(x)$ is continuous and satisfies $\sigma_{cj}(x) = 0$ when $x = 0$. Moreover, the set $\{\sigma_{cj}(x)\}$ with $j = 1, 2, \dots, l$ is linearly independent.

The weight vector of the neural network will be trained to minimize the residual error in a least squares sense. Substituting (18) into $\text{GHJB}(V, \mu) = 0$, we obtain the discrete-time GHJB equation with a residual error as

$$\text{GHJB}\left(\hat{V}(x) = \sum_{j=1}^l \hat{\omega}_{cj} \sigma_{cj}(x), \mu\right) \triangleq e_c(x). \quad (19)$$

Here, the method of weighted residuals is adopted in order to derive the least squares solution. For any $x \in \Omega$, the weight vector $\hat{\omega}_c$ can be acquired by projecting the residual error onto $\partial e_c(x)/\partial \hat{\omega}_c$ and letting the result be zero, which can be expressed as

$$\left\langle \frac{\partial e_c(x)}{\partial \hat{\omega}_c}, e_c(x) \right\rangle = 0. \quad (20)$$

In (20), the inner product $\langle a(x), b(x) \rangle$ is defined as the Lebesgue integral, i.e., $\langle a(x), b(x) \rangle = \int_{\Omega} a(x)b(x)dx$.

When solving $\hat{\omega}_c$ by expanding (20), the integration is computationally difficult to acquire directly. Instead, an approximate result using the definition of Riemann integration is available. To do this, a mesh with p points over the integral region on Ω is introduced and they are x^1, x^2, \dots, x^p . Here, the size of the mesh is denoted as Δx , which should be chosen as a tradeoff between accuracy and computational complexity. In addition, the number of points in the mesh should satisfy $p > l$. Define

$$\begin{aligned}A &= \left[\delta^T(x)|_{x=x^1}, \delta^T(x)|_{x=x^2}, \dots, \delta^T(x)|_{x=x^p}\right]^T \\ B &= [\eta(x)|_{x=x^1}, \eta(x)|_{x=x^2}, \dots, \eta(x)|_{x=x^p}]^T.\end{aligned}\quad (21)$$

In (21), $\delta(x) = [\delta_1(x), \delta_2(x), \dots, \delta_l(x)]^T$, where

$$\begin{aligned}\delta_i(x) &= \nabla \sigma_{ci}^T(x)(f(x) + g(x)\mu(x) - x) \\ &\quad + \frac{1}{2}(f(x) + g(x)\mu(x) - x)^T \nabla^2 \sigma_{ci}(x)(f(x) \\ &\quad + g(x)\mu(x) - x), \quad i = 1, 2, \dots, l\end{aligned}\quad (22)$$

and $\eta(x) = d_M^2(x) + U(x, \mu(x))$. Then, we have

$$\hat{\omega}_c = -\left(A^T A\right)^{-1} \left(A^T B\right) \quad (23)$$

which implies that the weight vector of the neural network can be obtained effectively. Additionally, by using (18), the control function related to $\hat{V}(x)$ can also be derived, that is

$$\begin{aligned}\hat{\mu}(x) &= -\left(g^T(x)\nabla^2 \hat{V}(x)g(x) + 2R\right)^{-1} g^T(x) \\ &\quad \times \left(\nabla \hat{V}(x) + \nabla^2 \hat{V}(x)(f(x) - x)\right)\end{aligned}\quad (24)$$

where $\hat{V}(x) = \hat{\omega}_c^\top \sigma_c(x)$. This is the approximate control function associated with the immediate weight vector. If we get the convergent weight vector, we can therefore acquire the approximate optimal control of system (3).

Now, we present in detail the design procedure of nonlinear optimal control scheme based on the discrete-time GHJB formulation and neural network.

Step 1: Specify a small positive constant ε and a sufficiently large integer i_{\max} . Construct a neural network to approximate the cost function as $\hat{V}(x) = \hat{\omega}_c^\top \sigma_c(x)$. Set $i = 0$ and select an initial admissible control $\mu^{(0)}(x)$.

Step 2: Apply the least squares method to deal with the equation $\text{GHJB}(\hat{V}^{(0)}, \mu^{(0)}) = 0$, and obtain the weight vector $\hat{\omega}_c^{(0)}$ and the cost function $\hat{V}^{(0)}(x)$. Then, update the control function by using

$$\hat{\mu}^{(i+1)}(x) = -\left(g^\top(x) \nabla^2 \hat{V}^{(i)}(x) g(x) + 2R\right)^{-1} g^\top(x) \times \left(\nabla \hat{V}^{(i)}(x) + \nabla^2 \hat{V}^{(i)}(x)(f(x) - x)\right) \quad (25)$$

where $\hat{V}^{(i)}(x) = \hat{\omega}_c^{(i)\top} \sigma_c(x)$.

Step 3: Set $i = i + 1$.

Step 4: Solve $\text{GHJB}(\hat{V}^{(i)}, \hat{\mu}^{(i)}) = 0$, obtain the weight vector $\hat{\omega}_c^{(i)}$ and the cost function $\hat{V}^{(i)}(x)$, and then derive the updated control $\hat{\mu}^{(i+1)}(x)$ based on (25).

Step 5: If $|\hat{V}^{(i)}(x) - \hat{V}^{(i-1)}(x)| \leq \varepsilon$, go to step 7; otherwise, go to step 6.

Step 6: If $i > i_{\max}$, go to step 7; otherwise, go to step 3.

Step 7: Stop.

After the neural network implementation process, the approximate optimal cost function \hat{V}^* and approximate optimal control $\hat{\mu}^*$ for the nominal system (3) are obtained. Next, we prove that $\hat{\mu}^*$ is a robust feedback control of system (1).

C. Derivation of Robust Controller

In this section, we show in theory that the approximate optimal controller of system (3) is a robust stabilizer of uncertain system (1). This is the main result of the paper.

Theorem 1: For the nominal system (3) with the cost function (4), assume the solution of the discrete-time HJB equation exists. Then, the control function $\hat{\mu}^*$ ensures closed-loop locally asymptotic stability of uncertain nonlinear system (1) if the condition

$$\rho d_M^2(x_k) \geq d^\top(x_k) R d(x_k) + \frac{1}{2} (g(x_k) d(x_k))^\top \nabla^2 \hat{V}^*(x_k) g(x_k) d(x_k) \quad (26)$$

is satisfied.

Proof: Let $\hat{V}^*(x)$ be the approximate solution of the discrete-time HJB equation and $\hat{\mu}^*(x)$ be the approximate optimal control by using the neural-network-based discrete-time GHJB formulation. Now, we prove that $\hat{\mu}^*(x)$ is a solution to the robust control problem, i.e., the equilibrium point $x_k = 0$ of system (1) is asymptotically stable for all possible uncertainties $d(x_k)$.

Since $\hat{V}^*(x)$ and $\hat{\mu}^*(x)$ satisfy the discrete-time HJB equation (16), we can regard $\hat{V}^*(x)$ as a positive definite function and we also obtain

$$\begin{aligned} & \left(\nabla \hat{V}^*(x_k) \right)^\top (f(x_k) + g(x_k) \hat{\mu}^*(x_k) - x_k) \\ & + \frac{1}{2} (f(x_k) + g(x_k) \hat{\mu}^*(x_k) - x_k)^\top \nabla^2 \hat{V}^*(x_k) \\ & \times (f(x_k) + g(x_k) \hat{\mu}^*(x_k) - x_k) \\ & = -\rho d_M^2(x_k) - U(x_k, \hat{\mu}^*(x_k)). \end{aligned} \quad (27)$$

In addition, when considering \hat{V}^* and $\hat{\mu}^*$, the formula (24) suggests that

$$\begin{aligned} & \left(g^\top(x_k) \nabla^2 \hat{V}^*(x_k) g(x_k) + 2R \right) \hat{\mu}^*(x_k) + g^\top(x_k) \\ & \times \left(\nabla \hat{V}^*(x_k) + \nabla^2 \hat{V}^*(x_k)(f(x_k) - x_k) \right) = 0. \end{aligned} \quad (28)$$

Obviously, (28) implies that

$$\begin{aligned} & (f(x_k) + g(x_k) \hat{\mu}^*(x_k) - x_k)^\top \nabla^2 \hat{V}^*(x_k) g(x_k) \\ & + \left(\nabla \hat{V}^*(x_k) \right)^\top g(x_k) + 2(\hat{\mu}^*(x_k))^\top R = 0. \end{aligned} \quad (29)$$

Note that the difference of the approximate optimal cost function \hat{V}^* is

$$\begin{aligned} \Delta \hat{V}^*(x_k) &= \hat{V}^*(x_{k+1}) - \hat{V}^*(x_k) \\ &\approx \left(\nabla \hat{V}^*(x_k) \right)^\top (f(x_k) + g(x_k) \hat{\mu}^*(x_k) + g(x_k) d(x_k) - x_k) \\ &+ \frac{1}{2} (f(x_k) + g(x_k) \hat{\mu}^*(x_k) + g(x_k) d(x_k) - x_k)^\top \\ &\times \nabla^2 \hat{V}^*(x_k) (f(x_k) + g(x_k) \hat{\mu}^*(x_k) + g(x_k) d(x_k) - x_k). \end{aligned} \quad (30)$$

Considering (27) and (29), we can obtain that (30) becomes

$$\begin{aligned} \Delta \hat{V}^*(x_k) &= -\rho d_M^2(x_k) - U(x_k, \hat{\mu}^*(x_k)) \\ &+ (f(x_k) + g(x_k) \hat{\mu}^*(x_k) - x_k)^\top \nabla^2 \hat{V}^*(x_k) \\ &\times g(x_k) d(x_k) + \left(\nabla \hat{V}^*(x_k) \right)^\top g(x_k) d(x_k) \\ &+ \frac{1}{2} (g(x_k) d(x_k))^\top \nabla^2 \hat{V}^*(x_k) g(x_k) d(x_k) \\ &= -\rho d_M^2(x_k) - U(x_k, \hat{\mu}^*(x_k)) - 2(\hat{\mu}^*(x_k))^\top R d(x_k) \\ &+ \frac{1}{2} (g(x_k) d(x_k))^\top \nabla^2 \hat{V}^*(x_k) g(x_k) d(x_k). \end{aligned} \quad (31)$$

By adding and subtracting $d^\top(x_k) R d(x_k)$ to (31), we can further obtain

$$\begin{aligned} \Delta \hat{V}^*(x_k) &= -\rho d_M^2(x_k) - x_k^\top Q x_k + d^\top(x_k) R d(x_k) \\ &+ \frac{1}{2} (g(x_k) d(x_k))^\top \nabla^2 \hat{V}^*(x_k) g(x_k) d(x_k) \\ &- (\hat{\mu}^*(x_k) + d(x_k))^\top R (\hat{\mu}^*(x_k) + d(x_k)) \\ &\leq -\left(\rho d_M^2(x_k) - d^\top(x_k) R d(x_k) \right. \\ &\quad \left. - \frac{1}{2} (g(x_k) d(x_k))^\top \nabla^2 \hat{V}^*(x_k) g(x_k) d(x_k) \right) \\ &\quad - x_k^\top Q x_k. \end{aligned} \quad (32)$$

Considering (26), we can conclude that $\Delta \hat{V}^*(x_k) \leq -x_k^\top Q x_k < 0$ for any $x_k \neq 0$. Therefore, \hat{V}^* is a Lyapunov function and the conditions for Lyapunov local stability theory are satisfied. The proof is completed. ■

The developed robust control strategy is feasible because of the powerfulness of the GHJB-formulation-based optimal control approach. In the following, an example will be taken to demonstrate its effectiveness.

IV. SIMULATION

Consider the following discrete-time nonlinear system:

$$x_{k+1} = \begin{bmatrix} -x_{1k} x_{2k} \\ x_{1k}^2 + 0.8 x_{2k} \end{bmatrix} + \begin{bmatrix} 0 \\ -1 \end{bmatrix} (u(x_k) + d(x_k)) \quad (33)$$

where $x_k = [x_{1k}, x_{2k}]^\top \in \mathbb{R}^2$ and $u(x_k) \in \mathbb{R}$ are the state and control variables, respectively. In (33), $d(x_k) = 2\vartheta x_{2k} \sin x_{1k}^2 \cos x_{2k}$,

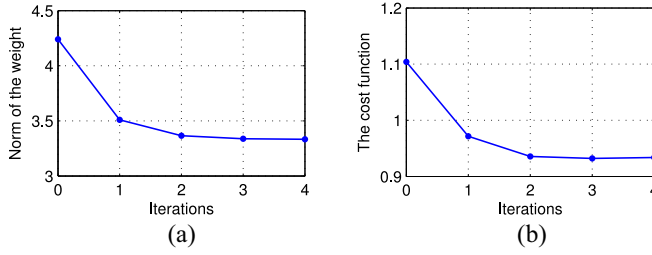


Fig. 1. Simulation results. Convergence process of the (a) norm of weight vector and (b) cost function sequence.

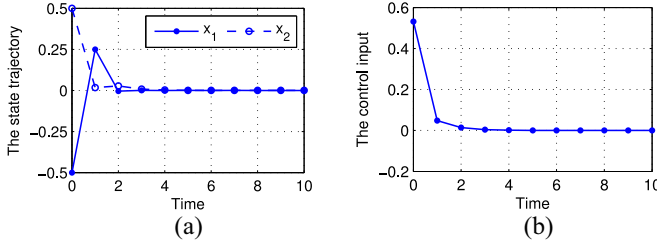


Fig. 2. Simulation results. (a) State trajectory under the action of robust control $\hat{\mu}^*(x)$ when setting $\vartheta = 0.5$. (b) Control trajectory under the action of robust control $\hat{\mu}^*(x)$ when setting $\vartheta = 0.5$.

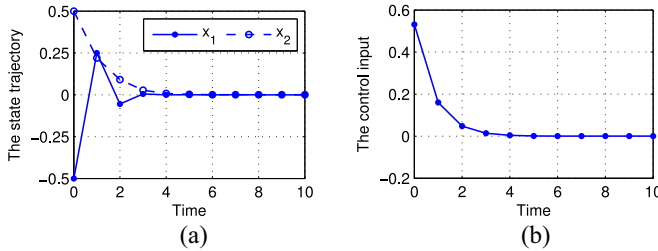


Fig. 3. Simulation results. (a) State trajectory under the action of robust control $\hat{\mu}^*(x)$ when setting $\vartheta = -0.5$. (b) Control trajectory under the action of robust control $\hat{\mu}^*(x)$ when setting $\vartheta = -0.5$.

which represents the unknown perturbation with $\vartheta \in [-0.5, 0.5]$ for simplicity. We choose $d_M(x_k) = \|x_k\|$, which clearly satisfies the bounded condition (2). According to the aforementioned results, the robust stabilization problem can be transformed into the optimal control problem of the nominal system. By selecting $\rho = 1$, $Q = I_2$, and $R = I$, where I denotes the identity matrix with a suitable dimension, the cost function is defined as

$$J(x_k, u) = \sum_{q=k}^{\infty} \left\{ \|x_q\|^2 + x_q^T x_q + u^T(x_q) u(x_q) \right\}. \quad (34)$$

In order to apply the discrete-time GHJB formulation to obtain the approximate optimal control, a neural network is constructed as follows:

$$\begin{aligned} \hat{V}(x) = & \hat{w}_{c1}x_1^2 + \hat{w}_{c2}x_1x_2 + \hat{w}_{c3}x_2^2 + \hat{w}_{c4}x_1^4 + \hat{w}_{c5}x_1^3x_2 + \hat{w}_{c6}x_1^2x_2^2 \\ & + \hat{w}_{c7}x_1x_2^3 + \hat{w}_{c8}x_2^4 + \hat{w}_{c9}x_1^6 + \hat{w}_{c10}x_1^5x_2 + \hat{w}_{c11}x_1^4x_2^2 \\ & + \hat{w}_{c12}x_1^3x_2^3 + \hat{w}_{c13}x_1^2x_2^4 + \hat{w}_{c14}x_1x_2^5 + \hat{w}_{c15}x_2^6. \end{aligned} \quad (35)$$

During the implementation process, the mesh size is set as $\Delta x = 0.02$. The initial state vector and the initial control are chosen as $x_0 = [-0.5, 0.5]^T$ and $\mu^{(0)}(x) = 0.5x_2$, respectively. After five iterations, the weight vector of the neural network converges to $\hat{w}_c^* = [2.5456, 0, 1.6993, -0.9181, 0, -0.5532, 0, -0.7105, 0.1940, 0, 0.1197, 0, 0.1080, 0, 0.1497]^T$. The convergence process of the norm of weight vector is depicted in Fig. 1(a). Besides, the convergence of the cost function sequence is displayed in Fig. 1(b). Then, the approximate

optimal control $\hat{\mu}^*$ of the nominal system is derived, which according to Theorem 1, is also the robust control of the uncertain system (33).

Now, we apply the robust control $\hat{\mu}^*$ to system (33) for ten time steps. Fig. 2(a) and (b) exhibits the state trajectory and control trajectory, respectively, when setting $\vartheta = 0.5$. Besides, Fig. 3(a) and (b) exhibits the state trajectory and control trajectory, respectively, when setting $\vartheta = -0.5$. The simulation results illustrate that the established approximate optimal control ensures closed-loop asymptotic stability of the controlled plant.

V. CONCLUSION

A robust control strategy of a class of affine discrete-time nonlinear systems with uncertainties is established based on the optimal control approach using discrete-time GHJB formulation. By transforming the robust control problem into the optimal control problem, the discrete-time GHJB equation is introduced and solved through the successive approximation method. Additionally, the detailed design procedure via neural network is given, while the numerical simulation is also provided to verify the control performance. In our future work, the robust control of affine discrete-time nonlinear systems under uncertain and unknown environment will be studied by employing the ADP-based optimal control approach. In addition, more comparisons with other traditional robust stabilization methods will be studied in the future.

REFERENCES

- [1] S. Hussain, S. Q. Xie, and P. K. Jamwal, "Robust nonlinear control of an intrinsically complicant robotic gait training orthosis," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 43, no. 3, pp. 655–665, May 2013.
- [2] D. M. Adhyaru, I. N. Kar, and M. Gopal, "Bounded robust control of nonlinear systems using neural network-based HJB solution," *Neural Comput. Appl.*, vol. 20, no. 1, pp. 91–103, Feb. 2011.
- [3] F. Lin, R. D. Brand, and J. Sun, "Robust control of nonlinear systems: Compensating for uncertainty," *Int. J. Control*, vol. 56, no. 6, pp. 1453–1459, Dec. 1992.
- [4] R. W. Beard, G. N. Saridis, and J. T. Wen, "Galerkin approximations of the generalized Hamilton–Jacobi–Bellman equation," *Automatica*, vol. 33, no. 12, pp. 2159–2177, Aug. 1996.
- [5] P. J. Werbos, "Approximate dynamic programming for real-time control and neural modeling," in *Handbook of Intelligent Control: Neural, Fuzzy, and Adaptive Approaches*, D. A. White and D. A. Sofge, Eds. New York, NY, USA: Van Nostrand Reinhold, 1992, ch. 13.
- [6] F. L. Lewis, D. Vrabie, and K. G. Vamvoudakis, "Reinforcement learning and feedback control: Using neural decision methods to design optimal adaptive controllers," *IEEE Control Syst. Mag.*, vol. 32, no. 6, pp. 76–105, Dec. 2012.
- [7] M. Abu-Khalaf and F. L. Lewis, "Nearly optimal state feedback of constrained nonlinear systems using a neural network HJB approach," *Annu. Rev. Control*, vol. 28, no. 2, pp. 239–251, Jan. 2004.
- [8] D. Wang and D. Liu, "Neuro-optimal control for a class of unknown nonlinear dynamic systems using SN-DHP technique," *Neurocomputing*, vol. 121, pp. 218–225, Dec. 2013.
- [9] D. Wang, D. Liu, and H. Li, "Policy iteration algorithm for online design of robust control for a class of continuous-time nonlinear systems," *IEEE Trans. Autom. Sci. Eng.*, vol. 11, no. 2, pp. 627–632, Apr. 2014.
- [10] D. Wang, D. Liu, H. Li, and H. Ma, "Neural-network-based robust optimal control design for a class of uncertain nonlinear systems via adaptive dynamic programming," *Inf. Sci.*, vol. 282, pp. 167–179, Oct. 2014.
- [11] S. Chen and S. Jagannathan, "Generalized Hamilton–Jacobi–Bellman formulation-based neural network control of affine nonlinear discrete-time systems," *IEEE Trans. Neural Netw.*, vol. 19, no. 1, pp. 90–106, Jan. 2008.
- [12] S. Jagannathan, "Neural network control in discrete-time using Hamilton–Jacobi–Bellman formulation," in *Neural Network Control of Nonlinear Discrete-Time Systems*, S. Jagannathan, Ed. Boca Raton, FL, USA: CRC Press, 2006, ch. 9.
- [13] L. Cui, H. Zhang, D. Liu, and Y. S. Kim, "Constrained optimal control of affine nonlinear discrete-time systems using GHJB method," in *Proc. IEEE Symp. Adapt. Dyn. Program. Reinf. Learn.*, Nashville, TN, USA, Mar. 2009, pp. 16–21.