# Output-Feedback Based Simplified Optimized Backstepping Control for Strict-Feedback Systems with Input and State Constraints

Jiaxin Zhang, Kewen Li, and Yongming Li, *Senior Member, IEEE*

***Abstract*—In this paper, an adaptive neural-network (NN) output feedback optimal control problem is studied for a class of strict-feedback nonlinear systems with unknown internal dynamics, input saturation and state constraints. Neural networks are used to approximate unknown internal dynamics and an adaptive NN state observer is developed to estimate immeasurable states. Under the framework of the backstepping design, by employing the actor-critic architecture and constructing the tan-type Barrier Lyapunov function (BLF), the virtual and actual optimal controllers are developed. In order to accomplish optimal control effectively, a simplified reinforcement learning (RL) algorithm is designed by deriving the updating laws from the negative gradient of a simple positive function, instead of employing existing optimal control methods. In addition, to ensure that all the signals in the closed-loop system are bounded and the output can follow the reference signal within a bounded error, all state variables are confined within their compact sets all times. Finally, a simulation example is given to illustrate the effectiveness of the proposed control strategy.**

***Index Terms*—Backstepping design, immeasurable states, neural-networks (NNs), optimal control, state constraints.**

## I. Introduction

IN the last decade, fuzzy logic systems (FLSs) and NNs were widely used in adaptive backstepping recursive control design [1]–[3]. In [1], direct adaptive NN control was presented for a class of nonlinear systems with unknown nonlinearities. The authors focused on adaptive fuzzy tracking control in [2] for a class of nonlinear systems. The result [3] developed two different backstepping NN control approaches for a class of strict-feedback systems with unknown nonlinearities. In [4], the fuzzy logic systems and error transformation-based method were used in online learning of completely unknown dynamics and prescribed performance tracking, respectively. The authors developed a finite-time adaptive fuzzy control strategy for a class of nonlinear strict-feedback systems in [5]. Furthermore, the authors in [6] proposed a global nested PID control method for nonlinear systems with unknown system nonlinearities without linearized approximators. However, it is worth mentioning that the above-mentioned adaptive backstepping control methods all assume that the states of the systems are measurable and can be used for control design directly.

As pointed out in [7]–[10], in practice, state variables were often unmeasured for many nonlinear systems. The authors in [7]–[10] designed different state observers, and some intelligent adaptive output feedback control approaches were developed for a class of uncertain nonlinear systems with immeasurable states. Although the great progress has been made in intelligent adaptive control for nonlinear systems, the constraint problems were not fully considered.

In engineering control, saturation, dead zones and time-delay are common phenomena, all stemming from the existence of control constraints. Once the control is constrained, the stability of the nonlinear system is often difficult to guarantee. In [11]–[18], the control problems for nonlinear systems with full-state constraints and partial state constraints were studied. The stability was guaranteed without violation of any constraints. In order to clarify the effect of control constraints on system stability, many scholars investigated such problems based on the BLF. The authors proposed an indirect adaptive fuzzy controller in [19] for a class of uncertain nonlinear systems with input and output constrains. In [20], an adaptive fault-tolerant control (FTC) scheme was proposed for a class of nonlinear systems with control inputs and system state constraints. The authors designed an adaptive fuzzy control scheme in [21] for a class of uncertain nonlinear systems with input saturation and output constraints. In [22], the authors addressed the cooperative control problem for multiple high-speed trains, which guaranteed that the speed and the position of high-speed trains were confined to specific speed limitations, and allowed distances ratified by the automatic train protection and the moving authority, respectively. Even though various intelligent control strategies [11]–[22] have been devised in the constraints problem for nonlinear dynamics, optimization in control design and stability analysis has not been considered therein.

As the foremost branch of modern control theory, optimal control was developed by Bellman [23] and Chambers [24] 50 years ago. Since then, some significant results were reported,

Manuscript received December 14, 2020; revised February 7, 2021; accepted March 14, 2021. This work was supported by National Natural Science Foundation of China (61822307, 61773188). Recommended by Associate Editor Weinan Gao. *(Corresponding author: Yongming Li.)*

Citation: J. X. Zhang, K. W. Li, and Y. M. Li, "Output-feedback based simplified optimized backstepping control for strict-feedback systems with input and state constraints," *IEEE/CAA J. Autom. Sinica*, vol. 8, no. 6, pp. 1119–1132, Jun. 2021.

J. X. Zhang and Y. M. Li are with the College of Science, Liaoning University of Technology, Jinzhou 121001, China (e-mail: z_j_x_2019@163.com; liyongming1981@163.com).

K. W. Li is with the Institute of Automation, Qufu Normal University, Qufu 273165, China (e-mail: likewen2018@163.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/JAS.2021.1004018

for example in [25]–[33]. In [25], a novel RL-based robust adaptive controller was developed for the continuous-time (CT) uncertain nonlinear systems with input constraints. The authors developed an adaptive RL solution in [26] for the infinite-horizon optimal control problem of constrained-input continuous-time nonlinear systems in the presence of nonlinearities with unknown structures. In [27], an optimal NN control scheme was presented for CT nonlinear systems with asymmetric input constraints. The authors in [28] proposed an integral reinforcement learning (IRL) algorithm on an actor-critic structure for a class of affine nonlinear systems, wherein the partially-unknown constrained-input was considered. The finite-time optimal control problem was studied in [29] for the high-order nonlinear systems whose powers were positive odd ratio numbers. However, all of the above adaptive optimal control methods are limited to affine nonlinear systems and thus cannot be applied to nonlinear systems with strict-feedback. To handle this issue, a control technique called optimized backstepping (OB) was first proposed in [30] by implementing tracking control for a class of strict-feedback systems. Recently, the authors in [31] investigated an adaptive RL optimal control design problem for a class of nonstrict-feedback discrete-time systems. In order to accomplish optimal control effectively, the authors designed a simplified RL algorithm in [32] instead of employing the existing RL-based optimal control methods.

Although an optimized control method was developed in [32] based on the OB technique using simplified RL for nonlinear systems, input saturation and state constraints under unpredictable systems states were not considered. Based on the above results, this paper proposes an optimal control scheme based on NN approximation for a class of strict-feedback systems with unknown dynamics, input saturation and state constraints. Compared with the existing works, the main contributions of this paper are listed in the following.

1) In this paper, an adaptive NN backstepping output feedback simplified optimal control method is proposed for a class of uncertain nonlinear systems with unmeasured states, input saturation and state constraints. The tan-type barrier optimal cost functions are constructed for subsystems. In contrast with [30], the method proposed here does not require priori knowledge due to the utilization of the state observer.

2) By separating the optimal value function into a novel error form, the proposed control strategy can effectively solve the optimal tracking control problem. Unlike [30] and [32], this paper adopts a stepwise optimization strategy to analyze the stability of each step of the system. Each controller is designed in this paper to be the optimal solution for the corresponding subsystem, thus optimizing the control of the whole system.

## II. Preliminaries

### A. Problem Statement

Consider the following strict-feedback nonlinear systems as:

$$\begin{cases} \dot{x}_i = f_i(\bar{x}_i) + x_{i+1}, 1 \le i \le n-1 \\ \dot{x}_n = f_n(\bar{x}_n) + u_s \\ y = x_1 \end{cases} \tag{1}$$

where the state $\bar{x}_n = [x_1, x_2, \ldots, x_n]^T \in \mathbb{R}^n$ and $y \in \mathbb{R}$ is the output of system, $f_i(\bar{x}_i)$ $(i = 1, 2, \ldots, n)$ are unknown smooth functions. $f_i(\bar{x}_i) + x_{i+1}$ and $f_n(\bar{x}_n) + u_s$ are assumed Lipschitz continuous and stabilizable on the sets containing the origin. $u_s \in \mathbb{R}$ denotes the plant input subjected to the saturation described by

$$u_s = \begin{cases} \text{sgn}(u)\bar{u}_s & |u| \ge \bar{u}_s \\ u & |u| < \bar{u}_s \end{cases} \tag{2}$$

where $\bar{u}_s$ is the saturation bound of $u$ and $u$ is the control input.

*Assumption 1 [33], [34]:* Assume that all the states (expect output $y$) are immeasurable and constrained in compact sets, i.e., $|x_i| < k_{ci}$, $(i = 1, \ldots, n)$, where $k_{ci} > 0$ is a known constant.

*Assumption 2 [35], [36]:* The neural networks approximation error $\varepsilon_f = [\varepsilon_{1,f}, \ldots, \varepsilon_{n,f}]^T$ is bounded, i.e., $\|\varepsilon_f\| \le \varepsilon_{fM}$. The neural network weight $W_f^*$ is bounded by a known positive constant $W_{fM}$, i.e., $\|W_f^*\| \le W_{fM}$.

*Control Objective:* The control objective of this paper is to obtain a NN backstepping output feedback optimal control that not only stabilizes system (1), but also minimizes the value function, while ensuring that all the closed-loop signals are guaranteed to be uniformly ultimately bounded (UUB). All the system states are ensured not to transgress their constrained sets so that the output $y$ can track the reference signal $y_r$.

### B. Neural Networks

It is well known that NNs can approximate an unknown continuous function $f(x): \mathbb{R}^n \to \mathbb{R}^m$ over a compact set $D$. Then, for any constant $\varepsilon > 0$, there exists a radial-basis-function NN (RBFNN) $W^T S(x)$ such that $\sup_{x \in D} |f(x) - W^T S(x)| < \varepsilon$, where $x \in \Omega_x \subset \mathbb{R}^q$ is the input vector, $n$ is a positive integer, $W \in \mathbb{R}^{r \times m}$ is the NN weight and the neuron number is $r$. Each element $S_i(x)$ $(i = 1, \ldots, r)$ of vector $S(x)$ is a basis function with

$$S_i(x) = \exp\Big(-\frac{(x-\mu_i)^T (x-\mu_i)}{\sigma_i^2}\Big)$$

where $\mu_i \in \mathbb{R}^n$ is the center vector and $\sigma_i$ is the width of Gaussian function.

*Lemma 1:* If the continuous function $V(t) \in \mathbb{R}$ satisfies $\dot{V}(t) \le -cV(t) + D$, where $c > 0$ and $D > 0$ are constants, then the following inequality holds:

$$V(t) \le V(t_0) e^{-c(t-t_0)} + \frac{D}{c}.$$

*Lemma 2 (Young's Inequality):* For any vectors $x, y \in \mathbb{R}^n$, the following Young's inequality holds:

$$x^T y \le (\eta^a / a)||x||^a + (1/b\eta^b)||y||^b$$

where $\eta > 0$, $a > 1$, $b > 1$, and $(a-1)(b-1) = 1$.

## III. Main Result

### A. State Observer Design

In this section, a state observer needs to be designed to estimate the unmeasured states. Then, under the actor-critic architecture, a NN adaptive backstepping output feedback optimal controller will be designed based on the designed state observer. Finally, a stability analysis of the closed-loop system is given to prove our main conclusions. Rewrite system (1) as the following state space expression form:

$$\dot{x} = Ax + Ly + \sum_{i=1}^{n} B_i f_i(\bar{x}_i) + B_n u_s$$
$$y = Cx \tag{3}$$

where $A = \begin{bmatrix} -l_1 & & \\ \vdots & I & \\ -l_n & 0 \; \cdots & 0 \end{bmatrix}$, $L = \begin{bmatrix} l_1 \\ \vdots \\ l_n \end{bmatrix}$, $x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$, $B_i = [0 \; \dots \; 1 \; \dots \; 0]^T$, $B_n = [0 \; \dots \; 1]^T$, $C = [1 \; 0 \; \dots \; 0]$, $A$ is a strict Hurwitz matrix and $l_i$ $(i = 1, ..., n)$ are observer gains.

Thus, for a given positive definite matrix $Q$, there exists a matrix $P > 0$ that satisfies the following equation:

$$A^T P + PA = -Q. \tag{4}$$

Since $f_i(\bar{x}_i)$ is an unknown continuous function, $f_i(\bar{x}_i)$ can be identified by the NNs $\hat{f}_i(\hat{\bar{x}}_i | \hat{W}_{i,f}) = \hat{W}_{i,f}^T S_{i,f}(\hat{\bar{x}}_i)$ $(1 \le i \le n)$, and we assume that

$$f_i(\bar{x}_i) = W_{i,f}^{*T} S_{i,f}(\bar{x}_i) + \varepsilon_{i,f}(\bar{x}_i) \tag{5}$$

where $W_{i,f}^{*T}$ and $\varepsilon_{i,f}(\bar{x}_i)$ are the ideal weight vector and the approximation error, respectively, and $\hat{W}_{i,f}$ is the estimate of $W_{i,f}^*$.

Since the state variables in the system are immeasurable, to achieve the purpose of output feedback control design, the nonlinear state observer is designed as follows:

$$\dot{\hat{x}}_i = \hat{f}_i(\hat{\bar{x}}_i | \hat{W}_{i,f}) + \hat{x}_{i+1} + l_i(y - \hat{y}), \; 1 \le i \le n$$
$$\dot{\hat{x}}_n = \hat{f}_n(\hat{\bar{x}}_n | \hat{W}_{n,f}) + h(u) + l_n(y - \hat{y})$$
$$\hat{y} = \hat{x}_1 \tag{6}$$

where $h(u) = \bar{u}_s \times \tanh(u/\bar{u}_s) = \bar{u}_s(e^{u/\bar{u}_s} - e^{-u/\bar{u}_s})/e^{u/\bar{u}_s} + e^{-u/\bar{u}_s}$ is a smooth function to approximate the saturation of the system. Therefore, (2) can be expressed as $u_s = h(u) + \rho(u) = \bar{u}_s \times \tanh(u/\bar{u}_s) + \rho(u)$, where $\rho(u) = u_s - h(u)$ is a bounded function, and $|\rho(u)| = |u_s - h(u)| \le \bar{u}_s(1 - \tanh(1)) = m$, $m > 0$ is a constant. Note that within the bound $0 \le |u| \le \bar{u}_s$, $\rho(u)$ grows from 0 to $m$, and $|u|$ changes from 0 to $\bar{u}_s$. Outside of this range, $\rho(u)$ decreases from $m$ to 0.

Then, rewrite (6) as the following form:

$$\dot{\hat{x}} = \sum_{i=1}^{n} B_i[\hat{f}_i(\hat{\bar{x}}_i | \hat{W}_{i,f})] + Ly + A\hat{x} + B_n h(u)$$
$$\hat{y} = C\hat{x} \tag{7}$$

where $\hat{x}_i$ is the estimate of $x_i$.

From (1) and (7), the following error equation can be obtained:

$$\dot{e} = B_n \rho(u) + Ae + F - \hat{W}_f^T S_f \tag{8}$$

where $F = [f_1(x_1), \dots, f_n(\bar{x}_n)]^T$, $e = [e_1, \dots, e_n]^T$ and $e_i = x_i - \hat{x}_i$, $i = 1, \dots, n$, $W_f^{*T} = \text{diag}\{W_{1,f}^{*T}, \dots, W_{n,f}^{*T}\}$ is estimated by $\hat{W}_f^T = \text{diag}\{\hat{W}_{1,f}^T, \dots, \hat{W}_{n,f}^T\}$ and $S_f = [S_{1,f}(\hat{x}_1), \dots, S_{n,f}(\hat{\bar{x}}_n)]^T$.

*Theorem 1:* The NN weight estimate $\hat{W}_f$ is updated by

$$\dot{\hat{W}}_f = \eta_W S_f e_1 C A^{-1} - \rho_W e_1 \hat{W}_f \tag{9}$$

where $\eta_w$ and $\rho_w$ are positive design parameters. As a result, the state observer error vector $e(t)$, the estimate errors of the NN weights $\tilde{W}_f = \hat{W}_f - W_f^*$ and $\hat{W}_f$ are ensured to be UUB. Moreover, the error vector $e(t)$ converges to the small compact set $\Omega_e$, i.e., $\{e : \|e\| \le k_e\}$, where $k_e$ can be made as small as desired by appropriately choosing design parameter $\tau$.

*Proof:* Consider the Lyapunov function candidate

$$V_0 = e^T pe + \frac{1}{2} tr(\tilde{W}_f^T \rho_W^{-1} \tilde{W}_f). \tag{10}$$

Taking the derivative of $V_0$ results in

$$\dot{V}_0 = e^T(A^T P + PA)e + 2e^T P(F + B_n \rho(u) - \hat{W}_f^T S_f) + tr(\tilde{W}_f^T \rho_W^{-1} \dot{\hat{W}}_f). \tag{11}$$

Substituting (9) into (11) yields

$$\dot{V}_0 = e^T(A^T P + PA)e + 2e^T P(F + B_n \rho(u) - \hat{W}_f^T S_f) + tr(\tilde{W}_f^T \rho_W^{-1} \eta_W S_f e_1 C A^{-1} - \tilde{W}_f^T |e_1| (\tilde{W}_f + W_f^*)). \tag{12}$$

Since $tr(XY^T) = tr(Y^T X) = Y^T X$, for $\forall X, Y \in \mathbb{R}^n$, we can obtain

$$tr(\rho_W^{-1} \eta_W \tilde{W}_f^T S_f e_1 C A^{-1}) = \rho_W^{-1} \eta_W e_1 C A^{-1} \tilde{W}_f^T S_f. \tag{13}$$

As $-tr(\tilde{W}_f^T(\tilde{W}_f + W_f^*)) \le \|\tilde{W}_f\| \|W_f^*\| - \|\tilde{W}_f\|^2$, (12) becomes

$$\dot{V}_0 \le -e^T Qe + 2e^T P(F + B_n \rho(u) - \hat{W}_f^T S_f) + \rho_W^{-1} \eta_W e_1 C A^{-1} \tilde{W}_f^T S_f + |e_1| \|\tilde{W}_f\| \|W_f^*\| - |e_1| \|\tilde{W}_f\|^2. \tag{14}$$

From Assumption 2, the following inequality holds true:

$$\begin{aligned} &2e^T P(F - \hat{W}_f^T S_f + B_n \rho(u)) \\ &= 2e^T P[W_f^{*T} S_f(\bar{x}_i) + \varepsilon_f - \hat{W}_f^T S_f(\hat{\bar{x}}_i) + B_n \rho(u)] \\ &\le 2\|e\| \|P\| [2W_{fM} S_{fM} + \varepsilon_{fM} + m + \|\tilde{W}_f\| S_{fM}] \end{aligned} \tag{15}$$

where $\|S_f\| \le S_{fM}$, and $S_{fM}$ is a positive constant.

From (14) and (15), it follows that:

$$\begin{aligned} \dot{V}_0 \le &-e^T Qe + 2\|e\| \|P\| [2W_{fM} S_{fM} + \varepsilon_{fM} \\ &+ \|B_n\| m + \|\tilde{W}_f\| S_{fM}] + \rho_W^{-1} \eta_W \|e\| C A^{-1} \\ &\times \|\tilde{W}_f\| S_f + \|e\| \|\tilde{W}_f\| W_{fM} - \|e\| \|\tilde{W}_f\|^2 \\ \le &-\tau \|e\|^2 + \|e\| \{d_0 + 2\|P\| (\varepsilon_{fM} + \|B_n\| m) \\ &+ \beta_W^2 - (\|\tilde{W}_f\| - \beta_W)^2\} \\ \le &(-\tau \|e\| + d_0 + 2\|P\| (\varepsilon_{fM} + m) + \beta_W^2) \|e\| \end{aligned} \tag{16}$$

where $\tau = \lambda_{\min}(Q)$, $\lambda_{\min}(Q)$ denotes the minimum eigenvalue of matrix $Q$; $d_0 = 4\|P\| W_{fM} S_{fM}$ and $\beta_W = [\rho_W^{-1} \eta_W \|CA^{-1}\| S_{fM} + 2\|P\| S_{fM} + W_{fM}]/2$.

Let $k_e = [d_0 + 2\|P\| (\varepsilon_{fM} + m) + \beta_W^2]/\tau$. $\dot{V}_0$ is negative only if

$\|e\| \ge k_e$. According to the Lyapunov extension theorem, both the system observer error $e(t)$, the neural network weights $\hat{W}_f$ and the estimate errors of the neural network weights $\tilde{W}_f$ are UUB. ■

*B. Output Feedback Optimized Controller Design and Stability Analysis*

In this section, the optimal tracking controller is designed under the framework of backstepping technology. An auxiliary design system is introduced to reduce the effect arisen from input saturation, and the tan-type BLF is introduced to handle the problem of state constraints. The actor-critic architecture was used to construct optimal virtual controllers $\hat{\alpha}_i(t)$ $(i = 1,\ldots,n-1)$ and updated weights $\dot{\hat{W}}_{ci}$, $\dot{\hat{W}}_{ai}$. A simplified RL algorithm is developed, which is generated from the partial derivative of the HJB equation. In the $n$-th step, the optimal actual controller and the updating weights for critic and actor NNs are obtained.

*Step 1:* Define the tracking error variable as

$$s_1 = y - y_r \tag{17}$$

where $y_r$ represents the tracking signal, and $y_r(t)$, $\dot{y}_r(t)$ are bounded.

Its time derivative along (17) is

$$\dot{s}_1 = \hat{x}_2 + \dot{e}_1 + \hat{W}_{1,f}^T S_{1,f}(\hat{x}_1) + l_1(y-\hat{y}) - \dot{y}_r \tag{18}$$

where $\hat{x}_2$ denotes the ideal optimal virtual controller $\alpha_1^*(s_1)$, i.e., $\hat{x}_2 \triangleq \alpha_1^*(s_1)$. Since $\|e\| \le k_e$, the state observe error $e_1$ is UUB and converges to the compact $\Omega_e$. Then, we can determine that $\dot{e}_1$ is bounded.

The optimal value function for the $s_1$-subsystem is defined as

$$\begin{aligned} J_1^*(s_1) &= \min_{\alpha_1 \in \Psi(\Omega_{s1})} \int_t^{\infty} (M_1(x_1) + r_1(\alpha_1(\tau))^2) d\tau \\ &= \int_t^{\infty} (M_1(x_1) + r_1(\alpha_1^*(\tau))^2) d\tau \end{aligned} \tag{19}$$

where $M_1(x_1) = (k_{b1}^2/\pi)\tan(\pi s_1^2/2k_{b1}^2)$, $\alpha_1^*(s_1)$ is the optimal virtual controller, and $\Omega_{s1} = \{s_1 : |s_1| < k_{b1}\}$ is a compact set containing origin. $\Psi(\Omega_{s1})$ is the admissible control set of $\alpha_1$, and $r_1 > 0$.

By decomposing (19) into the following form:

$$\begin{aligned} J_1^*(s_1) &= \eta_1 \frac{S_1(n_1)}{2} + 2\bar{\eta}_1 \frac{k_{b1}^2}{\pi} \tan(\frac{\pi s_1^2}{2k_{b1}^2}) - \eta_1 \frac{S_1(n)}{2} \\ &\quad - 2\bar{\eta}_1 \frac{k_{b1}^2}{\pi} \tan(\frac{\pi s_1^2}{2k_{b1}^2}) + J_1^*(s_1) \\ &= \eta_1 \frac{S_1(n_1)}{2} + 2\bar{\eta}_1 \frac{k_{b1}^2}{\pi} \tan(\frac{\pi s_1^2}{2k_{b1}^2}) + J_{c1}(s_1) \end{aligned} \tag{20}$$

where $J_{c1}(s_1) = -\eta_1 S_1(n_1)/2 - 2\bar{\eta}_1 k_{b1}^2 \tan(\pi s_1^2/2k_{b1}^2)/\pi + J_1^*(s_1)$ is a real scalar-value function, $S_1(n_1) = \int_0^{n_1} (\sin n_1/n_1) dn_1$ (where $n_1 = \pi/k_{b1}^2 s_1^2$) and $\eta_1 > 0$, $\bar{\eta}_1 > 0$ are constants. For the value function $J_1^*(s_1)$ and the optimal virtual controller $\alpha_1^*(s_1)$, the HJB equation of the $s_1$-subsystem is defined as

$$\begin{aligned} &H_1(s_1, \alpha_1^*, \frac{\partial J_1^*(s_1)}{\partial s_1}) \\ &= \frac{k_{b1}^2}{\pi} \tan(\frac{\pi s_1^2}{2k_{b1}^2}) + r_1(\alpha_1^*)^2 + (\frac{2\eta_1}{s_1} \sin(\frac{\pi s_1^2}{2k_{b1}^2}) \cos(\frac{\pi s_1^2}{2k_{b1}^2}) \\ &\quad + \frac{\partial J_{c1}(s_1)}{\partial s_1} + \frac{2\bar{\eta}_1 s_1}{\cos^2(\pi s_1^2/2k_{b1}^2)})(\alpha_1^* + \hat{W}_{1,f}^T S_{1,f}(\hat{x}_1) \\ &\quad + l_1(y-\hat{y}) - \dot{y}_r + \dot{e}_1) = 0. \end{aligned} \tag{21}$$

By solving $\partial H_1/\partial \alpha_1^* = 0$, $\alpha_1^*$ can be obtained as

$$\begin{aligned} \alpha_1^* = &-\frac{\eta_1}{r_1 s_1} \sin(\frac{\pi s_1^2}{2k_{b1}^2}) \cos(\frac{\pi s_1^2}{2k_{b1}^2}) - \frac{\bar{\eta}_1 s_1}{r_1 \cos^2(\pi s_1^2/2k_{b1}^2)} \\ &- \frac{1}{2r_1} \frac{\partial J_{c1}(s_1)}{\partial s_1}. \end{aligned} \tag{22}$$

Note that $\partial J_{c1}(s_1)/\partial s_1$ is an unknown function of variable $s_1$. It can be approximated by a neural network on the compact set $\Omega_{s1}$ as

$$\frac{\partial J_{c1}(s_1)}{\partial s_1} = W_1^{*T} S_{J1}(s_1) + \varepsilon_1(s_1) \tag{23}$$

where $W_1^*$ and $S_{J1}(s_1)$ are the ideal weight vector and the basis function vector, respectively. $\varepsilon_1(s_1)$ is the approximation error and $|\varepsilon_1(s_1)| \le \bar{\delta}_1$ ($\bar{\delta}_1 > 0$ is a constant).

Using (23), the ideal optimal virtual controller $\alpha_1^*$ becomes

$$\begin{aligned} \alpha_1^* = &-\frac{\eta_1}{r_1 s_1} \sin(\frac{\pi s_1^2}{2k_{b1}^2}) \cos(\frac{\pi s_1^2}{2k_{b1}^2}) - \frac{\bar{\eta}_1 s_1}{r_1 \cos^2(\pi s_1^2/2k_{b1}^2)} \\ &- \frac{1}{2r_1} (W_1^{*T} S_{J1}(s_1) + \varepsilon_1(s_1)). \end{aligned} \tag{24}$$

Since $W_1^*$ is an unknown constant vector, the estimation vector $\hat{W}_{c1}$ is used to approximate $W_1^*$, namely,

$$\frac{\partial \hat{J}_{c1}(s_1)}{\partial s_1} = \hat{W}_{c1}^T S_{J1}(s_1). \tag{25}$$

Based on (24), we use $\hat{W}_{a1}$ to approximate $W_1^*$ in actor neural networks. The optimal virtual controller $\hat{\alpha}_1(t)$ becomes

$$\begin{aligned} \hat{\alpha}_1 = &-\frac{\eta_1}{r_1 s_1} \sin(\frac{\pi s_1^2}{2k_{b1}^2}) \cos(\frac{\pi s_1^2}{2k_{b1}^2}) - \frac{\bar{\eta}_1 s_1}{r_1 \cos^2(\pi s_1^2/2k_{b1}^2)} \\ &- \frac{1}{2r_1} \hat{W}_{a1}^T S_{J1}(s_1). \end{aligned} \tag{26}$$

*Remark 1:* In order to ensure that the term $\cos^2(\pi s_1^2/2k_{b1}^2)$ in (26) is not zero, i.e., $\pi s_1^2/2k_{b1}^2 \ne \pi\Upsilon/2$ $(\Upsilon = 1,2,\ldots)$, one can obtain $\mathrm{sgn}(s_1)s_1 \ne k_{b1}\sqrt{\Upsilon}$. Since $|s_1| < k_{b1}$, $\mathrm{sgn}(s_1)s_1 \ne k_{b1}\sqrt{\Upsilon}$ is obvious. In addition, the equivalent infinitesimal form of $\sin(\pi s_1^2/2k_{b1}^2)$ is $\pi s_1^2/2k_{b1}^2$ when the error vector $s_1 \to 0$. We can then get $\lim_{s_1 \to 0} \sin(\pi s_1^2/2k_{b1}^2)\cos(\pi s_1^2/2k_{b1}^2)/s_1 \to 0$. The singularity problem in the optimal virtual controller $\hat{\alpha}_1$ is effectively avoided.

Based on (26), the approximate HJB equation is obtained as

$$\begin{aligned} &H_1(s_1, \hat{\alpha}_1, \frac{\partial \hat{J}_1(s_1)}{\partial s_1}) \\ &= \frac{k_{b1}^2}{\pi} \tan(\frac{\pi s_1^2}{2k_{b1}^2}) + r_1(\frac{\bar{\eta}_1 s_1}{r_1 \cos^2(\pi s_1^2/2k_{b1}^2)} + \frac{\eta_1}{r_1 s_1} \end{aligned}$$

$$\begin{aligned}&\times\sin(\frac{\pi s_1^2}{2k_{b1}^2})\cos(\frac{\pi s_1^2}{2k_{b1}^2})+\frac{1}{2r_1}\hat{W}_{a1}^T S_{J1}(s_1))^2\\&+(\frac{2\eta_1}{s_1}\sin(\frac{\pi s_1^2}{2k_{b1}^2})\cos(\frac{\pi s_1^2}{2k_{b1}^2})+\frac{2\bar{\eta}_1 s_1}{\cos^2(\pi s_1^2/2k_{b1}^2)}\\&+\frac{\partial J_{c1}(s_1)}{\partial s_1})(-\frac{\eta_1}{r_1 s_1}\sin(\frac{\pi s_1^2}{2k_{b1}^2})\cos(\frac{\pi s_1^2}{2k_{b1}^2})-\dot{y}_r\\&-\frac{\bar{\eta}_1 s_1}{r_1\cos^2(\pi s_1^2/2k_{b1}^2)}-\frac{1}{2r_1}\hat{W}_{a1}^T S_{J1}(s_1)+\dot{e}_1\\&+\hat{W}_{1,f}^T S_{1,f}(\hat{x}_1)+l_1(y-\hat{y})).\end{aligned}\tag{27}$$

Define the Hamiltonian's approximation error as

$$\begin{aligned}E_1&=H_1(s_1,\hat{\alpha}_1,\frac{\partial\hat{J}_{c1}(s_1)}{\partial s_1})-H_1(s_1,\alpha_1^*,\frac{\partial J_1^*(s_1)}{\partial s_1})\\&=H_1(s_1,\hat{\alpha}_1,\frac{\partial\hat{J}_{c1}(s_1)}{\partial s_1}).\end{aligned}\tag{28}$$

The critic NN adaptive law is designed as

$$\dot{\hat{W}}_{c1}(t)=-\gamma_{c1}S_{J1}(s_1)S_{J1}^T(s_1)\hat{W}_{c1}^T(t)\tag{29}$$

where $\gamma_{c1}>0$ is the critic designed constant.

The actor NN adaptive law is given as

$$\begin{aligned}\dot{\hat{W}}_{a1}(t)=&-S_{J1}(s_1)S_{J1}^T(s_1)(\gamma_{a1}(\hat{W}_{a1}(t)-\hat{W}_{c1}(t))\\&+\gamma_{c1}\hat{W}_{c1}(t))\end{aligned}\tag{30}$$

where $\gamma_{a1}>0$ is the actor designed constant.

According to the above analysis, the optimized solution $\hat{\alpha}_1(s_1)$ is expected to satisfy $E_1(t)=H_1(s_1,\hat{\alpha}_1,\partial\hat{J}_{c1}(s_1)/\partial s_1)\to 0$. If $H_1(s_1,\hat{\alpha}_1,\partial\hat{J}_{c1}(s_1)/\partial s_1)=0$ is held and has the unique solution, then it is equivalent to the following:

$$\frac{\partial H_1(s_1,\hat{\alpha}_1,\frac{\partial\hat{J}_{c1}(s_1)}{\partial s_1})}{\partial\hat{W}_{a1}}=\frac{1}{2r_1}S_{J1}(s_1)S_{J1}^T(s_1)(\hat{W}_{a1}(t)-\hat{W}_{c1}(t))=0.\tag{31}$$

In order to derive the adaptive laws to guarantee (31), the following positive function is constructed:

$$P_1(t)=(\hat{W}_{a1}(t)-\hat{W}_{c1}(t))^T(\hat{W}_{a1}(t)-\hat{W}_{c1}(t)).\tag{32}$$

Clearly, $P_1(t)=0$ is equivalent to (31). Since $r_1\frac{\partial P_1(t)}{\partial\hat{W}_{a1}(t)}=-\frac{\partial P_1(t)}{\partial\hat{W}_{c1}(t)}=2(\hat{W}_{a1}(t)/r_1-\hat{W}_{c1}(t))$, we can get

$$\begin{aligned}\frac{dP_1(t)}{dt}&=\frac{\partial P_1(t)}{\partial\hat{W}_{c1}(t)}\dot{\hat{W}}_{c1}(t)+\frac{\partial P_1(t)}{\partial\hat{W}_{a1}(t)}\dot{\hat{W}}_{a1}(t)\\&=-\gamma_{c1}\frac{\partial P_1(t)}{\partial\hat{W}_{c1}(t)}S_{J1}(s_1)S_{J1}^T(s_1)\hat{W}_{c1}(t)-\frac{\partial P_1(t)}{\partial\hat{W}_{a1}}\\&\quad\times S_{J1}S_{J1}^T(\gamma_{a1}(\hat{W}_{a1}(t)-\hat{W}_{c1}(t))+\gamma_{c1}\hat{W}_{c1}(t))\\&=-\frac{\gamma_{a1}}{2}\frac{\partial P_1(t)}{\partial\hat{W}_{a1}(t)}S_{J1}(s_1)S_{J1}^T(s_1)\frac{\partial P_1(t)}{\partial\hat{W}_{a1}(t)}\le 0.\end{aligned}\tag{33}$$

Consider the tan-type barrier Lyapunov function candidate for $s_1$-subsystem

$$V_1(t)=\frac{k_{b1}^2}{\pi}\tan(\frac{\pi s_1^2}{2k_{b1}^2})+\frac{1}{2}\tilde{W}_{c1}^T(t)\tilde{W}_{c1}(t)+\frac{1}{2}\tilde{W}_{a1}^T(t)\tilde{W}_{a1}(t)\tag{34}$$

where $\tilde{W}_{c1}=\hat{W}_{c1}-W_1^*$, $\tilde{W}_{a1}=\hat{W}_{a1}-W_1^*$ are critic and actor NNs approximation errors, respectively.

From $s_2=\hat{x}_2-\hat{\alpha}_1$ and $s_1=x_1-y_r$, we have

$$\dot{s}_1=s_2+\hat{\alpha}_1+\dot{e}_1+\hat{W}_{1,f}^T S_{1,f}(\hat{x}_1)+l_1e_1-\dot{y}_r.\tag{35}$$

The time derivative of $V_1$ is

$$\begin{aligned}\dot{V}_1(t)&=\frac{s_1}{\cos^2(\pi s_1^2/2k_{b1}^2)}\dot{s}_1+\tilde{W}_{c1}^T(t)\dot{\hat{W}}_{c1}(t)+\tilde{W}_{a1}^T(t)\dot{\hat{W}}_{a1}\\&=\frac{s_1}{\cos^2(\pi s_1^2/2k_{b1}^2)}(s_2-\frac{\eta_1}{r_1s_1}\sin(\frac{\pi s_1^2}{2k_{b1}^2})\cos(\frac{\pi s_1^2}{2k_{b1}^2})-\dot{y}_r\\&\quad-\frac{1}{2r_1}\hat{W}_{a1}^T S_{J1}+l_1e_1+\dot{e}_1-\frac{\bar{\eta}_1 s_1}{r_1}\frac{1}{\cos^2(\pi s_1^2/2k_{b1}^2)}\\&\quad+\hat{W}_{1,f}^T S_{1,f}(\hat{x}_1))+\tilde{W}_{c1}^T\dot{\hat{W}}_{c1}+\tilde{W}_{a1}^T\dot{\hat{W}}_{a1}.\end{aligned}\tag{36}$$

Letting $\vartheta_{s1}=s_1/\cos^2(\pi s_1^2/2k_{b1}^2)$ and substituting (26), (29), (30), and (35) into (36) reaches

$$\begin{aligned}\dot{V}_1(t)&=\vartheta_{s1}s_2-\frac{\eta_1}{r_1}\tan(\frac{\pi s_1^2}{2k_{b1}^2})-\frac{\bar{\eta}_1}{r_1}\vartheta_{s1}^2-\frac{\vartheta_{s1}}{2r_1}\hat{W}_{a1}^T S_{J1}(s_1)\\&\quad+\vartheta_{s1}(\hat{W}_{1,f}^T S_{1,f}(\hat{x}_1)+l_1e_1-\dot{y}_r+\dot{e}_1)-\gamma_{c1}\tilde{W}_{c1}^T(t)\\&\quad\times S_{J1}(s_1)S_{J1}^T(s_1)\hat{W}_{c1}^T(t)-\tilde{W}_{a1}^T(t)(S_{J1}(s_1)S_{J1}^T(s_1)\\&\quad\times(\gamma_{a1}(\hat{W}_{a1}(t)-\hat{W}_{c1}(t))+\gamma_{c1}\hat{W}_{c1}))\\&=\vartheta_{s1}s_2-\frac{\eta_1}{r_1}\tan(\frac{\pi s_1^2}{2k_{b1}^2})-\frac{\bar{\eta}_1}{r_1}\vartheta_{s1}^2-\frac{\vartheta_{s1}}{2r_1}\hat{W}_{a1}^T S_{J1}(s_1)\\&\quad+\vartheta_{s1}(\hat{W}_{1,f}^T S_{1,f}(\hat{x}_1)+l_1e_1-\dot{y}_r+\dot{e}_1)-\gamma_{c1}\tilde{W}_{c1}^T(t)\\&\quad\times S_{J1}(s_1)S_{J1}^T(s_1)\hat{W}_{c1}^T(t)-\gamma_{a1}\tilde{W}_{a1}^T(t)S_{J1}S_{J1}^T\\&\quad\times\hat{W}_{a1}(t)+(\gamma_{a1}-\gamma_{c1})\tilde{W}_{a1}^T(t)S_{J1}(s_1)S_{J1}^T(s_1)\hat{W}_{c1}.\end{aligned}\tag{37}$$

Similarly, by using $\tilde{W}_{a1}(t)=\hat{W}_{a1}(t)-W_1^*$ and $\tilde{W}_{c1}(t)=\hat{W}_{c1}(t)-W_1^*$, there are the following equations:

$$\begin{aligned}&\tilde{W}_{c1}^T(t)S_{J1}(s_1)S_{J1}^T(s_1)\hat{W}_{c1}(t)\\&\quad=\frac{1}{2}\tilde{W}_{c1}^T(t)S_{J1}(s_1)S_{J1}^T(s_1)\tilde{W}_{c1}(t)-\frac{1}{2}(W_{J1}^{*T}S_1(s_1))^2\\&\quad+\frac{1}{2}\hat{W}_{c1}^T(t)S_{J1}(s_1)S_{J1}^T(s_1)\hat{W}_{c1}(t)\end{aligned}\tag{38}$$

$$\begin{aligned}&\tilde{W}_{a1}^T(t)S_{J1}(s_1)S_{J1}^T(s_1)\hat{W}_{a1}(t)\\&\quad=\frac{1}{2}\tilde{W}_{a1}^T(t)S_{J1}(s_1)S_{J1}^T(s_1)\tilde{W}_{a1}(t)-\frac{1}{2}(W_{J1}^{*T}S_{J1}(s_1))^2\\&\quad+\frac{1}{2}\hat{W}_{a1}^T(t)S_{J1}(s_1)S_{J1}^T(s_1)\hat{W}_{a1}(t).\end{aligned}\tag{39}$$

Substituting (38) and (39) into (37), one has

$$\begin{aligned}\dot{V}_1(t)=&\vartheta_{s1}s_2-\frac{\eta_1}{r_1}\tan(\frac{\pi s_1^2}{2k_{b1}^2})-\frac{\bar{\eta}_1}{r_1}\vartheta_{s1}^2-\frac{\vartheta_{s1}}{2r_1}\hat{W}_{a1}^TS_{J1}(s_1)\\&+\vartheta_{s1}(-\dot{y}_r+l_1e_1+\dot{e}_1+\hat{W}_{1,f}^TS_{1,f}(\hat{x}_1))\\&-\frac{\gamma_{a1}}{2}\tilde{W}_{a1}^T(t)S_{J1}(s_1)S_{J1}^T(s_1)\tilde{W}_{a1}(t)-\frac{\gamma_{c1}}{2}\tilde{W}_{c1}^T(t)\\&\times S_{J1}(s_1)S_{J1}^T(s_1)\tilde{W}_{c1}(t)-\frac{\gamma_{c1}}{2}\hat{W}_{c1}^T(t)S_{J1}(s_1)\\&\times S_{J1}^T(s_1)\hat{W}_{c1}(t)-\frac{\gamma_{a1}}{2}\hat{W}_{a1}^T(t)S_{J1}(s_1)S_{J1}^T(s_1)\hat{W}_{a1}(t)\\&+(\frac{\gamma_{a1}}{2}+\frac{\gamma_{c1}}{2})(W_{J1}^{*T}S_{J1}(s_1))^2+(\gamma_{a1}-\gamma_{c1})\tilde{W}_{a1}^T(t)\\&\times S_{J1}(s_1)S_{J1}^T(s_1)\hat{W}_{c1}(t).\end{aligned}\tag{40}$$

Using Young's inequality, there is the following fact that:

$$\begin{aligned}(\gamma_{a1}-\gamma_{c1})&\tilde{W}_{a1}^T(t)S_{J1}(s_1)S_{J1}^T(s_1)\hat{W}_{c1}(t)\\&\le\frac{\gamma_{a1}-\gamma_{c1}}{2}\tilde{W}_{a1}^T(t)S_{J1}(s_1)S_{J1}^T(s_1)\tilde{W}_{a1}(t)\\&\quad+\frac{\gamma_{a1}-\gamma_{c1}}{2}\hat{W}_{c1}^T(t)S_{J1}(s_1)S_{J1}^T(s_1)\hat{W}_{c1}(t).\end{aligned}\tag{41}$$

Substituting (41) into (40), one has

$$\begin{aligned}\dot{V}_1(t)\le&\vartheta_{s1}s_2-\frac{\eta_1}{r_1}\tan(\frac{\pi s_1^2}{2k_{b1}^2})-\frac{\bar{\eta}_1}{r_1}\vartheta_{s1}^2-\frac{\vartheta_{s1}}{2r_1}\hat{W}_{a1}^TS_{J1}\\&+\vartheta_{s1}(l_1e_1+\dot{e}_1-\dot{y}_r+\hat{W}_{1,f}^TS_{1,f})-\frac{\gamma_{a1}}{2}\tilde{W}_{a1}^T\\&\times S_{J1}(s_1)S_{J1}^T(s_1)\tilde{W}_{a1}(t)-\frac{\gamma_{c1}}{2}\tilde{W}_{c1}^T(t)S_{J1}(s_1)\\&\times S_{J1}^T(s_1)\tilde{W}_{c1}(t)-\frac{\gamma_{c1}}{2}\hat{W}_{c1}^T(t)S_{J1}(s_1)S_{J1}^T(s_1)\\&\times\hat{W}_{c1}(t)-\frac{\gamma_{a1}}{2}\hat{W}_{a1}^T(t)S_{J1}(s_1)S_{J1}^T(s_1)\hat{W}_{a1}(t)\\&+\frac{\gamma_{a1}+\gamma_{c1}}{2}(W_{J1}^{*T}S_{J1}(s_1))^2+\frac{\gamma_{a1}-\gamma_{c1}}{2}\tilde{W}_{a1}^T\\&\times S_{J1}S_{J1}^T\tilde{W}_{a1}+\frac{\gamma_{a1}-\gamma_{c1}}{2}\hat{W}_{c1}^TS_{J1}S_{J1}^T\hat{W}_{c1}\\\le&\vartheta_{s1}s_2-\frac{\eta_1}{r_1}\tan(\frac{\pi s_1^2}{2k_{b1}^2})-\frac{\bar{\eta}_1}{r_1}\vartheta_{s1}^2-\frac{\vartheta_{s1}}{2r_1}\hat{W}_{a1}^T\\&\times S_{J1}(s_1)+\vartheta_{s1}(l_1e_1+\dot{e}_1-\dot{y}_r+\hat{W}_{1,f}^TS_{1,f}\\&-\frac{\gamma_{c1}}{2}\tilde{W}_{c1}^T(t)S_{J1}(s_1)S_{J1}^T(s_1)\tilde{W}_{c1}(t)-\frac{\gamma_{c1}}{2}\\&\times\tilde{W}_{a1}^T(t)S_{J1}S_{J1}^T\tilde{W}_{a1}(t)-(\gamma_{c1}-\frac{\gamma_{a1}}{2})(\hat{W}_{c1}^T(t)\\&\times S_{J1}(s_1))^2-\frac{\gamma_{a1}}{2}(\hat{W}_{a1}^T(t)S_{J1}(s_1))^2+(\frac{\gamma_{a1}}{2}\\&+\frac{\gamma_{c1}}{2})(W_{J1}^{*T}(t)S_{J1}(s_1))^2.\end{aligned}\tag{42}$$

According to the Young's inequality, one has

$$\vartheta_{s1}s_2\le\frac{\vartheta_{s1}^2}{2}+\frac{1}{2}k_{b2}^2\tag{43}$$

$$\begin{aligned}\vartheta_{s1}&(\hat{W}_{1,f}^TS_{1,f}(\hat{x}_1)+l_1e_1-\dot{y}_r+\dot{e}_1)\\&\le2\vartheta_{s1}^2+\frac{1}{2}l_1^2k_e^2+\frac{1}{2}\dot{e}_1^2+\frac{1}{2}\dot{y}_r^2+\frac{1}{2}\hat{W}_{1,f}^TS_{1,f}S_{1,f}^T\hat{W}_{1,f}^T\end{aligned}\tag{44}$$

$$-\frac{\vartheta_{s1}}{2r_1}\hat{W}_{a1}^T(t)S_{J1}\le\frac{\vartheta_{s1}^2}{4r_1}+\frac{1}{4r_1}\hat{W}_{a1}^TS_{J1}(s_1)S_{J1}^T(s_1)\hat{W}_{a1}^T\tag{45}$$

where $|s_2|<k_{b2}$, Substituting (43)–(45) into (42)

$$\begin{aligned}\dot{V}_1(t)\le&-\frac{\eta_1}{r_1}\tan(\frac{\pi s_1^2}{2k_{b1}^2})-(\frac{\bar{\eta}_1}{r_1}-\frac{1}{4r_1}-\frac{5}{2})\vartheta_{s1}^2+\frac{1}{2}l_1^2k_e^2\\&+\frac{1}{2}\dot{e}_1^2+\frac{1}{2}\dot{y}_r^2+\frac{1}{2}k_{b2}^2+\frac{1}{2}\hat{W}_{1,f}^TS_{1,f}(\hat{x}_1)\\&\times S_{1,f}^T(\hat{x}_1)\hat{W}_{1,f}^T-\frac{\gamma_{c1}}{2}\tilde{W}_{c1}^T(t)S_{J1}(s_1)S_{J1}^T(s_1)\\&\times\tilde{W}_{c1}(t)-\frac{\gamma_{c1}}{2}\tilde{W}_{a1}^T(t)S_{J1}S_{J1}^T\tilde{W}_{a1}(t)-(\gamma_{c1}\\&-\frac{\gamma_{a1}}{2})(\hat{W}_{c1}^T(t)S_{J1}(s_1))^2-(\frac{\gamma_{a1}}{2}-\frac{1}{4r_1})(\hat{W}_{a1}^T(t)\\&\times S_{J1}(s_1))^2+(\frac{\gamma_{a1}}{2}+\frac{\gamma_{c1}}{2})(W_{J1}^{*T}(t)S_{J1}(s_1))^2.\end{aligned}\tag{46}$$

Let $\lambda_{S_{1,f}}^{\max}$ be the maximal eigenvalue of $S_{1,f}(\hat{x}_1)S_{1,f}(\hat{x}_1)$ and $\lambda_{S_{J1}}^{\min}$ be the minimal eigenvalue of $S_{J1}(s_1)S_{J1}^T(s_1)$. Inequality (46) can become

$$\begin{aligned}\dot{V}_1(t)\le&-\frac{\eta_1}{r_1}\tan(\frac{\pi s_1^2}{2k_{b1}^2})-(\frac{\bar{\eta}_1}{r_1}-\frac{1}{4r_1}-\frac{5}{2})\vartheta_{s1}^2\\&-\frac{\gamma_{c1}}{2}\lambda_{S_{J1}}^{\min}\tilde{W}_{a1}^T\tilde{W}_{a1}-\frac{\gamma_{c1}}{2}\lambda_{S_{J1}}^{\min}\tilde{W}_{c1}^T\tilde{W}_{c1}\\&-(\gamma_{c1}-\frac{\gamma_{a1}}{2})(\hat{W}_{c1}^T(t)S_{J1}(s_1))^2-(\frac{\gamma_{a1}}{2}\\&-\frac{1}{4r_1})(\hat{W}_{a1}^T(t)S_{J1}(s_1))^2+D_1\end{aligned}\tag{47}$$

where $D_1=\sup_{t\ge0}\{D_1(t)\}$ and $D_1(t)=\frac{1}{2}\lambda_{S_{1,f}}^{\max}\hat{W}_{1,f}^T(\hat{x}_1)\hat{W}_{1,f}(\hat{x}_1)+\frac{1}{2}\dot{e}_1^2+\frac{1}{2}l_1^2k_e^2+\frac{1}{2}k_{b2}^2+(\frac{\gamma_{a1}}{2}+\frac{\gamma_{c1}}{2})(W_{J1}^{*T}(t)S_{J1}(s_1))^2+\frac{1}{2}\dot{y}_r^2$.

We then design the parameters $\gamma_{c1}$, $\gamma_{a1}$, $r_1$, and $\bar{\eta}_1$, which satisfy the following inequalities:

$$\gamma_{c1}-\frac{\gamma_{a1}}{2}>0\tag{48}$$

$$\frac{\gamma_{a1}}{2}-\frac{1}{4r_1}>0\tag{49}$$

$$\frac{\bar{\eta}_1}{r_1}-\frac{1}{4r_1}-\frac{5}{2}>0.\tag{50}$$

Denote $\eta_{10}=\eta_1\pi/k_{b1}^2$, (50) can then be rewritten as

$$\begin{aligned}\dot{V}_1(t)\le&-\frac{k_{b1}^2\eta_{10}}{\pi r_1}\tan(\frac{\pi s_1^2}{2k_{b1}^2})-\frac{\gamma_{c1}}{2}\lambda_{S_{J1}}^{\min}\tilde{W}_{a1}^T(t)\tilde{W}_{a1}(t)\\&-\frac{\gamma_{c1}}{2}\lambda_{S_{J1}}^{\min}\tilde{W}_{c1}^T(t)\tilde{W}_{c1}(t)+D_1.\end{aligned}\tag{51}$$

Let $c_1=\min\{\eta_{10}/r_1,\gamma_{c1}\lambda_{S_{J1}}^{\min},\gamma_{c1}\lambda_{S_{J1}}^{\min}\}$. Then, (51) becomes

$$\dot{V}_1\le-c_1V_1+D_1.\tag{52}$$

From (52), we can have

$$V_1(t)\le V_1(t_0)e^{-c_1(t-t_0)}+\frac{D_1}{c_1}.\tag{53}$$

Since $|s_1|<k_{b1}$ and $s_1=y-y_r$, we have $|x_1(t)|\le|s_1(t)|+$

$|y_r(t)| < k_{b1} + |y_r(t)|$. Define $k_{b1} = k_{c1} - |y_r|$, where $|x_1| \le k_{c1}$. From (53), as $t \to \infty$, $e^{-(t-t_0)} \to 0$. It follows that there exists $T_1$, when $t \ge T_1$, $\|\tilde{W}_{a1}(t)\| \le \sqrt{2D_1/c_1}$, and $\|\tilde{W}_{c1}(t)\| \le \sqrt{2D_1/c_1}$. Clearly, the reduction of $\sqrt{2D_1/c_1}$ can be made arbitrarily small by increasing $c_1$, while decreasing $D_1$. Therefore, we can determine that $\|\hat{W}_{a1}(t)\|$ and $\|\hat{W}_{c1}(t)\|$ are bounded. In addition, we know the boundedness of $e_1$, $s_1$, $\|\hat{W}_{a1}(t)\|$, $\|\hat{W}_{c1}(t)\|$, and $\|S_1\|$. Thus, $\hat{\alpha}_1$, $\|\dot{\hat{W}}_{a1}(t)\|$, and $\dot{s}_1$ are bounded (where $|\hat{\alpha}_1| \le A_2$, $A_2$ is a positive constant), and $\dot{\hat{\alpha}}_1$ is bounded.

*Step i ($2 \le i \le n-1$):* Similarly, define $s_i = \hat{x}_i - \hat{\alpha}_{i-1}$ where the time derivative is

$$\dot{s}_i = \hat{x}_{i+1} - \dot{\hat{\alpha}}_{i-1} + \hat{W}_{i,f}^T S_{i,f}(\hat{\bar{x}}_i) + l_i(y - \hat{y}). \tag{54}$$

The optimal value function for the $s_i$-subsystem is defined as

$$\begin{aligned} J_i^*(s_i) &= \min_{\alpha_i \in \Psi(\Omega_{si})} \int_t^\infty (M_i(x_i) + r_i(\alpha_i(\tau))^2) d\tau \\ &= \int_t^\infty (M_i(x_i) + r_i(\alpha_i^*(\tau))^2) d\tau \end{aligned} \tag{55}$$

where $M_i(x_i) = k_{bi}^2 \tan(\pi s_i^2/2k_{bi}^2)/\pi$, $\alpha_i^*(s_i)$ is the optimal virtual controller, $\Omega_{si} = \{s_i : |s_i| < k_{bi}\}$ is a compact set containing origin. $\Psi(\Omega_{si})$ is the admissible control set of $\alpha_i$, and $r_i > 0$ is a constant.

The optimal value function for $s_i$-subsystem satisfies the following equation:

$$\begin{aligned} J_i^*(s_i) &= \eta_i \frac{S_i(n_i)}{2} + 2\bar{\eta}_i \frac{k_{bi}^2}{\pi} \tan\left(\frac{\pi s_i^2}{2k_{bi}^2}\right) - \eta_i \frac{S_i(n_i)}{2} \\ &\quad - 2\bar{\eta}_i \frac{k_{bi}^2}{\pi} \tan\left(\frac{\pi s_i^2}{2k_{bi}^2}\right) + J_i^*(s_i) \\ &= \eta_i \frac{S_i(n_i)}{2} + 2\bar{\eta}_i \frac{k_{bi}^2}{\pi} \tan\left(\frac{\pi s_i^2}{2k_{bi}^2}\right) + J_{ci}(s_i) \end{aligned} \tag{56}$$

where $J_{ci}(s_i) = -\eta_i S_i(n_i)/2 - 2\bar{\eta}_i k_{bi}^2 \tan(\pi s_i^2/2k_{bi}^2)/\pi + J_i^*(s_i)$. Similarly, $S_i(n_i) = \int_0^{n_i} (\sin n_i/n_i) d_{n_i}$, $n_i = \pi/k_{bi}^2 s_i^2$, and $\eta_i$, $\bar{\eta}_i > 0$ are constants. For $\alpha_i^*(s_i)$ and $J_i^*(s_i)$, the HJB equation of the $s_i$-subsystem is defined as

$$\begin{aligned} &H_i\left(s_i, \alpha_i^*, \frac{\partial J_i^*(s_i)}{\partial s_i}\right) \\ &= \frac{k_{bi}^2}{\pi} \tan\left(\frac{\pi s_i^2}{2k_{bi}^2}\right) + r_i(\alpha_i^*)^2 + \left(\frac{2\eta_i}{s_i} \sin\left(\frac{\pi s_i^2}{2k_{bi}^2}\right) \cos\left(\frac{\pi s_i^2}{2k_{bi}^2}\right)\right. \\ &\quad + \frac{\partial J_{ci}(s_i)}{\partial s_i} + \left.\frac{2\bar{\eta}_i s_i}{\cos^2(\pi s_i^2/2k_{bi}^2)}\right)(\alpha_i^* + \hat{W}_{i,f}^T S_{i,f}(\hat{x}_i) \\ &\quad + l_i(y - \hat{y}) - \dot{\hat{\alpha}}_{i-1}) = 0. \end{aligned} \tag{57}$$

Similarly, solving $\partial H_i / \partial \alpha_i^* = 0$, yields

$$\alpha_i^* = -\frac{\eta_i}{r_i s_i} \sin\left(\frac{\pi s_i^2}{2k_{bi}^2}\right) \cos\left(\frac{\pi s_i^2}{2k_{bi}^2}\right) - \frac{\bar{\eta}_i s_i}{r_i \cos^2(\pi s_i^2/2k_{bi}^2)} - \frac{1}{2r_i} \frac{\partial J_{ci}(s_i)}{\partial s_i} \tag{58}$$

where $\partial J_{ci}(s_i)/\partial s_i$ can be approximated by the following NN on the compact set $\Omega_{si}$:

$$\frac{\partial J_{ci}(s_i)}{\partial s_i} = W_i^{*T} S_{Ji}(s_i) + \varepsilon_i(s_i) \tag{59}$$

where $W_i^*$ is an ideal weight vector and $S_{Ji}(s_i)$ is the basis function vector. $\varepsilon_i(s_i)$ is the approximation error satisfying $|\varepsilon_i(s_i)| \le \bar{\delta}_i$ where $\bar{\delta}_i > 0$ is a real constant. By (58) and (59), the ideal optimal virtual controller $\alpha_i^*$ can be acquired as

$$\begin{aligned} \alpha_i^* = &-\frac{\eta_i}{r_i s_i} \sin\left(\frac{\pi s_i^2}{2k_{bi}^2}\right) \cos\left(\frac{\pi s_i^2}{2k_{bi}^2}\right) - \frac{\bar{\eta}_i s_i}{r_i \cos^2(\pi s_i^2/2k_{bi}^2)} \\ &- \frac{1}{2r_i}(W_i^{*T} S_{Ji}(s_i) + \varepsilon_i(s_i)). \end{aligned} \tag{60}$$

Similarly, we can get

$$\frac{\partial \hat{J}_{ci}(s_i)}{\partial s_i} = \hat{W}_{ci}^T S_{Ji}(s_i) \tag{61}$$

$$\begin{aligned} \hat{\alpha}_i = &-\frac{\eta_i}{r_i s_i} \sin\left(\frac{\pi s_i^2}{2k_{bi}^2}\right) \cos\left(\frac{\pi s_i^2}{2k_{bi}^2}\right) - \frac{\bar{\eta}_i s_i}{r_i \cos^2(\pi s_i^2/2k_{bi}^2)} \\ &- \frac{1}{2r_i} \hat{W}_{ai}^T S_{Ji}(s_i) \end{aligned} \tag{62}$$

where $\hat{W}_{ci}$ and $\hat{W}_{ai}$ are the critic and actor NN weights, respectively. Similarly, $\mathrm{sgn}(s_i)s_i \neq k_{bi}\sqrt{\Upsilon}$ ($\Upsilon = 1, 2, \ldots$). Thus, we can get $\lim_{s_i \to 0} \sin(\pi s_i^2/2k_{bi}^2)\cos(\pi s_i^2/2k_{bi}^2)/s_i \to 0$ and the singularity problem in the optimal virtual controller $\hat{\alpha}_i$ is effectively avoided.

From (62), the approximate HJB equation is obtained as

$$\begin{aligned} &H_i\left(s_i, \hat{\alpha}_i, \frac{\partial \hat{J}_i(s_i)}{\partial s_i}\right) \\ &= \frac{k_{bi}^2}{\pi} \tan\left(\frac{\pi s_i^2}{2k_{bi}^2}\right) + r_i\left(\frac{\eta_i}{r_i s_i} \sin\left(\frac{\pi s_i^2}{2k_{bi}^2}\right) \cos\left(\frac{\pi s_i^2}{2k_{bi}^2}\right) + \frac{1}{2r_i} \hat{W}_{ai}^T\right. \\ &\quad \times S_{Ji}(s_i) + \left.\frac{\bar{\eta}_i s_i}{r_i \cos^2(\pi s_i^2/2k_{bi}^2)}\right)^2 + \left(\frac{2\eta_i}{s_i} \sin\left(\frac{\pi s_i^2}{2k_{bi}^2}\right) \cos\left(\frac{\pi s_i^2}{2k_{bi}^2}\right)\right. \\ &\quad + \left.\frac{2\bar{\eta}_i s_i}{\cos^2(\pi s_i^2/2k_{bi}^2)} + \frac{\partial J_{ci}(s_i)}{\partial s_i}\right)\left(-\frac{\eta_i}{r_i s_i} \sin\left(\frac{\pi s_i^2}{2k_{bi}^2}\right) \cos\left(\frac{\pi s_i^2}{2k_{bi}^2}\right)\right. \\ &\quad - \frac{\bar{\eta}_i s_i}{r_i \cos^2(\pi s_i^2/2k_{bi}^2)} - \frac{1}{2r_i} \hat{W}_{ai}^T S_i(s_i) + \hat{W}_{i,f}^T S_{i,f}(\hat{\bar{x}}_i) \\ &\quad + \left.l_i(y - \hat{y}) - \dot{\hat{\alpha}}_{i-1}\right). \end{aligned} \tag{63}$$

Define the Bellman error $E_i$ as

$$\begin{aligned} E_i &= H_i\left(s_i, \hat{\alpha}_i, \frac{\partial \hat{J}_{ci}(s_i)}{\partial s_i}\right) - H_i\left(s_i, \alpha_i^*, \frac{\partial J_i^*(s_i)}{\partial s_i}\right) \\ &= H_i\left(s_i, \hat{\alpha}_i, \frac{\partial \hat{J}_{ci}(s_i)}{\partial s_i}\right). \end{aligned} \tag{64}$$

The actor and critic NN adaptive laws are given as

$$\dot{\hat{W}}_{ci}(t) = -\gamma_{ci} S_{Ji}(s_i) S_{Ji}^T(s_i) \hat{W}_{ci}^T(t) \tag{65}$$

$$\begin{aligned} \dot{\hat{W}}_{ai}(t) = &-S_{Ji}(s_i) S_{Ji}^T(s_i)(\gamma_{ai}(\hat{W}_{ai}(t) - \hat{W}_{ci}(t)) \\ &+ \gamma_{ci} \hat{W}_{ci}(t)) \end{aligned} \tag{66}$$

where $\gamma_{ci} > 0$ and $\gamma_{ai} > 0$ are critic and actor designed constants, respectively.

Consider the barrier Lyapunov function candidate for the $s_i$ subsystem

$$V_i(t) = \frac{k_{bi}^2}{\pi}\tan(\frac{\pi s_i^2}{2k_{bi}^2}) + \frac{1}{2}\tilde{W}_{ci}^T(t)\tilde{W}_{ci}(t) + \frac{1}{2}\tilde{W}_{ai}^T(t)\tilde{W}_{ai}(t) \quad (67)$$

where $\tilde{W}_{ci} = \hat{W}_{ci} - W_i^*$ and $\tilde{W}_{ai} = \hat{W}_{ai} - W_i^*$ are critic and actor NNS approximation errors, respectively.

From the definitions of $s_{i+1} = \hat{x}_{i+1} - \hat{\alpha}_i$, we have

$$\dot{s}_i = s_{i+1} + \hat{\alpha}_i + \hat{W}_{i,f}^T S_{i,f}(\hat{\bar{x}}_i) + l_i e_1 - \dot{\hat{\alpha}}_{i-1}. \quad (68)$$

The time derivative of $V_i$ is

$$\dot{V}_i(t) = \frac{s_i}{\cos^2(\pi s_i^2/2k_{bi}^2)}\dot{s}_i + \tilde{W}_{ci}^T(t)\dot{\tilde{W}}_{ci}(t) + \tilde{W}_{ai}^T(t)\dot{\tilde{W}}_{ai}(t). \quad (69)$$

Let $\vartheta_{si} = s_i/\cos^2(\pi s_i^2/2k_{bi}^2)$. As a result, (69) becomes

$$\begin{aligned}\dot{V}_i(t) = {} & \vartheta_{si}s_{i+1} - \frac{\eta_i}{r_i}\tan(\frac{\pi s_i^2}{2k_{bi}^2}) - \frac{\bar{\eta}_i}{r_i}\vartheta_{si}^2 - \frac{\vartheta_{si}}{2r_i}\hat{W}_{ai}^T S_{Ji}(s_i)\\ & + \vartheta_{si}(\hat{W}_{i,f}^T S_{i,f}(\hat{x}_i) + l_i e_1 - \dot{\hat{\alpha}}_{i-1}) - \gamma_{ci}\tilde{W}_{ci}^T(t)\\ & \times S_{Ji}(s_i)S_{Ji}^T(s_i)\hat{W}_{ci}^T(t) - \gamma_{ai}\tilde{W}_{ai}^T(t)S_{Ji}(s_i)\\ & \times S_{Ji}^T(s_i)\hat{W}_{ai}(t) + (\gamma_{ai} - \gamma_{ci})\tilde{W}_{ai}^T(t)S_{Ji}(s_i)\\ & \times S_{Ji}^T(s_i)\hat{W}_{ci}(t). \end{aligned} \quad (70)$$

Similarly, by using $\tilde{W}_{ci}(t) = \hat{W}_{ci}(t) - W_i^*$ and $\tilde{W}_{ai}(t) = \hat{W}_{ai}(t) - W_i^*$, there are the following equations:

$$\begin{aligned}&\tilde{W}_{ci}^T(t)S_{Ji}(s_i)S_{Ji}^T(s_i)\hat{W}_{ci}(t)\\ &\quad = \frac{1}{2}\tilde{W}_{ci}^T(t)S_{Ji}(s_i)S_{Ji}^T(s_i)\tilde{W}_{ci}(t) - \frac{1}{2}(W_{Ji}^{*T}S_i(s_i))^2\\ &\qquad + \frac{1}{2}\hat{W}_{ci}^T(t)S_{Ji}(s_i)S_{Ji}^T(s_i)\hat{W}_{ci}(t)\end{aligned} \quad (71)$$

$$\begin{aligned}&\tilde{W}_{ai}^T(t)S_{Ji}(s_i)S_{Ji}^T(s_i)\hat{W}_{ai}(t)\\ &\quad = \frac{1}{2}\tilde{W}_{ai}^T(t)S_{Ji}(s_i)S_{Ji}^T(s_i)\tilde{W}_{ai}(t) - \frac{1}{2}(W_{Ji}^{*T}S_{Ji}(s_i))^2\\ &\qquad + \frac{1}{2}\hat{W}_{ai}^T(t)S_{Ji}(s_i)S_{Ji}^T(s_i)\hat{W}_{ai}(t).\end{aligned} \quad (72)$$

Substituting (71) and (72) into (70), one has

$$\begin{aligned}\dot{V}_i(t) = {} & \vartheta_{si}s_{i+1} - \frac{\eta_i}{r_i}\tan(\frac{\pi s_i^2}{2k_{bi}^2}) - \frac{\bar{\eta}_i}{r_i}\vartheta_{si}^2 - \frac{\vartheta_{si}}{2r_i}\hat{W}_{ai}^T S_{Ji}\\ & + \vartheta_{si}(\hat{W}_{i,f}^T S_{i,f}(\hat{x}_i) + l_i e_1 - \dot{\hat{\alpha}}_{i-1}) - \frac{\gamma_{ci}}{2}\tilde{W}_{ci}^T S_{Ji}\\ & \times S_{Ji}^T\tilde{W}_{ci}(t) + (\gamma_{ai} - \gamma_{ci})\tilde{W}_{ai}^T S_{Ji}S_{Ji}^T\hat{W}_{ci} - \frac{\gamma_{ci}}{2}\\ & \times \hat{W}_{ci}^T(t)S_{Ji}(s_i)S_{Ji}^T(s_i)\hat{W}_{ci}(t) - \frac{\gamma_{ai}}{2}\tilde{W}_{ai}^T(t)S_{Ji}(s_i)\\ & \times S_{Ji}^T(s_i)\tilde{W}_{ai}(t) - \frac{\gamma_{ai}}{2}\hat{W}_{ai}^T(t)S_{Ji}(s_i)S_{Ji}^T(s_i)\hat{W}_{ai}(t)\\ & + (\frac{\gamma_{ai}}{2} + \frac{\gamma_{ci}}{2})(W_{Ji}^{*T}S_{Ji}(s_i))^2.\end{aligned} \quad (73)$$

Similarly, we can get

$$\begin{aligned}&(\gamma_{ai}-\gamma_{ci})\tilde{W}_{ai}^T(t)S_{Ji}(s_i)S_{Ji}^T(s_i)\hat{W}_{ci}(t)\\ &\quad \le \frac{\gamma_{ai}-\gamma_{ci}}{2}\tilde{W}_{ai}^T(t)S_{Ji}(s_i)S_{Ji}^T(s_i)\tilde{W}_{ai}(t)\\ &\qquad + \frac{\gamma_{ai}-\gamma_{ci}}{2}\hat{W}_{ci}^T(t)S_{Ji}(s_i)S_{Ji}^T(s_i)\hat{W}_{ci}(t)\end{aligned} \quad (74)$$

$$\vartheta_{si}s_{i+1} \le \frac{\vartheta_{si}^2}{2} + \frac{1}{2}k_{bi+1}^2 \quad (75)$$

$$\begin{aligned}&\vartheta_{si}(\hat{W}_{i,f}^T S_{i,f}(\hat{x}_i) + l_i e_1 - \dot{\hat{\alpha}}_{i-1})\\ &\quad \le \frac{3}{2}\vartheta_{si}^2 + \frac{1}{2}l_i^2 k_e^2 + \frac{1}{2}\dot{\hat{\alpha}}_{i-1}^2 + \frac{1}{2}\hat{W}_{i,f}^T S_{i,f}(\hat{x}_i)S_{i,f}(\hat{x}_i)\hat{W}_{i,f}^T\end{aligned} \quad (76)$$

$$-\frac{\vartheta_{si}}{2r_i}\hat{W}_{ai}^T(t)S_{Ji} \le \frac{\vartheta_{si}^2}{4r_i} + \frac{1}{4r_i}\hat{W}_{ai}^T(t)S_{Ji}(s_i)S_{Ji}^T(s_i)\hat{W}_{ai}^T(t). \quad (77)$$

Substituting (74)–(77) into (73), one has

$$\begin{aligned}\dot{V}_i(t) \le {} & -\frac{\eta_i}{r_i}\tan(\frac{\pi s_i^2}{2k_{bi}^2}) - (\frac{\bar{\eta}_i}{r_i} - \frac{1}{4r_i} - 2)\vartheta_{si}^2\\ & + \frac{1}{4r_i}\hat{W}_{ai}^T(t)S_{Ji}S_{Ji}^T\hat{W}_{ai}^T(t) + \frac{l_i^2k_e^2}{2} + \frac{k_{bi+1}^2}{2}\\ & + \frac{\dot{\hat{\alpha}}_{i-1}^2}{2} + \frac{1}{2}\hat{W}_{i,f}^T S_{i,f}(\hat{x}_i)S_{i,f}^T(\hat{x}_i)\hat{W}_{i,f}^T\\ & - \frac{\gamma_{ci}}{2}\tilde{W}_{ci}^T(t)S_{Ji}S_{Ji}^T\tilde{W}_{ci}(t) - \frac{\gamma_{ci}}{2}\tilde{W}_{ai}^T(t)\\ & \times S_{Ji}S_{Ji}^T\tilde{W}_{ai}(t) - (\gamma_{ci} - \frac{\gamma_{ai}}{2})(\hat{W}_{ci}^T(t)S_{Ji})^2\\ & - \frac{\gamma_{ai}}{2}(\hat{W}_{ai}^T(t)S_{Ji}(s_i))^2 + (\frac{\gamma_{ai}}{2} + \frac{\gamma_{ci}}{2})\\ & \times (W_{Ji}^{*T}(t)S_{Ji}(s_i))^2.\end{aligned} \quad (78)$$

Let $\lambda_{S_{i,f}}^{\max}$ be the maximal eigenvalue of $S_{i,f}(\hat{x}_i)S_{i,f}(\hat{x}_i)$, and $\lambda_{S_{Ji}}^{\min}$ be the minimal eigenvalue of $S_{Ji}(s_i)S_{Ji}^T(s_i)$. Inequality (78) can then become

$$\begin{aligned}\dot{V}_i(t) \le {} & -\frac{\eta_i}{r_i}\tan(\frac{\pi s_i^2}{2k_{bi}^2}) - (\frac{\bar{\eta}_i}{r_i} - \frac{1}{4r_i} - 2)\vartheta_{si}^2\\ & - \frac{\gamma_{ci}}{2}\lambda_{S_{Ji}}^{\min}\tilde{W}_{ai}^T(t)\tilde{W}_{ai}(t) - \frac{\gamma_{ci}}{2}\lambda_{S_{Ji}}^{\min}\tilde{W}_{ci}^T(t)\tilde{W}_{ci}(t)\\ & - (\frac{\gamma_{ai}}{2} - \frac{1}{4r_i})(\hat{W}_{ai}^T(t)S_{Ji}(s_i))^2 + D_i\\ & - (\gamma_{ci} - \frac{\gamma_{ai}}{2})(\hat{W}_{ci}^T(t)S_{Ji})^2\end{aligned} \quad (79)$$

where $D_i = \sup_{t\ge0}\{D_i(t)\}$ and $D_i(t) = \frac{l_i^2k_e^2}{2} + \frac{k_{bi+1}^2}{2} + \frac{\dot{\hat{\alpha}}_{i-1}^2}{2} + \frac{\gamma_{ai}+\gamma_{ci}}{2}(W_{Ji}^{*T}S_{Ji})^2 + \frac{1}{2}\lambda_{S_{i,f}}^{\max}\hat{W}_{i,f}^T(\hat{x}_i)\hat{W}_{i,f}(\hat{x}_i)$.

We design the parameters $\gamma_{ai}$, $\gamma_{ci}$, $r_i$, and $\bar{\eta}_i$, which satisfy the following inequalities:

$$\gamma_{ci} - \frac{\gamma_{ai}}{2} > 0 \quad (80)$$

$$\frac{\gamma_{ai}}{2} - \frac{1}{4r_i} > 0 \quad (81)$$

$$\frac{\bar{\eta}_i}{r_i}-\frac{1}{4r_i}-2>0. \tag{82}$$

From (80)–(82) and Lemma 1, we have

$$\dot{V}_i(t)\le -\frac{\eta_i}{r_i}\tan\left(\frac{\pi s_i^2}{2k_{bi}^2}\right)-\frac{\gamma_{ci}}{2}\lambda_{S_{Ji}}^{\min}\tilde{W}_{ai}^T(t)\tilde{W}_{ai}(t) -\frac{\gamma_{ci}}{2}\lambda_{S_{Ji}}^{\min}\tilde{W}_{ci}^T(t)\tilde{W}_{ci}(t)+D_i. \tag{83}$$

Denote $\eta_{i0}=\eta_i\pi/k_{bi}^2$, (83) can then be rewritten as

$$\dot{V}_i(t)\le -\frac{k_{bi}^2\eta_{i0}}{\pi r_i}\tan\left(\frac{\pi s_i^2}{2k_{bi}^2}\right)-\frac{\gamma_{ci}}{2}\lambda_{S_{Ji}}^{\min}\tilde{W}_{ai}^T(t)\tilde{W}_{ai}(t) -\frac{\gamma_{ci}}{2}\lambda_{S_{Ji}}^{\min}\tilde{W}_{ci}^T(t)\tilde{W}_{ci}(t)+D_i. \tag{84}$$

Let $c_i=\min\{\eta_{i0}/r_i,\gamma_{ci}\lambda_{S_{Ji}}^{\min},\gamma_{ci}\lambda_{S_{Ji}}^{\min}\}$, then (84) becomes

$$\dot{V}_i\le -c_iV_i+D_i. \tag{85}$$

From (85), we have that

$$V_i(t)\le V_i(t_0)e^{-c_i(t-t_0)}+\frac{D_i}{c_i}. \tag{86}$$

Since $s_i=\hat{x}_i-\hat{\alpha}_{i-1}$, we have $|\hat{x}_i|=|s_i+\hat{\alpha}_{i-1}|<k_{bi}+A_i$, ($|\hat{\alpha}_{i-1}|\le A_i$, $A_i$ is a positive constant). Since $e_i(t)=x_i-\hat{x}_i$, it has $|x_i|=|e_i+\hat{x}_i|<k_e+k_{bi}+A_i$. Therefore, if we define $k_{bi}<k_{ci}-A_i-k_e$, then we can prove $|x_i(t)|\le k_{ci}$. From (86), as $t\to\infty$, $e^{-(t-t_0)}\to 0$. It follows that there exists $T_i$, when $t\ge T_i$, $\|\tilde{W}_{ai}(t)\|\le\sqrt{2D_i/c_i}$ and $\|\tilde{W}_{ci}(t)\|\le\sqrt{2D_i/c_i}$. Then, we can obtain that $e_i$, $s_i$, $\|\hat{W}_{ai}(t)\|$, $\|\hat{W}_{ci}(t)\|$, and $\|S_i\|$ are bounded, so $\hat{\alpha}_i$, $\|\dot{\hat{W}}_{ai}(t)\|$, and $\dot{s}_i$ are bounded ($|\hat{\alpha}_i|\le A_{i+1}$, $A_{i+1}>0$ is a constant), and then $\dot{\hat{\alpha}}_i$ is bounded.

*Step n:* Define the error variable as $s_n=\hat{x}_n-\hat{\alpha}_{n-1}-\lambda$ for the $s_n$-system. In order to compensate for the effect of the saturation, the following system is constructed to generate signal:

$$\dot{\lambda}=-k\lambda+\triangle u \tag{87}$$

where $k$ is a positive constant and $\triangle u=h(u)-u$.

The following change of coordinates is made:

$$\dot{s}_n=u+k\lambda+\hat{W}_{n,f}^TS_{n,f}(\hat{\bar{x}}_n)+l_n(y-\hat{y})-\dot{\hat{\alpha}}_{n-1}. \tag{88}$$

Considering the auxiliary dynamic system (87), the optimal value function for the $s_n$-subsystem is expressed as

$$J_n^*(s_n)=\min_{u\in\Psi(\Omega_{sn})}\int_t^\infty\left(\frac{k_{bn}^2}{\pi}\tan\left(\frac{\pi s_n^2}{2k_{bn}^2}\right)+r_n(u(\tau))^2\right)d\tau =\int_t^\infty\left(\frac{k_{bn}^2}{\pi}\tan\left(\frac{\pi s_n^2}{2k_{bn}^2}\right)+r_n(u^*(\tau))^2\right)d\tau \tag{89}$$

where $u$ is the optimal controller, $\Psi(\Omega_{sn})$ is the admissible control set of $u$, $\Omega_{sn}=\{s_n:|s_n|<k_{bn}\}$ and $r_n>0$ is a constant.

The optimal value function (89) can be rewritten as the following equation:

$$\begin{aligned}J_n^*(s_n)&=\eta_n\frac{S_n(n)}{2}+2\bar{\eta}_n\frac{k_{bn}^2}{\pi}\tan\left(\frac{\pi s_n^2}{2k_{bn}^2}\right)-\eta_n\frac{S_n(n)}{2}\\&\quad-2\bar{\eta}_n\frac{k_{bn}^2}{\pi}\tan\left(\frac{\pi s_n^2}{2k_{bn}^2}\right)+J_n^*(s_n)\\&=\eta_n\frac{S_n(n)}{2}+2\bar{\eta}_n\frac{k_{bn}^2}{\pi}\tan\left(\frac{\pi s_n^2}{2k_{bn}^2}\right)+J_{cn}(s_n)\end{aligned} \tag{90}$$

where $J_{cn}(s_n)=-\eta_nS_n(n)/2-2\bar{\eta}_nk_{bn}^2\tan(\pi s_n^2/2k_{bn}^2)/\pi+J_n^*(s_n)$.

The HJB equation of the $s_n$-subsystem is defined as

$$\begin{aligned}&H_n\left(s_n,u^*,\frac{\partial J_n^*(s_n)}{\partial s_n}\right)\\&=\frac{k_{bn}^2}{\pi}\tan\left(\frac{\pi s_n^2}{2k_{bn}^2}\right)+r_n(u^*)^2+\left(\frac{2\eta_n}{s_n}\sin\left(\frac{\pi s_n^2}{2k_{bn}^2}\right)\cos\left(\frac{\pi s_n^2}{2k_{bn}^2}\right)\right.\\&\quad\left.+\frac{\partial J_{cn}(s_n)}{\partial s_n}+\frac{2\bar{\eta}_ns_n}{\cos^2(\pi s_n^2/2k_{bn}^2)}\right)(u^*+k\lambda+\hat{W}_{n,f}^T\\&\quad\times S_{n,f}(\hat{\bar{x}}_n)+l_n(y-\hat{y})-\dot{\hat{\alpha}}_{n-1})=0.\end{aligned} \tag{91}$$

By solving $\partial H_n/\partial u^*=0$, we can obtain

$$u^*=-\frac{\eta_n}{r_ns_n}\sin\left(\frac{\pi s_n^2}{2k_{bn}^2}\right)\cos\left(\frac{\pi s_n^2}{2k_{bn}^2}\right)-\frac{\bar{\eta}_ns_n}{r_n\cos^2(\pi s_n^2/2k_{bn}^2)} -\frac{1}{2r_n}\frac{\partial J_{cn}(s_n)}{\partial s_n}. \tag{92}$$

Note that $\partial J_{cn}(s_n)/\partial s_n$ is an unknown function of variable $s_n$. It can be approximated as follows:

$$\frac{\partial J_{cn}(s_n)}{\partial s_n}=W_n^{*T}S_{Jn}(s_n)+\varepsilon_n(s_n) \tag{93}$$

where $W_n^*$ is an ideal weight vector, $S_{Jn}(s_n)$ is the basis function vector. $\varepsilon_n(s_n)$ is the approximation error satisfying $|\varepsilon_n(s_n)|\le\bar{\delta}_n$ and $\bar{\delta}_n$ is a positive real constant. The ideal optimal virtual controller $u^*$ can be devised as

$$u^*=-\frac{\eta_n}{r_ns_n}\sin\left(\frac{\pi s_n^2}{2k_{bn}^2}\right)\cos\left(\frac{\pi s_n^2}{2k_{bn}^2}\right)-\frac{\bar{\eta}_ns_n}{r_n\cos^2(\pi s_n^2/2k_{bn}^2)} -\frac{1}{2r_n}(W_n^{*T}S_{Jn}(s_n)+\varepsilon_n(s_n)). \tag{94}$$

From (92) and (93), we can get

$$\frac{\partial\hat{J}_{cn}(s_n)}{\partial s_n}=\hat{W}_{cn}^TS_{Jn}(s_n) \tag{95}$$

$$u=-\frac{\eta_n}{r_ns_n}\sin\left(\frac{\pi s_n^2}{2k_{bn}^2}\right)\cos\left(\frac{\pi s_n^2}{2k_{bn}^2}\right)-\frac{\bar{\eta}_ns_n}{r_n\cos^2(\pi s_n^2/2k_{bn}^2)} -\frac{1}{2r_n}\hat{W}_{an}^TS_{Jn}(s_n) \tag{96}$$

where $\hat{W}_{cn}$ and $\hat{W}_{an}$ are the critic and actor NN weights, respectively. Similarly, $\mathrm{sgn}(s_n)s_n\ne k_{bn}\sqrt{\Upsilon}$ ($\Upsilon=1,2,\ldots$) is obvious. Then, we can get $\lim_{s_n\to 0}\sin(\pi s_n^2/2k_{bn}^2)\cos(\pi s_n^2/2k_{bn}^2)/s_n\to 0$. The singularity problem in the optimal controller $u$ is effectively avoided.

From (91), (95), and (96), the approximate Hamiltonian is

$$\begin{aligned}
&H_n(s_n,u,\frac{\partial\hat{J}_n(s_n)}{\partial s_n})\\
&=\frac{k_{bn}^2}{\pi}\tan(\frac{\pi s_n^2}{2k_{bn}^2})+r_n(\frac{\eta_n}{r_n s_n}\sin(\frac{\pi s_n^2}{2k_{bn}^2})\cos(\frac{\pi s_n^2}{2k_{bn}^2})\\
&\quad+\frac{\bar{\eta}_n s_n}{r_n\cos^2(\pi s_n^2/2k_{bn}^2)}+\frac{1}{2r_n}\hat{W}_{an}^T S_{Jn})^2+(\frac{2\eta_n}{s_n}\sin(\frac{\pi s_n^2}{2k_{bn}^2})\\
&\quad\times\cos(\frac{\pi s_n^2}{2k_{bn}^2})+\frac{2\bar{\eta}_n s_n}{\cos^2(\pi s_n^2/2k_{bn}^2)}+\frac{\partial J_{cn}}{\partial s_n})(u+\hat{W}_{n,f}^T S_{n,f}\\
&\quad+l_n(y-\hat{y})-\dot{\hat{\alpha}}_{n-1}+k\lambda).
\end{aligned}\tag{97}$$

The Bellman error is defined as

$$\begin{aligned}
E_n&=H_n(s_n,u,\frac{\partial\hat{J}_{cn}(s_n)}{\partial s_n})-H_n(s_n,u^*,\frac{\partial J_n^*(s_n)}{\partial s_n})\\
&=H_n(s_n,u,\frac{\partial\hat{J}_{cn}(s_n)}{\partial s_n}).
\end{aligned}\tag{98}$$

The critic NN adaptive law is given as

$$\dot{\hat{W}}_{cn}(t)=-\gamma_{cn}S_{Jn}(s_n)S_{Jn}^T(s_n)\hat{W}_{cn}^T(t)\tag{99}$$

where $\gamma_{cn}>0$ is the critic designed constant.

In order to ensure the stability and optimal performance of the nonlinear system, the actor NN adaptive law is designed as

$$\begin{aligned}
\dot{\hat{W}}_{an}(t)=&-S_{Jn}(s_n)S_{Jn}^T(s_n)(\gamma_{an}(\hat{W}_{an}(t)-\hat{W}_{cn}(t))\\
&+\gamma_{cn}\hat{W}_{cn}(t))
\end{aligned}\tag{100}$$

where $\gamma_{an}>0$ is the critic designed constant.

Consider the overall Lyapunov function candidate for the final step as

$$\begin{aligned}
V(t)=&\sum_{i=1}^{n-1}V_i+\frac{1}{2}\lambda^2+\frac{k_{bn}^2}{\pi}\tan(\frac{\pi s_n^2}{2k_{bn}^2})+\frac{1}{2}\tilde{W}_{cn}^T(t)\tilde{W}_{cn}(t)\\
&+\frac{1}{2}\tilde{W}_{an}^T(t)\tilde{W}_{an}(t)
\end{aligned}\tag{101}$$

where $\tilde{W}_{cn}=\hat{W}_{cn}-W_n^*$ and $\tilde{W}_{an}=\hat{W}_{an}-W_n^*$ are critic and actor NN approximation errors, respectively.

The time derivative of Lyapunov function $V_n$ is

$$\begin{aligned}
\dot{V}(t)=&\sum_{i=1}^{n-1}\dot{V}_i+\lambda\dot{\lambda}+\frac{s_n}{\cos^2(\pi s_n^2/2k_{bn}^2)}\dot{s}_n+\tilde{W}_{cn}^T(t)\dot{\hat{W}}_{cn}(t)\\
&+\tilde{W}_{an}^T(t)\dot{\hat{W}}_{an}(t).
\end{aligned}\tag{102}$$

Similarly, we can get

$$\begin{aligned}
\dot{V}(t)=&\sum_{i=1}^{n-1}\dot{V}_i-\frac{\eta_n}{r_n}\tan(\frac{\pi s_n^2}{2k_{bn}^2})-\frac{\bar{\eta}_n}{r_n}\vartheta_{sn}^2-\frac{\vartheta_{sn}}{2r_n}\hat{W}_{an}^T S_{Jn}\\
&+\vartheta_{sn}(k\lambda-\dot{\hat{\alpha}}_{n-1}+l_n e_1+\hat{W}_{n,f}^T S_{n,f}(\hat{x}_n))\\
&-\lambda(-k\lambda+\Delta u)-\gamma_{cn}\tilde{W}_{cn}^T(t)S_{Jn}S_{Jn}^T\hat{W}_{cn}^T(t)\\
&+(\gamma_{an}-\gamma_{cn})\tilde{W}_{an}^T(t)S_{Jn}(s_n)S_{Jn}^T(s_n)\hat{W}_{cn}(t)\\
&-\gamma_{an}\tilde{W}_{an}^T(t)S_{Jn}(s_n)S_{Jn}^T(s_n)\hat{W}_{an}(t).
\end{aligned}\tag{103}$$

By using $\tilde{W}_{cn}(t)=\hat{W}_{cn}(t)-W_n^*$ and $\tilde{W}_{an}(t)=\hat{W}_{an}(t)-W_n^*$, there are the following equations:

$$\begin{aligned}
&\tilde{W}_{cn}^T(t)S_{Jn}(s_n)S_{Jn}^T(s_n)\hat{W}_{cn}(t)\\
&=\frac{1}{2}\tilde{W}_{cn}^T(t)S_{Jn}(s_n)S_{Jn}^T(s_n)\tilde{W}_{cn}(t)+\frac{1}{2}\hat{W}_{cn}^T(t)S_{Jn}(s_n)\\
&\quad\times S_{Jn}^T(s_n)\hat{W}_{cn}(t)-\frac{1}{2}(W_{Jn}^{*T}S_{Jn}(s_n))^2
\end{aligned}\tag{104}$$

$$\begin{aligned}
&\tilde{W}_{an}^T(t)S_{Jn}(s_n)S_{Jn}^T(s_n)\hat{W}_{an}(t)\\
&=\frac{1}{2}\tilde{W}_{an}^T(t)S_{Jn}(s_n)S_{Jn}^T(s_n)\tilde{W}_{an}(t)+\frac{1}{2}\hat{W}_{an}^T(t)S_{Jn}(s_n)\\
&\quad\times S_{Jn}^T(s_n)\hat{W}_{an}(t)-\frac{1}{2}(W_{Jn}^{*T}S_{Jn}(s_n))^2.
\end{aligned}\tag{105}$$

Substituting (104) and (105) into (103), one has

$$\begin{aligned}
\dot{V}(t)=&\sum_{i=1}^{n-1}\dot{V}_i-\frac{\eta_n}{r_n}\tan(\frac{\pi s_n^2}{2k_{bn}^2})-\frac{\bar{\eta}_n}{r_n}\vartheta_{sn}^2-\frac{\vartheta_{sn}}{2r_n}\hat{W}_{an}^T S_{Jn}\\
&+\vartheta_{sn}(\hat{W}_{n,f}^T S_{n,f}(\hat{x}_n)+k\lambda+l_n e_1-\dot{\hat{\alpha}}_{n-1})\\
&-\lambda(-k\lambda+\Delta u)-\frac{\gamma_{cn}}{2}\tilde{W}_{cn}^T(t)S_{Jn}(s_n)\\
&\times S_{Jn}^T(s_n)\tilde{W}_{cn}(t)-\frac{\gamma_{cn}}{2}\hat{W}_{cn}^T(t)S_{Jn}(s_n)\\
&\times S_{Jn}^T(s_n)\hat{W}_{cn}(t)-\frac{\gamma_{an}}{2}\tilde{W}_{an}^T(t)S_{Jn}(s_n)\\
&\times S_{Jn}^T(s_n)\tilde{W}_{an}(t)-\frac{\gamma_{an}}{2}\hat{W}_{an}^T(t)S_{Jn}(s_n)\\
&\times S_{Jn}^T(s_n)\hat{W}_{an}(t)+(\frac{\gamma_{an}}{2}+\frac{\gamma_{cn}}{2})(W_{Jn}^{*T}S_{Jn}(s_n))^2\\
&+(\gamma_{an}-\gamma_{cn})\tilde{W}_{an}^T(t)S_{Jn}(s_n)S_{Jn}^T(s_n)\hat{W}_{cn}(t).
\end{aligned}\tag{106}$$

Using Young's inequality, there is the following fact that:

$$\begin{aligned}
&(\gamma_{an}-\gamma_{cn})\tilde{W}_{an}^T(t)S_{Jn}(s_n)S_{Jn}^T(s_n)\hat{W}_{cn}(t)\\
&\quad\le\frac{\gamma_{an}-\gamma_{cn}}{2}\tilde{W}_{an}^T(t)S_{Jn}(s_n)S_{Jn}^T(s_n)\tilde{W}_{an}(t)\\
&\qquad+\frac{\gamma_{an}-\gamma_{cn}}{2}\hat{W}_{cn}^T(t)S_{Jn}(s_n)S_{Jn}^T(s_n)\hat{W}_{cn}(t)
\end{aligned}\tag{107}$$

$$\begin{aligned}
&\vartheta_{sn}(\hat{W}_{n,f}^T S_{n,f}(\hat{x}_n)+k\lambda+l_n e_1-\dot{\hat{\alpha}}_{n-1})\\
&\quad\le\frac{3}{2}\vartheta_{sn}^2+\frac{k\vartheta_{sn}^2}{2}+\frac{k}{2}\lambda^2+\frac{1}{2}l_n^2k_e^2+\frac{1}{2}\dot{\hat{\alpha}}_{n-1}^2\\
&\qquad+\frac{1}{2}\hat{W}_{n,f}^T S_{n,f}(\hat{x}_n)S_{n,f}(\hat{x}_n)\hat{W}_{n,f}^T
\end{aligned}\tag{108}$$

$$\begin{aligned}
&-\frac{\vartheta_{sn}}{2r_i}\hat{W}_{an}^T(t)S_{Jn}\\
&\quad\le\frac{\vartheta_{sn}^2}{4r_n}+\frac{1}{4r_n}\hat{W}_{an}^T(t)S_{Jn}(s_n)S_{Jn}^T(s_n)\hat{W}_{an}^T(t)
\end{aligned}\tag{109}$$

$$-\lambda(-k\lambda+\Delta u)\le-k\lambda^2+\frac{1}{2}\lambda^2+\frac{1}{2}\Delta u^2.\tag{110}$$

Substituting (107)–(110) into (106), one has

$$\begin{aligned}\dot{V}(t) \le & \sum_{i=1}^{n-1} \dot{V}_i - \frac{\eta_n}{r_n}\tan(\frac{\pi s_n^2}{2k_{bn}^2}) - (\frac{\bar{\eta}_n}{r_n} - \frac{1}{4r_n} - \frac{3}{2} - \frac{k}{2})\vartheta_{sn}^2 \\ & + \frac{l_n^2 k_e^2}{2} + \frac{\Delta u^2}{2} + \frac{\lambda^2}{2} - \frac{\gamma_{cn}}{2}\tilde{W}_{cn}^T(t) S_{Jn} S_{Jn}^T \tilde{W}_{cn}(t) \\ & - \frac{\gamma_{cn}}{2}\tilde{W}_{an}^T(t) S_{Jn} S_{Jn}^T \tilde{W}_{an}(t) - \frac{k\lambda^2}{2} + \frac{\dot{\hat{\alpha}}_{n-1}^2}{2} \\ & - (\gamma_{cn} - \frac{\gamma_{an}}{2})(\hat{W}_{cn}^T(t) S_{Jn}(s_n))^2 - (\frac{\gamma_{an}}{2} \\ & - \frac{1}{4r_n})(\hat{W}_{an}^T(t) S_{Jn}(s_n))^2 + \frac{\gamma_{an} + \gamma_{cn}}{2} \\ & \times (W_{Jn}^{*T}(t) S_{Jn}(s_n))^2 + \frac{1}{2}\hat{W}_{n,f}^T S_{n,f}(\hat{x}_n) \\ & \times S_{n,f}^T(\hat{x}_n)\hat{W}_{n,f}^T. \end{aligned} \tag{111}$$

Let $\lambda_{S_{n,f}}^{\max}$ be the maximal eigenvalue of $S_{n,f}(\hat{x}_n)S_{n,f}^T(\hat{x}_n)$ and $\lambda_{S_{Jn}}^{\min}$ be the minimal eigenvalue of $S_{Jn}(s_n)S_{Jn}^T(s_n)$, $\dot{V}_n(t)$ can be rewritten as

$$\begin{aligned}\dot{V}(t) \le & \sum_{i=1}^{n-1} \dot{V}_i - \frac{\eta_n}{r_n}\tan(\frac{\pi s_n^2}{2k_{bn}^2}) - (\frac{\bar{\eta}_n}{r_n} - \frac{1}{4r_n} - \frac{3}{2} - \frac{k}{2})\vartheta_{sn}^2 \\ & + \frac{1}{2}l_n^2 k_e^2 + \frac{1}{2}\dot{\hat{\alpha}}_{n-1}^2 + \frac{1}{2}\lambda_{S_{n,f}}^{\max}\hat{W}_{n,f}^T(\hat{x}_n)\hat{W}_{n,f}(\hat{x}_n) \\ & - (\frac{k}{2} - \frac{1}{2})\lambda^2 + \frac{1}{2}\Delta u^2 - (\gamma_{cn} - \frac{\gamma_{an}}{2})(\hat{W}_{cn}^T(t) S_{Jn})^2 \\ & - \frac{\gamma_{cn}}{2}\lambda_{S_{Jn}}^{\min}\tilde{W}_{an}^T(t)\tilde{W}_{an}(t) - \frac{\gamma_{cn}}{2}\lambda_{S_{Jn}}^{\min}\tilde{W}_{cn}^T(t)\tilde{W}_{cn}(t) \\ & - (\frac{\gamma_{an}}{2} - \frac{1}{4r_n})(\hat{W}_{an}^T(t) S_{Jn})^2 + (\frac{\gamma_{an}}{2} + \frac{\gamma_{cn}}{2}) \\ & \times (W_{Jn}^{*T}(t) S_{Jn})^2. \end{aligned} \tag{112}$$

By designing the parameters $\gamma_{an}$, $\gamma_{cn}$, $r_n$, $\bar{\eta}_n$, and $k$, which satisfy the following inequalities:

$$\gamma_{cn} - \frac{\gamma_{an}}{2} > 0 \tag{113}$$

$$\frac{\gamma_{an}}{2} - \frac{1}{4r_n} > 0 \tag{114}$$

$$\frac{\bar{\eta}_n}{r_n} - \frac{1}{4r_n} - \frac{3}{2} - \frac{k}{2} > 0 \tag{115}$$

$$\frac{k}{2} - \frac{1}{2} > 0. \tag{116}$$

Inequality (112) is rewritten as

$$\begin{aligned}\dot{V}(t) \le & \sum_{i=1}^{n-1} \dot{V}_i - \frac{\eta_n}{r_n}\tan(\frac{\pi s_n^2}{2k_{bn}^2}) - (\frac{k}{2} - \frac{1}{2})\lambda^2 \\ & - \frac{\gamma_{cn}}{2}\lambda_{S_{Jn}}^{\min}\tilde{W}_{an}^T(t)\tilde{W}_{an}(t) - \frac{\gamma_{cn}}{2}\lambda_{S_{Jn}}^{\min}\tilde{W}_{cn}^T(t) \\ & \times \tilde{W}_{cn}(t) + D_n \end{aligned} \tag{117}$$

where $D_n = \sup_{t\ge0}\{D_n(t)\}$ and $D_n(t) = \frac{l_n^2 k_e^2}{2} + \frac{\dot{\hat{\alpha}}_{n-1}^2}{2} + \frac{1}{2}\lambda_{S_{n,f}}^{\max}\hat{W}_{n,f}^T(\hat{x}_n)\hat{W}_{n,f}(\hat{x}_n) + \frac{\Delta u^2}{2} + \frac{\gamma_{an}+\gamma_{cn}}{2}(W_{Jn}^{*T}(t) S_{Jn})^2$.

Denote $\eta_{n0} = \eta_n \pi / k_{bn}^2$, then (117) can be rewritten as

$$\begin{aligned}\dot{V}(t) \le & \sum_{i=1}^{n-1} \dot{V}_i - \frac{k_{bn}^2 \eta_{n0}}{\pi r_n}\tan(\frac{\pi s_n^2}{2k_{bn}^2}) - (\frac{k}{2} - \frac{1}{2})\lambda^2 + D_n \\ & - \frac{\gamma_{cn}}{2}\lambda_{S_{Jn}}^{\min}\tilde{W}_{an}^T(t)\tilde{W}_{an}(t) - \frac{\gamma_{cn}}{2}\lambda_{S_{Jn}}^{\min}\tilde{W}_{cn}^T(t)\tilde{W}_{cn}(t). \end{aligned} \tag{118}$$

Let $c_n = \min\{\eta_{n0}/r_n, k-1, \gamma_{cn}\lambda_{S_{Jn}}^{\min}, \gamma_{cn}\lambda_{S_{Jn}}^{\min}\}$, then (118) becomes

$$\dot{V} \le -cV + D. \tag{119}$$

Let $c = \min_{1\le i\le n}\{c_i\}$ and $D = \sum_{i=1}^n D_i$. Then, (119) becomes

$$V(t) \le V(t_0)e^{-c(t-t_0)} + \frac{D}{c}. \tag{120}$$

From (120), it follows that there exists $T = \max_{1\le i\le n}\{T_i\}$, when $t \ge T$, $|s_i| \le \sqrt{2D_n/c_n}$, $\|\tilde{W}_{ai}(t)\| \le \sqrt{2D_n/c_n}$, and $\|\tilde{W}_{ci}(t)\| \le \sqrt{2D_n/c_n}$ $(i = 1, ..., n)$. Clearly, the reduction of $\sqrt{2D_n/c_n}$ can be achieved by increasing $c_n$ or decreasing $D_n$. Therefore, the parameter $c_n$ can be chosen to be large enough to render the tracking error and $|s_i(t)|$, $\|\tilde{W}_{ai}(t)\|$, and $\|\tilde{W}_{ci}(t)\|$ sufficiently small. Then, we can obtain that $\hat{W}_{ai}(t)$, $\hat{W}_{ci}(t)$, $\hat{W}_{i,f}(t)$, and $\hat{x}_i$ are bounded, and from Theorem 1, $x_i$ is UUB and $|x_i| \le k_{ci} (i = 1, ..., n)$.

## IV. Simulation Example

In this section, an example will be used to test the effectiveness of the proposed controller. Consider the following strict-feedback nonlinear systems as:

$$\dot{x}_1(t) = x_2(t) - \sin(2x_1)\cos(2x_1)$$
$$\dot{x}_2(t) = (1 - (2 + \sin(x_1)\cos(x_2))^2 + u_s$$

where $x_1(t)$ and $x_2(t) \in \mathbb{R}$ are the system states and $u_s \in \mathbb{R}$ represents the saturation form of the control input. The reference signal is given as $y_r = 2.5\sin(t-2) + 1$.

Then, the state observer is designed as

$$\dot{\hat{x}}_1(t) = \hat{x}_2(t) - \hat{W}_{1,f}^T S_{1,f}(\hat{x}_1) + 4(y - \hat{y})$$
$$\dot{\hat{x}}_2(t) = \hat{W}_{2,f}^T S_{2,f}(\hat{x}_1, \hat{x}_2) + h(u) + 8(y - \hat{y})$$
$$\hat{y} = \hat{x}_1.$$

Letting $Q = I$ and solving (4), we can get a positive-definite matrix

$$P = \begin{bmatrix} 1.49 & -0.5 \\ -0.5 & 0.1567 \end{bmatrix}.$$

In the auxiliary dynamic system (87), the parameter $k = 5$. The design parameters of $\hat{\alpha}_1(t)$ (26), $u(t)$ (96), $\dot{\hat{W}}_f$ (9), $\dot{\hat{W}}_{c1}$ (29), $\dot{\hat{W}}_{a1}$ (30), $\dot{\hat{W}}_{c2}$ (99), and $\dot{\hat{W}}_{a2}$ (100) are chosen as $\eta_1 = 80$, $\eta_2 = 20$, $\bar{\eta}_1 = 58$, $\bar{\eta}_2 = 4.75$, $k_{b1} = 8$, $k_{b2} = 22$, $r_1 = 20$, $r_2 = 25$, $\gamma_{a1} = 5$, $\gamma_{c1} = 0.5$, $\gamma_{a2} = 1.7$, $\gamma_{c2} = 2$, $\eta_{W1} = 0.5$, $\eta_{W2} = 0.2$, $\rho_{W1} = 15$, $\rho_{W2} = 7$, $\rho = 5$. The constrained boundaries are $k_{c1} = 4$ and $k_{c2} = 20$.

The initial values are set as $x_1(0) = -0.3$, $\hat{W}_{c1}(0) = [1,1,1,1,1]^T$, $\hat{W}_{a1}(0) = [0.1, 0.1, 0.1, 0.1, 0.1]^T$, $\hat{W}_{c2}(0) = [1,1,1,1,1,1]^T$, $\hat{W}_{a2}(0) = [2,2,2,2,2]^T$, $\hat{W}_{f1}(0) = [0.2, 0.2, 0.2, 0.2, 0.2]^T$, $\hat{W}_{f2}(0) = [0.5, 0.5, 0.5, 0.5, 0.5]^T$, $\lambda(0) = -1$, and other initial values are zeros.

The simulation results are shown by Figs. 1–8. Fig. 1 shows the control output $y$ and the reference signal $y_r$, it is clear that an ideal tracking performance can be obtained. Figs. 2 and 3 show the trajectories of states $x_i$ and their estimates $\hat{x}_i$, $(i=1,2)$ along with $|x_1| \leq k_{c1}$ and $|x_2| \leq k_{c2}$, respectively. Figs. 4–6 profile the 2-norm of the weights for the critic, actor

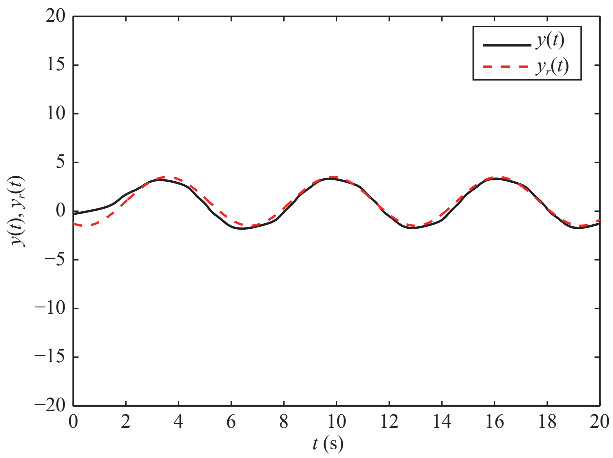


Fig. 1. The trajectories of $y$ and $y_r$.

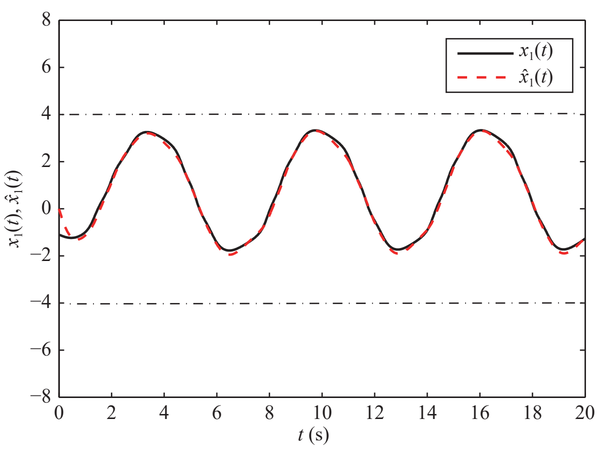


Fig. 2. The trajectories of $x_1$ and $\hat{x}_1$.

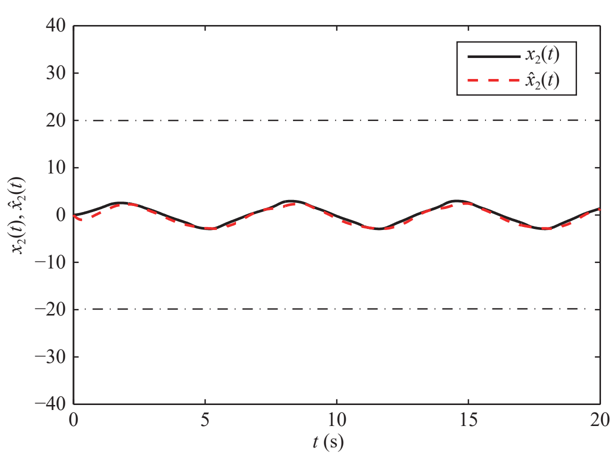


Fig. 3. The trajectories of $x_2$ and $\hat{x}_2$.

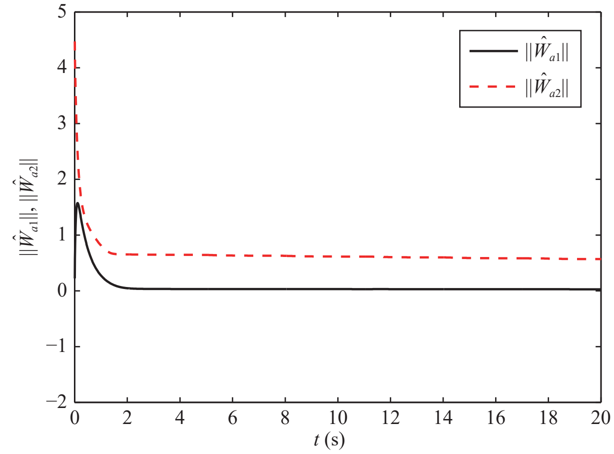


Fig. 4. The trajectories of NN weights $\|\hat{W}_{a1}\|$ and $\|\hat{W}_{a2}\|$.

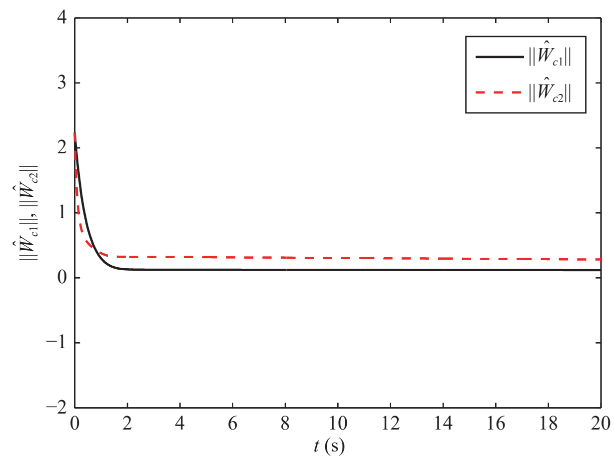


Fig. 5. The trajectories of NN weights $\|\hat{W}_{c1}\|$ and $\|\hat{W}_{c2}\|$.

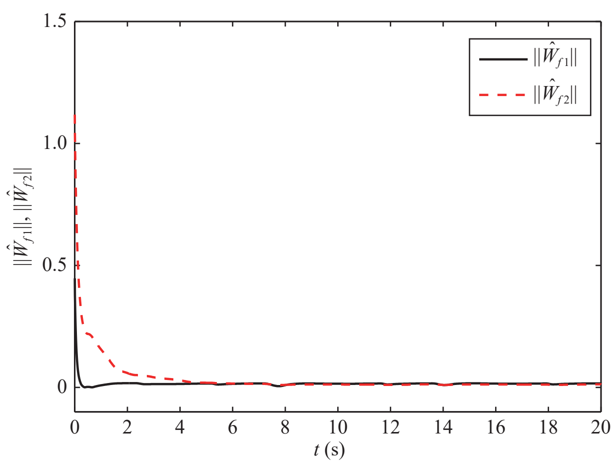


Fig. 6. The trajectories of NN weights $\|\hat{W}_{f1}\|$ and $\|\hat{W}_{f2}\|$.

and observer NN; Figs. 7 and 8 display the trajectories of controller $u$ without input saturation and with input saturation, respectively.

It can be clearly observed from the simulation results that the proposed control method ensures all signals in the closed-loop system are UUB, that the system output $y$ can track the

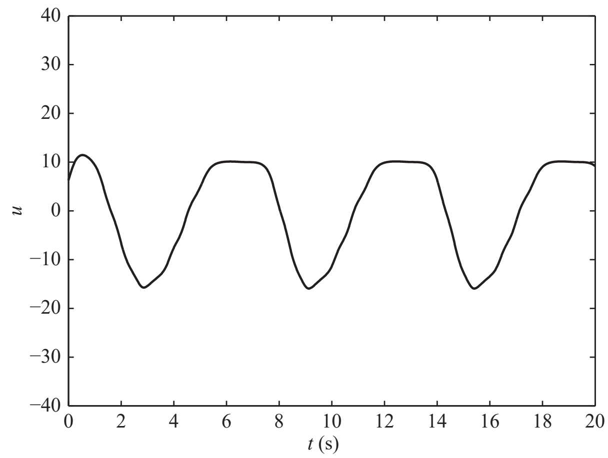


Fig. 7. The trajectory of controller $u(t)$ without input saturation.

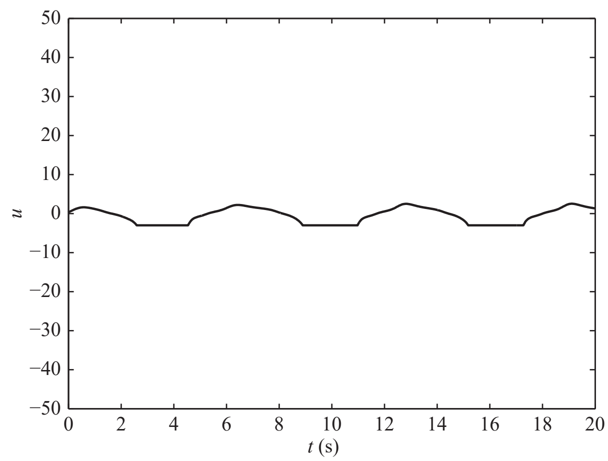


Fig. 8. The trajectory of controller $u(t)$ with input saturation.

given reference signal, and that all the system states are ensured not to violate any constraints.

## V. Conclusion

In this paper, an optimal control has been developed based on the backstepping technique using a simplified RL for a class of uncertain nonlinear systems with unmeasured states, input saturation and state constraints. The immeasurable states were approximated by the state-observer. At the same time, the tan-type BLF has been introduced to vary the constraint boundary. Meanwhile, the control design can also release the condition of persistent excitation. Based on the Lyapunov method, it was proven that the proposed adaptive NN optimal controller can ensure that the closed-loop system is UUB. In addition, the tracking error of the system converges to a small neighborhood of the origin and all states did not violate their constraints. Finally, the simulation further demonstrated the effectiveness of the proposed control method. One possible research point for future research is to extend the SISO system in this work to the MIMO case with milder assumptions.

## References

[1] S. S. Ge and C. Wang, "Direct adaptive NN control of a class of nonlinear systems," *IEEE Trans. Neural Networks*, vol. 13, no. 1, pp. 214–221, 2002.

[2] B. Chen, X. P. Liu, K. F. Liu, and C. Lin, "Direct adaptive fuzzy control of nonlinear strict-feedback systems," *Automatica*, vol. 45, no. 6, pp. 1530–1535, 2009.

[3] Y. H. Li, S. Qiang, X. Y. Zhuang, and O. Kaynak, "Robust and adaptive backstepping control for nonlinear systems using RBF neural networks," *IEEE Trans. Neural Networks*, vol. 15, no. 3, pp. 693–701, 2004.

[4] R. Dong, S. G. Gao, B. Ning, T. Tang, Y. D. Li, and K. P. Valavanis, "Error-driven nonlinear feedback design for fuzzy adaptive dynamic surface control of nonlinear systems with prescribed tracking performance," *IEEE Trans. Systems, Man, and Cybernetics: Systems*, vol. 50, no. 3, pp. 1013–1023, 2020.

[5] B. Chen and C. Lin, "Finite-time stabilization-based adaptive fuzzy control design," *IEEE Trans. Fuzzy Systems*, 2020. DOI: 10.1109/TFUZZ.2020.2991153

[6] S. G. Gao, Y. H. Hou, H. R. Dong, Y. X. Yue, and S. Y. Li, "Global nested PID control of strict-feedback nonlinear systems with prescribed output and virtual tracking performance," *IEEE Trans. Circuits and Systems II: Express Briefs*, vol. 67, no. 2, pp. 325–329, 2020.

[7] S. C. Tong, Y. M. Li, G. Feng, and T. S. Li, "Observer-based adaptive fuzzy backstepping dynamic surface control for a class of MIMO nonlinear systems," *IEEE Trans. Systems, Man, and Cybernetics, Part B*, vol. 41, no. 4, pp. 1124–1135, 2011.

[8] D. R. Ding, Z. D. Wang, and Q. L. Han, "Neural-network-based output feedback control with stochastic communication protocols," *Automatica*, vol. 106, pp. 221–229, 2019.

[9] S. C. Tong, X. Min, and Y. X. Li, "Observer-based adaptive fuzzy tracking control for strict-feedback nonlinear systems with unknown control gain functions," *IEEE Trans. Cybernetics*, vol. 50, no. 9, pp. 3901–3913, 2020.

[10] B. Chen, H. G. Zhang, X. P. Liu, and C. Lin, "Neural observer and adaptive neural control design for a class of nonlinear systems," *IEEE Trans. Neural Networks and Learning Systems*, vol. 29, no. 9, pp. 4261–4271, 2018.

[11] K. P. Tee and S. S. Ge, "Control of nonlinear systems with partial state constraints using a Barrier Lyapunov function," *Int. J. Control*, vol. 84, no. 12, pp. 2008–2023, 2011.

[12] K. P. Tee and S. S. Ge, "Control of state-constrained nonlinear systems using integral Barrier Lyapunov functionals," in *Proc. IEEE Conf. Decision and Control*, pp. 3239–3244, 2012. DOI: 10.1109/CDC.2012.6426196.

[13] Y. J. Liu and S. C. Tong, "Barrier Lyapunov functions-based adaptive control for a class of nonlinear pure-feedback systems with full state constraints," *Automatica*, vol. 64, pp. 70–75, 2016.

[14] T. Zhang, M. Xia, and Y. Yi, "Adaptive neural dynamic surface control of strict-feedback nonlinear systems with full state constraints and unmodeled dynamics," *Automatica*, vol. 81, pp. 232–239, 2017.

[15] T. Gao, Y. J. Liu, L. Liu, and D. Li, "Adaptive neural network-based control for a class of nonlinear pure-feedback systems with time-varying full state constraints," *IEEE/CAA J. Autom. Sinica*, vol. 5, no. 5, pp. 923–933, 2018.

[16] Y. J. Liu, M. Z. Gong, S. C. Tong, C. L. Philip. Chen, and D. J. Li, "Adaptive fuzzy output feedback control for a class of nonlinear systems with full state constraints," *IEEE Trans. Fuzzy System*, vol. 26, no. 5, pp. 2607–2617, 2018.

[17] K. Sun, S. Mou, J. Qiu, T. Wang, and H. Gao, "Adaptive fuzzy control for nontriangular structural stochastic switched nonlinear systems with full state constraints," *IEEE Trans. Fuzzy Systems*, vol. 27, no. 8, pp. 1587–1601, 2019.

[18] W. C. Meng, Q. M. Yang, J. Si, and Y. X. Sun, "Adaptive neural control of a class of output-constrained nonaffine systems," *IEEE Trans. Cybernetics*, vol. 46, no. 1, pp. 85–95, 2016.

[19] Y. M. Li, T. S. Li, and X. J. Jing, "Indirect adaptive fuzzy control for input and output constrained nonlinear systems using a barrier Lyapunov function," *Int. J. Adaptive Control and Signal Processing*, vol. 28, no. 2, pp. 184–199, 2014.

[20] X. Jin, "Adaptive fault tolerant control for a class of input and state constrained MIMO nonlinear systems," *Int. J. Robust and Nonlinear Control*, vol. 26, no. 2, pp. 286–302, 2016.

[21] Q. Zhou, L. Wang, C. Wu, H. Li, and H. Du, "Adaptive fuzzy control for nonstrict-feedback systems with input saturation and output constraint," *IEEE Trans. Systems, Man, and Cybernetics: Systems*, vol. 47, no. 1, pp. 1–12, 2017.

[22] S. G. Gao, H. R. Dong, B. Ning, and Q. Zhang, "Cooperative prescribed performance tracking control for multiple high-speed trains in moving block signaling system," *IEEE Trans. Intelligent Transportation Systems*, vol. 20, no. 7, pp. 2740–2749, 2019.

[23] R. E. Bellman, "Dynamic programming," *Princeton University Press*, Princeton, NJ, USA, 1957.

[24] M. L. Chambers, "The mathematical theory of optimal processes," *J. Operational Research Society*, vol. 16, no. 4, pp. 493–494, 1965.

[25] D. R. Liu, X. Yang, D. Wang, and Q. Wei, "Reinforcement-learningbased robust controller design for continuous-time uncertain nonlinear systems subject to input constraints," *IEEE Trans. Cybernetics*, vol. 45, no. 7, pp. 1372–1385, 2015.

[26] X. Yang, D. R. Liu, and D. Wang, "Reinforcement learning for adaptive optimal control of unknown continuous-time nonlinear systems with input constraints," *Int. J. Control*, vol. 87, no. 3, pp. 553–566, 2013.

[27] X. Yang and B. Zhao, "Optimal neuro-control strategy for nonlinear systems with asymmetric input constraints," *IEEE/CAA J. Autom. Sinica*, vol. 7, no. 2, pp. 575–583, 2020.

[28] H. Modares, F. L. Lewis, and M. Naghibi-Sistani, "Integral reinforcement learning and experience replay for adaptive optimal control of partially unknown constrained-input continuous-time systems," *Automatica*, vol. 50, no. 1, pp. 193–202, 2014.

[29] Y. M. Li, T. T. Yang, and S. C. Tong, "Adaptive neural networks finite-time optimal control for a class of nonlinear systems," *IEEE Trans. Neural Networks and Learning Systems*, vol. 31, no. 11, pp. 4451–4460, 2020.

[30] G. X. Wen, S. S. Ge, and F. W. Tu, "Optimized backstepping for tracking control of strict-feedback systems," *IEEE Trans. Neural Networks and Learning Systems*, vol. 29, no. 8, pp. 3850–3862, 2018.

[31] W. W. Bai, T. S. Li, and S. C. Tong, "NN reinforcement learning adaptive control for a class of nonstrict-feedback discrete-time systems," *IEEE Trans. Cybernetics*, vol. 50, no. 11, pp. 4573–4584, 2020.

[32] G. X. Wen, C. Chen, and S. S. Ge, "Simplified optimized backstepping control for a class of nonlinear strict-feedback systems with unknown dynamic functions," *IEEE Trans. Cybernetics*, 2020. DOI: 10.1109/TCYB.2020.3002108

[33] W. Sun, S. F. Su, Y. Q. Wu, J. W. Xia, and V. T. Nguyen, "Adaptive fuzzy control with high-order barrier Lyapunov functions for highorder uncertain nonlinear systems with full-state constraints," *IEEE Trans. Cybernetics*, vol. 50, no. 8, pp. 3424–3432, 2019.

[34] H. G. Zhang, L. L. Cui, and Y. H. Luo, "Near-optimal control for nonzerosum differential games of continuous-time nonlinear systems using single network ADP," *IEEE Trans. Cybernetics*, vol. 43, no. 1, pp. 206–216, 2013.

[35] X. Yang, D. R. Liu, and Q. L. Wei, "Online approximate optimal control for affine non-linear systems with unknown internal dynamics using adaptive dynamic programming," *IET Control Theory and Applications*, vol. 8, no. 16, pp. 1676–1688, 2014.

[36] T. C. Wang, S. Sui, and S. C. Tong, "Data-based adaptive neural network optimal output feedback control for nonlinear systems with actuator saturation," *Neurocomputing*, vol. 247, pp. 192–201, 2017.

**Jiaxin Zhang** received the B.S. degree in information and computing science from Liaoning University of Technology in 2019. She is working towards the M.S. degree in operational research and cybernetics from Liaoning University of Technology. Her current research interests include optimal control and adaptive control.

**Kewen Li** received the B.S. and M.S. degrees in applied mathematics from the Liaoning University of Technology in 2016 and 2019, respectively. He is currently pursuing the Ph.D. degree in Institute of Automation, Qufu Normal University. His current research interests include finite time control, fuzzy control, and adaptive control for nonlinear systems.

**Yongming Li** (SM'16) received the B.S. degree and the M.S. degree in applied mathematics from Liaoning University of Technology in 2004 and 2007, respectively. He received the Ph.D. degree in transportation information engineering & control from Dalian Maritime University in 2014. He is currently a Professor in the College of Science, Liaoning University of Technology. His current research interests include adaptive control, fuzzy control, and neural networks control for nonlinear systems.