A Discrete-Time Multivariable Neuro-Adaptive Control for Nonlinear Unknown Dynamic Systems

Chih-Lyang Hwang, *Associate Member, IEEE,* and Ching-Hung Lin

***Abstract*—First, we assume that the controlled systems contain a nonlinear matrix gain before a linear discrete-time multivariable dynamic system. Then, a forward control based on a nominal system is employed to cancel the system nonlinear matrix gain and track the desired trajectory. A novel recurrent-neural-network (RNN) with a compensation of upper bound of its residue is applied to model the remained uncertainties in a compact subset Ω. The linearly parameterized connection weight for the function approximation error of the proposed network is also derived. An e-modification updating law with projection for weight matrix is employed to guarantee its boundedness and the stability of network without the requirement of persistent excitation. Then a discrete-time multivariable neuro-adaptive variable structure control is designed to improve the system performances. The semi-global (i.e., for a compact subset Ω) stability of the overall system is then verified by the Lyapunov stability theory. Finally, simulations are given to demonstrate the usefulness of the proposed controller.**

***Index Terms*—Discrete-time system, Lyapunov stability theory, multivariable variable structure control, recurrent-neural-network.**

I. Introduction

It is known that a conventionally designed linear controller does not achieve an adequate performance over a variety of operating regimes, especially if the system is highly nonlinear [1]. It is also known that a robust controller design based on a nominal system is not enough to stabilize the system with huge uncertainty [2]. There is a wide class of nonlinear multivariable systems that can be modeled by an interconnected nonlinear matrix gain and a linear dynamic system, e.g., Wiener model, Hammerstein model, hysteresis model [3]–[5]. Due to the progress of microcomputer, a digital control is more implementable [7], [8]. In this paper, a nonlinear matrix gain that is not necessarily memoryless, combines with a linear discrete-time multivariable dynamic system to represent a class of nonlinear multivariable systems.

In the beginning, a forward control based on a nominal system is applied to cancel the system nonlinearity and track the desired trajectory. A forward control is simple and very acceptable for industry. In general, a suitable design of forward control can make a poor dynamic system become a well-controlled system [4]–[6], [8]–[10]. The more accurate a nominal model for the controlled system is used, the more excellent tracking performance is achieved. If the tracking performance using a forward control or a robust control can't satisfy the specific requirement, a neural-network for a compact subset Ω is used to approximate the remained uncertainties which are dynamic.

Most people used a multilayer-neural-network, combined with the tapped delays for the input, and a backpropagation training algorithm to deal with the dynamic problem [11]–[13]. On the other hand, recurrent-neural-networks (RNNs) have important capabilities not found in multilayer, such as dynamic mapping. Therefore, the RNN is better suited for dynamic systems than the multilayer-neural-network [14]–[18]. An RNN can cope with time-variant input or output through its own natural temporal operation. Hence, the less number of neurons for the RNN can model the uncertainties to achieve the required accuracy (e.g., [14]–[18] or Remark 6). To compensate the residue of RNN, a simple network is established to estimate its upper bound for the controller design. The proposed network is called "Recurrent-Neural-Network-with-Residue-Upper-Bound-Compensation (RNNRUBC)" (see Fig. 1). Besides, the stability of recurrent-neural-network-based control system is difficult because the corresponding dynamics are nonlinear in adjustable connection weights. However, the linearly parameterized weight for the function approximation error of the proposed network is derived. As the authors know, this is the first time to construct a neural-network with these features.

Furthermore, an e-modification updating law with projection [2], [19], [20] for the weight matrix is designed to ensure its boundedness and the stability of network without the requirement of persistent excitation. The purpose of using a projection algorithm for the feedback weight matrix between the neurons is to ensure a stable neural-network. Because the variable structure control possesses the following advantages: fast response, less sensitivity to uncertainties and easy implementation [21]–[24], a discrete-time multivariable neuro-adaptive variable structure control (DMNAVSC, see Fig. 2) is designed to improve the system performances. Under mild conditions, the convergent region for tracking error of the proposed control can be smaller than that of robust variable structure control [compare Fig. 4(d) and Fig. 5(d)]. As compared with the robust variable structure control, the proposed controller can deal with extra uncertainties to attain an excellent tracking result without the occurrence of chattering control input [see Figs. 4(c) and 5(d)]. Furthermore, without the prior estimation of weight matrix (e.g., off-line training) the initial value of weight matrices can be set to zero (i.e., no compensation for excess uncertainties with respect

Manuscript received October 29, 1999; revised March 22, 2000, April 16, 2000, and July 11, 2000. This work was supported in part by the National Science Council of Taiwan of R.O.C. under Grant NSC-89-2212-E-036-003. This paper was recommended by Associate Editor C. P. Neuman.

The authors are with Department of Mechanical Engineering of Tatung University, Taipei, 10451 Taiwan, R.O.C.

Publisher Item Identifier S 1083-4419(00)08799-9.

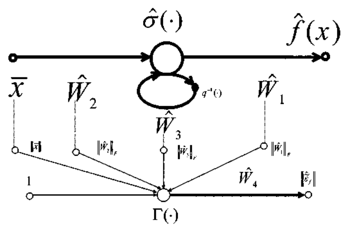


Fig. 1. Structure of RNNRUBC.

to robust variable structure control). This characteristic makes the suggested control more practical because many neural network controls have a problem to assign an initial value for the connection weight (e.g., [11]–[15], [17], [18], [20]).

Finally, the semi-global stability (i.e., for a compact subset Ω) of the overall system is verified by the Lyapunov stability theory. Simulations are also given to confirm the usefulness of the proposed control. The proposed scheme indeed improves the performances of some robust control schemes.

II. Problem Formulation

Consider a class of nonlinear discrete-time multivariable systems:

$$A_r(q^{-1})Y(t) = q^{-d}B_r(q^{-1})V(t), V(t) = F_r(U) \quad (1)$$

where the signals $Y(t)$, $V(t)$ and $U(t) \in \Re^m$, the mapping $F_r(\cdot) = F(\cdot) + \Delta F(\cdot)$: $\Re^m \rightarrow \Re^m$ is continuous, the mapping $F(\cdot)$ is known and invertible, q^{-1} is a backward-time shift operator (i.e., $q^{-1}Y(t) = Y(t-1)$), q^{-d} denotes d-step delay operator and the polynomial matrices $A_r(q^{-1}) = A(q^{-1}) + \Delta A(q^{-1})$, $B_r(q^{-1}) = B(q^{-1}) + \Delta B(q^{-1})$ are described as follows [7]:

$$\begin{aligned} A(q^{-1}) &= I + A_1q^{-1} + \cdots + A_\alpha q^{-\alpha}, \\ \Delta A(q^{-1}) &= \Delta A_1 q^{-1} + \cdots + \Delta A_{\alpha'} q^{-\alpha'}, \qquad \alpha' \geq \alpha \quad (2a) \\ B(q^{-1}) &= B_0 + B_1 q^{-1} + \cdots + B_\beta q^{-\beta}, \\ \Delta B(q^{-1}) &= \Delta B_0 + \Delta B_1 q^{-1} + \cdots + \Delta B_{\beta'} q^{-\beta'}, \\ &\qquad \beta' \geq \beta \quad (2b) \end{aligned}$$

where matrix parameters A_i, B_j for $1 \leq i \leq \alpha$, $0 \leq j \leq \beta$ are known, ΔA_i, $\Delta B_j (1 \leq i \leq \alpha', 0 \leq j \leq \beta')$ are known but bounded. Let $\gamma = \max[\alpha, \beta + 1]$. It is assumed that:

A1: m and $d \geq 1$ are known.

A2: $A(q^{-1})$, $B(q^{-1})$ are left coprime, $A(q^{-1})$, $A_r(q^{-1})$, $B(q^{-1})$ are stable matrices, and $\det\{B_0\} \neq 0$.

A3: The unknown mapping $\Delta F(\cdot)$ satisfies the following equality:

$$\Delta F F^{-1}(U) = \Delta F_0 U_b + \Delta F_1 U(t) + \Delta F_2(U) \quad (2c)$$

where $U_b = [1 \cdots 1]^T$, ΔF_0, $\Delta F_1 \in \Re^{m \times m}$ are bounded, $\Delta F_2(\cdot)$: $\Re^m \rightarrow \Re^m$, $\|\Delta F_2(U)\| < \delta$, where δ is known, as $U(t) \in \overline{\Omega} = \{U(t) \in \Re^m | \, \|U(t-d+1) - U_0\| < \overline{\gamma}\}$ which is a compact set. The symbol $\|\cdot\|$ denotes usual Euclidean-norm.

Remark 1: The signal $V(t)$ is not available. It is not necessarily to assume that the real linear system must be minimum phase [5]. Assumption A2 ensures that the linear system is controllable and observable [7] and that the existence of a forward control and the boundedness of uncertainties [see (15) and (18)]. Assumption A3 shows that in a compact set the nonlinear vector of uncertainty relative to the nominal value is linearized and has a norm-bounded value for the order equal to or greater than two.

Remark 2: For example, the known matrix mapping $F(\cdot)$ has the following entries:

$$f_{ii}(u_i) = \frac{1 + \gamma_{i1}\gamma_{i3} - \exp[-2\gamma_{i1}(u_i(t) - \gamma_{i2})]}{\gamma_{i1}\{1 + \exp[-2\gamma_{i1}(u_i(t) - \gamma_{i2})]\}}, \\ 1 \leq i \leq m, f_{ij}(\cdot) = 0 \quad \text{for } i \neq j \quad (3)$$

where $\gamma_{ij}(1 \leq i \leq m, 1 \leq j \leq 3)$ are known and $u_i(t)$ denotes an input signal for the mapping $f_{ii}(\cdot)$. Another example is a linear mapping, e.g., the argument in (2c) greater than and equal to two are set to zero, i.e., $\Delta F_2(U) = 0$. These two examples satisfy Assumption A3 for $\overline{\Omega} = \Re^m$

The objective of this paper is to construct a discrete-time multivariable neuro-adaptive control $U(t)$ (DMNAC) for a class of nonlinear multivariable discrete-time systems in the presence of uncertainties caused by ΔA_i, $\Delta B_j (1 \leq i \leq \alpha', 0 \leq j \leq \beta')$, $\Delta F(\cdot)$ which are not necessarily small. A forward control, i.e., $U(t) = U_f(t)$ as $U_\nu(t) = 0$, is first employed to cancel the mapping $F(\cdot)$ and track the desired trajectory $R(t)$ that is known and bounded. If the tracking performance using a forward control or a robust control does not satisfy the specific requirement, a RNNRUBC (see Fig. 1) is applied to model the remained uncertainties for their input belonging to a compact subset $\Omega \subseteq \overline{\Omega}$. Then a discrete-time multivariable neuro-adaptive variable structure control $U_\nu(t)$ (DMNAVSC) is designed to enhance the tracking performance (allude to Fig. 2). Finally, the semi-global stability (i.e., $U(t) \in \Omega$) of the overall system is verified by the Lyapunov stability theory (see Fig. 2).

III. Recurrent Neural Network with Residue Upper Bound Compensation

In this section, the RNNRUBC will be introduced such that the nonlinearly parameterized connection weight for the function approximation error is changed into a linearly parameterized form. The structure of RNNRUBC is presented in Fig. 1 that performs as an approximator described in the following matrix form:

$$\begin{aligned} &\hat{f}(\overline{x}, \hat{W}_1, \hat{W}_2, \hat{W}_3) \\ &\quad = \hat{W}_1^T(t)\sigma(\hat{W}_2^T(t)\overline{x}(t) + \hat{W}_3^T(t)q^{-1}(\hat{\sigma})) \quad (4) \end{aligned}$$

where

$$\begin{aligned} &\hat{W}_1^T(t) \in \Re^{m \times p}, \quad \hat{W}_2^T(t) \in \Re^{p \times (n+1)}, \\ &\hat{W}_3^T(t) \in \Re^{p \times p}, \quad \overline{x}(t) = [x^T(t) \quad b]^T \in \Re^{(n+1) \times 1} \end{aligned}$$

$q^{-1}(\cdot)$ and $\sigma(\cdot) \in \Re^{p \times 1}$ denote output-hidden weight matrix, hidden-input weight matrix, recurrent weight matrix, input vector with a known constant b, backward-time shift operator, i.e.,

$$q^{-1}(\hat{\sigma}) = q^{-1}\{\hat{\sigma}[\overline{x}(t)]\} = \hat{\sigma}[\overline{x}(t-1)]$$

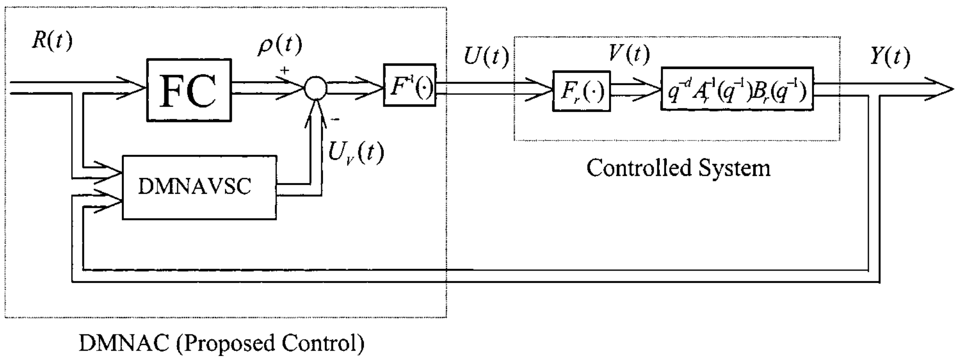


Fig. 2. Control block diagram of the overall system.

where

$$\hat{\sigma}(\overline{x}) = \sigma(\hat{W}_2^T(t)\overline{x}(t) + \hat{W}_3^T(t)q^{-1}(\hat{\sigma}))$$

and sigmoid function with entry

$$\hat{\sigma}_i(z_i) = 1/[1+\exp(-z_i)], \qquad 1 \le i \le p$$

where

$$z_i(t) = (\hat{W}_2^T(t)\overline{x}(t) + \hat{W}_3^T(t)q^{-1}(\hat{\sigma}))_i$$

respectively. If $\hat{W}_3(t) = 0$, then the network becomes a multi-layer-neural-network.

The universal approximation theory is stated as follows:

Theorem 1 (e.g., [11]–[13], [17]): Suppose $x(t) \in \Omega$ (a compact subset of $\Re^n$), $f(x)\colon \Omega \rightarrow \Re^m$ is a continuous function vector. For an arbitrary constant $\varepsilon > 0$, there exists an integer p (the number of hidden neurons) and real constant matrices $\overline{W}_1$, $\overline{W}_2$ and $\overline{W}_3$, where $0 < \overline{\varepsilon} \le \|\overline{W}_3\|_F \le W_{3m} < \sqrt{p}$ such that

$$f(x) = \overline{W}_1^T \sigma(\overline{W}_2^T \overline{x}(t) + \overline{W}_3^T q^{-1}(\sigma)) + \varepsilon_f(\overline{x}) \tag{5}$$

where $\|\cdot\|_F$ denotes the Frobenius norm (i.e., $\|W\|_F^2 = \mathrm{tr}[W^TW] = \mathrm{tr}[WW^T]$, $W \in \Re^{m\times m}$), $\varepsilon_f(\overline{x})$ denotes the approximation error vector satisfying $\|\varepsilon_f(\overline{x})\| \le \varepsilon, \forall x(t) \in \Omega$, and $\overline{x}(t)$ is the same as (4).

Remark 3: The constant matrices in Theorem 1 are not unique and satisfy the following inequalities:

$$\|\overline{W}_1\|_F \le W_{1m}, \quad \|\overline{W}_2\|_F \le W_{2m},$$
$$0 < \overline{\varepsilon} \le \|\overline{W}_3\|_F \le W_{3m} < \sqrt{p}$$

where $\overline{\varepsilon}$, W_{1m}, W_{2m} and W_{3m} are known. The RNNRUBC is a dynamic mapping. However, the multilayer-neural-network is a static mapping. Furthermore, the condition $\|\overline{W}_3\|_F < \sqrt{p}$ is required for a stable neural-network. Therefore, an updating law for $\hat{W}_3(t)$ can be modified as a project algorithm [19], [20] to ensure an effective learning.

The following lemma expresses the function approximation error $\tilde{f} = f - \hat{f}$ in a linearly parameterized form.

Lemma 1: Consider the function approximation error:

$$\begin{aligned}&\tilde{f}(\overline{x}, \hat{W}_1, \hat{W}_2, \hat{W}_3)\\ &= \tilde{W}_1^T(t)\{\hat{\sigma}(\overline{x}) - \hat{\sigma}'(\overline{x})\hat{W}_2^T(t)\overline{x}(t) - 2\hat{\sigma}'(\overline{x})\hat{W}_3^T(t)q^{-1}(\hat{\sigma})\}\\ &\quad + \hat{W}_1^T(t)\hat{\sigma}'(\overline{x})\{\tilde{W}_2^T(t)\overline{x}(t) + 2\tilde{W}_3^T(t)q^{-1}(\hat{\sigma})\} + \tilde{\varepsilon}_f(t, \overline{x})\end{aligned} \tag{6}$$

where

$$\begin{aligned}\tilde{\varepsilon}_f(t, \overline{x}) = &\tilde{W}_1^T(t)\hat{\sigma}'(\overline{x})\{\overline{W}_2^T\overline{x}(t) + \tilde{W}_3^T(t)q^{-1}(\tilde{\sigma})\\ &+ \overline{W}_3^T q^{-1}(\hat{\sigma}) + \hat{W}_3^T(t)q^{-1}(\sigma)\}\\ &+ \hat{W}_1^T(t)\hat{\sigma}'(\overline{x})\tilde{W}_3^T(t)q^{-1}(\sigma)\\ &+ \overline{W}_1^T O[\tilde{W}_2^T(t)\overline{x}(t) + \tilde{W}_3^T(t)q^{-1}(\sigma)\\ &+ \hat{W}_3^T(t)q^{-1}(\tilde{\sigma})]^2 + \varepsilon_f(\overline{x})\end{aligned} \tag{7}$$

with $O[\cdot]^2$ denoting a sum of order terms equal to and greater than two of the argument

$$\begin{aligned}&\hat{\sigma}'(\overline{x}) = \mathrm{diag}\{\hat{\sigma}_1'(\overline{x}), \cdots, \hat{\sigma}_p'(\overline{x})\} \in \Re^{p\times p},\\ &\hat{\sigma}_i'(\overline{x}) = d\hat{\sigma}_i(z)/dz|_{z=\overline{x}} = \hat{\sigma}_i(\overline{x})\,(1 - \hat{\sigma}_i(\overline{x})),\\ &\qquad 1 \le i \le p.\end{aligned}$$

Proof: Subtracting (5) from (4) gives

$$\begin{aligned}\tilde{f} &= (\tilde{W}_1^T + \hat{W}_1^T)(\tilde{\sigma} + \hat{\sigma}) + \varepsilon_f - \hat{W}_1^T\hat{\sigma}\\ &= \tilde{W}_1^T\tilde{\sigma} + \tilde{W}_1^T\hat{\sigma} + \hat{W}_1^T\tilde{\sigma} + \varepsilon_f\end{aligned} \tag{8}$$

where $\tilde{W}_1 = \overline{W}_1 - \hat{W}_1$ and $\tilde{\sigma} = \sigma - \hat{\sigma}$. Taylor expansion $\sigma(\overline{W}_2^T\overline{x} + \overline{W}_3^T q^{-1}(\sigma))$ about $\hat{W}_2^T\overline{x} + \hat{W}_3^T q^{-1}(\hat{\sigma})$ yields

$$\begin{aligned}&\sigma(\overline{W}_2^T\overline{x} + \overline{W}_3^T q^{-1}(\sigma))\\ &\quad = \sigma(\hat{W}_2^T\overline{x} + \hat{W}_3^T q^{-1}(\hat{\sigma})) + \hat{\sigma}'(\hat{W}_2^T\overline{x} + \hat{W}_3^T q^{-1}(\hat{\sigma}))\\ &\qquad \cdot [\tilde{W}_2^T\overline{x} + \tilde{W}_3^T q^{-1}(\sigma) + \hat{W}_3^T q^{-1}(\tilde{\sigma})]\\ &\qquad + O[\tilde{W}_2^T\overline{x} + \tilde{W}_3^T q^{-1}(\sigma) + \hat{W}_3^T q^{-1}(\tilde{\sigma})]^2.\end{aligned} \tag{9}$$

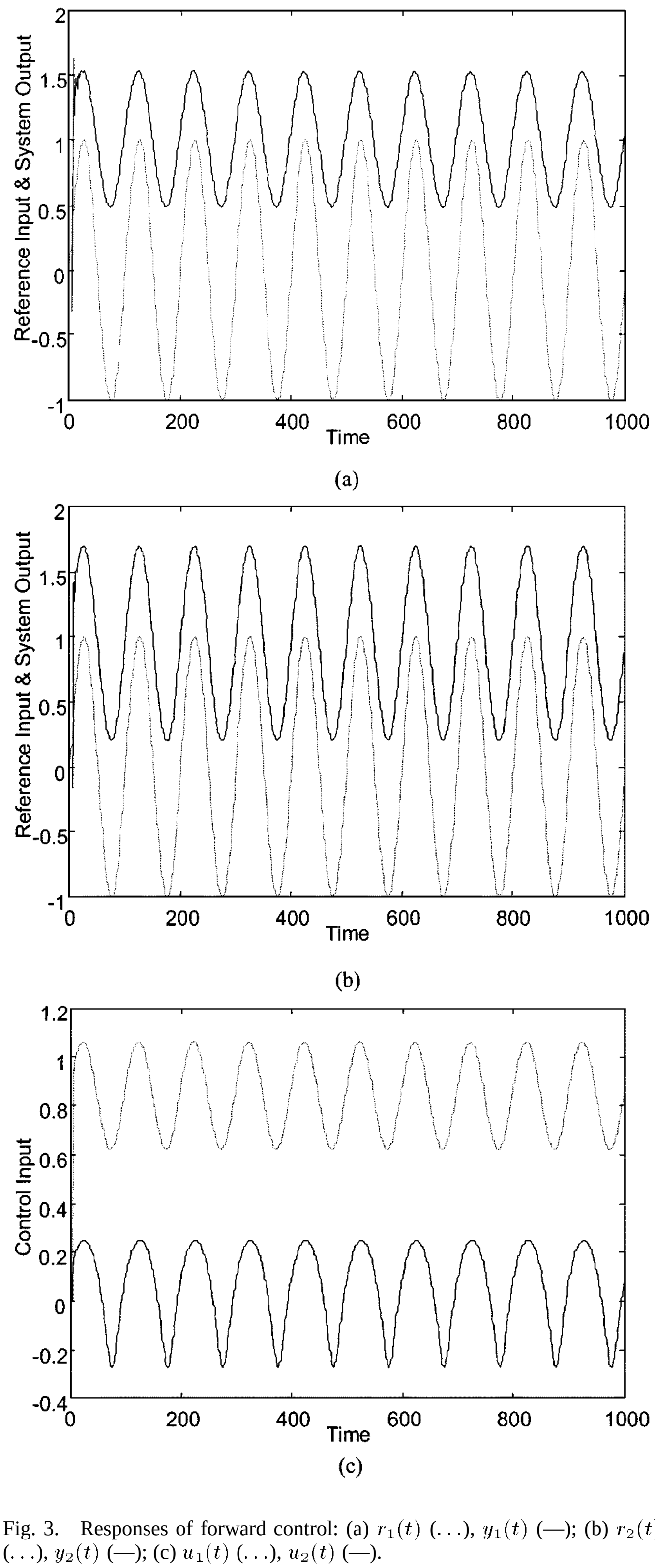


Fig. 3. Responses of forward control: (a) $r_1(t)$ (. . .), $y_1(t)$ (—); (b) $r_2(t)$ (. . .), $y_2(t)$ (—); (c) $u_1(t)$ (. . .), $u_2(t)$ (—).

Substituting $\tilde{\sigma}$ from (9) into (8) gives

$$\begin{aligned}\tilde{f} &= \tilde{W}_1^T\hat{\sigma} + \varepsilon_f + [\tilde{W}_1^T + \hat{W}_1^T] \\ &\quad \cdot \{\hat{\sigma}'[\tilde{W}_2^T\overline{x} + \tilde{W}_3^T q^{-1}(\sigma) + \hat{W}_3^T q^{-1}(\tilde{\sigma})] \\ &\quad + O[\tilde{W}_2^T\overline{x} + \tilde{W}_3^T q^{-1}(\sigma) + \hat{W}_3^T q^{-1}(\tilde{\sigma})]^2\}. \quad (10)\end{aligned}$$

Substituting the relation $\tilde{W}_i = \overline{W}_i - \hat{W}_i$, $1 \le i \le 3$, and $\tilde{\sigma} = \sigma - \hat{\sigma}$ into (10) yields

$$\begin{aligned}\tilde{f} &= \tilde{W}_1^T\hat{\sigma}'(\overline{W}_2^T - \hat{W}_2^T)\overline{x} + \tilde{W}_1^T\hat{\sigma}'(\overline{W}_3^T - \hat{W}_3^T) \\ &\quad \cdot q^{-1}(\tilde{\sigma} + \hat{\sigma}) + \tilde{W}_1^T\hat{\sigma}'\hat{W}_3^T q^{-1}(\sigma - \hat{\sigma}) \\ &\quad + \hat{W}_1^T\hat{\sigma}'\tilde{W}_2^T\overline{x} + \hat{W}_1^T\hat{\sigma}'\tilde{W}_3^T q^{-1}(\tilde{\sigma} + \hat{\sigma}) \\ &\quad + \hat{W}_1^T\hat{\sigma}'(\overline{W}_3^T - \tilde{W}_3^T)q^{-1}(\sigma - \hat{\sigma}) + \tilde{W}_1^T\hat{\sigma} \\ &\quad + (\tilde{W}_1^T + \hat{W}_1^T)O[\tilde{W}_2^T\overline{x} + \tilde{W}_3^T q^{-1}(\sigma) + \hat{W}_3^T q^{-1}(\tilde{\sigma})]^2 \\ &\quad + \varepsilon_f. \quad (11)\end{aligned}$$

Simplifying (11) gives the results (6) and (7). Q.E.D.

Remark 4: The approximation $\hat{f}(\overline{x}, \hat{W}_1, \hat{W}_2, \hat{W}_3)$ is used in the control law to cancel the unknown nonlinear function $f(x)$. The first and second term of $\tilde{f}(\overline{x}, \hat{W}_1, \hat{W}_2, \hat{W}_3)$ are canceled by the updating law of weights $\hat{W}_1(t)$, $\hat{W}_2(t)$ and $\hat{W}_3(t)$. Although the residue $\tilde{\varepsilon}_f(t, \overline{x})$ is unknown, an upper bound of $||\tilde{\varepsilon}_f(t, \overline{x})||$ can be achieved by the following lemma.

Lemma 2: The residue term $||\tilde{\varepsilon}_f(t, \overline{x})||$ can be bounded by the following inequality:

$$||\tilde{\varepsilon}_f(t, \overline{x})|| \le \overline{W}_4^T \Gamma(t) \quad (12)$$

where the unknown parameter vector $\overline{W}_4 > 0 \in \Re^{7\times 1}$ is bounded and the known function

$$\begin{aligned}\Gamma(t) &= \Gamma(\overline{x}, ||\hat{W}_1||_F, ||\hat{W}_2||_F, ||\hat{W}_3||_F) \\ &= [1 \quad ||\overline{x}(t)|| \quad ||\hat{W}_1(t)||_F \quad ||\hat{W}_3(t)||_F \\ &\quad ||\overline{x}(t)||\,||\hat{W}_1(t)||_F \quad ||\overline{x}(t)||\,||\hat{W}_2(t)||_F \\ &\quad ||\hat{W}_1(t)||_F||\hat{W}_3(t)||_F]^T.\end{aligned}$$

Proof: From (9) and triangular inequality, the high-order terms are bounded by

$$\begin{aligned}&||O[\tilde{W}_2^T\overline{x} + \tilde{W}_3^T q^{-1}(\sigma) + \hat{W}_3^T q^{-1}(\tilde{\sigma})]^2|| \\ &\quad = ||\tilde{\sigma} - \hat{\sigma}'[\tilde{W}_2^T\overline{x} + \tilde{W}_3^T q^{-1}(\sigma) + \hat{W}_3^T q^{-1}(\tilde{\sigma})]|| \\ &\quad \le ||\tilde{\sigma}|| + ||\hat{\sigma}'||\,[\,||\tilde{W}_2^T||_F||\overline{x}|| + ||\tilde{W}_3^T||_F||\sigma|| \\ &\qquad + ||\hat{W}_3^T||_F||\tilde{\sigma}||\,] \\ &\quad \le c_1 + c_2[(W_{2m} + ||\hat{W}_2||_F)||\overline{x}|| \\ &\qquad + c_3(W_{3m} + ||\hat{W}_3||_F) + c_1||\hat{W}_3||_F] \\ &\quad \le (c_1 + c_2c_3W_{3m}) + c_2W_{2m}||\overline{x}|| \\ &\qquad + c_2(c_1 + c_3)||\hat{W}_3||_F + c_2||\overline{x}||\,||\hat{W}_2||_F \quad (13)\end{aligned}$$

where

$$||\tilde{\sigma}|| \le c_1, \quad ||\hat{\sigma}'|| \le c_2, \quad ||\sigma|| \le c_3,$$

$$||\tilde{W}_2||_F \le W_{2m} + ||\hat{W}_2||_F \quad \text{and} \quad ||\tilde{W}_3||_F \le W_{3m} + ||\hat{W}_3||_F.$$

From (7) and (13), $||\tilde{\varepsilon}_f||$ is then bounded by

$$\begin{aligned}||\tilde{\varepsilon}_f|| &\le (W_{1m} + ||\hat{W}_1||_F)c_2\{W_{2m}||\overline{x}|| + (W_{3m} + ||\hat{W}_3||_F)c_1 \\ &\quad + W_{3m}c_4 + ||\hat{W}_3||_F c_3\} \\ &\quad + ||\hat{W}_1||_F c_2(W_{3m} + ||\hat{W}_3||_F)c_3 \\ &\quad + W_{1m}\{(c_1 + c_2c_3W_{3m}) + c_2W_{2m}||\overline{x}|| \\ &\quad + c_2(c_1 + c_3)\,||\hat{W}_3||_F + c_2||\overline{x}||\,||\hat{W}_2||_F\} + \varepsilon \\ &\le \{\,[c_1 + c_2(c_1 + c_3 + c_4)W_{3m}]W_{1m} + \varepsilon\} \\ &\quad + \{2c_2W_{1m}W_{2m}\}\,||\overline{x}|| + \{c_2(c_1 + c_3)W_{3m}\}\,||\hat{W}_1||_F \\ &\quad + \{2c_2(c_1 + c_3)W_{1m}\}\,||\hat{W}_3||_F \\ &\quad + \{c_2W_{2m}\}\,||\overline{x}||\,||\hat{W}_1||_F + \{c_1W_{1m}\}\,||\overline{x}||\,||\hat{W}_2||_F \\ &\quad + \{c_2(c_1 + 2c_3)\}\,||\hat{W}_1||_F||\hat{W}_3||_F \\ &= \overline{W}_4^T\Gamma \quad (14)\end{aligned}$$

where $||\tilde{W}_1||_F \le W_{1m} + ||\hat{W}_1||_F$ and $||\hat{\sigma}|| \le c_4$. Q.E.D.

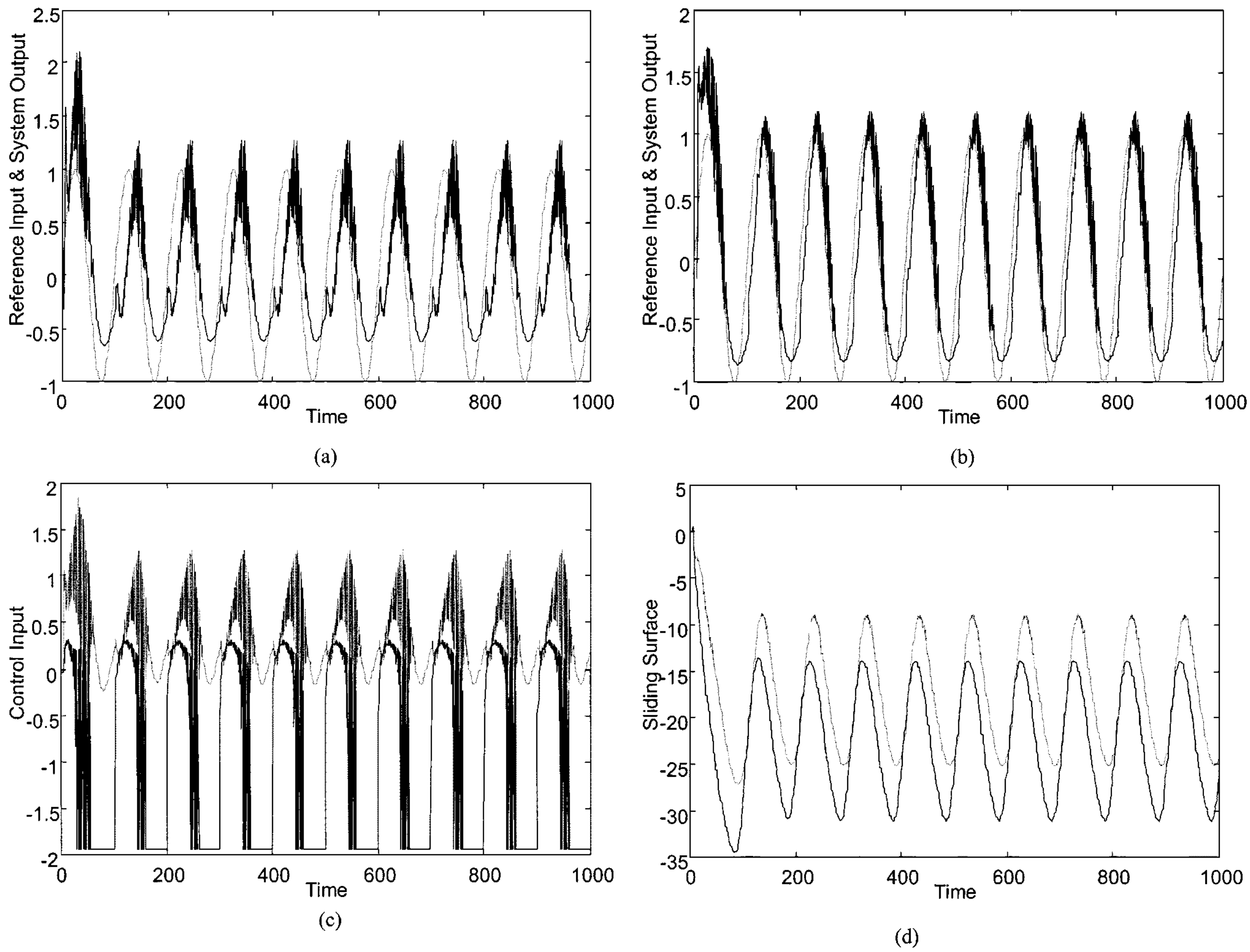


Fig. 4. Responses of robust variable structure control:. (a) $r_1(t)$ (. . .), $y_1(t)$ (—); (b) $r_2(t)$ (. . .), $y_2(t)$ (—); (c) $u_1(t)$ (. . .), $u_2(t)$ (—); (d) $s_1(t)$ (. . .), $s_2(t)$ (—).

Remark 5: Although $\tilde{\varepsilon}_f(t,\overline{x})$ is unknown, an updating law for $\hat{W}_4(t)$ is designed to compensate the upper bound of this residue, i.e.,$\|\hat{\tilde{\varepsilon}}_f(t,\overline{x})\| = \hat{W}_4^T(t)\Gamma(t)$. Similar as Remark 3, $\|\overline{W}_4\| = \|\overline{W}_4\|_F \le W_{4m}$. Furthermore, without using an updating law of $\hat{W}_4(t)$ but using the estimation of $\overline{W}_4$ from (14) to compensate the residue $\tilde{\varepsilon}_f(t,\overline{x})$ is too conservative to attain an excellent response.

Remark 6: The total number of connection weight of dynamic mapping for vector function approximation (i.e., $f(x) \in \Re^m$, $m \ge 2$) using recurrent, multilayer and radial basis function (e.g., Gaussian function) neural-network are discussed as follows [2], [11]–[13]:

$$
\begin{aligned}
N_{RNN} &= m \times p + p \times (n_{RNN} + 1) + p \times p, \\
n_{RNN} &= k, \qquad p \ge 2 \\
N_{MLN} &= m \times p + p \times (n_{MLN} + 1), \\
n_{MLN} &= k \times n_{ord}, \qquad n_{ord} \ge 2 \\
N_{RBFN} &= n_{seg}^{n_{RBFN}} + 1 \\
n_{RBFN} &= k \times n_{ord}, \qquad n_{seg} \ge 3
\end{aligned}
$$

where n_{RNN}, n_{MLN} and n_{RBFN} denote the number of input layer for recurrent, multilayer and radial basis function neural-network, respectively; p, n_{ord} and n_{seq} represent the number of hidden neuron, the order of nonlinear function, the number of segment for every input signal (in general, it is odd), respectively. For example, if $m = 2$, $k = 2m$, $p = 2$, $n_{ord} = 2$ and $n_{seg} = 3$ (in general, it is a minimum requirement) are selected, then $N_{RNN} = 18$, $N_{MLN} = 22$ and $N_{RBFN} = 7562$, respectively. If the number or order of nonlinear function or input layer or hidden neuron is large, the total number of recurrent neural-network is still smallest. From the above discussions, the RNN is more suitable for the approximation of complex (or coupled) and dynamic mapping.

IV. Controller Design

In this section, four subsections are arranged to discuss the controller design. In the beginning, a forward controller according to the nominal system is constructed such that the system nonlinearity is canceled and the tracking error is reduced. Then an error model resulting from the forward control is achieved in the second subsection. The updating law for weight matrix is given in the third subsection. Finally, a DMNAVSC is designed to improve the tracking performance.

A. Forward Control

A forward control to achieve an acceptable tracking performance for the system (1) with nonlinearity (3) is described as follows (e.g., [4]–[6], [8]–[10]):

$$U_f(t) = F^{-1}(\rho) \tag{15a}$$

with

$$f_{ii}^{-1}(\rho_i) = \gamma_{i2} - \frac{1}{2\gamma_{i1}} \log\left[\frac{1+\gamma_{i1}\gamma_{i3} - \gamma_{i1}\rho_i(t)}{1+\gamma_{i1}\rho_i(t)}\right], \quad 1 \le i \le m, f_{ij}^{-1}(\rho_j) = 0 \quad \text{for } i \ne j \tag{15b}$$

$$\rho(t) = \{q^d Adj[B(q^{-1})]A(q^{-1})R(t)\}/\det[B(q^{-1})] \tag{15c}$$

where $f_{ii}^{-1}(\cdot)$, $1 \le i \le m$, represents the diagonal entries of $F^{-1}(\cdot)$, the input signal $\rho_i(t)$ must satisfy the inequality $[1 + \gamma_{i1}\gamma_{i3} - \gamma_{i1}\rho_i(t)]/[1+\gamma_{i1}\rho_i(t)] > 0\ \forall t$. Alternatively, if $\gamma_{i1} > 0$, then $-1/\gamma_{i1} < \rho_i(t) < (1+\gamma_{i1}\gamma_{i3})/\gamma_{i1}\ \forall t$; if $\gamma_{i1} < 0$, then $(1+\gamma_{i1}\gamma_{i3})/\gamma_{i1} < \rho_i(t) < -1/\gamma_{i1}\ \forall t$.

Lemma 3: Applying the controller (15) to the system (1) gives the following tracking error:

$$\begin{aligned} E(t) = R(t) - Y(t) = -A^{-1}(q^{-1})q^{-d}\{&B(q^{-1})\Delta F(U) \\ &+ [\Delta B(q^{-1}) - (I + \Delta A(q^{-1})A^{-1}(q^{-1}))^{-1} \\ &\cdot \Delta A(q^{-1})A^{-1}(q^{-1})B_r(q^{-1})]F_r(U)\} \end{aligned} \tag{16}$$

where I denotes the unit matrix of dimension m and $U(t) = U_f(t)$.

Proof: For simplicity, it is omitted.

B. Error Model Resulting from Forward Control

Because an accurate modeling of system with huge uncertainties is difficult to attain, the tracking error based on a forward control is generally not superior. To compensate these phenomena, a DMNAVSC is affiliated with the previous forward control (see Fig. 2). The error-model of closed-loop system in Fig. 2 is depicted as follows:

$$A(q^{-1})E(t) = q^{-d}B(q^{-1})U_\nu(t) + A(q^{-1})\Phi(t) \tag{17}$$

where

$$\begin{aligned} \Phi(t) = -A^{-1}(q^{-1})q^{-d}\{&B(q^{-1})\Delta F(U) + [\Delta B(q^{-1}) \\ &- (I+\Delta A(q^{-1})A^{-1}(q^{-1}))^{-1}\Delta A(q^{-1}) \\ &\cdot A^{-1}(q^{-1})B_r(q^{-1})]F_r(U)\}. \end{aligned} \tag{18}$$

If $\Phi(t)$ is imperfectly known or difficult to estimate and adverse to a robust control, a neural-network based on the theory of universal approximator in last section will be employed to model it. Furthermore, the vector function $\Phi(t)$ is function of $U(t-d-i)$ for $i = 0, 1, \cdots$; i.e., the mapping between input $U(t-d)$ (or $\rho(t-d)$, $U_\nu(t-d)$) and output $\Phi(t)$ is dynamic. Hence, it is not suitable to use a multilayer-neural-network to approximate the vector function $\Phi(t)$. This is one of important motivation for this study. Equation (17) can be rewritten as the following observable canonical form [7]:

$$X(t+1) = AX(t) + BU_\nu(t-d+1) \tag{19a}$$

$$E(t) = CX(t) + \Phi(t) \tag{19b}$$

where

$$X(t) = [x_{11}(t) \quad x_{12}(t) \quad \cdots \quad x_{1m}(t) \quad \cdots \quad x_{\gamma 1}(t) \quad \cdots \quad x_{\gamma m}(t)]^T \in \Re^n$$

the pair (A, B, C) denotes the nominal system of (17) which is known, controllable and observable. Furthermore, the system state can be represented by the combination of $e_i(t), \cdots, e_i(t-\gamma+1)$, $u_{\nu i}(t-d), \cdots, u_{\nu i}(t-d-\gamma+2)$, and uncertainties $\phi_i(t)$, $1 \le i \le m$. That is,

$$X(t) = K_1 Z(t) + K_2 Z_\Phi(t) \tag{20a}$$

where $K_1 \in \Re^{n\times(n+m)}$ is function of A and B, $K_2 \in \Re^{n\times n}$ is the first n columns of K_1, and

$$\begin{aligned} Z(t) = [&e_1(t) \quad e_1(t-1) \quad \cdots \quad e_1(t-\gamma+1) \\ &\cdots \quad e_m(t) \quad e_m(t-1) \quad \cdots \quad e_m(t-\gamma+1) \\ &u_{\nu 1}(t-d) \quad \cdots \quad u_{\nu 1}(t-d-\gamma+2) \\ &\cdots \quad u_{\nu m}(t-d) \quad \cdots \quad u_{\nu m}(t-d-\gamma+2)]^T \\ \subset\ &\Re^{(n+m)\times 1} \end{aligned} \tag{20b}$$

$$\begin{aligned} Z_\Phi(t) = [&\phi_1(t) \quad \phi_1(t-1) \quad \cdots \quad \phi_1(t-\gamma+1) \\ &\cdots \quad \phi_m(t) \quad \phi_m(t-1) \quad \cdots \quad \phi_m(t-\gamma+1)]^T \\ \subset\ &\Re^{n\times 1}. \end{aligned} \tag{20c}$$

Due to the fact (20a), no state estimator is required for the controller [8]. This makes the controller design more practical; however, the uncertainties $K_2 Z_\Phi(t)$ occur. For the convenience of controller design, the uncertainties $K_2 Z_\Phi(t)$ rewrite as

$$K_2 Z_\Phi(t) = \overline{K}_2(q^{-1})\Phi(t) \tag{20d}$$

$$\overline{K}_2(q^{-1}) = [\overline{k}_{21}^T \quad \overline{k}_{22}^T q^{-1} \quad \cdots \quad \overline{k}_{2\gamma}^T q^{-\gamma+1}]^T \in \Re^{n\times m} \tag{20e}$$

and $\overline{k}_{2i}^T (1 \le i \le \gamma) \in \Re^{m\times m}$ are function of K_2.

C. Updating Law for Weight Matrix

First, the following sliding surface with integral feature is defined.

$$S(t) = \lambda S(t-1) + D_1 E(t) - D_2 E(t-1) \tag{21}$$

where

$$S(t) = [s_1(t) \quad s_2(t) \quad \cdots \quad s_m(t)]^T \in \Re^{m\times 1}, \quad 1^- \le \lambda \le 1,$$
$$D_i = \text{diag}(d_{ijj}), \quad |d_{2ii}/d_{1ii}| < 1, \quad i = 1, 2 \text{ and } 1 \le j \le m.$$

It is assumed that

$$\text{A4:} \quad \|(CB)^{-1}[I - q^{-1}CA\overline{K}_2(q^{-1})]A_r^{-1}(q^{-1})q^{-d+1} \cdot B_r(q^{-1})\Delta F_1\|_\infty = \overline{\beta} < 1 \tag{22}$$

where

$$\|A(q^{-1})\|_\infty = \text{ess.} \sup_{0\le\theta\le 2\pi} \overline{\lambda}[A(e^{-j\theta})], \overline{\lambda}[\cdot]$$

denotes the maximum singular eigenvalue [25]. Then we consider the following updating law for the weight matrix:

$$\hat{W}_i(t+1) = \hat{W}_i(t) + \Lambda_i(t) - \eta_i \hat{W}_i(t), \quad i = 1, 2, 4 \tag{23a}$$

$$\hat{W}_3(t+1) = \hat{W}_3(t) + \Lambda_3(t) - \eta_3\hat{W}_3(t) - P_r(t)\hat{W}_3(t) \tag{23b}$$

where

$$\begin{aligned}\Lambda_1(t) = \alpha_1[&\hat{\sigma}(\overline{\Psi}) - \hat{\sigma}'(\overline{\Psi})\hat{W}_2^T(t)\overline{\Psi}(t) \\ &- 2\hat{\sigma}'(\overline{\Psi})\hat{W}_3^T(t)q^{-1}(\hat{\sigma})]S^T(t)/2\end{aligned} \tag{24}$$

$$\Lambda_2(t) = \alpha_2\overline{\Psi}(t)S^T(t)\hat{W}_1^T(t)\hat{\sigma}'(\overline{\Psi})/2 \tag{25}$$

$$\Lambda_3(t) = \alpha_3 q^{-1}(\hat{\sigma})S^T(t)\hat{W}_1^T(t)\hat{\sigma}'(\overline{\Psi}) \tag{26a}$$

$$P_r(t) = \begin{cases} 0, & \text{if } \|\hat{W}_3(t)\|_F < \sqrt{p}, \text{ or} \\ & \text{if} \|\hat{W}_3(t)\|_F \geq \sqrt{p}, \\ & \text{tr}\{[\Lambda_3(t) - \eta_3 \hat{W}_3(t)]^T \hat{W}_3(t)\} \leq 0 \\ \text{tr}\{[\Lambda_3(t) - \eta_3 \hat{W}_3(t)]^T \hat{W}_3(t)\}/[\kappa + \|\hat{W}_3(t)\|_F^2], & \\ & \text{if} \|\hat{W}_3(t)\|_F \geq \sqrt{p} \\ & \text{tr}\{[\Lambda_3(t) - \eta_3 \hat{W}_3(t)]^T \hat{W}_3(t)\} > 0 \end{cases} \tag{26b}$$

$$\Lambda_4(t) = \alpha_4 \|S(t)\| \Gamma(t)/2 \tag{27}$$

$$\overline{\Psi}(t) = [\rho^T(t-d) \quad U_\nu^T(t-d) \quad b]^T \tag{28}$$

where $0 \leq \kappa, 0 < \alpha_i, 0 < \eta_i < 1 (1 \leq i \leq 4)$, and b is a known constant. The updating laws (23) have learning rates $\alpha_i (1 \leq i \leq 4)$, error function $S(t)$, and specific basis functions in $\Lambda_i(t) (1 \leq i \leq 4)$ except α_i and $S(t)$. Generally, the learning rate should be chosen small enough to avoid the instability of closed-loop system. The selection of $\eta_i \hat{W}_i(t)$ in (23) is the reason for the boundedness of estimated weight matrix without the requirement of PE condition. Otherwise, the drift of estimated weight probably occurs [2], [26]. In general, $\eta_i (1 \leq i \leq 4)$ are very small to allow a possibility of effective learning of $\hat{W}_i(t) (1 \leq i \leq 4)$. Too large values of $\eta_i (1 \leq i \leq 4)$ will force $\hat{W}_i(t) (1 \leq i \leq 4)$ converge into the neighborhood of zero. Under the circumstance, if $\overline{W}_i (1 \leq i \leq 4)$ are not very small, a poor learning of $\hat{W}_i(t) (1 \leq i \leq 4)$ or approximation of nonlinear function occurs. The projection term $P_r(t)\hat{W}_3(t)$ in (23b) ensures that $\|\hat{W}_3(t)\|_F < \sqrt{p}$ as $t \to \infty$.

D. Discrete-Time Multivariable Neuro-Adaptive Variable Structure Control

Based on the result of RNNRUBC in Section III, the following approximation of uncertainties as $U(t) \in \Omega$ (see Fig. 3) can be achieved.

$$\begin{aligned} P(U) = & -D_1[I + q^{-1} CA\overline{K}_2(q^{-1})] A^{-1}(q^{-1}) q^{-d+1} \{B(q^{-1}) \\ & \cdot [\Delta F_0 U_b + \Delta F_1(\rho(t) - U_{eq}(t)) + \Delta F_2(\rho - U_\nu)] \\ & + [\Delta B(q^{-1}) - (I + \Delta A(q^{-1}) A^{-1}(q^{-1}))^{-1} \Delta A(q^{-1}) \\ & \cdot A^{-1}(q^{-1}) B_r(q^{-1})] [\Delta F_0 U_b + (I + \Delta F_1) \\ & \cdot (\rho(t) - U_{eq}(t)) + \Delta F_2(\rho - U_\nu)]\} \\ = & \overline{W}_1^T \sigma(\overline{W}_2^T \overline{\Psi} + \overline{W}_3^T q^{-1}(\sigma)) + \varepsilon_f(\overline{\Psi}) \end{aligned} \tag{29}$$

where $\|\varepsilon_f(\overline{\Psi})\| \leq \varepsilon, \Omega = \{U(t) \in \overline{\Omega} | \, \|U(t-d+1) - U_0\| \leq \gamma < \overline{\gamma}\}$ and $U_{eq}(t-d+1)$ is described in (33). Because only the uncertainties as $U(t) \in \Omega$ are approximated by the RNNRUBC, the number (or order) of $\overline{W}_1$, $\overline{W}_2$ and $\overline{W}_3$ the value of $\varepsilon_f(\overline{\Psi})$ (or ε) are not necessarily large. The following signal connected with the system uncertainties without showing its argument is given:

$$\begin{aligned} Q = & \tilde{W}_1^T [\hat{\sigma} - \hat{\sigma}' \hat{W}_2^T \overline{\Psi} - 2\hat{\sigma}' \hat{W}_3^T q^{-1}(\hat{\sigma})] \\ & + \hat{W}_1^T \hat{\sigma}' [\tilde{W}_2^T \overline{\Psi} + 2\tilde{W}_3^T q^{-1}(\hat{\sigma})] + \tilde{\varepsilon}_f \\ & - [\Lambda_1^T/2 - \eta_1 \hat{W}_1^T] [\hat{\sigma} - \hat{\sigma}' \hat{W}_2^T \overline{\Psi} - 2\hat{\sigma}' \hat{W}_3^T q^{-1}(\hat{\sigma})] \\ & - \hat{W}_1^T \hat{\sigma}' [\Lambda_2^T/2 - \eta_2 \hat{W}_2^T] \overline{\Psi} \\ & - \hat{W}_1^T \hat{\sigma}' [\Lambda_3^T - 2\eta_3 \hat{W}_3^T] q^{-1}(\hat{\sigma}) \\ & - \alpha_4 S \Gamma^T \Gamma/4 - (1-\eta_4) S \hat{W}_4^T \Gamma / \|S\|. \end{aligned} \tag{30}$$

Because the signals $\tilde{W}_i(t) (1 \leq i \leq 4)$ are unavailable, the inequalities $\|\tilde{W}_i(t)\| \leq W_{im} + \|\hat{W}_i(t)\|_F (1 \leq i \leq 4)$ must be used to obtain the upper bound of (30)

$$\begin{aligned} \|Q\| \leq f_q = & [W_{1m} + (1-\eta_1) \|\hat{W}_1\|_F] \|\hat{\sigma} - \hat{\sigma}' W_2^T \overline{\Psi} \\ & - 2\hat{\sigma}' \hat{W}_3^T q^{-1}(\hat{\sigma})\| \\ & + [W_{2m} + (1-\eta_2) \|\hat{W}_2\|_F] \|\hat{W}_1^T \hat{\sigma}'\| \|\overline{\Psi}\| \\ & + 2[W_{3m} + (1-\eta_3) \|\hat{W}_3\|_F] \|\hat{W}_1^T \hat{\sigma}'\| \|q^{-1}(\hat{\sigma})\| \\ & + [W_{4m} + (1-\eta_4) \|\hat{W}_4\|] \|\Gamma\| \\ & + \|S\| \{\alpha_1 \|\hat{\sigma} - \hat{\sigma}' \hat{W}_2^T \overline{\Psi} - 2\hat{\sigma}' \hat{W}_3^T q^{-1}(\hat{\sigma})\|^2/4 \\ & + \alpha_2 \|\hat{W}_1^T \hat{\sigma}'\|^2 \|\overline{\Psi}\|^2/4 + \alpha_3 \|\hat{W}_1^T \hat{\sigma}'\|^2 \|q^{-1}(\hat{\sigma})\|^2 \\ & + \alpha_4 \|\Gamma\|^2/4\}. \end{aligned} \tag{31}$$

Some papers have used the upper bound of $\tilde{W}_i(t) (1 \leq i \leq 4)$ as the control parameters, e.g., [27]. They are unreasonable. If $\eta_i = 0 (1 \leq i \leq 4)$ in (31) is assigned, the signal $f_q(t)$ becomes a function of $S(t)$ and $\tilde{W}_i(t) (1 \leq i \leq 4)$. That is, the upper bound of uncertainties can be arbitrarily small if $S(t)$ and $\tilde{W}_i(t) (1 \leq i \leq 4)$ converge to the vicinity of zero. This feature is different from robust control that its upper bound of uncertainty always exists. Before designing the proposed control, the following lemma about the properties of trace operator is given.

Lemma 4: Define the trace operator as tr$[\cdot]$ and tr$[A] = \sum_i^n a_{ii}$, where $A \in \Re^{n \times n}$. Then

(a) tr$[ABC]$ = tr$[CAB]$ = tr$[BCA]$, where A, B and C are three compatible square (or nonsquare) matrices

(b) tr$[\tilde{W}^T \hat{W}] = -[\|\tilde{W}\|_F^2 + \|\hat{W}\|_F^2 - \|\overline{W}\|_F^2]/2$.

Proof:

(a) Using the definition of trace gives the result.

(b) Using the definitions $\tilde{W} = \overline{W} - \hat{W}$ and tr$[W^T W] = \|W\|_F^2$ yields the result.

The following theorem is the main theorem of this paper.

Theorem 2: Consider the system (20) with (23) and the following controller (32):

$$U_\nu(t-d+1) = U_{eq}(t-d+1) + U_{sw}(t-d+1) \tag{32}$$

where

$$\begin{aligned} & U_{eq}(t-d+1) \\ & \quad = \overline{U}_{eq}(t-d+1) - (D_1 CB)^{-1} \{\hat{W}_1^T(t) \hat{\sigma}(t) \\ & \qquad + [\Lambda_1^T(t)/2 - \eta_1 \hat{W}_1^T(t)] [\hat{\sigma}(t) - \hat{\sigma}'(t) \hat{W}_2^T(t) \overline{\Psi}(t) \\ & \qquad + 2\hat{\sigma}'(t) \hat{W}_3^T(t) q^{-1}(\hat{\sigma})] \\ & \qquad + \hat{W}_1^T(t) \hat{\sigma}'(t) [\Lambda_2^T(t)/2 - \eta_2 \hat{W}_2^T(t)] \overline{\Psi}(t) \\ & \qquad + \hat{W}_1^T(t) \hat{\sigma}'(t) [\Lambda_3^T(t) - 2\eta_3 \hat{W}_3^T(t)] q^{-1}(\hat{\sigma}) \\ & \qquad + \alpha_4 S(t) \Gamma^T(t) \Gamma(t)/4 + (1-\eta_4) S(t) \hat{W}_4^T(t) \\ & \qquad \cdot \Gamma(t)/\|S(t)\|\} \end{aligned} \tag{33a}$$

$$\begin{aligned} & \overline{U}_{eq}(t-d+1) \\ & \quad = -(D_1 CB)^{-1} \{D_1 CAK_1 Z(t) - D_2 E(t) \\ & \qquad - (1-\lambda) S(t)\} \end{aligned} \tag{33b}$$

$$\begin{aligned} & U_{sw}(t-d+1) \\ & \quad = \begin{cases} -\xi(t) f_q(t) (D_1 CB)^{-1} S(t) / \\ \quad [(1-\overline{\beta})(1+\overline{\beta}) \|S(t)\|], \\ \qquad \text{if } \|S(t)\| > 2(1-\overline{\beta})(1+\overline{\beta}) f_q(t) / \\ \qquad [(1-\overline{\beta})^2 - (1+\overline{\beta})^2 \mu] \\ 0, \quad \text{otherwise} \end{cases} \end{aligned} \tag{34}$$

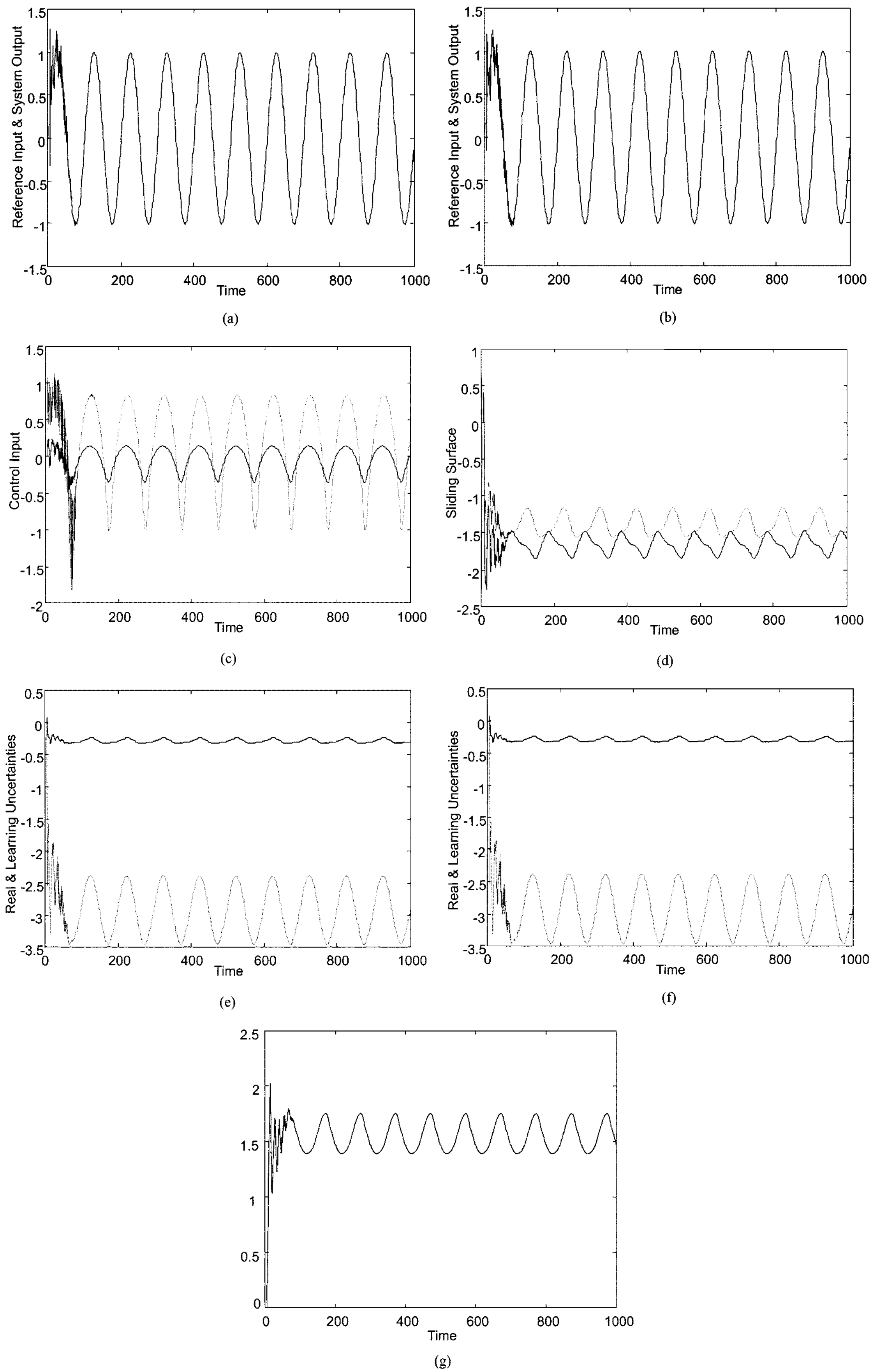


Fig. 5. Responses of neuro-adaptive control: (a) $r_1(t)$ $(\ldots)$, $y_1(t)$ (—); (b) $r_2(t)$ $(\ldots)$, $y_2(t)$ (—); (c) $u_1(t)$ $(\ldots)$, $u_2(t)$ (—); (d) $s_1(t)$ $(\ldots)$, $s_2(t)$ (—); (e) $p_1(U)$ $(\ldots)$ $p_1(U)$; (—); (f) $p_2(U)$ $(\ldots)$, $\hat{p}_2(U)$ (—); (g) $\hat{W}_4^T(t)\Gamma(t)$.

where $0 < \mu < (1-\overline{\beta})^2/(1+\overline{\beta})^2 < 1$. The amplitude of switching gains $\xi(t)$ is achieved from the following inequality:

$$\xi_2(t) > \xi(t) > \xi_1(t) \geq 0 \tag{35}$$

where

$$\xi_{1,2}(t) = h_1(t) \pm \sqrt{h_1^2(t) - h_2(t)} \tag{36}$$

$$h_1(t) = (1-\overline{\beta})^2\|S(t)\|/[(1+\overline{\beta})f_q(t)] - (1-\overline{\beta}) \tag{37}$$

$$h_2(t) = (1-\overline{\beta})^2[f_q^2(t) + 2f_q\|S(t)\| + \mu\|S(t)\|^2]/f_q^2(t). \tag{38}$$

If the overall system satisfies Assumptions A1–A4 and $\|\hat{W}_3(0)\|_F \leq W_{3m} < \sqrt{p}$, then $\{\hat{W}_1(t), \hat{W}_2(t), \hat{W}_3(t), \hat{W}_4(t), S(t), U_\nu(t-d+1)\}$ are bounded, $\|\hat{W}_3(t)\|_F < \sqrt{p}$ and

$$\|S(t)\| < 2(1-\overline{\beta})(1+\overline{\beta})f_q(t)/[(1-\overline{\beta})^2 - (1+\overline{\beta})^2\mu]$$

as $t \to \infty$.

Proof: Consider the case of $U \in \Omega$. Define the following Lyapunov function:

$$\begin{aligned} V(\tilde{W}_1, \tilde{W}_2, \tilde{W}_3, \tilde{W}_4, S) &= \sum_{i=1}^{4} \text{tr}[\tilde{W}_i^T\tilde{W}_i]/\alpha_i + S^TS/2 \\ &= \overline{Z}^T\overline{P}\,\overline{Z} \end{aligned} \tag{39}$$

where

$$\begin{aligned} \overline{Z}^T &= [\,\|\tilde{W}_1\|_F \quad \|\tilde{W}_2\|_F \quad \|\tilde{W}_3\|_F \quad \|\tilde{W}_4\|_F \quad \|S\|\,], \\ \overline{P} &= \text{diag}[1/\alpha_1 \quad 1/\alpha_2 \quad 1/\alpha_3 \quad 1/\alpha_4 \quad 1/2]. \end{aligned} \tag{40}$$

First, the case of $P_r = 0$ is examined. By using (23), the change rate of $V(\tilde{W}_1, \tilde{W}_2, \tilde{W}_3, \tilde{W}_4, S)$ is described as follows:

$$\begin{aligned} \Delta V = \sum_{i=1}^{5}\Delta V_i = \sum_{i=1}^{4} &\text{tr}[(\tilde{W}_i - \Lambda_i + \eta_i\hat{W}_i)^T \\ &\cdot (\tilde{W}_i - \Lambda_i + \eta_i\hat{W}_i) - \tilde{W}_i^T\tilde{W}_i]/\alpha_i \\ &+ \Delta S^T\Delta S/2 + S^T\Delta S. \end{aligned} \tag{41}$$

Similarly,

$$\begin{aligned} \Delta S &= D_1[CA\overline{K}_2 + qI]\Phi(t) + (D_1CB)U_{sw} \\ &= Q + (D_1CB)\Sigma U_{sw} \end{aligned} \tag{42}$$

where Q is shown in (30) and

$$\Sigma = I + (CB)^{-1}[I + q^{-1}CA\overline{K}_2]A_r^{-1}q^{-d+1}B_r\Delta F_1. \tag{43}$$

Then ΔV_i for $1 \leq i \leq 4$ are considered

$$\begin{aligned} \Delta V_i = \{&-2\,\text{tr}[\tilde{W}_i^T\Lambda_i] + 2\eta_i\,\text{tr}[\tilde{W}_i^T\hat{W}_i]\,\text{tr}[\Lambda_i^T\,\Lambda_i] \\ &- 2\eta_i\,\text{tr}[\hat{W}_i^T\Lambda_i] + \eta_i^2\,\text{tr}[\hat{W}_i^T\hat{W}_i]\}/\alpha_i. \end{aligned} \tag{44}$$

By inspecting (40)–(44) and (30), using the result of Lemma 4 gives the following equalities:

$$\begin{aligned} &-2\,\text{tr}[\tilde{W}_1^T\Lambda_1]/\alpha_1 = -S^T\tilde{W}_1^T[\hat{\sigma} - \hat{\sigma}'\hat{W}_2^T\overline{\Psi} - 2\hat{\sigma}'\hat{W}_3^Tq^{-1}(\hat{\sigma})] \\ &-2\,\text{tr}[\tilde{W}_2^T\Lambda_2]/\alpha_2 = -S^T\hat{W}_1^T\hat{\sigma}'\tilde{W}_2^T\overline{\Psi} \\ &-2\,\text{tr}[\tilde{W}_3^T\Lambda_3]/\alpha_3 = -2S^T\hat{W}_1^T\hat{\sigma}'\tilde{W}_3^Tq^{-1}(\hat{\sigma}) \\ &\{\text{tr}[\Lambda_1^T\Lambda_1] - 2\eta_1\,\text{tr}[\hat{W}_1^T\Lambda_1]\}/\alpha_1 \\ &\quad = S^T[\Lambda_1^T/2 - \eta_1\hat{W}_1^T]\,[\hat{\sigma} - \hat{\sigma}'\hat{W}_2^T\overline{\Psi} - 2\hat{\sigma}'\hat{W}_3^Tq^{-1}(\hat{\sigma})] \\ &\{\text{tr}[\Lambda_2^T\Lambda_2] - 2\eta_2\,\text{tr}[\hat{W}_2^T\Lambda_2]\}/\alpha_2 \\ &\quad = S^T\hat{W}_1^T\hat{\sigma}'[\Lambda_2^T/2 - \eta_2\hat{W}_2^T]\overline{\Psi} \\ &\{\text{tr}[\Lambda_3^T\Lambda_3] - 2\eta_3\,\text{tr}[\hat{W}_3^T\Lambda_3]\}/\alpha_3 \\ &\quad = S^T\hat{W}_1^T\hat{\sigma}'[\Lambda_3^T - 2\eta_3\hat{W}_3^T]q^{-1}(\hat{\sigma}) \\ &\{\text{tr}[\Lambda_4^T\Lambda_4] - 2\eta_4\,\text{tr}[\hat{W}_4^T\Lambda_4]\}/\alpha_4 \\ &\quad = S^T[\alpha_4S\Gamma^T\Gamma/4 - \eta_4S\hat{W}_4^T\Gamma/\|S\|\,]. \end{aligned} \tag{45}$$

In addition, the following inequality is achieved from (12) and Remark 3:

$$\begin{aligned} &S^T[\tilde{\varepsilon}_f - S\hat{W}_4^T\Gamma/\|S\|\,] - 2\,\text{tr}[\tilde{W}_4^T\Lambda_4]/\alpha_4 \\ &\quad \leq \|S\|\,\|\tilde{\varepsilon}_f\| - \|S\|\hat{W}_4^T\Gamma - \|S\|\tilde{W}_4^T\Gamma \leq 0. \end{aligned} \tag{46}$$

Using (41)–(46) and Remark 3 yields

$$\begin{aligned} \Delta V = \sum_{i=1}^{4} &\{2\eta_i\,\text{tr}[\tilde{W}_i^T(\overline{W}_i - \tilde{W}_i)] \\ &+ \eta_i^2\,\text{tr}[(\overline{W}_i - \tilde{W}_i)^T(\overline{W}_i - \tilde{W}_i)]\}/\alpha_i \\ &+ \Delta S^T\Delta S/2 + S^TD_1CB\Sigma U_{sw} \\ &+ S^T[\tilde{\varepsilon}_f - S\hat{W}_4^T\Gamma/\|S\|\,] - 2\,\text{tr}[\tilde{W}_4^T\Lambda_4]/\alpha_4 \\ \leq -\sum_{i=1}^{4} &\{\eta_i(2-\eta_i)\,\|\tilde{W}_i\|_F^2 - 2\eta_i(1-\eta_i)W_{im}\|\tilde{W}_i\|_F \\ &- \eta_i^2W_{im}^2\}/\alpha_i + \Delta S^T\Delta S/2 + S^TD_1CB\Sigma U_{sw}. \end{aligned} \tag{47}$$

Assume that

$$\begin{aligned} &-[\eta_i(2-\eta_i)\|\tilde{W}_i\|_F^2 - 2\eta_i(1-\eta_i)W_{im}\|\tilde{W}_i\|_F \\ &\quad - \eta_i^2W_{im}^2]/\alpha_i \leq -\lambda_i\|\tilde{W}_i\|_F^2/\alpha_i, \qquad 1 \leq i \leq 4. \end{aligned} \tag{48}$$

Hence, $\Delta V_i \leq -\lambda_iV_i$ for $1 \leq i \leq 4$ imply that $F_i/\alpha_i \leq 0$, where

$$\begin{aligned} F_i = &-[\eta_i(2-\eta_i) - \lambda_i] \\ &\cdot \left\{\left[\|\tilde{W}_i\|_F - \frac{\eta_i(1-\eta_i)W_{im}}{\eta_i(2-\eta_i) - \lambda_i}\right]^2 \right. \\ &\left. - \frac{(1-\lambda_i)\eta_i^2W_{im}^2}{[\eta_i(2-\eta_i) - \lambda_i]^2}\right\} \end{aligned} \tag{49}$$

and $\eta_i(2-\eta_i) - \lambda_i > 0, 1 > \eta_i > \lambda_i > 0$ for $1 \leq i \leq 4$. If

$$\|\tilde{W}_i\|_F \geq [(1-\eta_i) + \sqrt{1-\lambda_i}\,]\eta_iW_{im}/[\eta_i(2-\eta_i) - \lambda_i]$$

for $1 \le i \le 4$, then $F_i(1 \le i \le 4) \le 0$. Similarly, $\Delta V_5 \le -\mu V_5$ (or $\Delta \overline{V}_5 = \Delta V_5 - \mu V_5 \le 0$) is assumed. Then,

$$\begin{aligned}\Delta \overline{V}_5 &= U_{sw}^T \Sigma^T (D_1 CB)^T (D_1 CB) \Sigma U_{sw}/2 \\ &\quad + Q^T (D_1 CB) \Sigma U_{sw} + Q^T Q/2 + S^T Q \\ &\quad + S^T (D_1 CB) \Sigma U_{sw} + \mu S^T S/2 \\ &\le \frac{\xi^2 f_q^2}{2(1-\overline{\beta})^2} + \frac{\xi f_q^2}{(1-\overline{\beta})} + \frac{f_q^2}{2} \\ &\quad + \|S\| f_q - \frac{\xi f_q \|S\|}{(1+\overline{\beta})} + \frac{\mu \|S\|^2}{2} \\ &= f_q^2 \{\xi^2 - 2h_1 \xi + h_2\} / \{2(1-\overline{\beta})^2\} \end{aligned} \tag{50}$$

where the inequality has used Assumption A4 and (43), the expressions of h_1 and h_2 are shown in (37) and (38). Because

$$\|S\| > 2(1-\overline{\beta})(1+\overline{\beta}) f_q / [(1-\overline{\beta})^2 - (1+\overline{\beta})^2 \mu]$$

$h_1 > 0$ and $h_1^2 - h_2 > 0$. The result (35) is achieved from the inequality $\xi^2 - 2h_1\xi + h_2 < 0$. Hence, the change rate of Lyapunov function becomes

$$\Delta V \le -\sum_{i=1}^{4} \lambda_i V_i - \mu V_5 \le -\min_{1 \le i \le 4} (\lambda_i, \mu) V = -\lambda_0 V \tag{51}$$

where $0 < \lambda_0 < 1$. Then, $V(t+1) - (1-\lambda_0)V(t) \le 0$. Hence, outside of the domain $\overline{D}$ making $\Delta V \le -\lambda_0 V$ is described as follows:

$$\begin{aligned}\overline{D} = \{\overline{Z} \in \Re^5 | 0 \le \|\tilde{W}_i\|_F \le W_i^*, 1 \le i \le 4 \text{ and} \\ 0 \le \|S\| \le S^*\}\end{aligned} \tag{52a}$$

where

$$W_i^* = [(1-\eta_i) + \sqrt{1-\lambda_i}]\eta_i W_{im} / [\eta_i(2-\eta_i) - \lambda_i], \quad 1 \le i \le 4 \tag{52b}$$

$$S^* > 2(1-\overline{\beta})(1+\overline{\beta}) f_q / [(1-\overline{\beta})^2 - (1+\overline{\beta})^2 \mu]. \tag{52c}$$

Finally, from (30)–(34), is bounded.

Then the case of $P_r \ne 0$ is examined. The above results are the same except that ΔV_3 in (44) has the extra second term in the right-hand side of (53)

$$\begin{aligned}\Delta V_3 \le &-\{\eta_3(2-\eta_3)\|\tilde{W}_3\|_F^2 - 2\eta_3(1-\eta_3)W_{3m} \\ &\cdot \|\tilde{W}_3\|_F - \eta_3^2 W_{3m}^2\}/\alpha_3 + \{2P_r \operatorname{tr}[\tilde{W}_3^T \hat{W}_3] \\ &- 2P_r \operatorname{tr}[(\Lambda_3 - \eta_3 \hat{W}_3)^T \hat{W}_3] \\ &+ P_r^2 \operatorname{tr}[\hat{W}_3^T \hat{W}_3]\}/\alpha_3 \le -\lambda_3 V_3. \end{aligned} \tag{53}$$

Substituting the P_r in (26b) and part (b) of Lemma 4 into the second term of (53) gives

$$\begin{aligned}F_{3p} &= 2P_r \operatorname{tr}[\tilde{W}_3^T \hat{W}_3] - 2P_r \operatorname{tr}[(\Lambda_3 - \eta_3 \hat{W}_3)^T \hat{W}_3] \\ &\quad + P_r^2 \operatorname{tr}[\hat{W}_3^T \hat{W}_3] \\ &= -[\|\tilde{W}_3\|_F^2 + \|\hat{W}_3\|_F^2 - \|\overline{W}_3\|_F^2] \\ &\quad \cdot \operatorname{tr}[(\Lambda_3 - \eta_3 \hat{W}_3)^T \hat{W}_3] / (\kappa + \|\hat{W}_3\|_F^2) \\ &\quad - 2 \operatorname{tr}[(\Lambda_3 - \eta_3 \hat{W}_3)^T \hat{W}_3]^2 / (\kappa + \|\hat{W}_3\|_F^2) \\ &\quad + \operatorname{tr}[(\Lambda_3 - \eta_3 \hat{W}_3)^T \hat{W}_3]^2 \\ &\quad \cdot \operatorname{tr}[\hat{W}_3^T \hat{W}_3] / (\kappa + \|\hat{W}_3\|_F^2)^2. \end{aligned} \tag{54}$$

Because $\|\overline{W}_3\|_F^2 \le W_{3m}^2$, if $\|\hat{W}_3\|_F \ge \sqrt{p} > W_{3m}$ and $\operatorname{tr}[(\Lambda_3 - \eta_3 \hat{W}_3)^T \hat{W}_3] > 0$, then $F_{3p} < 0$. That is, the projection algorithm (26b) makes the original $\Delta V_3 \le -\lambda_3 V_3$ more negative [7], [19]. Because $0 < \overline{\varepsilon} \le \|\hat{W}_3(0)\| \le W_{3m} \le \sqrt{p}$ and the value $0 < \overline{\varepsilon} \le \|\overline{W}_3\| \le W_{3m} \le \sqrt{p}$ exists, then $\|\hat{W}_3\| \le \sqrt{p}$ as $t \to \infty$. Q.E.D.

Remark 7: The proposed $U_{eq}(t-d+1)$ contains the term $(1-\eta_4)S(t)\hat{W}_4^T(t)\Gamma(t)/\|S(t)\|$ to compensate the upper bound of residue of RNN. This feature does not appear in the previous studies [2], [11]–[18], [20], [26], [27]. The size of dead-zone for $U_{sw}(t-d+1)$ (i.e., $\|S(t)\| > 2(1-\overline{\beta})(1+\overline{\beta})f_q(t)/[(1-\overline{\beta})^2 - (1+\overline{\beta})^2\mu]$) is generally smaller than that in robust variable structure control (i.e., $\|S(t)\| > 4(1+\overline{\beta})\overline{f}_q(t)/[(1-\overline{\beta})^2 - (1+\overline{\beta})^2\mu]$, where $\overline{f}_q(t) = \beta_0 + \overline{\beta}\|\overline{U}_{eq}(t-d+1)\|$). The responses in Figs. 4(d) and 5(d) indicate these features. Hence, tracking performance of proposed control (i.e., Fig. 5(a) and (b) is better than that of robust control (i.e., Fig. 4(a) and (b)).

V. Computer Simulations

Consider two inputs and two outputs system with the following system matrices:

$$\begin{aligned}A_r(q^{-1}) &= \begin{bmatrix}1 & 0\\0 & 1\end{bmatrix} + \begin{bmatrix}a_{111}+\Delta a_{111} & a_{112}+\Delta a_{112}\\a_{121}+\Delta a_{121} & a_{122}+\Delta a_{122}\end{bmatrix} q^{-1} \\ &\quad + \begin{bmatrix}a_{211}+\Delta a_{211} & a_{212}+\Delta a_{212}\\a_{221}+\Delta a_{221} & a_{222}+\Delta a_{222}\end{bmatrix} q^{-2} \\ &\quad + \begin{bmatrix}\Delta a_{311} & \Delta a_{312}\\\Delta a_{321} & \Delta a_{322}\end{bmatrix} q^{-3} \\ &\quad + \begin{bmatrix}\Delta a_{411} & \Delta a_{412}\\\Delta a_{421} & \Delta a_{422}\end{bmatrix} q^{-4} \\ B_r(q^{-1}) &= \begin{bmatrix}b_{011}+\Delta b_{011} & b_{012}+\Delta b_{012}\\b_{021}+\Delta b_{021} & a_{022}+\Delta b_{022}\end{bmatrix} \\ &\quad + \begin{bmatrix}b_{111}+\Delta b_{111} & a_{112}+\Delta b_{112}\\b_{121}+\Delta b_{121} & a_{122}+\Delta b_{122}\end{bmatrix} q^{-1} \\ &\quad + \begin{bmatrix}\Delta b_{211} & \Delta b_{212}\\\Delta b_{221} & \Delta b_{222}\end{bmatrix} q^{-2} \\ &\quad + \begin{bmatrix}\Delta b_{311} & \Delta b_{312}\\\Delta b_{321} & \Delta b_{322}\end{bmatrix} q^{-3} \\ F_r(U) &= F(U) + \Delta F(U)\end{aligned}$$

where the nominal coefficients are described as follows:

$a_{111} = -0.6$, $a_{211} = 0.2$, $a_{112} = 0.4$, $a_{212} = -0.2$, $b_{011} = 1$, $b_{111} = 0.5$, $b_{012} = 0.8$, $b_{112} = 0.5$ (1st channel), $a_{121} = -0.5$, $a_{221} = -0.2$, $a_{122} = 0.6$, $a_{222} = 0.3$, $b_{021} = 0.5$, $b_{121} = 1$, $b_{022} = 2$, $b_{122} = 0.5$ (2nd channel), $\gamma_{11} = 1$, $\gamma_{12} = 1$, $\gamma_{31} = 0.35$ (1st channel); $\gamma_{21} = 2.5$, $\gamma_{22} = 0.5$, $\gamma_{11} = 2.5$ (2nd channel), and $d = 1$.

The poles and zeros of nominal system are: $-0.1338\pm0.5808j$, $0.1338 \pm 0.1959j$ (stable) and -0.5602, 0.2789 (minimum phase), respectively. The reference input is assigned as $R(t) = [\sin(2\pi t) \quad \sin(2\pi t)]^T$. Consider the following uncertainties those are not very small.

$$\begin{aligned}
&\Delta a_{111} = 0.3, \quad \Delta a_{211} = 0.25, \quad \Delta a_{311} = 0.05,\\
&\Delta a_{411} = -0.02, \quad \Delta a_{112} = -0.2, \quad \Delta a_{212} = -0.1,\\
&\Delta a_{312} = 0.02, \quad \Delta a_{412} = 0.01, \quad \Delta b_{011} = 0.2,\\
&\Delta b_{111} = 0.5, \quad \Delta b_{211} = -0.02, \quad \Delta b_{311} = -0.01,\\
&\Delta b_{012} = 0.2, \quad \Delta b_{112} = -1.8, \quad \Delta b_{212} = 0.01,\\
&\Delta b_{312} = 0.005, \quad \Delta a_{121} = 0.2, \quad \Delta a_{221} = -0.12,\\
&\Delta a_{312} = 0.03, \quad \Delta a_{421} = -0.02, \quad \Delta a_{122} = -0.12,\\
&\Delta a_{222} = -0.1, \quad \Delta a_{322} = 0.03, \quad \Delta a_{422} = -0.01,\\
&\Delta b_{021} = 0.1, \quad \Delta b_{121} = -0.2, \quad \Delta b_{221} = 0.025,\\
&\Delta b_{321} = 0.02, \quad \Delta b_{022} = -1.46, \quad \Delta b_{122} = 0.5,\\
&\Delta b_{222} = 0.04, \quad \Delta b_{322} = 0.02.
\end{aligned}$$

Then the poles and zeros of real system are: $0.0432 \pm 0.7795j$, $0.0373 \pm 0.3873j$, -0.5287, 0.2430, 0.1581, -0.2134 (stable) and $-0.0108\pm0.1481j$, -34.6021, -1.2250, -0.0401, 0.0556 (nonminimum phase), respectively. Furthermore, the vector function of nonlinear uncertainty is described as follows (see equation at the bottom of the page):

$$\Delta F_2(U) = \Delta F F^{-1}(U) - (\Delta F_0 U_b + \Delta F_1 U(t))$$

where

$$\Delta F_0 = \begin{bmatrix} 0.22 & 0.326 \\ 0.02 & 0.026 \end{bmatrix}, \quad \Delta F_1 = \begin{bmatrix} 0.06 & 0.02 \\ 0.0168 & 0.012 \end{bmatrix}.$$

The responses of forward control are presented in Fig. 3. The robust variable structure control is described as follows: $U(t-d+1) = \overline{U}_{eq}(t-d+1) + U_{sw}(t-d+1)$ (see Remark 7). The control parameters are chosen as follows:

$$\begin{aligned}
&\beta_0 = /0.001, \quad \overline{\beta} = 0.01, \quad \mu = 0.0085,\\
&d_{111} = d_{122} = 1, \quad d_{211} = 0.3, \quad d_{222} = 0.25,\\
&\lambda = 1, \quad \xi(t) = \xi_1(t) + \xi_0\sqrt{h_2^2 - h_1}\,(1 - 0.98e^{-\xi_3\|S(t)\|})
\end{aligned}$$

where $\xi_0 = 0.053$ and $\xi_3 = 100$. The simulation results are shown in Fig. 4. As the uncertainties become large, the range for the selection of control parameter in robust variable structure control becomes small. In other words, the rude choose of control parameter probably causes the instability of closed-loop system. From Figs. 3 and 4, one realizes that the responses of forward control are poor, and those of robust variable structure control are not excellent. The main reason of the result is that the

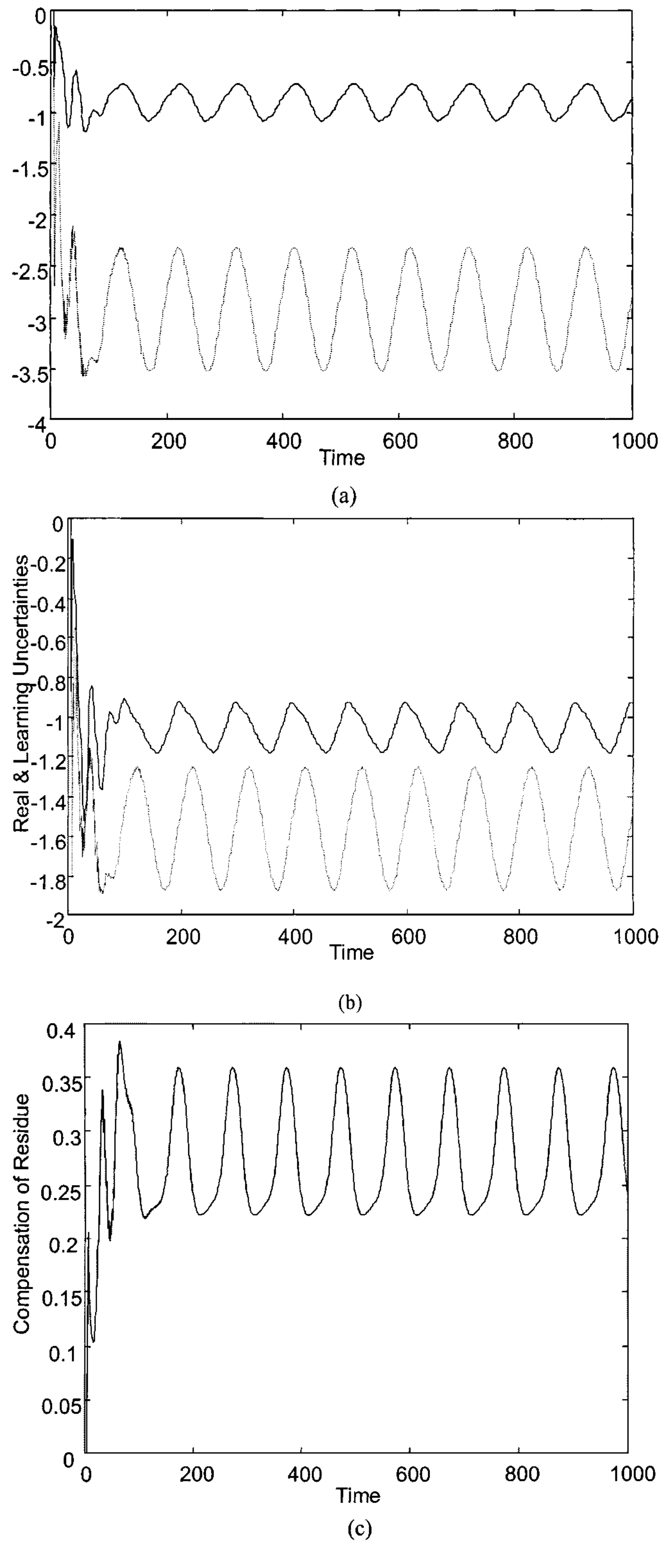


Fig. 6. Responses of neuro-adaptive control with small effect of $\hat{W}_4(t)$. (a) $p_1(U)$ (. . .), $\hat{p}_1(U)$ (—); (b) $p_2(U)$ (. . .), $p_2(U)$(—); (c) $\hat{W}_4^T(t)\Gamma(t)$.

$$\Delta F(U) = \begin{bmatrix} 0.666 + 0.06u_1 + 0.02u_2 + 0.02\sin(2u_1u_2) + 0.12\sin(0.5u_1u_2) \\ 0.132 + 0.0168u_1 + 0.012u_2 + 0.15\sin(u_1u_2) \end{bmatrix}$$

only use of forward control or robust variable structure control can't achieve an excellent tracking result as the system is subject to large uncertainties. This is one of important motivation for the study. Then the control parameters of proposed control are chosen as follows: $\mu = 0.03$, $d_{211} = 0.22$, $d_{222} = 0.01$, $\xi_0 = 0.7$, $p = 11$, $b = 1$, $W_{1m} = 1$, $W_{2m} = 1$, $W_{3m} = 0.9$, $W_{4m} = 1$ (upper bound of connection weight), $\alpha_1 = 0.098$, $\alpha_2 = 0.012$, $\alpha_3 = 0.0008$, $\alpha_4 = 0.312$(learning rate), $\eta_1 = 0.66$, $\eta_2 = 0.03$, $\eta_3 = 0.04$, $\eta_4 = 0.56$ (e-modification) and $\kappa = 0.001$. The total number of weights is 198. Weights can be initialized as random numbers but small. Without loss of generality, the initial values of weight are set to zero (i.e., no compensation for excess uncertainties with respect to robust variable structure control). Although too large values of learning rate make the closed-loop system unstable, it is permitted to have large learning rate properly with suitable e-modification parameter to yield a better transient response. The output responses of neuro-adaptive control are presented in Fig. 5(a) and (b). The maximum steady-state errors for channels 1 and 2 are 0.024 and 0.018 (or 2.4% and 1.8% of the amplitude of reference input), respectively. The corresponding control inputs and sliding surfaces are shown in Fig. 5(c) and (d), respectively. It can be seen that after transient response the control input and sliding surface of proposed control are smooth. The response of Fig. 5(c) is smoother and smaller than that in Fig. 4(c). Similarly, the response of Fig. 5(d) is smaller and smoother than that in Fig. 4(d). It reveals that the convergent region of sliding surface (or tracking error) of proposed control indeed smaller that of robust control. Furthermore, the real and learning uncertainties are shown in Fig. 5(e) and (f). Because of the strong compensation of the upper bound of residue of RNN (compare Fig. 5(g) and 6(c)), the learning uncertainties $\hat{P}(U)$ do not approximate $P(U)$ perfectly. However, the controller has the ability of reducing the uncertainties affecting the system performance by the compensation of residue (i.e., $\hat{W}_4^T(t)\Gamma(t)$). Comparing the control inputs of forward control (i.e., Fig. 3(c)) and proposed control (i.e., Fig. 5(c)) gives that only the second channel is adjusted to achieve an excellent tracking performance. It indicates that the proposed control is effective. Furthermore, the responses for other control parameters (e.g., W_{im}, α_i, η_i $(1 \le i \le 4)$, d_{211}, d_{222}, μ, ξ_0, ξ_3) different from the above are similar with Fig. 5. For simplicity, those are omitted. It reveals that the range for control parameter of the proposed control is larger than that of robust control.

Due to strong compensation of the upper bound of residue of RNN (i.e., $\hat{W}_4^T(t)\Gamma(b)$), the learning uncertainties $\hat{P}(U)$ have a poor approximation of uncertainties [see Fig. 5(e) and (f)]. On the contrary, when the effect of $\hat{W}_4^T(t)\Gamma(t)$ [see Fig. 6(c)] is reduced (i.e., the initial value of weight matrix $\hat{W}_4^T(t)$ and learning rate α_4 are small), $\hat{P}(U)$ have a better approximation of uncertainties [see Fig. 6(a) and (b)]. Some control parameters chosen different from the above are $\alpha_1 = 0.036$, $\alpha_2 = 0.7$, $\alpha_3 = 0.008$, $\alpha_4 = 0.000\,28$ and $\eta_= 0.4366$, $\eta_4 = 0.1$. The maximum steady-state errors for channels 1 and 2 are 0.0605 and 0.0524 (or 6.05% and 5.24% of the amplitude of reference input), respectively. Those are a little larger than the results shown in Fig. 5(a) and (b). The corresponding control inputs and sliding surfaces are also smooth. For brevity, those are omitted.

The results reveal that as compared with the responses of forward control, robust variable structure control and the DMNAC, the proposed control indeed improves the system performance including convergent region of tracking error, the smoothness of control input and the easy selection of control parameter.

VI. Conclusions

The proposed controller includes a forward control based on a nominal system and a discrete-time multivariable neuro-adaptive variable structure control (DMNAVSC) based on an on-line approximation of huge uncertainties using a new recurrent-neural-network-with-residue-upper-bound-compensation (RNNRUBC) for a compact subset. A projection algorithm for the updating of feedback weight matrix is employed to ensure a stable RNNRUBC. No state estimator or persistent excitation is required for a DMNAVSC. Under some conditions, the semi-globally ultimately bounded tracking with the boundedness of estimated weight matrix is accomplished by Lyapunov stability theory. An appropriate use of neural-network in control application can increase benefits. The proposed scheme indeed improves the performance of some robust control schemes, e.g., the convergent region of tracking error, the smoothness of control input and the easy selection of control parameter. The results of simulation confirm that the developed theory is used for the control of a class of multivariable systems with the existence of nonlinearities and enormous uncertainties.

Acknowledgment

The authors would like to thank Prof. Y.-H. Chen in Mechanical Engineering, Georgia Institute of Technology, Atlanta, for his valuable opinions.

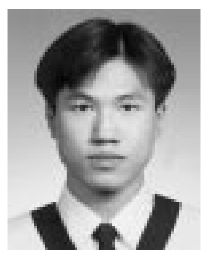
References

[1] H. K. Khalil, *Nonlinear Systems*, 2nd ed. Englewood Cliffs, NJ: Prentice-Hall, 1996.
[2] C. L. Hwang, "Neural-network-based variable structure control of electrohydraulic servosystems subject to huge uncertainties without persistent excitation," *IEEE/ASME Trans. Mechatron.*, vol. 4, no. 1, pp. 50–59, 1999.
[3] C. Banyasz, R. Haber, and L. Keviczky, "Some estimation methods for nonlinear discrete time identification," in *3rd IFAC Symp. Identification and System Parameter Estimation*, 1973, pp. 93–102.
[4] C. L. Hwang and C. J. Hsu, "Nonlinear control design for a Hammerstein model system," *Proc. Inst. Elect. Eng.*, vol. 142, pt. D, no. 4, pp. 277–285, 1995.
[5] G. Tao and P. V. Kokotovic, *Adaptive Control of Systems with Actuator and Sensor Nonlinearity*. New York: Wiley, 1996.
[6] B. M. Chen, T. H. Lee, C. C. Hang, Y. Guo, and S. Weerasooriya, "An H^∞ almost disturbance decoupling robust controller design for a piezoelectric bimorph actuator with hysteresis," *IEEE Trans. Contr. Syst. Technol.*, vol. 7, no. 2, pp. 160–174, 1999.
[7] G. C. Goodwin and K. S. Sin, *Adaptive Filtering, Prediction and Control*. Englewood Cliffs, NJ: Prentice-Hall, 1984.
[8] K. J. Astrom and B. Wittenmark, *Computer-Controlled Systems—Theory and Design*, 2nd ed. Englewood Cliffs, NJ: Prentice-Hall, 1997.
[9] P. Ge and M. Jouaneh, "Tracking control of a piezoceramic actuator," *IEEE Trans. Contr. Syst. Technol.*, vol. 4, no. 3, pp. 209–216, 1996.
[10] C. L. Hwang, M. H. Wei, and W. J. Jieng, "Noncircular cutting with lathe using three-stage intelligent controller," *J. Robotics & Computer-Integrated Manufact.*, vol. 13, no. 3, pp. 181–191, 1997.
[11] K. S. Narendra and K. Parthasarathy, "Identification and control of dynamical systems using neural networks," *IEEE Trans. Neural Networks*, vol. 1, no. 1, pp. 4–27, 1990.

[12] F. C. Chen and H. K. Khalil, “Adaptive control of a class of nonlinear discrete-time systems using neural networks,” *IEEE Trans. Automat. Contr.*, vol. 40, no. 5, pp. 791–801, 1995.
[13] S. Jagannathan and F. L. Lewis, “Multilayer discrete-time neural-net controller with guaranteed performance,” *IEEE Trans. Neural Networks*, vol. 7, no. 1, pp. 107–130, 1996.
[14] L. Jin, P. N. Nikiforuk, and M. M. Gupta, “Absolute stability conditions for discrete-time recurrent neural networks,” *IEEE Trans. Neural Networks*, vol. 5, no. 6, pp. 954–964, 1994.
[15] C. C. Ku and K. Y. Lee, “Diagonal recurrent neural networks for dynamic systems control,” *IEEE Trans. Neural Networks*, vol. 6, no. 1, pp. 144–156, 1995.
[16] M. K. Sundareshan and T. A. Condarcure, “Recurrent neural-network training by a learning automation approach for trajectory learning and control system design,” *IEEE Trans. Neural Networks*, vol. 9, no. 3, pp. 354–368, 1998.
[17] G. A. Rovithakis, “Tracking control of multi-input affine nonlinear dynamica systems with unknown nonlinearities using dynamical neural network,” *IEEE Trans. Syst., Man, Cybern. B*, vol. 29, no. 2, pp. 179–189, 1999.
[18] Y. Fang, T. W. S. Chow, and X. D. Li, “Use of a recurrent neural network in discrete sliding-mode control,” *Proc. Inst. Elect. Eng.*, vol. 146, pt. D, no. 1, pp. 84–90, 1999.
[19] H. K. Khalil, “Adaptive output feedback control of nonlinear systems represented by input-output models,” *IEEE Trans. Automat. Contr.*, vol. 41, no. 2, pp. 177–188, 1996.
[20] G. A. Rovithakis and M. A. Christodoulou, “Neural adaptive regulation of unknown nonlinear dynamical systems,” *IEEE Trans. Syst., Man, Cybern. B*, vol. 27, no. 5, pp. 810–822, 1997.
[21] K. Furuta, “VSS type self-tuning control,” *IEEE Trans. Ind. Electron.*, vol. 40, no. 1, pp. 37–44, 1993.
[22] P. Korondi, H. Hashimoto, and V. Utkin, “Direct control of flexible shaft in an observer-based discrete-time sliding mode,” *IEEE Trans. Ind. Electron.*, vol. 45, no. 2, pp. 291–296, 1998.
[23] K. D. Young, V. I. Utkin, and Ümit Özgüner, “A control engineer’s guide to sliding mode control,” *IEEE Trans. Contr. Syst. Technol.*, vol. 7, no. 3, pp. 328–342, 1999.
[24] S. Hui and S. H. Zak, “On discrete-time variable structure sliding mode control,” *Syst. Control Lett.*, vol. 38, pp. 283–288, 1999.
[25] M. Vidyasagar, *Control System Synthesis—A Factorization Approach*. Cambridge, MA: MIT Press, 1985.
[26] M. Zhihong, H. R. Wu, and M. Palaniswami, “An adaptive tracking controller using neural networks for a class of nonlinear systems,” *IEEE Trans. Neural Networks*, vol. 9, no. 5, pp. 947–955, 1998.
[27] F. Sun, Z. Sun, and P. Y. Woo, “S neural-network-based adaptive control for sampled-data nonlinear systems,” *IEEE Trans. Neural Networks*, vol. 9, no. 5, pp. 956–968, 1998.

Chih-Lyang Hwang (M’90) received the B.E. degree in aeronautical engineering from Tamkang University, Taiwan, R.O.C., in 1981, and the M.E. and Ph.D. degrees in mechanical engineering from Tatung Institute of Technology, Taiwan, in 1986 and 1990, respectively.

Since 1990, he has been with the Department of Mechanical Engineering, Tatung Institute of Technology, where he is engaged in teaching and research in the area of servocontrol and control of manufacturing systems. Since 1996, he has been an referee of patent for National Standard Bureau in Ministry of Economy of Taiwan. Since 1996, he has been a Professor of mechanical engineering at Tatung Institute of Technology. He is the author or coauthor of about 25 journal papers in the related field. His current research interests include neural-network modeling and control, variable structure control, fuzzy control, mechatronics and robotics. In 1998 and 1999, he was a Research Scholar of George W. Woodruf School of Mechanical Engineering of Georgia Institute of Technology.



Ching-Hung Lin was born in Taipei, Taiwan, R.O.C., in 1975. He received the B.E. degree in mechanical engineering from National Central University, Chungli, Taiwan, in 1997, and the M.E. degree in mechanical engineering from Tatung Institute of Technology, Taipei, in 1999. He is currently pursuing the Ph.D. degree at National Taiwan University, Taipei. His research interests include adaptive and learning systems, neural networks, fuzzy system, and their application to control problems.