ELSEVIER

Journal of Process Control 9 (1998) 313–323

JOURNAL OF PROCESS CONTROL

Nonlinear adaptive control using neural networks and its application to CSTR systems

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

Department of Electrical Engineering, National University of Singapore, Singapore 119260

Abstract

In this paper, adaptive tracking control is considered for a class of general nonlinear systems using multilayer neural networks (MNNs). Firstly, the existence of an ideal implicit feedback linearization control (IFLC) is established based on implicit function theory. Then, MNNs are introduced to reconstruct this ideal IFLC to approximately realize feedback linearization. The proposed adaptive controller ensures that the system output tracks a given bounded reference signal and the tracking error converges to an ε-neighborhood of zero with ε being a small design parameter, while stability of the closed-loop system is guaranteed. The effectiveness of the proposed controller is illustrated through an application to composition control in a continuously stirred tank reactor (CSTR) system.

Keywords: Nonlinear systems; Input–output feedback linearization; Adaptive control; Multilayer neural networks; CSTR

1. Introduction

Recently, active research has been carried out on neural network control for nonlinear dynamic systems [1–16]. The massive parallelism, natural fault tolerance and implicit programming of neural network computing architectures suggest that they are good candidates for implementing real-time adaptive control for nonlinear dynamical systems. In the NN control design based on backpropagation (BP) learning algorithm [1–3], sufficient off-line training is required before control action is initiated for achieving a stable closed-loop system. To deal with such a requirement, stable adaptive NN design methods have been proposed based on Lyapunov stability theory in [4–9], which are applicable to affine nonlinear system described by

$$\begin{cases} \dot{x} = f_1(x) + g_1(x)u \\ y = h(x) \end{cases} \tag{1}$$

where $f_1(x)$ and $g_1(x)$ are smooth vector fields. In practice, many physical systems such as chemical reactions, PH neutralization and distillation columns are inherently nonlinear, whose input variables may enter in the systems nonlinearly as described by the general form

$$\begin{cases} \dot{x} = f(x, u) \\ y = h(x) \end{cases} \tag{2}$$

where $x \in R^n$ is the state vector, $u \in R$ is the input and $y \in R$ is the output, and $f(\cdot,\cdot)$ and $h(\cdot)$ are smooth vector fields. To solve the control problem for general nonlinear systems, neural networks have been used in [10–13] for emulating the 'inverse controller' to achieve tracking control through off-line training. It has been shown that the adaptive controller proposed in [13] can guarantee the stability of the closed-loop system when the uniformly persistently exciting (PE) condition is satisfied.

Multilayer neural networks are general tools for modelling nonlinear functions since they can approximate any continuous nonlinear function to any desired accuracy over a compact set [17]. In this work, a new direct adaptive controller is developed for non-affine system (2) using MNNs and Lyapunov stability technique. It is shown that the tracking error converges to a small neighborhood of zero and the closed-loop stability is guaranteed without the requirement for off-line training and PE condition. The paper is organized as follows. Section 2 describes the class of nonlinear systems under study and the concept of input–output

* Corresponding author. Tel: +65-874-6821; fax: +65-779-1103; e-mail: elegesz@nus.edu.sg

0959-1524/98/$ - see front matter
PII: S0959-1524(98)00054-7

linearization. In Section 3, structure and properties of MNNs are presented. An adaptive NN controller and stability analysis of the closed-loop system are discussed in Section 4. In Section 5, the proposed method is used to control a nonlinear CSTR system, in which the control input appears nonlinearly. Finally, Section 6 contains the conclusion.

2. Input–output feedback linearization

In the following study, $\|\cdot\|$ denotes the 2-norm and $\|\cdot\|_F$ is the Frobenius norm, i.e. given $A=[a_{ij}]\in R^{m\times n}$, $\| A \|_F^2 = tr\{A^TA\}=\sum_{i,j}a_{ij}^2$. We have the following properties for these definitions

$$\| Ax \| \leqslant \| A \|_F \cdot \| x \|, \quad \| A \|_F \leqslant \sum_{i,j}|a_{ij}| \tag{3}$$

2.1. Problem formulation

The control objective of this paper is described as: Given a desired trajectory, $y_d(t)$, find a control, $u(t)$, such that the system output tracks the desired trajectory with an acceptable accuracy, while all the closed-loop signals remain bounded. In the case of controlling affine nonlinear system (1), the problem is solvable if there exists a control of the form

$$u=\alpha_1(x)+\alpha_2(x)w \tag{4}$$

which results in a linear map from w to y with w being a new control variable. The existence of such a linearizing feedback is in turn guaranteed if system (1) possesses a so-called "relative degree" from u to y [18]. Nevertheless, for non-affine system (2) it is in general not possible to find such an explicit linearizing feedback to achieve feedback linearization. We next review the 'relative degree' introduced by Tsinias et al. [19] and establish an implicit feedback linearization control for general system (2).

Let L_fh denote the Lie derivate of the function $h(x)$ with respect to the vector field $f(x,u)$

$$L_fh=\frac{\partial[h(x)]}{\partial x}f(x,u)$$

Higher-order Lie derivatives can be defined recursively as $L_f^kh=L_f(L_f^{k-1}h)$, $k>1$.

Definition 1. System (2) is said to have relative degree ρ at (x_0,u_0) if there exists a positive integer $1\leqslant\rho<\infty$ such that

$$\frac{\partial[L_f^ih]}{\partial u}=0,\ i=0,1,\ldots,\rho-1$$

$$\frac{\partial[L_f^\rho h]}{\partial u}\neq 0$$

for all $(x,u)\in B(x_0,u_0)$, a ball centered at (x_0,u_0) [19].

Let $\Omega_x\subset R^n$ and $\Omega_u\subset R$ be compact subsets containing x_0 and u_0, respectively. System (2) is said to have a strong relative degree ρ in a compact set $U=\Omega_x\times\Omega_u$ if it has relative degree ρ at every point $(x_0,u_0)\in U$.

Assumption 1. System (2) possesses a strong relative degree $\rho=n$, $\forall(x,u)\in U$.

Under Assumption 1, system (2) is feedback linearizable and the mapping $\Phi(x)=[\phi_1(x),\phi_2(x),\ldots,\phi_n(x)$ with $\phi_j(x)=L_f^{j-1}h$, $j=1,2,\ldots,\rho$ has a Jacobian matrix which is nonsingular for all $x\in\Omega_x$ [18]. Therefore, $\Phi(x)$ is a diffeomorphism on Ω_x. By setting $\xi=\Phi(x)$, system (2) can be transformed into a normal form

$$\begin{cases}\dot{\xi}_i=\xi_{i+1}, & i=1,\ldots,n-1\\ \dot{\xi}_n=b(\xi,u) & \\ y=\xi_1 & \end{cases} \tag{5}$$

with $b(\xi,u)=L_f^nh$ and $x=\Phi^{-1}(\xi)$. Define the domain of normal system (5) as

$$\bar{U}:=\{(\xi,u)|\xi\in\Phi(\Omega_x);\ u\in\Omega_u\}$$

Let the smooth function $b_u:=\partial[b(\xi,u)]/\partial u$ and the continuous function $\dot{b}_u:=\mathrm{d}(b_u)/\mathrm{d}t$ with an understanding that u constructed later belongs to C^1. According to Assumption 1 and Definition 1, we have $\partial[b(\xi,u)]/\partial u\neq 0$, $\forall(\xi,u)\in\bar{U}$. This implies that the smooth function b_u is strictly either positive or negative for all $(\xi,u)\in\bar{U}$. We suppose that the sign of b_u is known, and without losing generality, we shall assume $b_u>0$ in the following discussion.

Assumption 2. There exist positive constants b_0, b_1 and b_2 such that $b_0\leqslant b_u\leqslant b_1$ and $|\dot{b}_u|\leqslant b_2$ for all $(\xi,u)\in\bar{U}$.

Remark 2.1. *From the definition of b_u, we know that b_u can be viewed as the control gain of normal system (5). The requirement for $b_u\geqslant b_0$ means that the control gain of the system is larger than a positive constant. Many feedback linearization methods for affine nonlinear systems need such an assumption [18,20,21]. We also require the absolute values of b_u and $\dot{b}_u$ being bounded by positive constants b_1 and b_2, respectively. In general, this does not pose a strong restriction upon the class of systems. The reason is that if the controller is continuous, the situation in which a finite input causes an infinitely large effect upon the system rarely happen in physical systems due to the smoothness of b_u.*

2.2. Ideal implicit feedback linearization control

Define vectors ξ_d, $\bar{\xi}_d$ and $\tilde{\xi}$ as

$$\xi_d = [y_d, \dot{y}_d, \ldots, y_d^{(n-1)}]^T \in R^n, \quad \bar{\xi}_d = [\xi_d^T, y_d^{(n)}]^T \in R^{n+1}$$
$$\tilde{\xi} = \xi - \xi_d = [\tilde{\xi}_1, \tilde{\xi}_2, \ldots, \tilde{\xi}_n]^T \tag{6}$$

and a filtered tracking error as

$$e_s = [\Lambda^T \; 1]\tilde{\xi} \tag{7}$$

where $\Lambda = [\lambda_1, \lambda_2, \ldots, \lambda_{n-1}]^T$ is an appropriately chosen coefficient vector such that $\tilde{\xi}(t) \to 0$ as $e_s \to 0$, (i.e. $s^{n-1} + \lambda_{n-1}s^{n-2} + \ldots + \lambda_1$ is Hurwitz).

Assumption 3. The desired trajectory vector $\bar{\xi}_d$ is continuous and available, and $\| \bar{\xi}_d \| \leqslant c$ with c being a known bound.

From system (5), the time derivative of the filtered tracking error can be written as

$$\dot{e}_s = b(\xi, u) - y_d^{(n)} + [0 \; \Lambda^T]\tilde{\xi}. \tag{8}$$

We have the following lemma to establish the existence of an ideal implicit feedback linearization control (IFLC) input, u^*, which can linearize system (2) and bring the system output to the desired trajectory $y_d(t)$.

Lemma. *Consider system (2) satisfying Assumptions 1 to 3. For a given small constant $\varepsilon > 0$, there exist a compact subset $\Phi_0 \subset \Phi(\Omega_x)$ and an unique ideal IFLC input $u^* \in \Omega_u$ such that for all $\xi(0) \in \Phi_0$, the error Eq. (8) can be written in a linear form*

$$\dot{e}_s = -\frac{e_s}{\varepsilon} \tag{9}$$

which leads to $\lim_{t\to\infty} |y(t) - y_d(t)| = 0$.

Proof. Adding and subtracting e_s/ε to the right-hand side of the error Eq. (8), we obtain

$$\dot{e}_s = b(\xi, u) + \nu - \frac{1}{\varepsilon}e_s \tag{10}$$

where ν is defined as

$$\nu = \frac{1}{\varepsilon}e_s - y_d^{(n)} + [0 \; \Lambda^T]\tilde{\xi} \tag{11}$$

Considering (6) and (11), we have $\partial\nu/\partial u = 0$. From Assumption 1, $\partial[b(\xi, u)]/\partial u \neq 0$ holds for all $(\xi, u) \in \bar{U}$. Thus

$$\frac{\partial[b(\xi, u) + \nu]}{\partial u} \neq 0, \forall(\xi, u) \in \bar{U}.$$

Using the implicit function theorem [22], for every desired trajectory $y_d(t)$ satisfying Assumption 3, there exist a compact subset $\Phi_0 \subset \Phi(\Omega_x)$ and a unique local solution $u = \alpha(\xi, \nu) \in \Omega_u$ such that $b(\xi, \alpha(\xi, \nu)) + \nu = 0$ holds for all $\xi(0) \in \Phi_0$. Thus, if we choose the ideal IFLC input as

$$u^*(z) = \alpha(\xi, \nu), z = [\xi^T, \xi_d^T, \nu]^T \in \Omega_z \subset R^{2n+1} \tag{12}$$

where the compact set Ω_z is defined as

$$\Omega_z = \Big\{(\xi, \xi_d, \nu) | \xi \in \Phi(\Omega_x), \| \bar{\xi}_d \| \leqslant c; \; \nu = \frac{1}{\varepsilon}e_s - y_d^{(n)} + [0\Lambda^T]\tilde{\xi}\Big\} \tag{13}$$

then

$$b(\xi, u^*) + \nu = 0. \tag{14}$$

Therefore, under the action of u^*, Eqs. (10) and (14) imply that (9) holds. Since $\varepsilon > 0$, Eq. (9) is asymptotically stable, i.e. $\lim_{t\to\infty} |e_s| = 0$. Because $s^{n-1} + \lambda_{n-1} s^{n-2} + \ldots + \lambda_1$ is Hurwitz, we know that $\lim_{t\to\infty} |y(t) - y_d(t)| = 0$. Q.E.D.

Remark 2.2. *In the above lemma, we assume that Ω_u is large enough to achieve the control objective, i.e. a sufficient strong control input can be realized. If the region of Ω_u is not large enough, the upper bound of the desired trajectory $\bar{\xi}_d$ and the lower bound of parameter ε should be restricted.*

The above lemma only suggests the existence of the ideal IFLC input u^*, it does not provide the method of constructing it. In Section 3, we will use multilayer neural networks to reconstruct the ideal IFLC u^* for realizing feedback linearization approximately.

3. Multilayer neural networks

A three-layer NN with one net output and $2n + 1$ inputs is given by [5]

$$g(z) = \sum_{j=1}^{l}\left[w_j s\left(\sum_{k=1}^{2n+1} v_{jk}z_k + \theta_{vj}\right) + \theta_w\right] \tag{15}$$

where the activation function $s(\cdot)$ is chosen as

$$s(z_a) = \frac{1}{1 + e^{-z_a}}, z_a \in R \tag{16}$$

v_{jk} are the first-to-second layer interconnection weights, w_j are the second-to-third layer interconnection weights, θ_w and θ_{vj} are the threshold offsets, the number of hidden-layer neurons is l, and the number of input-layer neurons is $2n + 1$. Define

$$\begin{aligned}\bar{z} &= [\bar{z}_1, \bar{z}_2, \ldots, \bar{z}_{2n+2}]^T = [z^T, 1]^T \in R^{2n+2} \\ W &= [w_1, w_2, \ldots, w_l]^T \in R^{l\times 1} \\ V &= [v_1, v_2, \ldots, v_l]^T \in R^{(2n+2)\times l}\end{aligned} \tag{17}$$

with $v_i = [v_{i1}, v_{i2}, \ldots, v_{i2n+2}]^T,\ i = 1, 2, \ldots, l$. The term $\bar{z}_{2n+2} = 1$ in input vector $\bar{z}$ allows one to include the threshold vector $[\theta_{v1}, \theta_{v2}, \ldots, \theta_{vl}]^T$ as the last column of V^T, so that V contains both the weights and thresholds of the first-to-second layer connections. Similarly, one can incorporate the threshold θ_w in the weight W. Any tuning of W and V then includes the tuning of the thresholds automatically. The three-layer NN (15) can be conveniently expressed as

$$g(z) = W^T S(V^T \bar{z})$$

$$S(V^T\bar{z}) = [s(v_1^T\bar{z}), s(v_2^T\bar{z}), \ldots, s(v_l^T\bar{z})]^T$$

It has been proven that neural network $W^T S(V^T\bar{z})$ satisfies the conditions of the Stone–Weierstrass theorem and can therefore approximate any continuous function to any desired accuracy over a compact set [17]. Because the ideal IFLC input $u^*(z)$ defined in (12) is a continuous function on the compact set Ω_z, for an arbitrary constant $\varepsilon_N > 0$, there exists an integer l (the number of hidden neurons) and ideal constant weight matrices $W^* \in R^{l\times 1}$ and $V^* \in R^{(2n+2)\times l}$, such that

$$u^*(z) = W^{*T} S(V^{*T}\bar{z}) + u_k(z) + \varepsilon_u(z), \forall z \in \Omega_z \tag{18}$$

where $\varepsilon_u(z)$ is called the NN approximation error satisfying $|\varepsilon_u(z)| \leqslant \varepsilon_N, \forall z \in \Omega_z$, and the term $u_k(z)$ in (18) is a prior continuous controller (possibly PI, PID or some other type of controller), which is perhaps specified via heuristics or past experience with the application of conventional direct control. The ideal constant weights W^* and V^* are defined as

$$(W^*, V^*) := \arg\min_{(W,V)} \left\{ \sup_{z\in\Omega_z} |W^T S(V^T\bar{z}) + u_k(z) - u^*(z)| \right\}$$

Assumption 4. On the compact set Ω_z, the ideal neural network weights W^*, V^* and the NN approximation error are bounded by

$$\| W^* \| \leqslant w_m, \| V^* \|_F \leqslant v_m, |\varepsilon_u(z)| \leqslant \varepsilon_l \tag{19}$$

with w_m, v_m and ε_l being positive constants.

Let $\hat{W}$ and $\hat{V}$ be the estimates of W^* and V^*, respectively. The estimation errors of the weight matrices are defined as

$$\tilde{W} = \hat{W} - W^*, \tilde{V} = \hat{V} - V^* \tag{20}$$

The Taylor series expansion $S(V^{*T}\bar{z})$ for a given $\hat{V}^T\bar{z}$ may be written as

$$S(V^{*T}\bar{z}) = S(c\hat{V}^T\bar{z}) - \hat{S}'\tilde{V}^T\bar{z} + O(\tilde{V}^T\bar{z})^2. \tag{21}$$

where $\hat{S}' = \mathrm{diag}\{\hat{s}_1', \hat{s}_2', \ldots, \hat{s}_l'\}$ with $\hat{s}_i' = s'(\hat{v}_i^T\bar{z}) = d[s(z_a)]/dz_a|_{z_a=\hat{v}_i^T\bar{z}}, i = 1, 2, \ldots, l$ and $O(\tilde{V}^T\bar{z})^2$ denotes the sum of the high order terms in the Taylor series expansion. Using (20) and (21), the function approximation error can be written as

$$\begin{aligned}&\hat{W}^T S(\hat{V}^T\bar{z}) - W^{*T} S(V^{*T}\bar{z}) \\ &= (\tilde{W} + W^*)^T S(\hat{V}^T\bar{z}) - W^{*T}[s(\hat{V}^T\bar{z}) \\ &- \hat{S}'\tilde{V}^T\bar{z} + O(\tilde{V}^T\bar{z})^2] \\ &= \tilde{W}^T S(\hat{V}^T\bar{z}) + (\hat{W} - \tilde{W})^T\hat{S}'\tilde{V}^T\bar{z} - W^{*T}O(\tilde{V}^T\bar{z})^2 \\ &= \tilde{W}^T S(\hat{V}^T\bar{z}) + \hat{W}^T\hat{S}'\tilde{V}^T\bar{z} \\ &- \tilde{W}^T\hat{S}'(\hat{V} - V^*)^T\bar{z} - W^{*T}O(\tilde{V}^T\bar{z})^2 \\ &= \tilde{W}^T(\hat{S} - \hat{S}'\hat{V}^T\bar{z}) + \hat{W}^T\hat{S}'\tilde{V}^T\bar{z} + d_u\end{aligned} \tag{22}$$

where $\hat{S} = S(\hat{V}^T\bar{z})$ and the residual term d_u is defined as

$$d_u = \tilde{W}^T\hat{S}'V^{*T}\bar{z} - W^{*T}O(\tilde{V}^T\bar{z})^2 \tag{23}$$

Lemma 3.1. *There exist positive constants* α_1 *to* α_4 *such that the residual term* d_u *is bounded by*

$$|d_u| \leqslant \alpha_1 + \alpha_2|e_s| + \alpha_3 \| \tilde{W} \| + \alpha_4 \| \tilde{W} \| |e_s| \tag{24}$$

Proof. See Appendix A.

4. Direct adaptive control using MNNs

Let the MNN controller take the form

$$u = \hat{W}^T S(\hat{V}^T\bar{z}) + u_k(z) + u_b \tag{25}$$

$$u_b = -k_s|e_s|e_s \tag{26}$$

where $k_s > 0$ is a design parameter. The first term $\hat{W}^T S(\hat{V}^T\bar{z})$ in controller (25) is used to approximate the ideal IFLC input u^* for feedback-linearization. The second term, $u_k(z)$, is a prior control term based on prior model or past experience to improve the initial control performance. If such a priori knowledge is not available, u_k can be simply set to zero. The third term, u_b, is called the bounding control term, which is applied for guaranteeing the boundedness of the system states.

By using Mean Value Theorem in [23], there exists $\lambda \in (0, 1)$ such that

$$b(\xi, u) = b(\xi, u^*) + b_{u_\lambda}(u - u^*) \tag{27}$$

where $b_{u_\lambda} = \partial[b(\xi, u)]/\partial u|_{u=u_\lambda}$ with $u_\lambda = \lambda u + (1-\lambda)u^*$. Considering (14) and (27), the error system (10) can be rewritten as

$$\dot{e}_s = -\frac{1}{\varepsilon}e_s + b_{u_\lambda}(u - u^*).$$

Since $b_1 \geqslant b_{u_\lambda} \geqslant b_0 > 0$ (Assumption 2), we have

$$b_{u_\lambda}^{-1}\dot{e}_s = -\frac{1}{\varepsilon}b_{u_\lambda}^{-1}e_s + u - u^* \tag{28}$$

Substituting (18) and (25) into (28) and noting (22), we obtain

$$\begin{aligned} b_{u_\lambda}^{-1}\dot{e}_s = &-\frac{1}{\varepsilon}b_{u_\lambda}^{-1}e_s + \tilde{W}^T(\hat{S} - \hat{S}'\hat{V}^T\bar{z}) \\ &+ \hat{W}^T\hat{S}'\tilde{V}^T\bar{z} - k_s|e_s|e_s - \varepsilon_u(z) + d_u \end{aligned} \tag{29}$$

The MNN weight updating algorithms are presented as follows

$$\dot{\hat{W}} = -\Gamma_w[(\hat{S} - \hat{S}'\hat{V}^T\bar{z})e_s + \delta_w(1 + e_s^2)\hat{W}] \tag{30}$$

$$\dot{\hat{V}} = -\Gamma_v[\bar{z}\hat{W}^T\hat{S}'e_s + \delta_v\hat{V}] \tag{31}$$

where $\Gamma_w = \Gamma_w^T > 0$, $\Gamma_v = \Gamma_v^T > 0$, $\delta_w > 0$ and $\delta_v > 0$ are constant design parameters. The first terms on the right-hand sides of (30) and (31) are the modified back-propagation algorithms and the last terms of them correspond to the $\sigma-$modification [25] terms in adaptive control, which are helpful to guarantee bounded parameter estimates in the presence of NN approximation error. The main results of this paper are given in the following theorem.

Theorem 4.1 *For the closed-loop system (2), (25), (30) and (31), there exist a compact subset* Ω_0 and positive constants c^* *and* k_s^* *such that if*

i. *all initial states* $\xi(0) \in \Phi_0$, $(\hat{W}(0), \hat{V}(0)) \in \Omega_0$, *and*
ii. $c \leqslant c^*$ *and* $k_s \geqslant k_s^*$,

then the tracking error converges to an $\varepsilon-$*neighborhood of the origin and all the states and control input of the system remain in the compact set* $\bar{U}$.

Proof. See Appendix B.

Remark 4.1. *Compared with the traditional exact linearization techniques, the proposed adaptive NN control method needs not to search for an explicit controller to cancel the nonlinearities of the system. In fact, even if functions* $f(x, u)$ *and* $h(x)$ *in (2) are known exactly, there does not always exist an explicit controller for feedback linearization. Instead of solving the implicit function to get the ideal controller* u^*, *multilayer neural networks are applied to reconstruct the ideal IFLC input* u^* *for achieving approximate feedback linearization.*

Remark 4.2. *The bounding control term,* u_b, *in controller (25) is used to limit the upper bounds of the system states, therefore guarantees the validities of the strong relative degree and the NN approximation in assumptions 1 and 4, respectively. The developed adaptive method also accommodates the existence of prior control term,* u_k, *such that control engineers can plug in the proposed NN technique to work in parallel to improve the control performance.*

Remark 4.3. *The choices of controller parameters are important for obtaining a good control performance and worth to be discussed. It can be seen from (53) that the larger the number of NN nodes is, the smaller the tracking error will be. Therefore, if a high tracking accuracy is required, small* ε, *large* k_s *and* l *should be chosen. However, the multilayer neural networks with a larger hidden node number* l *usually make the controller more complex and increase the computation burden for the control system. Hence, there is a design trade-off between MNN complexity and system tracking error accuracy.*

Remark 4.4. *Theorem 4.1 shows that the boundedness of the system states and the convergence of tracking error are guaranteed without the requirement of PE condition, which is very important for practical applications because PE condition is difficult to check/guarantee in adaptive systems. The results indicate that although convergence of the NN weight estimates is not achieved in the absence of PE condition, the tracking error is guaranteed to converge to an* $\varepsilon-$*neighborhood of zero.*

5. Application

The continuously stirred tank reactor (CSTR) system given by Lightbody and Irwin [16] is shown in Fig. 1. This system consists of a constant volume reactor cooled by a single coolant stream flowing in a cocurrent fashion. An irreversible, exothermic reaction, $A \to B$, occurs in the tank. The objective is to control the effluent concentration, C_a, by manipulating the coolant flow rate, q_c. The process is described by the following differential equations

$$\dot{C}_a = \frac{q}{V}(C_{a0} - C_a) - a_0 C_a \mathrm{e}^{-\frac{E}{RT_a}} \tag{32}$$

$$\begin{aligned} \dot{T}_a = &\frac{q}{V}(T_f - T_a) + a_1 C_a \mathrm{e}^{-\frac{E}{RT_a}} \\ &+ a_3 q_c[1 - \mathrm{e}^{-\frac{a_2}{q_c}}](T_{cf} - T_a) \end{aligned} \tag{33}$$

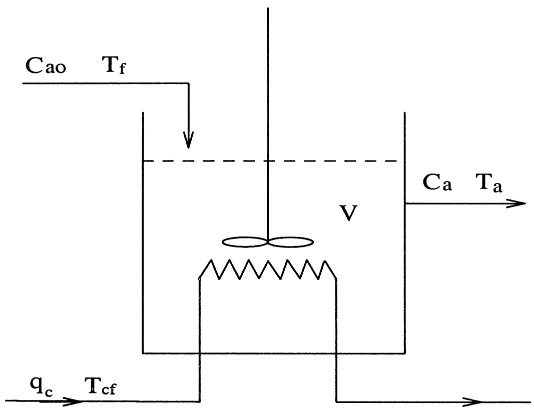

Fig. 1. Continuously stirred tank reactor (CSTR) system.

where variables C_a and T_a are the concentration and temperature of the tank, respectively; the coolant flow rate q_c is the control input; and the parameters of the system are given in Table 1.

Define the state variables, input and output as

$$x = [x_1, x_2]^T = [C_a, T_a]^T, u = q_c, y = C_a$$

System (32)–(33) can be written in the form of system (2):

$$\dot{x} = f(x, u) = \begin{bmatrix} 1 - x_1 - a_0 x_1 e^{-\frac{10^4}{x_2}} \\ 350 - x_2 + a_1 x_1 e^{-\frac{10^4}{x_2}} + a_3 u (1 - e^{-\frac{a_2}{u}})(350 - x_2) \end{bmatrix} \tag{34}$$

$$y = h(x) = x_1$$

Obviously, the model is not in affine form, because the control input, u, appears nonlinearly in (34).

From the parameters given in Table 1 and the irreversible exothermic property of the chemical process, we obtain the operating region of the states and control input as follows

$$0 < x_1 < 1, h_1 \geqslant x_2 > 350, 0 \leqslant u \leqslant h_2 \tag{35}$$

where constant h_1 is the highest temperature of the reactor and constant h_2 is the maximum value of the coolant flow rate. Since

$$\dot{y} = L_f h = 1 - x_1 - a_0 x_1 e^{-\frac{10^4}{x_2}}, \frac{\partial \dot{y}}{\partial u} = 0,$$

$$\ddot{y} = L_f^2 h = -\dot{x}_1 - a_0 \left(\dot{x}_1 + \frac{10^4 x_1 \dot{x}_2}{x_2^2} \right) e^{-\frac{10^4}{x_2}}$$

and

$$b_u = \frac{\partial L_f^2 h}{\partial u} = 10^4 a_0 a_3 x_1 e^{-\frac{10^4}{x_2}} \frac{(x_2 - 350)}{x_2^2} \left(1 - e^{-\frac{a_2}{u}} - \frac{a_2}{u} e^{-\frac{a_2}{u}} \right)$$

we have

$$1 - e^{-\frac{a_2}{u}} - \frac{a_2}{u} e^{-\frac{a_2}{u}} = 1 - \left(1 + \frac{a_2}{u} \right) e^{-(1+\frac{a_2}{u})} e > 0$$

by noticing the fact that $(1 + w)e^{-(1+w)} < e^{-1}$ for all $w > 0$. Consider the parameters shown in Table 1 and the operating region (35), we know that there indeed exist positive constants b_0 and b_1 such that $b_0 \leqslant b_u \leqslant b_1$. Therefore the plant is of relative degree 2 and Assumptions 1 and 2 are satisfied under operating condition (35). The above analysis is just to check the validity of the assumptions made in this paper. Since the exact model of CSTR is difficult to be known in practice, in the following simulation study, we design an adaptive NN controller without the prior model of the CSTR. The control objective is to make the concentration, y, track the set-point step change signal $r(t)$.

To show the effectiveness of the proposed method, two controllers are studied for comparison. A fixed-gain proportional plus integral (PI) control law is first used as follows

$$u_{pi} = k_c(r - y) + \frac{k_c}{T_i} \int_0^t (r - y) d\tau \tag{36}$$

with $k_c = 440\ L^2/\text{mol min}$ and $T_i = 0.25$ min [16]. The parameters of the PI controller are selected to give an adequate response for step changes $r(t)$ in the set-points of ± 0.02 mol/l about the nominal product concentration of 0.1 mol/l. The dashed lines in Figs. 2–4 indicate the output, tracking error and input of the system using PI controller (36), respectively. The PI controller cannot provide a good tracking response due to the effects of the nonlinearities.

The adaptive controller based on MNNs proposed in Section 4 is applied to the CSTR system. In order to obtain a smooth reference signal, we use a linear reference model to shape the discontinuous reference signal $r(t)$ to get the desired signals y_d, $\dot{y}_d$ and $\ddot{y}_d$. The reference model is taken as

$$\frac{y_d(s)}{r(s)} = \frac{w_n^2}{s^2 + 2\zeta_n w_n s + w_n^2}$$

where $w_n = 5.0$ rad/min and $\zeta_n = 1.0$. The signals e_s and ν are generated by

Table 1
CSTR parameters

Parameter	Description	Nominal value
q	Process flowrate	100 ml/min
C_{a0}	Concentration of component A	1 mol/l
T_f	Feed temperature	350 K
T_{cf}	Inlet coolant temperature	350 K
V	Volume of tank	100 l
h_a	Heat transfer coefficient	7×10^5 J/min K
a_0	Preexponential factor	7.2×10^{10} min^{-1}
E/R	Activation energy	1×10^4 K
$(-\Delta H)$	Heat of reaction	2×10^4 cal/mol
ρ_1, ρ_c	Liquid densities	1×10^3 g/l
C_p, C_{pc}	Heat capacities	1 cal/g K
$a_1 = \frac{(-\Delta H)a_0}{\rho_1 C_p} = 1.44\times10^{13}$	$a_2 = \frac{h_a}{\rho_c C_{pc}} = 6.987\times10^2$	$a_3 = \frac{\rho_c C_{pc}}{\rho_1 C_p V} = 0.01$

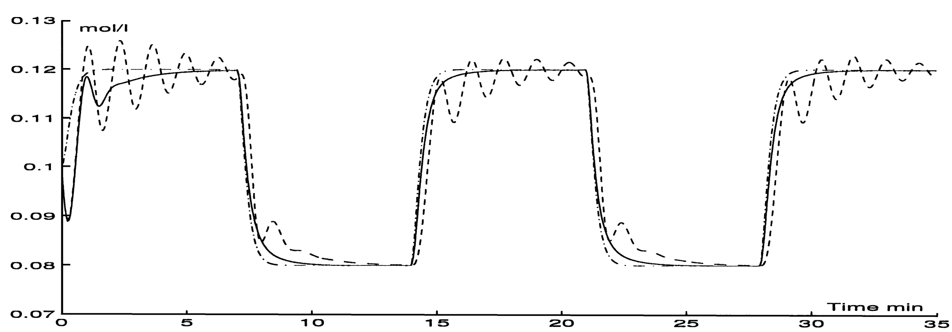

Fig. 2. Output tracking (y_d, ---; PI control - -; NN control —-).

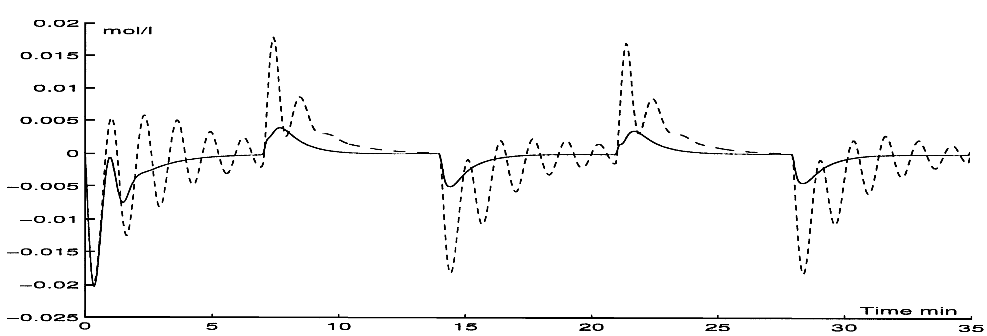

Fig. 3. Tracking errors $y - y_d$ (PI control - -; NN control —-).

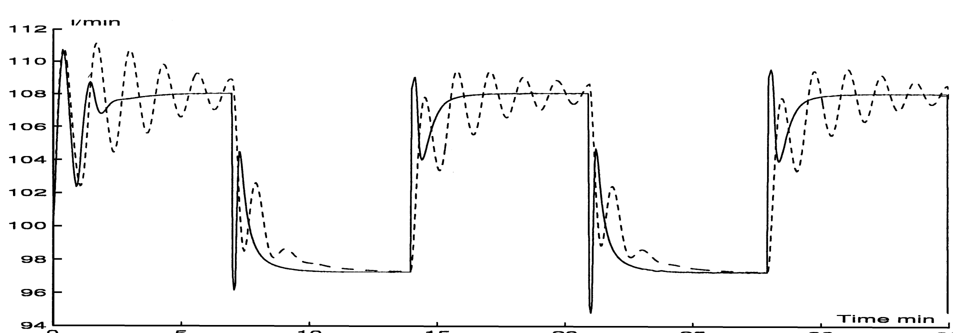

Fig. 4. Control input (PI control - -, NN control —-).

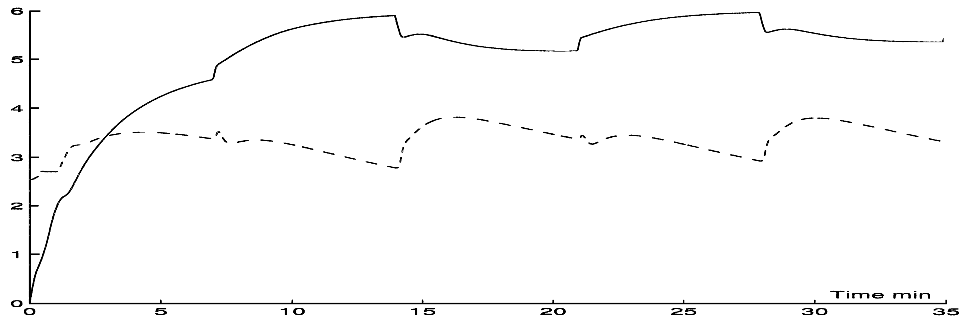

Fig. 5. NN weights $\| \hat{W} \|$ (—) and $|\hat{V}|_F$ (- -).

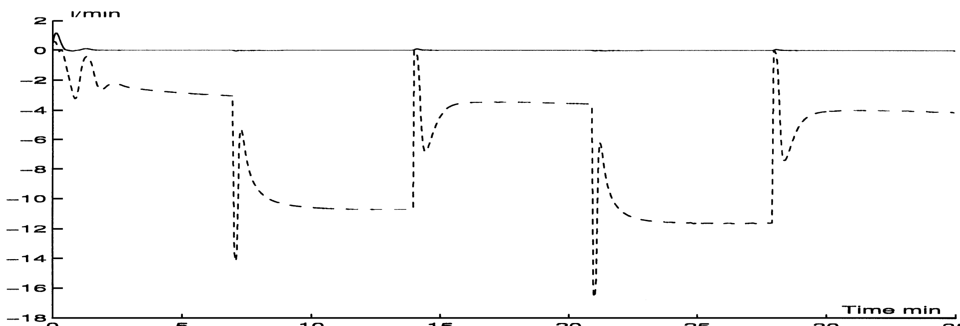

Fig. 6. Responses of the bounding control u_b (—) and NN output (- -).

$$e_s = \lambda_1(y - y_d) + \dot{y} - \dot{y}_d$$

$$\nu = \frac{1}{\varepsilon} e_s - \ddot{y}_d + \lambda_1(\dot{y} - \dot{y}_d)$$

with $\lambda_1 = 3.0$ and $\varepsilon = 0.1$. The controller is taken as

$$u(t) = \hat{W}^T S(\hat{V}^T \bar{z}) + u_{pi} - k_s |e_s| e_s \tag{37}$$

The input vector of MNNs is chosen as $\bar{z} = [y, \dot{y}, y_d, \dot{y}_d, \nu, 1]^T$. The gain $k_s = 100.0$ and the NN node number $l = 10$ are used in the simulation. The parameters in the weight updating laws (30) and (31) are chosen as $\Gamma_w = \text{diag}\{2.0\}$, $\Gamma_v = \text{diag}\{100.0\}$, $\delta_w = 0.2$ and $\delta_v = 4.0 \times 10^{-4}$. The initial conditions are $x(0) = [0.1, 440.0]^T$, $u_{pi}(0) = 100.0$, $y_d(0) = 0.1$, $\dot{y}_d(0) = 0.0$, $\ddot{y}_d(0) = 0.0$ and $\hat{W}(0) = (0)$. The initial weight $\hat{V}(0)$ is taken randomly in the interval [−1,1].

The solid lines in Figs. 2–4 shows the output, tracking error and input of the system using adaptive MNN controller (37). Comparing the set-point responses and the tracking errors in Figs. 2 and 3, respectively, the adaptive NN controller presents better control performance than that of the PI controller. It should be noticed that though we choose the initial weight $\hat{W}(0) = 0.0$ and $\hat{V}(0)$ randomly, the set-point tracking of the NN controller is still guaranteed during the initial period. This is due to the choice of the PI control, u_{pi}, as the prior control term. Fig. 4 indicates that a stronger control action is needed when using NN controller (37), which is not surprising from a practical point of view. The boundedness of the NN weight estimates are indicated in Fig. 5. To show the effectiveness of neural networks further, the responses of the bounding controller and neural network outputs are shown in Fig. 6. It is found that the magnitude of u_b is much smaller than that of neural networks. Therefore, the neural networks plays a major role in improving the control performance.

6. Conclusion

In this paper, a novel adaptive control design approach has been developed for a general class of nonlinear systems based on multilayer neural networks. The proposed adaptive controller guarantees the tracking error converging to a small neighborhood of zero, while the closed-loop system is proven to be regionally stable. The rigorous theoretical analysis and application study show that the developed approach possesses great potential for the control of general nonlinear systems.

Appendix A. Proof of Lemma 3.1

The derivative of the sigmoid activation function (16) with respect to z_a is

$$s'(z_a) = \frac{d[s(z_a)]}{dz_a} = \frac{e^{-z_a}}{(1+e^{-z_a})^2} \tag{38}$$

It is easy to check that $0 \leqslant s'(z_a) \leqslant 0.25$ and $|z_a s'(z_a)| \leqslant 0.2239$, $\forall z_a \in R$. Using the properties of the Frobenius norm in (3), we have

$$\begin{aligned} \| \hat{S}' \| &\leqslant \sum_{i=1l}^{l} s'(\hat{v}_i^T \bar{z}) \leqslant 0.25l, \\ \| \hat{S}'\hat{V}^T\bar{z} \| &\leqslant \sum_{i=1}^{l} |\hat{v}_i^T \bar{z} s'(\hat{v}_i^T \bar{z})| \leqslant 0.2239l \end{aligned} \tag{39}$$

From (21), we know that the high order term $O(\tilde{V}^T\bar{z})^2$ is bounded by

$$\begin{aligned} \| O(\tilde{V}^T\bar{z})^2 \| &\leqslant \| \hat{S}'\tilde{V}^T\bar{z} \| + \| S(V^{*T}\bar{z}) - S(\hat{V}^T\bar{z}) \| \\ &\leqslant \| \hat{S}'\hat{V}^T\bar{z} \| + \| \hat{S}'V^{*T}\bar{z} \| + \| S(V^{*T}\bar{z}) - S(\hat{V}^T\bar{z}) \| \\ &\leqslant \| \hat{S}'\hat{V}^T\bar{z} \| + \| \hat{S}' \|_F \cdot \| V^* \|_F \cdot \bar{z} \| + \| S(V^{*T}\bar{z}) - S(\hat{V}^T\bar{z}) \| \end{aligned} \tag{40}$$

Considering (39), $\| V^* \|_F \leqslant v_m$ (Assumption 4) and $\| S(V^{*T}\bar{z}) \| \leqslant l$, we have

$$\| O(\tilde{V}^T\bar{z})^2 \| \leqslant 1.2239l + 0.25v_m l \| \bar{z} \| \tag{41}$$

By (23), it can be seen that

$$\begin{aligned} |d_u| &\leqslant |\tilde{W}^T\hat{S}'V^{*T}\bar{z}| + |W^{*T}O(\tilde{V}^T\bar{z})^2| \\ &\leqslant v_m \| \tilde{W} \| \cdot \| \hat{S}' \|_F \cdot \| \bar{z} \| + w_m \| O(\tilde{V}^T\bar{z})^2 \| \end{aligned} \tag{42}$$

Using (39)–(41), inequality (42) can be further written as

$$|d_u| \leqslant 0.25v_m l \| \tilde{W} \| \| \bar{z} \| + w_m(1.2239l + 0.25v_m \| \bar{z} \|) \tag{43}$$

By choosing $c_0 = 1.2239w_m l$, $c_1 = 0.25w_m v_m$ and $c_2 = 0.25v_m l$, it follows that

$$|d_u| \leqslant c_0 + c_1 \| \bar{z} \| + c_2 \| \tilde{W} \| \| \bar{z} \| \tag{44}$$

According to the definitions of z and $\bar{z}$ in (12) and (17), respectively, $\bar{z}$ is bounded by

$$\| \bar{z} \| \leqslant \| z \| + 1 \leqslant \| \xi \| + \| \xi_d \| + |\nu| + 1 \tag{45}$$

Considering (6), (7) and Assumption 3, we can derive $\| \xi \| \leqslant d_1 \| \xi_d \| + d_2|e_s|$ with constants $d_1, d_2 > 0$. Using (11), there exist computable positive constants c_3 and c_4, such that $\| \bar{z} \| \leqslant c_3 + c_4|e_s|$. Substituting it into (44) and letting $\alpha_1 = c_0 + c_1c_3, \alpha_2 = c_1c_4, \alpha_3 = c_2c_3$ and $\alpha_4 = c_2c_4$, we conclude that (24) holds. Q.E.D.

Appendix B. Proof of Theorem 4.1

The proof includes two parts. We first assume that the system trajectories remain in the compact set $\bar{U}$, in which the NN approximation in Assumption 4 is valid. With this assumption, we prove the tracking error converging to an ε–neighborhood of the origin. Later, we show that for a suitable reference signal $y_d(t)$ and the proper choice of controller parameters, the trajectories do remain in the compact set $\bar{U}$ for all time.

Part 1: Choosing the Lyapunov function candidate

$$V_1 = \frac{1}{2}[b_{u_\lambda}^{-1}e_s^2 + \tilde{W}^T\Gamma_w^{-1}\tilde{W} + \mathrm{tr}\{\tilde{V}^T\Gamma_v^{-1}\tilde{V}\}] \tag{46}$$

and differentiating (46) along (29)–(31), we have

$$\begin{aligned} \dot{V}_1 &= e_s\Big[-\frac{1}{\varepsilon}b_{u_\lambda}^{-1}e_s + \tilde{W}^T(\hat{S} - \hat{S}'\hat{V}^T\bar{z}) + \hat{W}^T\hat{S}'\tilde{V}^T\bar{z} + d_u \\ &\quad - k_s|e_s|e_s - \varepsilon_u(z)\Big] + \frac{1}{2}\frac{d(b_{u_\lambda}^{-1})}{dt}e_s^2 + \tilde{W}^T\Gamma_w^{-1}\dot{\hat{W}} \\ &\quad + \mathrm{tr}\left\{\tilde{V}^T\Gamma_v^{-1}\dot{\hat{V}}\right\} = -k_s|e_s|e_s^2 - \frac{1}{\varepsilon}b_{u_\lambda}^{-1}e_s^2 - \frac{\dot{b}_{u_\lambda}}{2b_{u_\lambda}^2}e_s^2 \\ &\quad + [d_u - \varepsilon_u(z)]e_s - \delta_w(1+e_s^2)\tilde{W}^T\hat{W} - \delta_v\mathrm{tr}\left\{\tilde{V}^T\hat{V}\right\} \end{aligned}$$

By completing the squares it is shown that $2\tilde{W}^T\hat{W} \leqslant \| \tilde{W} \|^2 - \| W^* \|^2$ and $2\mathrm{tr}\left\{\tilde{V}^T\hat{V}\right\} \leqslant \| \tilde{V} \|_F^2 - \| V^* \|_F^2$. Noticing (19) and (24), we obtain

$$\begin{aligned} \dot{V}_1 &\leqslant -k_s|e_s|^3 - \frac{1}{\varepsilon}b_{u\lambda}^{-1}e_s^2 - \frac{\dot{b}_{u_\lambda}}{2b_{u_\lambda}^2}e_s^2 - \frac{\delta_w}{2}(e_s^2+1)(\| \tilde{W} \|^2 \\ &\quad - \| W^* \|^2) - \frac{\delta_v}{2}(\| \tilde{V} \|_F^2 - \| V^* \|_F^2) + |e_s|(\alpha_1 + \alpha_2|e_s| \\ &\quad + \alpha_3 \| \tilde{W} \| + \alpha_4 \| \tilde{W} \| |e_s| + \varepsilon_1) \leqslant -k_s|e_s|^3 \\ &\quad - \frac{1}{\varepsilon}b_{u_\lambda}^{-1}e_s^2 - \frac{\delta_w}{2}e_s^2 \| \tilde{W} \|^2 - \frac{\delta_w}{2} \| \tilde{W} \|^2 - \frac{\delta_w}{2} \| \tilde{V} \|_F^2 \\ &\quad + \left(\frac{|\dot{b}_{u_\lambda}|}{2b_{u_\lambda}^2} + \alpha_2 + \frac{\delta_w}{2} \| W^* \|^2\right)e_s^2 + (\varepsilon_l + \alpha_1)|e_s| \\ &\quad + \alpha_3 \| \tilde{W} \| |e_s| + \alpha_4 \| \tilde{W} \| e_s^2 + \frac{\delta_w}{2} \| W^* \|^2 + \frac{\delta_v}{2} \| V^* \|_F^2 \end{aligned} \tag{47}$$

Since $\alpha_3 \| \tilde{W} \| |e_s| \leqslant \frac{\delta_w}{4} \| \tilde{W} \|^2 e_s^2 + \frac{\alpha_3^2}{\delta_w}$, $\alpha_4 \| \tilde{W} \| e_s^2 \leqslant \frac{\delta_w}{4} \| \tilde{W} \|^2 e_s^2 + \frac{\alpha_4^2}{\delta_w}e_s^2$, $|\dot{b}_{u_\lambda}|/2b_{u_\lambda}^2 \leqslant b_2/2b_0^2$ and $b_{u_\lambda}^{-1} \geqslant b_1^{-1}$ (Assumption 2), inequality (47) can further be written as

$$\dot{V}_1 \leqslant -k_s|e_s|^3 - \frac{1}{\varepsilon}b_1^{-1}e_s^2 - \frac{\delta_w}{2} \| \tilde{W} \|^2 - \frac{\delta_v}{2} \| \tilde{V} \|_F^2 + \beta_0 e_s^2 + \beta_1|e_s| + \beta_2 \tag{48}$$

where β_0 to β_2 are positive constants defined by

$$\beta_0 = \frac{b_2}{2b_0^2} + \alpha_2 + \frac{\delta_w}{2} \| W^* \|^2 + \frac{\alpha_4^2}{\delta_w} \tag{49}$$

$$\beta_1 = \varepsilon_1 + \alpha_1 \tag{50}$$

$$\beta_2 = \frac{\delta_w}{2} \| W^* \|^2 + \frac{\delta_v}{2} \| V^* \|_F^2 + \frac{\alpha_3^2}{\delta_w} \tag{51}$$

Noticing $2\beta_0 e_s^2 \leqslant \frac{k_s}{2}|e_s|^3 + \frac{2\beta_0^2}{k_s}|e_s|$ and $\left(\frac{2\beta_0^2}{k_s} + \beta_1\right)|e_s| \leqslant \beta_0 e_s^2 + \frac{(2\beta_0^2+\beta_1 k_s)^2}{4k_s^2\beta_0}$, we have

$$\beta_0 e_s^2 + \beta_1|e_s| \leqslant \frac{k_s}{2}|e_s|^3 + \frac{(2\beta_0^2 + \beta_1 k_s)^2}{4k_s^2\beta_0}$$

Thus inequality (48) can be written as

$$\dot{V}_1 \leqslant -\frac{k_s}{2}|e_s|^3 - \frac{1}{\varepsilon b_1}e_s^2 - \frac{\delta_w}{2} \| \tilde{W} \|^2 - \frac{\delta_v}{2} \| \tilde{V} \|_F^2 + D \tag{52}$$

where $D = \beta_2 + (2\beta_0^2 + \beta_1 k_s)^2/4k_s^2\beta_0$. Now define

$$\Theta_e := \left\{ e_s \middle| |e_s| \leqslant \min\left\{ \left(\frac{2D}{k_s}\right)^{\frac{1}{3}}, \sqrt{\varepsilon D b_1} \right\} \right\} \tag{53}$$

$$\Theta_w := \left\{ (\tilde{W}, \tilde{V}) \middle| \| \tilde{W} \| \leqslant \sqrt{\frac{2D}{\delta_w}}, \| \tilde{V} \|_F \leqslant \sqrt{\frac{2D}{\delta_v}} \right\} \tag{54}$$

$$\Theta := \left\{ (e_s, \tilde{W}, \tilde{V}) \middle| \frac{k_s}{2}|e_s|^3 + \frac{1}{\varepsilon b_1}e_s^2 + \frac{\delta_w}{2} \| \tilde{W} \|^2 + \frac{\delta_v}{2} \| \tilde{V} \|_F^2 \leqslant D \right\} \tag{55}$$

Since D, ε and k_s are positive constants, we know that Θ_e, Θ_w and Θ are compact sets. Thus $\dot{V}_1 < 0$ as long as V_1 is outside the compact set Θ. According to a standard Lyapunov theorem [24], we conclude that e_s, $\tilde{W}$ and $\tilde{V}$ are bounded. It follows from (52) and (53) that $\dot{V}_1$ is strictly negative as long as e_s is outside the compact set Θ_e. Therefore, there exists a constant T_0 such that for $t > T_0$, e_s converges to Θ_e.

Let the vector $\zeta = [\tilde{\xi}_1, \tilde{\xi}_2, \ldots, \tilde{\xi}_{\rho-1}]^T$, then a state representation for the mapping (7) is $\dot{\zeta} = A\zeta + be_s$ with

$$A = \begin{bmatrix} 0 & 1 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & 1 \\ -\lambda_1 & -\lambda_2 & \ldots & -\lambda_{\rho-1} \end{bmatrix} \in R^{(\rho-1)\times(\rho-1)}, \quad b = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix} \in R^{\rho-1} \tag{56}$$

As $s^{\rho-1} + \lambda_{\rho-1}s^{\rho-2} + \ldots + \lambda_1$ is Hurwitz, we know that A is a stable matrix. Therefore there exist constants $k_0 > 0$ and $\lambda_0 > 0$ such that $\| e^{At} \| \leqslant k_0 e^{-\lambda_0 t}$. The solution for ζ can be expressed as $\zeta(t) = \zeta(0)e^{At} + \int_0^t e^{A(t-\tau)}be_s d\tau$. Noting $\| e_s \| \leqslant E_m$ for $t > T_1$, we conclude that

$$|\zeta(t)| \leqslant k_0 e^{-\lambda_0 t} \| \zeta(0) \| + \frac{1 - e^{-\lambda_0 t}}{\lambda_0}k_0 E_m \leqslant k_0 e^{-\lambda_0 t} \| \zeta(0) \| + \frac{k_0 E_m}{\lambda_0}, \forall t > T_0 \tag{57}$$

where $E_m = \min\left\{ \left(\frac{2D}{K_s}\right)^{\frac{1}{3}}, \sqrt{\varepsilon D b_1} \right\}$ and $k_0 e^{-\lambda_0 t} \| \zeta(0) \|$ decays exponentially. This confirms that the tracking error $y - y_d = \tilde{\xi}_1$ will converge to an ε–neighborhood of the origin.

Part 2. To complete the proof, we need to prove that for a suitable tracking signal $y_d(t)$ and the appropriate design parameters, the trajectories do remain in the compact set $\bar{U}$. From (7) and $\tilde{\xi} = [\zeta^T, \tilde{\xi}_\rho]^T$, it is shown that $\tilde{\xi}_\rho = e_s - \Lambda^T\zeta$. Considering (53) and (57), there exist positive constants $\bar{k}_0$ and $\bar{k}_1$ such that

$$\| \tilde{\xi}(t) \| \leqslant \| \zeta(t) \| + |\tilde{\xi}_\rho(t)| \leqslant \bar{k}_0 e^{-\lambda_0 t} \| \zeta(0) \| + \bar{k}_1 E_m \tag{58}$$

From the analysis of Part 1, we know that $\dot{V}_1 \leqslant 0$ as long as $(e_s, \tilde{W}, \tilde{V})$ is outside the compact set Θ. Let

$$\Omega_{D1} := \left\{ (e_s, \tilde{W}, \tilde{V}) \middle| V_1 \leqslant D_1 \right\} \tag{59}$$

$$E := \inf\left\{ D_1 \middle| \Omega_{D_1} \supset \Theta \right\}$$

Clearly, all trajectories starting in Ω_E will remain in Ω_E for all time. We can also see that E is a function of c, ε, k_s and ε_1, and

i. ε_1 can be made arbitrarily small by increasing the number of the MNN nodes l, and
ii. inequality (58) implies that $\| \xi \| \leqslant \bar{k}_0 e^{-\lambda_0 t} \| \bar{\zeta}(0) \| + \bar{k}_1 E_m + \| \xi_d \|$,

Hence, there do exist a compact set Ω_0, and positive constants c^* and k_s^* such that for all $\xi(0) \in \Phi_0$, $(\hat{W}(0), \hat{V}(0)) \in \Omega_0$, $c \leqslant c^*$ and $k_s \geqslant k_s^*$, we have

$$(e_s, \tilde{W}, \tilde{V}) \in \Omega_E \Rightarrow (\xi, u) \in \bar{U}$$

This completes the proof. Q.E.D.

References

[1] K.S. Narendra, K. Parthasarathy, Identification and control of dynamic systems using neural networks, IEEE Trans. Neural Networks 1 (1) (1990) 4–27.

[2] F.C. Chen, C.C. Liu, Adaptively controlling nonlinear continuous-time systems using multilayer neural networks, IEEE Trans. Automat. Control 39 (6) (1994) 1306–1310.

[3] L. Jin, P.N. Nikiforuk, M.M. Gupta, Fast neural learning and control of discrete-time nonlinear systems, IEEE Trans. Syst, Man, and Cybernetic 25 (3) (1995) 478–488.

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

[5] F.L. Lewis, A. Yesildirek, K. Liu, Multilayer neural-net robot controller with guaranteed tracking performance, IEEE Trans. Neural Networks 7 (2) (1996) 388–398.

[6] J.T. Spooner, K.M. Passino, Stable adaptive control using fuzzy systems and neural networks, IEEE Trans. Fuzzy Systems 4 (3) (1996) 339–359.

[7] M.M. Polycarpou, Stable adaptive neural control scheme for nonlinear systems, IEEE Trans. Automat. Control 41 (3) (1996) 447–451.

[8] R.M. Sanner, J.J.E. Slotine, Gaussian networks for direct adaptive control, IEEE Trans. Neural Networks 3 (6) (1992) 837–863.

[9] S.S. Ge, Robust adaptive NN feedback linearization control of nonlinear systems, International Journal of Systems Science 27 (12) (1996) 1327–1338.

[10] D. Psaltis, A. Sideris, A. Yamamura, A multilayered neural network controller, IEEE Control Sys. Magazine 8 (1988) 17–21.

[11] C.J. Goh, Model reference control of non-linear systems via implicit function emulation, Int. J. Control (60) (1994) 91–115.

[12] A.U. Levin, K.S. Narendra, Control of nonlinear dynamical systems using neural networks Part II: observability, identification, and control, IEEE Trans. Neural Networks 7 (1) (1996) 30–42.

[13] C.J. Goh, T.H. Lee, Direct adaptive control of nonlinear systems via implicit function emulation, Control-Theory and Advanced Technology 10 (3) (1994) 539–552.

[14] T. Zhang, S.S. Ge, C.C. Hang, Neural-based direct adaptive control for a class of general nonlinear systems, International Journal of Systems Science 28 (10) (1997) 1011–1020.

[15] T. Zhang, S.S. Ge, C.C. Hang, Direct adaptive control of non-affine nonlinear systems using multilayer neural networks, Proc. American Control Conference, Philadelphia, PA, June 1998, pp. 515–519.

[16] G. Lightbody, G.W. Irwin, Direct neural model reference adaptive control, IEE Proc-Control Theory Appl. 142 (1) (1995) 31–43.

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

[18] A. Isidori, Nonlinear Control System, second ed., Springer–Verlag, Berlin, 1989.

[19] J. Tsinias, N. Kalouptsidis, Invertability of nonlinear analytic single-input systems, IEEE Trans. Automat. Control 28 (9) (1983) 931–933.

[20] S. Behtash, Robust output tracking for nonlinear system, Int. J. Control 51 (6) (1990) 1381–1407.

[21] A. Teel, R. Kadiyala, P.V. Kokotovic, S.S. Sastry, Indirect techniques for adaptive input output linearization of nonlinear systems, Int. J. Control 53 (1991) 193–222.

[22] S. Lang, Real Analysis, Addison–Wesley, Reading, MA, 1983.

[23] T.M. Apostol, Mathematical Analysis, Addison–Wesley, Reading, MA, 1957.

[24] K.S. Narendra, A.M. Annaswamy, Stable Adaptive System, Prentice Hall, Englewood Cliffs, NJ, 1989.

[25] P.A. Ioannou, J. Sun, Robust Adaptive Control, Prentice Hall, Englewood Cliffs, NJ, 1995.