Robust and Adaptive Backstepping Control for Nonlinear Systems Using RBF Neural Networks

Yahui Li, Sheng Qiang, Xianyi Zhuang, and Okyay Kaynak, *Fellow, IEEE*

***Abstract*—In this paper, two different backstepping neural network (NN) control approaches are presented for a class of affine nonlinear systems in the strict-feedback form with unknown nonlinearities. By a special design scheme, the controller singularity problem is avoided perfectly in both approaches. Furthermore, the closed loop signals are guaranteed to be semiglobally uniformly ultimately bounded and the outputs of the system are proved to converge to a small neighborhood of the desired trajectory. The control performances of the closed-loop systems can be shaped as desired by suitably choosing the design parameters. Simulation results obtained demonstrate the effectiveness of the approaches proposed. The differences observed between the inputs of the two controllers are analyzed briefly.**

***Index Terms*—Adaptive control, backstepping, neural network (NN), robust adaptive control, uncertain strict-feedback system.**

I. INTRODUCTION

IN recent adaptive and robust control literature, numerous approaches have been proposed for the design of nonlinear control systems. Among these, adaptive backstepping constitutes a major design methodology [1]–[3]. The idea behind backstepping design is that some appropriate functions of state variables are selected recursively as pseudocontrol inputs for lower dimension subsystems of the overall system. Each backstepping stage results in a new pseudocontrol design, expressed in terms of the pseudocontrol designs from the preceding design stages. When the procedure terminates, a feedback design for the true control input results, which achieves the original design objective by virtue of a final Lyapunov function, formed by summing the Lyapunov functions associated with each individual design stage [1].

The backstepping design provides a systematic framework for the design of tracking and regulation strategies, suitable for a large class of state feedback linearizable nonlinear systems. Integrator backstepping is used to systematically design controllers for systems with known nonlinearities with mismatched conditions [2]. The approach can be extended to handle systems with unknown parameters via adaptive backstepping [2], [4], [5]. Apart from the systematic approach used, another important feature of it is its ability to shape performance [1], [5].

Manuscript received September 10, 2002; revised December 23, 2003. This work was supported by the Youth Scientific Foundation of Heilongjiang Province, China, by Project QC02C40.

Y. Li, S. Qiang, and X. Zhuang are with Department of Control Science and Engineering, Harbin Institute of Technology, Harbin 150001, China (e-mail: lyahui2002@yahoo.com.cn; zhx@hit.edu.cn).

O. Kaynak is with Department of Electrical and Electronic Engineering, Bogazici University, 80815 Bebek-Istanbul, Turkey (e-mail: kaynak@boun.edu.tr).

Digital Object Identifier 10.1109/TNN.2004.826215

However, in spite of the merits explained, there are some problems in the backstepping design method. A major constraint is that certain functions must be "linear in the unknown parameters" [6], [7], which may not be satisfied in practise. Furthermore, some very tedious analysis is needed to determine "regression matrices"[6]. In the case of backstepping adaptive control, the problem of determining and computing the regression matrices becomes even more acute. In [8], for example, for the relatively simple application to dc motor control, one will notice that the regression matrix almost covers one full page in the IEEE TRANSACTIONS ON CONTROL SYSTEM TECHNOLOGY.

The very rapid developments described in adaptive and robust control techniques are accompanied by an increasing in the use of neural networks (NNs) for system identification [9], [10] or identification-based control [11]–[13]. With the help of neural networks, the linearity-in-the-parameter assumption of nonlinear function and the determination of regression matrices can be avoided. It is due to this that in the last few years, a large number of backstepping design schemes are reported that combine the backstepping technique with adaptive NNs [6], [7], [14]–[20].

Although significant progress has been made by combining backstepping methodology with NN technologies, there are still some problems that need to be solved for practical implementations. For example. in almost all the approaches reported in the literature, in order to avoid the controller singularity problem it is assumed that the gain functions $g_i(\overline{x}_i)(i = 1, 2, \ldots\ldots, n)$ (see (1) in Section II) are constants [6], [7], [14]–[17] or known functions [6]. However, this assumption cannot be satisfied in many cases. In some work, for example in [18], the authors assume the gain functions to be unknown and adopt NN structures to approximate them. In order to avoid the possible divergence of the weights of the neural networks during on-line tuning, the discontinuous projections with fictitious bounds have to be used in design. In [19], gain functions are assumed to be unknown and a backstepping design is proposed that incorporates adaptive neural network techniques. However, due to the integral-type Lyapunov function introduced, this approach is complicated and difficult to use in practice. In [20], Ge *et al.* propose another simpler scheme under the same conditions. Nevertheless, because the derivatives of the virtual controllers are included in NNs, the NNs are difficult to realize and calculate.

Variable structure control is one of the main methods used in the literature to overcome the uncertainty of systems. In backstepping design, this approach is difficult to use, because the derivatives of virtual controllers are all included in the actual controller. Therefore, robustness issues are rarely addressed in backstepping design. In this paper, an alternative, simpler design is proposed, which can completely avoid the singularity problem. Based on this, a class of continuous functions is introduced to ensure robustness of in backstepping design.

The paper is organized as follows. In Section II, a description of the system is given and radial basis function (RBF) neural networks are briefly explained. Adaptive backstepping design is described in Section III and simulated on a nonlinear system. Section IV proposes an approach to ensure a degree of robustness in the controller and the effectiveness of the approach is demonstrated by the simulation studies on the same system as used in Section III. Finally, conclusions are given in Section V.

II. Problem Formulation

A. System Description

The model of many practical nonlinear systems can be expressed in or transformed into a special state-space form

$$\begin{aligned} \dot{x}_i &= f_i(\overline{x}_i) + g_i(\overline{x}_i)x_{i+1}, \quad 1 \le i \le n-1 \\ \dot{x}_n &= f_n(\overline{x}_n) + g_n(\overline{x}_n)u, \quad n \ge 2 \\ y &= x_1 \end{aligned} \tag{1}$$

where $\overline{x}_i[x_1, x_2, \cdots, x_i]^T \in R^i$, $i = 1, \ldots, n$, $u \in R$, $y \in R$ are state variables, system input and output, respectively. The control objective is to design an adaptive NN controller for system (1) such that 1) all the signals in the closed-loop remain semiglobally uniformly ultimately bounded and 2) the output y follows a desired trajectory y_d, which and whose derivatives up to the $(m+1)$th order are bounded.

Note that in the following derivation of the adaptive neural controller, NN approximation is only guaranteed with some compact sets. Accordingly, the stability results obtained in this work are semiglobal in the sense that, as long as desired, there exists controllers with sufficient large number of NN nodes such that all the signals in the closed-loop remain bounded.

Since $g_i(\cdot), i = 1, \ldots, n$ are smooth functions, they are therefore bounded within some compact set. Accordingly, we can make the following two assumptions as commonly being done in the literature.

Assumption 1: The signs of $g_i(\cdot)$ are bounded, i.e., there exist constants $g_{i1} \ge g_{i0} > 0$ such that ,$g_{i1} \ge |g_i(\cdot)| \ge g_{i0}$, $\forall \overline{x}_n \in \Omega \in R^n$.

The above assumption implies that the smooth functions $g_i(\cdot)$ are strictly either positive or negative. Without losing generality, we shall assume, $g_{i1} \ge g_i(\cdot) \ge g_{i0} > 0$, $\forall \overline{x}_n \in \Omega \in R^n$.

Assumption 2: There exist constants $g_{id} > 0$ such that, $|\dot{g}_i(\cdot)| \le g_{id}$, $\forall \overline{x}_n \in \Omega \in R^n$.

B. Function Approximation Using RBF Neural Networks

The control design presented in this paper employs RBF NNs to approximate the nonlinear functions in system (1). They are of the general form $\hat{F}(\cdot) = \theta^T \xi(\cdot)$, where $\theta \in R^p$ is a vector of adjustable weights and $\xi(\cdot)$ a vector of RBF's. Their ability to uniformly approximate smooth functions over compact sets is well documented in the literature (for example [21]). In general, it has been shown that given a smooth function $F : \Omega \to R$, where Ω is a compact subset of R^m (m is an appropriate integer) and $\varepsilon > 0$, there exists a RBF vector $\xi : R^m \to R^p$ and a weight vector $\theta^* \in R^p$ such that $|F(x) - \theta^{*T}\xi(x)| \le \varepsilon$, $\forall x \in \Omega$. The quantity $F(x) - \theta^{*T}\xi(x) \doteq d_F(x)$ is called the network reconstruction error and obviously $|d_F(x)| \le \varepsilon$.

The optimal weight vector θ^* defined above is a quantity only for analytical purposes. Typically θ^* is chosen as the value of θ that minimizes $d_F(x)$ over Ω, that is

$$\theta^* = \arg \min_{\theta \in R^p} \left\{ \sup_{x \in \Omega} \left| F(x) - \theta^T \xi(x) \right| \right\}. \tag{2}$$

The Gaussian functions are employed as basis functions, in the same form as in [20], which are located on a regular grid that contains the subset of interest of the state space.

III. Adaptive NN Control

A. Controller Design

The detailed design procedure is described in the following steps. For clarity and conciseness, Steps 1 and 2 are described with detailed explanations, while Step i and Step n are simplified, with the relevant equations and the explanations being omitted.

Step 1: Let $x_{1d} = y_d$ and define $e_1 = x_1 - x_{1d}$. Its derivative is

$$\dot{e}_1 = f_1(x_1) + g_1(x_1)x_2 - \dot{x}_{1d} \tag{3}$$

by viewing x_2 as a virtual control input. Equation (3) can be transformed into the following form:

$$\dot{e}_1 = g_1(x_1)\left[g_1^{-1}(x_1)f_1(x_1) + x_2 - g_1^{-1}(x_1)\dot{x}_{1d}\right]. \tag{4}$$

Let us choose controller x_{2d} as follows:

$$x_{2d} = x_2 = -g_1^{-1}(x_1)f_1(x_1) + g_1^{-1}(x_1)\dot{x}_{1d} - k_1 e_1 \tag{5}$$

where $k_1 > 0$ is constant. Substituting (5) into (4), $\dot{e}_1 = -g_1(x_1)k_1e_1$ is obtained. So, there exists a Lyapunov function $V_1 = (1/2)e_1^2$, such that $\dot{V}_1 = -g_1(x_1)k_1e_1^2 \le -g_{10}k_1e_1^2 \le 0$. Therefore, e_1 is asymptotically stable.

However, since the functions $f_1(x_1)$ and $g_1(x_1)$ are unknown, the desired controller cannot be implemented in practice. Instead, a NN-based virtual controller can be used as follows:

$$x_{2d} = -\theta_1^T \xi_1(x_1) + \delta_1^T \eta_1(x_1)\dot{x}_{1d} - k_1 e_1 \tag{6}$$

where $\theta_1^T \xi_1(x_1)$ and $\delta_1^T \eta_1(x_1)$ are RBF NNs used to approximate $g_1^{-1}(x_1)f_1(x_1)$ and $g_1^{-1}(x_1)$, respectively.

Defining $e_2 = x_2 - x_{2d}$, $\dot{e}_1$ can be obtained as

$$\begin{aligned} \dot{e}_1 &= f_1(x_1) + g_1(x_1)(e_2 + x_{2d}) - \dot{x}_{1d} \\ &= g_1(x_1)\left[g_1^{-1}(x_1)f_1(x_1) + e_2 + x_{2d} - g_1^{-1}(x_1)\dot{x}_{1d}\right] \\ &= g_1(x_1)\left[\theta_1^{*T}\xi_1(x_1) - \delta_1^{*T}\eta_1(x_1)\dot{x}_{1d} + e_2 + x_{2d} + d_1\right] \end{aligned} \tag{7}$$

where θ_1^* and δ_1^* are the optimal weight vectors of $g_1^{-1}f_1$ and g_1^{-1}. The neural reconstruction error

$$d_1 = \left(g_1^{-1}f_1 - \theta_1^{*T}\xi_1(x_1)\right) + \left(g_1^{-1} - \delta_1^{*T}\eta_1(x_1)\right)\dot{x}_{1d}$$

is bounded, i.e., there exists a constant $\varepsilon_1 > 0$ such that $|d_1| < \varepsilon_1$. Throughout the paper, we introduce $\theta_i^T \xi_i(\cdot)$ and $\delta_i^T \eta_i(\cdot)$ as neural networks and define their reconstruction errors as

$$d_i = \left(g_i^{-1} f_i - \theta_i^{*T} \xi_i(\overline{x}_i)\right) + \left(g_i^{-1} - \delta_i^{*T} \eta_i(\overline{x}_i)\right) \dot{x}_{id}$$

where $i = 1, \ldots, n$. Like in the case of d_1, d_i is bounded, i.e., $d_i| < \varepsilon_i$.

Substituting (6) into (7), we get

$$\dot{e}_1 = g_1(x_1)\left[\widetilde{\theta}_1^T \xi_1(x_1) - \widetilde{\delta}_1^T \eta_1(x_1)\dot{x}_{1d} - k_1 e_1 + e_2 + d_1\right] \tag{8}$$

where $\widetilde{\theta}_1 = \theta_1^* - \theta_1$, $\widetilde{\delta}_1 = \delta_1^* - \delta_1$. Through out this paper, we shall define $\widetilde{(\cdot)} = (\cdot)^* - (\cdot)$.

Consider the following Lyapunov candidate:

$$V_1 = \frac{1}{2g_1(x_1)} e_1^2 + \frac{1}{2}\widetilde{\theta}_1^T \Gamma_{11}^{-1} \widetilde{\theta}_1 + \frac{1}{2}\widetilde{\delta}_1^T \Gamma_{12}^{-1} \widetilde{\delta}_1 \tag{9}$$

where $\Gamma_{11} = \Gamma_{11}^T > 0$, $\Gamma_{12} = \Gamma_{12}^T > 0$ are adaptive gain matrices.

The derivative of V_1 is

$$\begin{aligned}\dot{V}_1 &= \frac{e_1 \dot{e}_1}{g_1(x_1)} - \frac{\dot{g}_1(x_1)}{2g_1^2(x_1)} e_1^2 - \widetilde{\theta}_1^T \Gamma_{11}^{-1} \dot{\theta}_1 - \widetilde{\delta}_1^T \Gamma_{12}^{-1} \dot{\delta}_1 \\ &= e_1 \widetilde{\theta}_1^T \xi_1(x_1) - e_1 \widetilde{\delta}_1^T \eta_1(x_1)\dot{x}_{1d} - \left(k_1 + \frac{\dot{g}_1}{2g_1^2}\right) e_1^2 \\ &\quad + e_1 e_2 + e_1 d_1 - \widetilde{\theta}_1^T \Gamma_{11}^{-1} \dot{\theta}_1 - \widetilde{\delta}_1^T \Gamma_{12}^{-1} \dot{\delta}_1 \\ &= \widetilde{\theta}_1^T \left[e_1 \xi_1(x_1) - \Gamma_{11}^{-1} \dot{\theta}_1\right] - \widetilde{\delta}_1^T \left[e_1 \eta_1(x_1)\dot{x}_{1d} + \Gamma_{12}^{-1} \dot{\delta}_1\right] \\ &\quad - \left(k_1 + \frac{\dot{g}_1}{2g_1^2}\right) e_1^2 + e_1 e_2 + e_1 d_1. \end{aligned} \tag{10}$$

Consider the following adaptation laws:

$$\begin{aligned}\dot{\theta}_1 &= \Gamma_{11}\left[e_1 \xi_1(x_1) - \sigma_1 \theta_1\right] \\ \dot{\delta}_1 &= \Gamma_{12}\left[-e_1 \xi_1(x_1)\dot{x}_{1d} - \gamma_1 \delta_1\right]\end{aligned} \tag{11}$$

where $\sigma_1 > 0$, $\gamma_1 > 0$ are small constants. Formulas (11) are so-called σ-modification, introduced to improve the robustness in the presence of the approximation error d_1 and avoid the weight parameters to drift to very large values.

Let $k_1 = k_{10} + k_{11}$, with k_{10} and $k_{11} > 0$. Then, (10) becomes

$$\dot{V}_1 = e_1 e_2 - \left(k_{10} + \frac{\dot{g}_1}{2g_1^2}\right) e_1^2 + \sigma_1 \widetilde{\theta}_1^T \theta_1 + \gamma_1 \widetilde{\delta}_1^T \delta_1 - k_{11} e_1^2 + e_1 d_1. \tag{12}$$

By completion of squares, we have

$$\begin{aligned}\sigma_1 \widetilde{\theta}_1^T \theta_1 &= \sigma_1 \widetilde{\theta}_1^T \left(\theta_1^* - \widetilde{\theta}_1\right) \\ &\le \sigma_1 \|\widetilde{\theta}_1\| \, \|\theta_1^*\| - \sigma_1 \|\widetilde{\theta}_1\|^2 \\ &\le \frac{\sigma_1 \|\theta_1^*\|^2}{2} - \frac{\sigma_1 \|\widetilde{\theta}_1\|^2}{2}\end{aligned} \tag{13}$$

$$\begin{aligned}\gamma_1 \widetilde{\delta}_1^T \delta_1 &= \gamma_1 \widetilde{\delta}_1^T \left(\delta_1^* - \widetilde{\delta}_1\right) \\ &\le \gamma_1 \|\widetilde{\delta}_1\| \, \|\delta_1^*\| - \gamma_1 \|\widetilde{\delta}_1\|^2 \\ &\le \frac{\gamma_1 \|\delta_1^*\|^2}{2} - \frac{\gamma_1 \|\widetilde{\delta}_1\|^2}{2}\end{aligned} \tag{14}$$

$$\begin{aligned}-k_{11} e_1^2 + e_1 d_1 &\le -k_{11} e_1^2 + e_1 |d_1| \\ &\le \frac{d_1^2}{4k_{11}} \\ &\le \frac{\varepsilon_1^2}{4k_{11}}.\end{aligned} \tag{15}$$

Because $-(k_{10} + (\dot{g}_1/2g_1^2))e_1^2 \le -(k_{10} - (g_{1d}/2g_1^2))e_1^2$, by choosing k_{10} such that $(k_{10}^* \doteq k_{10} - (g_{1d}/2g_1^2)) > 0$, we have the following inequality:

$$\begin{aligned}\dot{V}_1 \le e_1 e_2 - k_{10}^* e_1^2 &+ \frac{\sigma_1 \|\theta_1^*\|^2}{2} - \frac{\sigma_1 \|\widetilde{\theta}_1\|^2}{2} \\ &+ \frac{\gamma_1 \|\delta_1^*\|^2}{2} - \frac{\gamma_1 \|\widetilde{\delta}_1\|^2}{2} + \frac{\varepsilon_1^2}{4k_{11}}\end{aligned} \tag{16}$$

where the coupling term $e_1 e_2$ will be canceled in the next step.

Step 2: This step is to make the error between x_2 and x_{2d} as small as possible. Differentiating e_2 gives

$$\begin{aligned}\dot{e}_2 &= \dot{x}_2 - \dot{x}_{2d} = f_2(\overline{x}_2) + g_2(\overline{x}_2)x_3 - \dot{x}_{2d} \\ &= g_2(\overline{x}_2)\left[g_2^{-1}(\overline{x}_2) f_2(\overline{x}_2) + x_3 - g_2^{-1}(\overline{x}_2)\dot{x}_{2d}\right] \\ &= g_2\left[\theta_2^{*T} \xi_2(\overline{x}_2) - \delta_2^{*T} \eta_2(\overline{x}_2)\dot{x}_{2d} + e_3 + x_{3d} + d_2\right].\end{aligned} \tag{17}$$

Similarly, taking the virtual controller x_{3d} as of the form

$$x_{3d} = -e_1 - \theta_2^T \xi_2(\overline{x}_2) + \delta_2^T \eta_2(\overline{x}_2)\dot{x}_{2d} - k_2 e_2 \tag{18}$$

and substituting it into (17), we will have

$$\dot{e}_2 = g_2(\overline{x}_2)\left[\widetilde{\theta}_2^T \xi_2(\overline{x}_2) - \widetilde{\delta}_2^T \eta_2(\overline{x}_2)\dot{x}_{2d} - e_1 - k_2 e_2 + e_3 + d_2\right]. \tag{19}$$

Consider the Lyapunov function candidate

$$V_2 = V_1 + \frac{1}{2g_2(\overline{x}_2)} e_2^2 + \frac{1}{2}\widetilde{\theta}_2^T \Gamma_{21}^{-1} \widetilde{\theta}_2 + \frac{1}{2}\widetilde{\delta}_2^T \Gamma_{22}^{-1} \widetilde{\delta}_2 \tag{20}$$

where $\Gamma_{11} = \Gamma_{11}^{\mathrm{T}} > 0$, $\Gamma_{12} = \Gamma_{12}^{\mathrm{T}} > 0$ are adaptive gain matrices.

The derivative of V_2 is

$$\begin{aligned}\dot{V}_2 &= \dot{V}_1 + \frac{e_2 \dot{e}_2}{g_2(\overline{x}_2)} - \frac{\dot{g}_2(\overline{x}_2)}{2g_2^2(\overline{x}_2)} e_2^2 - \widetilde{\theta}_2^T \Gamma_{21}^{-1} \dot{\theta}_2 - \widetilde{\delta}_2^T \Gamma_{22}^{-1} \dot{\delta}_2 \\ &= \dot{V}_1 + e_2 \widetilde{\theta}_2^T \xi_2(\overline{x}_2) - e_2 \widetilde{\delta}_2^T \eta_2(\overline{x}_2)\dot{x}_{2d} - \left(k_2 + \frac{\dot{g}_2}{2g_2^2}\right) e_2^2 \\ &\quad - e_1 e_2 + e_2 e_3 + e_2 d_2 - \widetilde{\theta}_2^T \Gamma_{21}^{-1} \dot{\theta}_2 - \widetilde{\delta}_2^T \Gamma_{22}^{-1} \dot{\delta}_2 \\ &= \widetilde{\theta}_2^T \left[e_2 \xi_2(\overline{x}_2) - \Gamma_{21}^{-1} \dot{\theta}_2\right] - \widetilde{\delta}_2^T \left[e_2 \eta_2(\overline{x}_2)\dot{x}_{2d} + \Gamma_{22}^{-1} \dot{\delta}_2\right] \\ &\quad - \left(k_2 + \frac{\dot{g}_2}{2g_2^2}\right) e_2^2 - e_1 e_2 + e_2 e_3 + e_2 d_2 + \dot{V}_1\end{aligned} \tag{21}$$

where $e_3 = x_3 - x_{3d}$.

Consider the following adaptation laws:

$$\begin{aligned}\dot{\theta}_2 &= \Gamma_{21}\left[e_2 \xi_2(\overline{x}_2) - \sigma_2 \theta_2\right] \\ \dot{\delta}_2 &= \Gamma_{22}\left[-e_2 \eta_2(\overline{x}_2)\dot{x}_{2d} - \gamma_2 \delta_2\right]\end{aligned} \tag{22}$$

where $\sigma_2 > 0$ and $\gamma_2 > 0$ are small constants. Let $k_2 = k_{20} + k_{21}$, where k_{20} and $k_{21} > 0$. Substituting (16) and (22) into (21), and with some completion of squares and straightforward derivation similar to those employed in Step 1, the derivative of V_2 is obtained as

$$\dot{V}_2 \le e_2 e_3 - \sum_{l=1}^{2} k_{l0}^* e_l^2 - \sum_{l=1}^{2}\left(\frac{\sigma_l\|\widetilde{\theta}_l\|^2}{2} + \frac{\gamma_l\|\widetilde{\delta}_l\|^2}{2}\right) + \sum_{l=1}^{2}\left(\frac{\sigma_l\|\theta_l^*\|^2}{2} + \frac{\gamma_l\|\delta_l^*\|^2}{2}\right) + \sum_{l=1}^{2}\frac{\varepsilon_l^2}{4k_{l1}} \quad (23)$$

where k_{20} is chosen such that

$$\left(k_{20}^* \doteq k_{20} - \frac{g_{2d}}{2g_{20}^2}\right) > 0.$$

Step $i(3 \le i \le n-1)$: In a similar fashion, we can design a virtual controller $x_{(i+1)d}$ to make the error $e_i = x_i - x_{id}$ as small as possible. Differentiating e_i gives

$$\begin{aligned}\dot{e}_i &= g_i(\overline{x}_i)\left[g_i^{-1}(\overline{x}_i)f_i(\overline{x}_i) + x_{i+1} - g_i^{-1}(\overline{x}_i)\dot{x}_{id}\right]\\ &= g_i\left[\theta_i^{*T}\xi_i(\overline{x}_i) - \delta_i^{*T}\eta_i(\overline{x}_i)\dot{x}_{id} + e_{i+1} + x_{(i+1)d} + d_i\right] \quad (24)\end{aligned}$$

where $e_{i+1} = x_{(i+1)} - x_{(i+1)d}$.

Similarly, let the virtual controller to be of the form

$$x_{(i+1)d} = -e_{i-1} - \theta_i^T\xi_i(\overline{x}_i) + \delta_i^T\eta_i(\overline{x}_i)\dot{x}_{id} - k_i e_i. \quad (25)$$

Then we have

$$\dot{e}_i = g_i(\overline{x}_i)\left[\widetilde{\theta}_i^T\xi_i(\overline{x}_i) - \widetilde{\delta}_i^T\eta_i(\overline{x}_i)\dot{x}_{id} - e_{i-1} - k_i e_i + e_{i+1} + d_i\right]. \quad (26)$$

Consider the Lyapunov function candidate

$$V_i = V_{i-1} + \frac{1}{2g_i(\overline{x}_i)}e_i^2 + \frac{1}{2}\widetilde{\theta}_i^T\Gamma_{i1}^{-1}\widetilde{\theta}_i + \frac{1}{2}\widetilde{\delta}_i^T\Gamma_{i2}^{-1}\widetilde{\delta}_i. \quad (27)$$

and consider the following adaptation laws:

$$\begin{aligned}\dot{\theta}_i &= \Gamma_{i1}\left[e_i\xi_i(\overline{x}_i) - \sigma_i\theta_i\right]\\ \dot{\delta}_i &= \Gamma_{i2}\left[-e_i\eta_i(\overline{x}_i)\dot{x}_{id} - \gamma_i\delta_i\right]. \quad (28)\end{aligned}$$

where $\sigma_i > 0$ and $\gamma_i > 0$ are small constants. Let $k_i = k_{i0} + k_{i1}$, where k_{i0} and $k_{i1} > 0$. By using (23), (26), and (28), and with some completion of squares and straightforward derivation similar to those employed in the former steps, the derivative of V_i becomes

$$\dot{V}_i \le e_i e_{i+1} - \sum_{l=1}^{i} k_{l0}^* e_l^2 - \sum_{l=1}^{i}\left(\frac{\sigma_l\|\widetilde{\theta}_l\|^2}{2} + \frac{\gamma_l\|\widetilde{\delta}_l\|^2}{2}\right) + \sum_{l=1}^{i}\left(\frac{\sigma_l\|\theta_l^*\|^2}{2} + \frac{\gamma_l\|\delta_l^*\|^2}{2}\right) + \sum_{l=1}^{i}\frac{\varepsilon_l^2}{4k_{l1}} \quad (29)$$

where k_{i0} is chosen such that

$$\left(k_{i0}^* \doteq k_{i0} - \frac{g_{id}}{2g_{i0}^2}\right) > 0.$$

Step n: This is the final step. Differentiating the error $e_n = x_n - x_{nd}$, we will have

$$\begin{aligned}\dot{e}_n &= g_n(\overline{x}_n)\left[g_n^{-1}(\overline{x}_n)f_n(\overline{x}_n) + u - g_n^{-1}(\overline{x}_n)\dot{x}_{nd}\right]\\ &= g_n\left[\theta_n^{*T}\xi_n(\overline{x}_n) - \delta_n^{*T}\eta_n(\overline{x}_n)\dot{x}_{nd} + u + d_n\right]. \quad (30)\end{aligned}$$

Similarly, letting

$$u = -e_{n-1} - \theta_n^T\xi_n(\overline{x}_n) + \delta_n^T\eta_n(\overline{x}_n)\dot{x}_{nd} - k_n e_n \quad (31)$$

and substituting it into (30), gives

$$\dot{e}_n = g_n(\overline{x}_n)\left[\widetilde{\theta}_n^T\xi_n(\overline{x}_n) - \widetilde{\delta}_n^T\eta_n(\overline{x}_n)\dot{x}_{nd} - e_{n-1} - k_n e_n + d_n\right]. \quad (32)$$

Consider the overall Lyapunov function candidate

$$V_n = V_{n-1} + \frac{1}{2g_n(\overline{x}_n)}e_n^2 + \frac{1}{2}\widetilde{\theta}_n^T\Gamma_{n1}^{-1}\widetilde{\theta}_n + \frac{1}{2}\widetilde{\delta}_n^T\Gamma_{n2}^{-1}\widetilde{\delta}_n. \quad (33)$$

And consider the following adaptation law:

$$\begin{aligned}\dot{\theta}_n &= \Gamma_{n1}\left[e_n\xi_n(\overline{x}_n) - \sigma_n\theta_n\right]\\ \dot{\delta}_n &= \Gamma_{n2}\left[-e_n\eta_n(\overline{x}_n)\dot{x}_{nd} - \gamma_n\delta_n\right] \quad (34)\end{aligned}$$

where $\sigma_n > 0$ and $\gamma_n > 0$ are small constants, $\Gamma_{n1} = \Gamma_{n1}^{\mathrm{T}} > 0$, $\Gamma_{n2} = \Gamma_{n2}^{\mathrm{T}} > 0$ are gain matrices. Let $k_n = k_{n0} + k_{n1}$, where k_{n0} and $k_{n1} > 0$. By using (29), (32) and (34), and with some completion of squares and straightforward derivation similar to those employed in the former steps, the derivative of V_n becomes

$$\dot{V}_n \le -\sum_{l=1}^{n} k_{l0}^* e_l^2 - \sum_{l=1}^{n}\left(\frac{\sigma_l\|\widetilde{\theta}_l\|^2}{2} + \frac{\gamma_l\|\widetilde{\delta}_l\|^2}{2}\right) + \sum_{l=1}^{n}\left(\frac{\sigma_l\|\theta_l^*\|^2}{2} + \frac{\gamma_l\|\delta_l^*\|^2}{2}\right) + \sum_{l=1}^{n}\frac{\varepsilon_l^2}{4k_{l1}} \quad (35)$$

where k_{n0} is chosen such that

$$\left(k_{n0}^* \doteq k_{n0} - \frac{g_{nd}}{2g_{n0}^2}\right) > 0.$$

Let $\phi \doteq \sum_{l=1}^{n}(\sigma_l\|\theta_l^*\|^2/2 + \gamma_l\|\delta_l^*\|^2/2 + \varepsilon_l^2/4c_{l1})$. If we choose k_{l0}^* such that $k_{l0}^* \ge (\mu/2g_{l0})$, i.e., choose k_{l0} such that $k_{l0}^* \ge (\mu/2g_{l0}) + (g_{ld}/2g_{l0}^2)$, where μ is a positive constant, and choose σ_l, γ_l, Γ_{l1} and Γ_{l2} such that

$$\sigma_l \ge \mu\lambda_{\max}\left\{\Gamma_{l1}^{-1}\right\}, \quad \gamma_l \ge \mu\lambda_{\max}\left\{\Gamma_{l2}^{-1}\right\}$$

where $\lambda_{\max}\{\cdot\}$ is the largest eigenvalue of matrices, and $k = 1, \ldots, n$. Then, from (35), we have the following inequality:

$$\begin{aligned}\dot{V}_n &\le -\sum_{l=1}^{n} k_{l0}^* e_l^2 - \sum_{l=1}^{n}\left(\frac{\sigma_l\|\widetilde{\theta}_l\|^2}{2} + \frac{\gamma_l\|\widetilde{\delta}_l\|^2}{2}\right) + \phi \quad (36)\\ &\le -\sum_{l=1}^{n}\frac{\mu}{2g_{k0}}e_l^2 - \mu\sum_{l=1}^{n}\left(\frac{\widetilde{\theta}_k^T\Gamma_{k1}^{-1}\widetilde{\theta}_k}{2} + \frac{\widetilde{\delta}_k^T\Gamma_{k1}^{-1}\widetilde{\delta}_k}{2}\right) + \phi\\ &\le -\mu\left[\sum_{l=1}^{n}\frac{1}{2g_{k0}}e_l^2 + \sum_{l=1}^{n}\left(\frac{\widetilde{\theta}_k^T\Gamma_{k1}^{-1}\widetilde{\theta}_k}{2} + \frac{\widetilde{\delta}_k^T\Gamma_{k1}^{-1}\widetilde{\delta}_k}{2}\right)\right] + \phi\\ &\le \mu V_n + \phi. \quad (37)\end{aligned}$$

The following theorem shows the stability and control performance of the closed-loop adaptive system.

Theorem 1: Consider the closed-loop system consisting of (1) and the known reference signal, the controller (31), and the NN weight updating laws (11), (22), (28) and (34). Assume that there exists sufficiently large compact sets $\Omega_{\theta 1}$, $\Omega_{\delta 1}$ and Ω_1 with proper dimensions, such that $\theta \in \Omega_{\theta 1}$, $\delta \in \Omega_{\delta 1}$ and $e \in \Omega_1$

for all $t \geq 0$. Then for bounded initial conditions, we have the following:

1) all signals in the closed-loop system remain bounded;
2) the output tracking error $e = y(t) - y_d(t)$ converges to a small neighborhood around zero by an appropriate choice of the design parameters.

Proof:

1) From (36), we have

$$\dot{V}_n \leq -k^*_{\min}\|e\|^2 - \frac{\left(\sigma_{\min}\|\widetilde{\theta}\|^2 + \gamma_{\min}\|\widetilde{\delta}\|^2\right)}{2} + \phi \tag{38}$$

where

$$\begin{aligned} e &= [e_1 \quad e_2 \quad \cdots \quad e_n]^T \\ \widetilde{\theta} &= \left[\widetilde{\theta}_1^T \quad \widetilde{\theta}_2^T \quad \cdots \quad \widetilde{\theta}_n^T\right]^T \\ \widetilde{\delta} &= \left[\widetilde{\delta}_1^T \quad \widetilde{\delta}_2^T \quad \cdots \quad \widetilde{\delta}_n^T\right]^T \end{aligned} \tag{39}$$

and $(\cdot)_{\min}$ is the minimum of $(\cdot)_i, i = 1, 2, \ldots, n$. Therefore, the derivative of global Lyapunov function is negative as long as

$$e \notin \Omega = \left\{ e \left| \|e\| \leq \sqrt{\frac{\phi}{k^*_{\min}}} \right. \right\} \tag{40}$$

or

$$\widetilde{\theta} \notin \Omega_\theta = \left\{ \widetilde{\theta} \left| \|\widetilde{\theta}\| \leq \sqrt{\frac{2\phi}{\sigma_{\min}}} \right. \right\} \tag{41}$$

or

$$\widetilde{\delta} \notin \Omega_\delta = \left\{ \widetilde{\delta} \left| \|\widetilde{\delta}\| \leq \sqrt{\frac{2\phi}{\gamma_{\min}}} \right. \right\}. \tag{42}$$

According to a standard Lyapunov theorem extension [22], these demonstrate the uniformly ultimately boundedness (UUB) of e, $\widetilde{\theta}$ and $\widetilde{\delta}$. Since $e_1 = x_1 - x_{1d}$, and x_{1d} are bounded, we have that x_1 is bounded. From $e_i = x_i - x_{id}$, $i = 2, 3, \ldots, n$, and the definitions of virtual controls x_{id} (6), (18), and (25), we have that x_{id} remain bounded. Using (31), we conclude that control u is also bounded. According to assumption of the system that the function $f(x)$ and $g(x)$ are continuous, the function are bounded in any certainty compact. Thus, the optimal weights θ_i^* and δ_i^*, $i = 1, 2, \ldots, n$, are bounded, and the weights θ_i and δ_i of the NNs are bounded because of (40) and (41). So, all the signals in the closed-loop system remain bounded.

2) Let $\rho = \phi/\mu > 0$, then (37) satisfies

$$0 \leq V_n(t) \leq \rho + (V_n(0) - \rho)\exp(-\mu t) \tag{43}$$

From (43), we have

$$\sum_{i=1}^{n} \frac{1}{2g_k} e_i^2 < \rho + (V_n(0) - \rho)\exp(-\mu t) < \rho + V_n(0)\exp(-\mu t). \tag{44}$$

Let $g_{\max} = \max_{1 \leq i \leq n}\{g_{i1}\}$. Then, we have

$$\frac{1}{2g_{\max}} \sum_{i=1}^{n} e_i^2 \leq \sum_{i=1}^{n} \frac{1}{2g_k} e_i^2 < \rho + V_n(0)\exp(-\mu t) \tag{45}$$

that is

$$\sum_{i=1}^{n} e_i^2 < 2g_{\max}\rho + 2g_{\max}V_n(0)\exp(-\mu t) \tag{46}$$

which implies that given $\lambda > \sqrt{2g_{\max}\rho}$, there exists T such that for all $t \geq T$, the tracking error satisfies

$$|e_1| = |x_1(t) - x_{1d}(t)| = |y(t) - y_d(t)| < \lambda \tag{47}$$

where λ is the size of a small residual set which depends on the NNs approximation error d_i and controller parameters k_i, σ_i, γ_i, and Γ_i. It is easily seen that the increase in the control gain k_i and adaptive gain Γ_i, and NN node numbers l_j will result in a better tracking performance.

Remark 1: In [20], one NN is adopted to approximate the nonlinear function $(f_i(x_i) - \dot{x}_{id})/g_i(x_i)$ in every design step. However, because the derivatives of the virtual control x_{id} are included in the NNs, the dimensions of the input vectors of the NNs become twice as much as those of the corresponding state vectors and these additional inputs must be computed online too. Therefore, the approach is still difficult to implement and apply in practice. In this paper, although two NNs are adopted to approximate the nonlinear functions $g_i^{-1}(x)f_i(x)$ and $g_i^{-1}(x)$, respectively, in every "step", there are no dimensional increments and no additional parameters must be calculated. Compared with the approach in [20], the method presented in this paper is much simpler to understand and apply in practice. For example, we assume that the $\overline{x}_i$ is two-dimensional (2-D). So, the input vector of nonlinear function $(f_i(\overline{x}_i) - \dot{x}_{id})/g_i(\overline{x}_i)$ in literature [20] is 3-D at least, while those of nonlinear functions $g_i^{-1}(\overline{x}_i)f_i(\overline{x}_i)$ and $g_i^{-1}(\overline{x}_i)$ are still 2-D. Therefore, if we are given five fuzzy sets for every term of the input vectors, the node number will be 125 (5^3) although just one neural network is adopted to approximate the function $(f_i(\overline{x}_i) - \dot{x}_{id})/g_i(\overline{x}_i)$. However, each of the neural networks adopted in this paper is 25 (5^2) node one, although two networks must be adopted to approximate function $g_i^{-1}(\overline{x}_i)f_i(\overline{x}_i)$ and $g_i^{-1}(\overline{x}_i)$, respectively. So, we can draw the conclusion that the controller is simpler to for practical realizations and computational burden is not excessive.

Remark 2: In the above analysis, it is clear that the uniform ultimate boundedness of all the signals are guaranteed by choosing $k_i = k_{i0} + k_{i1}$ large enough, such that $k_{i0}^* = k_{i0} - (g_{id}/2g_{id}^2) > 0$. Moreover, it can be seen that 1) increasing k_{i0} might lead to larger μ, and increasing k_{i1} will reduce δ, thus, increasing k_i will lead to smaller Ω and 2) decreasing σ_i and γ_i will help to reduce ε_i, both of which will help to reduce the size of Ω. However, increasing k_i will lead to a high gain control scheme. On the other hand, though σ_i and γ_i is required to be chosen as a small positive constant when applying σ-modification [23], a very small σ_i and γ_i may not be enough to prevent the NN weight estimates from drifting to very large values in the presence of NN approximation errors [24], where the large

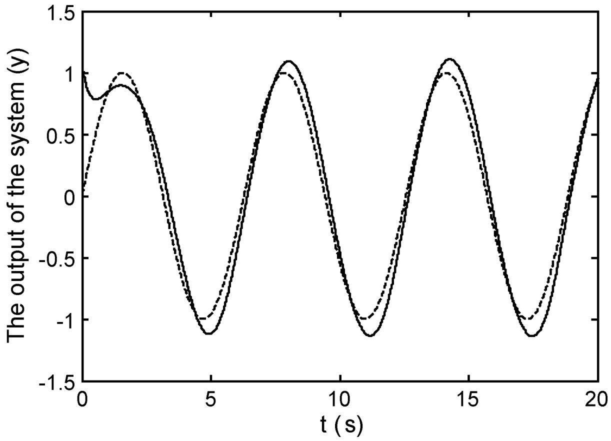


Fig. 1. The output of system under the adaptive controller.

θ_i and δ_i might result in a variation of a high gain control. Therefore, in practical applications, the design parameters should be adjusted carefully for achieving suitable transient performance and control action.

Remark 3: The adaptive NN controller (31) and the adaptation laws (11), (22), (28), and (34) are highly structural, and independent of the complexities of the system nonlinearities. Thus, it can be applied to other similar plants without repeating the complex controller design procedure for different system nonlinearities. In addition, such a property is particularly suitable for parallel processing and hardware implementation in practical applications.

B. Simulation

Here, a simple simulation is presented to show the effectiveness of the approach proposed above. The model of the system is given as

$$\begin{aligned} \dot{x}_1 &= 0.5x_1 + \left(1 + 0.1x_1^2\right) x_2 \\ \dot{x}_2 &= x_1 x_2 + [2 + \cos(x_1)]\, u \\ y &= x_1 \end{aligned} \tag{48}$$

where x_1 and x_2 are states, and y is the output of the system, respectively. The initial conditions is $x_0 = [x_{10}, x_{20}]^T = [1, 0]^T$ and the desired reference signal of the system is $y_d = \sin(t)$.

In this paper, all the basis functions of the NNs have the form [20]

$$G(\overline{x}_i) = \exp\left[-\frac{(\overline{x}_i - u_i)^T(\overline{x}_i - u_i)}{v_i^2}\right] \tag{49}$$

where $u_i = [u_{i1}, u_{i2}, \ldots, u_{ij}]^T$ is the center of the receptive field and v_i is the width of the Gaussian function. The NNs $\theta_1^T\xi_1(x_1)$, $\delta_1^T\eta_1(x_1)$ and $\delta_2^T\eta_1(x_1)$ all contain 13 nodes (i.e., $l_1 = 13$), with centers u_j evenly spaced in $[-6, 6]$, and widths $v_j = 1$ $(j = 1, 2, \ldots, l_1)$. NN $\theta_2^T\xi_2(\overline{x}_2)$ contains 169 nodes (i.e., $l_2 = 169$), with centers u_j evenly spaced in $[-6, 6] \times [-6, 6]$, and widths $v_j = 1$ $(j = 1, 2, \ldots, l_1)$. The design parameters of the above controller are $k_1 = k_2 = 3.5$, $\Gamma_1 = \Gamma_2 = \text{diag}\{2\}$, $\sigma_1 = \sigma_2 = \gamma_1 = \gamma_2 = 0.2$. The initial weights θ_1 and θ_2 are all given arbitrarily in $[-1, 1]$, and δ_1 and δ_2 in $[0, 1]$.

Figs. 1 and 2 show the simulation results of applying controller (31) to (47) for tracking desired signal y_d. From Fig. 1, we can see that good tracking performance is obtained. Fig. 2(a) shows the trajectory of the controller and Fig. 2(b) is magnification of the partition of it when the system is in tracking steady-state phase.

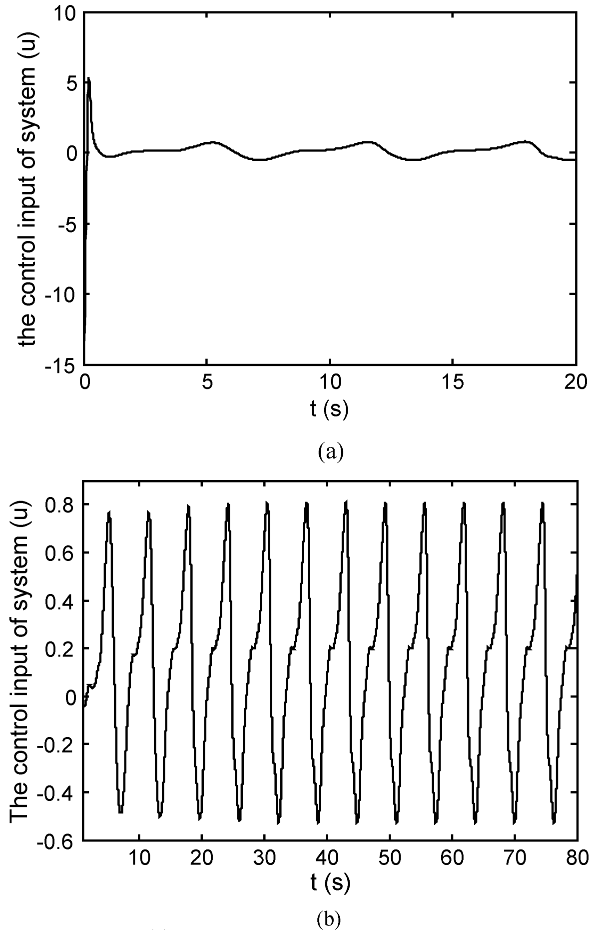


Fig. 2. The trajectory of the adaptive controller. (a) controller values in the interval [0, 20]. (b) controller values in the interval [1, 80].

IV. Robust Adaptive NN Control

A. System Description

From the proof of *Theorem 1* and the result of simulations, we can conclude that there exists a certain amount of error between the output of system and the desired signal, although tracking performance of this system is fairly good. It may also be that the tracking requirements are not met under some conditions. So, in this section, we will present another design approach, which can guarantee much higher tracking accuracy and have the property of "robustness."

Here, virtual controllers are denoted by the x_{ir}, $i = 1, 2, \ldots, n$, which satisfies $x_{1d} = x_{1r}$ and

$$\begin{aligned} x_{2r} &= -\theta_1^T\xi_1(x_1) + \delta_1^T\eta_1(x_1)\dot{x}_{1r} \\ &\quad - k_1 e_1 - \beta_1 \tanh(\alpha_1 e_1) \\ x_{(i+1)r} &= -e_{i-1} - \theta_i^T\xi_i(\overline{x}_i) + \delta_i^T\eta_i(\overline{x}_i)\dot{x}_{ir} \\ &\quad - k_i e_i - \beta_i \tanh(\alpha_i e_i) \\ u &= -e_{n-1} - \theta_n^T\xi_n(\overline{x}_n) + \delta_n^T\eta_n(\overline{x}_n)\dot{x}_{nr} \\ &\quad - k_n e_n - \beta_n \tanh(\alpha_n e_n). \end{aligned} \tag{50}$$

Equation (50) can be obtained by a similar design procedure as used in Section III, except that a term $u_{ir} = \tanh(\alpha_i e_i)$ is introduced in every virtual and actual controller for robustness. During the stability analysis of the closed-loop system, for simplicity, we will replace the function $\tanh(\cdot)$ with the following saturation function $\mathrm{sat}(\cdot)$, depicted in Fig. 3

$$\mathrm{sat}(\alpha_i e_i) = \begin{cases} 0.9, & \|e_i\| \geq \frac{\tanh^{-1}(0.9)}{\alpha_i} \\ \alpha_i e_i, & \|e_i\| < -\frac{\tanh^{-1}(0.9)}{\alpha_i} \end{cases} \tag{51}$$

where α_i and β_i are all positive constants, $i = 1, 2, \ldots, n$.

Obviously, virtual controllers x_{ir} are continuous, So they can be applied in backstepping design.

Assumptions 3: The ideal weights of NNs are bounded by known positive values so that

$$\begin{matrix} \|\theta_l^*\| \leq \Theta_l \\ \|\delta_l^*\| \leq \Delta_l \end{matrix} \quad \text{or} \quad \begin{matrix} \|\theta^*\| \leq \Theta \\ \|\delta^*\| \leq \Delta \end{matrix}$$

where $l = 1, 2, \ldots, n$.

Let us suppose that Assumptions 1, 2, and 3 are all satisfied. The following theorem can then be stated.

Theorem 2: Take the control input (49) with NN weight tuning be provided by

$$\begin{aligned} \dot{\theta}_i &= \Gamma_{i1}\left[e_i \xi_i(\overline{x}_i) - \sigma\|e\|\theta_i\right] \\ \dot{\delta}_i &= \Gamma_{i2}\left[-e_i \eta_i(\overline{x}_i)\dot{x}_{ir} - \gamma\|e\|\delta_i\right] \end{aligned} \tag{52}$$

where σ and δ are positive constants. Then the NN weights are UUB, and the errors $\mathrm{e}_i(t)$, $i = 1, 2, \ldots, n$, are semiglobally asymptotic stable. By choosing suitable parameters k_i, α_i and β_i, the tracking error of the system can converge into an arbitrarily small compact set.

Proof: In Step 1, replacing x_{2d} with x_{2r} in (7) and substituting (7) and (51) into (10), then choosing k_1 in (10) such that $(k_1^* \doteq k_1 - (g_{1d}/2g_1^2)) > 0$, we can obtain

$$\begin{aligned} \dot{V}_1 \leq\ & e_1 e_2 - k_1^* e_1^2 + \sigma\|e\|\widetilde{\theta}_1^T \theta_1 \\ & + \gamma\|e\|\widetilde{\delta}_1^T \delta_1 - \beta_1 e_1 \tanh(\alpha_1 e_1) + e_1 d_1 \end{aligned} \tag{53}$$

In Step 2, replacing x_{3d} with x_{3r} in (17) and substituting (17) and (51) into (21), then choosing k_2 in (21) such that $(k_2^* \doteq k_2 - (g_{2d}/2g_2^2)) > 0$, we can obtain

$$\begin{aligned} \dot{V}_2 \leq\ & e_2 e_3 - \sum_{l=1}^{2} k_l^* e_l^2 + \|e\| \sum_{l=1}^{2} \left(\sigma\widetilde{\theta}_l^T \theta_l + \gamma\widetilde{\delta}_l^T \delta_l\right) \\ & - \sum_{l=1}^{2} \beta_l e_l \tanh(\alpha_l e_l) + \sum_{l=1}^{2} e_l d_l. \end{aligned} \tag{54}$$

In a similar fashion, in Step n, we can obtain that the overall Lyapunov function candidate, satisfying the following inequality:

$$\begin{aligned} \dot{V}_n \leq\ & -\sum_{l=1}^{n} k_l^* e_l^2 + \|e\| \sum_{l=1}^{n} \left(\sigma\widetilde{\theta}_l^T \theta_l + \gamma\widetilde{\delta}_l^T \delta_l\right) \\ & - \sum_{l=1}^{n} \beta_l e_l \tanh(\alpha_l e_l) + \sum_{l=1}^{n} e_l d_l \end{aligned} \tag{55}$$

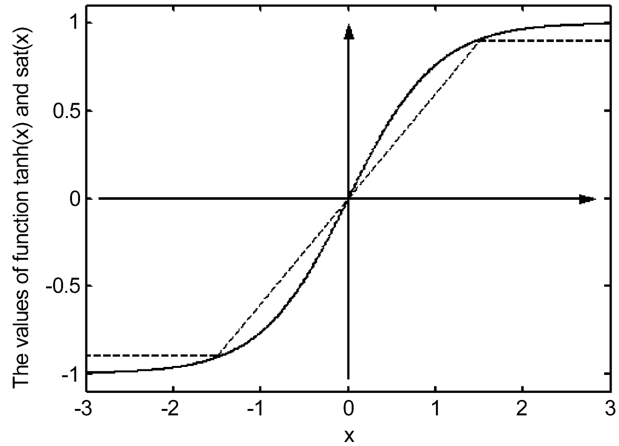


Fig. 3. The function $\tanh(x)$ and $\mathrm{sat}(x)$.

By combining (39), $d = [d_1, d_2, \ldots, d_n]^T$ and $\tanh(\alpha^T e) = [\tanh(\alpha_1 e_1), \tanh(\alpha_2 e_2), \ldots, \tanh(\alpha_n e_n)]^T$, (55) can be transformed into

$$\begin{aligned} \dot{V}_n \leq\ & -k_{\min}^*\|e\|^2 + \|e\|(\sigma\widetilde{\theta}^T\theta + \gamma\widetilde{\delta}^T\delta) \\ & - \beta_{\min} e^T \tanh(\alpha^T e) + e^T d \\ \leq\ & -k_{\min}^*\|e\|^2 + \|e\|\left[\sigma\widetilde{\theta}^T(\theta^* - \widetilde{\theta}) + \gamma\widetilde{\delta}^T(\delta^* - \widetilde{\delta})\right] \\ & - \beta_{\min} e^T \tanh(\alpha^T e) + \|e\|\,\|d\| \end{aligned} \tag{56}$$

where $(\cdot)_{\min}$ is the minimum of $(\cdot)_i$, $i = 1, 2, \ldots, n$.

Applying *Schwartz* inequality to (56)

$$\begin{aligned} \widetilde{\theta}^T(\theta^* - \widetilde{\theta}) &= \leq \|\widetilde{\theta}\|\,\|\theta^*\| - \|\widetilde{\theta}\|^2 \leq \|\widetilde{\theta}\|\Theta - \|\widetilde{\theta}\|^2 \\ \widetilde{\delta}^T(\delta^* - \widetilde{\delta}) &= \leq \|\widetilde{\delta}\|\,\|\delta^*\| - \|\widetilde{\delta}\|^2 \leq \|\widetilde{\delta}\|\Delta - \|\widetilde{\delta}\|^2 \end{aligned} \tag{57}$$

we have

$$\begin{aligned} \dot{V}_n \leq\ & \|e\|\left[\sigma\left(\|\widetilde{\theta}\|\Theta - \|\widetilde{\theta}\|^2\right) + \gamma\left(\|\widetilde{\delta}\|\Delta - \|\widetilde{\delta}\|^2\right)\right] \\ & + k_{\min}^*\|e\|^2 - \beta_{\min} e^T \tanh(\alpha^T e) + \|e\|\,\|d\| \\ \leq\ & -\|e\|\left[\sigma\left(\|\widetilde{\theta}\| - \frac{\Theta}{2}\right)^2 + \gamma\left(\|\widetilde{\delta}\| - \frac{\Delta}{2}\right)^2 \right. \\ & \left. - \frac{\Theta^2}{4} - \frac{\Delta^2}{4} + k_{\min}^*\|e\| - \|d\|\right] \\ & - \beta_{\min} e^T \tanh(\alpha^T e) \end{aligned} \tag{58}$$

Obviously, if

$$\begin{aligned} \|\widetilde{\theta}\| &\geq \frac{\Theta}{2} + \sqrt{\frac{\Theta^2}{4} + \frac{\Delta^2}{4} + \frac{\|d\|}{\sigma}} \\ \|\widetilde{\delta}\| &\geq \frac{\Delta}{2} + \sqrt{\frac{\Theta^2}{4} + \frac{\Delta^2}{4} + \frac{\|d\|}{\gamma}} \end{aligned} \tag{59}$$

the derivative of Lyapunov candidate is negative. Thus, the NN's weights are UUB. From (51), when all the error signals satisfy

$$|e_i| \geq \frac{\tanh^{-1}(0.9)}{\alpha_i}, \quad i = 1, 2, \ldots, n$$

the derivative of V_n becomes

$$\begin{aligned} \dot{V}_n \leq -\|e\|\left[\sigma\left(\|\widetilde{\theta}\| - \frac{\Theta}{2}\right)^2 + \gamma\left(\|\widetilde{\delta}\| - \frac{\Delta}{2}\right)^2 - \frac{\Theta^2}{4} - \frac{\Delta^2}{4} - \|d\|\right] \\ - 0.9\beta_{\min}\|e\|\mathrm{sgn}(e) - k_{\min}^*\|e\|^2. \end{aligned} \tag{60}$$

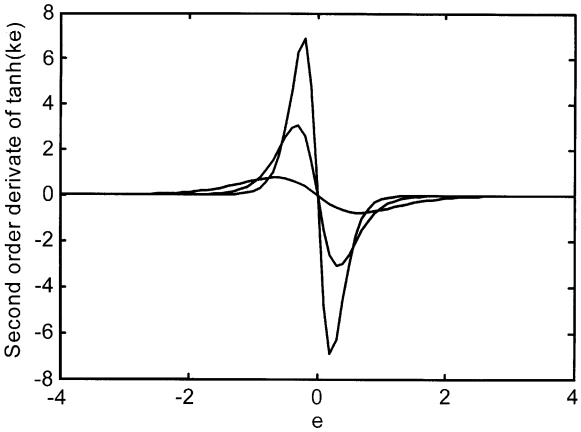


Fig. 4. The second order derivatives of $\tanh(kx)$ (dashed line: $k = 1$, solid line: $k = 2$, dash-dot line: $k = 3$).

So, if

$$\beta_{\min} \geq \frac{\left(\frac{\Theta^2}{4} + \frac{\Delta^2}{4} + \|d\|\right)}{0.9} \tag{61}$$

is satisfied, the derivative of V_n is negative. So, at least, the errors converge into a compact set

$$\Omega_e = \left\{ e \in R^n \middle| e_i \geq \frac{\tanh^{-1}(0.9)}{\alpha_i} \quad i = 1, 2, \cdots, n \right\} \tag{62}$$

Remark 4: In theorem 2, the term "semiglobally" asymptotically stable isused because the NNs can only approximate the nonlinear function in some compact set. If we can find NNs which approximate the nonlinear function in all system state space, the approach can guarantee global asymptotic stability.

Remark 5: The procedure of proof indicate that the approach is also effective for some other kinds of uncertainties (for example $\Delta f(x)$) or disturbances d_i, $i = 1, 2, \ldots, \mathrm{n}$, besides neural reconstruction errors. So, it can be set to be "robust" to some extent.

Remark 6: It is seen [from the equation before (60)] that the larger the parameters α_i, $i = 1, 2, \ldots, n$, are, the smaller will the tracking errors be. However, for high-order systems, due to the existence of high-order derivatives in the virtual controller, too large values α_i will require a quick response of the controller and the output will take large values when the error vector is in a small compact set (see Fig. 4). This is difficult to realize in practice. Besides, parameters k_i, effect the tracking performance. So, both the parameters α_i and k_i should be chosen suitably during the design stage.

B. Simulation

Here, the plant simulated is again (48) and the objective is to guarantee that the output y tracks the desired signal $y_d = \sin(t)$. For the simulation studies, all the parameters and the NNs are kept the same as those in Section III, except for $\alpha_1 = \alpha_2 = 1$ and $\beta_1 = \beta_2 = 3$.

Fig. 5 and Fig. 6 show the simulation results. From Fig. 5, we can draw the conclusion that this robust approach can guarantee much higher tracking accuracy than the adaptive one. Comparing Fig. 2 to Fig. 6, it can be seen that: 1) the robust controller has a larger maximum than the adaptive case during the tracking transient-state phase and can make the system to track the desired signal faster due to the terms introduced for robustness; and 2) the robust controller has smaller values when the system is in steady-state phase due to the smaller steady-state errors.

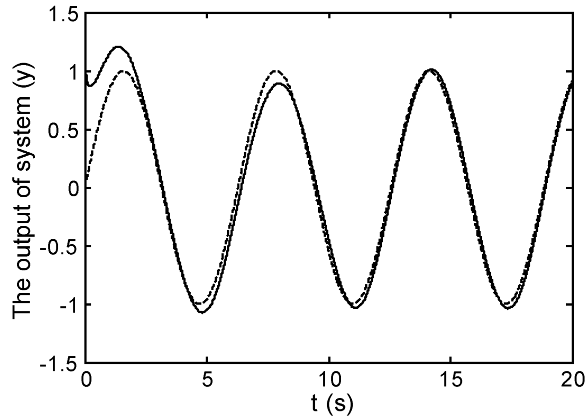
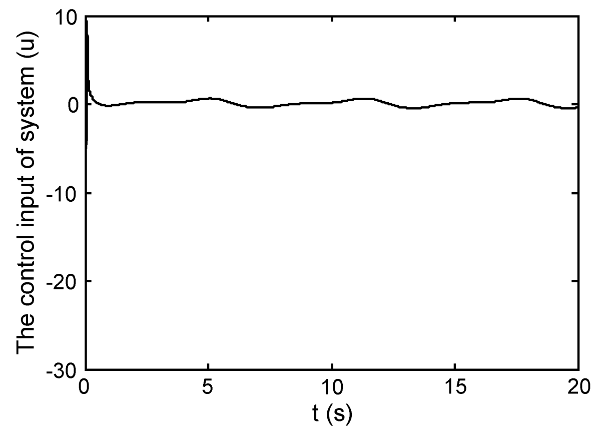
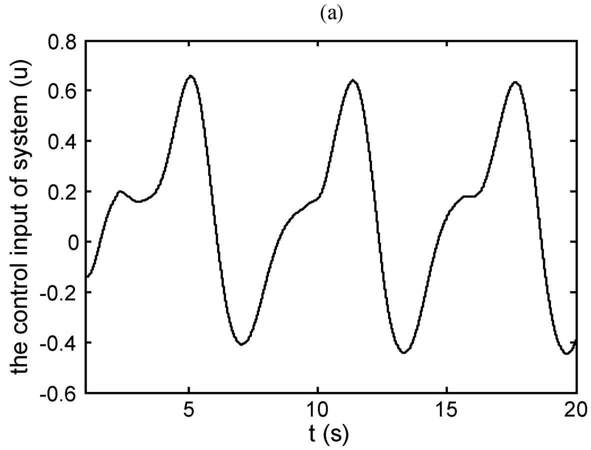


Fig. 5. The output of system under the robust adaptive controller.



(a)



(b)

Fig. 6. The trajectory of the adaptive controller. (a) Controller values in the interval [0, 20]. (b) controller values in the interval [1, 20].

V. Conclusion

In this paper, two backstepping NN control approaches are presented for a class of affine nonlinear systems in strict feedback form with unknown nonlinearities. By a special design scheme, both of the approaches avoid the controller singularity problem perfectly. The signals in the closed loop in the two methods are all guaranteed to be semiglobally uniformly ultimately bounded and the outputs of the system are both proved to converge to a small neighborhood of the desired trajectory. The control performances of the closed-loop systems can shaped by suitably choosing the design parameters. Simulation results carried out on the same nonlinear system demonstrate the effectiveness of the proposed approaches. The differences observed are analyzed briefly.

REFERENCES

[1] M. Krstic, I. Kanellakopoulos, and P. V. Kokotovic, *Nonlinear and Adaptive Control Design*. New York: Wiley, 1995.
[2] I. Kanellakopoulos, P. V. Kokotovic, and A. S. Morse, "Systematic design of adaptive controller for feedback linearizable system," *IEEE Trans. Automat. Contr.*, vol. 36, pp. 1241–1253, 1991.
[3] P. V. Kokotovic, "The joy of feedback: nonlinear and adaptive," *IEEE Contr. Syst. Mag.*, vol. 12, pp. 7–17, 1992.
[4] M. Krstic, I. Kanellakopoulos, and P. V. Kokotovic, "Adaptive nonlinear control without overparameterization," *Syst. Contr. Lett.*, vol. 19, pp. 177–185, 1992.
[5] M. Krstic and P. V. Kokotovic, "Adaptive nonlinear design with controller-identifier eparation and swapping," *IEEE Trans. Automat. Contr.*, vol. 40, pp. 426–440, 1995.
[6] C. Kwan and F. L. Lewis, "Robust backstepping control of nonlinear systems using neural networks," *IEEE Trans. Syst., Man, Cybern. A*, vol. 30, pp. 753–765, 2000.
[7] T. Knohl and H. Unbehauen, "ANNNAC—extension of adaptive backstepping algorithm with artificial neural networks," *Inst. Elect. Eng. Proc. Contr. Theory Appl.*, vol. 147, pp. 177–183, 2000.
[8] D. M. Dawson, J. J. Carroll, and M. Schneider, "Integrator backstepping control of a brush DC motor turning a robotic load," *IEEE Trans. Contr. Syst. Technol.*, vol. 2, pp. 233–244, 1994.
[9] S. R. Chu and R. Shoureshi, "Neural-based adaptive nonlinear ystem identification," in *Proc. Intelligent Control Systems, ASME Winter Annu. Meeting*, vol. DSC-45, 1992.
[10] B. Horn, D. Hush, and C. Abdallah, "The state space recurrent neural network for robot identification," in *Proc. Advanced Control Issues for Robot Manipulators, ASME Winter Annu. Meeting*, vol. DSC-39, 1992.
[11] F. C. Chen and H. K. Khalil, "Adaptive control of nonlinear systems using neural networks," *Int. J. Contr.*, vol. 55, no. 6, pp. 1299–1317, 1992.
[12] K. S. Narendra, "Adaptive control using neural networks," in *Neural Networks for Control*. Cambridge, MA: MIT Press, 1991, pp. 115–142.
[13] T. Ozaki, T. Suzuki, T. Furuhashi, S. Okuma, and Y. Ushikawa, "Trajectory control of robotic manipulators," *IEEE Trans. Ind. Electron.*, vol. 38, pp. 195–202, 1991.
[14] Y. Zhang, P. Y. Peng, and Z. P. Jiang, "Stable neural controller design for unknown nonlinear systems using backstepping," *IEEE Trans. Neural Networks*, vol. 11, pp. 1347–1360, 2000.
[15] C. M. Kwan and F. L. Lewis, "Robust backstepping control of induction motors using neural networks," *IEEE Trans. Neural Networks*, vol. 11, pp. 1178–1187, 2000.
[16] J. Y. Choi and J. A. Farrell, "Adaptive observer backstepping control using neural networks," *IEEE Trans. Neural Networks*, vol. 12, pp. 1103–1112, 2001.
[17] S. Jagannathan, "Control of a class of nonlinear discrete time systems using multilayer neural networks," *IEEE Trans. Neural Networks*, vol. 12, pp. 1113–1120, 2001.
[18] J. Q. Gong and B.Bin Yao, "Neural network adaptive robust control of nonlinear systems in semi-strict feedback form," *Automatica*, vol. 37, pp. 1149–1160, 2001.
[19] T. Zhang, S. S. Ge, and C. C. Hang, "Adaptive neural network control for strict-feedback nonlinear systems using backstepping design," *Automatica*, vol. 36, pp. 1835–1846, 2000.
[20] S. S.Shuzhi S. Ge and C. Wang, "Direct adaptive NN control of a class of nonlinear systems," *IEEE Trans. Neural Networks*, vol. 13, 2002.
[21] R. M. Sanner and J. J. E. Slotine, "Gussian networks for direct adaptive control," *IEEE Trans. Neural Networks*, vol. 3, pp. 837–863, 1992.
[22] K. S. Narendra and A. M. Annaswamy, "A new adaptive law for robust adaptation withour persistent excitation," *IEEE Trans. Automat. Contr.*, vol. 32, pp. 134–145, 1987.
[23] P. V. Kokotovic, "Bode lecture: the joy of feedback," *IEEE Contr. Syst. Mag.*, vol. 3, pp. 7–17, June 1992.
[24] G. Cybenko, "Approximation by superpositions of a sigmoidal function," *Math. Contr., Signals, Syst.*, vol. 2, no. 4, pp. 303–314, 1989.

Yahui Li was born in Hebei Province, China, on Feburary 1, 1974. He received the M.S. degree in material science, in 2000 and the Ph.D. degree in control theory and control engineering, in 2003, both from the Harbin Institute of Technology (HIT), China.

His research focuses on nonlinear system control, adaptive control, and intelligent control.

Sheng Qiang received the M.S. degree in pattern recognition and intelligent control in, 1998, and the Ph.D. degree in control theory and control engineering, in 2004, both from the Harbin Institute of Technology (HIT), China.

He is an Associate Professor of Control Science and Engineering Department, HIT. He has published more than 20 papers in journals and international conference proceedings. His research interests include soft computing methods, computer control, and process optimization.

Dr. Qiang is a member of the IPC of the 12th Mediterranean Conference on Control and Automation and the ITC of the 2005 IEEE Midsummer Workshop on Soft Computing in Industrial Applications.

Xianyi Zhuang was born in Jilin Province, China, in 1936. He received the B.Sc. degree in mathematics science from the University of Northeast Normal, China.

He is a Professor with the Department of Control Science and Engineering, Harbin Institute of Technology, China. His research focuses on adaptive control, intelligent control, and the development of high degree accuracy servo systems.

Okyay Kaynak (M'80–M'83–SM'90–F'03) received the B.Sc. degree with first class honors and the Ph.D. degree in electronic and electrical engineering from the University of Birmingham, U.K., in 1969 and 1972, respectively.

From 1972 to 1979, he held various positions within the industry. In 1979, he joined the Department of Electrical and Electronics Engineering, Bogazici University, Istanbul, Turkey, where he is presently a Full Professor. He served as the Chairman of the Computer Engineering Department for three years, of the Electrical and Electronic Engineering Department for two years, and was the Director of Biomedical Engineering Institute for one year. Currently, he is the UNESCO Chair on Mechatronics and the Director of Mechatronics Research and Application Centre. He has held long-term Visiting Professor/Scholar positions at various institutions in Japan, Germany, the U.S., and Singapore. His current research interests are in the fields of intelligent control and mechatronics. He has authored three books and edited five. He has also authored or coauthored more than 200 papers that have appeared in various journals and conference proceedings.

Dr. Kaynak was the President of the IEEE Industrial Electronics Society during 2002–2003. He is now of the Vice Presidents of the Neural Networks Society. He serves as an Associate Editor of both the IEEE TRANSACTIONS ON INDUSTRIAL ELECTRONICS and the IEEE TRANSACTIONS ON NEURAL NETWORKS and additionally he is on the Editorial or Advisory Boards of a number of scholarly journals.