PERGAMON

Automatica 36 (2000) 1835–1846

automatica

www.elsevier.com/locate/automatica

Adaptive neural network control for strict-feedback nonlinear systems using backstepping design☆

T. Zhang, S. S. Ge*, C. C. Hang

Department of Electrical Engineering, National University of Singapore, 10 Kent Ridge Crescent, Singapore 119260, Singapore

Received 7 May 1998; revised 13 October 1999; received in final form 20 May 2000

Abstract

This paper focuses on adaptive control of strict-feedback nonlinear systems using multilayer neural networks (MNNs). By introducing a modified Lyapunov function, a smooth and singularity-free adaptive controller is firstly designed for a first-order plant. Then, an extension is made to high-order nonlinear systems using neural network approximation and adaptive backstepping techniques. The developed control scheme guarantees the uniform ultimate boundedness of the closed-loop adaptive systems. In addition, the relationship between the transient performance and the design parameters is explicitly given to guide the tuning of the controller. One important feature of the proposed NN controller is the highly structural property which makes it particularly suitable for parallel processing in actual implementation. Simulation studies are included to illustrate the effectiveness of the proposed approach. © 2000 Elsevier Science Ltd. All rights reserved.

Keywords: Nonlinear systems; Adaptive control; Neural networks; Lyapunov stability

1. Introduction

Recently, interest in adaptive control of nonlinear systems has been ever increasing, and many significant developments have been achieved. In the early stage of the research, Sastry and Isidori (1989), Teel, Kadiyala, Kokotovic and Sastry (1991), Nam and Arapostations (1988), Taylor, Kokotovic, Marino and Kanellkopoulos (1989), Kanellakopoulos, Kokotovic and Marino (1989) and Pomet and Praly (1992) presented several important results on adaptive nonlinear control. In order to guarantee the global stability, some restrictions on the plants had to be made such as matching condition, extended matching condition, or growth conditions on system nonlinearities. In an attempt to overcome these restrictions, a novel recursive design procedure, adaptive backstepping, was provided, and a globally stable and asymptotic tracking adaptive controller was developed for parametric strict-feedback systems by Kanellakopoulos, Kokotovic and Morse (1991). In an effort to extend the backstepping idea to a larger class of nonlinear systems, Kanellakopoulos et al. (1991) studied the adaptive control problem of pure-feedback systems and obtained regionally stable results, Seto, Annaswamy and Baillienl (1994) proposed several adaptive approaches for nonlinear systems with triangular structures, and Krstic, Kanellakopoulos and Kokotovic (1995) further extended the adaptive backstepping technique to parametric strict-feedback systems with unknown virtual control coefficients. Recently, robust adaptive controllers have been studied for a class of semi-strict feedback systems by combining backstepping technique with robust control strategy (Polycarpou & Ioannou, 1996; Yao & Tomizuka, 1997), which guarantee global uniform ultimate boundedness in the presence of both parametric uncertainties and unknown nonlinearities.

As an alternative, intensive research has been carried out on neural network (NN) control of unknown nonlinear dynamic systems. For stable and efficient NN control (Narendra & Parthasarathy, 1990; Jin, Nikifornk & Gupta, 1993; Chen & Liu, 1994), off-line training phases were required before the NN controllers can be put into operation. To overcome such a problem, Lyapunov's stability theory was applied in the controller design and

☆This paper was not presented at any IFAC meeting. This paper was recommended for publication in revised form by Associate Editor Hassan Khalil under the direction of Editor Tamer Basar.

* Corresponding author. Tel.: + 65-874-6821; fax.: + 65-779-1103.
E-mail address: elegesz@nus.edu.sg (S. S. Ge).

0005-1098/00/$ - see front matter © 2000 Elsevier Science Ltd. All rights reserved.
PII: S0005-1098(00)00116-3

several stable adaptive NN control approaches were developed (Polycarpou & Ioannou, 1991; Sanner & Slotine, 1992; Yesidirek & Lewis, 1995; Spooner & Passino, 1996; Ge, Lee & Harris, 1998; Ge, Hang & Zhang, 1999a,b; Fabri & Kadirkamanathan, 1996; Zhang, Ge & Hang, 1999,2000). It is worth noting that most of these neural control schemes need some types of matching conditions, i.e., the unknown nonlinearities appear in the same equation as the control input in a state-space representation. Recently, using the idea of adaptive backstepping, Polycarpou (1996) presented an interesting neural-based adaptive controller for a class of nonlinear systems without satisfying matching condition, which ensures the semi-global stability of the closed-loop system. The research works (Kwan & Lewis, 1995a,b; Kwan, Lewis & Dawson, 1998) studied robust NN control for induction motor and robot systems via backstepping techniques. Adaptive backstepping control was extended to the nonautonomous strict-feedback system in the work (Ge, Wang & Lee, 2000).

In this paper, we deal with the control problem of nonlinear systems transformable to the following strict-feedback canonical form:

$$\begin{aligned}\dot{x}_i &= f_i(x_1, x_2, \dots, x_i) + g_i(x_1, x_2, \dots, x_i)x_{i+1}, \\ &\quad 1 \le i \le n-1, \\ \dot{x}_n &= f_n(x) + g_n(x)u, \\ y &= x_1, \end{aligned} \tag{1}$$

where $x = [x_1, x_2, \dots, x_n]^{\mathrm{T}} \in R^n$, $u \in R$, $y \in R$ are the state variables, system input and output, respectively; $f_i(\cdot)$ and $g_i(\cdot)$, $i = 1,2,\dots,n$ are unknown smooth functions and may not be linearly parameterized. For this control problem, one of the main difficulties comes from uncertainties $g_i(\cdot)$. When $g_i(\cdot)$ are known exactly, the design scheme provided by Polycarpou (1996) can be applied directly. Without such a knowledge, no effective method is available in the literature at the present stage. In this paper, by utilizing a novel kind of Lyapunov functions, we develop a semi-globally stable adaptive controller which does not require to estimate the unknown functions $g_i(\cdot)$, and therefore avoids the possible controller singularity problem usually met in feedback linearization design. Other features of the proposed method lie in the use of MNNs as function approximators and the guaranteed stability of the closed-loop neural adaptive systems.

This paper is organized as follows. Section 2 presents some notations and multilayer NNs used in the controller design. Section 3 provides a NN adaptive controller for a first-order system using an integral-type Lyapunov function. This scheme is then extended to n-dimensional systems in Section 4, and both stabilities and control performance of the closed-loop systems are discussed as well. Section 5 contains several simulation examples to show the effectiveness of the proposed controller.

2. Preliminaries

The control objective is to design an adaptive controller for strict-feedback system (1) such that the output y follows a desired trajectory y_{d}, while all signals in the closed-loop systems are bounded. Let $\mathbf{x}_i = [x_1, x_2, \dots, x_i]^{\mathrm{T}}$, $\mathbf{x}_{\mathrm{d}(i+1)} = [y_{\mathrm{d}}, \dot{y}_{\mathrm{d}}, \dots, y_{\mathrm{d}}^{(i)}]^{\mathrm{T}}$, $i = 1, \dots, n$, $\|\cdot\|$ denote the 2-norm, $\|\cdot\|_{\mathrm{F}}$ denote the Frobenius norm, $|A|_1 = \sum_{i=1}^{m} |a_i|$ with $A = [a_1, a_2, \dots, a_m]^{\mathrm{T}} \in R^m$, and $\lambda_{\max}(B)$ and $\lambda_{\min}(B)$ denote the largest and smallest eigenvalues of a square matrix B, respectively.

Assumption 1. The signs of $g_i(\mathbf{x}_i)$ are known, and there exist constants $g_{i0} > 0$ and known smooth functions $\mathbf{g}_i(\mathbf{x}_i)$ such that $\mathbf{g}_i(\mathbf{x}_i) \ge |g_i(\mathbf{x}_i)| \ge g_{i0}$, $\forall \mathbf{x}_i \in R^i$.

Remark 2.1. The above assumption implies that smooth functions $g_i(\mathbf{x}_i)$ are strictly either positive or negative. From now onwards, without losing generality, we shall assume $\mathbf{g}_i(\mathbf{x}_i) \ge g_i(\mathbf{x}_i) \ge g_{i0} > 0$, $\forall \mathbf{x}_i \in R^i$. Assumption 1 is reasonable because $g_i(\mathbf{x}_i)$ being away from zeros are controllable conditions of system (1), which is made in most of control schemes (Krstic et al., 1995; Sepulchre, Jankovic & Kokotovic, 1997). For a given practical system, the upper bounds of $g_i(\mathbf{x}_i)$ are not difficult to determine by choosing functions $\mathbf{g}_i(\mathbf{x}_i)$ large enough. It should be emphasized that the low bounds g_{i0} are only required for analytical purposes, their true values are not necessarily known.

Assumption 2. The desired trajectory vectors $\mathbf{x}_{\mathrm{d}i}$ with $i = 2, \dots, n+1$ are continuous and available, and $\mathbf{x}_{\mathrm{d}i} \in \Omega_{\mathrm{d}i} \subset R^i$ with $\Omega_{\mathrm{d}i}$ known compact sets.

Multilayer neural networks are usually used as a tool for modelling nonlinear functions because of their good capabilities in function approximation. In this paper, the following three-layer NNs (Lewis, Yesildirek & Liu, 1996) are used to approximate a smooth function $h(Z): R^m \to R$,

$$g_{\mathrm{nn}}(Z) = W^{\mathrm{T}} S(V^{\mathrm{T}} \mathbf{Z}), \tag{2}$$

where $\mathbf{Z} = [Z^{\mathrm{T}}, 1]^{\mathrm{T}}$ is the input vector; $W = [w_1, w_2, \dots, w_l]^{\mathrm{T}} \in R^l$ and $V = [v_1, v_2, \dots, v_l] \in R^{(m+1)\times l}$ are the first-to-second layer and the second-to-third layer weights, respectively; the NN node number $l > 1$; and $S(V^{\mathrm{T}}\mathbf{Z}) = [s(v_1^{\mathrm{T}}\mathbf{Z}), s(v_2^{\mathrm{T}}\mathbf{Z}), \dots, s(v_{l-1}^{\mathrm{T}}\mathbf{Z}), 1]^{\mathrm{T}}$ with $s(z_a) = 1/(1 + \mathrm{e}^{-\gamma z_a})$, constant $\gamma > 0$. It has been proven that neural network (2) satisfies the conditions of Stone–Weierstrass Theorem and can approximate any continuous function over a compact set (Funahashi, 1989). Therefore,

$$h(Z) = W^{*T}S(V^{*T}\mathbf{Z}) + \mu, \quad \forall Z \in \Omega_z \subset R^m, \tag{3}$$

where μ is the NN approximation error and Ω_z is a compact set.

Assumption 3. For a given smooth function $h(Z)$ and NN approximator (2), there exist ideal constant weights W^* and V^* such that $|\mu| \leq \boldsymbol{\mu}$ with constant $\boldsymbol{\mu} > 0$ for all $Z \in \Omega_z$.

In general, the ideal weights W^* and V^* are unknown and need to be estimated in controller design. Let $\hat{W}$ and $\hat{V}$ be the estimates of W^* and V^*, respectively, and the weight estimation errors are $\tilde{W} = \hat{W} - W^*$ and $\tilde{V} = \hat{V} - V^*$.

Lemma 2.1. *For neural network approximator* (2), *the NN estimation error can be expressed as*

$$\begin{aligned} &\hat{W}^T S(\hat{V}^T\mathbf{Z}) - W^{*T}S(V^{*T}\mathbf{Z}) \\ &\quad = \tilde{W}^T(\hat{S} - \hat{S}'\hat{V}^T\mathbf{Z}) + \hat{W}^T\hat{S}'\tilde{V}^T\mathbf{Z} + d_u, \end{aligned} \tag{4}$$

where $\hat{S} = S(\hat{V}^T\mathbf{Z})$, $\hat{S}' = diag\{\hat{s}'_1, \hat{s}'_2, \ldots, \hat{s}'_l\}$ *with* $\hat{s}'_i = s'(\hat{v}_i^T\mathbf{Z}) = d[s(z_a)]/dz_a|_{z_a = \hat{v}_i^T\mathbf{Z}}$, $i = 1,2,\ldots,l$, *and the residual term* d_u *is bounded by*

$$|d_u| \leq \|V^*\|_F \|\mathbf{Z}\hat{W}^T\hat{S}'\|_F + \|W^*\|\|\hat{S}'\hat{V}^T\mathbf{Z}\| + |W^*|_1. \tag{5}$$

The proof of Lemma 2.1 can be found in the reference (Zhang et al., 1999).

3. Adaptive control for first-order systems

In Lyapunov-based control design of nonlinear systems, the construction of Lyapunov functions is a crucial part and is usually an intractable problem. Different choices of Lyapunov functions may result in different control structures and closed-loop performance. In this section, an integral-type Lyapunov function is proposed for developing a smooth adaptive NN controller which is free from control singularity.

To illustrate the design methodology, we consider a first-order system

$$\dot{x}_1 = f_1(x_1) + g_1(x_1)u_1 \tag{6}$$

with u_1 as control input. Define $z_1 = x_1 - y_d$, $\beta_1(x_1) = \mathbf{g}_1(x_1)/g_1(x_1)$ and a smooth scalar function

$$V_{z1} = \int_0^{z_1} \sigma\beta_1(\sigma + y_d)\,d\sigma. \tag{7}$$

By changing the variable $\sigma = \theta z_1$, we may rewrite V_{z1} as $V_{z1} = z_1^2\int_0^1 \theta\beta_1(\theta z_1 + y_d)\,d\theta$. Noting that $1 \leq \beta_1(\theta z_1 + y_d) \leq \mathbf{g}_1(\theta z_1 + y_d)/g_{10}$ (Assumption 1), we have

$$\frac{z_1^2}{2} \leq V_{z1} \leq \frac{z_1^2}{g_{10}} \int_0^1 \theta\mathbf{g}_1(\theta z_1 + y_d)\,d\theta. \tag{8}$$

Therefore, V_{z1} is a positive-definite function with respect to z_1. In the following, Lyapunov function candidate V_{z1} shall be utilized to develop a desired feedback control (DFC).

Lemma 3.1. *For first-order system* (6), *if the DFC is chosen as*

$$u_1^* = \frac{1}{\mathbf{g}_1(x_1)}[-k(t)z_1 - h_1(Z_1)], \tag{9}$$

where the smooth function

$$\begin{aligned} &h_1(Z_1) = \beta_1(x_1)f_1(x_1) - \dot{y}_d\int_0^1 \beta_1(\theta z_1 + y_d)\,d\theta, \\ &\quad Z_1 = [x_1, y_d, \dot{y}_d]^T \in R^3 \end{aligned} \tag{10}$$

and $k(t) \geq k^* > 0$ *with* k^* *being any positive constant, then the system tracking error converges to zero asymptotically.*

Proof. Taking V_{z1} given in (7) as a Lyapunov function candidate, its time derivative along (6) is

$$\begin{aligned} \dot{V}_{z1} &= z_1\beta_1(x_1)\dot{z}_1 + \dot{y}_d\int_0^{z_1} \sigma\frac{\partial\beta_1(\sigma + y_d)}{\partial\sigma}\,d\sigma \\ &= z_1\beta_1(x_1)[g_1(x_1)u_1 + f_1(x_1) - \dot{y}_d] \\ &\quad + \dot{y}_d\left[\sigma\beta_1(\sigma + y_d)\Big|_0^{z_1} - \int_0^{z_1}\beta_1(\sigma + y_d)\,d\sigma\right] \\ &= z_1\Bigg[\mathbf{g}_1(x_1)u_1 + \beta_1(x_1)f_1(x_1) \\ &\quad - \dot{y}_d\int_0^1\beta_1(\theta z_1 + y_d)\,d\theta\Bigg]. \end{aligned} \tag{11}$$

Substituting $u_1 = u_1^*$ into (11), we obtain $\dot{V}_{z1} = -k(t)z_1^2 \leq -k^*z_1^2 \leq 0$. Hence, V_{z1} is a Lyapunov function and the tracking error $z_1 \to 0$ as $t \to \infty$ asymptotically. □

In the case that nonlinearities $f_1(x_1)$ and $g_1(x_1)$ are unknown, the desired controller u_1^* is not available due to the unknown function $h_1(Z_1)$. However, $h_1(Z_1)$ is a smooth function, and may be approximated by MNNs provided in (2). We choose the following controller for the first-order system (6)

$$u_1 = \frac{1}{\mathbf{g}_1(x_1)}[-k_1(t)z_1 - \hat{W}_1^T S_1(\hat{V}_1^T\mathbf{Z}_1)], \tag{12}$$

where neural network $\hat{W}_1^T S_1(\hat{V}_1^T\mathbf{Z}_1)$ is introduced to approximate $h_1(Z_1)$. We now specify the gain $k_1(t)$ and the adaptive algorithms for adjusting the NN weights to guarantee the system stability.

Theorem 3.1. *Consider the closed-loop system consisting of the first-order plant* (6) *and controller* (12), *if gain*

$$k_1(t) = \frac{1}{\varepsilon_1}\left(1 + \int_0^1 \theta \mathbf{g}_1(\theta z_1 + y_d)\,d\theta + \|\mathbf{Z}_1 \hat{W}_1^T \hat{S}_1'\|_F^2 + \|\hat{S}_1' \hat{V}_1^T \mathbf{Z}_1\|^2\right) \quad (13)$$

with constant $\varepsilon_1 > 0$, *and the NN weights are updated by*

$$\dot{\hat{W}}_1 = \Gamma_{w1}[(\hat{S}_1 - \hat{S}_1' \hat{V}_1^T \mathbf{Z}_1) z_1 - \sigma_{w1} \hat{W}_1], \quad (14)$$

$$\dot{\hat{V}}_1 = \Gamma_{v1}[\mathbf{Z}_1 \hat{W}_1^T \hat{S}_1' z_1 - \sigma_{v1} \hat{V}_1] \quad (15)$$

with $\Gamma_{w1} = \Gamma_{w1}^T > 0$, $\Gamma_{v1} = \Gamma_{v1}^T > 0$, *and* $\sigma_{w1}, \sigma_{v1} > 0$, *then for bounded initial conditions* $x_1(0), \hat{W}_1(0)$ and $\hat{V}_1(0)$, *all signals in the closed-loop system are bounded, and the vector* Z_1 *remains in*

$$\Omega_{z1} = \left\{(x_1, y_d, \dot{y}_d) \,\middle|\, |z_1(t)| \le \sqrt{2c_0 e^{-\lambda_1 t} + 2c_1/\lambda_1},\ \mathbf{x}_{d2} \in \Omega_{d2}\right\} \quad (16)$$

with positive constants c_0, c_1 *and* λ_1.

Proof. See Appendix A.

It is interesting to note that most of the available NN controllers in the literature are based on feedback linearization techniques, whose structures are usually taken the form $u = [-\hat{f}_1 + v]/\hat{g}_1$ with $\hat{f}_1$ and $\hat{g}_1$ be the estimates of f_1 and g_1, respectively, and v be a new control variable. To avoid singularity problem when $\hat{g}_1 \to 0$, several modified adaptive methods were provided (Wang, 1994; Spooner & Passino, 1996; Yesidirek & Lewis, 1995; Kosmatopoulos, 1996). However, either discontinuous controllers or projection adaptive algorithms have to be used, which becomes a main obstacle for applying backstepping technique to extend these schemes to general strict-feedback systems. The key point of the proposed design is to utilize the Lyapunov function V_{z1} in (7) for constructing the NN controller (12), which completely removes the possible controller singularity. In addition, both controller (12) and updating laws (14) and (15) are smooth. This makes the use of backstepping design possible for extending the approach to high-order systems.

4. Adaptive NN control for strict-feedback systems

Before proceeding the backstepping design, some notations are presented below. Define positive-definite functions $\beta_i(\mathbf{x}_i) = \mathbf{g}_i(\mathbf{x}_i)/g_i(\mathbf{x}_i)$, $i = 2, \ldots, n$, and let $h_i(Z_i)$ with input vectors Z_i be smooth functions on compact sets Ω_{zi}. According to the NN approximation properties shown in Section 2, we have

$$h_i(Z_i) = W_i^{*T} S_i(V_i^{*T} \mathbf{Z}_i) + \mu_i, \quad \forall Z_i \in \Omega_{zi},\ i = 1, 2, \ldots, n, \quad (17)$$

where W_i^* and V_i^* are ideal constant weights, and $|\mu_i| \le \boldsymbol{\mu}_i$ with constants $\boldsymbol{\mu}_i > 0$. It follows from Lemma 2.1 that the function estimation error ψ_i can be expressed as

$$\psi_i = \hat{W}_i^T S_i(\hat{V}_i^T \mathbf{Z}_i) - h_i(Z_i) = \tilde{W}_i^T(\hat{S}_i - \hat{S}_i' \hat{V}_i^T \mathbf{Z}_i) + \hat{W}_i^T \hat{S}_i' \tilde{V}_i^T \mathbf{Z}_i + d_{ui} - \mu_i, \quad (18)$$

where

$$\hat{W}_i = [\hat{w}_{i,1}, \hat{w}_{i,2}, \ldots, \hat{w}_{i,l_i}]^T \in R^{l_i}$$

and

$$\hat{V}_i = [\hat{v}_{i,1}, \hat{v}_{i,2}, \ldots, \hat{v}_{i,l_i}] \in R^{(m_i+1)\times l_i}$$

denote the estimates of W_i^* and V_i^*, respectively; $\hat{S}_i = S_i(\hat{V}_i^T \mathbf{Z}_i)$; $\hat{S}_i' = diag\{\hat{s}_{i,1}', \hat{s}'_{i,2}, \ldots, \hat{s}'_{i,l_i}\}$ with $\hat{s}'_{i,j} = s'(\hat{v}_{i,j}^T \mathbf{Z}_i) = d[s(z_a)]/dz_a|_{z_a = \hat{v}_{i,j}^T \mathbf{Z}_i}$, $j = 1, 2, \ldots, l_i$, and $l_i > 1$; and the residual terms d_{ui} are bounded by

$$|d_{ui}| \le \|V_i^*\|_F \|\mathbf{Z}_i \hat{W}_i^T \hat{S}_i'\|_F + \|W_i^*\| \|\hat{S}_i' \hat{V}_i^T \mathbf{Z}_i\| + |W_i^*|_1. \quad (19)$$

4.1. Backstepping design

The backstepping design procedure contains n steps. At each step, an intermediate control function α_k shall be developed using an appropriate Lyapunov function V_{sk} and the method being similar to that of Section 3.

Step 1: Let us firstly consider the equation in (1) when $i = 1$, i.e.,

$$\dot{x}_1 = f_1(x_1) + g_1(x_1) x_2.$$

By viewing x_2 as a virtual control input, we choose a new error variable $z_2 = x_2 - \alpha_1$ with $\alpha_1 = u_1$ defined in (12), then

$$\dot{z}_1 = f_1(x_1) + g_1(x_1)(z_2 + \alpha_1) - \dot{y}_d.$$

Taking V_{z1} given in (7) as a Lyapunov function candidate, along the similar procedure of (11), its time derivative can be expressed as

$$\dot{V}_{z1} = z_1[\mathbf{g}_1(x_1)(z_2 + \alpha_1) + h_1(Z_1)].$$

Using (12) and (18), we have

$$\dot{V}_{z1} = -k_1(t) z_1^2 - \psi_1 z_1 + z_1 \mathbf{g}_1(x_1) z_2. \quad (20)$$

Step 2: The equation in system (1) for $i = 2$ is given by

$$\dot{x}_2 = f_2(\mathbf{x}_2) + g_2(\mathbf{x}_2) x_3. \quad (21)$$

Again, by viewing x_3 as a virtual control, we may design a control input α_2 for (21). Define $z_3 = x_3 - \alpha_2$, we have

$$\dot{z}_2 = \dot{x}_2 - \dot{\alpha}_1 = f_2(\mathbf{x}_2) + g_2(\mathbf{x}_2)(z_3 + \alpha_2) - \dot{\alpha}_1. \quad (22)$$

Choosing a positive-definite function

$$V_{s2} = V_{z1} + \int_0^{z_2} \sigma\beta_2(\mathbf{x}_1,\sigma + \alpha_1)\,d\sigma,$$

its time derivative becomes

$$\dot{V}_{s2} = \dot{V}_{z1} + z_2\beta_2(\mathbf{x}_2)\dot{z}_2 + \int_0^{z_2} \sigma\left[\frac{\partial\beta_2(\mathbf{x}_1,\sigma + \alpha_1)}{\partial\mathbf{x}_1}\dot{\mathbf{x}}_1 + \frac{\partial\beta_2(\mathbf{x}_1,\sigma + \alpha_1)}{\partial\alpha_1}\dot{\alpha}_1\right]d\sigma. \quad (23)$$

Using (20), (22) and the facts that

(i)

$$\int_0^{z_2} \sigma\frac{\partial\beta_2(\mathbf{x}_1,\sigma + \alpha_1)}{\partial\alpha_1}\dot{\alpha}_1\,d\sigma = \dot{\alpha}_1\int_0^{z_2} \sigma\frac{\partial\beta_2(\mathbf{x}_1,\sigma + \alpha_1)}{\partial\sigma}\,d\sigma = \dot{\alpha}_1\left[z_2\beta_2(\mathbf{x}_2) - \int_0^{z_2}\beta_2(\mathbf{x}_1,\sigma + \alpha_1)\,d\sigma\right],$$

(ii)

$$\dot{\alpha}_1 = \frac{\partial\alpha_1}{\partial\mathbf{x}_1}\dot{\mathbf{x}}_1 + \omega_1,$$
$$\omega_1 = \frac{\partial\alpha_1}{\partial\mathbf{x}_{d2}}\dot{\mathbf{x}}_{d2} + \frac{\partial\alpha_1}{\partial\hat{W}_1}\dot{\hat{W}}_1 + \sum_{\rho=1}^{l_1}\left[\frac{\partial\alpha_1}{\partial\hat{v}_{1\rho}}\dot{\hat{v}}_{1\rho}\right] \quad (24)$$

with $\dot{\hat{W}}_1$ and $\dot{\hat{v}}_{1\rho}$ defined in (14) and (15), we obtain

$$\dot{V}_{s2} = -k_1(t)z_1^2 - \psi_1 z_1 + z_1\mathbf{g}_1(x_1)z_2 + z_2[\mathbf{g}_2(\mathbf{x}_2)(z_3 + \alpha_2) + h_2(Z_2)], \quad (25)$$

where

$$h_2(Z_2) = \beta_2(\mathbf{x}_2)f_2(\mathbf{x}_2) + \frac{\dot{\mathbf{x}}_1}{z_2}\int_0^{z_2}\sigma\frac{\partial\beta_2(\mathbf{x}_1,\sigma + \alpha_1)}{\partial\mathbf{x}_1}\,d\sigma - \frac{\dot{\alpha}_1}{z_2}\int_0^{z_2}\beta_2(\mathbf{x}_1,\sigma + \alpha_1)\,d\sigma$$
$$= \beta_2(\mathbf{x}_2)f_2(\mathbf{x}_2) + \dot{\mathbf{x}}_1 z_2\int_0^1\theta\frac{\partial\beta_2(\mathbf{x}_1,\theta z_2 + \alpha_1)}{\partial\mathbf{x}_1}\,d\theta - \dot{\alpha}_1\int_0^1\beta_2(\mathbf{x}_1,\theta z_2 + \alpha_1)\,d\theta$$

with $Z_2 = [\mathbf{x}_2^{\mathrm{T}},\alpha_1,\partial\alpha_1/\partial x_1,\omega_1]^{\mathrm{T}} \in \Omega_{z2} \subset R^5$.

Now, choose the control function

$$\alpha_2 = \frac{1}{\mathbf{g}_2(\mathbf{x}_2)}[-\mathbf{g}_1(x_1)z_1 - k_2(t)z_2 - \hat{W}_2^{\mathrm{T}}S_2(\hat{V}_2^{\mathrm{T}}\mathbf{Z}_2)], \quad (26)$$

where

$$k_2(t) = \frac{1}{\varepsilon_2}\left(1 + \int_0^1\theta\mathbf{g}_2(\mathbf{x}_1,\theta z_2 + \alpha_1)\,d\theta + \|\mathbf{Z}_2^{\mathrm{T}}\hat{W}_2^{\mathrm{T}}\hat{S}_2'\|_F^2 + \|\hat{S}_2'\hat{V}_2^{\mathrm{T}}\mathbf{Z}_2\|^2\right)$$

with constant $\varepsilon_2 > 0$, and network weights are updated by

$$\dot{\hat{W}}_2 = \Gamma_{w2}[(\hat{S}_2 - \hat{S}_2'\hat{V}_2^{\mathrm{T}}\mathbf{Z}_2)z_2 - \sigma_{w2}\hat{W}_2], \quad (27)$$

$$\dot{\hat{V}}_2 = \Gamma_{v2}[\mathbf{Z}_2\hat{W}_2^{\mathrm{T}}\hat{S}_2'z_2 - \sigma_{v2}\hat{V}_2] \quad (28)$$

with $\Gamma_{w2} = \Gamma_{w2}^{\mathrm{T}} > 0$, $\Gamma_{v2} = \Gamma_{v2}^{\mathrm{T}} > 0$, and $\sigma_{w2},\sigma_{v2} > 0$. Using the above controller and noting (18), we obtain

$$\dot{V}_{s2} = -\sum_{j=1}^{2}[k_j(t)z_j^2 + \psi_j z_j] + z_2\mathbf{g}_2(\mathbf{x}_2)z_3.$$

A similar procedure is employed recursively for each step k ($3 \le k \le n-1$). By considering the equation of system (1) for $i = k$, $\dot{x}_k = f_k(\mathbf{x}_k) + g_k(\mathbf{x}_k)x_{k+1}$, and the Lyapunov function candidate

$$V_{sk} = V_{s(k-1)} + \int_0^{z_k}\sigma\beta_k(\mathbf{x}_{k-1},\sigma + \alpha_{k-1})\,d\sigma,$$

we may design the control function α_k, and learning laws for $\hat{W}_k$ and $\hat{V}_k$, which take similar forms of (26), (27) and (28), respectively. The controller u for system (1) shall be constructed in step n.

Step n: Consider $z_n = x_n - \alpha_{n-1}$. We have

$$\dot{z}_n = \dot{x}_n - \dot{\alpha}_{n-1} = f_n(x) + g_n(x)u - \dot{\alpha}_{n-1}.$$

Taking the following Lyapunov function candidate:

$$V_{sn} = V_{s(n-1)} + \int_0^{z_n}\sigma\beta_n(\mathbf{x}_{n-1},\sigma + \alpha_{n-1})\,d\sigma, \quad (29)$$

its time derivative is

$$\dot{V}_{sn} = \dot{V}_{s(n-1)} + z_n\beta_n(\mathbf{x}_n)\dot{z}_n + \int_0^{z_n}\sigma\left[\frac{\partial\beta_n(\mathbf{x}_{n-1},\sigma + \alpha_{n-1})}{\partial\mathbf{x}_{n-1}}\dot{\mathbf{x}}_{n-1} + \frac{\partial\beta_n(\mathbf{x}_{n-1},\sigma + \alpha_{n-1})}{\partial\alpha_{n-1}}\dot{\alpha}_{n-1}\right]d\sigma.$$

Noting (27) and using the similar way as (24) in Step 2, we have

$$\dot{\alpha}_{n-1} = \frac{\partial\alpha_{n-1}}{\partial\mathbf{x}_j}\dot{\mathbf{x}}_j + \omega_{n-1} = \sum_{j=1}^{n-1}\left\{\frac{\partial\alpha_{n-1}}{\partial x_j}[f_j(\mathbf{x}_j) + g_j(\mathbf{x}_j)x_{j+1}]\right\} + \omega_{n-1},$$

where

$$\omega_{n-1} = \sum_{j=1}^{n-1} \left(\frac{\partial \alpha_{n-1}}{\partial \mathbf{x}_{d(j+1)}} \dot{\mathbf{x}}_{d(j+1)} + \frac{\partial \alpha_{n-1}}{\partial \hat{W}_j} \dot{\hat{W}}_j + \sum_{\rho=1}^{l_j} \frac{\partial \alpha_{n-1}}{\partial \hat{v}_{j,\rho}} \dot{\hat{v}}_{j,\rho} \right)$$

with $\dot{\hat{W}}_j$ and $\dot{\hat{v}}_{j,\rho}$ for $j = 1,2,\ldots,n-1$ being designed in the previous $n-1$ steps. Following the procedure of (23)–(25) in Step 2, we obtain

$$\dot{V}_{sn} = -\sum_{j=1}^{n-1} [k_j(t) z_j^2 + \psi_j z_j] + z_{n-1} \mathbf{g}_{n-1}(\mathbf{x}_{n-1}) z_n + z_n[\mathbf{g}_n(x)u + h_n(Z_n)], \quad (30)$$

where

$$h_n(Z_n) = \beta_n(\mathbf{x}_n) f_n(\mathbf{x}_n) + z_n \int_0^1 \theta \frac{\partial \beta_n(\mathbf{x}_{n-1}, \theta z_n + \alpha_{n-1})}{\partial \mathbf{x}_{n-1}} \dot{\mathbf{x}}_{n-1} \, d\theta - \dot{\alpha}_{n-1} \int_0^1 \beta_n(\mathbf{x}_{n-1}, \theta z_n + \alpha_{n-1}) \, d\theta$$

$$Z_n = \left[\mathbf{x}_n^T, \alpha_{n-1}, \frac{\partial \alpha_{n-1}}{\partial x_1}, \frac{\partial \alpha_{n-1}}{\partial x_2}, \ldots, \frac{\partial \alpha_{n-1}}{\partial x_{n-1}}, \omega_{n-1} \right]^T \in \Omega_{zn} \subset R^{2n+1}.$$

Now, we are ready to choose the controller as

$$u = \frac{1}{\mathbf{g}_n(x)} [-\mathbf{g}_{n-1}(\mathbf{x}_{n-1}) z_{n-1} - k_n(t) z_n - \hat{W}_n^T S_n(\hat{V}_n^T \mathbf{Z}_n)], \quad (31)$$

where

$$k_n(t) = \frac{1}{\varepsilon_n} \left(1 + \int_0^1 \theta \mathbf{g}_n(\mathbf{x}_{n-1}, \theta z_n + \alpha_{n-1}) \, d\theta + \|\mathbf{Z}_n \hat{W}_n^T \hat{S}_n'\|_F^2 + \|\hat{S}_n' \hat{V}_n^T \mathbf{Z}_n\|^2 \right) \quad (32)$$

with constant $\varepsilon_n > 0$, and neural network learning laws are

$$\dot{\hat{W}}_n = \Gamma_{wn}[(\hat{S}_n - \hat{S}_n' \hat{V}_n^T \mathbf{Z}_n) z_n - \sigma_{wn} \hat{W}_n], \quad (33)$$

$$\dot{\hat{V}}_n = \Gamma_{vn}[\mathbf{Z}_n \hat{W}_n^T \hat{S}_n' z_n - \sigma_{vn} \hat{V}_n] \quad (34)$$

with $\Gamma_{wn} = \Gamma_{wn}^T > 0$, $\Gamma_{vn} = \Gamma_{vn}^T > 0$, and $\sigma_{wn}, \sigma_{vn} > 0$. Substituting controller (31) into (30) and using approximation (18), we finally have

$$\dot{V}_{sn} = -\sum_{j=1}^{n} [k_j(t) z_j^2 + \psi_j z_j]. \quad (35)$$

4.2. Stability analysis

It is worth noticing that at each step the following type of positive-definite functions:

$$V_{zi} = \int_0^{z_i} \sigma \beta_i(\mathbf{x}_{i-1}, \sigma + \alpha_{i-1}) \, d\sigma, \quad i = 2,3,\ldots,n \quad (36)$$

has been used, which is the key point of the proposed method. According to Assumption 1, we know that $1 \le \beta_i(\mathbf{x}_{i-1}, \sigma + \alpha_{i-1}) \le \mathbf{g}_i(\mathbf{x}_{i-1}, \sigma + \alpha_{i-1})/g_{i0}$ and the following properties hold:

(i)

$$V_{zi} = z_i^2 \int_0^1 \theta \beta_i(\mathbf{x}_{i-1}, \theta z_i + \alpha_{i-1}) \, d\theta \ge z_i^2 \int_0^1 \theta \, d\theta = \frac{z_i^2}{2}, \quad (37)$$

(ii)

$$V_{zi} = z_i^2 \int_0^1 \theta \beta_i(\mathbf{x}_{i-1}, \theta z_i + \alpha_{i-1}) \, d\theta \le \frac{z_i^2}{g_{i0}} \int_0^1 \theta \mathbf{g}_i(\mathbf{x}_{i-1}, \theta z_i + \alpha_{i-1}) \, d\theta. \quad (38)$$

The stability and control performance of the closed-loop system are given in the following theorem.

Theorem 4.1. *Consider the closed-loop system consisting of strict-feedback system* (1) *satisfying Assumption* 1, *controller* (31) *and the NN weight updating laws* (33) *and* (34). *For bounded initial conditions,*

(i) *all signals in the closed-loop system are bounded, and the vectors Z_j remain in the compact sets*

$$\Omega_{zj} = \left\{ Z_j \,\middle|\, \sum_{i=1}^{n} z_i^2(t) \le C_0,\ \sum_{i=1}^{n} \|\tilde{W}_i\|^2 \le \frac{C_0}{\lambda_{\min}(\Gamma_w^{-1})},\ \sum_{i=1}^{n} \|\tilde{V}_i\|_F^2 \le \frac{C_0}{\lambda_{\min}(\Gamma_v^{-1})},\ \mathbf{x}_{d(j+1)} \in \Omega_{d(j+1)} \right\} \quad (39)$$

with constant $C_0 > 0$, and

(ii) *the following inequalities hold*:

$$\lim_{t\to\infty} \frac{1}{t} \int_0^t z_j^2(\tau) \, d\tau \le \frac{2\varepsilon_j}{1 + g_{j0}} \sum_{i=1}^{n} c_i \quad (40)$$

and

$$\sum_{i=1}^{n} z_i^2(t) \le 2V_s(0) e^{-\lambda_s t} + \frac{2}{\lambda_s} \sum_{i=1}^{n} c_i, \quad \forall t \ge 0 \quad (41)$$

with positive constants c_i, $V_s(0)$ and λ_s.

Proof. (i) For notation simplicity, let $\Gamma_w = \Gamma_{wj}$, $\Gamma_v = \Gamma_{vj}$, $\sigma_w = \sigma_{wj}$ and $\sigma_v = \sigma_{vj}$ for $j = 1,2,\ldots,n$. Consider the Lyapunov function candidate

$$V_s = V_{sn} + \frac{1}{2} \sum_{j=1}^{n} [\tilde{W}_j^T \Gamma_w^{-1} \tilde{W}_j + tr\{\tilde{V}_j^T \Gamma_v^{-1} \tilde{V}_j\}]. \quad (42)$$

By taking its time derivative along (35) and noting (18), it follows that

$$\dot{V}_s = -\sum_{j=1}^{n} [k_j(t)z_j^2 + \tilde{W}_j^{\mathrm{T}}(\hat{S}_j - \hat{S}'_j\hat{V}_j^{\mathrm{T}}\mathbf{Z}_j)z_j + \hat{W}_j^{\mathrm{T}}\hat{S}'_j\tilde{V}_j^{\mathrm{T}}\mathbf{Z}_j z_j + (d_{uj} - \mu_j)z_j - \tilde{W}_j^{\mathrm{T}}\Gamma_w^{-1}\dot{\hat{W}}_j - tr\{\tilde{V}_j^{\mathrm{T}}\Gamma_v^{-1}\dot{\hat{V}}_j\}].$$

Using adaptive tuning laws (14), (15), (27), (28) and (33), (34), and the fact that $\hat{W}_j^{\mathrm{T}}\hat{S}'_j\tilde{V}_j^{\mathrm{T}}\mathbf{Z}_j = tr\{\tilde{V}_j^{\mathrm{T}}\mathbf{Z}_j\hat{W}_j^{\mathrm{T}}\hat{S}'_j\}$, we obtain

$$\dot{V}_s = -\sum_{j=1}^{n} [k_j(t)z_j^2 + (d_{uj} - \mu_j)z_j + \sigma_w\tilde{W}_j^{\mathrm{T}}\hat{W}_j + \sigma_v\, tr\{\tilde{V}_j^{\mathrm{T}}\hat{V}_j\}].$$

Noting (32), (19) and following the same procedure in the proof of Theorem 3.1 (from (A.3)–(A.7)), we further obtain

$$\dot{V}_s \le -\sum_{j=1}^{n}\left\{\frac{z_j^2}{\varepsilon_j}\left[\frac{1}{2} + \int_0^1 \theta\mathbf{g}_j(\mathbf{x}_{j-1},\theta z_j + \alpha_{j-1})\,\mathrm{d}\theta\right] + \frac{\sigma_w}{2}\|\tilde{W}_j\|^2 + \frac{\sigma_v}{2}\|\tilde{V}_j\|_{\mathrm{F}}^2 - c_j\right\}, \tag{43}$$

where $\mathbf{g}_j(\mathbf{x}_{j-1},\theta z_j + \alpha_{j-1}) = \mathbf{g}_1(\theta z_1 + y_{\mathrm{d}})$ for $j = 1$ and

$$c_j = \varepsilon_j\left(\frac{1}{4}\|W_j^*\|^2 + \frac{1}{4}\|V_j^*\|_{\mathrm{F}}^2 + |W_j^*|_1^2 + \boldsymbol{\mu}_j^2\right) + \frac{\sigma_w^2}{2}\|W_j^*\|^2 + \frac{\sigma_v^2}{2}\|V_j^*\|_{\mathrm{F}}^2. \tag{44}$$

Considering (38), (42) and (43), we have

$$\dot{V}_s \le -\lambda_s V_s + \sum_{j=1}^{n} c_j, \tag{45}$$

where $\lambda_s = \min\{g_{10}/\varepsilon_1,\ g_{20}/\varepsilon_2, \dots, g_{n0}/\varepsilon_n,\ \sigma_w/\lambda_{\max}(\Gamma_w^{-1}),\ \sigma_v/\lambda_{\max}(\Gamma_v^{-1})\}$. It follows from (45) that

$$V_s(t) \le V_s(0)\mathrm{e}^{-\lambda_s t} + \frac{1}{\lambda_s}\sum_{j=1}^{n} c_j, \quad \forall t \ge 0, \tag{46}$$

where constant

$$V_s(0) = \sum_{j=1}^{n}\left\{\int_0^{z_j(0)} \sigma\beta_j(\mathbf{x}_{j-1}(0),\sigma + \alpha_{j-1}(0))\,\mathrm{d}\sigma + \frac{1}{2}[\tilde{W}_j^{\mathrm{T}}(0)\Gamma_w^{-1}\tilde{W}_j(0) + tr\{\tilde{V}_j^{\mathrm{T}}(0)\Gamma_v^{-1}\tilde{V}_j(0)\}]\right\} \tag{47}$$

with $\beta_j(\mathbf{x}_{j-1}(0),\sigma + \alpha_{j-1}(0)) = \beta_1(\sigma + y_{\mathrm{d}}(0))$ for $j = 1$. Considering (42), we know that

$$\sum_{j=1}^{n}\|\tilde{W}_j\|^2 \le \frac{2V_s(t)}{\lambda_{\min}(\Gamma_w^{-1})}, \qquad \sum_{j=1}^{n}\|\tilde{V}_j\|_{\mathrm{F}}^2 \le \frac{2V_s(t)}{\lambda_{\min}(\Gamma_v^{-1})}. \tag{48}$$

It follows from (37) and (42) that

$$V_s(t) \ge V_{sn} = \sum_{j=1}^{n} V_{zj} \ge \frac{1}{2}\sum_{j=1}^{n} z_j^2(t). \tag{49}$$

Inequalities (46), (48) and (49) confirm that all the signals x, z_j, $\hat{W}_j$ and $\hat{V}_j$ in the closed-loop system are bounded. Let $C_0 = 2V_s(0) + (2/\lambda_s)\sum_{i=1}^{n} c_i$, we conclude that there do exist compact sets Ω_{zj} defined in (39) such that the vectors $Z_j \in \Omega_{zj}$ for all time.

(ii) It can be seen from $\mathbf{g}_j(\mathbf{x}_j) \ge g_{j0}$ (Assumption 1) that $\int_0^1 \theta\mathbf{g}_j(\mathbf{x}_{j-1},\theta z_j + \alpha_{j-1})\,\mathrm{d}\theta \ge g_{j0}/2$. Then inequality (43) may be re-written as

$$\dot{V}_s \le -\sum_{j=1}^{n}\frac{z_j^2}{2\varepsilon_j}(1 + g_{j0}) + \sum_{j=1}^{n} c_j.$$

Integrating the above inequality over $[0,t]$ leads to

$$\int_0^t z_j^2(\tau)\,\mathrm{d}\tau \le \frac{2\varepsilon_j}{1 + g_{j0}}\left[V_s(0) + t\sum_{i=1}^{n} c_i\right], \quad j = 1,2,\dots,n \tag{50}$$

which implies that inequality (40) holds. The error bound (41) can be derived from (46) and (49) directly. □

Remark 4.1. Since the function approximation property (3) of neural networks is only guaranteed within a compact set, the stability result proposed in this work is semi-global in the sense that, for any compact set, there exists a controller with sufficiently large number of NN nodes such that all the closed-loop signals are bounded when the initial states are within this compact set. In practical applications, the NN node number usually cannot be chosen too large due to the possible computation problem. This implies that the NN approximation capability is limited, and some constraints on Ω_{zj} are necessary to guarantee the NN approximation. It has been shown in (39) that the sizes of Ω_{zj} depend on $x(0), V_j(0), W_j(0), W_j^*, V_j^*, \mu_j$, and all design parameters. Since there is no analytical result in the NN literature to quantify the relationship of l_j, W_j^*, V_j^*, μ_j, an explicit expression of the stability condition is not available at the present stage. Some suggestions are given to guide the choices of initial conditions and control parameters: (i) small initial errors $z_j(0)$ and large learning gains Γ_w and Γ_v will help to reduce $V_s(0)$, subsequently lead to small Ω_{zj}, (ii) decreasing ε_j, σ_w and σ_v help to reduce Ω_{zj}, and (iii) large the NN node number l_j reduces $\boldsymbol{\mu}_j$ in (44), and improves both stability and performance of the adaptive system.

5. Simulation studies

The following three second-order plants are considered in the simulation:

$$\Sigma_1: \quad \begin{aligned} \dot{x}_1 &= x_2, \\ \dot{x}_2 &= (1 - x_1^2)x_2 - x_1 + (1 + x_1^2 + x_2^2)u, \end{aligned}$$

$$\Sigma_2: \quad \begin{aligned} \dot{x}_1 &= x_1 e^{-0.5x_1} + (1 + x_1^2)x_2, \\ \dot{x}_2 &= x_1 x_2^2 + [3 + \cos(x_1 x_2)]u, \end{aligned}$$

$$\Sigma_3: \quad \begin{aligned} \dot{x}_1 &= 0.5x_1^3 + \ln(10 + x_1^2)x_2, \\ \dot{x}_2 &= \frac{x_1 x_2}{1 + x_1^2 + x_2^2} + (1 + e^{-x_1^2 - x_2^2})u \end{aligned}$$

with the output $y = x_1$. Clearly, Plants Σ_1–Σ_3 are in strict-feedback form and satisfy Assumption 1. For simplifying the controller design and easily comparing the control performance, we choose initial condition $[x_1(0), x_2(0)]^T = [0, 0]^T$, and $\mathbf{g}_i(\mathbf{x}_i) = 1$ for all controllers. It should be pointed out that the choice of $\mathbf{g}_i(\mathbf{x}_i) = 1$ may violate Assumption 1. In this simulation, we are intended to relax Assumption 1 to check the robustness of the proposed controller. In fact, Assumption 1 is a sufficient condition, not a necessary one, there are some tolerance in the choice of $\mathbf{g}_i(\mathbf{x}_i)$ under certain conditions.

As plants Σ_1–Σ_3 are second-order systems, according to (31) the adaptive NN controller shall be chosen as follows:

$$u = -z_1 - k_2(t)z_2 - \hat{W}_2^T S_2(\hat{V}_2^T \mathbf{Z}_2), \tag{51}$$

where

$$z_1 = x_1 - y_d, z_2 = x_2 - \alpha_1$$

and

$$\mathbf{Z}_2 = [x_1, x_2, \alpha_1, \partial\alpha_1/\partial x_1, \omega_1, 1]^T$$

with

$$\alpha_1 = -k_1(t)z_1 - \hat{W}_1^T S_1(\hat{V}_1^T \mathbf{Z}_1), \quad \mathbf{Z}_1 = [x_1, y_d, \dot{y}_d, 1]^T,$$

$$\omega_1 = \frac{\partial\alpha_1}{\partial y_d}\dot{y}_d + \frac{\partial\alpha_1}{\partial \dot{y}_d}\ddot{y}_d + \frac{\partial\alpha_1}{\partial \hat{W}_1}\dot{\hat{W}}_1 + \sum_{\rho=1}^{l_1}\left[\frac{\partial\alpha_1}{\partial \hat{v}_{1\rho}}\dot{\hat{v}}_{1\rho}\right],$$

$$k_j(t) = \frac{1}{\varepsilon_j}\left(\tfrac{3}{2} + \|\mathbf{Z}_j \hat{W}_j^T \hat{S}_j'\|_F^2 + \|\hat{S}_j' \hat{V}_j^T \mathbf{Z}_j\|^2\right), \quad j = 1, 2$$

and NN weights $\hat{W}_1$, $\hat{V}_1$, $\hat{W}_2$, and $\hat{V}_2$ are updated by (14), (15) and (27), (28) correspondingly. Both neural networks $\hat{W}_j^T S_j(\hat{V}_j^T \mathbf{Z}_j)$, $j = 1, 2$ contain 10 hidden nodes (i.e., $l_1 = l_2 = 10$) and the coefficient in activation function $s(\cdot)$ is taken as $\gamma = 3.0$. The design parameters of the above controller are $\varepsilon_1 = 1.0$, $\varepsilon_2 = 5.0$, $\Gamma_{w1} = \Gamma_{w2} = diag\{1.0\}$, $\Gamma_{v1} = \Gamma_{v2} = diag\{10.0\}$, $\sigma_{w1} = \sigma_{w2} = 1 \times 10^{-2}$, $\sigma_{v1} = 1 \times 10^{-4}$ and $\sigma_{v2} = 1 \times 10^{-3}$. The initial weights $\hat{W}_1(0) = 0.0$, $\hat{W}_2(0) = 0.0$, and the elements of $\hat{V}_1(0)$ and $\hat{V}_2(0)$ are taken randomly in the interval $[-0.5, 0.5]$. To show the learning ability of the proposed scheme, without modifying any controller parameters, we apply controller (51) to three plants Σ_1–Σ_3 whose dynamics and control objectives are quite different.

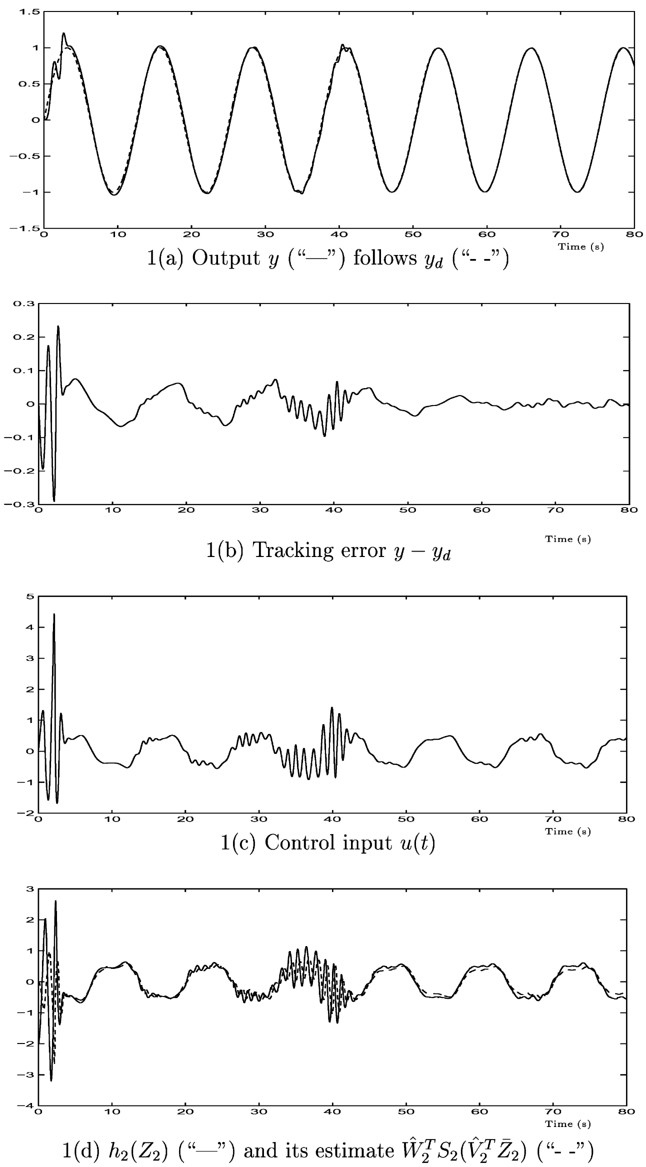


Fig. 1. Simulation results for Plant Σ_1 with $y_d = \sin(0.5t)$.

Fig. 1 shows the simulation results for plant Σ_1 (which is known as Van der Pol system) with the reference signal $y_d = \sin(0.5t)$. From Figs. 1(a) and (b), we can see that a large tracking error exists during the first 5 seconds. This is due to the lack of knowledge about the plant nonlinearities and the initial weights are chosen randomly. Through the NN learning phase, better tracking performance is obtained after 10 s. To illustrate the learning ability of neural networks, the nonlinear function $h_2(Z_2)$

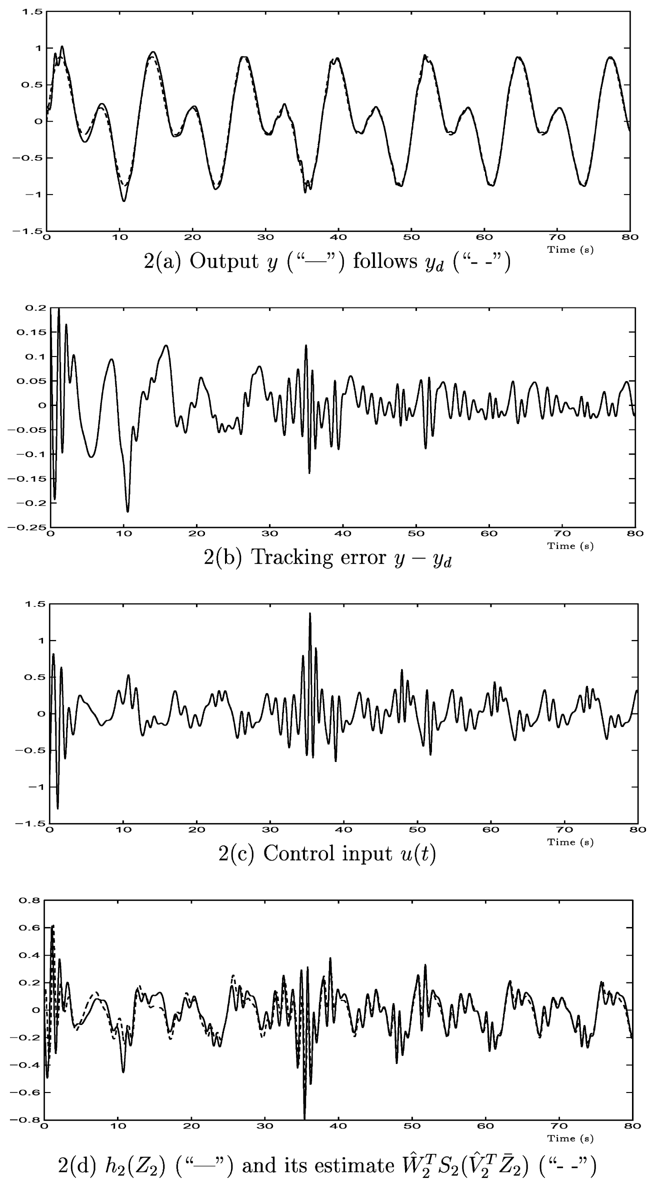


Fig. 2. Simulation results for Plant Σ_2 with $y_\mathrm{d} = 0.5[\sin(t) + \sin(0.5t)]$.

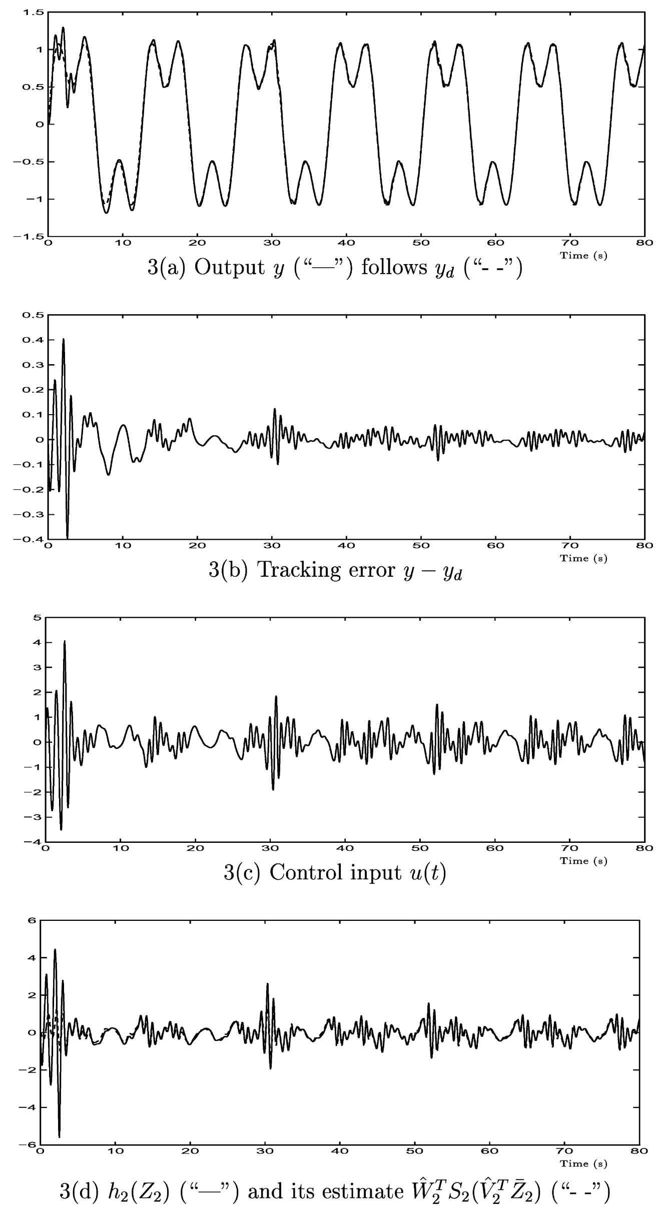


Fig. 3. Simulation results for Plant Σ_3 with $y_\mathrm{d} = \sin(0.5t) + 0.5\sin(1.5t)$.

and its estimate $\hat{W}_2^\mathrm{T} S_2(\hat{V}_2^\mathrm{T} \mathbf{Z}_2)$ are shown in Fig. 1(d). It can be seen that they match very well after 40 s. Hence, the proposed adaptive controller possesses the abilities of learning and controlling the unknown nonlinear system.

Keeping all design parameters as before, we apply controller (51) to plant Σ_2 for tracking a desired signal $y_\mathrm{d} = 0.5[\sin(t) + \sin(0.5t)]$. Again, the initial response shown in Fig. 2(a) is not satisfactory, and a small tracking error provided in Fig. 2(b) is achieved after several learning periods. Figs. 2(c) and (d) indicate the boundedness of control input and the responses of $h_2(Z_2)$ and $\hat{W}_2^\mathrm{T} S_2(\hat{V}_2^\mathrm{T} \mathbf{Z}_2)$. Finally, the same adaptive controller is used to control Plant Σ_3 for forcing the output y track the reference $y_\mathrm{d} = \sin(0.5t) + 0.5\sin(1.5t)$. Fig. 3 shows the simulation results and confirms the effectiveness of the developed method.

Although the dynamics, operation regions and the reference signals of three plants Σ_1–Σ_3 are quite different in the above experiment tests, the simulation results indicate that the developed neural controller has strong adaptability and achieves good control performance for all three nonlinear plants. It is interesting to notice that adaptive NN controller (51) is highly structural, and independent of the complexities of the system nonlinearities. Such a structural property is particularly suitable for parallel processing and hardware implementation in practical applications. In addition, the controller, once configurated, can be applied to other similar plants without repeating the complex controller design procedure for different system nonlinearities.

6. Conclusion

In this paper, we have presented an adaptive NN control scheme for strict-feedback nonlinear systems using backstepping design. The main feature of the proposed approach is the application of the novel Lyapunov functions to construct the NN-based backstepping adaptive controller. Semi-global stability results are obtained and the tracking error converges to a small residual set which is adjustable by tuning the design parameters. Simulation experiments have been given to demonstrate the theoretical analysis and the learning ability of neural networks used in the controller. Further investigation can be directed to the robustness of the scheme under external disturbances and the extension to other types of nonlinear systems.

Appendix

Proof of Theorem 3.1. The proof includes two parts. We first suppose that there exists a compact set Ω_{z1} such that $Z_1 \in \Omega_{z1}$, $\forall t \geq 0$, and the following function approximation holds:

$$h_1(Z_1) = W_1^{*\mathrm{T}} S_1(V_1^{*\mathrm{T}} \mathbf{Z}_1) + \mu_1, \quad \forall Z_1 \in \Omega_{z1}, \tag{A.1}$$

where W_1^* and V_1^* are the ideal constant weights and $|\mu_1| \leq \boldsymbol{\mu}_1$ with constant $\boldsymbol{\mu}_1 > 0$. Then, we show that this compact set Ω_{z1} do exist for bounded initial conditions.

Consider the Lyapunov function candidate

$$V_0 = V_{z1} + \tfrac{1}{2}[\tilde{W}_1^{\mathrm{T}} \Gamma_{w1}^{-1} \tilde{W}_1 + tr\{\tilde{V}_1^{\mathrm{T}} \Gamma_{v1}^{-1} \tilde{V}_1\}], \tag{A.2}$$

where $\tilde{W}_1 = \hat{W}_1 - W_1^*$, $\tilde{V}_1 = \hat{V}_1 - V_1^*$ and V_{z1} is defined by (7). Taking its time derivative along (11) and noting (12), we have

$$\dot{V}_0 = z_1[-k_1(t) z_1 - \hat{W}_1^{\mathrm{T}} S_1(\hat{V}_1^{\mathrm{T}} \mathbf{Z}_1) + h_1(Z_1)]$$
$$+ \tilde{W}_1^{\mathrm{T}} \Gamma_{w1}^{-1} \dot{\hat{W}}_1 + tr\{\tilde{V}_1^{\mathrm{T}} \Gamma_{v1}^{-1} \dot{\hat{V}}_1\}.$$

Using (A.1) and Lemma 2.1, we obtain

$$\dot{V}_0 = z_1[-k_1(t) z_1 - \tilde{W}_1^{\mathrm{T}}(\hat{S}_1 - \hat{S}'_1 \hat{V}_1^{\mathrm{T}} \mathbf{Z}_1)$$
$$- \hat{W}_1^{\mathrm{T}} \hat{S}'_1 \tilde{V}_1^{\mathrm{T}} \mathbf{Z}_1 - d_{u1} + \mu_1]$$
$$+ \tilde{W}_1^{\mathrm{T}} \Gamma_{w1}^{-1} \dot{\hat{W}}_1 + tr\{\tilde{V}_1^{\mathrm{T}} \Gamma_{v1}^{-1} \dot{\hat{V}}_1\}.$$

Considering adaptive laws (14) and (15), and the fact that $\hat{W}_1^{\mathrm{T}} \hat{S}'_1 \tilde{V}_1^{\mathrm{T}} \mathbf{Z}_1 = tr\{\tilde{V}_1^{\mathrm{T}} \mathbf{Z}_1 \hat{W}_1^{\mathrm{T}} \hat{S}'_1\}$, the above equation can be further written as

$$\dot{V}_0 = -k_1(t) z_1^2 + (\mu_1 - d_{u1}) z_1 - \sigma_{w1} \tilde{W}_1^{\mathrm{T}} \hat{W}_1$$
$$- \sigma_{v1}\, tr\{\tilde{V}_1^{\mathrm{T}} \hat{V}_1\}. \tag{A.3}$$

By noting (5) in Lemma 2.1, (13) and the properties that $2\tilde{W}_1^{\mathrm{T}} \hat{W}_1 \geq \|\tilde{W}_1\|^2 - \|W_1^*\|^2$ and $2\,tr\{\tilde{V}_1^{\mathrm{T}} \hat{V}_1\} \geq \|\tilde{V}_1\|_{\mathrm{F}}^2 - \|V_1^*\|_{\mathrm{F}}^2$, the following inequality follows:

$$\dot{V}_0 \leq -\frac{z_1^2}{\varepsilon_1}\left(1 + \int_0^1 \theta \mathbf{g}_1(\theta z_1 + y_{\mathrm{d}})\,\mathrm{d}\theta\right.$$
$$\left. + \|\mathbf{Z}_1 \hat{W}_1^{\mathrm{T}} \hat{S}'_1\|_{\mathrm{F}}^2 + \|\hat{S}'_1 \hat{V}_1^{\mathrm{T}} \mathbf{Z}_1\|^2\right)$$
$$+ (\|V_1^*\|_{\mathrm{F}} \|\mathbf{Z}_1 \hat{W}_1^{\mathrm{T}} \hat{S}'_1\|_{\mathrm{F}} + \|W_1^*\| \|\hat{S}'_1 \hat{V}_1^{\mathrm{T}} \mathbf{Z}_1\|$$
$$+ |W_1^*|_1 + |\mu_1|)|z_1| - \frac{\sigma_{w1}}{2}(\|\tilde{W}_1\|^2 - \|W_1^*\|^2)$$
$$- \frac{\sigma_{v1}}{2}(\|\tilde{V}_1\|_{\mathrm{F}}^2 - \|V_1^*\|_{\mathrm{F}}^2). \tag{A.4}$$

Since

$$\|V_1^*\|_{\mathrm{F}} \|\mathbf{Z}_1 \hat{W}_1^{\mathrm{T}} \hat{S}'_1\|_{\mathrm{F}} |z_1| \leq \frac{z_1^2}{\varepsilon_1} \|\mathbf{Z}_1 \hat{W}_1^{\mathrm{T}} \hat{S}'_1\|_{\mathrm{F}}^2 + \frac{\varepsilon_1}{4} \|V_1^*\|_{\mathrm{F}}^2, \tag{A.5}$$

$$(|W_1^*|_1 + |\mu_1|)|z_1| \leq \frac{z_1^2}{2\varepsilon_1} + \varepsilon_1(|W_1^*|_1^2 + \mu_1^2),$$

$$\|W_1^*\| \|\hat{S}'_1 \hat{V}_1^{\mathrm{T}} \mathbf{Z}_1\| |z_1| \leq \frac{z_1^2}{\varepsilon_1} \|\hat{S}'_1 \hat{V}_1^{\mathrm{T}} \mathbf{Z}_1\|^2 + \frac{\varepsilon_1}{4} \|W_1^*\|^2 \tag{A.6}$$

and $|\mu_1| \leq \boldsymbol{\mu}_1$, inequality (A.4) can be re-written as

$$\dot{V}_0 \leq -\frac{z_1^2}{\varepsilon_1}\left[\frac{1}{2} + \int_0^1 \theta \mathbf{g}_1(\theta z_1 + y_{\mathrm{d}})\,\mathrm{d}\theta\right] - \frac{\sigma_{w1}}{2}\|\tilde{W}_1\|^2$$
$$- \frac{\sigma_{v1}}{2}\|\tilde{V}_1\|_{\mathrm{F}}^2 + c_1 \tag{A.7}$$

with constant

$$c_1 = \varepsilon_1\left(\frac{1}{4}\|W_1^*\|^2 + \frac{1}{4}\|V_1^*\|_{\mathrm{F}}^2 + |W_1^*|_1^2 + \mu_1^2\right)$$
$$+ \frac{\sigma_{w1}^2}{2}\|W_1^*\|^2 + \frac{\sigma_{v1}^2}{2}\|V_1^*\|_{\mathrm{F}}^2.$$

From (8) and (A.7), we further have

$$\dot{V}_0 \leq -\frac{g_{10}}{\varepsilon_1} V_{z1} - \frac{\sigma_{w1}}{2}\|\tilde{W}_1\|^2 - \frac{\sigma_{v1}}{2}\|\tilde{V}_1\|_{\mathrm{F}}^2 + c_1.$$

Therefore,

$$\dot{V}_0 \leq -\lambda_1 V_0 + c_1, \tag{A.8}$$

where constant $\lambda_1 = \min\{g_{10}/\varepsilon_1, \sigma_{w1}/\lambda_{\max}(\Gamma_{w1}^{-1}), \sigma_{v1}/\lambda_{\max}(\Gamma_{v1}^{-1})\}$. Solving inequality (A.8), we obtain

$$V_0(t) \le V_0(0)e^{-\lambda_1 t} + c_1 \int_0^t e^{-\lambda_1 (t-\tau)} d\tau$$

$$\le c_0 e^{-\lambda_1 t} + \frac{c_1}{\lambda_1}, \quad \forall t \ge 0, \tag{A.9}$$

where positive constant $c_0 = V_0(0)$. It follows from (8), (A.2) and (A.10) that

$$\frac{z_1^2(t)}{2} \le V_{z1}(t) \le V_0(t) \le c_0 e^{-\lambda_1 t} + \frac{c_1}{\lambda_1}, \quad \forall t \ge 0. \tag{A.10}$$

This confirms that for bounded initial conditions, all signals z_1, $\hat{W}_1$ and $\hat{V}_1$ of the closed-loop system are bounded and there do exist a compact set Ω_{z1} such that $Z_1 \in \Omega_{z1}$ for all time. □

References

Chen, F. C., & Liu, C. C. (1994). Adaptively controlling nonlinear continuous-time systems using multilayer neural networks. *IEEE Transactions on Automatic Control*, *AC-39*, 1306–1310.

Fabri, S., & Kadirkamanathan, V. (1996). Dynamic structure neural networks for stable adaptive control of nonlinear systems. *IEEE Transactions on Neural Networks*, *7*, 1151–1167.

Funahashi, K. I. (1989). On the approximate realization of continuous mappings by neural networks. *Neural Networks*, *2*, 183–192.

Ge, S. S., Lee, T. H., & Harris, C. J. (1998). *Adaptive neural network control of robotic manipulators*. London: World Scientific.

Ge, S. S., Hang, C. C., & Zhang, T. (1999a). Nonlinear adaptive control using neural networks and its application to CSTR systems. *Journal of Process Control*, *9*(4), 313–323.

Ge, S. S., Hang, C. C., & Zhang, T. (1999b). Adaptive neural network control of nonlinear systems by state and output feedback. *IEEE Transactions on Systems, Man, and Cybernetics (Part B)*, *29*, 818–828.

Ge, S. S., Wang, C., & Lee, T. H. (2000). Adaptive backstepping control of a class of chaotic systems. *International Journal of Bifurcation and Chaos*, *10*(5) 1149–1156.

Jin, L., Nikiforuk, P. N., & Gupta, M. M. (1993). Direct adaptive output tracking control using multilayered neural networks. *IEE Proceedings D*, *140*, 393–398.

Kanellakopoulos, I., Kokotovic, P. V., & Marino, R. (1989). Robustness of adaptive nonlinear control under an extended matching condition. In *Proceedings of IFAC symposium on nonlinear control systems and design*, Capri, Italy (pp. 192–197).

Kanellakopoulos, I., Kokotovic, P. V., & Morse, A. S. (1991). Systematic design of adaptive controller for feedback linearizable systems. *IEEE Transactions on Automatic Control*, *AC-36*, 1241–1253.

Kosmatopoulos, E. B. (1996). Universal stabilization using control Lyapunov functions adaptive derivative feedback and neural network approximators. In *Proceedings of 35th CDC*, Kobe, Japan, December (pp. 2444–2449).

Krstic, M., Kanellakopoulos, I., & Kokotovic, P. V. (1995). *Nonlinear and adaptive control design*. New York: Wiley.

Kwan, C., & Lewis, F. L. (1995a). Robust backstepping control of induction motors using neural networks. *Proceedings of IEEE Mediterranean symposium on new directions in control and automation*, Cyptus (pp. 323–330).

Kwan, C., & Lewis, F. L. (1995b). Robust backstepping control of nonlinear systems using neural networks. *Proceedings of European control conference*, Rome (pp. 2772–2777).

Kwan, C., Lewis, F. L., & Dawson, D. M. (1998). Robust neural-network control of rigid-link electrically driven robots. *IEEE Transactions on Neural Networks*, *9*, 581–588.

Lewis, F. L., Yesildirek, A., & Liu, K. (1996). Multilayer neural net robot controller with guaranteed tracking performance. *IEEE Transactions on Neural Networks*, *7*, 388–398.

Nam, K., & Arapostations, A. (1988). A model-reference adaptive control scheme for pure-feedback nonlinear systems. *IEEE Transactions on Automatic Control*, *AC-33*, 803–811.

Narendra, K. S., & Parthasarathy, K. (1990). Identification and control of dynamic systems using neural networks. *IEEE Transactions on Neural Networks*, *1*, 4–27.

Pomet, J. B., & Praly, L. (1992). Adaptive nonlinear regulation: Estimation from the Lyapunov equation. *IEEE Transactions on Automatic Control*, *AC-37*, 729–740.

Polycarpou, M. M. (1996). Stable adaptive neural control scheme for nonlinear systems. *IEEE Transactions on Automatic Control*, *AC-41*, 447–451.

Polycarpou, M. M., & Ioannou, P. A. (1991). *Identification and control using neural network models*: *Design and stability analysis*. Technical Report 91-09-01, Department of Electrical Engineering Systems, University of Southern California.

Polycarpou, M. M., & Ioannou, P. A. (1996). A robust adaptive nonlinear control design. *Automatica*, *32*, 423–427.

Sanner, R. M., & Slotine, J. E. (1992). Gaussian networks for direct adaptive control. *IEEE Transactions on Neural Networks*, *3*, 837–863.

Sastry, S. S., & Isidori, A. (1989). Adaptive control of linearizable systems. *IEEE Transactions on Automatic Control*, *AC-34*, 1123–1131.

Sepulchre, R., Jankovic, M., & Kokotovic, P. V. (1997). *Constructive nonlinear control*. London: Springer.

Seto, D., Annaswamy, A. M., & Baillieul, J. (1994). Adaptive control of nonlinear systems with a triangular structure. *IEEE Transactions on Automatic Control*, *AC-39*, 1411–1428.

Spooner, J. T., & Passino, K. M. (1996). Stable adaptive control using fuzzy systems and neural networks. *IEEE Transactions on Fuzzy Systems*, *4*, 339–359.

Taylor, D. G., Kokotovic, P. V., Marino, R., & Kanellakopoulos, I. (1989). Adaptive regulation of nonlinear systems with unmodeled dynamics. *IEEE Transactions on Automatic Control*, *AC-34*, 405–412.

Teel, A. R., Kadiyala, R. R., Kokotovic, P. V., & Sastry, S. S. (1991). Indirect techniques for adaptive input-output linearization of nonlinear systems. *International Journal of control*, *53*, 193–222.

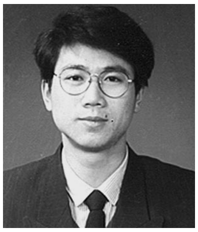
Wang, L. X. (1994). *Adaptive fuzzy systems and control: Design and analysis*. Englewood Cliff, NJ: Prentice-Hall.

Yao, B., & Tomizuka, M. (1997). Adaptive robust control of SISO nonlinear systems in a semi-strict feedback form. *Automatica*, *33*, 893–900.

Yesidirek, A., & Lewis, F. L. (1995). Feedback linearization using neural networks. *Automatica*, *31*, 1659–1664.

Zhang, T., Ge, S. S., & Hang, C. C. (1999). Design and performance analysis of a direct adaptive controller for nonlinear systems. *Automatica*, *35*, 1809–1817.

Zhang, T., Ge, S. S., & Hang, C. C. (2000). Stable adaptive control for a class of nonlinear systems using a modified Lyapunov function. *IEEE Transactions on Automatic Control*, *AC-45*, 129–132.



Tao Zhang received the B.Eng. and the M.Eng. in the Department of Automatic Control in 1990 and 1993, respectively, both from Northeastern University, P.R. China. He received the Ph.D. degree in the Department of Electrical Engineering, the National University of Singapore in 2000. He has been with Seagate Technology International from 1999. His technical paper "Adaptive neural network control for strict-feedback nonlinear systems using backstepping design" co-authored with S. S. Ge and C. C. Hang was selected in the finalist for the Best Student Paper Award of the 1999 American Control Conference. His research interests include adaptive nonlinear control, robust adaptive control, neural network control, disk drive control, PID auto-tuning and control applications.

S. S. Ge received the B.Sc. degree in Control Engineering from Beijing University of Aeronautics and Astronautics, China, in July 1986, and the Ph.D. degree and the Diploma of Imperial College (DIC) in Mechanical/Electrical Engineering from Imperial College of Science, Technology and Medicine, University of London, United Kingdom, in January 1993. From May 1992 to June 1993, he was a postdoctoral research associate at Leicester University, United Kingdom. He has been with the Department of Electrical Engineering, the National University of Singapore as a lecturer from July 1993 to June 1998, a Senior Lecturer from July 1998 to June 2000, and an Associate Professor from July 2000. He was a visiting staff in Laboratoire de'Automatique de Grenoble, France in 1996 and Department of Electrical and Electronics Engineering, the University of Melbourne, Australia in 1998 and 1999. He served as an Associate Editor on the Conference Editorial Board of the IEEE Control Systems Society in 1998 and 1999. He currently serves as an Associate Editor, IEEE Transactions on Control Systems Technology. He was the winner of the 1999 National Technology Award of National Science & Technology Board, Singapore. He serves as a technical consultant to local industry.

His current research interests are Adaptive Control, Neural Networks and Fuzzy Logic, Robot Control, Real-Time Implementation, Genetic Algorithms, Friction Compensation and Sensor Fusion. He has authored and co-authored over 100 international journal and conference papers, one monograph and co-invented two patents.

C. C. Hang graduated with a First Class Honours Degree in Electrical Engineering from the University of Singapore in 1970. He received the Ph.D. degree in control engineering from the University of Warwick, England, in 1973. From 1974 to 1977, he worked as a Computer and Systems Technologist in the Shell Eastern Petroleum Company (Singapore) and the Shell International Petroleum Company (The Netherlands). Since 1977, he has been with the National University of Singapore, serving in various positions including being the Vice-Dean of the Faculty of Engineering and Head of the Department of Electrical Engineering. Since October 1994, he has been appointed Deputy Vice-Chancellor. His major area of research is adaptive control in which he has published one book, 170 international journal and conference papers and 4 patents. He was a Visiting Scientist in Yale University in 1983, and in Lund University in 1987 and 1992. Since March 1992, he has been appointed Principal Editor (Adaptive Control) of the Automatica Journal. In January 1998, he was elected a Fellow of IEEE.