Adaptive Neural Control of Uncertain MIMO Nonlinear Systems

Shuzhi Sam Ge, *Senior Member, IEEE,* and Cong Wang, *Member, IEEE*

***Abstract*—In this paper, adaptive neural control schemes are proposed for two classes of uncertain multi-input/multi-output (MIMO) nonlinear systems in block-triangular forms. The MIMO systems consist of interconnected subsystems, with couplings in the forms of unknown nonlinearities and/or parametric uncertainties in the input matrices, as well as in the system interconnections without any bounding restrictions. Using the block-triangular structure properties, the stability analyses of the closed-loop MIMO systems are shown in a nested iterative manner for all the states. By exploiting the special properties of the affine terms of the two classes of MIMO systems, the developed neural control schemes avoid the controller singularity problem completely without using projection algorithms. Semiglobal uniform ultimate boundedness (SGUUB) of all the signals in the closed-loop of MIMO nonlinear systems is achieved. The outputs of the systems are proven to converge to a small neighborhood of the desired trajectories. The control performance of the closed-loop system is guaranteed by suitably choosing the design parameters. The proposed schemes offer systematic design procedures for the control of the two classes of uncertain MIMO nonlinear systems. Simulation results are presented to show the effectiveness of the approach.**

***Index Terms*—Adaptive neural control, backstepping, block-triangular form, multi-input/multi-output (MIMO) nonlinear systems, neural networks (NN).**

I. INTRODUCTION

In practice, most practical systems considered are nonlinear and multivariable in character. It is of certainty that the control theory for nonlinear multivariable systems will find immediate and wide applications. For multi-input/multi-output (MIMO) nonlinear systems, the control problem is very complicated due to the couplings among various inputs and outputs. It becomes in general very difficult to deal with when there exist uncertain parameters and/or unknown nonlinear functions in the input coupling matrix. Due to these difficulties, it is noticed that in comparison with the vast amount of results on controller design for single-input/single-output (SISO) nonlinear systems in the control literature, there are relatively fewer results available for the broader class of MIMO nonlinear systems. Based on feedback linearization [36], several adaptive control schemes have been proposed for certain classes of MIMO nonlinear systems with parametric uncertainties in the input coupling matrix (see, e.g., [15], [31], [32]). In these schemes, to remove the couplings of system inputs, an estimate of "the decoupling matrix" is usually needed and required to be invertible during parameter adaptation. Therefore, additional efforts have to be made to avoid the possible singularity problem when calculating the inverse of the estimated "decoupling matrix," e.g., by using projection algorithm [31] to keep the estimated parameters inside a feasible set in which the singularity problem does not happen. The disadvantage of using projection is that, it usually requires *a priori* knowledge for the feasible parameter set and no systematic procedure is available for constructing such a set for a general plant [34].

Since the 1990s, backstepping [7], [9] has become one of the most popular design methods for a large class of SISO nonlinear systems. The advantages of backstepping methodology include that: i) global stability can be achieved with ease; ii) the transient performance can be guaranteed and explicitly analyzed; and iii) it has the flexibility to avoid cancellations of useful nonlinearities compared with the feedback linearization techniques [36]. It is noticed that in comparison with the large amount of work on backstepping design for SISO nonlinear triangular systems, only a few results (e.g., [9], [27]–[30], etc.) are available in applying backstepping design to uncertain MIMO nonlinear systems. This is mainly due to the difficulties in coping with the uncertainties in the input and output coupling matrices. As indicated in [29], sometimes even the presentation of MIMO forms in a meaningful manner becomes a difficult task. Backstepping adaptive control is presented in [9] for MIMO nonlinear systems in a parametric strict feedback form but with no parametric uncertainties in the input matrix. In [27], backstepping design was extended to multiinput nonlinear systems in the generalized normal form. Linear high-gain control is investigated for semiglobal robust stabilization of MIMO nonlinear system, which is more general than the generalized normal form in [30]. However, it is well known that high gain control is undesirable in practice, because it may excite the unmodeled dynamics and destroy the stability of the closed-loop system, and it may not even be achievable by actuation [24]. Recently, adaptive robust control schemes were proposed for MIMO nonlinear systems in semistrict feedback forms with known input matrix [29]. In all of these works, the system interconnections are either known functions [27], [28], [30], or bounded by known nonlinear functions [29], and there is no unknown nonlinear function in the input coupling matrices. Moreover, for the case when unknown nonlinear functions exist in both the input coupling matrices and the system interconnections (with no bounding restrictions), very few results are available in the literature. These results indicate that the control of coupled uncertain MIMO nonlinear systems remains to be a difficult problem, and is still open in the control area.

Manuscript received February 14, 2001; revised July 9, 2002 and December 20, 2003.

S. S. Ge is with the Department of Electrical and Computer Engineering, National University of Singapore, Singapore 117576 Republic of Singapore (e-mail: elegesz@nus.edu.sg).

C. Wang is with the College of Automation, South China University of Technology.

Digital Object Identifier 10.1109/TNN.2004.826130

As an alternative, following the pioneering works [1], [10], [11] on controlling nonlinear dynamical systems using neural networks, there has been tremendous interest in the study of adaptive neural control of uncertain nonlinear systems with unknown nonlinearities, and a great deal of progress has been made both in theory and practical applications, see, e.g., [2], [3], [5], [6] and the references therein for a survey of recent development.

In the literature of adaptive neural control, neural networks are mostly used as approximation models for the unknown nonlinearities due to their inherent approximation capabilities. With the help of neural networks (NN) approximation, it is not necessary to spend much effort on system modeling which might be very difficult in some cases. In the earlier neural control schemes, optimization techniques were mainly used to derive parameter adaptation laws, and the feasibility of such neural control schemes were demonstrated via numerous empirical studies with little analytical results for stability and performance. To overcome these problems, some elegant adaptive neural control approaches based on Lyapunov's stability theory have been proposed for nonlinear systems with certain types of matching conditions [10]–[19], as well as nonlinear triangular systems without the requirement of matching conditions [20]–[24]. The advantage of these approaches is that stability of the closed-loop system is guaranteed, and the performance and robustness properties are readily determined. Semiglobal uniform ultimate boundedness of all the signals in the closed-loop is achieved and the output of the system is proven to converge to a small neighborhood of the desired trajectory.

The aforementioned works demonstrate that adaptive neural control is particularly suitable for controlling highly uncertain, nonlinear, and complex systems. However, similar to the situation in the control literature without using neural networks or any other universal function approximators, there are only a few results available in the literature on adaptive neural control of nonlinear MIMO systems (e.g., [3], [4], [25]). In [3], [4], adaptive neural network controllers were proposed for special classes of MIMO nonlinear robotic systems, using several nice properties of the robotic systems. Energy-type Lyapunov functions are chosen to develop the stable neural controller and adaptation laws. In [25], an adaptive neural control approach was proposed for a class of MIMO nonlinear systems with a triangular structure in control inputs. By using the triangular property, integral-type Lyapunov functions are introduced to construct a Lyapunov-based control structure, which does not try to cancel the "decoupling matrix" to approximately linearize the system. However, due to the integral operation, the obtained controller is complicated and difficult to use in practice.

In this paper, we consider adaptive neural control of uncertain MIMO nonlinear systems in block-triangular forms. The purpose of this paper is to investigate how far adaptive neural control can achieve for MIMO nonlinear systems. First, we present two classes of uncertain MIMO nonlinear systems in block-triangular forms. Denoted as Σ_{M_1} and Σ_{M_2} (as will be detailed in Sections II and IV, respectively), the MIMO systems considered are composed of interconnected subsystems, with couplings in the forms of unknown nonlinearities in the input matrices, as well as in the system interconnections without any bounding restrictions. Such MIMO systems cannot be controlled by the adaptive or robust control schemes in the classical control literature.

Compared with the MIMO system considered in [25], the block-triangular MIMO systems in this paper are more general in system state interconnections. Specifically, system interconnections in [25] only appear in the last equation of each subsystem, while in this paper, they appear in every equation of each subsystem. The more general form makes it difficult to conclude the stability of the whole system by stability analysis of individual subsystem separately. Tto conduct stability analysis for the whole closed-loop system, we make full use of the block-triangular structure properties. We first design for each subsystem a full state feedback controller, and then conclude the stability of the entire system states in a nested iterative manner. In other words, since the state variables of one subsystem may be embedded in another subsystem, and no bounding restrictions are imposed on the uncertain system interconnections, the proof of the stability analysis does not follow immediately after the controllers have been designed as in the standard backstepping design. We can only prove the stability by following a specific order from one state variable in a subsystem to another state variable in another subsystem. With the deployment of NN approximation, the uncertain MIMO systems, with complicated couplings and unknown nonlinearities, can be controlled using adaptive neural design. By exploiting the special properties of the affine terms of the two classes of MIMO systems Σ_{M_1} and Σ_{M_2}, the developed schemes avoid the controller singularity problem completely without using projection algorithms. The developed scheme achieves semiglobal uniform ultimate boundedness of all the signals in the closed-loop of the MIMO systems. The outputs of the system are proven to converge to a small neighborhood of the desired trajectories. The control performance of the closed-loop system is guaranteed by suitably choosing the design parameters.

The rest of the paper is organized as follows: Section II describes the first class of uncertain MIMO systems Σ_1 and the control problem. In Section III, an adaptive neural control scheme is presented for uncertain MIMO nonlinear system Σ_1. For the second class of partially known MIMO nonlinear system Σ_2, which contains both unknown nonlinear functions and constant parametric uncertainties, another adaptive neural control scheme is proposed in Section IV. Simulation results are performed to demonstrate the effectiveness of the approach in Section V. Section VI contains the conclusions.

Throughout this paper, the following notations are used.

- $\|\cdot\|$ stands for Euclidean norm of vectors and induced norm of matrices.
- $\lambda_{\max}(B)$ and $\lambda_{\min}(B)$ denote the largest and smallest eigenvalues of a square matrix B, respectively.
- W^* and $\hat{W}$ denote the ideal neural weights, respectively, and the estimates of neural weights, $\tilde{W} \triangleq \hat{W} - W^*$ denotes the error between $\hat{W}$ and W^*.
- i and j are integer indices, ρ_j denotes the order of the jth subsystem, $\varrho_{jl} = \rho_j - \rho_l$ is the order difference between the jth and lth subsystems, and i_j denotes the subscription

of the i_jth component of the corresponding items in the jth subsystem.

- x_{j,i_j} denotes the i_jth state of the jth subsystem, $\bar{x}_{j,i_j} = [x_{j,1},\ldots,x_{j,i_j}]^T \in R^{i_j}, i_j = 1,\ldots,\rho_j$ and $j = 1,\ldots,m$ denotes the vector of partial state variables in the jth subsystem as defined, and $X = [\bar{x}^T_{1,\rho_1},\ldots,\bar{x}^T_{m,\rho_m}]^T$ denotes the state variables of the complete system.
- u_j and y_j are the input and output of the jth subsystem.

For the convenience of stability analysis, let us present the the following technical lemmas of bounded-input and bounded output property for stable dynamic inequalities.

Lemma 1.1: Let function $V(t) \geq 0$ be a continuous function defined $\forall\, t \in R^+$ and $V(0)$ bounded, and $\rho(t) \in L_\infty$ be real-valued function. If the following inequality holds:

$$\dot{V}(t) \leq -c_1 V(t) + c_2\rho(t) \tag{1}$$

where $c_1 > 0$, c_2 are constants, then we can conclude that $V(t)$ is bounded.

Proof: Multiplying both side by $e^{c_1 t}$, (1) becomes

$$\frac{d}{dt}(V(t)e^{c_1 t}) \leq c_2\rho(t)e^{c_1 t}. \tag{2}$$

Integrating it over $[0,t]$, we have

$$V(t) \leq V(0)e^{-c_1 t} + c_2\int_0^t e^{-c_1(t-\tau)}\rho(\tau)d\tau. \tag{3}$$

By noting the following inequality:

$$\begin{aligned} c_2\int_0^t e^{-c_1(t-\tau)}\rho(\tau)d\tau &\leq c_2 e^{-c_1 t}\int_0^t |\rho(\tau)|e^{c_1\tau}d\tau \\ &\leq c_2 e^{-c_1 t}\sup_{\tau\in[0,t]}[|\rho(\tau)|]\int_0^t e^{c_1\tau}d\tau \\ &\leq \frac{c_2}{c_1}\sup_{\tau\in[0,t]}[|\rho(\tau)|](1-e^{-c_1 t}) \\ &\leq \frac{c_2}{c_1}\sup_{\tau\in[0,t]}[|\rho(\tau)|] \end{aligned} \tag{4}$$

we know that if $\rho(t)$ is bounded, then $c_2\int_0^t e^{-c_1(t-\tau)}\rho(\tau)d\tau$ is bounded. Let $c_0 = (c_2/c_1)\sup_{\tau\in[0,t]}[|\rho(\tau)|]$, then (3) becomes

$$V(t) \leq c_0 + V(0)e^{-c_1 t} \leq c_0 + V(0). \tag{5}$$

We can readily conclude the boundedness of $V(t)$. ◇

Remark 1.1: Apparently, $V(t)$ might be larger than $V(0)$, but it will never explode up as it is bounded by $(c_2/c_1)\sup_{\tau\in[0,t]}[|\rho(\tau)|] + V(0)$. The actual maximum size of the bound is a function of its initial value $V(0)$, the stable coefficient c_1, the input gain c_2, and the actual input $\rho(t)$. This result can be regarded as the bounded-input–bounded-output property of stable dynamic inequality (1).

Corollary 1.1: Let function $V(t) \geq 0$ be a continuous function defined $\forall\, t \in R^+$ and $V(0)$ bounded. If the following inequality holds:

$$\dot{V}(t) \leq -c_1 V(t) + c_2 \tag{6}$$

where $c_1 > 0$, c_2 are constants, then we can conclude that $V(t)$ is bounded.

Lemma 1.2: Consider the positive function given by

$$V(t) = \frac{1}{2}e^T(t)Q(t)e(t) + \frac{1}{2}\tilde{W}^T(t)\Gamma^{-1}(t)\tilde{W}(t) \tag{7}$$

where $e(t) = x(t) - x_d(t)$ and $\tilde{W}(t) = \hat{W}(t) - W^*$ with $x(t) \in R^n$, $x_d(t) \in \Omega_d \subset R^n$, $\hat{W}(t) \in R^m$, and constants $W^* \in R^m$, $Q(t) = Q^T(t) > 0$ and $\Gamma(t) = \Gamma^T(t) > 0$ are dimensionally compatible matrices. If the following inequality holds:

$$\dot{V}(t) \leq -c_1 V(t) + c_2 \tag{8}$$

then, given any initial compact set defined by

$$\Omega_0 = \left\{x(0), x_d(0), \hat{W}(0) \mid x(0), \hat{W}(0) \text{ finite }, x_d(0) \in \Omega_d\right\} \tag{9}$$

we can conclude that

i) the states and weights in the closed-loop system will remain in the compact set defined by

$$\Omega = \left\{x(t), \hat{W}(t)\big|\|x(t)\| \leq c_{e\max} + \max_{\tau\in[0,t]}\{\|x_d(\tau)\|\},\; x_d(t) \in \Omega_d, \|\hat{W}\| \leq c_{\tilde{W}\max} + \|W^*\|\right\}$$

ii) the states and weights will eventually converge to the compact sets defined by

$$\Omega_s = \left\{x(t), \hat{W}(t)\big|\lim_{t\to\infty}\|e(t)\| = \mu_e^*, \lim_{t\to\infty}\|\tilde{W}\| = \mu_{\tilde{W}}^*\right\} \tag{10}$$

where constants

$$c_{e\max} = \sqrt{\frac{2V(0) + \frac{2c_2}{c_1}}{\lambda_{Q\min}}} \qquad c_{\tilde{W}\max} = \sqrt{\frac{2V(0) + \frac{2c_2}{c_1}}{\lambda_{\Gamma\min}}} \tag{11}$$

$$\mu_e^* = \sqrt{\frac{2c_2}{c_1\lambda_{Q\min}}} \qquad \mu_{\tilde{W}}^* = \sqrt{\frac{2c_2}{c_1\lambda_{\Gamma\min}}} \tag{12}$$

with $\lambda_{Q\min} = \min_{\tau\in[0,t]}\lambda_{\min}(Q(\tau))$, and $\lambda_{\Gamma\min} = \min_{\tau\in[0,t]}\lambda_{\min}(\Gamma^{-1}(\tau))$.

Proof: In fact, the results of this Lemma is quite apparent by noticing the results of Lemma 1.1 or the Corollary 1.1 and the particular definition of $V(t)$ in (7). For completeness, it is shown in details below. Multiplying (8) by $e^{c_1 t}$ yields

$$\frac{d}{dt}(V(t)e^{c_1 t}) \leq c_2 e^{c_1 t}. \tag{13}$$

Integrating (13) over $[0,t]$ leads to

$$0 \leq V(t) \leq \left[V(0) - \frac{c_2}{c_1}\right]e^{-c_1 t} + \frac{c_2}{c_1} \tag{14}$$

where $V(0) = (1/2)e^T(0)Q(0)e(0) + (1/2)\tilde{W}^T(0)\Gamma^{-1}(0)\tilde{W}(0)$.

i) Uniform Boundedness (UB)

From (14), we have

$$0 \le V(t) \le \left[V(0) - \frac{c_2}{c_1}\right] e^{-c_1 t} + \frac{c_2}{c_1} \le V(0) + \frac{c_2}{c_1}. \quad (15)$$

From (7), we have

$$\begin{aligned}\frac{1}{2}\lambda_{Q\min}\|e(t)\|^2 &\le \frac{1}{2}\lambda_{\min}(Q(t))\|e(t)\|^2 \\ &\le \frac{1}{2}e^T(t)Q(t)e(t) \\ &\le V(t) \quad (16)\end{aligned}$$

$$\begin{aligned}\frac{1}{2}\lambda_{\Gamma\min}\|\tilde{W}(t)\|^2 &\le \frac{1}{2}\lambda_{\min}(\Gamma^{-1}(t))\|\tilde{W}(t)\|^2 \\ &\le \frac{1}{2}\tilde{W}^T(t)\Gamma^{-1}(t)\tilde{W}(t) \\ &\le V(t) \quad (17)\end{aligned}$$

then, by combining with (15), we have

$$\|e(t)\| \le c_{e\max}, \quad \|\tilde{W}(t)\| \le c_{\tilde{W}\max} \quad (18)$$

where $c_{e\max}$ and $c_{\tilde{W}\max}$ are given in (11). Since $e(t) = x(t) - x_d(t)$ and $\tilde{W}(t) = \hat{W}(t) - W^*$, we have

$$\|x(t)\| - \|x_d(t)\| \le \|x(t) - x_d(t)\| \le c_{e\max} \quad (19)$$

$$\|\hat{W}(t)\| - \|W^*\| \le \|\hat{W}(t) - W^*\| \le c_{\tilde{W}\max} \quad (20)$$

i.e.,

$$\begin{aligned}\|x(t)\| &\le c_{e\max} + \|x_d(t)\| \le c_{e\max} \\ &\quad + \max_{\tau\in[0,t]}\{\|x_d(\tau)\|\} \quad (21)\end{aligned}$$

$$\|\hat{W}(t)\| \le c_{\tilde{W}\max} + \|W^*\| \quad (22)$$

ii) Uniformly Ultimate Boundedness (UUB)

From (14), (16) and (17), we have

$$\|e(t)\| \le \frac{\sqrt{2\left[V(0) - \frac{c_2}{c_1}\right]e^{-c_1 t} + \frac{2c_2}{c_1}}}{\lambda_{Q\min}} \quad (23)$$

$$\|\tilde{W}(t)\| \le \frac{\sqrt{2\left[V(0) - \frac{c_2}{c_1}\right]e^{-c_1 t} + \frac{2c_2}{c_1}}}{\lambda_{\Gamma\min}}. \quad (24)$$

If it so happens that $V(0) = c_2/c_1$, then $\|e(t)\| \le \mu_e^*, \ \forall\, t \ge 0$.

If $V(0) \ne c_2/c_1$, from (23), we can conclude that given any $\mu_e > \mu_e^*$, there exists T_e, such that for any $t > T_e$, we have $\|e(t)\| \le \mu_e$. Specifically, given any μ_e

$$\mu_e = \sqrt{\frac{2\left[V(0) - \frac{c_2}{c_1}\right]e^{-c_1 T_e} + \frac{2c_2}{c_1}}{\lambda_{Q\min}}}, \quad V(0) \ne \frac{c_2}{c_1} \quad (25)$$

then

$$T_e = -\frac{1}{c_1}\ln\left(\frac{\mu_e^2\lambda_{Q\min} - \frac{2c_2}{c_1}}{2\left[V(0) - \frac{c_2}{c_1}\right]}\right) \quad (26)$$

and

$$\lim_{t\to\infty}\|e(t)\| = \mu_e^* \quad (27)$$

where μ_e^* is defined in (12).

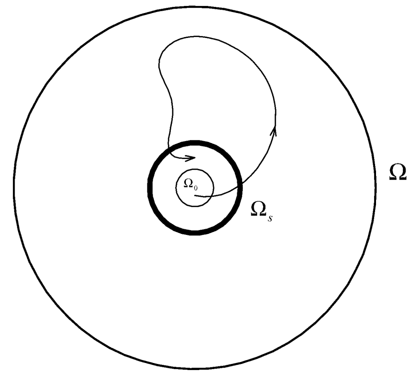

Fig. 1. Compact sets in Lemma 1.2.

Similar conclusion can be made about $\|\tilde{W}\|$ and $\lim_{t\to\infty}\|\tilde{W}\| = \mu_{\tilde{W}}^*$ as defined in (12). ◇

Remark 1.2: In Lemma 1.2, there are three compact sets: initial compact set, Ω_0, bounding compact setsΩ, and the steady state compact set Ω_s. The relationship among the three compact sets is illustrated in Fig. 1. From the detailed analysis, we know that: i) the size of Ω_0 affacts Ω, but not Ω_s, and ii) Ω_s (i.e., the steady state errors e and $\tilde{W}$) can be made smaller by changing the appropriate parameters c_1, c_2, $Q(t)$, and $\Gamma^{-1}(t)$.

Remark 1.3: Lemma 1.2 gives an explicit theoretical explanation of approximation based control techniques in the literature. Give any initial set of control parameters $e(0), \hat{W}(0) \in \Omega_0$, if inequality (8) holds, then we know that $x(t)$ and $\hat{W}(t)$ are bounded in Ω. Accordingly, we can construct a stable adaptive neural network controller easily if the neural network is chosen large enough to cover Ω for bounded initial conditions. The nice thing is that the larger the Ω is, the larger the Ω_0 is. As the actual sizes of Ω_0 and Ω are not specified in advance, they can really be made as large as deemed necessary in practical applications for any given size of initial condition. Theoretically speaking, it follows the definition of SGUUB in the sense that bounded initial conditions guarantee the boundedness of all the signals in the closed-loop system provided the neural network is chosen to cover a compact set of sufficiently large size. For clarity, it will not be repeated again and again in the paper, but is understood as such.

II. Problem Formulation and Preliminaries

Here, we describe the first class of uncertain MIMO nonlinear system in the following block-triangular form as shown in (28) at the bottom of the next page, where x_{j,i_j}, $i_j = 1, \ldots, \rho_j$ are the states of the jth subsystem, $u_j \in R$ and $y_j \in R$ are the input and output of the jth subsystem, respectively; $f_{j,i_j}(\cdot)$ and $g_{j,i_j}(\cdot)$ ($i_j = 1, \ldots, \rho_j$ and $j = 1, \ldots, m$) are unknown nonlinear smooth functions, j, i_j, ρ_j and m are positive integers, and ϱ_{jl} is defined as

$$\varrho_{jl} \triangleq \rho_j - \rho_l.$$

Remark 2.1: Note that ϱ_{jl} is the order difference between the jth and lth subsystems. The introduction of the notation ϱ_{jl} ($j, l = 1, \ldots, m$) is very important in analyzing the MIMO system (28). To help readers to understand the notation better, let us consider the following two cases:

i) When $j = l$, we have $\varrho_{jj} = 0$. Accordingly, $\bar{x}_{j,(i_j-\varrho_{jj})} = \bar{x}_{j,i_j}$, which is exactly the state variables of the jth subsystem.

ii) When $j \neq l$, we have two situations to consider. If $i_j - \varrho_{jl} \leq 0$, then the corresponding variable vector $\bar{x}_{l,(i_j-\varrho_{jl})}$ does not exist, and does not appear in the functions in (28). If $i_j - \varrho_{jl} > 0$, then $\bar{x}_{l,(i_j-\varrho_{jl})}$ represents the maximum state variables of the lth subsystem which are embedded in the jth subsystem.

Remark 2.2: As examples, consider the following two block-triangular MIMO systems

$$\Sigma_{S_1}: \begin{cases} \dot{x}_{1,1} = f_{1,1}(\bar{x}_{1,1}, \bar{x}_{2,1}) + g_{1,1}(\bar{x}_{1,1}, \bar{x}_{2,1})x_{1,2} \\ \dot{x}_{1,2} = f_{1,2}(X) + g_{1,2}(\bar{x}_{1,1}, \bar{x}_{2,1})u_1 \\ \dot{x}_{2,1} = f_{2,1}(\bar{x}_{1,1}, \bar{x}_{2,1}) + g_{2,1}(\bar{x}_{1,1}, \bar{x}_{2,1})x_{2,2} \\ \dot{x}_{2,2} = f_{2,2}(X, u_1) + g_{2,2}(\bar{x}_{1,1}, \bar{x}_{2,1})u_2 \\ y_j = x_{j,1}, \quad j = 1, 2 \end{cases} \tag{29}$$

where $X = \left[\bar{x}_{1,2}^T, \bar{x}_{2,2}^T\right]^T$ with $\bar{x}_{j,2} = [x_{j,1}, x_{j,2}]^T$, $j = 1, 2$, and

$$\Sigma_{S_2}: \begin{cases} \dot{x}_{1,1} = f_{1,1}(\bar{x}_{1,1}, \bar{x}_{2,3}) + g_{1,1}(\bar{x}_{1,1}, \bar{x}_{2,3})x_{1,2} \\ \dot{x}_{1,2} = f_{1,2}(X) + g_{1,2}(\bar{x}_{1,1}, \bar{x}_{2,3})u_1 \\ \dot{x}_{2,1} = f_{2,1}(\bar{x}_{2,1}) + g_{2,1}(\bar{x}_{2,1})x_{2,2} \\ \dot{x}_{2,2} = f_{2,2}(\bar{x}_{2,2}) + g_{2,2}(\bar{x}_{2,2})x_{2,3} \\ \dot{x}_{2,3} = f_{2,3}(\bar{x}_{1,1}, \bar{x}_{2,3}) + g_{2,2}(\bar{x}_{1,1}, \bar{x}_{2,3})x_{2,4} \\ \dot{x}_{2,4} = f_{2,4}(X, u_1) + g_{2,4}(\bar{x}_{1,1}, \bar{x}_{2,3})u_2 \\ y_j = x_{j,1}, \quad j = 1, 2 \end{cases} \tag{30}$$

where $\bar{x}_{j,i_j} = [x_{j,1}, \ldots, x_{j,i_j}]^T$, $j = 1, 2$, $i_1 = 1, 2$, $i_2 = 1, \ldots, 4$, and $X = \left[\bar{x}_{1,2}^T, \bar{x}_{2,4}^T\right]^T$.

System Σ_{S_1} shows the case when the orders of the subsystems are the same ($\rho_1 = \rho_2 = 2$), while system Σ_{S_2} demonstrates the situation when the orders of the subsystems are different ($\rho_1 = 2$, $\rho_2 = 4$). These two cases show how the state variables of one subsystem are embedded into the other subsystem according to the values of $i_j - \varrho_{jl}$ as stated above in Remark 1.

Compared with the MIMO system considered in [25], where the system interconnections are only limited to $f_{j,\rho_j}(X)$, the above block-triangular MIMO system (28) has the most system interconnections that the approach, which we are going to present, can handle due to the introduction of $\bar{x}_{l,(i_j-\varrho_{jl})}$. The system interconnections are represented by terms $f_{j,i_j}\left(\bar{x}_{1,(i_j-\varrho_{j1})}, \ldots, \bar{x}_{m,(i_j-\varrho_{jm})}\right)$ and $g_{j,i_j}\left(\bar{x}_{1,(i_j-\varrho_{j1})}, \ldots, \bar{x}_{m,(i_j-\varrho_{jm})}\right)$, $i_j = 1, \ldots, \rho_j - 1$, $j = 1, \ldots, m$, as well as by terms $f_{j,\rho_j}(X, u_1, \ldots, u_{j-1})$ and $g_{j,\rho_j}(\bar{x}_{1,\rho_1-1}, \ldots, \bar{x}_{m,\rho_m-1})$. Note that $f_{j,\rho_j}(X, u_1, \ldots, u_{j-1})(j = 1, \ldots, m)$ are very general in the sense that they include not only the system states X, but also the system inputs $u_1, \ldots, u_{j-1}$, where $u_1, \ldots, u_{j-1}$ are the control signals of the $(1, \ldots, j-1)$th subsystems directly passing through to the jth subsystem.

Notice that the affine terms $g_{j,\rho_j}(\bar{x}_{1,\rho_1-1}, \ldots, \bar{x}_{j,\rho_m-1})$, $j = 1, \ldots, m$ are assumed to be independent of the states $x_{1,\rho_1}, \ldots, x_{j,\rho_j}$. By utilizing this special structure property, which can be found in many practical systems (see, e.g., [5] and the references therein), the controller singularity problem is avoided, and the stability of the resulting adaptive system is guaranteed without the requirement for the integral-type Lyapunov function [25].

The control objective is to design adaptive neural controller for system (28) such that: i) all the signals in the closed-loop remain semiglobally uniformly ultimately bounded and ii) the output y_j follows the desired trajectories y_{dj} generated from the following smooth, bounded reference model

$$\begin{aligned} \dot{x}_{di} &= f_{di}(x_d), \quad 1 \leq i \leq n \\ y_{dj} &= x_{dj}, \quad 1 \leq j \leq m \leq n \end{aligned} \tag{31}$$

where $x_d = [x_{d1}, x_{d2}, \ldots, x_{dm}]^T \in R^m$ are the states, $y_{dj} \in R$, $1 \leq j \leq m \leq n$, are the system outputs, $f_{di}(\cdot)$, $i = 1, 2, \ldots, m$ are known smooth nonlinear functions.

In control engineering, radial basis function (RBF) NNs are usually used as a tool for modeling nonlinear functions because of their good capabilities in function approximation [35]. In this paper, the following RBF NN [16] is used to approximate the continuous function $h(Z): R^q \to R$

$$h_{nn}(Z) = W^T S(Z) \tag{32}$$

where the input vector $Z \in \Omega \subset R^q$, weight vector $W = [w_1, w_2, \ldots, w_l]^T \in R^l$, the NN node number $l > 1$; and $S(Z) = [s_1(Z), \ldots, s_l(Z)]^T$ with

$$s_i(Z) = \exp\left[\frac{-(Z - \mu_i)^T (Z - \mu_i)}{\eta_i^2}\right], \quad i = 1, 2, \ldots, l \tag{33}$$

where $\mu_i = [\mu_{i1}, \mu_{i2}, \ldots, \mu_{iq}]^T$ is the center of the receptive field and η_i is the width of the Gaussian function.

It has been proven that network (32) can approximate any smooth function over a compact set $\Omega_Z \subset R^q$ to arbitrarily any accuracy as

$$h(Z) = W^{*T} S(Z) + \epsilon, \quad \forall\, Z \in \Omega_Z \tag{34}$$

where W^* is ideal constant weights vector, and ϵ is the approximation error. The stability results obtained in NN control literature are semiglobal in the sense that, as long as the input variables Z of the NNs remain within some prefixed compact set $\Omega_Z \subset R^q$ where the compact set Ω_Z can be made as large as desired, there exists controller(s) with sufficiently large number of NN nodes such that all the signals in the closed-loop remain bounded.

$$\Sigma_1: \begin{cases} \dot{x}_{1,i_1} = f_{1,i_1}\left(\bar{x}_{1,(i_1-\varrho_{11})}, \ldots, \bar{x}_{m,(i_1-\varrho_{1m})}\right) + g_{1,i_1}\left(\bar{x}_{1,(i_1-\varrho_{11})}, \ldots, \bar{x}_{m,(i_1-\varrho_{1m})}\right) x_{1,i_1+1}, & 1 \leq i_1 \leq \rho_1 - 1 \\ \dot{x}_{1,\rho_1} = f_{1,\rho_1}(X) + g_{1,\rho_1}\left(\bar{x}_{1,\rho_1-1}, \ldots, \bar{x}_{m,\rho_m-1}\right) u_1 \\ \cdots \\ \dot{x}_{j,i_j} = f_{j,i_j}\left(\bar{x}_{1,(i_j-\varrho_{j1})}, \ldots, \bar{x}_{m,(i_j-\varrho_{jm})}\right) + g_{j,i_j}\left(\bar{x}_{1,(i_j-\varrho_{j1})}, \ldots, \bar{x}_{m,(i_j-\varrho_{jm})}\right) x_{j,i_j+1}, & 1 \leq i_j \leq \rho_j - 1 \\ \dot{x}_{j,\rho_j} = f_{j,\rho_j}(X, u_1, \ldots, u_{j-1}) + g_{j,\rho_j}\left(\bar{x}_{1,\rho_1-1}, \ldots, \bar{x}_{m,\rho_m-1}\right) u_j \\ y_j = x_{j,1}, & 1 \leq j \leq m \end{cases} \tag{28}$$

Assumption 2.1: There exist ideal constant weights W^* such that $|\epsilon| \le \epsilon^*$ with constant $\epsilon^* > 0$ for all $Z \in \Omega_Z$.

The ideal weight vector W^* is an "artificial" quantity required for analytical purposes. W^* is defined as the value of W that minimizes $|\epsilon|$ for all $Z \in \Omega_Z \subset R^q$, i.e.,

$$W^* \triangleq \arg\min_{W \in R^l} \left\{ \sup_{Z \in \Omega_Z} |h(Z) - W^T S(Z)| \right\}. \quad (35)$$

RBF NN represents a class of linearly parameterized approximators, and can be replaced by any other linearly parameterized approximators such as spline functions [33] or fuzzy systems [18]. Moreover, nonlinearly parameterized approximators, such as multilayer neural network (MNN), can be linearized as linearly parameterized approximators, with the higher order terms of Taylor series expansions being taken as part of the modeling error, as shown in [5], [17]. The stability and performance properties of the adaptive system using nonlinearly parameterized approximators can be analyzed following the similar procedures therein. It is omitted here for clarity and conciseness.

III. Adaptive Neural Control Design for Σ_1

In the following derivation of the adaptive neural controller, NN approximation is only guaranteed within a compact set. Accordingly, the stability results obtained in this work are semiglobal in the sense that, as long as the input variables of the NNs remain within some compact sets, where the compact sets can be made as large as desired, there exists controller(s) with sufficiently large number of NN nodes that guarantees all the signals in the closed-loop remain bounded.

For the control of the uncertain MIMO system Σ_1 (28), we make the following assumption as commonly being done in the literature.

Assumption 3.1: The signs of $g_{j,i_j}(\cdot)$ are known, and there exist constants $\bar{g}_{j,i_j} \ge \underline{g}_{j,i_j} > 0$, $i_j = 1, \ldots, \rho_j$, $j = 1, \ldots, m$ such that $\bar{g}_{j,i_j} \ge |g_{j,i_j}(\cdot)| \ge \underline{g}_{j,i_j}$.

The above assumption implies that smooth functions $g_{j,i_j}(\cdot)$ are strictly either positive or negative. Without losing generality, we shall assume $\bar{g}_{j,i_j} \ge g_{j,i_j}(\cdot) \ge \underline{g}_{j,i_j}$.

The derivatives of $g_{j,i_j}(\cdot)$ are given by

$$\begin{aligned}\dot{g}_{j,i_j}(\bar{x}_{1,(i_j-\varrho_{j1})}, \ldots, \bar{x}_{m,(i_j-\varrho_{jm})}) &= \sum_{l=1}^{m} \sum_{k=1}^{i_j-\varrho_{jl}} \frac{\partial g_{j,i_j}(\cdot)}{\partial x_{l,k}} \dot{x}_{l,k} \\ &= \sum_{l=1}^{m} \sum_{k=1}^{i_j-\varrho_{jl}} \frac{\partial g_{j,i_j}(\cdot)}{\partial x_{l,k}} \\ &\quad \times [g_{l,k}(\cdot) x_{l,k+1} + f_{l,k}(\cdot)] \\ i_j &= 1, \ldots, \rho_j - 1, \\ j &= 1, \ldots, m \end{aligned} \quad (36)$$

$$\begin{aligned}\dot{g}_{j,\rho_j}(\bar{x}_{1,\rho_1-1}, \ldots, \bar{x}_{j,\rho_m-1}) &= \sum_{l=1}^{m} \sum_{k=1}^{\rho_l-1} \frac{\partial g_{j,\rho_j}(\cdot)}{\partial x_{l,k}} \dot{x}_{l,k} \\ &= \sum_{l=1}^{m} \sum_{k=1}^{\rho_l-1} \frac{\partial g_{j,\rho_j}(\cdot)}{\partial x_{l,k}} \\ &\quad \times [g_{l,k}(\cdot) x_{l,k+1} + f_{l,k}(\cdot)] \\ j &= 1, \ldots, m. \end{aligned} \quad (37)$$

Clearly, they only depend on states X. Because $f_{j,i_j}(\cdot)$ and $g_{j,i_j}(\cdot)$ are assumed to be smooth functions, they are therefore bounded within the compact set Ω. Thus, we have the following assumption.

Assumption 3.2: There exist constants $g^d_{j,i_j} > 0$, $i_j = 1, \ldots, \rho_j$, $j = 1, \ldots, m$ such that $|\dot{g}_{j,i_j}(\cdot)| \le g^d_{j,i_j}$ in the compact set Ω_{j,i_j} $\forall\, t \ge 0$.

For uncertain MIMO nonlinear system (28), we employ the idea of backstepping to design controllers for all the subsystems of (28). Note that because all subsystems in system (28) are interconnected, it is difficult to conclude the stability of the whole system by stability analysis of the individual subsystem separately. However, due to the block-triangular structure property, it is feasible to design for each subsystem a full state feedback controller, and prove the stability of the closed-loop MIMO system in a nested iterative manner.

For the controller design of the jth subsystem of (28), an intermediate desired feedback control α^*_{j,i_j} is first shown to exist which possesses some desired stabilizing properties at the recursive i_jth step, and then the i_jth-order subsubsystem of the jth subsystem is stabilized with respect to a Lyapunov function V_{j,i_j} by the design of a stabilizing function α_{j,i_j}, where an RBF neural network is employed to approximate the unknown part in intermediate desired feedback control α^*_{j,i_j}. The control law u_j for the jth subsystem is designed in the ρ_jth step.

Step 1: Define $z_{j,1} = x_{j,1} - x_{d1}$. Its derivative is

$$\begin{aligned}\dot{z}_{j,1} = f_{j,1}&\left(\bar{x}_{1,(1-\varrho_{j1})}, \ldots, \bar{x}_{m,(1-\varrho_{jm})}\right) \\ &+ g_{j,1}\left(\bar{x}_{1,(1-\varrho_{j1})}, \ldots, \bar{x}_{j,(m-\varrho_{jm})}\right) x_{j,2} - \dot{x}_{d1} \end{aligned} \quad (38)$$

where $\varrho_{jl} = \rho_j - \rho_l$, $l = 1, \ldots, m$. If $1 - \varrho_{jl} \le 0$, then the corresponding variable vector $\bar{x}_{j,(1-\varrho_{jl})}$ does not exist.

By viewing $x_{j,2}$ as a virtual control input, apparently there exists a desired feedback control

$$\alpha^*_{j,1} = -c_{j,1} z_{j,1} - \frac{1}{g_{j,1}} (f_{j,1} - \dot{x}_{d1})$$

where $c_{j,1}$ is a positive design constant to be specified later, $g_{j,1}(x_{j,1})$ and $f_{j,1}(x_{j,1})$ are unknown smooth functions of $x_{j,1}$.

Let $h_{j,1}(Z_{j,1}) = 1/g_{j,1}(f_{j,1} - \dot{x}_{d1})$ denote the unknown part of $\alpha^*_{j,1}$, with

$$Z_{j,1} = \left[\bar{x}^T_{1,(1-\varrho_{j1})}, \ldots, \bar{x}^T_{m,(1-\varrho_{jm})}, \dot{x}_{d1}\right]^T \in \Omega_{j,1}. \quad (39)$$

By employing an RBF neural network $W^T_{j,1} S_{j,1}(Z_{j,1})$ to approximate $h_{j,1}(Z_{j,1})$, $\alpha^*_{j,1}$ can be expressed as

$$\alpha^*_{j,1} = -c_{j,1} z_{j,1} - W^{*T}_{j,1} S_{j,1}(Z_{j,1}) - \epsilon_{j,1}$$

where $W^*_{j,1}$ denotes the ideal constant weights, and $|\epsilon_{j,1}| \le \epsilon^*_{j,1}$ is the approximation error with constant $\epsilon^*_{j,1} > 0$.

Since $W^*_{j,1}$ is unknown, $\alpha^*_{j,1}$ cannot be realized in practice. Define $z_{j,2} = x_{j,2} - \alpha_{j,1}$, and let

$$\alpha_{j,1} = -c_{j,1} z_{j,1} - \hat{W}^T_{j,1} S_{j,1}(Z_{j,1}). \quad (40)$$

Then, we have

$$\begin{aligned}\dot{z}_{j,1} &= f_{j,1} + g_{j,1}(z_{j,2} + \alpha_{j,1}) - \dot{x}_{d1} \\ &= g_{j,1}\left[z_{j,2} - c_{j,1} z_{j,1} - \tilde{W}^T_{j,1} S_{j,1}(Z_{j,1}) + \epsilon_{j,1}\right]. \end{aligned} \quad (41)$$

Consider the Lyapunov function candidate

$$V_{j,1} = \frac{1}{2g_{j,1}} z_{j,1}^2 + \frac{1}{2}\tilde{W}_{j,1}^T \Gamma_{j,1}^{-1} \tilde{W}_{j,1}. \tag{42}$$

The derivative of $V_{j,1}$ is

$$\begin{aligned}\dot{V}_{j,1} &= \frac{z_{j,1}\dot{z}_{j,1}}{g_{j,1}} - \frac{\dot{g}_{j,1} z_{j,1}^2}{2g_{j,1}^2} + \tilde{W}_{j,1}^T \Gamma_{j,1}^{-1} \dot{\hat{W}}_{j,1} \\ &= z_{j,1}z_{j,2} - c_{j,1}z_{j,1}^2 - \frac{\dot{g}_{j,1}}{2g_{j,1}^2} z_{j,1}^2 + z_{j,1}\epsilon_{j,1} \\ &\quad - \tilde{W}_{j,1}^T S_{j,1}(Z_{j,1}) z_{j,1} + \tilde{W}_{j,1}^T \Gamma_{j,1}^{-1} \dot{\hat{W}}_{j,1}. \end{aligned} \tag{43}$$

Consider the adaptation law for $\hat{W}_{j,1}$ as

$$\dot{\hat{W}}_{j,1} = \dot{\tilde{W}}_{j,1} = \Gamma_{W_{j,1}} [S_{j,1}(Z_{j,1}) z_{j,1} - \sigma_{j,1}\hat{W}_{j,1}] \tag{44}$$

where $\sigma_{j,1} > 0$ and $\Gamma_{W_{j,1}} = \Gamma_{W_{j,1}}^T > 0$ are design constants, $\tilde{W}_{j,1} = \hat{W}_{j,1} - W_{j,1}^*$.

Let $c_{j,1} = c_{j,10} + c_{j,11}$, with $c_{j,10}$ and $c_{j,11} > 0$. Then, (43) becomes

$$\begin{aligned}\dot{V}_{j,1} = z_{j,1}z_{j,2} - \left(c_{j,10} + \frac{\dot{g}_{j,1}}{2g_{j,1}^2}\right) z_{j,1}^2 - c_{j,11} z_{j,1}^2 \\ + z_{j,1}\epsilon_{j,1} - \sigma_{j,1}\tilde{W}_{j,1}^T \hat{W}_{j,1}. \end{aligned} \tag{45}$$

By completion of squares, we have

$$\begin{aligned}-\sigma_{j,1}\tilde{W}_{j,1}^T\hat{W}_{j,1} &= -\sigma_{j,1}\tilde{W}_{j,1}^T\left(\tilde{W}_{j,1} + W_{j,1}^*\right) \\ &\le -\sigma_{j,1}\left\|\tilde{W}_{j,1}\right\|^2 + \sigma_{j,1}\left\|\tilde{W}_{j,1}\right\|\left\|W_{j,1}^*\right\| \\ &\le -\frac{\sigma_{j,1}\left\|\tilde{W}_{j,1}\right\|^2}{2} + \frac{\sigma_{j,1}\left\|W_{j,1}^*\right\|^2}{2}\end{aligned}$$

$$-c_{j,11}z_{j,1}^2 + z_{j,1}\epsilon_{j,1} \le \frac{\epsilon_{j,1}^2}{4c_{j,11}} \le \frac{\epsilon_{j,1}^{*2}}{4c_{j,11}}. \tag{46}$$

Because $-(c_{j,10} + (\dot{g}_{j,1}/2g_{j,1}^2))z_{j,1}^2 \le -(c_{j,10} - (g_{j,1}^d/2\underline{g}_{j,1}^2))$ $z_{j,1}^2$, by choosing $c_{j,10}$ such that $c_{j,10}^* \triangleq c_{j,10} - \left(g_{j,1}^d/2\underline{g}_{j,1}^2\right) > 0$, we have the following inequality

$$\begin{aligned}\dot{V}_{j,1} < z_{j,1}z_{j,2} - c_{j,10}^* z_{j,1}^2 - \frac{\sigma_{j,1}\left\|\tilde{W}_{j,1}\right\|^2}{2} \\ + \frac{\sigma_{j,1}\left\|W_{j,1}^*\right\|^2}{2} + \frac{\epsilon_{j,1}^{*2}}{4c_{j,11}}. \end{aligned} \tag{47}$$

Step $i_j (2 \le i_j \le \rho_j - 1)$: Define $z_{j,i_j} = x_{j,i_j} - \alpha_{j,i_j-1}$. Its derivative is

$$\begin{aligned}\dot{z}_{j,i_j} &= f_{j,i_j}\left(\bar{x}_{1,(i_j-\varrho_{j1})}, \ldots, \bar{x}_{m,(i_j-\varrho_{jm})}\right) \\ &\quad + g_{j,i_j}\left(\bar{x}_{1,(i_j-\varrho_{j1})}, \ldots, \bar{x}_{m,(i_j-\varrho_{jm})}\right) x_{j,i_j+1} - \dot{\alpha}_{j,i_j-1}\end{aligned} \tag{48}$$

where $\varrho_{jl} = \rho_j - \rho_l, l = 1, \ldots, m$, and if $i_j - \varrho_{jl} \le 0$, then the corresponding variable vector $\bar{x}_{j,(i_j-\varrho_{jl})}$ does not exist.

By viewing x_{j,i_j+1} as a virtual control to stabilize the $(z_{j,1}, \ldots, z_{j,i_j})$-subsubsystem of the jth subsystem, there exists a desired feedback control

$$\alpha_{j,i_j}^* = -z_{j,i_j-1} - c_{j,i_j} z_{j,i_j} - \frac{1}{g_{j,i_j}}(f_{j,i_j} - \dot{\alpha}_{j,i_j-1})$$

where $c_{j,i_j} > 0$ is a design constant to be specified later, α_{j,i_j-1} is a function of $\bar{x}_{1,(i_j-1-\varrho_{j1})}, \ldots, \bar{x}_{m,(i_j-1-\varrho_{jm})}, x_d$ and $\hat{W}_{j,1}, \ldots, \hat{W}_{j,i_j-1}$. Therefore, $\dot{\alpha}_{j,i_j-1}$ can be expressed as

$$\dot{\alpha}_{j,i_j-1} = \sum_{l=1}^{m} \sum_{k=1}^{i_j-1-\varrho_{jl}} \frac{\partial \alpha_{j,i_j-1}}{\partial x_{l,k}} (g_{l,k}x_{l,k+1} + f_{l,k}) + \phi_{j,i_j-1}$$

where

$$\begin{aligned}\phi_{j,i_j-1} &= \frac{\partial \alpha_{j,i_j-1}}{\partial x_d}\dot{x}_d \\ &\quad + \sum_{k=1}^{i_j-1} \frac{\partial \alpha_{j,i_j-1}}{\partial \hat{W}_{j,k}} \left[\Gamma_{j,k}\left(S_{j,k}(Z_{j,k})z_{j,k} - \sigma_{j,k}\hat{W}_{j,k}\right)\right]\end{aligned} \tag{49}$$

is computable.

For the desired feedback control α_{j,i_j}^*, let $h_{j,i_j}(Z_{j,i_j}) = 1/g_{j,i_j}\left(f_{j,i_j} - \dot{\alpha}_{j,i_j-1}\right)$ denote the unknown part of α_{j,i_j}^*, where

$$\begin{aligned}Z_{j,i_j} = \Bigg[&\bar{x}_{1,(i_j-\varrho_{j1})}^T, \ldots, \bar{x}_{m,(i_j-\varrho_{jm})}^T, \left(\frac{\partial \alpha_{j,i_j-1}}{\partial \bar{x}_{1,(i_j-1-\varrho_{j1})}}\right)^T \\ &, \ldots, \left(\frac{\partial \alpha_{j,i_j-1}}{\partial \bar{x}_{m,(i_j-1-\varrho_{jm})}}\right)^T, \phi_{j,i_j-1}\Bigg]^T \in \Omega_{j,i_j}.\end{aligned} \tag{50}$$

By employing an RBF neural network $W_{j,i_j}^T S_{j,i_j}(Z_{j,i_j})$ to approximate $h_{j,i_j}(Z_{j,i_j})$, α_{j,i_j}^* can be expressed as

$$\alpha_{j,i_j}^* = -z_{j,i_j-1} - c_{j,i_j}z_{j,i_j} - W_{j,i_j}^{*T} S_{j,i_j}(Z_{j,i_j}) - \epsilon_{j,i_j}$$

where W_{j,i_j}^* denotes the ideal constant weights, and $|\epsilon_{j,i_j}| \le \epsilon_{j,i_j}^*$ is the approximation error with constant $\epsilon_{j,i_j}^* > 0$.

Since W_{j,i_j}^* is unknown, α_{j,i_j}^* cannot be realized in practice. Let us define $z_{j,i_j+1} = x_{j,i_j+1} - \alpha_{j,i_j}$ and

$$\alpha_{j,i_j} = -z_{j,i_j-1} - c_{j,i_j}z_{j,i_j} - \hat{W}_{j,i_j}^T S_{j,i_j}(Z_{j,i_j}). \tag{51}$$

Then, we have

$$\begin{aligned}\dot{z}_{j,i_j} &= f_{j,i_j} + g_{j,i_j}(z_{j,i_j+1} + \alpha_{j,i_j}) - \dot{\alpha}_{j,i_j-1} \\ &= g_{j,i_j}\left[z_{j,i_j+1} - z_{j,i_j-1} - c_{j,i_j}z_{j,i_j}\right. \\ &\quad \left. - \tilde{W}_{j,i_j}^T S_{j,i_j}(Z_{j,i_j}) + \epsilon_{j,i_j}\right].\end{aligned} \tag{52}$$

Remark 3.1: The principle of designing the neural network $W_{j,i_j}^T S_{j,i_j}(Z_{j,i_j})$ as such is to use as few neurons as possible to approximate the unknown function $h_{j,i_j}(Z_{j,i_j})$. Though $h_{j,i_j}(Z_{j,i_j})$ is a function of $\bar{x}_{1,(i_j-\varrho_{j1})}, \ldots, \bar{x}_{m,(i_j-\varrho_{jm})}, x_d$ and $\hat{W}_{j,1}, \ldots, \hat{W}_{j,i_j-1}$, the weight estimates $\hat{W}_{j,1}, \ldots, \hat{W}_{j,i_j-1}$ are not recommended to be taken as inputs to the NN because of the curse of dimensionality of RBF NN (see, e.g., [26]). By defining intermediate variables $\left(\partial\alpha_{j,i_j-1}/\partial\bar{x}_{1,(i_j-1-\varrho_{j1})}\right)^T, \ldots, \left(\partial\alpha_{j,i_j-1}/\partial\bar{x}_{m,(i_j-1-\varrho_{jm})}\right)^T$ and ϕ_{j,i_j-1}, which are available through the computation of known information, the NN approximation $\hat{W}_{j,i_j}^T S_{j,i_j}(Z_{j,i_j})$ of the unknown function $h_{j,i_j}(Z_{j,i_j})$ can be computed by using the minimal number of NN inputs Z_{j,i_j} as given in (50). The same idea of choosing the inputs of NN is also used in the following design steps.

Consider the Lyapunov function candidate

$$V_{j,i_j} = V_{j,i_j-1} + \frac{1}{2g_{j,i_j}} z_{j,i_j}^2 + \frac{1}{2}\tilde{W}_{j,i_j}^T \Gamma_{j,i_j}^{-1} \tilde{W}_{j,i_j}. \tag{53}$$

The derivative of V_{j,i_j} is

$$\begin{aligned}\dot{V}_{j,i_j} =& \dot{V}_{j,i_j-1} + \frac{z_{j,i_j}\dot{z}_{j,i_j}}{g_{j,i_j}} - \frac{\dot{g}_{j,i_j} z_{j,i_j}^2}{2g_{j,i_j}^2} + \tilde{W}_{j,i_j}^T \Gamma_{j,i_j}^{-1} \dot{\hat{W}}_{j,i_j} \\ =& \dot{V}_{j,i_j-1} - z_{j,i_j-1} z_{j,i_j} + z_{j,i_j} z_{j,i_j+1} - c_{j,i_j} z_{j,i_j}^2 \\ & - \frac{\dot{g}_{j,i_j}}{2g_{j,i_j}^2} z_{j,i_j}^2 + z_{j,i_j}\epsilon_{j,i_j} - \tilde{W}_{j,i_j}^T S_{j,i_j}(Z_{j,i_j}) z_{j,i_j} \\ & + \tilde{W}_{j,i_j}^T \Gamma_{j,i_j}^{-1} \dot{\hat{W}}_{j,i_j}. \end{aligned} \tag{54}$$

Consider the adaptation law for $\hat{W}_{j,i_j}$ as

$$\dot{\hat{W}}_{j,i_j} = \dot{\tilde{W}}_{j,i_j} = \Gamma_{W_{j,i_j}} \left[S_{j,i_j}(Z_{j,i_j}) z_{j,i_j} - \sigma_{j,i_j} \hat{W}_{j,i_j} \right] \tag{55}$$

where $\sigma_{j,i_j} > 0$ and $\Gamma_{W_{j,i_j}} = \Gamma_{W_{j,i_j}}^T > 0$ are design constants, $\tilde{W}_{j,i_j} = \hat{W}_{j,i_j} - W_{j,i_j}^*$.

Let $c_{j,i_j} = c_{j,i_j0} + c_{j,i_j1}$, where c_{j,i_j0} and $c_{j,i_j1} > 0$. Then, (54) becomes

$$\begin{aligned}\dot{V}_{j,i_j} =& \dot{V}_{j,i_j-1} - z_{j,i_j-1} z_{j,i_j} + z_{j,i_j} z_{j,i_j+1} \\ & - \left(c_{j,i_j0} + \frac{\dot{g}_{j,i_j}}{2g_{j,i_j}^2} \right) z_{j,i_j}^2 - c_{j,i_j1} z_{j,i_j}^2 + z_{j,i_j}\epsilon_{j,i_j} \\ & - \sigma_{j,i_j} \tilde{W}_{j,i_j}^T \hat{W}_{j,i_j}. \end{aligned} \tag{56}$$

By completion of squares, we have

$$\begin{aligned}-\sigma_{j,i_j} \tilde{W}_{j,i_j}^T \hat{W}_{j,i_j} =& -\sigma_{j,i_j} \tilde{W}_{j,i_j}^T \left(\tilde{W}_{j,i_j} + W_{j,i_j}^* \right) \\ \le& -\sigma_{j,i_j} \left\| \tilde{W}_{j,i_j} \right\|^2 + \sigma_{j,i_j} \left\| \tilde{W}_{j,i_j} \right\| \left\| W_{j,i_j}^* \right\| \\ \le& -\frac{\sigma_{j,i_j} \left\| \tilde{W}_{j,i_j} \right\|^2}{2} + \frac{\sigma_{j,i_j} \left\| W_{j,i_j}^* \right\|^2}{2} \\ & - c_{j,i_j1} z_{j,i_j}^2 + z_{j,i_j}\epsilon_{j,i_j} \\ \le& \frac{\epsilon_{j,i_j}^2}{4c_{j,i_j1}} \le \frac{\epsilon_{j,i_j}^{*2}}{4c_{j,i_j1}}. \end{aligned} \tag{57}$$

Because $-(c_{j,i_j0} + (\dot{g}_{j,i_j}/2g_{j,i_j}^2)) z_{j,i_j}^2 \le -(c_{j,i_j0} - (g_{j,i_j}^d/2\underline{g}_{j,i_j}^2)) z_{j,i_j}^2$, by choosing c_{j,i_j0} such that $c_{j,i_j0}^* \triangleq c_{j,i_j0} - \left(g_{j,i_j}^d / 2\underline{g}_{j,i_j}^2 \right) > 0$, we have the following inequality:

$$\begin{aligned}\dot{V}_{j,i_j} <& z_{j,i_j} z_{j,i_j+1} - \sum_{k=1}^{i_j} c_{j,k0}^* z_{j,k}^2 - \sum_{k=1}^{i_j} \frac{\sigma_{j,k} \left\| \tilde{W}_{j,k} \right\|^2}{2} \\ & + \sum_{k=1}^{i_j} \frac{\sigma_{j,k} \left\| W_{j,k}^* \right\|^2}{2} + \sum_{k=1}^{i_j} \frac{\epsilon_{j,k}^{*2}}{4c_{j,k1}}. \end{aligned} \tag{58}$$

Step ρ_j: The derivative of $z_{j,\rho_j} = x_{j,\rho_j} - \alpha_{j,\rho_j-1}$ is

$$\begin{aligned}\dot{z}_{j,\rho_j} =& f_{j,\rho_j}(X, u_1, \ldots, u_{j-1}) \\ & + g_{j,\rho_j}\left(\bar{x}_{1,\rho_1-1}, \ldots, \bar{x}_{j,\rho_j-1} \right) u_j - \dot{\alpha}_{j,\rho_j-1}. \end{aligned} \tag{59}$$

To stabilize the jth subsystem $(z_{j,1}, \ldots, z_{j,\rho_j})$, there exists a desired feedback control

$$u_j^* = -z_{j,\rho_j-1} - c_{j,\rho_j} z_{j,\rho_j} - \frac{1}{g_{j,\rho_j}} (f_{j,\rho_j} - \dot{\alpha}_{j,\rho_j-1})$$

where $c_{j,\rho_j} > 0$ is a design constant to be specified later, α_{j,ρ_j-1} is a function of $\bar{x}_{1,(\rho_1-1)}, \ldots, \bar{x}_{m,(\rho_m-1)}, x_d$ and $\hat{W}_{j,1}, \ldots, \hat{W}_{j,\rho_j-1}$. Therefore, $\dot{\alpha}_{j,\rho_j-1}$ can be expressed as

$$\dot{\alpha}_{j,\rho_j-1} = \sum_{l=1}^{m} \sum_{k=1}^{\rho_l-1} \frac{\partial \alpha_{j,\rho_j-1}}{\partial x_{l,k}} (g_{l,k} x_{l,k+1} + f_{l,k}) + \phi_{j,\rho_j-1}$$

where

$$\begin{aligned}\phi_{j,\rho_j-1} =& \frac{\partial \alpha_{j,\rho_j-1}}{\partial x_d} \dot{x}_d \\ & + \sum_{k=1}^{\rho_j-1} \frac{\partial \alpha_{j,\rho_j-1}}{\partial \hat{W}_{j,k}} \left[\Gamma_{j,k} \left(S_{j,k}(Z_{j,k}) z_{j,k} - \sigma_{j,k} \hat{W}_{j,k} \right) \right] \end{aligned} \tag{60}$$

is computable.

For the desired feedback control u_j^*, let $h_{j,\rho_j}(Z_{j,\rho_j}) = 1/g_{j,\rho_j}(f_{j,\rho_j} - \dot{\alpha}_{j,\rho_j-1})$ denote the unknown part of u_j^*, where

$$\begin{aligned}Z_{j,\rho_j} =& \left[X^T, u_1, \ldots, u_{j-1}, \left(\frac{\partial \alpha_{j,\rho_j-1}}{\partial \bar{x}_{1,(\rho_1-1)}} \right)^T, \ldots, \right. \\ & \left. \left(\frac{\partial \alpha_{j,\rho_j-1}}{\partial \bar{x}_{m,(\rho_m-1)}} \right)^T, \phi_{j,\rho_j-1} \right]^T \in \Omega_{j,\rho_j}. \end{aligned} \tag{61}$$

By employing an RBF neural network $W_{j,\rho_j}^T S_{j,\rho_j}(Z_{j,\rho_j})$ to approximate $h_{j,\rho_j}(Z_{j,\rho_j})$, u_j^* can be expressed

$$u_j^* = -z_{j,\rho_j-1} - c_{j,\rho_j} z_{j,\rho_j} - W_{j,\rho_j}^{*T} S_{j,\rho_j}(Z_{j,\rho_j}) - \epsilon_{j,\rho_j}$$

where W_{j,ρ_j}^* denotes the ideal constant weights, and $|\epsilon_{j,\rho_j}| \le \epsilon_{j,\rho_j}^*$ is the approximation error with constant $\epsilon_{j,\rho_j}^* > 0$.

Since W_{j,ρ_j}^* is unknown, u_j^* cannot be realized in practice. Consider

$$u_j = -z_{j,\rho_j-1} - c_{j,\rho_j} z_{j,\rho_j} - \hat{W}_{j,\rho_j}^T S_{j,\rho_j}(Z_{j,\rho_j}). \tag{62}$$

Then, we have

$$\begin{aligned}\dot{z}_{j,\rho_j} =& f_{j,\rho_j} + g_{j,\rho_j} u_j - \dot{\alpha}_{j,\rho_j-1} \\ =& g_{j,\rho_j} \left[-z_{j,\rho_j-1} - c_{j,\rho_j} z_{j,\rho_j} \right. \\ & \left. - \tilde{W}_{j,\rho_j}^T S_{j,\rho_j}(Z_{j,\rho_j}) + \epsilon_{j,\rho_j} \right]. \end{aligned}$$

Consider the Lyapunov function candidate

$$V_{j,\rho_j} = V_{j,\rho_j-1} + \frac{1}{2g_{j,\rho_j}} z_{j,\rho_j}^2 + \frac{1}{2} \tilde{W}_{j,\rho_j}^T \Gamma_{j,\rho_j}^{-1} \tilde{W}_{j,\rho_j}. \tag{63}$$

The derivative of V_{j,ρ_j} is

$$\begin{aligned}\dot{V}_{j,\rho_j} =& \dot{V}_{j,\rho_j-1} + \frac{z_{j,\rho_j}\dot{z}_{j,\rho_j}}{g_{j,\rho_j}} - \frac{\dot{g}_{j,\rho_j} z_{j,\rho_j}^2}{2g_{j,\rho_j}^2} + \tilde{W}_{j,\rho_j}^T \Gamma_{j,\rho_j}^{-1} \dot{\hat{W}}_{j,\rho_j} \\ =& \dot{V}_{j,\rho_j-1} - z_{j,\rho_j-1} z_{j,\rho_j} - c_{j,\rho_j} z_{j,\rho_j}^2 - \frac{\dot{g}_{j,\rho_j}}{2g_{j,\rho_j}^2} z_{j,\rho_j}^2 \\ & + z_{j,\rho_j}\epsilon_{j,\rho_j} - \tilde{W}_{j,\rho_j}^T S_{j,\rho_j}(Z_{j,\rho_j}) z_{j,\rho_j} \\ & + \tilde{W}_{j,\rho_j}^T \Gamma_{j,\rho_j}^{-1} \dot{\hat{W}}_{j,\rho_j}. \end{aligned} \tag{64}$$

Consider the adaptation law for $\hat{W}_{j,\rho_j}$ as

$$\dot{\hat{W}}_{j,\rho_j} = \dot{\tilde{W}}_{j,\rho_j} = \Gamma_{W_{j,\rho_j}} \left[S_{j,\rho_j}(Z_{j,\rho_j}) z_{j,\rho_j} - \sigma_{j,\rho_j} \hat{W}_{j,\rho_j} \right] \tag{65}$$

where $\sigma_{j,\rho_j} > 0$ and $\Gamma_{W_{j,\rho_j}} = \Gamma_{W_{j,\rho_j}}^T > 0$ are design constants, $\tilde{W}_{j,\rho_j} = \hat{W}_{j,\rho_j} - W_{j,\rho_j}^*$.

Let $c_{j,\rho_j} = c_{j,\rho_j 0} + c_{j,\rho_j 1}$, where $c_{j,\rho_j 0}$ and $c_{j,\rho_j 1} > 0$. Then, (64) becomes

$$\dot{V}_{j,\rho_j} = \dot{V}_{j,\rho_j-1} - z_{j,\rho_j-1} z_{j,\rho_j} - \left(c_{j,\rho_j 0} + \frac{\dot{g}_{j,\rho_j}}{2g_{j,\rho_j}^2} \right) z_{j,\rho_j}^2 - c_{j,\rho_j 1} z_{j,\rho_j}^2 + z_{j,\rho_j} \epsilon_{j,\rho_j} - \sigma_{j,\rho_j} \tilde{W}_{j,\rho_j}^T \hat{W}_{j,\rho_j}. \quad (66)$$

By completion of squares, we have

$$\begin{aligned} -\sigma_{j,\rho_j} \tilde{W}_{j,\rho_j}^T \hat{W}_{j,\rho_j} &= -\sigma_{j,\rho_j} \tilde{W}_{j,\rho_j}^T \left(\tilde{W}_{j,\rho_j} + W_{j,\rho_j}^* \right) \\ &\le -\sigma_{j,\rho_j} \left\| \tilde{W}_{j,\rho_j} \right\|^2 + \sigma_{j,\rho_j} \left\| \tilde{W}_{j,\rho_j} \right\| \left\| W_{j,\rho_j}^* \right\| \\ &\le -\frac{\sigma_{j,\rho_j} \left\| \tilde{W}_{j,\rho_j} \right\|^2}{2} + \frac{\sigma_{j,\rho_j} \left\| W_{j,\rho_j}^* \right\|^2}{2} \\ &\quad - c_{j,\rho_j 1} z_{j,\rho_j}^2 + z_{j,\rho_j} \epsilon_{j,\rho j} \\ &\le \frac{\epsilon_{j,\rho_j}^2}{4c_{j,\rho_j 1}} \le \frac{\epsilon_{j,\rho_j}^{*2}}{4c_{j,\rho_j 1}}. \end{aligned} \quad (67)$$

Because $-(c_{j,\rho_j 0} + (\dot{g}_{j,\rho_j}/2g_{j,\rho_j}^2)) z_{j,\rho_j}^2 \le -(c_{j,\rho_j 0} - (g_{j,\rho_j}^d/2\underline{g}_{j,\rho_j}^2)) z_{j,\rho_j}^2$, by choosing $c_{j,\rho_j 0}$ such that $c_{j,\rho_j 0}^* \triangleq c_{j,\rho_j 0} - \left(g_j, \rho_j^d/2\underline{g}_j^2, \rho_j \right) > 0$, we have the following inequality:

$$\dot{V}_{j,\rho_j} < -\sum_{k=1}^{\rho_j} c_{j,k0}^* z_{j,k}^2 - \sum_{k=1}^{\rho_j} \frac{\sigma_{j,k} \left\| \tilde{W}_{j,k} \right\|^2}{2} + \delta_j \quad (68)$$

where

$$\delta_j \le \sum_{k=1}^{\rho_j} \frac{\sigma_{j,k} \left\| W_{j,k}^* \right\|^2}{2} + \sum_{k=1}^{\rho_j} \frac{\epsilon_{j,k}^{*2}}{4c_{j,k1}}.$$

If we choose $c_{j,k0}^*$ such that $c_{j,k0}^* \ge \gamma_j/2\underline{g}_{j,k}$, i.e., $c_{j,k0} > (\gamma_j/2\underline{g}_{j,k}) + (g_{j,k}^d/2\underline{g}_{j,k}^2)$, $k = 1, \ldots, \rho_j$, where γ_j is a positive constant, and choose $\sigma_{j,k}$ and $\Gamma_{j,k}$ such that $\sigma_{j,k} \ge \gamma_j \lambda_{\max} \left\{ \Gamma_{j,k}^{-1} \right\}$, $k = 1, \ldots, n$, then from (68) we have the following inequality:

$$\begin{aligned} \dot{V}_{j,\rho_j} &< -\sum_{k=1}^{\rho_j} c_{j,k0}^* z_{j,k}^2 - \sum_{k=1}^{\rho_j} \frac{\sigma_{j,k} \left\| \tilde{W}_{j,k} \right\|^2}{2} + \delta_j \\ &\le -\sum_{k=1}^{\rho_j} \frac{\gamma_j}{2\underline{g}_{j,k}} z_{j,k}^2 - \sum_{k=1}^{\rho_j} \frac{\gamma_j \tilde{W}_{j,k}^T \Gamma_{j,k}^{-1} \tilde{W}_{j,k}}{2} + \delta_j \\ &\le -\gamma_j \left[\sum_{k=1}^{\rho_j} \frac{1}{2g_{j,k}} z_{j,k}^2 + \sum_{k=1}^{\rho_j} \frac{\tilde{W}_{j,k}^T \Gamma_{j,k}^{-1} \tilde{W}_{j,k}}{2} \right] + \delta_j \\ &\le -\gamma_j V_{j,\rho_j} + \delta_j. \end{aligned} \quad (69)$$

Let $V = \sum_{j=1}^m V_{j,\rho_j}$. Its derivative is

$$\begin{aligned} \dot{V} = \sum_{j=1}^m \dot{V}_{j,\rho_j} &< \sum_{j=1}^m (-\gamma_j V_{j,\rho_j} + \delta_j) \\ &< -\gamma V + \delta \end{aligned} \quad (70)$$

where $\gamma = \min\{\gamma_1, \ldots, \gamma_m\}$ and $\delta = \sum_{j=1}^m \delta_j$ are positive constants.

Therefore, according to Lemma 1.2, all z_{j,i_j} and $\hat{W}_{j,i_j}$ $(j = 1, \ldots, m,\ i_j = 1, \ldots, \rho_j)$ are uniformly ultimately bounded for bounded initial conditions. Because of the interconnections between the subsystems in system (28), we cannot conclude the stability of the whole closed-loop system by stability analysis of individual subsystem separately. However, due to the structure property of system (28), we can prove the stability of the states X of the whole MIMO system in a nested iterative manner, as will be shown in the proof of the following theorem.

Theorem 3.1: Consider the closed-loop system consisting of the plant (28), the reference model (31), the controllers (62) and the NN weight updating laws (44), (55), and (65) for all the subsystems $(j = 1, \ldots, m)$. Then, for bounded initial conditions,

i) there exist sufficiently large compact sets Ω_{j,i_j} such that $Z_{j,i_j} \in \Omega_{j,i_j}$, $i_j = 1, \ldots, \rho_j$, $j = 1, \ldots, m$ for all $t \ge 0$, and all signals in the closed-loop system remain bounded,

ii) the states X and the neural weights $\hat{W}_j = \left[\hat{W}_{j,1}^T, \ldots, \hat{W}_{j,\rho_j}^T \right]^T$ $(j = 1, \ldots, m)$ eventually converge to the compact set

$$\Omega_{s1} \triangleq \left\{ X, \hat{W}_1, \ldots, \hat{W}_m \,\middle|\, \lim_{t\to\infty} \|z(t)\| = \mu_z^*, \lim_{t\to\infty} \|\tilde{W}\| = \mu_{\tilde{W}}^*, x_d \in \Omega_d \right\} \quad (71)$$

where constants $\mu_z^* = \sqrt{(2c_2/c_1 \lambda_{Q\min})}$, $\mu_{\tilde{W}}^* = \sqrt{(2c_2/c_1 \lambda_{\Gamma\min})}$, $\lambda_{Q\min} = \min_{\tau\in[0,t]} \lambda_{\min}(Q(\tau))$, and $\lambda_{\Gamma\min} = \lambda_{\min}\Gamma^{-1}$ with $Q(t) = \mathrm{diag}[1/g_{j,i_j}]$ and $\Gamma^{-1} = \mathrm{diag}[\Gamma_{j,i_j}]$, $j = 1, \ldots, n$, $i_j = 1, \ldots, \rho_j$.

In the following, we will prove the semiglobal stability of the closed-loop system in part i), and the convergence of output tracking error in part ii), respectively.

Proof: i) According to Lemma 1.2, we know from (69), that all z_{j,i_j} and $\hat{W}_{j,i_j}$ $(j = 1, \ldots, m, i_j = 1, \ldots, \rho_j)$ are uniformly bounded. Since $z_{j,1} = x_{j,1} - x_{dj}$ and x_d are bounded, we have that the first state $x_{j,1}$ of all the subsystems $(j = 1, \ldots, m)$ remain bounded. To prove $x_{j,2} = z_{j,2} + \alpha_{j,1}$, $j = 1, \ldots, m$ remain bounded, we need to show that $\alpha_{j,1}$ is bounded. Because $\alpha_{j,1}$ in (40) is a function of $z_{j,1}$, $Z_{j,1}$ and $\hat{W}_{j,1}$, i.e., a function of $\bar{x}_{1,(1-\varrho_{j1})}, \ldots, \bar{x}_{m,(1-\varrho_{jm})}$, x_d and $\hat{W}_{j,1}$, we may not conclude the boundedness of $\alpha_{j,1}$ immediately as not all the states $\bar{x}_{l,(1-\varrho_{jl})}$, $l = 1, \ldots, m$, $l \ne j$ have been proven bounded yet. As stated in Remark 1, ϱ_{jl} is the order difference between the jth and lth subsystems, and $\bar{x}_{l,(i_j-\varrho_{jl})}$ represents the state variables of the lth subsystem which are embedded in the jth subsystem. For clarity of the presentation, we consider the following two cases.

1) All the subsystems are of the same order, i.e., $\rho_1 = \rho_2 = \cdots = \rho_m$. Accordingly, for the order difference between the jth and lth subsystems, $\varrho_{jl} = \rho_j - \rho_l = 0, \forall\, j, l = 1, \ldots, m$, and $\bar{x}_{l,(i_j-\varrho_{jl})} = \bar{x}_{l,i_j}$, $l = 1, \ldots, m$, which means that $\alpha_{j,1}$ in (40) is a function of $x_{1,1}, \ldots, x_{m,1}, x_d$ and $\hat{W}_{j,1}$. Since the first state $x_{j,1}$ of all the subsystems $(j = 1, \ldots, m)$ have been proven bounded, we conclude that $\alpha_{j,1}$, $j = 1, \ldots, m$ remain bounded, which in turn leads to the boundedness of $x_{j,2}$, $j = 1, \ldots, m$. Following the same way, we can prove one by one that all α_{j,i_j-1} and x_{j,i_j}, $j = 1, \ldots, m$, $i_j = 3, \ldots, \rho_j$ remain bounded. Therefore, the states X of the interconnected MIMO system (28) remain bounded.

2) Not all the subsystems are of the same order, i.e., there exist at least two subsystems (e.g., the jth and the lth subsystems), such that $\rho_j \neq \rho_l$, and $\varrho_{jl} = \rho_j - \rho_l \neq 0$. Note that in this case, we may have $i_j - \varrho_{jl} \leq 0$. As shown in Section II, if $i_j - \varrho_{jl} \leq 0$, then the corresponding variable vector $\bar{x}_{l,(i_j-\varrho_{jl})}$ does not exist, and does not appear in the functions in (28). If $i_j - \varrho_{jl} > 0$, then state variables $\bar{x}_{l,(i_j-\varrho_{jl})}$ of the lth subsystem will be embedded in the jth subsystem.

In this case, we cannot prove the boundedness of $\alpha_{j,1}$, $j = 1, \ldots, m$ in one single step as in case 1). We can only prove the boundedness of some $\alpha_{q,1}$ in those qth subsystem(s) first, where q is (or qs are) determined by ($i_q = i_j = 2, j = 1, \ldots, m$ in this step)

$$(q, i_q) = \arg \max_{(j,i_j) \in U_r} \{\rho_j - i_j\} \quad (72)$$

with

$$U_r \triangleq \{(j, i_j) | x_{j,i_j} \text{ not proven bounded yet}\}.$$

By choosing the qth subsystem(s) in this way, we can pick only those $\alpha_{q,1}$ which are functions of bounded variables x_d, $\hat{W}_{q,1}$, and $x_{q,1}$. Accordingly, we conclude that $\alpha_{q,1}$ are bounded. From the boundedness of z_2, and $\alpha_{q,1}$, we obtain the boundedness of $x_{q,2}$. Next, we will proceed to prove the boundedness of α_{q,i_q-1}, and subsequently x_{q,i_q} by definition, where (q, i_q) are determined again by (72). In fact, according to the block-triangular structure of system (28), checking condition (72) indicates that the virtual control α_{q,i_q-1} has the minimum variable embedding from other subsystems and thus is a function of state variables which have all been proven bounded already. Accordingly, α_{q,i_q-1} can be proven bounded next. In this way, we can prove one by one that all α_{j,i_j-1} in (51) and subsequently, $x_{j,i_j}, j = 1, \ldots, m, i_j = 2, \ldots, \rho_j$, remain bounded. Thus, the states X of the interconnected MIMO system (28) remain bounded.

Using (62), we conclude that u_1 is bounded because it is a function of bounded variables X, x_d and $\hat{W}_{1,1}, \ldots, \hat{W}_{1,\rho_1}$. Similarly, we have that controls $u_j, j = 2, \ldots, m$, are also bounded. Thus, all the signals in the closed-loop system remain bounded. For better understanding to the above stability analysis, two exemplar MIMO systems will be provided in Remark 3.2 for demonstration.

Thus, for bounded initial conditions, all signals in the closed-loop system remain bounded.

ii) From the derivations, we know that

$$V = \frac{1}{2}\sum_{j=1}^{m}\sum_{i_j=1}^{\rho_j} \frac{1}{g_{j,i_j}} z_{j,i_j}^2 + \frac{1}{2}\sum_{j=1}^{m}\sum_{i_j=1}^{\rho_j} \tilde{W}_{j,i_j}^T \Gamma_{j,i_j}^{-1} \tilde{W}_{j,i_j} \quad (73)$$

which can be easily written into the form of (7) as follows:

$$V(t) = \frac{1}{2} e^T(t) Q(t) e(t) + \frac{1}{2} \tilde{W}^T(t) \Gamma^{-1}(t) \tilde{W}(t) \quad (74)$$

where $Q(t) = \text{diag}[1/g_{j,i_j}]$ and $\Gamma^{-1} = \text{diag}[\Gamma_{j,i_j}]$, $j = 1, \ldots, n, i_j = 1, \ldots, \rho_j$.

The uniformly ultimately boundedness follows directly from Lemma 1.2. The actual size of a residual set depends on the NN approximation error ϵ_{j,i_j} and controller parameters c_{j,i_j}, σ_{j,i_j} and Γ_{j,i_j}. It is easily seen that the increase in the control gain c_{j,i_j}, adaptation gain Γ_{j,i_j} and NN node number l will result in a better tracking performance. ◇

Remark 3.2: For MIMO system (28) with many subsystems, it might become very involved when trying to find out which state to be proven bounded next. By introducing the checking condition (72), the determination of q and i_q, and subsequently, α_{q,i_q-1} and x_{q,i_q} can be easily done. To help readers understand the above stability analysis for the whole closed-loop system better, let us consider systems Σ_{S_1} (29) and Σ_{S_2} (30) as examples for cases 1) and 2), respectively.

a) System Σ_{S_1}(29) for case 1)

After proving the boundedness of $x_{1,1}$ and $x_{2,1}$, we can proceed to prove the boundedness of $\alpha_{j,1}$, $j = 1, 2$, because $\alpha_{j,1}$(40) is a function of $x_{1,1}, x_{2,1}, x_d$ and $\hat{W}_{j,1}$. Since the states $x_{1,1}$ and $x_{2,1}$ have been proven bounded, we conclude that $\alpha_{j,1}$, $j = 1, 2$ remain bounded. This leads to the boundedness of $x_{1,2}$ and $x_{2,2}$ by definition. Thus, the states X of the overall system (29) are proven bounded.

b) System Σ_{S_2}(30) for case 2)

As stated in case 2), after proving the boundedness of $x_{1,1}$ and $x_{2,1}$, we cannot proceed to prove the boundedness of $\alpha_{1,1}$ because as can be seen from (39) and (40), $\alpha_{1,1}$ is also a function of $\bar{x}_{2,3}$, where $\bar{x}_{2,3} = [x_{2,1}, x_{2,2}, x_{2,3}]^T$, and $x_{2,2}, x_{2,3}$ have not been proven bounded yet. According to (72), we know that $q = 2$ ($i_q = 2$ in this step), which means that we can prove the boundedness of $\alpha_{2,1}$ of the second subsystem in Σ_{S2} first. (From (40), it is clear that $\alpha_{2,1}$ is a function of $x_{2,1}$, x_d, $\hat{W}_{2,1}$, which have been proven bounded already). This will lead to the boundedness of $x_{2,2}$ by definition.

By using (72), we find that $(q, i_q) = (2, 3)$, which means that we can prove the boundedness of $\alpha_{2,2}$ and $x_{2,3}$. From (51) and (50), we can see that $\alpha_{2,2}$ is bounded because it is a function of $x_{2,1}$, $x_{2,2}$, x_d, $\hat{W}_{2,1}$ and $\hat{W}_{2,2}$, which have all been proven bounded. The boundedness of $\alpha_{2,2}$ will lead to the boundedness of $x_{2,3}$ by definition. However, we cannot prove the boundedness of $\alpha_{1,1}$ and $x_{1,2}$ in this step, because we still need to prove the boundedness of the state variable $x_{2,3}$ first. By using (72) again, we have $(q, i_q) = (1, 2)$ and $(q, i_q) = (2, 4)$. At this time, we can prove the boundedness of $\alpha_{1,1}$(40) and $\alpha_{2,3}$(51), and subsequently, $x_{1,2}$ and $x_{2,4}$ simultaneously. Therefore, the states X of the overall system (30) are proven bounded. The sequence of the state variables being proven bounded is shown as follows:

$$(x_{1,1}, x_{2,1}) \rightarrow x_{2,2} \rightarrow x_{2,3} \rightarrow (x_{1,2}, x_{2,4}).$$

Remark 3.3: In the above analysis, it can be seen from (69)–(71) that the size of Ω_{s1} depends on $W_{j,k}^*$, $\epsilon_{j,k}^*$, and all design parameters. Since there is no analytical result in the NN literature to quantify the relationship of the NN node numbers l_j, the ideal neural network weights W_{j,i_j}^*, and the bounding approximation error ϵ_{j,i_j}^*, an explicit expression of the stability condition is not available at present. However, it is clear that: i) increasing c_{j,i_j0} might lead to larger γ_j, and subsequently larger γ, and increasing $c_{j,i_j 1}$ will reduce δ_{j,i_j}, and subsequently δ, thus, increasing c_{j,i_j} will lead to smaller Ω_{s1}; ii) decreasing σ_{j,i_j} will help to reduce δ_{j,i_j}, and increasing the NN node number l_j will help to reduce ϵ^*_{j,i_j}, both of which will help to reduce the size of Ω_{s1}.

IV. Adaptive Neural Control of Partially Known MIMO Systems

Note that in Section III, adaptive neural controller (62) is designed for general system Σ_1(28), where all the terms $f_{j,i_j}(\cdot)$ and $g_{j,i_j}(\cdot)$ ($i_j = 1,\ldots,\rho_j$ and $j = 1,\ldots,m$) are unknown nonlinear functions. Additional *a priori* information about the systems could not be fully exploited/utilized, say, $f_{j,i_j}(\cdot)$ contain LIP terms, and g_{j,i_j} ($i_j = 1,\ldots,\rho_j$ and $j = 1,\ldots,m$) are unknown constants. In this section, we consider a special case of partially known MIMO system in the following block-triangular form shown in (75) at the bottom of the page where x_{j,i_j}, $i_j = 1,\ldots,\rho_j$ are the states of the jth subsystem, $X = \left[\bar{x}^T_{1,\rho_1},\ldots,\bar{x}^T_{m,\rho_m}\right]^T$ with $\bar{x}_{j,i_j} = [x_{j,1},\ldots,x_{j,i_j}]^T \in R^{i_j}$ ($i_j = 1,\ldots,\rho_j$ and $j = 1,\ldots,m$) represents the state variables of the whole system, $u_j \in R$ and $y_j \in R$ are the system inputs and outputs, respectively; $f_{j,i_j}(\cdot)$ are unknown smooth functions, $F_{j,i_j}(\cdot)$ and $G_{j,i_j}(\cdot)$ are known smooth nonlinear functions, g_{j,i_j} and θ_{j,i_j} are unknown constant parameters, j, i_j, ρ_j and m are positive integers, and $\varrho_{jl} \triangleq \rho_j - \rho_l$ is the order difference between the jth and lth subsystems, as discussed in Remark 1.

System Σ_2 (75), shown at the bottom of the page, has unknown nonlinear functions $f_{j,i_j}(\cdot)$, parametric uncertainties g_{j,i_j}, θ_{j,i_j}, and known nonlinear functions $F_{j,i_j}(\cdot)$ and $G_{j,i_j}(\cdot)$. This problem formulation is very general in the sense that any *a priori* information represented by $F_{j,i_j}(\cdot)$ and $G_{j,i_j}(\cdot)$, which may be available through the laws of physics, properties of materials, or any identification methods, can be employed and incorporated into the controller design.

The control objective here is to design adaptive neural controller for system (75) such that: i) all the signals in the closed-loop remain semiglobally uniformly ultimately bounded; ii) the output y_j follows the desired trajectories y_{dj} generated from the smooth, bounded reference model (31); and iii) the overparametrization problem [7] caused by the parametric uncertainties g_{j,i_j} and θ_{j,i_j} in adaptive backstepping is avoided without using tuning functions [8].

For the control of the partially known MIMO system Σ_2 (75), we make the following assumption as commonly being done in the literature.

Assumption 4.1: The signs of unknown constants g_{j,i_j} ($i_j = 1,\ldots,\rho_j$) are known, and there exist constants $\bar{g}_{j,i_j} \geq \underline{g}_{j,i_j} > 0$ such that $\bar{g}_{j,i_j} \geq |g_{j,i_j}| \geq \underline{g}_{j,i_j}$.

The above assumption implies that g_{j,i_j} are strictly either positive or negative constants. Without losing generality, we shall assume $\bar{g}_{j,i_j} \geq g_{j,i_j} \geq \underline{g}_{j,i_j}$.

For the jth subsystem of the MIMO nonlinear system (75), we design for each subsystem a full state feedback controller as follows. The closed-loop stability can be proved as a whole, which is similar to the proof of *Theorem 3.1*

$$u_j = -z_{j,\rho_j-1} - c_{j,\rho_j} z_{j,\rho_j} - \hat{W}^T_{j,\rho_j} S_{j,\rho_j}(Z_{j,\rho_j}) \tag{76}$$

with Coordinate transformation

$$z_{j,1} = x_{j,1} - x_{dj}, \quad z_{j,i_j+1} = x_{j,i_j+1} - \alpha_{j,i_j}, \quad 1 \leq i_j \leq \rho_j - 1. \tag{77}$$

Virtual control functions

$$\alpha_{j,1} = -c_{j,1} z_{j,1} - \hat{\eta}^T_{j,1} F_{\eta_{j,1}} - \hat{W}^T_{j,1} S_{j,1}(Z_{j,1}) \tag{78}$$

$$\alpha_{j,i_j} = -z_{j,i_j-1} - c_{j,i_j} z_{j,i_j} - \hat{\eta}^T_{j,i_j} F_{\eta_{j,i_j}} - \hat{W}^T_{j,i_j} S_{j,i_j}(Z_{j,i_j}). \tag{79}$$

NN inputs

$$Z_{j,1} = \left[\bar{x}^T_{j,(1-\varrho_{j1})},\ldots,\bar{x}^T_{j,(1-\varrho_{jm})}\right]^T \in \Omega_{j,1} \tag{80}$$

$$Z_{j,i_j} = \left[\bar{x}^T_{1,(i_j-\varrho_{j1})},\ldots,\bar{x}^T_{m,(i_j-\varrho_{jm})}, \left(\frac{\partial \alpha_{j,i_j-1}}{\partial \bar{x}_{1,(i_j-1-\varrho_{j1})}}\right)^T \ldots, \left(\frac{\partial \alpha_{j,i_j-1}}{\partial \bar{x}_{m,(i_j-1-\varrho_{jm})}}\right)^T\right]^T \in \Omega_{j,i_j} \tag{81}$$

$$1 \leq i_j \leq \rho_j - 1$$

$$Z_{j,\rho_j} = \left[X^T, u_1,\ldots,u_{j-1}, \left(\frac{\partial \alpha_{j,\rho_j-1}}{\partial \bar{x}_{1,(\rho_1-1)}}\right)^T \ldots, \left(\frac{\partial \alpha_{j,\rho_j-1}}{\partial \bar{x}_{m,(\rho_m-1)}}\right)^T\right]^T \in \Omega_{j,\rho_j}. \tag{82}$$

$$\Sigma_2 : \begin{cases} \dot{x}_{1,i_1} = f_{1,i_1}(\bar{x}_{1,(i_1-\varrho_{j1})},\ldots,\bar{x}_{m,(i_1-\varrho_{jm})}) + \theta^T_{1,i_1} F_{1,i_1}(\bar{x}_{1,(i_1-\varrho_{j1})},\ldots,\bar{x}_{m,(i_1-\varrho_{jm})}) \\ \qquad + G_{1,i_1}(\bar{x}_{1,(i_1-\varrho_{j1})},\ldots,\bar{x}_{m,(i_1-\varrho_{jm})}) + g_{1,i_1} x_{1,i_1+1} & 1 \leq i_1 \leq \rho_1 - 1 \\ \dot{x}_{1,\rho_1} = f_{1,\rho_1}(X) + \theta^T_{1,\rho_1} F_{1,\rho_1}(X) + G_{1,\rho_1}(X) + g_{1,\rho_1} u_1 \\ \cdots \\ \dot{x}_{j,i_j} = f_{j,i_j}(\bar{x}_{1,(i_j-\varrho_{j1})},\ldots,\bar{x}_{m,(i_j-\varrho_{jm})}) + \theta^T_{j,i_j} F_{j,i_j}(\bar{x}_{1,(i_j-\varrho_{j1})},\ldots,\bar{x}_{m,(i_j-\varrho_{jm})}) \\ \qquad + G_{j,i_j}(\bar{x}_{1,(i_j-\varrho_{j1})},\ldots,\bar{x}_{m,(i_j-\varrho_{jm})}) + g_{j,i_j} x_{j,i_j+1}, & 1 \leq i_j \leq \rho_j - 1 \\ \dot{x}_{j,\rho_j} = f_{j,\rho_j}(X,u_1,\ldots,u_{j-1}) + \theta^T_{j,\rho_j} F_{j,\rho_j}(X,u_1,\ldots,u_{j-1}) + G_{j,\rho_j}(X,u_1,\ldots,u_{j-1}) + g_{j,\rho_j} u_j \\ y_j = x_{j,1}, & 1 \leq j \leq m \end{cases} \tag{75}$$

Intermediate variable

$$\phi_{j,i_j-1} = \sum_{l=1}^{m}\sum_{k=1}^{i_j-1-\varrho_{jl}} \frac{\partial \alpha_{j,i_j-1}}{\partial x_{l,k}} G_{l,k} + \frac{\partial \alpha_{j,i_j-1}}{\partial x_d}\dot{x}_d + \sum_{k=1}^{i_j-1}\frac{\partial \alpha_{j,i_j-1}}{\partial \hat{\eta}_{j,k}}\Gamma_{\eta_{j,k}}\left[F_{\eta_{j,k}} z_{j,k} - \sigma_{\eta_{j,k}}\hat{\eta}_{j,k}\right] + \sum_{k=1}^{i_j-1}\frac{\partial \alpha_{j,i_j-1}}{\partial \hat{W}_{j,k}}\Gamma_{W_{j,k}} \times \left[S_{j,k}(Z_{j,k})z_{j,k} - \sigma_{W_{j,i_j}}\hat{W}_{j,k}\right] \quad 2 \le i_j \le \rho_j \tag{83}$$

and Adaptation laws

$$\dot{\hat{\eta}}_{j,i_j} = \dot{\tilde{\eta}}_{j,i_j} = \Gamma_{\eta_{j,i_j}}\left[F_{\eta_{j,i_j}} z_{j,i_j} - \sigma_{\eta_{j,i_j}}\hat{\eta}_{j,i_j}\right]$$
$$\dot{\hat{W}}_{j,i_j} = \dot{\tilde{W}}_{j,i_j} = \Gamma_{W_{j,i_j}}\left[S_{j,i_j}(Z_{j,i_j})z_{j,i_j} - \sigma_{W_{j,i_j}}\hat{W}_{j,i_j}\right] \quad i_j = 1,\ldots,\rho_j \tag{84}$$

where $c_{j,i_j} > 0$ are design constants, RBF NNs $W_{j,i_j}^{*T} S_{j,i_j}(Z_{j,i_j})$ are used to approximate the unknown functions in the controller design, with $\hat{W}_{j,i_j}$ being the estimates to W_{j,i_j}^*, $\eta_{j,i_j} = \left[\left(\theta_{j,i_j}^T/g_{j,i_j}\right), (1/g_{j,i_j})\right]^T$ is an unknown constant vector with $\hat{\eta}_{j,i_j}$ being the estimates to η_{j,i_j}, $F_{\eta_{j,i_j}} = \left[F_{j,i_j}^T, (f_{j,i_j} - \phi_{j,i_j-1})\right]^T$ is a known function vector, $\sigma_{\eta_{j,i_j}}, \sigma_{W_{j,i_j}} > 0$, $\Gamma_{W_{j,i_j}} = \Gamma_{W_{j,i_j}}^T > 0$, $\Gamma_{\eta_{j,i_j}} = \Gamma_{\eta_{j,i_j}}^T > 0$ are design constants, $\tilde{W}_{j,i_j} = \hat{W}_{j,i_j} - W_{j,i_j}^*$, $\tilde{\eta}_{j,i_j} = \hat{\eta}_{j,i_j} - \eta_{j,i_j}^*$.

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

Theorem 4.1: Consider the closed-loop system consisting of the plant (75), the reference model (31), the controller (76) and the NN weight updating laws (84). Then, for bounded initial conditions, there exists sufficiently large compact sets Ω_{j,i_j}, $i_j = 1,\ldots,\rho_j$, $j = 1,\ldots,m$ such that $Z_{j,i_j} \in \Omega_{j,i_j}$ for all $t \ge 0$; all signals in the closed-loop system remain bounded; and the tracking errors converge to a compact set whose size can be reduced by appropriately choosing design parameters.

Proof: Theorem 4.1 can be proved following the similar procedures of Theorem 3.1 for system Σ_1 in Section III. In the following, we will mainly describe the differences of the design procedures of Σ_2 from that of Σ_1.

Step 1: The derivative of $z_{j,1} = x_{j,1} - x_{d1}$ is

$$\dot{z}_{j,1} = f_{j,1}\left(\bar{x}_{1,(i_1-\varrho_{j1})},\ldots,\bar{x}_{m,(i_1-\varrho_{jm})}\right) + \theta_{j,1}^T F_{j,1}\left(\bar{x}_{1,(i_1-\varrho_{j1})},\ldots,\bar{x}_{m,(i_1-\varrho_{jm})}\right) + G_{j,1}\left(\bar{x}_{1,(i_1-\varrho_{j1})},\ldots,\bar{x}_{m,(i_1-\varrho_{jm})}\right) + g_{j,1}x_{j,2} - f_{d1}. \tag{85}$$

By viewing $x_{j,2}$ as a virtual control input, apparently there exists a desired feedback control

$$\alpha_{j,1}^* = -c_{j,1}z_{j,1} - \frac{1}{g_{j,1}}\left[f_{j,1} + \theta_{j,1}^T F_{j,1} + G_{j,1} - f_{d1}\right] = -c_{j,1}z_{j,1} - \eta_{j,1}^T F_{\eta_{j,1}} - \frac{1}{g_{j,1}} f_{j,1}$$

where $\eta_{j,1} \triangleq [(\theta_{j,1}^T/g_{j,1}), (1/g_{j,1})]^T$ is introduced as a new unknown constant vector, and $F_{\eta_{j,1}} \triangleq [F_{j,1}^T, (G_{j,1} - f_{d1})]^T$ as a known function vector. Note that because of the introduction of $\eta_{j,1}$, we only need to estimate $\eta_{j,1}$ rather than $g_{j,1}$.

For uniformity of presentation, let $h_{j,1}(Z_{j,1}) = (1/g_{j,1})f_{j,1}$ denote the unknown function of $\alpha_{j,1}^*$, where $Z_{j,1} = [\bar{x}_{1,(i_1-\varrho_{j1})}^T,\ldots,\bar{x}_{m,(i_1-\varrho_{jm})}^T]^T \in \Omega_{j,1}$. By employing a RBF neural network $W_{j,1}{}^T S_{j,1}(Z_{j,1})$ to approximate $h_{j,1}(Z_{j,1})$, $\alpha_{j,1}^*$ can be expressed as

$$\alpha_{j,1}^* = -c_{j,1}z_{j,1} - \eta_{j,1}^T F_{\eta_{j,1}} - W_{j,1}^{*\,T} S_{j,1}(Z_{j,1}) - \epsilon_{j,1}$$

.

Since $\eta_{j,1}$ and $W_{j,1}^*$ are unknown, $\alpha_{j,1}^*$ cannot be realized in practice. Choose $\alpha_{j,1}$ as in (78), we have

$$\dot{z}_{j,1} = g_{j,1}(z_{j,2} + \alpha_{j,1}) + \theta_{j,1}^T F_{j,1} + f_{j,1} + G_{j,1} - f_{d1} = g_{j,1}[z_{j,2} - c_{j,1}z_{j,1} - (\hat{\eta}_{j,1} - \eta_{j,1})^T F_{\eta_{j,1}} - (\hat{W}_{j,1} - W_{j,1}^*)^T S_{j,1}(Z_{j,1}) + \epsilon_{j,1}]. \tag{86}$$

Consider the Lyapunov function candidate

$$V_{j,1} = \frac{1}{2g_{j,1}}z_{j,1}^2 + \frac{1}{2}\tilde{\eta}_{j,1}^T\Gamma_{\eta_{j,1}}^{-1}\tilde{\eta}_{j,1} + \frac{1}{2}\tilde{W}_{j,1}^T\Gamma_{W_{j,1}}^{-1}\tilde{W}_{j,1}. \tag{87}$$

By using (86) and (84), and with some completion of squares and straightforward derivation similar to those employed in Section III, the derivative of $V_{j,1}$ becomes

$$\dot{V}_{j,1} < z_{j,1}z_{j,2} - c_{j,10}z_{j,1}^2 - \frac{\sigma_{\eta_{j,1}}\|\tilde{\eta}_{j,1}\|^2}{2} - \frac{\sigma_{W_{j,1}}\|\tilde{W}_{j,1}\|^2}{2} + \frac{\sigma_{\eta_{j,1}}\|\eta_{j,1}\|^2}{2} + \frac{\sigma_{W_{j,1}}\|W_{j,1}^*\|^2}{2} + \frac{\epsilon_{j,1}^{*\,2}}{4c_{j,11}} \tag{88}$$

where $c_{j,10} > 0$ is a design constant.

Step i_j $(2 \le i_j \le n-1)$: The derivative of $z_{j,i_j} = x_{j,i_j} - \alpha_{j,i_j-1}$ is

$$\dot{z}_{j,i_j} = f_{j,i_j}(\bar{x}_{1,(i_j-\varrho_{j1})},\ldots,\bar{x}_{m,(i_j-\varrho_{jm})}) + \theta_{j,i_j}^T F_{j,i_j}(\bar{x}_{1,(i_j-\varrho_{j1})},\ldots,\bar{x}_{m,(i_j-\varrho_{jm})}) + G_{j,i_j}(\bar{x}_{1,(i_j-\varrho_{j1})},\ldots,\bar{x}_{m,(i_j-\varrho_{jm})}) + g_{j,i_j}x_{j,i_j+1} - \dot{\alpha}_{j,i_j-1}. \tag{89}$$

By viewing x_{j,i_j+1} as a virtual control input to stabilize the $(z_{j,1},\ldots,z_{j,i_j})$-subsystem, there exists a desired feedback control

$$\alpha_{j,i_j}^* = -z_{j,i_j-1} - c_{j,i_j}z_{j,i_j} - \frac{1}{g_{j,i_j}}[f_{j,i_j} + \theta_{j,i_j}^T F_{j,i_j} + G_{j,i_j} - \dot{\alpha}_{j,i_j-1}] \tag{90}$$

where α_{j,i_j-1} is a function of $\bar{x}_{1,(i_j-1-\varrho_{j1})},\ldots,\bar{x}_{m,(i_j-1-\varrho_{jm})}$, x_d, $\hat{\eta}_{j,1},\ldots,\hat{\eta}_{j,i_j-1}$ and $\hat{W}_{j,1},\ldots,\hat{W}_{j,i_j-1}$. Thus, $\dot{\alpha}_{j,i_j-1}$ can be expressed as

$$\dot{\alpha}_{j,i_j-1} = \sum_{l=1}^{m}\sum_{k=1}^{i_j-1-\varrho_{jl}}\frac{\partial \alpha_{j,i_j-1}}{\partial x_{l,k}}\dot{x}_{l,k} + \frac{\partial \alpha_{j,i_j-1}}{\partial x_d}\dot{x}_d + \sum_{k=1}^{j,i_j-1}\frac{\partial \alpha_{j,i_j-1}}{\partial \hat{\eta}_{j,k}}\dot{\hat{\eta}}_{j,k} + \sum_{k=1}^{i_j-1}\frac{\partial \alpha_{j,i_j-1}}{\partial \hat{W}_{j,k}}\dot{\hat{W}}_{j,k} = \sum_{l=1}^{m}\sum_{k=1}^{i_j-1-\varrho_{jl}}\frac{\partial \alpha_{j,i_j-1}}{\partial x_{l,k}}[g_{l,k}x_{l,k+1} + f_{l,k} + \theta_{l,k}^T F_{l,k}] + \phi_{j,i_j-1} \tag{91}$$

with ϕ_{j,i_j-1} as given in (83). Then, α^*_{j,i_j} is

$$\alpha^*_{j,i_j} = -z_{j,i_j-1} - c_{j,i_j}z_{j,i_j} - \eta^T_{j,i_j}F_{\eta_{j,i_j}} - \frac{1}{g_{j,i_j}} \times \left[f_{j,i_j} - \sum_{l=1}^{m}\sum_{k=1}^{i_j-1-\varrho_{jl}} \frac{\partial \alpha_{j,i_j-1}}{\partial x_{l,k}} \times (g_{l,k}x_{l,k+1} + \theta^T_{l,k}F_{l,k} + f_{l,k})\right] \quad (92)$$

where $\eta_{j,i_j} \triangleq [(\theta^T_{j,i_j}/g_{j,i_j}), (1/g_{j,i_j})]^T$ is an unknown constant vector, $F_{\eta_{j,i_j}} \triangleq [F^T_{j,i_j}, (G_{j,i_j} - \phi_{j,i_j-1})]^T$ is a known function vector.

Let

$$h_{j,i_j}(Z_{j,i_j}) = \frac{1}{g_{j,i_j}}\left[f_{j,i_j}(\bar{x}_{j,i_j}) - \sum_{l=1}^{m}\sum_{k=1}^{i_j-1-\varrho_{jl}} \frac{\partial \alpha_{j,i_j-1}}{\partial x_{l,k}}\left(g_{l,k}x_{l,k+1} + \theta^T_{l,k}F_{l,k} + f_{l,k}\right)\right] \quad (93)$$

denote the unknown function in α^*_{j,i_j}, where Z_{j,i_j} is given in (80). By employing a RBF neural network $W^T_{j,i_j}S_{j,i_j}(Z_{j,i_j})$ to approximate $h_{j,i_j}(Z_{j,i_j})$, α^*_{j,i_j} can be expressed as

$$\alpha^*_{j,i_j} = -z_{j,i_j-1} - c_{j,i_j}z_{j,i_j} - \eta^T_{j,i_j}F_{\eta_{j,i_j}} - W^{*T}_{j,i_j}S_{j,i_j}(Z_{j,i_j}) - \epsilon_{j,i_j}$$

Remark 4.1: Note that in this step, the parametric uncertainties $g_{j,1},\ldots g_{j,i_j-1}$ and $\theta_{j,1},\ldots \theta_{j,i_j-1}$ have been combined into the unknown term $h_{j,i_j}(Z_{j,i_j})$ which is to be approximated by neural network $W^T_{j,i_j}S_{j,i_j}(Z_{j,i_j})$. By doing so, there is no need to estimate $g_{j,1},\ldots,g_{j,i_j-1}$ and $\theta_{j,1},\ldots,\theta_{j,i_j-1}$ repeatedly in the controller, which is known as overparametrization [7], or to use tuning functions to deal with the partial derivative terms [8]. Therefore, the overparametrization problem in adaptive backstepping design is avoided without using tuning functions.

Since η_{j,i_j} and W^*_{j,i_j} are unknown, α^*_{j,i_j} cannot be realized in practice. Choose α_{j,i_j} as in (79), we have

$$\begin{aligned}\dot{z}_{j,i_j} &= g_{j,i_j}(z_{j,i_j+1} + \alpha_{j,i_j}) + f_{j,i_j} + \theta^T_{j,i_j}F_{j,i_j} + G_{j,i_j} - \dot{\alpha}_{j,i_j-1} \\ &= g_{j,i_j}\Big[z_{j,i_j+1} - z_{j,i_j-1} - c_{j,i_j}z_{j,i_j} - \tilde{\eta}^T_{j,i_j}F_{\eta_{j,i_j}} - \tilde{W}^T_{j,i_j}S_{j,i_j}(Z_{j,i_j}) + \epsilon_{j,i_j}\Big].\end{aligned} \quad (94)$$

Consider the Lyapunov function candidate

$$V_{j,i_j} = V_{j,i_j-1} + \frac{1}{2g_{j,i_j}}z^2_{j,i_j} + \frac{1}{2}\tilde{\eta}^T_{j,i_j}\Gamma^{-1}_{\eta_{j,i_j}}\tilde{\eta}_{j,i_j} + \frac{1}{2}\tilde{W}^T_{j,i_j}\Gamma^{-1}_{W_{j,i_j}}\tilde{W}_{j,i_j}. \quad (95)$$

By using (94) and (84), and with some completion of squares and straightforward derivation similar to those employed in Section III, the derivative of V_{j,i_j} becomes

$$\dot{V}_{j,i_j} < z_{j,i_j}z_{j,i_j+1} - \sum_{k=1}^{i_j}c_{j,k0}z^2_{j,k} - \sum_{k=1}^{i_j}\frac{\sigma_{\eta_{j,k}}\|\tilde{\eta}_{j,k}\|^2}{2} - \sum_{k=1}^{i_j}\frac{\sigma_{W_{j,k}}\|\tilde{W}_{j,k}\|^2}{2} + \sum_{k=1}^{i_j}\frac{\sigma_{\eta_{j,k}}\|\eta_{j,k}\|^2}{2} + \sum_{k=1}^{i_j}\frac{\sigma_{W_{j,k}}\|W^*_{j,k}\|^2}{2} + \sum_{k=1}^{i_j}\frac{{\epsilon^*_{j,k}}^2}{4c_{k1}}. \quad (96)$$

Step ρ_j: This is the final step. The derivative of z_{j,ρ_j} is

$$\begin{aligned}\dot{z}_{j,\rho_j} &= g_{j,\rho_j}u + f_{j,\rho_j}(X, u_1,\ldots,u_{j-1}) + \theta^T_{j,\rho_j}F_{j,\rho_j}(X, u_1,\ldots,u_{j-1}) \\ &\quad + G_{j,\rho_j}(X, u_1,\ldots,u_{j-1}) - \dot{\alpha}_{j,\rho_j-1}.\end{aligned} \quad (97)$$

To stabilize the $(z_{j,1},\ldots,z_{j,\rho_j})$ subsystem, there exists a desired feedback control

$$u^* = -z_{j,\rho_j-1} - c_{j,\rho_j}z_{j,\rho_j} - \frac{1}{g_{j,\rho_j}}\left[f_{j,\rho_j} + \theta^T_{j,\rho_j}F_{j,\rho_j} + G_{j,\rho_j} - \dot{\alpha}_{j,\rho_j-1}\right] \quad (98)$$

where α_{j,ρ_j-1} is a function of $\bar{x}_{1,(\rho_1-1)},\ldots,\bar{x}_{m,(\rho_m-1)}$, x_d, $\hat{\eta}_{j,1},\ldots,\hat{\eta}_{j,\rho_j-1}$ and $\hat{W}_{j,1},\ldots,\hat{W}_{j,\rho_j-1}$. Thus, $\dot{\alpha}_{j,\rho_j-1}$ can be expressed as

$$\begin{aligned}\dot{\alpha}_{j,\rho_j-1} &= \sum_{l=1}^{m}\sum_{k=1}^{\rho_l-1}\frac{\partial \alpha_{j,\rho_j-1}}{\partial x_{l,k}}\dot{x}_{l,k} + \frac{\partial \alpha_{j,\rho_j-1}}{\partial x_d}\dot{x}_d + \sum_{k=1}^{\rho_j-1}\frac{\partial \alpha_{j,\rho_j-1}}{\partial \hat{\eta}_{j,k}}\dot{\hat{\eta}}_{j,k} + \sum_{k=1}^{\rho_l-1}\frac{\partial \alpha_{j,\rho_j-1}}{\partial \hat{W}_{j,k}}\dot{\hat{W}}_{j,k} \\ &= \sum_{l=1}^{m}\sum_{k=1}^{\rho_l-1}\frac{\partial \alpha_{j,\rho_j-1}}{\partial x_{l,k}}\left[g_{l,k}x_{l,k+1} + f_{l,k} + \theta^T_{l,k}F_{l,k}\right] + \phi_{j,\rho_j-1}\end{aligned} \quad (99)$$

with ϕ_{j,ρ_j-1} given in (83). Then, u^* is given by

$$u^* = -z_{j,\rho_j-1} - c_{j,\rho_j}z_{j,\rho_j} - \eta^T_{j,\rho_j}F_{\eta_{j,\rho_j}} - \frac{1}{g_{j,\rho_j}} \times \left[f_{j,\rho_j} - \sum_{l=1}^{m}\sum_{k=1}^{\rho_l-1}\frac{\partial \alpha_{j,\rho_j-1}}{\partial x_{l,k}} \times (g_{l,k}x_{l,k+1} + f_{l,k} + \theta^T_{l,k}F_{l,k})\right] \quad (100)$$

where $\eta_{j,\rho_j} \triangleq \left[\theta^T_{j,\rho_j}/g_{j,\rho_j}, 1/g_{j,\rho_j}\right]^T$ is an unknown constant vector, $F_{\eta_{j,\rho_j}} \triangleq \left[F^T_{j,\rho_j}, (G_{j,\rho_j} - \phi_{j,\rho_j-1})\right]^T$ is a known function vector.

For the desired feedback control u^*, let

$$h_{j,\rho_j}(Z_{j,\rho_j}) \triangleq \frac{1}{g_{j,\rho_j}}\Bigg[f_{j,\rho_j} - \sum_{l=1}^{m}\sum_{k=1}^{\rho_l-1}\frac{\partial \alpha_{j,\rho_j-1}}{\partial x_{l,k}}(g_{l,k}x_{l,k+1}+f_{l,k}+\theta_{l,k}^T F_{l,k})\Bigg] \quad (101)$$

where Z_{j,ρ_j} is given in (82). By employing a RBF neural network $W_{j,\rho_j}^T S_{j,\rho_j}(Z_{j,\rho_j})$ to approximate $h_{j,\rho_j}(Z_{j,\rho_j})$, u^* can be expressed as

$$u^* = -z_{j,\rho_j-1} - c_{j,\rho_j} z_{j,\rho_j} - \eta_{j,\rho_j}^T F_{\eta_{j,\rho_j}} - W_{j,\rho_j}^{*T} S_{j,\rho_j}(Z_{j,\rho_j}) - \epsilon_{j,\rho_j}.$$

Choose the adaptive neural controller as in (76), we have

$$\begin{aligned}\dot{z}_{j,\rho_j} &= g_{j,\rho_j} u + f_{j,\rho_j} + \theta_{j,\rho_j}^T F_{j,\rho_j} + G_{j,\rho_j} - \dot{\alpha}_{j,\rho_j-1}\\ &= g_{j,\rho_j}\Big[-z_{j,\rho_j-1} - c_{j,\rho_j} z_{j,\rho_j} - \tilde{\eta}_{j,\rho_j}^T F_{\eta_{j,\rho_j}} - \tilde{W}_{j,\rho_j}^T S_{j,\rho_j}(Z_{j,\rho_j}) + \epsilon_{j,\rho_j}\Big].\end{aligned} \quad (102)$$

Consider the Lyapunov function candidate

$$V_{j,\rho_j} = V_{j,\rho_j-1} + \frac{1}{2g_{j,\rho_j}} z_{j,\rho_j}^2 + \frac{1}{2}\tilde{\eta}_{j,\rho_j}^T \Gamma_{\eta_{j,\rho_j}}^{-1}\tilde{\eta}_{j,\rho_j} + \frac{1}{2}\tilde{W}_{j,\rho_j}^T \Gamma_{W_{j,\rho_j}}^{-1}\tilde{W}_{j,\rho_j}. \quad (103)$$

By using (102) and (84), and with some completion of squares and straightforward derivation similar to those employed in Section III, the derivative of V_{j,ρ_j} becomes

$$\dot{V}_{j,\rho_j} < -\sum_{k=1}^{\rho_j} c_{j,k0} z_{j,k}^2 - \sum_{k=1}^{\rho_j}\frac{\sigma_{\eta_{j,k}}\|\tilde{\eta}_{j,k}\|^2}{2} - \sum_{k=1}^{\rho_j}\frac{\sigma_{W_{j,k}}\|\tilde{W}_{j,k}\|^2}{2} + \delta_j \quad (104)$$

where $\delta_j \triangleq \sum_{k=1}^{\rho_j}(\sigma_{\eta_{j,k}}\|\eta_{j,k}\|^2/2) + \sum_{k=1}^{\rho_j}(\sigma_{W_{j,k}}\|W_{j,k}^*\|^2/2) + \sum_{k=1}^{\rho_j}({\epsilon_{j,k}^*}^2/4c_{k1})$.

If we choose $c_{j,k0}$ such that $c_{j,k0} \geq \gamma_j/2\underline{g}_{j,k}$, $k=1,\ldots,\rho_j$, where γ_j is a positive constant, and choose $\sigma_{\eta_{j,k}}$, $\sigma_{W_{j,k}}$ and $\Gamma_{\eta_{j,k}}$, $\Gamma_{W_{j,k}}$ such that $\sigma_{\eta_{j,k}} \geq \gamma_j \lambda_{\max}\left\{\Gamma_{\eta_{j,k}}^{-1}\right\}$, $\sigma_{W_{j,k}} \geq \gamma_j\lambda_{\max}\left\{\Gamma_{W_{j,k}}^{-1}\right\}$, $k=1,\ldots,\rho_j$, then from (104) we have the following

$$\begin{aligned}\dot{V}_{j,\rho_j} &< -\sum_{k=1}^{\rho_j} c_{j,k0} z_{j,k}^2 - \sum_{k=1}^{\rho_j}\frac{\sigma_{\eta_{j,k}}\|\tilde{\eta}_{j,k}\|^2}{2} - \sum_{k=1}^{\rho_j}\frac{\sigma_{W_{j,k}}\|\tilde{W}_{j,k}\|^2}{2} + \delta_j\\ &\leq -\gamma_j\left[\sum_{k=1}^{\rho_j}\frac{1}{2g_{j,k}}z_{j,k}^2 + \sum_{k=1}^{\rho_j}\frac{\gamma_j\tilde{\eta}_{j,k}^T\Gamma_{\eta_{j,k}}^{-1}\tilde{\eta}_{j,k}}{2} + \sum_{k=1}^{\rho_j}\frac{\gamma_j\tilde{W}_{j,k}^T\Gamma_{W_{j,k}}^{-1}\tilde{W}_{j,k}}{2}\right] + \delta_j\\ &\leq -\gamma_j V_{j,\rho_j} + \delta_j.\end{aligned} \quad (105)$$

Similar to the proof of Theorem 3.1, it can be shown that: i) all the variables in the closed-loop, including the states x_{j,i_j}, the parameter estimates $\hat{\eta}_{j,i_j}$, the weight estimates $\hat{W}_{j,i_j}$ and the control u_j, $j=1,\ldots,m$, $i_j=1,\ldots,\rho_j$, remain uniformly bounded and ii) the states and parameter estimation errors converge to compact sets whose sizes can be reduced by choosing appropriate design parameters. ◇

Remark 4.2: Compare the controller design for systems Σ_1 and Σ_2, we notice that the dimensions of the input variables Z_{j,i_j} of NNs in (80)–(82) for system Σ_2 are one less than that in (39), (50), (61) for system Σ_1, respectively. Therefore, due to the exponentially increasing number of NN nodes with respect to NN inputs, much less neurons may be needed to approximate the unknown function $h_{j,i_j}(Z_{j,i_j})$ for the adaptive neural control of system Σ_2. This can be regarded as one benefit obtained from the *a priori* information and the affine terms g_{j,i_j} in system Σ_2.

On the other hand, the complicated backstepping controller, which is mainly caused by the analytical derivation using tuning functions [8], is much simplified. With the help of NN approximation, the parametric uncertainties $g_{j,1},\ldots,g_{j,i_j-1}$ and $\theta_{j,1},\ldots,\theta_{j,i_j-1}$ in system Σ_2 have been combined into the unknown term $h_{j,i_j}(Z_{j,i_j})$, which can be approximated by neural network $W_{j,i_j}^T S_{j,i_j}(Z_{j,i_j})$. By doing so, there is no need to estimate $g_{j,1},\ldots,g_{j,i_j-1}$ and $\theta_{j,1},\ldots,\theta_{j,i_j-1}$ repeatedly in the controller design, which is known as overparametrization [7], or to use tuning functions to deal with the partial derivative terms [8].

V. Simulation Studies

To verify the effectiveness of the proposed approach, the developed adaptive NN controller is applied to the following fairly complicated MIMO nonlinear system: (see (106) at the bottom of the page). The reference model is taken as the famous van der Pol oscillator [40]

$$\begin{cases}\dot{x}_{d1} = x_{d2}\\ \dot{x}_{d2} = -x_{d1} + \beta(1-x_{d1}^2)x_{d2}\\ y_{dj} = x_{dj}, \qquad j=1,2.\end{cases} \quad (107)$$

As shown in [40], the phase-plane trajectories of the van der Pol oscillator, starting from an initial state other than $(0,0)$, approach a limit cycle when $\beta > 0$.

$$\Sigma_{S_1}: \begin{cases}\dot{x}_{1,1} = 0.5(x_{1,1}+x_{2,1}) + \left(1+0.1x_{1,1}^2x_{2,1}^2\right)x_{1,2}\\ \dot{x}_{1,2} = (x_{1,1}x_{1,2}+x_{2,1}x_{2,2}) + [2+\cos(x_{1,1}x_{2,1})]u_1\\ \dot{x}_{2,1} = x_{1,1}x_{2,1} + [2+\sin(x_{1,1}x_{2,1})]x_{2,2}\\ \dot{x}_{2,2} = (x_{1,1}x_{1,2}+x_{2,1}x_{2,2}+u_1)^2 + (e^{x_{1,1}}+e^{-x_{2,1}})u_2\\ y_j = x_{j,1}, \qquad j=1,2.\end{cases} \quad (106)$$

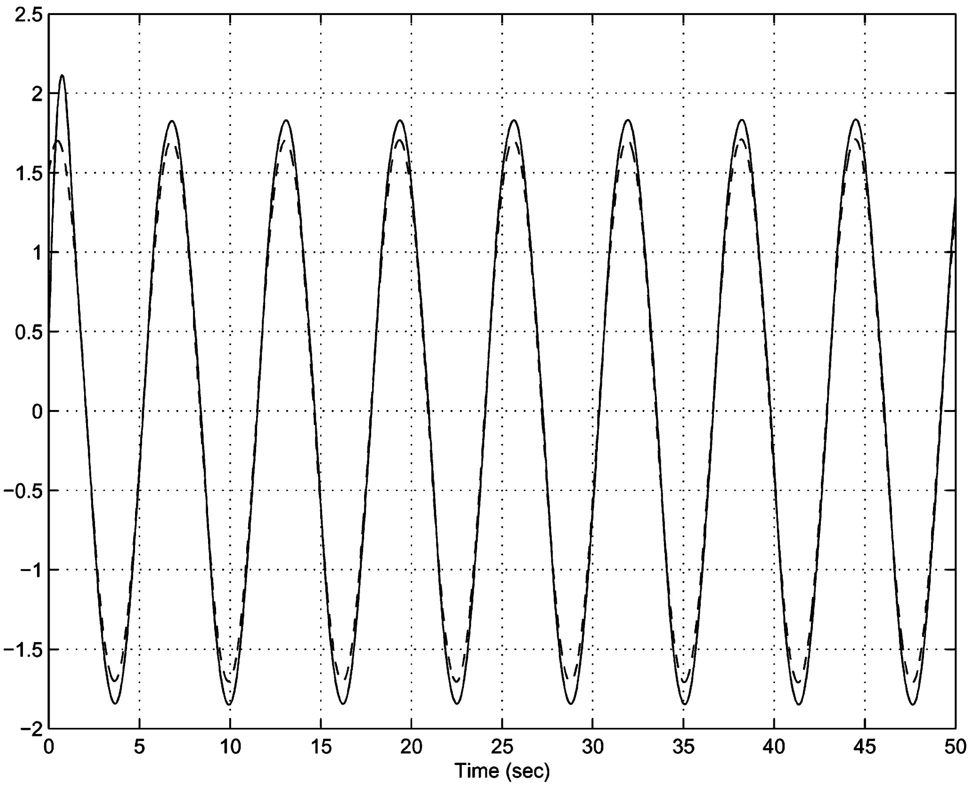

Fig. 2. Output y_1("-") follows reference y_{d1}("–").

The control objective is to design controller for system Σ_{S_1} such that: i) all the signals in the closed-loop system remain bounded, and ii) the output y_j, $j = 1, 2$ of system Σ_{S_1} follows the desired trajectory y_{dj}, $j = 1, 2$ generated from the van der Pol oscillator.

Clearly, system Σ_{S_1} is in the block-triangular form (28) and satisfy Assumptions 2–3. As system Σ_{S_1} consists of two second order subsystems ($\rho_1 = \rho_2 = 2$), the adaptive NN controllers for both subsystems are chosen according to (62) as follows:

$$u_j = -z_{j,1} - c_{j,2}z_{j,2} + \hat{W}_{j,2}^T S_{j,2}(Z_{j,2}), \quad j = 1, 2. \quad (108)$$

where $z_{j,1} = x_{j,1} - y_{dj}$, $z_{j,2} = x_{j,2} - \alpha_{j,1}$, $j = 1, 2$ and

$$Z_{1,2} = \left[x_{1,1}, x_{1,2}, x_{2,1}, x_{2,2}, \frac{\partial \alpha_{1,1}}{\partial x_{1,1}}, \frac{\partial \alpha_{1,1}}{\partial x_{2,1}}, \phi_{1,1}\right]^T \in R^7$$

$$Z_{2,2} = \left[x_{1,1}, x_{1,2}, x_{2,1}, x_{2,2}, u_1, \frac{\partial \alpha_{2,1}}{\partial x_{1,1}}, \frac{\partial \alpha_{2,1}}{\partial x_{2,1}}, \phi_{2,1}\right]^T \in R^8$$

with

$$\alpha_{1,1} = -c_{1,1}z_{1,1} + \hat{W}_{1,1}^T S_{1,1}(Z_{1,1}),$$
$$Z_{1,1} = [x_{1,1}, x_{2,1}, \dot{x}_{d1}]^T \in R^3$$
$$\phi_{1,1} = \frac{\partial \alpha_{1,1}}{\partial x_{d1}}\dot{x}_{d1} + \frac{\partial \alpha_{1,1}}{\partial x_{d2}}\dot{x}_{d2} + \frac{\partial \alpha_{1,1}}{\partial \hat{W}_{1,1}}\dot{\hat{W}}_{1,1}$$
$$\alpha_{2,1} = -c_{2,1}z_{2,1} - \hat{W}_{2,1}^T S_{2,1}(Z_{2,1}),$$
$$Z_{2,1} = [x_{1,1}, x_{2,1}, \dot{x}_{d2}]^T \in R^3$$
$$\phi_{2,1} = \frac{\partial \alpha_{2,1}}{\partial x_{d1}}\dot{x}_{d1} + \frac{\partial \alpha_{2,1}}{\partial x_{d2}}\dot{x}_{d2} + \frac{\partial \alpha_{2,1}}{\partial \hat{W}_{2,1}}\dot{\hat{W}}_{2,1}$$

and NN weights $\hat{W}_{j,i_j}$, $j = 1, 2$, $i_j = 1, 2$ are updated by (55) as

$$\dot{\hat{W}}_{j,i_j} = \Gamma_{W_{j,i_j}}\left[S_{j,i_j}(Z_{j,i_j})z_{j,i_j} - \sigma_{j,i_j}\hat{W}_{j,i_j}\right]. \quad (109)$$

In practice, the selection of the centers and widths of RBF has a great influence on the performance of the designed controller. According to [11], Gaussian RBF NNs arranged on a regular lattice on R^n can uniformly approximate sufficiently smooth functions on closed, bounded subsets. Accordingly, in the following simulation studies, the centers and widths are chosen on a regular lattice in the respective compact sets. Specifically, neural networks $\hat{W}_{1,1}^T S_{1,1}(Z_{1,1})$ contains 27 nodes (i.e., $l_{1,1} = 27$), with centers $\mu_l(l = 1, \ldots, l_{1,1})$ evenly spaced in $[-2.5, 2.5] \times [-2.5, 2.5] \times [-2, 2]$, and widths $\eta_l = 3$ ($l = 1, \ldots, l_{1,1}$). Neural networks $\hat{W}_{1,2}^T S_{1,2}(Z_{1,2})$ contains 2187 nodes (i.e., $l_{1,2} = 2187$), with centers $\mu_l(l = 1, \ldots, l_{1,2})$ evenly spaced in $[-2.5, 2.5] \times [-1.5, 1.5] \times [-2.5, 2.5] \times [-1.5, 1.5] \times [-1.9, -1.6] \times [-0.2, 0.1] \times [-4, 4]$, and widths $\eta_l = 5$ ($l = 1, \ldots, l_{1,2}$). As the number of nodes increases, it is desirable to assign the centers automatically rather than mannually. This can be easily done using nested multiloops similar to the assignments of μ_C on page 361 of [4]. Neural networks $\hat{W}_{2,1}^T S_{2,1}(Z_{2,1})$ contains 27 nodes (i.e., $l_{2,1} = 27$), with centers μ_l ($l = 1, \ldots, l_{2,1}$) evenly spaced in $[-2.5, 2.5] \times [-2.5, 2.5] \times [-2.5, 2.5]$ and widths $\eta_l = 3$ ($l = 1, \ldots, l_{2,1}$). Neural networks $\hat{W}_{2,2}^T S_{2,2}(Z_{2,2})$ contains 6567 nodes (i.e., $l_{2,2} = 6567$), with centers $\mu_l(l = 1, \ldots, l_{2,2})$ evenly spaced in $[-2.5, 2.5] \times [-1.5, 1.5] \times [-2.5, 2.5] \times [-1.5, 1.5] \times [-2, 2] \times [-0.3, 0] \times [-1.7, -1.4] \times [-4, 4]$, and widths $\eta_l = 8(l = 1, \ldots, l_{2,2})$. The design parameters of the

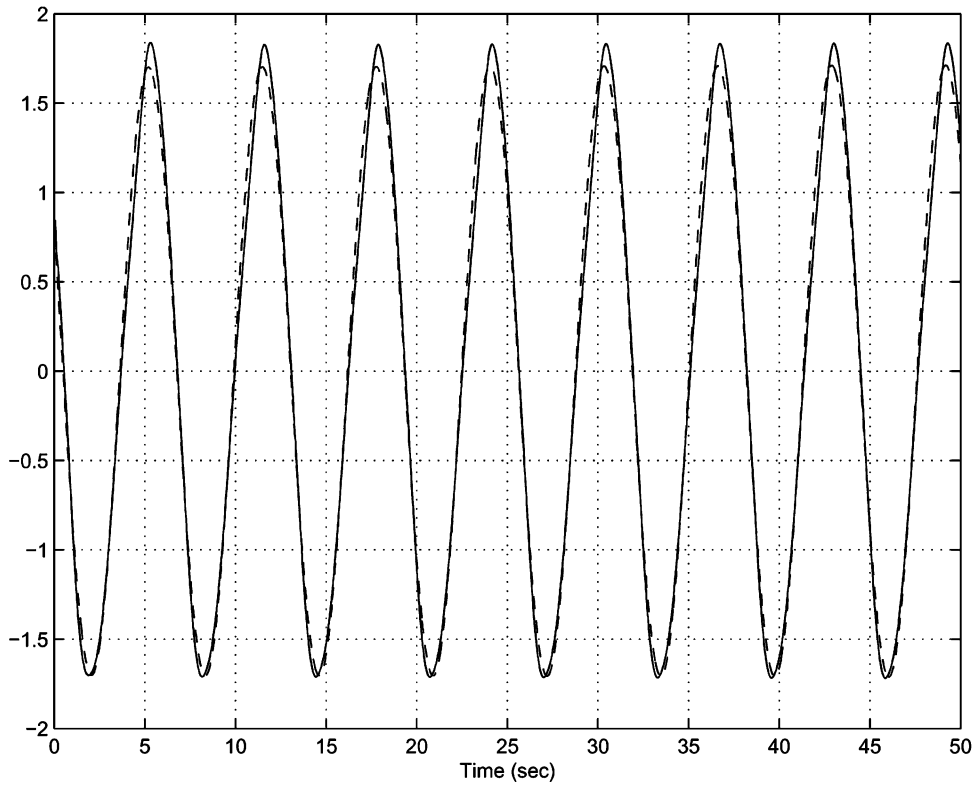

Fig. 3. Output y_2("-") follows reference y_{d2}("–").

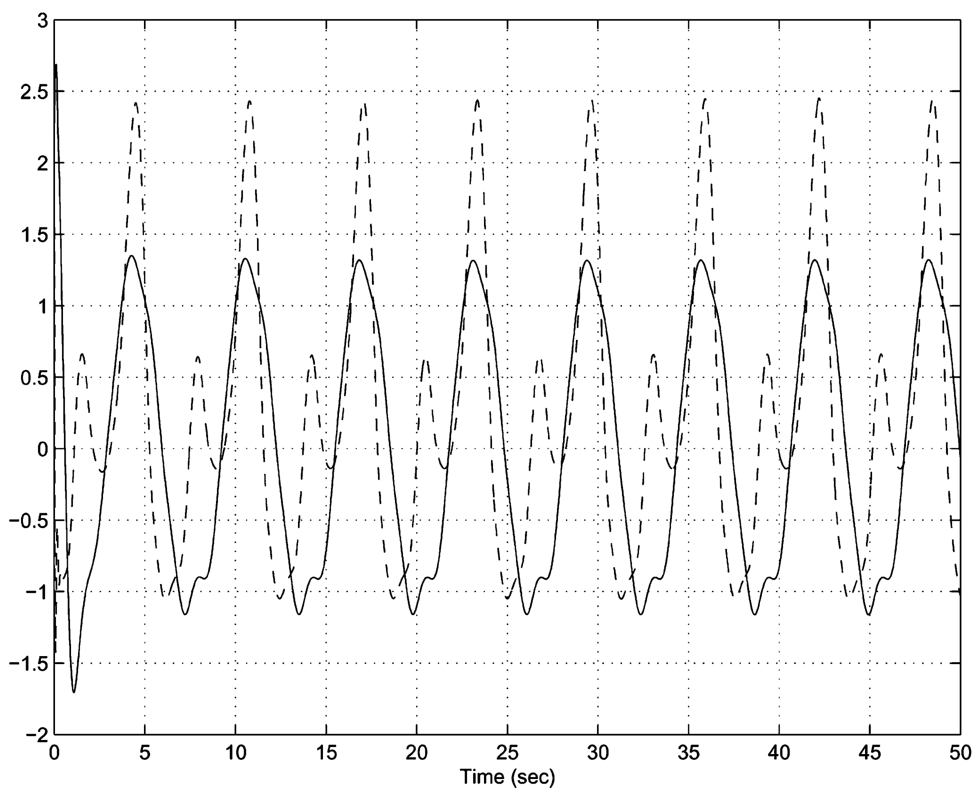

Fig. 4. States $x_{1,2}$("-") and $x_{2,2}$("–").

above controller are $c_{1,1} = 2.5$, $c_{1,2} = 10$, $c_{2,1} = 6$, $c_{2,2} = 20$, $\Gamma_{1,1} = \Gamma_{1,2} = \Gamma_{2,1} = \Gamma_{2,2} = \mathrm{diag}\{2\}$, $\sigma_{1,1} = \sigma_{1,2} = \sigma_{2,1} = \sigma_{2,2} = 0.1$. The initial conditions $[x_{1,1}(0), x_{1,2}(0), x_{2,1}(0), x_{2,2}(0)]^T = [0.5, 2, 0.7, 1]^T$ and $[x_{d1}(0), x_{d2}(0)]^T = [1.5, 0.8]^T$. The initial weights $\hat{W}_{1,1}(0) = \hat{W}_{1,2}(0) = \hat{W}_{2,1}(0) = \hat{W}_{2,2}(0) = 0$.

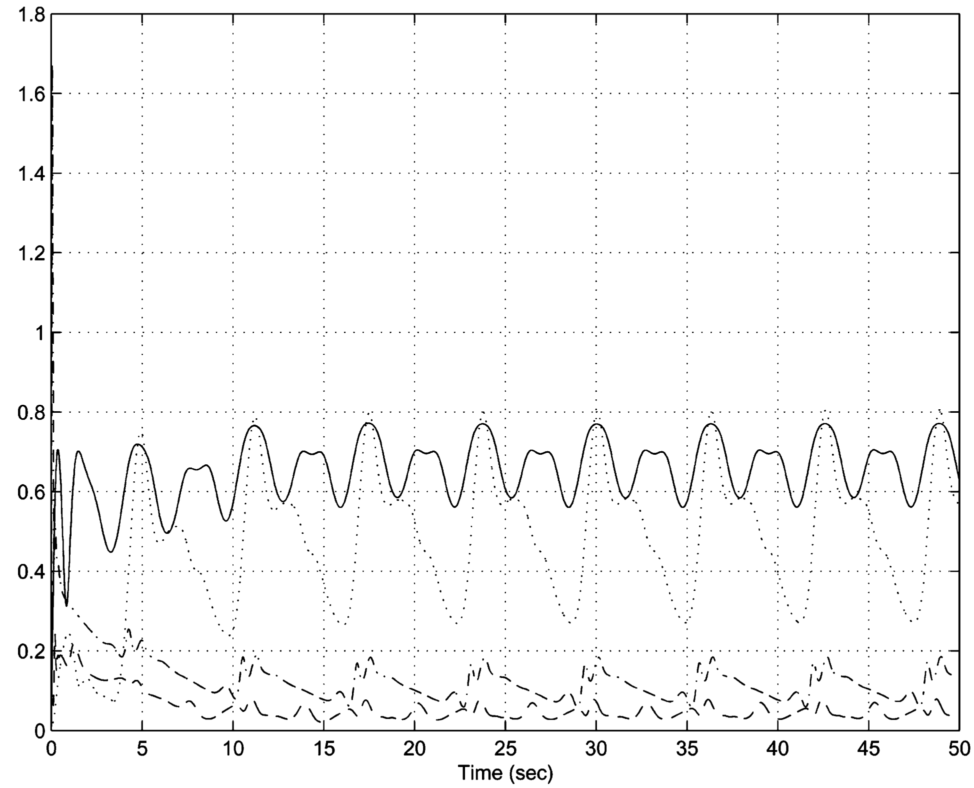

Fig. 5. L_2 norms of the NN weights: $\hat{W}_{1,1}$("-"), $\hat{W}_{1,2}$("--"), $\hat{W}_{2,1}$("$\cdots$"), $\hat{W}_{2,2}$("-").

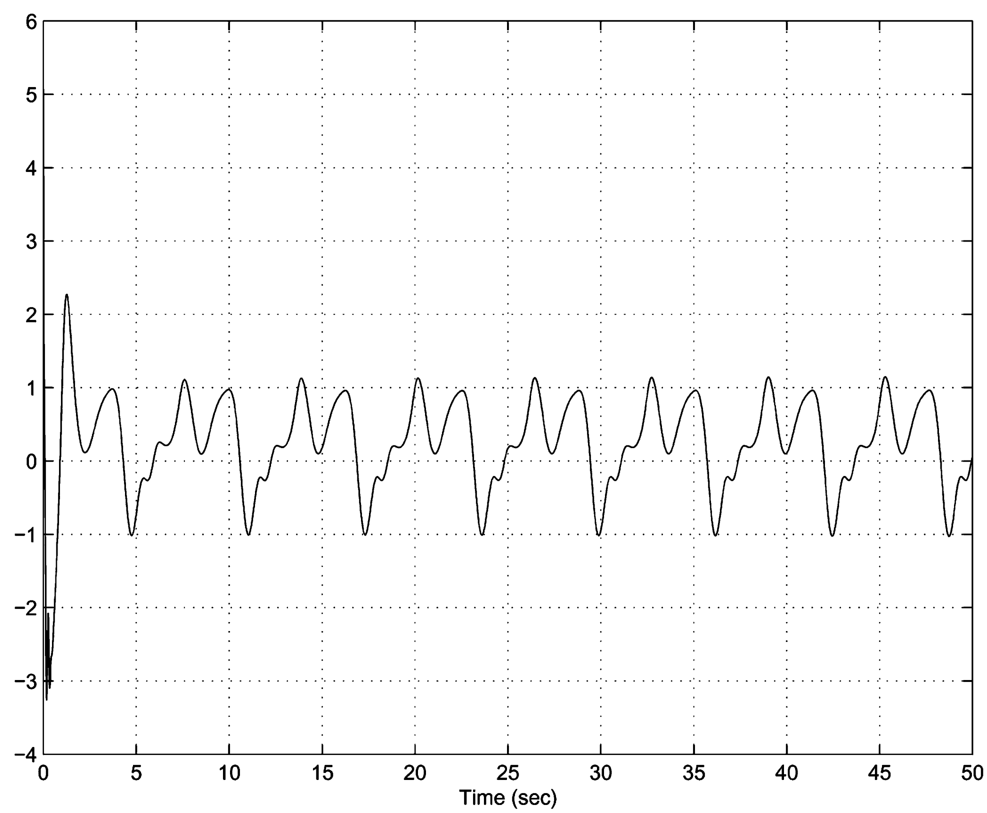

Fig. 6. Control torque u_1.

Figs. 2–7 show the simulation results of applying controller (108) to system Σ_{S_1} (106) for tracking desired signal y_{dj}, $j = 1, 2$ with $\beta = 0.001$. From Figs. 1 and 2, it can be seen that fairly good tracking performance is obtained. The boundedness of other system states $x_{1,2}, x_{2,2}$, NN weights $\hat{W}_{1,1}, \hat{W}_{1,2}, \hat{W}_{2,1}, \hat{W}_{2,2}$ and control signals u_1, u_2 are shown in Figs. 3–6, respec-

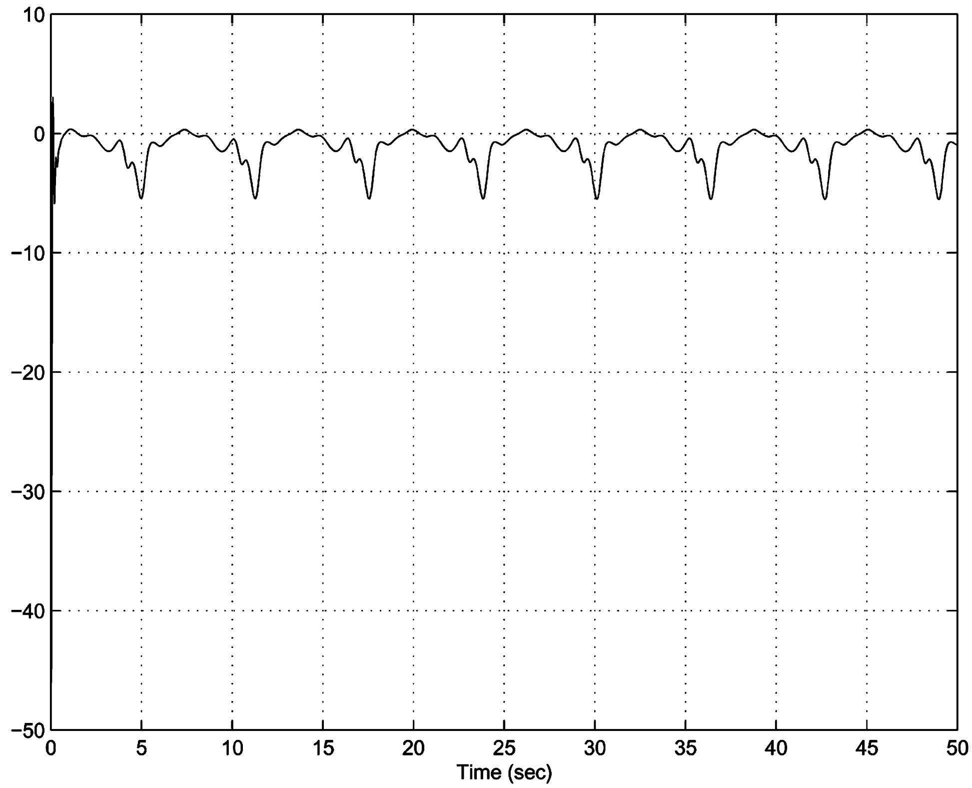

Fig. 7. Control torque u_2.

tively. As for each input variable to the neural network, only three nodes are selected, it is by any standard very sparse and rough. Control performance will become better if the size of the neural network is increased.

Simulation study for system Σ_{S_2} can be conducted similarly, and it is ommitted here for clarity. However, as the system order increases, there will be more inputs to the neural networks. For example, according to (50), there are 12 inputs for the neural network $\hat{W}_{2,4}^T S_{2,4}(Z_{2,4})$ in total. Suppose that there are at least three evenly spaced centers for each input, the neural network $\hat{W}_{2,4}^T S_{2,4}(Z_{2,4})$ contains at least $3^{12}(= 531\,441)$ nodes, which requires large computational power and is a restriction for real-time implementation. Further studies are needed to be carried to solve this problem.

VI. Conclusion

In this paper, adaptive neural control schemes have been proposed for two classes of uncertain MIMO nonlinear systems in block-triangular forms. By exploiting the special structure properties of the two classes of MIMO systems Σ_1 and Σ_2, the developed schemes avoid the controller singularity problem completely without using projection algorithms. With the help of NNs to approximate all the uncertain nonlinear functions in the controllers design, the developed schemes achieve semiglobal uniform ultimate boundedness of all the signals in the closed-loop of MIMO nonlinear systems. The outputs of the systems are proven to converge to small neighborhoods of the desired trajectories. The control performance of the closed-loop system is guaranteed by suitably choosing the design parameters. The proposed scheme can be applied to uncertain MIMO nonlinear systems without repeating the complex controller design procedure for different system nonlinearities.

Acknowledgment

The authors are very grateful to the many constructive comments from the anonymous Associate Editor and reviewers. The authors also thank F. Hong for her help in doing the simulation studies.

References

[1] K. S. Narendra and K. Parthasarathy, "Identification and control of dynamic systems using neural networks," *IEEE Trans. Neural Networks*, vol. 1, pp. 4–27, 1990.
[2] M. M. Gupta and D. H. Rao, *Neuro-Control Systems: Theory and Applications*. New York: IEEE Neural Networks Council, 1994.
[3] F. L. Lewis, S. Jagannathan, and A. Yeildirek, *Neural Network Control of Robot Manipulators and Nonlinear Systems*. London, U.K.: Taylor Francis, 1999.
[4] S. S. Ge, T. H. Lee, and C. J. Harris, *Adaptive Neural Network Control of Robotic Manipulators*. London, U.K.: World Scientific, 1998.
[5] S. S. Ge, C. C. Hang, T. H. Lee, and T. Zhang, *Stable Adaptive Neural Network Control*. Norwell, MA: Kluwer, 2001.
[6] S. G. Fabri and V. Kadirkamanathan, *Functional Adaptive Control: An Intelligent Systems Approach*. London, U.K.: Springer-Verlag, 2001.
[7] I. Kanellakopoulos, P. V. Kokotovic, and A. S. Morse, "Systematic design of adaptive controller for feedback linearizable systems," *IEEE Trans. Automat. Contr.*, vol. 36, pp. 1241–1253, 1991.
[8] M. Krstic, I. Kanellakopoulos, and P. V. Kokotovic, "Adaptive nonlinear control without overparameterization," *Syst. Contr. Lett.*, vol. 19, pp. 177–185, 1992.
[9] M. Krstić, I. Kanellakopoulos, and P. Kokotović, *Nonlinear and Adaptive Control Design*. New York: Wiley, 1995.
[10] M. M. Polycarpou and P. A. Ioannou, "Identification and Control Using Neural Network Models: Design and Stability Analysis," Dept. Elect. Eng. Syst. Univ. Southern Calif., Tech. Rep. 91-09-01, 1991.

[11] R. M. Sanner and J. E. Slotine, "Gaussian networks for direct adaptive control," *IEEE Trans. Neural Networks*, vol. 3, pp. 837–863, 1992.
[12] F. C. Chen and C. C. Liu, "Adaptively controlling nonlinear continuous-time systems using multilayer neural networks," *IEEE Trans. Automat. Contr.*, vol. 39, no. 6, pp. 1306–1310, 1994.
[13] G. A. Rovithakis and M. A. Christodoulou, "Adaptive control of unknown plants using dynamical neural networks," *IEEE Trans. Syst., Man, Cybern.*, vol. 24, pp. 400–412, Mar. 1994.
[14] K. S. Narendra and S. Mukhopadhyay, "Adaptive control of nonlinear multivariable system using neural networks," *Neural Networks*, vol. 7, no. 5, pp. 737–752, 1994.
[15] 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.
[16] A. Yesidirek and F. L. Lewis, "Feedback linearization using neural networks," *Automatica*, vol. 31, no. 11, pp. 1659–1664, 1995.
[17] F. L. Lewis, A. Yesildirek, and K. Liu, "Multilayer neural-net robot controller with guaranteed tracking performance," *IEEE Trans. Neural Networks*, vol. 7, pp. 388–398, 1996.
[18] J. T. Spooner and K. M. Passino, "Stable adaptive control using fuzzy systems and neural networks," *IEEE Trans. Fuzzy Syst.*, vol. 4, pp. 339–359, 1996.
[19] G. A. Rovithakis, "Tracking control of multi-input affine nonlinear dynamical systems with unknown nonlinearities using dynamical neural networks," *IEEE Trans. Syst., Man, Cybern.*, vol. 29, pp. 179–189, 1999.
[20] M. M. Polycarpou, "Stable adaptive neural control scheme for nonlinear systems," *IEEE Trans. Automat. Contr.*, vol. 41, pp. 447–451, 1996.
[21] M. M. Polycarpou and M. J. Mears, "Stable adaptive tracking of uncertain systems using nonlinearly parametrized on-line approximators," *Int. J. Contr.*, vol. 70, no. 3, pp. 363–384, 1998.
[22] F. L. Lewis, A. Yesildirek, and K. Liu, "Robust backstepping control of induction motors using neural networks," *IEEE Trans. Neural Networks*, vol. 11, pp. 1178–1187, 2000.
[23] C. Kwan and F. L. Lewis, "Robust backstepping control of nonlinear systems using neural networks," *IEEE Trans. Syst., Man, Cybern., A*, vol. 30, pp. 753–766, 2000.
[24] Y. Zhang, P. Y. Peng, and Z. P. Jiang, "Stable neural controller design for unknown nonlinear systems using backstepping," *IEEE Trans. Neural Networks*, vol. 11, pp. 1347–1359, 2000.
[25] S. S. Ge, C. C. Hang, and T. Zhang, "Stable adaptive control for multivariable systems with a triangular control structure," *IEEE Trans. Automat. Contr.*, vol. 45, pp. 1221–1225, 2000.
[26] S. Haykin, *Neural Networks: A Comprehensive Foundation*, 2nd ed. Englewood Cliffs, NJ: Prentice-Hall, 1999.
[27] B. Schwartz, A. Isidori, and T. J. Tarn, "Global normal forms for MIMO nonlinear systems, with applications to stabilization and disturbance attenuation," *Math. Contr. Signals Syst.*, vol. 12, no. 2, pp. 121–142, 1999.
[28] X. P. Liu, G. X. Gu, and K. M. Zhou, "Robust stabilization of MIMO nonlinear systems by backstepping," *Automatica*, vol. 35, no. 5, pp. 987–992, 1999.
[29] B. Yao and M. Tomizuka, "Adaptive robust control of MIMO nonlinear systems in semi-strict feedback forms," *Automatica*, vol. 37, no. 9, pp. 1305–1321, 2001.
[30] W. Lin and C. J. Qian, "Semi-global robust stabilization of MIMO nonlinear systems by partial state and dynamic output feedback," *Automatica*, vol. 37, no. 7, pp. 1093–1101, 2001.
[31] S. S. Sastry and A. Isidori, "Adaptive control of linearizable systems," *IEEE Trans. Automat. Contr.*, vol. 34, pp. 1123–1131, 1989.
[32] K. Nam and A. Arapostations, "A model-reference adaptive control scheme for pure-feedback nonlinear systems," *IEEE Trans. Automat. Contr.*, vol. 33, pp. 803–811, 1988.
[33] G. Nurnberger, *Approximation by Spline Functions*. New York: Springer-Verlag, 1989.
[34] P. A. Ioannou and J. Sun, *Robust Adaptive Control*. Englewood Cliffs, NJ: Prentice-Hall, 1995.
[35] J. Park and I. W. Sandberg, "Universal approximation using radial basis function networks," *Neural Comput.*, vol. 3, no. 2, pp. 246–257, 1991.
[36] A. Isidori, *Nonlinear Control System*, 2 ed. Berlin, Germany: Springer-Verlag, 1989.
[37] Z. Lin and A. Saberi, "Robust semi-global stabilization of minimum-phase input–output linearizable systems via partial state and output feedback," *IEEE Trans. Automat. Contr.*, vol. 40, pp. 1029–1041, 1995.
[38] R. Sepulchre, M. Jankovic, and P. V. Kokotovic, *Constructive Nonlinear Control*. London, U.K.: Springer-Verlag, London Ltd., 1997.
[39] Z. Qu, *Robust Control of Nonlinear Uncertain Systems*. New York: Wiley, 1998.
[40] M. Vidyasagar, *Nonlinear Systems Analysis*, 2nd ed. Englewood Cliffs, NJ: Prentice-Hall, 1993.

Shuzhi Sam Ge (S'90–M'92–SM'00) received the B.Sc. degree from Beijing University of Aeronautics and Astronautics (BUAA), Beijing, China, in 1986, and the Ph.D. degree and the Diploma of Imperial College (DIC) from the Imperial College of Science, Technology and Medicine, University of London, U.K., in 1993.

From 1992 to 1993, he was a Postdoctoral Research with Leicester University, U.K. He visited the Laboratoire de'Automatique de Grenoble, France, in 1996, the University of Melbourne, Australia, in 1998 and 1999, the University of Petroleum, Shanghai Jiaotong University, China, in 2001, and City University, Hong Kong, in 2002. He has been with the Department of Electrical and Computer Engineering, National University of Singapore, since 1993 where he is currently an Associate Professor. He serves as a technical consultant local industry. He has authored and coauthored more than 100 international journal and conference papers, two monographs, and coinvented two patents. His current research interests are intelligent control, neural and fuzzy systems, hybrid systems, and sensor fusion.

Dr. Ge has been serving as an Associate Editor of the IEEE TRANSACTIONS ON CONTROL SYSTEMS TECHNOLOGY since 1999. He has been a member of the Technical Committee on Intelligent Control of the IEEE Control System Society since 2000. He was the recipient of the 1999 National Technology Award, the 2001 University Young Research Award, and the 2002 Temasek Young Investigator Award, Singapore.

Cong Wang (S'90–S'91–S'93–M'93) received the B.E. and M.E. degrees from the Department of Automatic Control, Beijing University of Aeronautic and Astronautics (BUAA), Beijing, China, in 1989 and 1997, respectively, and the Ph.D. degree from the Department of Electrical and Computer Engineering, National University of Singapore, in 2002.

He was a Postdoctoral Researcher with the City University of Hong Kong. He is currently a Faculty Member at the College of Automation, South China University of China. His research interests include adaptive neural control, neural learning control, and control applications.