PERGAMON

Available online at www.sciencedirect.com

SCIENCE DIRECT®

Automatica 39 (2003) 807–819

automatica

www.elsevier.com/locate/automatica

Adaptive NN control for a class of strict-feedback discrete-time nonlinear systems☆

S.S. Ge*, G.Y. Li, T.H. Lee

Department of Electrical and Computer Engineering, National University of Singapore, 117576 Singapore

Received 12 January 2001; received in revised form 3 May 2002; accepted 22 January 2003

Abstract

In this paper, both full state and output feedback adaptive neural network (NN) controllers are presented for a class of strict-feedback discrete-time nonlinear systems. Firstly, Lyapunov-based full-state adaptive NN control is presented via backstepping, which avoids the possible controller singularity problem in adaptive nonlinear control and solves the noncausal problem in the discrete-time backstepping design procedure. After the strict-feedback form is transformed into a cascade form, another relatively simple Lyapunov-based direct output feedback control is developed. The closed-loop systems for both control schemes are proven to be semi-globally uniformly ultimately bounded.

Keywords: Adaptive control; Neural networks; Discrete time; Backstepping

1. Introduction

Active research has been carried out in neural network (NN) control by using the fact that NN can approximate a wide range of nonlinear functions to any desired degree of accuracy under certain conditions. It was shown that for stable and efficient on-line control using the backpropagation (BP) learning algorithm, the identification must be sufficiently accurate before control action is initiated (Hunt, Sbarbaro, Zbikowski, & Gawthrop, 1992; Levin & Narendra, 1996; Narendra & Parthasarathy, 1990). Recently, several good NN control approaches have been proposed based on Lyapunov's stability theory (Lewis, Yesildirek, & Liu, 1996; Polycarpou, 1996; Yesidirek & Lewis, 1995; Ge, Lee, & Harris, 1998; Ge, Hang, Lee, & Zhang, 2001). One main advantage of these schemes is that the adaptive laws are derived based on the Lyapunov synthesis method and therefore guarantee the stability of continuous-time systems without the requirement of off-line training.

Adaptive control for strict-feedback nonlinear systems is still an active topic of research. Using backstepping, systematic approaches for adaptive controller design have been presented for a class of nonlinear systems transformable to a parametric strict-feedback canonical form, and guarantee the closed-loop stability (Ge et al., 2001; Kanellakopoulos, Kokotovic, & Morse, 1991; Krstic, Kanellakopoulos, & Kokotovic, 1995; Seto, Annaswamy, & Baillieul, 1994). However, all these elegant methods in continuous-time domain are not directly applicable to discrete-time systems due to the noncausal problem in the controller design procedure via backstepping, as detailed in Section 4. Recently, for discrete-time systems transformable to the parametric-strict-feedback form and the parametric-pure-feedback form, the noncausal problem was elegantly solved in backstepping using a time-varying mapping (Yeh & Kokotovic, 1995), which was further extended to cases with time-varying parameters and nonparametric uncertainties in Zhang, Wen, and Soh (2000). However, for strict-feedback nonlinear systems in a more general description form, the control construction still remains an open problem.

In this paper, both full state and output feedback adaptive NN controllers are presented for a class of unknown discrete-time nonlinear systems in strict-feedback form by using high-order neural networks (HONNs). The closed-loop systems are proven to be semi-globally uniformly ultimately bounded (SGUUB), and the tracking

☆ This paper was not presented at any IFAC meeting. This paper was recommended for publication in revised form by Associate Editor Antonio E. de Barros Ruano under the direction of Editor Robert R. Bitmead.

* Corresponding author.

E-mail address: elegesz@nus.edu.sg (S.S. Ge).

0005-1098/03/$ - see front matter © 2003 Published by Elsevier Science Ltd.
doi:10.1016/S0005-1098(03)00032-3

error converges to a small neighborhood of the origin. The main contributions of this paper are as follows:

(i) By transforming the original one-step ahead descriptor description into an equivalent n-step ahead predictor description, an elegant approximation-based controller can be constructed without the causality contradiction problem which has been regarded as the main obstacle encountered in controller design for general strict-feedback nonlinear system through backstepping.

(ii) By choosing Lyapunov functions consisting of plant states and NN parameter deviation from the ideal values W^*, the proposed controllers are free from the controller singularity problem usually encountered in adaptive control using feedback linearization, and at the same time avoids the problem of indirect adaptive control, where one relies on independently showing that the NN modeling converges to a small neighborhood and then appealing to stability arguments.

(iii) Through a diffeomorphism transformation, the strict-feedback form is first transformed into a cascade form, Lyapunov-based direct output feedback control is then developed, which is relatively easy to implement as it only requires the measurement of inputs and outputs, and the controller structure is less complex than that of the proposed backstepping design—this is especially true for high-order systems.

Fundamentally, the results presented in the paper cover nicely some territory between function approximation NNs and the much more historical Direct Model Reference Adaptive Control using "tuned systems" (Kosut & Friedlander, 1985; Anderson et al., 1986). In fact, there are strong connections to the tuned system ideas (where W^* would be a parameter value for which the control works well, rather than that value for which the modeling error is least) of direct adaptive control via Lyapunov methods. The results also open the door to the study of robust disturbance rejection subject to ultimate model quality requirements.

2. System dynamics and stability notions

Consider the following single-input and single-output (SISO) discrete-time nonlinear system in strict-feedback form:

$$\xi_i(k+1) = f_i(\bar{\xi}_i(k)) + g_i(\bar{\xi}_i(k))\xi_{i+1}(k),$$

$$i = 1, 2, \ldots, n-1,$$

$$\xi_n(k+1) = f_n(\bar{\xi}_n(k)) + g_n(\bar{\xi}_n(k))u(k) + d_1(k),$$

$$y_k = \xi_1(k), \tag{1}$$

where $\bar{\xi}_i(k) = [\xi_1(k), \xi_2(k), \ldots, \xi_i(k)]^{\mathrm{T}} \in R^i$, $i = 1, 2, \ldots, n$, $u(k) \in R$, $y_k \in R$ are the state variables, system input and output respectively; $d_1(k)$ denotes the external disturbance bounded by a known constant $d_{10} > 0$, i.e., $|d_1(k)| \leqslant d_{10}$; $f_i(\bar{\xi}_i(k))$ and $g_i(\bar{\xi}_i(k))$, $i = 1, 2, \ldots, n$ are unknown smooth functions.

Assumption 1. The signs of $g_i(\bar{\xi}_i(k))$, $i = 1, 2, \ldots, n$ are known and there exist two constants $\underline{g}_i, \bar{g}_i > 0$ such that $\underline{g}_i \leqslant |g_i(\bar{\xi}_i(k))| \leqslant \bar{g}_i$ $\forall \bar{\xi}_n(k) \in \Omega \subset R^n$.

Without losing generality, we shall assume that $g_i(\bar{\xi}_i(k))$ and $g_n(\bar{\xi}_n(k))$ are positive in this paper. The objective is to design control $u(k)$ to make the system output y_k follow a known and bounded trajectory $y_d(k)$.

Assumption 2. The desired trajectory $y_d(k) \in \Omega_y, \forall k > 0$ is smooth and known, where $\Omega_y := \{\chi \mid \chi = \xi_1\}$.

Definition 2.1. The solution of (1) is semi-globally uniformly ultimately bounded (SGUUB), if for any Ω, a compact subset of R^n and all $\bar{\xi}_n(k_0) \in \Omega$, there exist an $\varepsilon > 0$ and a number $N(\varepsilon, \bar{\xi}_n(k_0))$ such that $\|\bar{\xi}_n(k)\| < \varepsilon$ for all $k \geqslant k_0 + N$ (Lin & Saberi, 1995).

Definition 2.2. The sequence $S(k)$ is said to be persistently exciting (PE) if there are $\bar{\lambda} > 0$ and integer $L > 0$ such that

$$\lambda_{\min}\left[\sum_{k=k_0}^{k_0+L-1} S(k)S^{\mathrm{T}}(k)\right] \geqslant \bar{\lambda} \quad \forall k_0 \geqslant 0, \tag{2}$$

where $\lambda_{\min}(M)$ denotes the smallest eigenvalue of M (Sadegh, 1993).

Consider the linear time-varying discrete-time system given by

$$x(k+1) = A(k)x(k) + Bu(k), \qquad y_k = Cx(k), \tag{3}$$

where $A(k), B$ and C are appropriately dimensional matrices with B and C being constant matrices.

Lemma 1. *Let $\Phi(k_1, k_0)$ be the state-transition matrix corresponding to $A(k)$ for system (3), i.e. $\Phi(k_1, k_0) = \prod_{k=k_0}^{k_1-1} A(k)$. If $\|\Phi(k_1, k_0)\| < 1, \forall k_1 > k_0 \geqslant 0$, then system* (3) *is* (i) *globally exponentially stable for the unforced system* (*i.e.* $u(k) = 0$); *and* (ii) *bounded-input–bounded-output* (*BIBO*) *stable.*

3. Function approximation by HONN

Under certain conditions, it has been proven that several approximation methods, such as polynomials, splines, NNs, fuzzy systems, have function approximation abilities, and have been frequently used as function approximators. There are several types of NNs that have been frequently used, which include linearly parametrized and nonlinearly parametrized networks (Ge et al., 2001). For clarity and simplicity, consider HONNs (Ge et al., 2001; Kosmatopoulos, Polycarpou, Christodoulou, & Ioannou, 1995)

$$\phi(W,z) = W^{\mathrm{T}}S(z), \quad W \text{ and } S(z) \in R^l,$$
$$S(z) = [s_1(z), s_2(z), \ldots, s_l(z)]^{\mathrm{T}}, \tag{4}$$

$$s_i(z) = \prod_{j \in I_i} [s(z_j)]^{d_j(i)}, \quad i = 1, 2, \ldots, l, \tag{5}$$

where $z = [z_1, z_2, \ldots, z_m]^{\mathrm{T}} \in \Omega_z \subset R^m$; positive integer l denotes the NN node number; $\{I_1, I_2, \ldots, I_l\}$ is a collection of l not-ordered subsets of $\{1, 2, \ldots, m\}$ and $d_j(i)$ are non-negative integers; W is an adjustable synaptic weight vector; and $s(z_j)$ is chosen as a hyperbolic tangent function $s(z_j) = (e^{z_j} - e^{-z_j})/(e^{z_j} + e^{-z_j})$.

It has been proven that HONNs satisfies the conditions of the Stone–Weierstrass Theorem and can approximate any continuous function to any desired accuracy over a compact set (Paretto & Niez, 1986). For a smooth function $\varphi(z)$, according to Girosi and Poggio (1989), there exist ideal weights W^* such that the smooth function $\varphi(z)$ can be approximated by an ideal NN on a compact set $\Omega_z \subset R^m$:

$$\varphi(z) = W^{*\mathrm{T}}S(z) + \varepsilon_z, \tag{6}$$

where ε_z is called the NN approximation error. The NN approximation error is a critical quantity, representing the minimum possible deviation of the ideal approximator $W^{*\mathrm{T}}S(z)$ from the unknown smooth function $\varphi(z)$. The NN approximation error can be reduced by increasing the number of the adjustable weights. Universal approximation results for NNs (Gupta & Rao, 1994) indicate that, if NN node number l is sufficiently large, then $|\varepsilon_z|$ can be made arbitrarily small on a compact region.

The ideal weight vector W^* is an "artificial" quantity required for analytical purposes. W^* is defined as the value of W that minimizes $|\varepsilon_z|$ for all $z \in \Omega_z \subset R^m$ in a compact region, i.e.,

$$W^* := \arg\min_{W \in R^l} \left\{ \sup_{z \in \Omega_z} |\varphi(z) - W^{\mathrm{T}}S(z)| \right\}, \quad \Omega_z \subset R^m. \tag{7}$$

In general, the ideal NN weight W^* is unknown and needs to be estimated. In this paper, we shall consider $\hat{W}$ being the estimate of the ideal NN weight W^*.

Assumption 3. On a compact set $\Omega_z \subset R^m$, the ideal NN weights W^* satisfies $\|W^*\| \leqslant w_m$ where w_m is a positive constant.

Consider the basis functions of HONN (4) with $z(k)$ being the input vector. The following properties of HONN will be used in the proof of closed-loop system stability. Let

$$S(z(k)) = [s_1(z(k)), s_2(z(k)), \ldots, s_l(z(k))]^{\mathrm{T}},$$

then $S(k)$ has the following properties:

$$\lambda_{\max}[S(k)S^{\mathrm{T}}(k)] < 1, \quad S^{\mathrm{T}}(k)S(k) < l. \tag{8}$$

4. Adaptive state feedback NN control

Consider the strict-feedback SISO nonlinear discrete-time system described in (1). Since Assumption 1 is only valid on the compact set Ω , it is necessary to guarantee the system's states remaining in Ω for all time. We will design an adaptive control $u(k)$ for system (1) which makes system output y_k follow the desired trajectory $y_d(k)$, and simultaneously guarantees $\bar{\xi}_n(k) \in \Omega$, $\forall k > 0$ under the condition that $\bar{\xi}_n(0) \in \Omega$.

The causality contradiction is one of the major problems that we will encounter in discrete-time domain when we construct a controller for the general strict-feedback nonlinear system through backstepping. Design an ideal fictitious control of the form

$$\xi_{2f}(k) = -\frac{1}{g_1(\bar{\xi}_1(k))}(f_1(\bar{\xi}_1(k)) - y_d(k+1))$$

to stabilize the equation corresponding to $i = 1$ in (1). Similarly, we can construct another ideal fictitious control in the form of

$$\xi_{3f}(k) = -\frac{1}{g_2(\bar{\xi}_2)}(f_2(\bar{\xi}_2) - \xi_{2f}(k+1)) \tag{9}$$

to stabilize the equation corresponding to $i = 2$ in (1). But unfortunately, $\xi_{2f}(k+1)$ in (9) is a fictitious control of the future. This means that the fictitious control $\xi_{3f}(k)$ is infeasible in practice. If we continue the process to construct the final control $u(k)$, we end up with a $u(k)$ which is infeasible again because it contains more future information. However, the above problem can be avoided if we transform the system equation into a special form which is suitable for backstepping design. The basic idea is as follows. If we consider the original system description as a one-step ahead predictor, and we can then transform the one-step ahead predictor into an equivalent maximum n-step ahead predictor, which can predict the future states, $\xi_1(k+n), \xi_2(k+n-1), \ldots, \xi_n(k+1)$, then the causality contradiction is avoided when controller is constructed based on the maximum n-step ahead predictor by backstepping. The transformation of the system is detailed below.

By carefully examining the first $(n-1)$ equations in (1), we conclude that $\xi_i(k+1)$ is a function of $\bar{\xi}_{i+1}(k)$ $\forall 1 \leqslant i \leqslant n-1$. For clarity, denote $\xi_i(k+1)$ as

$$\xi_i(k+1) = f^c_{n,i}(\bar{\xi}_{i+1}(k)), \quad i = 1, 2, \ldots, n-1, \tag{10}$$

where $f^c_{n,i}(\bar{\xi}_{i+1}(k)) = f_i(\bar{\xi}_i(k)) + g_i(\bar{\xi}_i(k))\xi_{i+1}(k)$, $i = 1,2,\ldots,n-1$. Accordingly, we have

$$\bar{\xi}_i(k+1) = \begin{bmatrix} \xi_1(k+1) \\ \vdots \\ \xi_i(k+1) \end{bmatrix} = \begin{bmatrix} f^c_{n,1}(\bar{\xi}_2(k)) \\ \vdots \\ f^c_{n,i}(\bar{\xi}_{i+1}(k)) \end{bmatrix},$$

$$i = 1,2,\ldots,n-1, \tag{11}$$

which is a vectored function of $\bar{\xi}_{i+1}(k)$ denoted by

$$\bar{\xi}_i(k+1) = F^c_{n,i}(\bar{\xi}_{i+1}(k)) \in R^i, \quad i = 1,2,\ldots,n-1. \tag{12}$$

It should be noted that $f^c_{n,i}(\bar{\xi}_{i+1}(k))$ is a scalar, but $F^c_{n,i}(\bar{\xi}_{i+1}(k))$ is a vector of dimension i. When $i = 1$, $F^c_{n,1}(\bar{\xi}_2(k))$ degrades into scalar $f^c_{n,1}(\bar{\xi}_2(k))$.

After one more step, the first $(n-1)$ equations in (1) become

$$\xi_i(k+2) = f_i(\bar{\xi}_i(k+1)) + g_i(\bar{\xi}_i(k+1))\xi_{i+1}(k+1),$$

$$i = 1,2,\ldots,n-2,$$

$$\xi_{n-1}(k+2) = f_{n-1}(\bar{\xi}_{n-1}(k+1)) + g_{n-1}(\bar{\xi}_{n-1}(k+1))\xi_n(k+1). \tag{13}$$

Substituting (10) and (12) into (13), we obtain

$$\xi_i(k+2) = f^c_{n-1,i}(\bar{\xi}_{i+2}(k)),$$

$$i = 1,2,\ldots,n-2,$$

$$\xi_{n-1}(k+2) = F_{n-1}(\bar{\xi}_n(k)) + G_{n-1}(\bar{\xi}_n(k))\xi_n(k+1), \tag{14}$$

where

$$f^c_{n-1,i}(\bar{\xi}_{i+2}(k)) = f_i(F^c_{n,i}(\bar{\xi}_{i+1}(k))) + g_i(F^c_{n,i}(\bar{\xi}_{i+1}(k)))f^c_{n,i+1}(\bar{\xi}_{i+2}(k)),$$
$$i = 1,2,\ldots,n-2,$$

$$F_{n-1}(\bar{\xi}_n(k)) = f_{n-1}(F^c_{n,n-1}(\bar{\xi}_n(k))),$$

$$G_{n-1}(\bar{\xi}_n(k)) = g_{n-1}(F^c_{n,n-1}(\bar{\xi}_n(k))).$$

Following the same procedure, the first $(n-2)$ equations in (14) can be described by

$$\bar{\xi}_i(k+2) = \begin{bmatrix} \xi_1(k+2) \\ \vdots \\ \xi_i(k+2) \end{bmatrix} = \begin{bmatrix} f^c_{n-1,1}(\bar{\xi}_3(k)) \\ \vdots \\ f^c_{n-1,i}(\bar{\xi}_{i+2}(k)) \end{bmatrix},$$

$$i = 1,2,\ldots,n-2,$$

which is a function of $\bar{\xi}_{i+2}(k)$ and denoted as

$$\bar{\xi}_i(k+2) = F^c_{n-1,i}(\bar{\xi}_{i+2}(k)),\ i = 1,2,\ldots,n-2. \tag{15}$$

Continuing the procedure as above recursively, after $(n-2)$ steps, the first two equations in (1) become

$$\xi_1(k+n-1) = f^c_{2,1}(\bar{\xi}_n(k)),$$

$$\xi_2(k+n-1) = F_2(\bar{\xi}_n(k)) + G_2(\bar{\xi}_n(k))\xi_3(k+n-2), \tag{16}$$

where $f^c_{2,1}(\bar{\xi}_n(k)) = f_1(F^c_{3,1}(\bar{\xi}_{n-1}(k))) + g_1(F^c_{3,1}(\bar{\xi}_{n-1}(k))) \times f^c_{3,2}(\bar{\xi}_n(k))$, $F_2(\bar{\xi}_n(k)) = f_2(F^c_{3,2}(\bar{\xi}_n(k)))$ and $G_2(\bar{\xi}_n(k)) = g_2(F^c_{3,2}(\bar{\xi}_n(k)))$.

After still another further step, the first equation in (1) becomes

$$\xi_1(k+n) = F_1(\bar{\xi}_n(k)) + G_1(\bar{\xi}_n(k))\xi_2(k+n-1), \tag{17}$$

where $F_1(\bar{\xi}_n(k)) = f_1(f^c_{2,1}(\bar{\xi}_n(k)))$ and $G_1(\bar{\xi}_n(k)) = g_1(f^c_{2,1}(\bar{\xi}_n(k)))$.

Since all the equations from (13) to (17) are derived from the original system, the original strict-feedback form (1) is equivalent to

$$\xi_1(k+n) = F_1(\bar{\xi}_n(k)) + G_1(\bar{\xi}_n(k))\xi_2(k+n-1),$$
$$\vdots$$
$$\xi_{n-1}(k+2) = F_{n-1}(\bar{\xi}_n(k)) + G_{n-1}(\bar{\xi}_n(k))\xi_n(k+1),$$
$$\xi_n(k+1) = f_n(\bar{\xi}_n(k)) + g_n(\bar{\xi}_n(k))u(k) + d_1(k),$$
$$y_k = \xi_1(k). \tag{18}$$

It should be noted that functions $F_i(\bar{\xi}_n(k))$, $i = 1,2,\ldots,n-1$, become highly nonlinear. In fact, as i decrease, $F_i(\bar{\xi}_n(k))$ and $G_i(\bar{\xi}_n(k))$ become more entangled and complex, because $F_{n-1}(\bar{\xi}_n(k))$ and $G_{n-1}(\bar{\xi}_n(k))$ are derived by one-step substitution, while $F_1(\bar{\xi}_n(k))$ and $G_1(\bar{\xi}_n(k))$ are derived by $(n-1)$-step substitution. As mentioned in Section 3, an HONN can emulate a given nonlinear function to a small error tolerance, therefore it is a good choice to construct our controller using HONN without knowing the exact structures of $F_i(\bar{\xi}_n(k))$ and $G_i(\bar{\xi}_n(k))$. In the following discussion, we will show how to construct our HONN controller by backstepping.

From the definition of $G_i(\bar{\xi}_n(k))$ in each step, it is clear that the value of $G_i(\bar{\xi}_n(k))$, $i = 1,2,\ldots,n-1$, is the same as $g_i(\bar{\xi}_i(k))$, therefore $G_i(\bar{\xi}_n(k))$ satisfy $\underline{g}_i \leqslant G_i(\bar{\xi}_n(k)) \leqslant \bar{g}_i, \forall \bar{\xi}_n(k) \in \Omega$ under Assumption 1.

Now we can construct control for (18) via backstepping method without the problem of causality contradiction. The design procedure is recursive. At the ith step, the ith-order subsystem is stabilized with respect to a Lyapunov function, V_i, by the design of adaptive NN function $\hat{W}_i S_i(z_i(k))$ as a fictitious control. The feedback control $u(k)$ is designed in the final step. For convenience, let $W_i^*, \hat{W}_i, \tilde{W}_i$, $i = 1,2,\ldots,n$, be the ideal value, estimate, and estimation error of the NN weights at the ith design step. After n recursions, we can construct the fictitious controls $\xi_{jf}(k)$, real control $u(k)$, and the NN weight updating law as follows:

$$\xi_{jf}(k) = \hat{W}_{j-1}^{\mathrm{T}}(k)S_{j-1}(z_{j-1}(k)), \quad j = 2,3,\ldots,n,$$

$$u(k) = \hat{W}_n^{\mathrm{T}}(k)S_n(z_n(k)), \tag{19}$$

$$\hat{W}_i(k+1) = \hat{W}_i(k_i) - \Gamma_i[S_i(z_i(k_i))\eta_i(k+1), +\sigma_i\hat{W}_i(k_i)],$$

$$k_i = k - n + i, \ \ i = 1,2,\ldots,n, \tag{20}$$

where diagonal gain matrix $\Gamma_i = \Gamma_i^{\mathrm{T}} > 0$ with $\lambda_{\max}(\Gamma_i) = \bar{\gamma}_i$, constant $\sigma_i > 0$, $z_1(k) = [\bar{\xi}_n^{\mathrm{T}}(k), y_d(k+n)]^{\mathrm{T}} \in \Omega_{z1} \subset R^{n+1}$, $z_j(k) = [\bar{\xi}_n^{\mathrm{T}}(k), \xi_{jf}(k)]^{\mathrm{T}} \in \Omega_{zj} \subset R^{n+1}$, $j = 2,3,\ldots,n$, $\hat{W}_i \in R^{l_i}$ and $S_i(z_i(k)) \in R^{l_i}$ with l_i denotes the neurons used. The error vector $\eta(k) = [\eta_1(k), \eta_2(k), \ldots, \eta_n(k)]^{\mathrm{T}}$ is defined as $\eta_1(k) = \xi_1(k) - y_d(k), \eta_2(k) = \xi_2(k) - \xi_{2f}(k-n+1), \ldots, \eta_n(k) = \xi_n(k) - \xi_{nf}(k-1)$.

In comparison with the standard parameter adaptation algorithms, it should be noted that parameter adaptation algorithm (20) is of $k+1-k_i$th order ($i = 1,2,\ldots,n$) in order to solve the control problem for the ith equation. In fact, the current weight estimate of the ith equation, $\hat{W}_i(k)$, is updated from the estimate, $\hat{W}_i(k_i-1)$, of $n+1-i$ steps earlier rather than that of the previous step for systems in a one-step ahead predictor form. From this observation, we deduce that high-order controllers are useful and effective in handling high-order nonlinear systems.

The stability results of the closed-loop system are summarized in Theorem 1.

Theorem 1. *The closed-loop adaptive system consisting of plant* (1), *controller* (19) *and update law* (20) *is SGUUB and has an equilibrium at* $\eta = [\eta_1, \eta_2, \ldots, \eta_n]^{\mathrm{T}} = 0$, *if* $\bar{\xi}_n(0)$ *is initialized in* Ω. *This guarantees that all the signals include the states* $\bar{\xi}_n = [\xi_1, \xi_2, \ldots, \xi_n]$, *the control* u *and NN weight estimates* $\hat{W}_i$, $i = 1,2,\ldots,n$ *are SGUUB, subsequently,*

$$\lim_{k\to\infty} |y_k - y_d(k)| \leqslant \varepsilon, \tag{21}$$

where ε *is a small positive number.*

Proof. See Appendix A.

5. Adaptive output feedback NN control

In Section 4, we have proposed an adaptive NN controller for strict-feedback nonlinear system via backstepping. This method is relatively easy to implement for relatively lower-order systems, but it becomes too complicated for high-order systems because it needs to update too many parameters. In addition, too many NNs used in the controller not only consume much computing time but are also hard to tune. Furthermore, the control scheme becomes infeasible when some system states are unmeasurable. In order to pursue an easy-to-implement controller for systems with a relatively high order, and with only output measurable, the strict-feedback form is transformed into a cascade form first. For the convenience of analysis and less technical complexity, we assume that there is no disturbance, i.e $d_1(k) = 0$, in this case.

For convenience of analysis, for $i = 1,2,\ldots,n-1$, let

$$F_i(k) = F_i(\bar{\xi}_n(k)), \quad G_i(k) = G_i(\bar{\xi}_n(k)),$$

$$f_n(k) = f_n(\bar{\xi}_n(k)), \quad g_n(k) = g_n(\bar{\xi}_n(k)).$$

They are functions of system state $\bar{\xi}_n(k)$ at the kth step. Substituting $\xi_2(k+n-1)$ in equation (16) into equation (17), we obtain

$$\xi_1(k+n) = F_1(k) + F_2(k)G_1(k) + G_1(k)G_2(k)\xi_3(k+n-2). \tag{22}$$

Continuing the substitution procedure until $u(k)$ appears, we obtain

$$\xi_1(k+n) = F(\bar{\xi}_n(k)) + G(\bar{\xi}_n(k))u(k), \tag{23}$$

where

$$F(\bar{\xi}_n(k)) = F_1(k) + F_2(k)G_1(k) + \cdots + F_{n-1}(k)\prod_{i=1}^{n-2} G_i(k) + f_n(k)\prod_{i=1}^{n-1} G_i(k), \tag{24}$$

$$G(\bar{\xi}_n(k)) = g_n(k)\prod_{i=1}^{n-1} G_i(k). \tag{25}$$

Define the new system states $x(k) = [x_1(k), x_2(k), \ldots, x_n(k)]^{\mathrm{T}}$ as

$$x_1(k) = \xi_1(k),$$

$$x_2(k) = \xi_1(k+1) = f_{n,1}^c(\bar{\xi}_2(k)),$$

$$\vdots$$

$$x_n(k) = \xi_1(k+n-1) = f_{2,1}^c(\bar{\xi}_n(k)), \tag{26}$$

which can be written as $x(k) = T(\bar{\xi}_n(k))$, where $T(\bar{\xi}_n(k))$ is a nonlinear coordinate transformation of the form

$$T(\bar{\xi}_n(k)) = [\xi_1(k), f_{n,1}^c(\bar{\xi}_2(k)), \ldots, f_{2,1}^c(\bar{\xi}_n(k))]^{\mathrm{T}}. \tag{27}$$

For equivalent transformation of coordinates, it is important to make sure that the transformation map is diffeomorphism. Lemmas 2 and 3 guarantee the coordinate transformation mapping T is diffeomorphism.

Lemma 2. *Let* U *be an open subset of* R^n *and let* $\varphi = (\varphi_1, \ldots, \varphi_n): U \to R^n$ *be a smooth map. If the Jacobian Matrix*

$$\frac{\mathrm{d}\varphi}{\mathrm{d}x} = \begin{bmatrix} \dfrac{\partial\varphi_1}{\partial x_1} & \cdots & \dfrac{\partial\varphi_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \dfrac{\partial\varphi_n}{\partial x_1} & \cdots & \dfrac{\partial\varphi_n}{\partial x_n} \end{bmatrix}$$

is nonsingular at some point $p \in U$, *or equivalently,* $\mathrm{Rank}(\mathrm{d}\varphi/\mathrm{d}x) = n$ *at some point* $p \in U$, *then there exists a neighborhood* $V \subset U$ *of* p *such that* $\varphi : V \to \varphi(V)$ *is a diffeomorphism* (Marino & Tomei, 1995).

Lemma 3. *For system* (1), *the nonlinear coordinate transformation* $T(\bar{\xi}_n(k))$ *in* (27), $T : \Omega \to \Omega_x \subset R^n$, *is a smooth map and a diffeomorphism, namely, there exists a unique transformation function* $T^{-1}(x(k))$ *such that* $\bar{\xi}_n(k) = T^{-1}(x(k))$, *and both* T *and* T^{-1} *are one-to-one.*

Proof. See Appendix B.

Based on the above analysis, we can obtain the cascade system description, which is equivalent to the original system (1), as follows:

$$
\begin{aligned}
&x_1(k+1) = x_2(k),\\
&x_2(k+1) = x_3(k),\\
&\quad\vdots\\
&x_n(k+1) = f(x(k)) + g(x(k))u(k),\\
&y_k = x_1(k),
\end{aligned} \tag{28}
$$

where $x(k) = [x_1(k), x_2(k), \ldots, x_n(k)]^{\mathrm{T}} \in R^n$, and $f(x(k)) = F(\bar{\xi}_n(k)) = F(T^{-1}(x(k))), g(x(k)) = G(\bar{\xi}_n(k)) = G(T^{-1}(x(k)))$ are unknown smooth functions. Our objective becomes to force the output y_k of system (28), whose state $x(k)$ belongs to a compact subset $\Omega_x \subset R^n$, to follow a desired trajectory $y_d(k)$.

It should be noted that although the new states $x_2, x_3, \ldots, x_n$ are not available in practice, we can predict them as will be detailed in the following discussion.

Let $x_d(k) = [y_d(k), y_d(k+1), \ldots, y_d(k+n-1)]^{\mathrm{T}}$ be the desired system states. Noting (25), the following statements are reasonable under Assumptions 1 and 2: (i) $x_d \in \Omega_d \subset \Omega_x$ and (ii) the sign of $g(x(k))$ is known and there exist two constant $\underline{g}$, $\bar{g} > 0$ such that $\underline{g} \leqslant |g(x(k))| \leqslant \bar{g}\ \forall x(k) \in \Omega_x$. Since the sign of $g(x(k))$ is known, we may, without losing generality, assume that $g(x(k))$ is positive in the following discussion.

Define error $e(k) = x(k) - x_d(k) = [e_1(k), e_2(k), \ldots, e_n(k)]^{\mathrm{T}}$. The equation of $e(k)$ can be written as

$$
\begin{aligned}
&e_1(k+1) = e_2(k),\\
&e_2(k+1) = e_3(k),\\
&\quad\vdots\\
&e_n(k+1) = f(x(k)) + g(x(k))u(k) - y_d(k+n).
\end{aligned} \tag{29}
$$

In order to develop the output feedback control clearly, define the following new variables $\underline{y}(k) = [y_{k-n+1}, \ldots, y_{k-1}, y_k]^{\mathrm{T}}$, $\underline{u}_{k-1}(k) = [u_{k-1}, \ldots, u_{k-n+1}]^{\mathrm{T}}$ and $\underline{z}(k) = [\underline{y}^{\mathrm{T}}(k), \underline{u}^{\mathrm{T}}_{k-1}(k)]^{\mathrm{T}} \in \Omega_{\underline{z}} \subset R^{2n-1}$.

According to the definition of the new states, we know that $\underline{y}(k) = [x_1(k-n+1), \ldots, x_1(k-1), x_1(k)]^{\mathrm{T}}$. From Eq. (28), we obtain

$$
\begin{aligned}
y_{k+1} &= x_2(k) = x_3(k-1) = \cdots\\
&= x_n(k-n+2) = f(\underline{y}(k)) + g(\underline{y}(k))u_{k-n+1}.
\end{aligned} \tag{30}
$$

It means that $x_2(k)$ is a function of $\underline{y}(k)$ and u_{k-n+1}. It should be noted that although the right-hand side of (30) does not contain all the elements of $\underline{z}(k)$, for convenience of analysis, we can denote (30) as follows without any ambiguity:

$$y_{k+1} = x_2(k) = \psi_2(\underline{z}(k)). \tag{31}$$

Similarly, we obtain the following equation from (28):

$$y_{k+2} = x_3(k) = f(\underline{y}(k+1)) + g(\underline{y}(k+1))u_{k-n+2}. \tag{32}$$

Substituting (31) into (32), we know that $x_3(k)$ is a function of $\underline{y}(k)$, u_{k-n+2} and u_{k-n+1}. Let

$$y_{k+2} = x_3(k) = \psi_3(\underline{z}(k)). \tag{33}$$

If we continue the substitution recursively, it is easy to prove that $x_n(k)$ is a function of $\underline{z}(k)$ as expressed below:

$$y_{k+n-1} = x_n(k) = \psi_n(\underline{z}(k)). \tag{34}$$

Up to this step, $\psi_n(*)$ contains all the elements of $\underline{z}(k)$. Substituting the predicted states into the last equation in (28), we obtain

$$y_{k+n} = x_n(k+1) = f_0(\underline{z}(k)) + g_0(\underline{z}(k))u_k, \tag{35}$$

where

$$
\begin{aligned}
f_0(\underline{z}(k)) &:= F_0(\underline{z}(k))\\
&= f([x_1(k), \psi_2(\underline{z}(k)), \ldots, \psi_n(\underline{z}(k))]^{\mathrm{T}}),\\
g_0(\underline{z}(k)) &:= G_0(\underline{z}(k))\\
&= g([x_1(k), \psi_2(\underline{z}(k)), \ldots, \psi_n(\underline{z}(k))]^{\mathrm{T}}).
\end{aligned}
$$

Define tracking error as $e_y(k) = y_k - y_d(k)$. The tracking error dynamics are given by

$$e_y(k+n) = -y_d(k+n) + f_0(\underline{z}(k)) + g_0(\underline{z}(k))u_k. \tag{36}$$

Supposing that the nonlinear functions $f_0(\underline{z}(k))$ and $g_0(\underline{z}(k))$ are known exactly, we present a desired control, $\bar{u}^*_k$, such that the output y_k follows the desired trajectory $y_d(k)$ in deadbeat step

$$\bar{u}^*_k = -\frac{1}{g_0(\underline{z}(k))}(f_0(\underline{z}(k)) - y_d(k+n)). \tag{37}$$

Substituting the desired control $\bar{u}^*_k$ into error dynamics equation (36), we obtain

$$e_y(k+n) = 0.$$

This means that after n steps, we have $e_y(k) = 0$. Therefore, $\bar{u}^*_k$ is a n-step deadbeat control.

Accordingly, the desired control $\bar{u}_k^*$ can be expressed as

$$\bar{u}_k^* = \bar{u}^*(\bar{z}(k)),$$

where $\bar{z}(k) = [\underline{z}^{\mathrm{T}}(k), y_d(k+n)]^{\mathrm{T}} \in \Omega_{\bar{z}} \subset R^{2n}$ with compact set $\Omega_{\bar{z}}$ defined as

$$\Omega_{\bar{z}} = \{(\underline{y}(k), \underline{u}_{k-1}, y_d) | \underline{u}_{k-1}(k) \in \Omega_u, \underline{y}(k) \in \Omega_y, y_d \in \Omega_y\}.$$

As mentioned in Section 3, there exist an integer l^* and an ideal constant weight vector W^*, such that for all $l \geqslant l^*$,

$$\bar{u}^*(\bar{z}) = W^{*\mathrm{T}} S(\bar{z}) + \varepsilon_{\bar{z}}, \quad \forall \bar{z} \in \Omega_{\bar{z}}, \tag{38}$$

where $\varepsilon_{\bar{z}}$ is the NN estimation error satisfying $|\varepsilon_{\bar{z}}| < \varepsilon_0$.

If we choose the control law and the weight updating law as

$$u_k = \hat{W}(k) S(\bar{z}(k)), \tag{39}$$

$$\hat{W}(k+1) = \hat{W}(k_1) + \Gamma[S(\bar{z}(k_1))(y_{k+1} - y_d(k+1)) + \sigma \hat{W}(k)],$$

$$k_1 = k - n + 1, \tag{40}$$

where diagonal gain matrix $\Gamma = \Gamma^{\mathrm{T}} > 0$, and $\sigma > 0$, the following theorem summarizes the closed-loop stability.

Theorem 2. *Consider the closed-loop system consisting of system* (1), *controller* (39) *and adaptation law* (40). *There exist compact sets $\Omega_{y0} \subset \Omega_y$, $\Omega_{w_0} \subset \Omega_w$ and positive constants l^*, γ^*, and σ^* such that if*

(i) *assumptions* 1–3 *being satisfied, the initial condition $y(0) \in \Omega_{y0}$, $\tilde{W}(0) \in \Omega_{w_0}$, and*
(ii) *the design parameters are suitably chosen such that $l > l^*, \sigma < \sigma^*$ and $\bar{\gamma} < \gamma^*$ with $\bar{\gamma}$ being the largest eigenvalue of Γ,*

then the closed-loop system is SGUUB. The tracking error can be made arbitrarily small by increasing the approximation accuracy of the NNs.

Proof. See Appendix C.

6. Simulation study

To demonstrate the effectiveness of the proposed schemes, consider a nonlinear discrete-time SISO plant described by

$$\xi_1(k+1) = f_1(\xi_1(k)) + 0.3\xi_2(k),$$

$$\xi_2(k+1) = f_2(\bar{\xi}_2(k)) + u(k) + d(k),$$

$$y_k = x_1(k),$$

where

$$f_1(\xi_1(k)) = \frac{1.4\xi_1^2(k)}{1+\xi_1^2(k)},$$

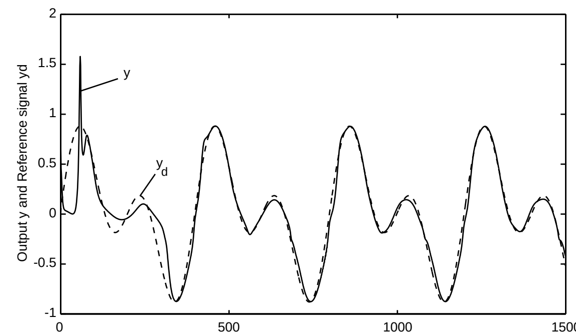

Fig. 1. Tracking performance of state feedback.

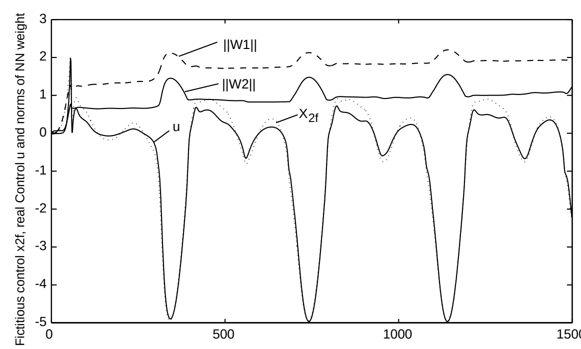

Fig. 2. Boundedness of $x_{2f}(k)$, $u(k)$, $\|\hat{W}_1(k)\|$ and $\|\hat{W}_2(k)\|$.

$$f_2(\bar{\xi}_2(k)) = \frac{\xi_1(k)}{1+\xi_1^2(k)+\xi_2^2(k)},$$

$$d(k) = 0.1\cos(0.05k)\cos(\xi_1(k)).$$

Suppose that there is no a priori knowledge of the system nonlinearities. It can be checked that Assumption 1 and 2 are satisfied. The tracking objective is to make the output y_k follow a desired reference signal $y_d(k) = \frac{1}{2}\sin(k\pi/20) + \frac{1}{2}\sin(k\pi/10)$.

6.1. State feedback control via backstepping

The initial condition for system states is $x(0) = [0\,0]^{\mathrm{T}}$, and those for NNs are $\hat{W}_1(0) = 0, \hat{W}_2(0) = 0$. Other controller parameters are chosen as $l_1 = 22, l_2 = 22, \Gamma_1 = 0.08I, \Gamma_2 = 0.08I$. The simulation results are presented in Figs. 1 and 2. Fig. 1 shows the output y_k and the reference signal $y_d(k)$. Fig. 2 illustrates the boundedness of the fictitious control $x_{2f}(k)$, real control $u(k)$, and the norms of NN weights $\|\hat{W}_1(k)\|$ and $\|\hat{W}_2(k)\|$.

6.2. Direct output feedback control

The initial condition for system states is $x(0) = [0\,0]^{\mathrm{T}}$, and that for NNs is $\hat{W}(0) = 0$. Other controller parameters are chosen as $l_1 = 29, \Gamma = 0.15I$. The simulation results are presented in Figs. 3 and 4. Fig. 3 shows the output y_k and the

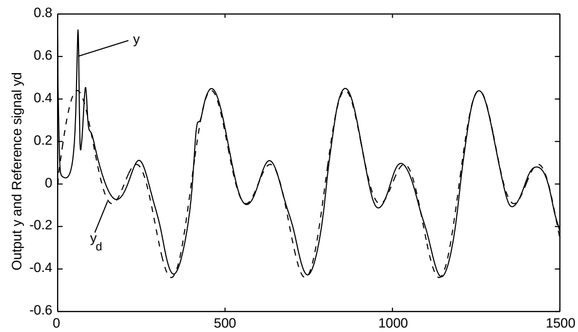

Fig. 3. Tracking performance of output feedback.

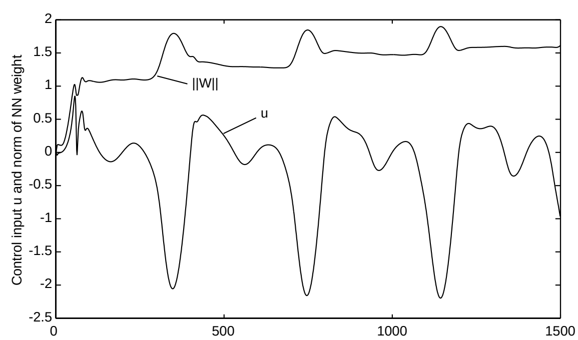

Fig. 4. Boundedness of $u(k)$ and $\|\hat{W}(k)\|$.

reference signal $y_d(k)$. Fig. 4 illustrates the boundedness of the control input $u(k)$ and the norm of NN weight $\|\hat{W}(k)\|$.

Remark.

- It should be pointed out that HONNs used in this paper may be replaced by any other linear approximator such as radial basis function networks (Sanner & Slotine, 1992), spline functions (Nurnberger, 1989) or fuzzy systems (Spooner & Passino, 1996), which have the similar properties as described in (8), while the stability and performance properties of the adaptive system are still valid.
- Since normally $S^{\mathrm{T}}(z)S(z) \ll l$, we can choose a little larger Γ for fast learning rate without destroying the stability of the closed-loop system.

7. Conclusion

In this paper, for a class of nonlinear unknown discrete-time systems in strict-feedback form, an adaptive full state feedback NN controller has been presented via backstepping, which avoids the possible controller singularity problem in adaptive nonlinear control and solves the noncausal problem in the discrete-time backstepping design procedure. After the system description form is transformed into the cascade form, another adaptive direct output feedback control scheme has also been presented using neural networks. It has been shown that for appropriately chosen controller parameters, stability of the closed-loop adaptive system can be guaranteed.

Acknowledgements

The authors are very grateful to the many positive and constructive comments from the Editor and anonymous reviewers. Without these comments, the paper would not be improved to its present quality.

Appendix A. Proof of Theorem 1

Proof. For convenience of analysis and discussion, for $i = 1,2,\ldots,n-1$, let

$$F_i(k) = F_i(\bar{\xi}_n(k)), \quad G_i(k) = G_i(\bar{\xi}_n(k)),$$

$$f_n(k) = f_n(\bar{\xi}_n(k)), \quad g_n(k) = g_n(\bar{\xi}_n(k)).$$

They are functions of system states $\bar{\xi}_n(k)$ at the kth step. Suppose that $\bar{\xi}_n(k) \in \Omega, \forall k \geqslant 0$, then NN approximation (6) is valid. Now we prove that $\bar{\xi}_n(k+1) \in \Omega$ and $u(k)$ is bounded by backstepping.

Before further going, let $k_i = k - n + i,\ i = 1,2,\ldots,n-1$ for convenient description.

Step 1: For $\eta_1(k) = \xi_1(k) - y_d(k)$, its nth difference is given by

$$\begin{aligned}\eta_1(k+n) &= \xi_1(k+n) - y_d(k+n) \\ &= F_1(k) + G_1(k)\xi_2(k+n-1) \\ &\quad - y_d(k+n). \end{aligned} \tag{A.1}$$

Considering $\xi_2(k+n-1)$ as a fictitious control for (A.1), it is obvious that $\eta_1(k+n) = 0$ if we choose

$$\begin{aligned}\xi_2(k+n-1) &= \xi_{2d}^*(k) \\ &= -\frac{1}{G_1(k)}[F_1(k) - y_d(k+n)]. \end{aligned} \tag{A.2}$$

Since $F_1(k)$ and $G_1(k)$ are unknown, they are not available for constructing a fictitious control $\xi_{2d}^*(k)$. However, $F_1(k)$ and $G_1(k)$ are function of system state $\bar{\xi}_n(k)$, therefore we can use HONN to approximate $\xi_{2d}^*(k)$ as follows:

$$\xi_{2d}^*(k) = W_1^{*\mathrm{T}} S_1(z_1(k)) + \varepsilon_{z_1}(z_1(k)),$$

$$z_1(k) = [\bar{\xi}_n^{\mathrm{T}}(k), y_d(k+n)]^{\mathrm{T}} \in \Omega_{z1} \subset R^{n+1}. \tag{A.3}$$

Letting $\hat{W}_1$ be the estimate of W_1^*, consider the direct adaptive fictitious control

$$\xi_2(k+n-1) = \xi_{2f}(k) = \hat{W}_1^{\mathrm{T}}(k) S_1(z_1(k)) \tag{A.4}$$

and the robust updating algorithm for NN weights as

$$\begin{aligned}\hat{W}_1(k+1) &= \hat{W}_1(k_1) - \Gamma_1[S_1(z_1(k_1))\eta_1(k+1) \\ &\quad + \sigma_1 \hat{W}_1(k_1)]. \end{aligned} \tag{A.5}$$

Substituting fictitious control (A.4) into (A.1), the error equation (A.1) is re-written as

$$\eta_1(k+n)=F_1(k)-y_d(k+n)+G_1(k)\hat{W}_1^{\mathrm{T}}(k)S_1(z_1(k)). \quad (A.6)$$

Adding and subtracting $G_1(k)\xi_{2d}^*(k)$ on the right-hand side of (A.6) and noting (A.3), we have

$$\begin{aligned}\eta_1(k+n)=&F_1(k)-y_d(k+n)+G_1(k)[\hat{W}_1^{\mathrm{T}}(k)S_1(z_1(k))\\&-W_1^{*\mathrm{T}}S_1(z_1(k))-\varepsilon_{z_1}(z_1(k))]\\&+G_1(k)\xi_{2d}^*(k),\quad \forall z_1(k)\in\Omega_{z1}.\end{aligned} \quad (A.7)$$

Substituting (A.2) into (A.7) leads to

$$\eta_1(k+n)=G_1(k)[\tilde{W}_1^{\mathrm{T}}(k)S_1(z_1(k))-\varepsilon_{z_1}]. \quad (A.8)$$

Choose the Lyapunov function candidate

$$V_1(k)=\frac{1}{\bar{g}_1}\eta_1^2(k)+\sum_{j=0}^{n-1}\tilde{W}_1^{\mathrm{T}}(k_1+j)\Gamma_1^{-1}\tilde{W}_1(k_1+j), \quad (A.9)$$

where $k_1=k-n+1$.

Noting the fact that $\tilde{W}_1^{\mathrm{T}}(k_1)S_1(z_1(k_1))=\eta_1(k+1)/G_1(k_1)+\varepsilon_{z_1}$, the first difference of (A.9) along (A.5) and (A.8) is given by

$$\begin{aligned}\Delta V_1 =& \frac{1}{\bar{g}_1}[\eta_1^2(k+1)-\eta_1^2(k)]+\tilde{W}_1^{\mathrm{T}}(k+1)\Gamma_1^{-1}\\&\times\tilde{W}_1(k+1)-\tilde{W}_1^{\mathrm{T}}(k_1)\Gamma_1^{-1}\tilde{W}_1(k_1)\\
=&\frac{1}{\bar{g}_1}[\eta_1^2(k+1)-\eta_1^2(k)]-2\tilde{W}_1^{\mathrm{T}}(k_1)[S_1(z_1(k_1))\\&\times\eta_1(k+1)+\sigma_1\hat{W}_1(k_1)]+[S_1(z_1(k_1))\eta_1(k+1)\\&+\sigma_1\hat{W}_1(k_1)]^{\mathrm{T}}\Gamma_1[S_1(z_1(k_1))\eta_1(k+1)\\&+\sigma_1\hat{W}_1(k_1)]\\
=&\frac{1}{\bar{g}_1}[\eta_1^2(k+1)-\eta_1^2(k)]-2\tilde{W}_1^{\mathrm{T}}(k_1)S_1(z_1(k_1))\\&\times\eta_1(k+1)-2\sigma_1\tilde{W}_1^{\mathrm{T}}(k_1)\hat{W}_1(k_1)\\&+S_1^{\mathrm{T}}(z_1(k_1))\Gamma_1S_1(z_1(k_1))\eta_1^2(k+1)\\&+2\sigma_1\hat{W}_1^{\mathrm{T}}(k_1)\Gamma_1S_1(z_1(k_1))\eta_1(k+1)\\&+\sigma_1^2\hat{W}_1^{\mathrm{T}}(k_1)\Gamma_1\hat{W}_1(k_1)\\
\leqslant&-\frac{1}{\bar{g}_1}\eta_1^2(k+1)-\frac{1}{\bar{g}_1}\eta_1^2(k)+2\varepsilon_{z_1}\eta_1(k+1)\\&-2\sigma_1\tilde{W}_1^{\mathrm{T}}(k_1)\hat{W}_1(k_1)\\&+S_1^{\mathrm{T}}(z_1(k_1))\Gamma_1S_1(z_1(k_1))\eta_1^2(k+1)\\&+2\sigma_1\hat{W}_1^{\mathrm{T}}(k_1)\Gamma_1S_1(z_1(k_1))\eta_1(k+1)\\&+\sigma_1^2\hat{W}_1^{\mathrm{T}}(k_1)\Gamma_1\hat{W}_1(k_1).\end{aligned} \quad (A.10)$$

Using the facts that

$$S_1^{\mathrm{T}}(z_1(k_1))S_1(z_1(k_1))<l_1,$$

$$S_1^{\mathrm{T}}(z_1(k_1))\Gamma_1S_1(z_1(k_1))\leqslant\bar{\gamma}_1S_1^{\mathrm{T}}(z_1(k_1))S_1(z_1(k_1))\leqslant\bar{\gamma}_1l_1,$$

$$2\varepsilon_{z_1}\eta_1(k+1)\leqslant\frac{\bar{\gamma}_1\eta_1^2(k+1)}{\bar{g}_1}+\frac{\bar{g}_1\varepsilon_{z_1}^2}{\bar{\gamma}_1},$$

$$2\sigma_1\hat{W}_1^{\mathrm{T}}(k_1)\Gamma_1S_1(z_1(k_1))\eta_1(k+1)\leqslant\frac{\bar{\gamma}_1l_1\eta_1^2(k+1)}{\bar{g}_1}+\bar{g}_1\sigma_1^2\bar{\gamma}_1\|\hat{W}_1\|^2,$$

$$2\tilde{W}_1^{\mathrm{T}}(k_1)\hat{W}_1(k_1)=\|\tilde{W}_1(k_1)\|^2+\|\hat{W}_1(k_1)\|^2-\|W_1^*\|^2,$$

we obtain

$$\Delta V_1\leqslant-\frac{\rho_1}{\bar{g}_1}\eta_1^2(k+1)-\frac{1}{\bar{g}_1}\eta_1^2(k)+\beta_1-\sigma_1(1-\sigma_1\bar{\gamma}_1-\bar{g}_1\sigma_1\bar{\gamma}_1)\|\hat{W}_1(k_1)\|^2,$$

where $\rho_1=1-\bar{\gamma}_1-\bar{\gamma}_1l_1-\bar{g}_1\bar{\gamma}_1l_1$, $\beta_1=\bar{g}_1\varepsilon_{z1}^2/\bar{\gamma}_1+\sigma_1\|W_1^*\|^2$. If we choose the design parameters as follows:

$$\bar{\gamma}_1<\frac{1}{1+l_1+\bar{g}_1l_1},\qquad \sigma_1<\frac{1}{(1+\bar{g}_1)\bar{\gamma}_1}, \quad (A.11)$$

then $\Delta V_1\leqslant 0$ once the error $\eta_1(k)$ is larger than $\sqrt{\bar{g}_1\beta_1}$. This implies the boundedness of $V_1(k)$ for all $k\geqslant 0$, which leads to the boundedness of $\eta_1(k)$ because $V_1(k)=V_1(0)+\sum_{j=0}^{k}\Delta V_1(j)<\infty$. Furthermore, the tracking error $\eta_1(k)$ will asymptotically converge to the compact set denoted by $\Omega_1\subset R$, where $\Omega_1:=\{\chi|\chi\leqslant\sqrt{\bar{g}_1\beta_1}\}$.

The adaptation dynamics (A.5) can be written as

$$\begin{aligned}\tilde{W}_1(k+1)=&\tilde{W}_1(k_1)-\Gamma_1[S_1(z_1(k_1))\eta_1(k+1)\\&+\sigma_1\tilde{W}_1(k_1)+\sigma_1W_1^*]\\
=&\tilde{W}_1(k_1)-\Gamma_1\{S_1(z_1(k_1))G_1(k_1)\\&\times[\tilde{W}_1^{\mathrm{T}}(k_1)S_1(z_1(k_1))-\varepsilon_{z_1}]\\&+\sigma_1\tilde{W}_1(k_1)+\sigma_1W^*\}\\
=&A_1(k)\tilde{W}_1(k_1)+\Gamma_1S_1(z_1(k_1))G_1(k_1)\varepsilon_{z_1}\\&-\sigma_1\Gamma_1W_1^*,\end{aligned}$$

where $A_1(k)=I-\sigma_1\Gamma_1-G_1(k_1)\Gamma_1S_1(z_1(k_1))S_1^{\mathrm{T}}(z_1(k_1))$, ε_{z_1} is the bounded functional estimation error, and function $G_1(k)$ is bounded from Assumption 1. From the proof in Ge, Lee, Li, and Zhang (2003), we know that the $A_1(k)$ so defined guarantees $\|\Phi(k_1,k_0)\|<1$. Applying Lemma 1, $\tilde{W}_1(k)$ is bounded in a compact set denoted by Ω_{w1}, and hence the boundedness of $\hat{W}_1(k)$ is assured, without the need for PE condition.

Step 2: As defined before, $\eta_2(k) = \xi_2(k) - \xi_{2f}(k_1)$. Its $(n-1)$th difference is given by

$$\eta_2(k+n-1) = \xi_2(k+n-1) - \xi_{2f}(k) = F_2(k) + G_2(k)\xi_3(k+n-2) - \xi_{2f}(k). \quad (A.12)$$

Similarly, consider $\xi_3(k+n-2)$ as a fictitious control for (A.12). It is obvious that $\eta_2(k+n-1) = 0$ is true when we choose

$$\xi_3(k+n-2) = \xi_{3d}^*(k) = -\frac{1}{G_2(k)}[F_2(k) - \xi_{2f}(k)]. \quad (A.13)$$

Accordingly, $\xi_{3d}^*(k)$ can be approximated by an ideal HONN

$$\xi_{3d}^*(k) = W_2^{*\mathrm{T}} S_2(z_2(k)) + \varepsilon_{z_2}(z_2(k)),$$
$$z_2(k) = [\bar{\xi}_n^{\mathrm{T}}(k), \xi_{2f}(k)]^{\mathrm{T}} \in \Omega_{z2} \subset R^{n+1}. \quad (A.14)$$

Consider the direct adaptive fictitious controller as

$$\xi_3(k+n-2) = \xi_{3f}(k) = \hat{W}_2^{\mathrm{T}}(k) S_2(z_2(k)) \quad (A.15)$$

and the robust updating algorithm for NN weights as

$$\hat{W}_2(k+1) = \hat{W}_2(k_2) - \Gamma_2[S_2(z_2(k_2))\eta_2(k+1) + \sigma_2 \hat{W}_2(k_2)]. \quad (A.16)$$

Following the same procedure and methods as in Step 1, we obtain the second step error equation

$$\eta_2(k+n-1) = G_2(k)[\tilde{W}_2^{\mathrm{T}}(k) S_2(z_2(k)) - \varepsilon_{z_2}]. \quad (A.17)$$

Choose the Lyapunov function candidate

$$V_2(k) = V_1(k) + \frac{1}{\bar{g}_2}\eta_2^2(k) + \sum_{j=0}^{n-2} \tilde{W}_2^{\mathrm{T}}(k_2+j)\Gamma_2^{-1}\tilde{W}_2(k_2+j), \quad (A.18)$$

where $k_2 = k - n + 2$. The first difference of (A.18) along (A.16) and (A.17) is given by

$$\Delta V_2 \leqslant -\frac{\rho_1}{\bar{g}_1}\eta_1^2(k+1) - \frac{1}{\bar{g}_1}\eta_1^2(k) - \frac{\rho_2}{\bar{g}_2}\eta_2^2(k+1) - \frac{1}{\bar{g}_2}\eta_2^2(k) + \beta_2 - \sigma_2(1 - \sigma_2\bar{\gamma}_2 - \bar{g}_2\sigma_2\bar{\gamma}_2)\|\hat{W}_2(k_2)\|^2,$$

where ρ_1 is defined as in Step 1, and $\rho_2 = 1 - \bar{\gamma}_2 - \bar{\gamma}_2 l_2 - \bar{g}_2\bar{\gamma}_2 l_2$, $\beta_2 = \beta_1 + \bar{g}_2\varepsilon_{z2}^2/\bar{\gamma}_2 + \sigma_2\|W_2^*\|^2$.

If we choose the design parameters as follows:

$$\bar{\gamma}_2 < \frac{1}{1 + l_2 + \bar{g}_2 l_2}, \quad \sigma_2 < \frac{1}{(1+\bar{g}_2)\bar{\gamma}_2}, \quad (A.19)$$

then $\Delta V_2 \leqslant 0$ once $|\eta_1(k)| > \sqrt{\bar{g}_1\beta_2}$ and/or $|\eta_2(k)| > \sqrt{\bar{g}_2\beta_2}$.

As explained in Step 1, $V_2(k)$ is bounded for all $k \geqslant 0$, and the tracking errors $\eta_1(k)$ and $\eta_2(k)$ are also bounded and will asymptotically converge to the compact set denoted by $\Omega_2 \subset R^2$, where $\Omega_2 := \{\chi | \chi = [\chi_1, \chi_2]^{\mathrm{T}}, \chi_1 \leqslant \sqrt{\bar{g}_1\beta_2}, \chi_2 \leqslant \sqrt{\bar{g}_2\beta_2}\}$. The boundedness of $\hat{W}_2(k)$, or equivalently of $\tilde{W}_2(k)$, can be proved as in Step 1.

Step i: Following the same procedure as in Step 2, for $\eta_i(k) = \xi_i(k) - \xi_{if}(k_{i-1})$, we have the direct adaptive fictitious controller and the robust updating algorithm for NN weights as follows:

$$\xi_{i+1}(k+n-i) = \xi_{i+1,f}(k) = \hat{W}_i^{\mathrm{T}}(k) S_i(z_i(k)), \quad (A.20)$$

$$\hat{W}_i(k+1) = \hat{W}_i(k_i) - \Gamma_i[S_i(z_i(k_i))\eta_i(k+1) + \sigma_i\hat{W}_i(k_i)],$$
$$z_i(k) = [\bar{\xi}_n^{\mathrm{T}}(k), \xi_{if}(k)]^{\mathrm{T}} \in \Omega_{zi} \subset R^{n+1}. \quad (A.21)$$

Accordingly, we obtain the ith error equation

$$\eta_i(k+n-i+1) = G_i(k)[\tilde{W}_i^{\mathrm{T}}(k) S_i(z_i(k)) - \varepsilon_{z_i}]. \quad (A.22)$$

Choose the Lyapunov function candidate

$$V_i(k) = \sum_{j=1}^{i-1} V_j(k) + \frac{1}{\bar{g}_i}\eta_i^2(k) + \sum_{j=0}^{n-i} \tilde{W}_i^{\mathrm{T}}(k_i+j)\Gamma_i^{-1}\tilde{W}_i(k_i+j), \quad (A.23)$$

where $k_i = k - n + i$. The first difference of (A.23) along (A.21) and (A.22) is given as

$$\Delta V_i \leqslant -\sum_{j=1}^{i}\frac{\rho_j}{\bar{g}_j}\eta_j^2(k+1) - \sum_{j=1}^{i}\frac{1}{\bar{g}_j}\eta_j^2(k) + \beta_i - \sigma_i(1 - \sigma_i\bar{\gamma}_i - \bar{g}_i\sigma_i\bar{\gamma}_i)\|\hat{W}_i(k_i)\|^2,$$

where ρ_j, $j = 1, 2, \ldots, i-1$, are defined as the previous $(i-1)$ steps, $\rho_i = 1 - \bar{\gamma}_i - \bar{\gamma}_i l_i - \bar{g}_i\bar{\gamma}_i l_i$ and $\beta_i = \beta_{i-1} + \bar{g}_i\varepsilon_{zi}^2/\bar{\gamma}_i + \sigma_i\|W_i^*\|^2$. If we choose the design parameters as follows:

$$\bar{\gamma}_i < \frac{1}{1 + l_i + \bar{g}_i l_i}, \quad \sigma_i < \frac{1}{(1+\bar{g}_i)\bar{\gamma}_i}, \quad (A.24)$$

then $\Delta V_i \leqslant 0$ once any one of the i errors satisfies $|\eta_j(k)| > \sqrt{\bar{g}_j\beta_i}$ $j = 1, 2, \ldots, i$. This demonstrates that the tracking error $\eta_1(k), \eta_2(k), \ldots, \eta_i(k)$ are bounded for all $k \geqslant 0$, and will asymptotically converge to the compact set denoted by $\Omega_i \subset R^i$, where $\Omega_i := \{\chi | \chi = [\chi_1, \chi_2, \ldots, \chi_i]^{\mathrm{T}}, \chi_j \leqslant \sqrt{\bar{g}_j\beta_i}$ $j = 1, 2, \ldots, i\}$. The boundedness of $\hat{W}_i(k)$, or equivalently of $\tilde{W}_i(k)$, can be proved as in Step 1.

Step n: For $\eta_n(k) = \xi_n(k) - \xi_{nf}(k-1)$, its first difference is given by

$$\eta_n(k+1) = \xi_n(k+1) - \xi_{nf}(k) = f_n(k) + g_n(k)u(k) + d_1(k) - \xi_{nf}(k). \quad (A.25)$$

It is obvious that $\eta_n(k+1)=0$, if we choose

$$u(k)=u^*(k)=-\frac{1}{g_n(k)}[f_n(k)-\xi_{nf}(k)] \tag{A.26}$$

and there are no disturbances, i.e. $d_1(k)=0$. If $d_1(k)\neq 0$, we obtain $\eta_n(k+1)=d_1(k)$. Thus exact tracking cannot be obtained though bounded due to the boundedness of the disturbances. Similarly, $u^*(k)$ can be approximated by an HONN

$$u^*(k)=W_n^{*\mathrm{T}}S_n(z_n(k))+\varepsilon_{z_n}(z_n(k)),$$

$$z_n(k)=[\bar{\xi}_n^{\mathrm{T}}(k),\xi_{nf}(k)]^{\mathrm{T}}\in\Omega_{zn}\subset R^{n+1}. \tag{A.27}$$

Following the same procedure as in Step i or 2, we choose the direct adaptive controller and robust undating algorithm for NN weights as

$$u(k)=\hat{W}_n^{\mathrm{T}}(k)S_n(z_n(k)), \tag{A.28}$$

$$\hat{W}_n(k+1)=\hat{W}_n(k)-\Gamma_n[S_n(z_n(k))\eta_n(k+1)+\sigma_n\hat{W}_n(k)]. \tag{A.29}$$

For the nth step error equation

$$\begin{aligned}\eta_n(k+1)&=g_n(k)[\tilde{W}_n^{\mathrm{T}}(k)S(z_n(k))-\varepsilon_{z_n}]+d_1(k)\\&=g_n(k)\left[\tilde{W}_n^{\mathrm{T}}(k)S(z_n(k))-\varepsilon_{z_n}+\frac{d_1(k)}{g_n(k)}\right]\\&=g_n(k)[\tilde{W}_n^{\mathrm{T}}(k)S(z_n(k))-\varepsilon'_{z_n}]\end{aligned} \tag{A.30}$$

with $\varepsilon'_{z_n}=\varepsilon_{z_n}-d_1(k)/g_n(k)$. It is obvious that ε'_{z_n} is bounded because of the boundedness of ε_{z_n}, $d_1(k)$ and $g_n(k)$. Choosing the Lyapunov function candidate,

$$V_n(k)=\sum_{j=1}^{n-1}V_j(k)+\frac{1}{\bar{g}_n}\eta_n^2(k)+\tilde{W}_n^{\mathrm{T}}(k)\Gamma_n^{-1}\tilde{W}_n(k). \tag{A.31}$$

The first difference of (A.31) along (A.29) and (A.30) is given as

$$\Delta V_n\leqslant-\sum_{j=1}^{n}\frac{\rho_j}{\bar{g}_j}\eta_j^2(k+1)-\sum_{j=1}^{n}\frac{1}{\bar{g}_j}\eta_j^2(k)+\beta_n-\sigma_n(1-\sigma_n\bar{\gamma}_n-\bar{g}_n\sigma_n\bar{\gamma}_n)\|\hat{W}_n(k)\|^2,$$

where ρ_j, $j=1,2,\ldots,n-1$, are defined as previous $(n-1)$ steps, and $\rho_n=1-\bar{\gamma}_n-\bar{\gamma}_n l_n-\bar{g}_n\bar{\gamma}_n l$, $\beta_n=\beta_{n-1}+\bar{g}_n\varepsilon'^2_{zn}/\bar{\gamma}_n+\sigma_n\|W_n^*\|^2$.

If we choose the design parameters as follows:

$$\bar{\gamma}_n<\frac{1}{1+l_n+\bar{g}_n l_n},\qquad \sigma_n<\frac{1}{(1+\bar{g}_n)\bar{\gamma}_n}, \tag{A.32}$$

then $\Delta V_n\leqslant 0$ once any one of the n errors satisfies $|\eta_j(k)|>\sqrt{\bar{g}_j\beta_n}$ $j=1,2,\ldots,n$. This demonstrates that the tracking error $\eta_1(k),\eta_2(k),\ldots,\eta_n(k)$ are bounded for all $k\geqslant 0$, and will asymptotically converge to the compact set denoted by Ω_n, where $\Omega_n:=\{\chi|\chi=[\chi_1,\chi_2,\ldots,\chi_n]^{\mathrm{T}},\chi_j\leqslant\sqrt{\bar{g}_j\beta_n}\ \ j=1,2,\ldots,n\}$. The boundedness of $\hat{W}_n(k)$, or equivalently of $\tilde{W}_n(k)$, can be proved as in Step 1.

Based on the procedure above, we can conclude that $\bar{\xi}_n(k+1)\in\Omega$ and $u(k)$ are bounded if $\bar{\xi}_n(k)\in\Omega$. Finally, if we initialize $\bar{\xi}_n(0)\in\Omega$, and choose the design parameters according to (A.11), (A.19), (A.24) and (A.32), there exists a k^*, such that all errors asymptotically converge to Ω_n, and NN weight errors are all bounded. This implies that the closed-loop system is SGUUB. Then $\bar{\xi}_n(k)\in\Omega$, $\hat{W}_i$, $i=1,2,\ldots,n$ will hold for all $k>0$. □

Appendix B. Proof of Lemma 3

Proof. From the new states definition in (26), $T(\bar{\xi}_n(k))$ is obviously a smooth map. However, due to the complexity of $f^c_{n,1}(\bar{\xi}_2(k)),\ldots,f^c_{2,1}(\bar{\xi}_n(k))$, it is not easy to check out whether the Jacobian Matrix of $T(\bar{\xi}_n(k))$ is of full rank. According to the definition of diffeomorphism, if we can find the unique transformation matrix $T^{-1}(x(k))$ and prove that the Jacobian Matrix of $T^{-1}(x(k))$ is of full rank, then $T(\bar{\xi}_n(k))$ must be a diffeomorphism.

From the definition of the new states in (26), it is clear that

$$\xi_1(k)=x_1(k), \tag{B.1}$$

$$\xi_1(k+1)=x_2(k). \tag{B.2}$$

From the first equation in (1), we obtain

$$\xi_2(k)=\frac{\xi_1(k+1)-f_1(\xi_1(k))}{g_1(\xi_1(k))}.$$

Substituting (B.1) and (B.2) into the above equation, we obtain

$$\xi_2(k)=t_2(x_1(k),x_2(k))\equiv t_2(k), \tag{B.3}$$

where

$$t_2(k)=\frac{x_2(k)-f_1(x_1(k))}{g_1(x_1(k))}.$$

It is clear that

$$\xi_2(k+1)=t_2(k+1), \tag{B.4}$$

$$\frac{\partial t_2(k)}{\partial x_2(k)}=\frac{1}{g_1(x_1(k))}, \tag{B.5}$$

$$\frac{\partial t_2(k+1)}{\partial x_3(k)}=\frac{1}{g_1(x_2(k))}. \tag{B.6}$$

From the second equation in (1), we obtain

$$\xi_3(k)=\frac{\xi_2(k+1)-f_2([\xi_1(k),\xi_2(k)]^{\mathrm{T}})}{g_2([\xi_1(k),\xi_2(k)]^{\mathrm{T}})}.$$

Substituting (B.1), (B.3) and (B.4) into the above equation, we obtain

$$\xi_3(k)=t_3(x_1(k),x_2(k),x_3(k))\equiv t_3(k), \tag{B.7}$$

where

$$t_3(k) = \frac{t_2(k+1) - f_2([x_1(k), t_2(k)]^{\mathrm{T}})}{g_2([x_1(k), t_2(k)]^{\mathrm{T}})}.$$

It is also clear that

$$\frac{\partial t_3(k)}{\partial x_3(k)} = \frac{\partial t_2(k+1)/\partial x_3(k)}{g_2([x_1(k), t_2(k)]^{\mathrm{T}})} = \frac{1}{g_1(x_2(k)) g_2([x_1(k), t_2(k)]^{\mathrm{T}})}.$$

Continue the process recursively until we find the final equation expressed as

$$\xi_n(k) = t_n(x_1(k), \ldots, x_n(k)) \equiv t_n(k), \tag{B.8}$$

$$\frac{\partial t_n(k)}{\partial x_n(k)} = \frac{1}{g_1(x_n(k)) \cdots g_n([x_1(k), \ldots, t_n(k)]^{\mathrm{T}})}. \tag{B.9}$$

Now, we can summarize that the unique transformation $T^{-1}(x(k))$ as $T^{-1}(x(k)) = [x_1(k), t_2(k), \ldots, t_n(k)]^{\mathrm{T}}$ and its Jacobian Matrix takes the triangular form of

$$\frac{\mathrm{d}T^{-1}(x(k))}{\mathrm{d}x(k)} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ * & \frac{1}{g_1(x_1(k))} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ * & * & \cdots & \frac{1}{g_1(x_n(k)) \cdots g_n([x_1(k), \ldots, t_n(k)]^{\mathrm{T}})} \end{bmatrix}. \tag{B.10}$$

Based on Assumption 1, Trace $(\mathrm{d}T^{-1}(x(k))/\mathrm{d}x(k)) \neq 0$, therefore T^{-1} is a diffeomorphism according to Lemma 2. Equivalently, T is a diffeomorphism. □

Appendix C. Proof of Theorem 2

Proof. Substituting controller (39) into (36), Eq. (36) can be re-written as

$$e_y(k+n) = -y_d(k+n) + f_0(\underline{z}(k)) + g_0(\underline{z}(k)) \hat{W}^{\mathrm{T}}(k) S(\bar{z}(k)). \tag{C.1}$$

Adding and subtracting $g_0(\underline{z}(k))\bar{u}^*(\bar{z}(k))$ on the right-hand side of (C.1) and noting (38), we have

$$e_y(k+n) = -y_d(k+n) + f_0(\underline{z}(k)) + g_0(\underline{z}(k))\bar{u}^*(\bar{z}(k)) + g_0(\underline{z}(k))[\hat{W}^{\mathrm{T}}(k) S(\bar{z}(k)) - W^{*\mathrm{T}} S(\bar{z}(k)) - \varepsilon_{\bar{z}}]. \tag{C.2}$$

Substituting (37) into (C.2) leads to

$$e_y(k+n) = g_0(\underline{z}(k))[\tilde{W}^{\mathrm{T}}(k) S(\bar{z}(k)) - \varepsilon_{\bar{z}}], \tag{C.3}$$

where $\tilde{W}(k) = \hat{W}(k) - W^*$.

Since NN approximation (6) and *Assumptions* 1–3 are only valid on the compact set Ω, it is necessary to guarantee the system's states remaining in Ω for all time. Noting that the system states are actually the unique transformation of the system outputs, therefore, $y_k \in \Omega_y\ \forall k$ can guarantee that $x(k) \in \Omega$ under Assumption 1.

Choose the following Lyapunov function candidate:

$$V(k) = \frac{1}{\bar{g}} e_y^2(k) + \sum_{j=0}^{n-1} \tilde{W}^{\mathrm{T}}(k_1 + j) \Gamma^{-1} \tilde{W}(k_1 + j). \tag{C.4}$$

Using the same techniques as in Appendix A, the first difference of (C.4) along (C.3) and (40) is given as

$$\Delta V \leqslant -\frac{\rho}{\bar{g}} e_y^2(k+1) - \frac{1}{\bar{g}} e_y^2(k) + \beta - \sigma(1 - \sigma\bar{\gamma} - \bar{g}\sigma\bar{\gamma}) \|\hat{W}(k_1)\|^2,$$

where $\rho = 1 - \bar{\gamma} - \bar{\gamma} l - \bar{g}\bar{\gamma} l$, and $\beta = \bar{g}\varepsilon_{\bar{z}}^2/\bar{\gamma} + \sigma\|W^*\|^2$.

If we choose the design parameters as follows:

$$\bar{\gamma} < \frac{1}{1 + l + \bar{g} l}, \quad \sigma < \frac{1}{(1 + \bar{g})\bar{\gamma}} \tag{C.5}$$

then $\Delta V \leqslant 0$ once the tracking error $e_y(k)$ is larger than $\sqrt{(\bar{g} + \Delta_{20})\beta}$. This implies the boundedness of $V(k)$ for all $k \geqslant 0$, which leads to the boundedness of $e_y(k)$. Furthermore, the tracking error $e_y(k)$ will asymptotically converge to the compact set denoted by $\varepsilon \leqslant \sqrt{\bar{g}\beta}$.

Due to negativity of ΔV, we can conclude that $y_{k+1} \in \Omega_y$ if all past outputs $y_{k-j} \in \Omega_y$, $j = 0, \ldots, n-1$, and compact set ε is small enough.

We have proved that $y_{k+1} \in \Omega_y$, we can use the same techniques as in Appendix A to show that the NN weight error stays in a small compact set Ω_{we}, and the control signal $u_k \in L_\infty$. Finally, if we initialize state $y_0 \in \Omega_{y0}$, $\tilde{W}(0) \in \Omega_{w_0}$, and we choose suitable parameters $\bar{\gamma}$, σ according to (C.5) to make ε small enough, there exists a constant k^* such that all tracking errors asymptotically converge to ε, and NN weight error asymptotically converges to Ω_{we} for all $k > k^*$. This implies that the closed-loop system is SGUUB. Then $y_k \in \Omega_y$ and $\hat{W}(k) \in L_\infty$ will hold for all $k > 0$. □

References

Anderson, B. D. O., Bitmead, R. R., Johnson, C. R. Jr., Kokotovic, P. V., Kosut, R. L., Mareels, I. M. Y., Praly, L., & Riedle, B. D. (1986). *Stability of adaptive systems: Passivity and averaging analysis*. Cambridge, MA: MIT Press.

Ge, S. S., Hang, C. C., Lee, T. H., & Zhang, T. (2001). *Stable adaptive neural network control*. Norwell, USA: Kluwer Academic.

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

Ge, S. S., Lee, T. H., Li, G. Y., & Zhang, J. (2003). Adaptive nn control for a class of discrete-time nonlinear systems. *International Journal of Control*, to appear.

Girosi, F., & Poggio, T. (1989). Networks and the best approximation property. *Artificial Intelligence Laboratory Memorandum* No. 1164, MIT, Cambridge.

Gupta, M. M., & Rao, D. H. (1994). Neuro-control systems: Theory and applications. New York:IEEE Press.

Hunt, K. J., Sbarbaro, D., Zbikowski, R., & Gawthrop, P. J. (1992). Neural networks for control system-a survey. *Automatica*, *28*(6), 1083–1112.

Kanellakopoulos, I., Kokotovic, P. V., & Morse, A. S. (1991). Systematic design of adaptive controller for feedback linearizable systems. *IEEE Transactions on Automatic Control*, *36*(11), 1241–1253.

Kosmatopoulos, E. B., Polycarpou, M. M., Christodoulou, M. A., & Ioannou, P. A. (1995). High-order neural network structures for identification of dynamical systems. *IEEE Transactions on Neural Networks*, *6*(2), 422–431.

Kosut, R. L., & Friedlander, B. (1985). Robust adaptive control: Conditions for global stability. *IEEE Transactions on Automatic Control*, *30*(7), 610–624.

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

Levin, A. U., & Narendra, K. S. (1996). Control of nonlinear dynamical systems using neural networks—Part II: Observability, identification, and control. *IEEE Transactions on Neural Networks*, *7*(1), 30–42.

Lewis, F. L., Yesildirek, A., & Liu, K. (1996). Multilayer neural-net robot controller with guaranteed tracking performance. *IEEE Transactions on Neural Networks*, *7*(2), 388–398.

Lin, Z., & Saberi, A. (1995). Robust semi-global stabilization of minimum-phase input-output linearizable systems via partial state and output feedback. *IEEE Transactions on Automatic Control*, *40*(6), 1029–1041.

Marino, R., & Tomei, P. (1995). *Nonlinear control design*. London: Prentice-Hall.

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

Nurnberger, G. (1989). *Approximation by spline functions*. New York: Springer.

Paretto, P., & Niez, J. J. (1986). Long term memory storage capacity of multiconnected neural networks. *Biological Cybernetics*, *54*, 53–63.

Polycarpou, M. M. (1996). Stable adaptive neural control scheme for nonlinear systems. *IEEE Transactions on Automatic Control*, *41*(3), 447–451.

Sadegh, N. (1993). A perception network for functional identification and control of nonlinear systems. *IEEE Transactions on Neural Networks*, *4*(6), 982–988.

Sanner, R. M., & Slotine, J. E. (1992). Gaussian networks for direct adaptive control. *IEEE Transactions on Neural Networks*, *3*(6), 837–863.

Seto, D., Annaswamy, A. M., & Baillieul, J. (1994). Adaptive control of nonlinear systems with a triangular structure. *IEEE Transactions on Automatic Control*, *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*(3), 339–359.

Yeh, P. C., & Kokotovic, P. V. (1995). Adaptive control of a class of nonlinear discrete-time systems. *International Journal of Control*, *62*(2), 303–324.

Yesidirek, A., & Lewis, F. L. (1995). Feedback linearization using neural networks. *Automatica*, *31*(11), 1659–1664.

Zhang, Y., Wen, C. Y., & Soh, Y. C. (2000). Discrete-time robust backstepping adaptive control for nonlinear time-varying systems. *IEEE Transactions on Automatic Control*, *45*(9), 1749–1755.