$x(T)$ is nonzero. Based on this remark, it was erroneously claimed[1] that W^n has no usual extension.

From Proposition 1 one obtains the following result.

Proposition 2: The norm of $E_T x$ in W^n is

$$\|E_T x\|_W = \left(\|x\|_{2,T}^2 + \|\dot{x}\|_{2,T}^2 + \|x(T)\|^2\right)^{1/2} \tag{1}$$

where $\|\cdot\|_{2,T}$ denotes the seminorm in $L_{2,e}^n$ defined by $\|\cdot\|_{2,T} = \|P_T \cdot\|_2$.

Set $\|\cdot\|_{W,T} := \|E_T \cdot\|_W$ (and notice the difference with the definition of $\|\cdot\|_{2,T}$). The family $(\|\cdot\|_{W,T})_{T\geq 0}$ is a nondecreasing family of seminorms; it defines over W^n a topology $\mathcal{T}$ which is strictly finer than the one defined by the norm $\|\cdot\|_W$ [because if a sequence $(x_p)_{p\geq 0}$ of W^n is such that $\|x_p\|_W$ tends to zero, then $\|x_p\|_{W,T}$ tends to zero for every $T \geq 0$, but not conversely]. In accordance with [2], we take the following definition.

Definition 1: The extended resolution space W_e^n is the completion of $(W^n, \mathcal{T})$.

In other words, W_e^n is the set of functions x such that $E_T x$ belongs to W^n for every $T \geq 0$; W_e^n is equipped with the family of seminorms $(\|\cdot\|_{W,T})_{T\geq 0}$ and is complete for the associated topology (so that W_e^n is a Fréchet space).

III. Relations Between Various Extended Resolution Spaces

The following replaces Proposition 1.[1]

Proposition 3: Let $x \in W_e^n$. Then for every $T \geq 0$ the following inequality holds:

$$\|x\|_{\infty,T} \leq \|x\|_{W,T}. \tag{2}$$

The proof is similar to that of Proposition,[1] taking into account the new expression (1). Taking $T \to \infty$, one obtains the inclusion $W^n \subset L_\infty^n$ and the inequality

$$\|x\|_\infty \leq \|x\|_W.$$

Similarly, the following result improves Proposition 2;[1] in addition, it clarifies the relation between $L_{2,e}^n$ and W_e^n. Let $\boldsymbol{K}$ be the causal convolution operator with transfer matrix $\boldsymbol{K}(s) = (1+s)^{-1} I_n$.

Proposition 4: $\boldsymbol{K}$ is an extended resolution space-isomorphism from $L_{2,e}^n$ onto W_e^n. More specifically, $\boldsymbol{K}$ is one-to-one from $L_{2,e}^n$ onto W_e^n and $\|\boldsymbol{K}x\|_{W,T} = \|x\|_{2,T}$.

IV. Concluding Remarks

Compared with Lebesgue spaces (especially L_2), the Sobolev space W is somewhat particular in that it contains only differentiable functions (in the sense of distributions). The consequence is that, in place of the usual truncation operator P_T, the resolution of the identity which has to be used is another operator E_T. Equipped with this operator, W can be "extended" in a classic manner. As shown by Propositions 3 and 4, the extended space W_e has nice relations with the extended spaces $L_{2,e}$ and $L_{\infty,e}$. Part II of the paper should be replaced by the above approach whereas the rest of the paper can be left unchanged.

References

[1] H. Brezis, *Analyse Fonctionnelle-Théorie et Applications*, Paris: Masson, 1983.

[2] A. Feintuch and R. Saeks, *System Theory: A Hilbert Space Approach*. New York: Academic, 1982.

Stable Adaptive Control for Nonlinear Multivariable Systems with a Triangular Control Structure

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

***Abstract*—A stable adaptive controller is developed for a class of nonlinear multivariable systems using nonlinearly parametrized function approximators. By utilizing the system triangular property, integral-type Lyapunov functions are introduced for deriving the control structure and adaptive laws without the need of estimating the "decoupling matrix" of the multivariable nonlinear system. It is shown that stability of the adaptive closed-loop system is guaranteed, and transient performance is analytically quantified by mean square and L_∞ tracking error bound criteria.**

***Index Terms*—Adaptive control, function approximation, Lyapunov function, multivariable systems, stability.**

I. Introduction

Recent years have seen many significant developments in adaptive control for single-input/single-output (SISO) nonlinear systems. For multi-input/multi-output (MIMO) nonlinear systems, due to the couplings among various inputs and outputs, the control problem is more complex and few results are available in the literature. One of the major difficulties comes from the uncertainty in the input coupling matrix. Based on feedback linearization techniques, a variety of adaptive controllers have been proposed for linearizable systems (e.g., [1]–[6]). In these methods, for removing 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. For example, the projection algorithm was applied in [1] to keep the estimated parameters inside a feasible set in which the singularity problem does not occur. Although such a projection is a standard technique in linear adaptive control, 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 [22]. In fact, even for SISO adaptive nonlinear control, solving the control singularity problem is far from trivial (see [12] and the references therein). For MIMO nonlinear systems, there is no effective scheme available at the present stage. As an alternative approach, a neural control design proposed in [2], [3] suggested that the initial values of neural network (NN) weights are chosen sufficiently close to the ideal values such that the inverse of the estimated decoupling matrix always exists and, therefore, an off-line training phase for NN's is necessary before the controller is put into the closed-loop system. In [6], an interesting adaptive controller was developed for multi-input parametric pure-feedback nonlinear systems. Global stability and convergence of the tracking errors were achieved for all initial parameter estimates lying in an open neighborhood of the true parameter space. For a special class of MIMO nonlinear plants, rigid robot systems, adaptive neural network control has been proposed based on energy-type Lyapunov functions in [11], [15].

As mentioned earlier, the couplings that exist among the inputs and outputs of MIMO nonlinear systems are the main difficulty in multivariable adaptive control. For systems without any input coupling, decentralized adaptive controllers were presented for a class of nonlinear systems with state interconnections [7]–[9]. Global stability and asymptotic tracking were obtained in [7], [8], provided that the system

Manuscript received December 14, 1998; revised October 23, 1999. Recommended by Associate Editor, M. Polycarpou.

The authors are with the Department of Electrical Engineering, National University of Singapore, Singapore 117576.

Publisher Item Identifier S 0018-9286(00)04231-8.

interconnections are bounded by a pth-order polynomial in states. Recently, this decentralized adaptive control idea was further extended to large-scale nonlinear systems transformable to a decentralized strict-feedback form in [9] under the similar interconnection conditions.

In this note, we shall consider a class of MIMO nonlinear systems having a triangular structure in control inputs. Using such a triangular property, novel Lyapunov functions are introduced to construct a Lyapunov-based control structure, which does not try to cancel the decoupling matrix for linearizing the system. The control singularity problem discussed above is eliminated without using projection algorithms or restricting the initial values of parameter estimates. Other features of the proposed scheme include that nonlinearly parametrized approximators, multilayer neural networks (MNN's), are used for function approximation, and the stability of the closed-loop adaptive system is guaranteed. Explicit transient bounds of output tracking errors are given analytically, and several possible schemes are provided for improving the control performance of the resulting adaptive system.

II. Problem Formulation

Consider MIMO nonlinear systems in the triangular control form, shown in (1), where $X = [x_1^T, x_2^T, \cdots, x_m^T]^T$ with $x_j = [x_{j,1}, x_{j,2}, \cdots, x_{j,\rho_j}]^T \in R^{\rho_j}$, $u_j \in R$ and $y_j \in R$ are the state variables, the system inputs and outputs, respectively; $f_j(\cdot)$ and $g_{j,i}(\cdot)$ are unknown smooth functions; and j, i_j, ρ_j, and m are positive integers. System (1)

$$\begin{cases} \dot{x}_{1,i_1} = x_{1,i_1+1}, \\ \dot{x}_{1,\rho_1} = f_1(X) + g_{1,1}(x_1)u_1, \\ \dot{x}_{2,i_2} = x_{2,i_2+1}, \\ \dot{x}_{2,\rho_2} = f_2(X) + g_{2,1}(X)u_1 + g_{2,2}(x_1,x_2)u_2, \\ \qquad 1 \le i_j \le \rho_j - 1 \\ \quad \vdots \\ \dot{x}_{j,i_j} = x_{j,i_j+1}, \qquad 1 \le j \le m \\ \dot{x}_{j,\rho_j} = f_j(X) + g_{j,1}(X)u_1 + g_{j,2}(X)u_2 + \cdots \\ \qquad + g_{j,j}(x_1, x_2, \cdots, x_j)u_j, \\ y_j = x_{j,1}, \end{cases} \tag{1}$$

possesses a special triangular form in control inputs, which covers a large class of plants including the decentralized systems studied in [7], [8]. In addition, MIMO system (1) is fully interconnected because the interconnection terms $f_j(\cdot)$ contain all the system states X, which is different from the hierarchical or triangular systems defined in [24]. The control objective is to design a stable adaptive controller such that the system outputs y_j follow the desired trajectories y_{dj}.

Let $X_j = [x_1^T, x_2^T, \cdots, x_j^T]^T \in R^{n_j}$ with $n_j = \sum_{i=1}^{j} \rho_i$; $\|\cdot\|$ denote the 2-norm; $\|\cdot\|_F$ denote the Frobenius norm; $|A|_1 = \sum_{i=1}^{p} |a_i|$ with $A = [a_1, a_2, \cdots, a_p]^T \in R^p$; and $\lambda_{\max}(B)$ and $\lambda_{\min}(B)$ denote the largest and smallest eigenvalues of positive definite real square matrix B, respectively.

Assumption 1: $g_{j,j}(X_j)$ are bounded away from zero with known signs, and known smooth functions $\overline{g}_j(X_j)$ exist such that $|g_{j,j}(X_j)| \le \overline{g}_j(X_j)$, $\forall X_j \in R^{n_j}$.

Remark 2.1: Assumption 1 implies that smooth functions $g_{j,j}(X_j)$ are strictly either positive or negative definite. From now on, we shall assume that $0 < g_{0j} \le g_{j,j}(X_j) \le \overline{g}_j(X_j)$, where the unknown constants g_{0j} are only used for analytical purpose; their true values are not necessarily known for controller design.

Define x_{dj}, e_j and filtered tracking errors e_{sj} as

$$\begin{aligned} x_{dj} &= [y_{dj}, \dot{y}_{dj}, \cdots, y_{dj}^{(\rho_j-1)}]^T \\ e_j &= x_j - x_{dj} = [e_{j,1}, e_{j,2}, \cdots, e_{j,\rho_j}]^T \\ e_{sj} &= \left(\frac{d}{dt} + \lambda_j\right)^{\rho_j-1} e_{j,1} = [\Lambda_j^T 1]e_j, \text{ with } \lambda_j > 0 \end{aligned} \tag{2}$$

where $\Lambda_j = [\lambda_j^{\rho_j-1}, (\rho_j-1)\lambda_j^{\rho_j-2}, \cdots, (\rho_j-1)\lambda_j]^T$.

Remark 2.2: It has been shown in [23] that definition (2) has the following properties: 1) $e_{sj} = 0$ define time-varying hyperplans in R^{ρ_j} on which the tracking errors $e_{j,1}$ converge to zeros asymptotically; 2) if the magnitudes of e_{sj} are bounded by constants $C_j > 0$, then $e_j(t) \in \Omega_{cj}$ for all $e_j(0) \in \Omega_{cj}$ with

$$\begin{aligned} \Omega_{cj} = \{e_j(t) | \, &|e_{j,i_j}(t)| \le 2^{i_j-1}\lambda_j^{i_j-\rho_j}C_j, \\ &i_j = 1, 2, \cdots, \rho_j, \qquad j = 1, 2, \cdots, m\} \end{aligned} \tag{3}$$

and 3) the tracking errors $e_{j,1} = H_j(s)e_{sj}$ with $H_j(s)$ proper stable transfer functions.

Assumption 2: The desired trajectory vectors $\overline{x}_{dj} = [x_{dj}^T, y_{dj}^{(\rho_j)}]^T$ are continuous and available, and $\overline{x}_{dj} \in \Omega_{dj} \subset R^{\rho_j+1}$ with Ω_{dj} being compact subsets.

Neural networks have been widely used in modeling and control of nonlinear systems because of their good capabilities of nonlinear function approximation, learning, and fault tolerance. The feasibility of applying NN's to unknown dynamic system control has been demonstrated in many studies [2]–[5], [10]–[15], [21]. In this paper, three-layer NN's will be used to approximate a continuous function $h(Z)$: $R^p \to R$ as described by [11], [13]

$$g_{nn}(Z) = W^T S(V^T \overline{Z}) \tag{4}$$

where $\overline{Z} = [Z^T, 1]^T$ is the input vector; $W = [w_1, w_2, \cdots, w_l]^T \in R^l$ and $V = [v_1, v_2, \cdots, v_l] \in R^{(p+1)\times l}$ are the first-to-second layer and the second-to-third layer weights, respectively; $S(V^T\overline{Z}) = [s(v_1^T\overline{Z}), s(v_2^T\overline{Z}), \cdots, s(v_{l-1}^T\overline{Z}), 1]^T$ with $s(z_a) = 1/(1 + e^{-\gamma z_a})$ and constant $\gamma > 0$; and the NN node number $l > 1$. It has been proven that neural network (4) satisfies the conditions of the Stone-Weierstrass Theorem and can approximate any continuous function over a compact set [11], [16], i.e., on a compact set Ω_z, we have

$$h(Z) = W^{*T} S(V^{*T}\overline{Z}) + \mu(Z) \quad \forall Z \in \Omega_z \subset R^p \tag{5}$$

where W^* and V^* are ideal NN weights and $\mu(Z)$ is the NN approximation error.

Assumption 3: For a given continuous function $h(Z)$, constant W^* and V^* exist such that function approximation (5) holds and $|\mu(Z)| \le \overline{\mu}$ with constant $\overline{\mu} > 0$ for all $Z \in \Omega_z$.

Let $\hat{W}$ and $\hat{V}$ be the estimates of W^* and V^*, respectively, and the weight estimation errors denoted as $\tilde{W} = \hat{W} - W^*$ and $\tilde{V} = \hat{V} - V^*$. In the analysis that follows, we will use the following lemma.

Lemma 2.1: (Reference [14]) For MNN's (4), the NN parametric approximation error can be expressed as

$$\begin{aligned} &\hat{W}^T S(\hat{V}^T\overline{Z}) - W^{*T} S(V^{*T}\overline{Z}) \\ &\quad = \tilde{W}^T(\hat{S} - \hat{S}'\hat{V}^T\overline{Z}) + \hat{W}^T\hat{S}'\tilde{V}^T\overline{Z} + d_u \end{aligned} \tag{6}$$

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

$$|d_u| \le \|V^*\|_F \|\overline{Z}\hat{W}^T\hat{S}'\|_F + \|W^*\| \, \|\hat{S}'\hat{V}^T\overline{Z}\| + |W^*|_1. \tag{7}$$

III. Integral Type Lyapunov Functions for Control Design

Define positive definite functions $\beta_j(X_j) = \overline{g}_j(X_j)/g_{j,j}(X_j)$ and

$$\begin{aligned} \overline{x}_j &= [x_{j,1}, x_{j,2}, \cdots, x_{j,\rho_j-1}]^T \in R^{\rho_j-1} \\ \nu_j &= y_{dj}^{(\rho_j-1)} - [\Lambda_j^T \quad 0]e_j \end{aligned} \tag{8}$$

$$\overline{X}_j = [x_1^T, x_2^T, \cdots, x_{j-1}^T, \overline{x}_j]^T \in R^{n_{j-1}+\rho_j-1} \tag{9}$$

with $\overline{X}_1 = \overline{x}_1$ and $n_{j-1} = \sum_{i=1}^{j-1} \rho_i$ for $j > 1$. From system (1) and definitions (2) and (9), the time derivatives of e_{sj} can be written as

$$\dot{e}_{sj} = f_j(X) + \sum_{i=1}^{j-1} g_{j,i}(X)u_i + g_{j,j}(X_j)u_j - \dot{\nu}_j \tag{10}$$

where $\sum_{i=1}^{j-1} (\cdot) = 0$ for $j = 1$. It is shown from (2) and (8) that $x_{j,\rho_j} = e_{sj} + \nu_j$. In the following, we denote $\beta_j(X_j) = \beta_j(\overline{X}_j, e_{sj} + \nu_j)$ for notational convenience. Define smooth scalar functions

$$\begin{aligned} V_{zj}(e_{sj}, \overline{X}_j, \nu_j) &= \int_0^{e_{sj}} \sigma\beta_j(\overline{X}_j, \sigma + \nu_j)\, d\sigma \\ &= e_{sj}^2 \int_0^1 \theta\beta_j(\overline{X}_j, \theta e_{sj} + \nu_j)\, d\theta. \end{aligned} \tag{11}$$

The second equality in the above equation can be obtained by the variable change $\sigma = \theta e_{sj}$. According to Assumption 1 and Remark 2.1, we have $1 \le \beta_j(\overline{X}_j, \theta e_{sj} + \nu_j) \le \overline{g}_j(\overline{X}_j, \theta e_{sj} + \nu_j)/g_{0j}$, and therefore

$$\frac{e_{sj}^2}{2} \le V_{zj} \le \frac{e_{sj}^2}{g_{0j}} \int_0^1 \theta\overline{g}_j(\overline{X}_j, \theta e_{sj} + \nu_j)\, d\theta. \tag{12}$$

Thus, V_{zj} are positive definite with respect to e_{sj}. Remark 2.2 shows that the boundedness of system states and convergence of tracking errors are completely determined by those of filtered tracking errors e_{sj}. We shall study the system stability and performance through investigating the properties of e_{sj}. It will be shown later that integral type functions V_{zj} are applied for constructing a Lyapunov-based control structure that does not contain the inverses of the unknown functions $g_{j,j}(X_j)$, and the subsequent adaptive design avoids the use of estimates $\hat{g}_{j,j}^{-1}(X_j)$. Taking V_{zj} as Lyapunov function candidates, their time derivatives are

$$\begin{aligned} \dot{V}_{zj} = e_{sj}\beta_j(X_j)\dot{e}_{sj} + \int_0^{e_{sj}} \sigma \Bigg[&\frac{\partial\beta_j(\overline{X}_j, \sigma + \nu_j)}{\partial\overline{X}_j}\, \dot{\overline{X}}_j \\ &+ \frac{\partial\beta_j(\overline{X}_j, \sigma + \nu_j)}{\partial\nu_j}\, \dot{\nu}_j \Bigg]\, d\sigma. \end{aligned} \tag{13}$$

The following two equalities hold:

1)
$$\begin{aligned} &\int_0^{e_{sj}} \left[\sigma \frac{\partial\beta_j(\overline{X}_j, \sigma + \nu_j)}{\partial\nu_j}\dot{\nu}_j\right] d\sigma \\ &= \dot{\nu}_j \int_0^{e_{sj}} \sigma \frac{\partial\beta_j(\overline{X}_j, \sigma + \nu_j)}{\partial\sigma}\, d\sigma \\ &= \dot{\nu}_j \left[e_{sj}\beta_j(X_j) - \int_0^{e_{sj}} \beta_j(\overline{X}_j, \sigma + \nu_j)\, d\sigma\right] \\ &= \dot{\nu}_j e_{sj} \left[\beta_j(X_j) - \int_0^1 \beta_j(\overline{X}_j, \theta e_{sj} + \nu_j)\, d\theta\right] \end{aligned} \tag{14}$$

2)
$$\begin{aligned} &\int_0^{e_{sj}} \sigma \left[\frac{\partial\beta_j(\overline{X}_j, \sigma + \nu_j)}{\partial\overline{X}_j}\, \dot{\overline{X}}_j\right] d\sigma \\ &= \int_0^{e_{sj}} \sigma \left[\sum_{i=1}^{j-1} \frac{\partial\beta_j(\overline{X}_j, \sigma + \nu_j)}{\partial x_i}\, \dot{x}_i + \frac{\partial\beta_j(\overline{X}_j, \sigma + \nu_j)}{\partial\overline{x}_j}\, \dot{\overline{x}}_j\right] d\sigma \\ &= \int_0^{e_{sj}} \sigma \left\{\sum_{i=1}^{j} \left(\sum_{k=1}^{\rho_i-1} \frac{\partial\beta_j(\overline{X}_j, \sigma + \nu_j)}{\partial x_{i,k}} x_{i,k+1}\right) + \sum_{i=1}^{j-1} \left(\frac{\partial\beta_j(\overline{X}_j, \sigma + \nu_j)}{\partial x_{i,\rho_i}}\, \dot{x}_{i,\rho_i}\right)\right\} d\sigma \\ &= e_{sj}^2\phi_j(Z_j) \end{aligned} \tag{15}$$

with

$$\begin{aligned} \phi_j(Z_j) = &\int_0^1 \theta \left\{\sum_{i=1}^{j} \left(\sum_{k=1}^{\rho_i-1} \frac{\partial\beta_j(\overline{X}_j, \theta e_{sj} + \nu_j)}{\partial x_{i,k}} x_{i,k+1}\right)\right\} d\theta \\ &+ \int_0^1 \theta \sum_{i=1}^{j-1} \left\{\frac{\partial\beta_j(\overline{X}_j, \theta e_{sj} + \nu_j)}{\partial x_{i,\rho_i}} \cdot \left[f_i(X) + \sum_{k=1}^{i-1} g_{i,k}(X)u_k + g_{i,i}(X_i)u_i\right]\right\} d\theta \end{aligned}$$

and $Z_j = [X^T, \nu_j, \dot{\nu}_j, u_1, u_2, \cdots, u_{j-1}]^T \in R^{n_m+j+1}$. Substituting (10) into (13) and noting (14) and (15), equalities (13) can be expressed as

$$\dot{V}_{zj} = -e_{sj}^2[\phi_j^+(Z_j) - \phi_j(Z_j)] + e_{sj}\overline{g}_j(X_j)u_j + h_j(Z_j)] \tag{16}$$

where

$$\begin{aligned} h_j(Z_j) = &\beta_j(X_j)\left[f_j(X) + \sum_{i=1}^{j-1} g_{j,i}(X_j)u_i\right] \\ &- \dot{\nu}_j \int_0^1 \beta_j(\overline{X}_j, \theta e_{sj} + \nu_j)\, d\theta + e_{sj}\phi_j^+(Z_j) \end{aligned} \tag{17}$$

and

$$\phi_j^+(Z_j) = \begin{cases} \phi_j(Z_j), & \phi_j(Z_j) \ge 0 \\ 0, & \text{otherwise.} \end{cases} \tag{18}$$

In the ideal case, that all nonlinearities of system (1) are known exactly, a possible controller takes the form $u_j^* = [-k_je_{sj} - h_j(Z_j)]/\overline{g}_j(X_j)$ with $k_j > 0$. Substituting $u_j = u_j^*$ into (16) and noting (18), we have $\dot{V}_{zj} = -[k_j + \phi_j^+(Z_j) - \phi_j(Z_j)]e_{sj}^2 \le -k_je_{sj}^2$. Therefore, V_{zj} are Lyapunov functions and $e_{sj} \to 0$ as $t \to \infty$.

In the case that no exact knowledge for the system nonlinearities is available, MNN's provided in (4) can be applied for the approximations

$$\begin{aligned} h_j(Z_j) = W_j^{*T} S_j(V_j^{*T}\overline{Z}_j) + \mu_j(Z_j) \quad \forall Z_j \in \Omega_{zj}, \\ j = 1, 2, \cdots, n \end{aligned} \tag{19}$$

where $\overline{Z}_j = [Z_j^T, 1]^T$ are the input vectors; Ω_{zj} are compact sets to be specified later; W_j^* and V_j^* are ideal constant weights; and $|\mu_j(Z_j)| \le \overline{\mu}_j$ with constants $\overline{\mu}_j > 0$. It follows from Lemma 2.1 that the NN estimation errors may be written as

$$\begin{aligned} &\hat{W}_j^T S_j(\hat{V}_j^T\overline{Z}_j) - h_j(Z_j) \\ &\quad = \tilde{W}_j^T(\hat{S}_j - \hat{S}'_j\hat{V}_j^T\overline{Z}_j) + \hat{W}_j^T\hat{S}'_j\tilde{V}_j^T\overline{Z}_j + d_{uj} - \mu_j(Z_j) \end{aligned} \tag{20}$$

where $\hat{W}_j = [\hat{w}_{j,1}, \hat{w}_{j,2}, \cdots, \hat{w}_{j,l_j}]^T \in R^{l_j}$ and $\hat{V}_j = [\hat{v}_{j,1}, \hat{v}_{j,2}, \cdots, \hat{v}_{j,l_j}] \in R^{(n_m+j+2)\times l_j}$ denote the estimates of W_j^* and V_j^*, respectively; $\hat{S}_j = S_j(\hat{V}_j^T\overline{Z}_j)$; $\hat{S}'_j = \mathrm{diag}\{\hat{s}'_{j,1}, \hat{s}'_{j,2}, \cdots, \hat{s}'_{j,l_j}\}$ with $\hat{s}'_{j,k} = s'(\hat{v}_{j,k}^T\overline{Z}_j) = d[s(z_a)]/dz_a|_{z_a=\hat{v}_{j,k}^T\overline{Z}_j}, k = 1, 2, \cdots, l_j$, and $l_j > 1$; and the residual terms d_{uj} are bounded by

$$|d_{uj}| \le \|V_j^*\|_F\|\overline{Z}_j\hat{W}_j^T\hat{S}'_j\|_F + \|W_j^*\|\,\|\hat{S}'_j\hat{V}_j^T\overline{Z}_j\| + |W_j^*|_1. \tag{21}$$

IV. Adaptive Controller and Stability Analysis

We now present the adaptive NN controller

$$u_j = \frac{1}{\overline{g}_j(X_j)}\,[-k_j(t)e_{sj} - \hat{W}_j^T S_j(\hat{V}_j^T\overline{Z}_j)], \quad j = 1, 2, \cdots, m \tag{22}$$

where $\overline{Z}_j = [X^T, \nu_j, \dot{\nu}_j, u_1, u_2, \cdots, u_{j-1}, 1]^T$ with $\overline{Z}_1 = [X^T, \nu_1, \dot{\nu}_1, 1]^T$ for $j = 1$ and

$$k_j(t) = \frac{1}{\varepsilon_j}\left(1 + \int_0^1 \theta\overline{g}_j(\overline{X}_j, \theta e_{sj} + \nu_j)\, d\theta + \|\overline{Z}_j\hat{W}_j^T\hat{S}'_j\|_F^2 + \|\hat{S}'_j\hat{V}_j^T\overline{Z}_j\|^2\right) \tag{23}$$

with design parameters $\varepsilon_j > 0$. From (22), it is clear that $k_j(t)$ can be viewed as the controller gains. Since $\int_0^1 \theta\overline{g}_j(\overline{X}_j, \theta e_{sj} + \nu_j)\, d\theta \geq g_{0j}/2 > 0$ (using conditions $\overline{g}_j(X_j) \geq g_{0j}$ in Remark 2.1, the integrals in $k_j(t)$ are always positive definite. It is shown later that these integral terms are essential for capturing the explicit upper bounds of the system signals. The following adaptive laws are used to update the NN weights

$$\dot{\hat{W}}_j = \Gamma_{wj}[(\hat{S}_j - \hat{S}'_j\hat{V}_j^T\overline{Z}_j)e_{sj} - \sigma_{wj}\hat{W}_j] \tag{24}$$

$$\dot{\hat{V}}_j = \Gamma_{vj}[\overline{Z}_j\hat{W}_j^T\hat{S}'_j e_{sj} - \sigma_{vj}\hat{V}_j] \tag{25}$$

where $\Gamma_{wj} = \Gamma_{wj}^T > 0$, $\Gamma_{vj} = \Gamma_{vj}^T > 0$, and $\sigma_{wj}, \sigma_{vj} > 0$ are constant design parameters. In the above adaptive algorithms, σ-modification [22] terms are introduced to improve the controller robustness in the presence of the NN approximation errors.

Theorem 4.1: Given nonlinear system (1) satisfying Assumptions 1–3, controller (22) and weight updating laws (24) and (25), then for bounded initial conditions,

1) all signals in the closed-loop system are bounded, and compact sets Ω_X and Ω_{wj} exist such that the vectors Z_j remain in

$$\begin{aligned}\Omega_{zj} = \{&(X, \nu_j, \dot{\nu}_j, u_1, u_2, \cdots, u_{j-1})|\\ &X \in \Omega_X, (\hat{W}_1, \hat{V}_1) \in \Omega_{w1},\\ &(\hat{W}_2, \hat{V}_2) \in \Omega_{w2}, \cdots, (\hat{W}_{j-1}, \hat{V}_{j-1}) \in \Omega_{w(j-1)},\\ &\overline{x}_{d(j+1)} \in \Omega_{d(j+1)}\}\end{aligned} \tag{26}$$

where $\Omega_{z1} = \{(X, \nu_1, \dot{\nu}_1)|X \in \Omega_X, \overline{x}_{d2} \in \Omega_{d2}\}$ for notational convenience, and

2) mean square tracking performance

$$\frac{1}{t}\int_0^t e_{j,1}^2(\tau)\, d\tau \leq \frac{2\varepsilon_j a_j c_j}{1 + g_{0j}} + \frac{1}{t}\left(\frac{2\varepsilon_j a_j V_{sj}(0)}{1 + g_{0j}} + b_j\right) \quad \forall t > 0 \tag{27}$$

and L_∞ tracking error bounds

$$|e_{j,1}(t)| \leq \lambda_j^{1-\rho_j}\left(2V_{sj}(0)e^{-\lambda_{sj}t} + \frac{2}{\lambda_{sj}}c_j\right)^{1/2} \tag{28}$$

with positive constants a_j, b_j, c_j, $V_{sj}(0)$ and λ_{sj}.

Proof: 1) The proof includes two parts. We first suppose that compact sets Ω_{zj} exist such that $Z_j \in \Omega_{zj}$, $\forall t \geq 0$, and the NN approximation (20) holds. Then, we prove that these compact sets Ω_{zj} exist for bounded initial conditions. Consider the augmented Lyapunov function candidates

$$V_{sj} = V_{zj} + \tfrac{1}{2}[\tilde{W}_j^T\Gamma_{wj}^{-1}\tilde{W}_j + \mathrm{tr}\{\tilde{V}_j^T\Gamma_{vj}^{-1}\tilde{V}_j\}] \tag{29}$$

Taking their time derivatives along (16) and noting (20) and (22), we have

$$\begin{aligned}\dot{V}_{sj} = &-e_{sj}[k_j(t)e_{sj} + \tilde{W}_j^T(\hat{S}_j - \hat{S}'_j\hat{V}_j^T\overline{Z}_j)\\ &+ \hat{W}_j^T\hat{S}'_j\tilde{V}_j^T\overline{Z}_j + d_{uj} - \mu_j(Z_j)]\\ &-e_{sj}^2[\phi_j^+(Z_j) - \phi_j(Z_j)] + \tilde{W}_j^T\Gamma_{wj}^{-1}\dot{\hat{W}}_j + \mathrm{tr}\{\tilde{V}_j^T\Gamma_{vj}^{-1}\dot{\hat{V}}_j\}\end{aligned}$$

Substituting (24) and (25) into the above equations and using the facts that $\hat{W}_j^T\hat{S}'_j\tilde{V}_j^T\overline{Z}_j = \mathrm{tr}\{\tilde{V}_j^T\overline{Z}_j\hat{W}_j^T\hat{S}'_j\}$ and $\phi_j^+(Z_j) - \phi_j(Z_j) \geq 0$, we obtain

$$\begin{aligned}\dot{V}_{sj} \leq &-k_j(t)e_{sj}^2 - [d_{uj} - \mu_j(Z_j)]e_{sj}\\ &- \sigma_{wj}\tilde{W}_j^T\hat{W}_j - \sigma_{vj}\mathrm{tr}\{\tilde{V}_j^T\hat{V}_j\}.\end{aligned}$$

Noting (21), (23) and the properties that $2\tilde{W}_j^T\hat{W}_j \geq \|\tilde{W}_j\|^2 - \|W_j^*\|^2$ and $2\mathrm{tr}\{\tilde{V}_j^T\hat{V}_j\} \geq \|\tilde{V}_j\|_F^2 - \|V_j^*\|_F^2$, the following inequalities hold:

$$\begin{aligned}\dot{V}_{sj} \leq &-\frac{e_{sj}^2}{\varepsilon_j}\left(1 + \int_0^1 \theta\overline{g}_j(\overline{X}_j, \theta e_{sj} + \nu_j)\, d\theta + \|\overline{Z}_j\hat{W}_j^T\hat{S}'_j\|_F^2 + \|\hat{S}'_j\hat{V}_j\overline{Z}_j\|^2\right)\\ &+ (\|V_j^*\|_F\|\overline{Z}_j\hat{W}_j^T\hat{S}'_j\|_F + \|W_j^*\|\,\|\hat{S}'_j\hat{V}_j^T\overline{Z}_j\|\\ &+ |W_j^*|_1 + |\mu_j(Z_j)|)|e_{sj}|\\ &- \frac{\sigma_{wj}}{2}(\|\tilde{W}_j\|^2 - \|W_j^*\|^2) - \frac{\sigma_{vj}}{2}(\|\tilde{V}_j\|_F^2 - \|V_j^*\|_F^2)\end{aligned} \tag{30}$$

Since

$$\|V_j^*\|_F\|\overline{Z}_j\hat{W}_j^T\hat{S}''_j\|_F|e_{sj}| \leq \frac{e_{sj}^2}{\varepsilon_j}\|\overline{Z}_j\hat{W}_j^T\hat{S}'_j\|_F^2 + \frac{\varepsilon_j}{4}\|V_j^*\|_F^2 \tag{31}$$

$$\begin{aligned}(|W_j^*|_1 + |\mu_j(Z_j)|)|e_{sj}| &\leq \frac{e_{sj}^2}{2\varepsilon_j} + \varepsilon_j[|W_j^*|_1^2 + \mu_j^2(Z_j)]\\ \|W_j^*\|\,\|\hat{S}'_j\hat{V}_j^T\overline{Z}_j\|\,|e_{sj}| &\leq \frac{e_{sj}^2}{\varepsilon_j}\|\hat{S}'_j\hat{V}_j^T\overline{Z}_j\|^2 + \frac{\varepsilon_j}{4}\|W_j^*\|^2\end{aligned} \tag{32}$$

and $|\mu_j(Z_j)| \leq \overline{\mu}_j$, inequality (30) can be rewritten as

$$\begin{aligned}\dot{V}_{sj} \leq &-\frac{e_{sj}^2}{\varepsilon_j}\left[\frac{1}{2} + \int_0^1 \theta\overline{g}_j(\overline{X}_j, \theta e_{sj} + \nu_j)\, d\theta\right]\\ &- \frac{\sigma_{wj}}{2}\|\tilde{W}_j\|^2 - \frac{\sigma_{vj}}{2}\|\tilde{V}_j\|_F^2 + c_j\end{aligned} \tag{33}$$

with constants

$$\begin{aligned}c_j = &\varepsilon_j\left(\frac{1}{4}\|W_j^*\|^2 + \frac{1}{4}\|V_j^*\|_F^2 + |W_j^*|_1^2 + \overline{\mu}_j^2\right)\\ &+ \frac{\sigma_{wj}^2}{2}\|W_j^*\|^2 + \frac{\sigma_{vj}^2}{2}\|V_j^*\|_F^2\end{aligned} \tag{34}$$

Considering (12), (29), and (33), we further have $\dot{V}_{sj} \leq -\lambda_{sj}V_{sj} + c_j$ with $\lambda_{sj} = \min\{g_{0j}/\varepsilon_j,\ \sigma_{wj}/\lambda_{\max}(\Gamma_{wj}^{-1}),\ \sigma_{vj}/\lambda_{\max}(\Gamma_{vj}^{-1})\}$. Therefore

$$V_{sj}(t) \leq V_{sj}(0)e^{-\lambda_{sj}t} + \frac{1}{\lambda_{sj}}c_j \quad \forall t \geq 0 \tag{35}$$

where constants

$$\begin{aligned}V_{sj}(0) = &\int_0^{e_{sj}(0)} \sigma\beta_j(\overline{X}_j(0), \sigma + \nu_j(0))\, d\sigma\\ &+ \tfrac{1}{2}[\tilde{W}_j^T(0)\Gamma_{wj}^{-1}\tilde{W}_j(0) + \mathrm{tr}\,\{\tilde{V}_j^T(0)\Gamma_{vj}^{-1}\tilde{V}_j(0)\}]\end{aligned} \tag{36}$$

Since $\|\tilde{W}_j\|^2 \leq V_{sj}(t)/\lambda_{\min}(\Gamma_{wj}^{-1})$ and $\|\tilde{V}_j\|^2 \leq V_{sj}(t)/\lambda_{\min}(\Gamma_{vj}^{-1})$, the NN weights $(\hat{W}_j, \hat{V}_j)$ belong to the following compact sets

$$\Omega_{wj} = \left\{(\hat{W}_j, \hat{V}_j) \mid \|\tilde{W}_j\|^2 \leq \frac{2[V_{sj}(0) + c_j/\lambda_{sj}]}{\lambda_{\min}(\Gamma_{wj}^{-1})},\ \|\tilde{V}_j\|_F^2 \leq \frac{2[V_{sj}(0) + c_j/\lambda_{sj}]}{\lambda_{\min}(\Gamma_{vj}^{-1})}\right\}. \tag{37}$$

It is shown from (11), (12), and (29) that $V_{sj}(t) \geq V_{zj} \geq e_{sj}^2(t)/2$. According to Remark 2.2, the system states remain in the compact set

$$\Omega_X = \left\{ X \mid |e_{j,i_j}(t)| \leq 2^{i_j-1}\lambda_j^{i_j-\rho_j}\left(2V_{sj}(0)e^{-\lambda_{sj}t} + \frac{2c_j}{\lambda_{sj}}\right)^{1/2}, \right. \left. i_j = 1,2,\cdots,\rho_j, j=1,2,\cdots,m, \overline{x}_{d(j+1)} \in \Omega_{d(j+1)} \right\}. \tag{38}$$

We conclude that for bounded initial conditions, compact sets Ω_{zj} defined in (26) exist such that the vectors $Z_j \in \Omega_{zj}$ for all time, which means all the signals $X, \hat{W}_j, \hat{V}_j$, and u_j are being bounded.

2) As $\int_0^1 \theta \overline{g}_j(\overline{X}_j, \theta e_{sj} + \nu_j)\, d\theta \geq g_{0j}/2$, inequality (33) can be further written as $\dot{V}_{sj} \leq -e_{sj}^2(1+g_{0j})/2\varepsilon_j + c_j$. Integrating them over $[0,t]$ leads to

$$\int_0^t e_{sj}^2(\tau)\, d\tau \leq \frac{2\varepsilon_j}{1+g_{0j}}[V_{sj}(0) + tc_j], \qquad j=1,2,\cdots,m. \tag{39}$$

Since tracking errors $e_{j,1} = H_j(s)e_{sj}$ with stable transfer functions $H_j(s)$ (see Remark 2.2) by applying [22, Lemma 4.8.2] we obtain

$$\int_0^t e_{j,1}^2(\tau)\, d\tau \leq a_j \int_0^t e_{sj}^2(\tau)\, d\tau + b_j \leq \frac{2t\varepsilon_j a_j c_j}{1+g_{0j}} + \frac{2\varepsilon_j a_j V_{sj}(0)}{1+g_{0j}} + b_j \tag{40}$$

with computable constants $a_j, b_j > 0$. Dividing (40) by t, we arrive at (27). Considering (35) and $|e_{sj}(t)| \leq \sqrt{2V_{sj}(t)}$, and using (3) for $i_j = 1$, inequality (28) follows. **Q.E.D.**

Remark 4.1: Noting (27) and (28), we can see that the transient responses of tracking errors are affected by the bounds of $V_{sj}(0)$ significantly. It is shown from (36) that larger adaptation gains Γ_{wj} and Γ_{vj} may result in smaller $V_{sj}(0)$; hence, fast adaptations are helpful for improving the transient performance. In practice, however, we do not suggest the use of high adaptation gains or very small design parameters ϵ_j in (23) because such a choice may result in a variation of high-gain control [21], and, therefore, increase the bandwidth of the adaptive system. Any small noise in the measurements might be amplified and cause large oscillations in the control outputs.

Remark 4.2: It should be mentioned that the integrals in control gain (23) might not be solved analytically for some functions $\overline{g}_j(X_j)$ and may make the implementation of the controller difficult. This problem can be dealt with by suitably choosing the design functions $\overline{g}_j(X_j)$. Since the choices of $\overline{g}_j(X_j)$ are only required to be larger than $|g_{j,j}(X_j)|$, the designer has the freedom to find suitable $\overline{g}_j(X_j)$ such that the integrals are analytically solvable. As an alternative scheme, one can also use on-line numerical approximation to calculate the integral, which however requires more computational power in practical applications.

Remark 4.3: In Theorem 4.1, only boundedness of the system states and the responses of output tracking errors are given; convergence of the NN weight estimates is not guaranteed because of the lack of persistent excitation (PE) condition and the local minima problems of nonlinear parametric approximators. In adaptive control systems, PE condition is important for parameter convergence and system robustness; however, it is usually difficult to check/guarantee in practical applications [22]. The works [17]–[19] have studied PE conditions of linearly parametrized networks and provided several practically applicable methods. The persistent excitation condition for nonlinearly parametrized networks is a current topic of research in neural network identification and control fields.

V. Conclusion

In this note, we have presented a stable adaptive control scheme for a class of multivariable nonlinear systems using multilayer neural networks. The triangular property of the studied systems has been fully utilized in developing the control structure and neural weight learning laws. The proposed design guarantees the stability of the closed-loop adaptive system and the mean square tracking errors converging to small residual sets, which are adjustable by tuning the design parameters.

References

[1] S. S. Sastry and A. Isidori, "Adaptive control of linearizable systems," *IEEE Trans. Automat. Contr.*, vol. 34, pp. 1123–1131, Nov. 1989.

[2] C. C. Liu and F. C. Chen, "Adaptive control of non-linear continuous-time systems using neural networks—General relative degree and MIMO cases," *Int. J. Contr.*, vol. 58, pp. 317–335, 1993.

[3] 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, pp. 791–801, May 1995.

[4] K. S. Narendra and S. Mukhopadhyay, "Adaptive control of nonlinear multivariable system using neural networks," *IEEE Trans. Neural Networks*, vol. 7, pp. 737–752, Sept. 1994.

[5] S. S. Ge, "Robust adaptive NN feedback linearization control of nonlinear systems," *Int. J. Syst. Sci.*, vol. 27, pp. 1327–1338, 1996.

[6] 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.

[7] L. Shi and S. K. Singh, "Decentralized adaptive controller design for large-scale systems with higher order interconnections," *IEEE Trans. Automat. Contr.*, vol. 37, pp. 1106–1118, Aug. 1992.

[8] J. T. Spooner and K. M. Passino, "Adaptive control of a class of decentralized nonlinear systems," *IEEE Trans. Automat. Contr.*, vol. 41, pp. 280–284, 1996.

[9] S. Jain and F. Khorrami, "Decentralized adaptive control of a class of large-scale interconnected nonlinear systems," *IEEE Trans. Automat. Contr.*, vol. 42, pp. 136–154, 1997.

[10] R. M. Sanner and J. E. Slotine, "Gaussian networks for direct adaptive control," *IEEE Trans. Neural Networks*, vol. 3, pp. 837–863, Nov. 1992.

[11] 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, Mar. 1996.

[12] A. Yesidirek and F. L. Lewis, "Feedback linearization using neural networks," *Automatica*, vol. 31, no. 11, pp. 1659–1664, 1995.

[13] T. Zhang, S. S. Ge, and C. C. Hang, "Design and performance analysis of a direct adaptive controller for nonlinear systems," *Automatica*, vol. 35, no. 11, pp. 1809–1817, 1999.

[14] ——, "Direct adaptive control of non-affine nonlinear systems using multilayer neural networks," in *Proc. Amer. Contr. Conf.*, Philadelphia, PA, June 1998, pp. 515–519.

[15] S. S. Ge, T. H. Lee, and C. J. Harris, *Adaptive Neural Network Control of Robotic Manipulators*. Singapore: World Scientific, Dec. 1998.

[16] K. I. Funahashi, "On the approximate realization of continuous mappings by neural networks," *IEEE Trans. Neural Networks*, vol. 2, pp. 183–192, 1989.

[17] J. Farrell and T. Berger, "On the effects of the training sample density in passive learning control," in *Proc. Amer. Contr. Conf.*, Seattle, WA, 1995, pp. 872–877.

[18] J. Farrell, "Persistence of excitation conditions in passive learning control," *Automatica*, vol. 33, pp. 699–703, 1997.

[19] ——, "Stability and approximator convergence in nonparametric nonlinear adaptive control," *IEEE Trans. Neural Networks*, vol. 9, pp. 1008–1020, 1998.

[20] D. Gorinevsky, "Ont he persistency of excitation in radial basis function network identification of nonlinear systems," *IEEE Trans. Neural Networks*, vol. 6, pp. 1237–1244, 1995.

[21] M. M. Polycarpou, "Stable adaptive neural control scheme for nonlinear systems," *IEEE Trans. Automat. Contr.*, vol. 41, no. 3, pp. 447–451, 1996.

[22] P. A. Ioannou and J. Sun, *Robust Adaptive Control*. Englewood Cliffs, NJ: Prentice-Hall, 1996.

[23] J. E. Slotine and W. Li, *Applied Nonlinear Control*. Englewood Cliffs, NJ: Prentice-Hall, 1991.

[24] M. Vidyasagar, *Nonlinear Systems Analysis*, 2nd ed. Englewood Cliffs, NJ: Prentice-Hall, 1993.