

Robust Tracking Control of Uncertain Nonlinear Systems Using Adaptive Dynamic Programming

Xiong Yang¹(✉), Derong Liu², and Qinglai Wei¹

¹ The State Key Laboratory of Management and Control for Complex Systems, Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China
{xiong.yang, qinglai.wei}@ia.ac.cn

² School of Automation and Electrical Engineering, University of Science and Technology Beijing, Beijing 100083, China
derong@ustb.edu.cn

Abstract. In this paper, we develop an adaptive dynamic programming-based robust tracking control for a class of continuous-time matched uncertain nonlinear systems. By selecting a discounted value function for the nominal augmented error system, we transform the robust tracking control problem into an optimal control problem. The control matrix is not required to be invertible by using the present method. Meanwhile, we employ a single critic neural network (NN) to approximate the solution of the Hamilton-Jacobi-Bellman equation. Based on the developed critic NN, we derive optimal tracking control without using policy iteration. Moreover, we prove that all signals in the closed-loop system are uniformly ultimately bounded via Lyapunov's direct method. Finally, we provide an example to show the effectiveness of the present approach.

Keywords: Adaptive dynamic programming · Robust control · Tracking control · Uncertain nonlinear system

1 Introduction

During the past several decades, robust tracking control of nonlinear systems has attracted considerable attention [1–3]. Many significant approaches have been proposed. Among these methods, the feedback linearization approach is often employed. However, to use the feedback linearization method, the control matrix needs to be invertible. This requirement is usually hard to satisfy in applications.

Recently, adaptive dynamic programming (ADP) [4] is applied to give the optimal tracking control of nonlinear systems. In [5], Heydari and Balakrishnan proposed a single network adaptive critic architecture to obtain the optimal tracking control for continuous-time (CT) nonlinear systems. By employing the architecture, the control matrix was no longer required to be invertible. After that, Modares and Lewis [6] introduced a discounted value function for the CT constrained-input optimal tracking control problem. By proposing an ADP algorithm, the optimal tracking control was obtained without requiring

the control matrix to be invertible either. Though the aforementioned literature provides important insights into deriving optimal tracking control for CT nonlinear systems, the ADP-based robust tracking control for CT uncertain nonlinear systems is not considered.

In this paper, an ADP-based robust tracking control is developed for CT matched uncertain nonlinear systems. By choosing a discounted value function for the nominal augmented error dynamical system, the robust tracking control problem is transformed into an optimal control problem. The control matrix is not required to be invertible in the present method. Meanwhile, a single critic neural network (NN) is used to approximate the solution of the Hamilton-Jacobi-Bellman (HJB) equation. Based on the developed critic NN, the optimal tracking control is obtained without using policy iteration. In addition, all signals in the closed-loop system are proved to be uniformly ultimately bounded (UUB) via Lyapunov's direct method.

The rest of the paper is organized as follows. Preliminaries are presented in Sect. 2. The problem transformation is given in Sect. 3. Approximating the HJB solution via ADP is shown in Sect. 4. Simulation results are provided in Sect. 5. Finally, several conclusions are given in Sect. 6.

2 Preliminaries

Consider the CT uncertain nonlinear system given by

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

where $x(t) \in \mathbb{R}^n$ is the state vector available for measurement, $u(t) \in \mathbb{R}^m$ is the control vector, $f(x(t)) \in \mathbb{R}^n$ and $g(x(t)) \in \mathbb{R}^{n \times m}$ are known functions with $f(0) = 0$, and $\Delta f(x(t)) \in \mathbb{R}^n$ is an unknown perturbation. $f(x) + g(x)u$ is Lipschitz continuous on a compact set $\Omega \subset \mathbb{R}^n$ containing the origin, and system (1) is assumed to be controllable.

Assumption 1. *There exists a constant $g_M > 0$ such that $0 < \|g(x)\| \leq g_M \forall x \in \mathbb{R}^n$. $\Delta f(x) = g(x)d(x)$, where $d(x) \in \mathbb{R}^m$ is unknown function bounded by a known function $d_M(x) > 0$. Moreover, $d(0) = 0$ and $d_M(0) = 0$.*

Assumption 2. *$x_d(t)$ is the desired trajectory of system (1). Meanwhile, $x_d(t)$ is bounded and produced by the command generator model $\dot{x}_d(t) = \eta(x_d(t))$, where $\eta: \mathbb{R}^n \rightarrow \mathbb{R}^n$ is a Lipschitz continuous function with $\eta(0) = 0$.*

Objective of Control: Without the requirement of the control matrix $g(x)$ to be invertible, a robust control scheme based on ADP is developed to keep the state of system (1) following the desired trajectory $x_d(t)$ to a small neighborhood of the origin in the presence of the unknown term $d(x)$.

3 Problem Transformation

Define the tracking error as $e_{\text{err}}(t) = x(t) - x_d(t)$. Then, the tracking error dynamics system is derived as

$$\begin{aligned} \dot{e}_{\text{err}}(t) &= f(x_d(t) + e_{\text{err}}(t)) + g(x_d(t) + e_{\text{err}}(t))u(t) \\ &\quad - \eta(x_d(t)) + \Delta f(x_d(t) + e_{\text{err}}(t)). \end{aligned} \quad (2)$$

In this sense, the robust tracking control can be obtained by giving a control such that, without the requirement of $g(\cdot)$ to be invertible, system (2) is stable in the sense of uniform ultimate boundedness and the ultimate bound is small.

Denote $z(t) = [e_{\text{err}}^{\text{T}}(t), x_d^{\text{T}}(t)]^{\text{T}} \in \mathbb{R}^{2n}$. By Assumption 2 and using (2), we derive an augmented system for the error dynamics as

$$\dot{z}(t) = F(z(t)) + G(z(t))u(t) + \Delta F(z(t)) \quad (3)$$

where $F: \mathbb{R}^{2n} \rightarrow \mathbb{R}^{2n}$ and $G: \mathbb{R}^{2n} \rightarrow \mathbb{R}^{2n \times m}$ are, respectively, defined as

$$F(z(t)) = \begin{bmatrix} f(x_d(t) + e_{\text{err}}(t)) - \eta(x_d(t)) \\ \eta(x_d(t)) \end{bmatrix}, \quad G(z(t)) = \begin{bmatrix} g(x_d(t) + e_{\text{err}}(t)) \\ 0 \end{bmatrix}$$

and $\Delta F(z(t)) = G(z(t))d(z(t))$ with $d(z(t)) \in \mathbb{R}^m$ and $\|d(z(t))\| \leq d_M(z(t))$.

In what follows we show that the robust tracking control problem can be transformed into the optimal control problem with a discounted value function for the nominal augmented error system (i.e., system (3) without uncertainty).

The nominal augmented system is given as

$$\dot{z}(t) = F(z(t)) + G(z(t))u(t). \quad (4)$$

The value function for system (4) is described by

$$V(z(t)) = \int_t^{\infty} e^{-\alpha(\tau-t)} [\rho d_M^2(z(\tau)) + \bar{U}(z(\tau), u(\tau))] d\tau \quad (5)$$

where $\alpha > 0$ is a discount factor, $\rho = \lambda_{\max}(R)$, and $\lambda_{\max}(R)$ denotes the maximum eigenvalue of R , $\bar{U}(z, u) = z^{\text{T}}\bar{Q}z + u^{\text{T}}Ru$ with $\bar{Q} = \text{diag}\{Q, 0_{n \times n}\}$, and $Q \in \mathbb{R}^{n \times n}$ and $R \in \mathbb{R}^{m \times m}$ are symmetric positive definite matrices.

According to [7], the optimal control for system (4) with the value function (5) is

$$u^*(z) = -(1/2)R^{-1}G^{\text{T}}(z)V_z^* \quad (6)$$

where $V_z^* = \partial V^*(z)/\partial z$ and $V^*(z)$ denotes the optimal value of $V(z)$ given in (5). Meanwhile, the corresponding HJB equation is derived as

$$\rho d_M^2(z) + V_z^{*\text{T}}(F(z) + G(z)u^*) - \alpha V^*(z) + z^{\text{T}}\bar{Q}z + u^{*\text{T}}Ru^* = 0. \quad (7)$$

Theorem 1. *Consider the CT nominal system described by (4) with the value function (5). Let Assumptions 1 and 2 hold. Then, the optimal control $u^*(x)$ given in (6) ensures system (2) to be stable in the sense of uniform ultimate boundedness and the ultimate bound can be kept small.*

Proof. Taking the derivative of $V^*(z)$ along the system trajectory $\dot{z} = F(z) + G(z)u^* + \Delta F(z)$, we have $\dot{V}^*(z) = V_z^{*\top}(F(z) + G(z)u^*) + V_z^{*\top}\Delta F(z)$. Noticing that $V_z^{*\top}\Delta F(z) = -2u^{*\top}Rd(z)$ and by (7), we obtain $\dot{V}^*(z) = -\rho d_M^2(z) - z^\top \bar{Q}z - u^{*\top}Ru^* - 2u^{*\top}Rd(z) + \alpha V^*(z)$. Then it can be rewritten as $\dot{V}^*(z) = -\rho d_M^2(z) - e_{\text{err}}^\top Q e_{\text{err}} - (u^* + d(z))^\top R(u^* + d(z)) + d^\top(z)Rd(z) + \alpha V^*(z)$. Observing that $\rho = \lambda_{\max}(R)$ and $\|d(z)\| \leq d_M$, we derive $\dot{V}^*(z) \leq -\lambda_{\min}(Q)\|e_{\text{err}}\|^2 + \alpha V^*(z)$. Because u^* is actually an admissible control, there exists a constant $b_{v^*} > 0$ such that $\|V^*(z)\| \leq b_{v^*}$. Thus, $\dot{V}^*(z) \leq -\lambda_{\min}(Q)\|e_{\text{err}}\|^2 + \alpha b_{v^*}$. So, $\dot{V}^*(z) < 0$ as long as e_{err} is out of the set $\tilde{\Omega}_{e_{\text{err}}} = \{e_{\text{err}} : \|e_{\text{err}}\| \leq \sqrt{\alpha b_{v^*}/\lambda_{\min}(Q)}\}$. By Lyapunov Extension Theorem [8], we obtain that the optimal control u^* guarantees $e_{\text{err}}(t)$ to be UUB with ultimate bound $\sqrt{\alpha b_{v^*}/\lambda_{\min}(Q)}$. Moreover, if α is selected to be very small, then $\sqrt{\alpha b_{v^*}/\lambda_{\min}(Q)}$ can be kept small enough.

From Theorem 1, we can find that the robust tracking control can be obtained by solving the optimal control problem (4) and (5). In other words, we need get the solution of (7). In what follows, a novel ADP-based control scheme is developed to obtain the approximate solution of (7). Before proceeding further, we present an assumption used in [9, 10].

Assumption 3. Let $L_1(z) \in C^1$ be a Lyapunov function candidate for system (4) and satisfied $\dot{L}_1(z) = L_{1z}^\top(F(z) + G(z)u^*) < 0$ with L_{1z} the partial derivative of $L_1(z)$ with respect to z . In addition, there exists a symmetric positive definite matrix $\Lambda(z) \in \mathbb{R}^{2n \times 2n}$ such that $L_{1z}^\top(F(z) + G(z)u^*) = -L_{1z}^\top\Lambda(z)L_{1z}$.

4 Approximate the HJB Solution via ADP

By using the universal approximation property of NNs, $V^*(z)$ given in (7) can be represented by a single-layer NN on a compact set Ω as

$$V^*(z) = W_c^\top \sigma(z) + \varepsilon(z) \quad (8)$$

where $W_c \in \mathbb{R}^{N_0}$ is the ideal NN weight, $\sigma(z) = [\sigma_1(z), \sigma_2(z), \dots, \sigma_{N_0}(z)]^\top \in \mathbb{R}^{N_0}$ is the activation function with $\sigma_j(z) \in C^1(\Omega)$ and $\sigma_j(0) = 0$, the set $\{\sigma_j(z)\}_1^{N_0}$ is often selected to be linearly independent, N_0 is the number of neurons, and $\varepsilon(z)$ is the NN function reconstruction error.

Substituting (8) into (6), we have

$$u^*(z) = -(1/2)R^{-1}G^\top(z)\nabla\sigma^\top W_c + \varepsilon_{u^*} \quad (9)$$

where $\nabla\sigma = \partial\sigma(z)/\partial z$ and $\varepsilon_{u^*} = -(1/2)R^{-1}G^\top(z)\nabla\varepsilon$. Meanwhile, by using (8), (7) becomes

$$W_c^\top \nabla\sigma F - \alpha W_c^\top \sigma + z^\top \bar{Q}z + \rho d_M^2(z) - (1/4)W_c^\top \nabla\sigma \mathcal{A} \nabla\sigma^\top W_c = \varepsilon_{\text{HJB}} \quad (10)$$

where $\mathcal{A} = G(z)R^{-1}G^\top(z)$ and $\varepsilon_{\text{HJB}} = -\nabla\varepsilon^\top F + \alpha\varepsilon + (1/2)W_c^\top \nabla\sigma \mathcal{A} \nabla\varepsilon + (1/4)\nabla\varepsilon^\top \mathcal{A} \nabla\varepsilon$ is the HJB approximation error [11].

Due to the unavailability of W_c , $u^*(z)$ given in (9) cannot be implemented in real control process. Therefore, we use a critic NN to approximate $V^*(z)$ as

$$\hat{V}(z) = \hat{W}_c^T \sigma(z) \quad (11)$$

where \hat{W}_c is the estimated weight of the ideal weight W_c . The weight estimation error for the critic NN is defined as $\tilde{W}_c = W_c - \hat{W}_c$.

Using (11), the estimated value of optimal control $u^*(z)$ is

$$\hat{u}(z) = -(1/2)R^{-1}G^T(z)\nabla\sigma^T\hat{W}_c. \quad (12)$$

Combining (7), (11) and (12), we obtain the residual error as $\delta = \hat{W}_c^T \nabla\sigma F - \alpha\hat{W}_c^T \sigma + z^T \bar{Q}z + \rho d_M^2(z) - (1/4)\hat{W}_c^T \nabla\sigma \mathcal{A} \nabla\sigma^T \hat{W}_c$. By utilizing (10), we have $\delta = -\tilde{W}_c^T \phi + (1/4)\tilde{W}_c^T \nabla\sigma \mathcal{A} \nabla\sigma^T \tilde{W}_c + \varepsilon_{\text{HJB}}$ with $\phi = \nabla\sigma(F(z) + G(z)\hat{u}) - \alpha\sigma(z)$.

To get the minimum value of δ , we develop a novel critic NN tuning law as

$$\begin{aligned} \dot{\hat{W}}_c = & -\gamma\bar{\phi}\left(Y(z) + \rho d_M^2(z) - (1/4)\hat{W}_c^T \nabla\sigma \mathcal{A} \nabla\sigma^T \hat{W}_c\right) + \frac{\gamma}{2}\Sigma(z, \hat{u})\nabla\sigma \mathcal{A} L_{1z} \\ & + \gamma\left((K_1\varphi^T - K_2)\hat{W}_c + (1/4)\nabla\sigma \mathcal{A} \nabla\sigma^T \hat{W}_c \frac{\varphi^T}{m_s} \hat{W}_c\right) \end{aligned} \quad (13)$$

where $\bar{\phi} = \phi/m_s^2$, $\varphi = \phi/m_s$, $m_s = 1 + \phi^T \phi$, $Y(z) = \hat{W}_c^T \nabla\sigma F - \alpha\hat{W}_c^T \sigma + z^T \bar{Q}z$, L_{1z} is given as in Assumption 3, K_1 and K_2 are tuning parameter matrices with suitable dimensions, and $\Sigma(z, \hat{u})$ is an indicator function defined as

$$\Sigma(z, \hat{u}) = \begin{cases} 0, & \text{if } L_{1z}^T(F(z) + G(z)\hat{u}) < 0, \\ 1, & \text{otherwise.} \end{cases} \quad (14)$$

Then, we obtain the weight estimation error dynamics of the critic NN as

$$\begin{aligned} \dot{\tilde{W}}_c = & \gamma\bar{\phi}\left(-\tilde{W}_c^T \phi + (1/4)\tilde{W}_c^T \nabla\sigma \mathcal{A} \nabla\sigma^T \tilde{W}_c + \varepsilon_{\text{HJB}}\right) - \frac{\gamma}{2}\Sigma(z, \hat{u})\nabla\sigma \mathcal{A} L_{1z} \\ & - \gamma\left((K_1\varphi^T - K_2) - (1/4)\nabla\sigma \mathcal{A} \nabla\sigma^T (W_c - \tilde{W}_c) \frac{\varphi^T}{m_s}\right) (W_c - \tilde{W}_c). \end{aligned} \quad (15)$$

In what follows we develop a theorem to show the stability of all signals in the closed-loop system. Before proceeding further, an assumption is provided as follows.

Assumption 4. W_c is bounded by a known constant $W_M > 0$. There exist constants $b_\varepsilon > 0$ and $b_{\varepsilon z} > 0$ such that $\|\varepsilon(z)\| < b_\varepsilon$ and $\|\nabla\varepsilon(z)\| < b_{\varepsilon z} \forall z \in \hat{\Omega}$. There exists a constant $b_{\varepsilon_{u^*}} > 0$ such that $\|\varepsilon_{u^*}\| \leq b_{\varepsilon_{u^*}}$. In addition, there exist constants $b_\sigma > 0$ and $b_{\sigma z} > 0$ such that $\|\sigma(z)\| \leq b_\sigma$ and $\|\nabla\sigma(z)\| \leq b_{\sigma z} \forall z \in \hat{\Omega}$.

Theorem 2. Consider the CT nominal system given by (4) with associated HJB equation (7). Let Assumptions 1-4 hold and take the control input for system (4) as given in (12). Meanwhile, let the critic NN weight tuning law be (13). Then, the function L_{1z} and the weight estimation error \tilde{W}_c are guaranteed to be UUB.

Proof. We provide an outline of the proof due to the space limit. Consider the Lyapunov function candidate $L(t) = L_1(z) + (1/2)\tilde{W}_c^\top \gamma^{-1} \tilde{W}_c$. Taking the time derivative of $L(t)$, we have $\dot{L}(t) = L_{1z}^\top (F(z) + G(z)\hat{u}) + \dot{\tilde{W}}_c^\top \gamma^{-1} \tilde{W}_c$. By using (15) and simplification, and noticing that $\dot{z} = F(z) + G(z)\hat{u}$, we obtain

$$\dot{L}(t) \leq L_{1z}^\top \dot{z} - \lambda_{\min}(M) \|\mathcal{Z}\|^2 + b_N \|\mathcal{Z}\| - (1/2)\Sigma(z, \hat{u}) L_{1z}^\top \mathcal{A} \nabla \sigma^\top \tilde{W}_c \quad (16)$$

where $\mathcal{Z} = [\tilde{W}_c^\top \varphi, \tilde{W}_c^\top]^\top$, b_N is the upper bound of $\|N\|$, M and N are, respectively, given as

$$M = \begin{bmatrix} I & -\frac{W_c^\top \mathcal{B}}{8m_s} - \frac{K_1^\top}{2} \\ -\frac{\mathcal{B}W_c}{8m_s} - \frac{K_1}{2} & K_2 - \frac{\varphi^\top W_c \mathcal{B}}{4m_s} \end{bmatrix}, \quad N = \begin{bmatrix} \frac{\varepsilon_{HJB}}{m_s} \\ \frac{\mathcal{B}W_c \varphi^\top W_c}{4m_s} + K_2 W_c - K_1 \varphi^\top W_c \end{bmatrix}$$

with $\mathcal{B} = \nabla \sigma \mathcal{A} \nabla \sigma^\top$. Due to the definition of $\Sigma(z, \hat{u})$ given in (14), we divide (16) into the following two cases for discussion:

- (i) $\Sigma(z, \hat{u}) = 0$. In this circumstance, we have $L_{1z}^\top \dot{z} < 0$. By employing *dense property* of \mathbb{R} [12], we can obtain a positive constant β such that $0 < \beta \leq \|\dot{z}\|$ implies $L_{1z}^\top \dot{z} \leq -\|L_{1z}\| \beta < 0$. Then (16) can be developed as $\dot{L}(t) \leq -\|L_{1z}\| \beta + (1/4)b_N^2 / \lambda_{\min}(M) - \lambda_{\min}(M) (\|\mathcal{Z}\| - (1/2)b_N / \lambda_{\min}(M))^2$. Notice that $\|\mathcal{Z}\| \leq \sqrt{1 + \|\varphi\|^2} \|\tilde{W}_c\| \leq (\sqrt{5}/2) \|\tilde{W}_c\|$. Therefore, $\dot{L}(t) < 0$ is valid only if $\|L_{1z}\| > b_N^2 / (4\beta \lambda_{\min}(M))$ or $\|\tilde{W}_c\| > 2b_N / (\sqrt{5} \lambda_{\min}(M))$.
- (ii) $\Sigma(z, \hat{u}) = 1$. In this case, (16) can be developed as $\dot{L}(t) \leq L_{1z}^\top (F(z) + G(z)u^*) + L_{1z}^\top G(z)(\hat{u} - u^*) - \lambda_{\min}(M) \|\mathcal{Z}\|^2 + b_N \|\mathcal{Z}\| - \frac{1}{2} L_{1z}^\top \mathcal{A} \nabla \sigma^\top \tilde{W}_c$. By using Assumption 3 and similar with (i), we can obtain that $\dot{L}(t) < 0$ is valid only if $\|L_{1z}\| > g_M b_{\varepsilon_{u^*}} / (2\lambda_{\min}(\Lambda(z))) + \sqrt{\ell} / \lambda_{\min}(\Lambda(z))$ or $\|\tilde{W}_c\| > b_N / (\sqrt{5} \lambda_{\min}(M)) + \sqrt{4\ell / (5\lambda_{\min}(M))}$, where $\ell = g_M^2 b_{\varepsilon_{u^*}}^2 / (4\lambda_{\min}(\Lambda(z))) + b_N^2 / (4\lambda_{\min}(M))$.

Combining (i) and (ii) and using the standard Lyapunov Extension Theorem [8], we derive that the function L_{1z} and the weight estimation error \tilde{W}_c are UUB.

5 Simulation Results

Consider the CT uncertain nonlinear system given by

$$\begin{aligned} \dot{x}_1 &= -x_1 + x_2 \\ \dot{x}_2 &= -(x_1 + 1)x_2 - 49x_1 + u + q \cos^3(x_1) \sin(x_2) \end{aligned} \quad (17)$$

where $x = [x_1, x_2]^\top \in \mathbb{R}^2$, and the uncertain term $d(x) = q \cos^3(x_1) \sin(x_2)$ with unknown parameter $q \in [-1, 1]$. We choose $d_M(x) = \|x\|$. The reference trajectory x_d is generated by $\dot{x}_{1d} = x_{2d}$ and $\dot{x}_{2d} = -49x_{1d}$ with the initial condition $x_d(0) = [0.2, 0.4]^\top$. Then the augmented tracking error system is derived as

$$\dot{z} = C(z) + D(z)(u + d(z)) \quad (18)$$

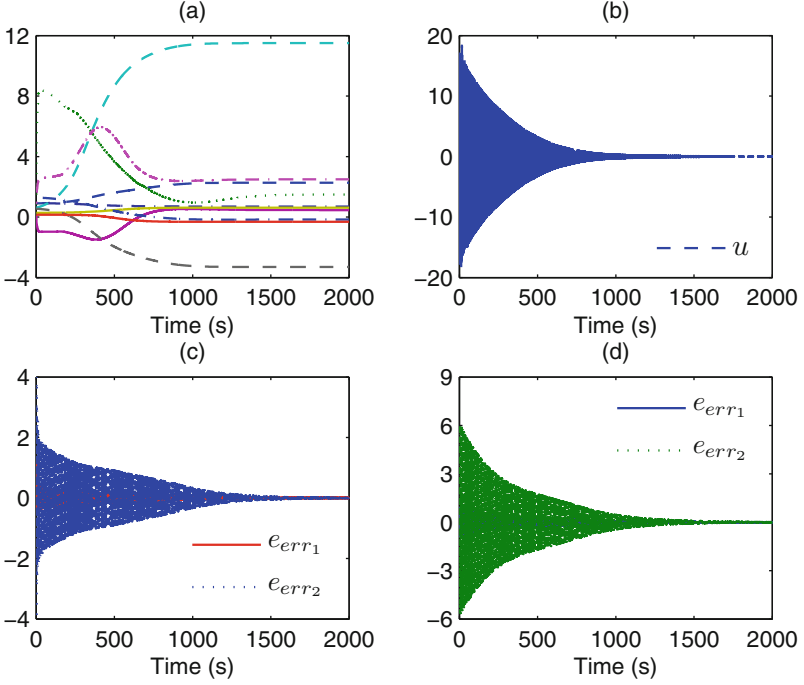


Fig. 1. (a) Convergence of critic NN weights (b) Control input u (c) Evolution of tracking errors $e_{err_i}(t)$ ($i = 1, 2$) during NN learning process (d) Tracking errors e_{err_i} ($i = 1, 2$) between the state of system (17) and the desired trajectory x_d under the approximate optimal control

where $z = [z_1, z_2, z_3, z_4]^T = [e_{err_1}, e_{err_2}, x_{1d}, x_{2d}]^T$ with $e_{err_i} = x_i - x_{id}$, and $D(z) = [0, 1, 0, 0]^T$, and $C(z) = [-z_1 + z_2 - z_3; -(z_1 + z_3)(z_2 + z_4) - 49z_1 - z_2 - z_4; z_4; -49z_3]$. The nominal augmented system is $\dot{z} = C(z) + D(z)u$ with $C(z)$ and $D(z)$ is given in (18). The cost function $V(z)$ for nominal augmented error system is given as (5), where $R = 1$ and $Q = 2I_2$. The activation function of the critic NN is chosen with $N_0 = 10$ as $\sigma(x) = [z_1^2, z_2^2, z_3^2, z_4^2, z_1z_2, z_1z_3, z_1z_4, z_2z_3, z_2z_4, z_3z_4]^T$, and the weight of the critic NN is written as $\hat{W}_c = [W_{c1}, W_{c2}, \dots, W_{c10}]^T$.

The initial system state is $x(0) = [0.5, -0.5]^T$, and the initial weight for the critic NN is selected randomly within an interval of $[0, 1]$, which implies that no initial stabilizing control is required. In addition, $\alpha = 0.15$ and $\gamma = 0.5$. The developed control algorithm is implemented via (12) and (13). The computer simulation results are shown by Fig. 1(a)–(d). Figure 1(a) and (b) indicate convergence of critic NN weights and control input u , respectively. Figure 1(c) shows the evolution of tracking errors e_{err_i} ($i = 1, 2$) during NN learning process. Figure 1(d) illustrates tracking errors e_{err_i} ($i = 1, 2$) between the state of system (17) and the desired trajectory x_d under the approximate optimal control. From simulation results, it is observed that the state $x(k)$ tracks the desired trajectory $x_d(k)$ very well, and all signals in the closed-loop system are bounded.

6 Conclusions

We have developed an ADP-based robust tracking control for CT matched uncertain nonlinear systems. The robust tracking control is obtained without the requirement of the control matrix to be invertible. By using Lyapunov's method, the stability of the closed-loop system is proved, and all signals involved are UUB. The computer simulation results show that the developed control scheme can perform successfully control and attain the desired performance. In our future work, we focus on studying robust control for CT nonaffine nonlinear systems.

Acknowledgments. This work was supported in part by the National Natural Science Foundation of China under Grants 61233001, 61273140, 61304086, and 61374105, in part by Beijing Natural Science Foundation under Grant 4132078, and in part by the Early Career Development Award of the State Key Laboratory of Management and Control for Complex Systems (SKLMCCS).

References

1. Godbole, D.N., Sastry, S.S.: Approximate decoupling and asymptotic tracking for MIMO systems. *IEEE Trans. Autom. Control* **40**(3), 441–450 (1995)
2. Chang, Y.C.: An adaptive H_∞ tracking control for a class of nonlinear multiple-input multiple-output (MIMO) systems. *IEEE Trans. Autom. Control* **46**(9), 1432–1437 (2001)
3. Liu, D., Yang, X., Wang, D., Wei, Q.: Reinforcement-learning-based robust controller design for continuous-time uncertain nonlinear systems subject to input constraints. *IEEE Trans. Cybern.* **45**(7), 1372–1385 (2015)
4. Werbos, P.J.: Beyond Regression: New Tools for Prediction and Analysis in the Behavioral Sciences. Ph.D. Thesis, Harvard University, Cambridge, MA (1974)
5. Heydari, A., Balakrishnan, S.: Fixed-final-time optimal tracking control of input-affine nonlinear systems. *Neurocomputing* **129**, 528–539 (2014)
6. Modares, H., Lewis, F.L.: Optimal tracking control of nonlinear partially-unknown constrained-input systems using integral reinforcement learning. *Automatica* **50**(7), 1780–1792 (2014)
7. Liu, D., Yang, X., Li, H.: Adaptive optimal control for a class of nonlinear partially uncertain dynamic systems via policy iteration. In: 3rd International Conference on Intelligent Control and Information Processing, Dalian, China, pp. 92–96 (2012)
8. Lewis, F.L., Jagannathan, S., Yesildirak, A.: *Neural Network Control of Robot Manipulators and Nonlinear Systems*. Taylor & Francis, London (1999)
9. Yang, X., Liu, D., Wei, Q.: Online approximate optimal control for affine nonlinear systems with unknown internal dynamics using adaptive dynamic programming. *IET Control Theor. Appl.* **8**(16), 1676–1688 (2014)
10. Dierks, T., Jagannathan, S.: Optimal control of affine nonlinear continuous-time systems. In: American Control Conference, Baltimore, MD, USA, pp. 1568–1573 (2010)
11. Abu-Khalaf, M., Lewis, F.L., Huang, J.: Neurodynamic programming and zero-sum games for constrained control systems. *IEEE Trans. Neural Netw.* **19**(7), 1243–1252 (2008)
12. Rudin, W.: *Principles of Mathematical Analysis*. McGraw-Hill Publishing Co., New York (1976)